Selected papers on

GEODESIC FIELDS

Translated and edited by

D. H. Delphenich

Table of contents

Introduction

This collection of selecta represents a tributary of evolution in the history of a mathematical idea, namely, the role played by geodesic fields in the calculus of variations for multiple integrals. Along the way, one also sees how the nature of the problem suggested the introduction of the methods of the calculus of exterior differential forms as a convenient tool, the basic concepts of contact geometry, and eventually the notion that what finite variations actually seem to represent are differentiable homotopies of the objects being varied. The infinitesimal variations are then simply the vector fields that one obtains by differentiating the homotopy parameter, and thus represent infinitesimal generators of differentiable one-parameter families of objects that begin with the object that one is varying.

As one learns in the calculus of a finite number of variables, the necessary and sufficient condition for a critical point of a twice-differentiable function to be a local minimum is that the Hessian matrix that is defined by the second partial derivatives of that function must be positive-definite. In the case of infinity variables, which is how one can regard the calculus of variations (at least, heuristically), one sees that since the matrix in question would be potentially infinite in its numbers of rows and columns (and probably not even countably infinite), some other approach must be taken.

Weierstrass and Legendre obtained in the case of extremal curves were sufficient conditions for a "strong" and "weak" local minimum, respectively. Indeed, the Legendre condition is the positive-definiteness of the Hessian for the Lagrangian function itself, at least when one considers the partial derivatives with respect to the velocities. Furthermore, the customary analysis [1] of sufficient conditions for an extremal curve to be a (strong or weak) minimum involved the constructions of Mayer fields of extremals in which the given extremal would be embedded, the Hilbert independent integral, and the Weierstrass excess function.

Apparently, the genesis of the chain of events that is documented here in translation was a paper [**3**] by Clebsch that attempted to extend the well-established theory of extremal curves to extremal submanifolds of higher dimension. In a series of brief notes [**4**], Hadamard later commented that the problem of establishing sufficient conditions for an extremal submanifold to represent a strong or weak minimum for the action functional in question was more involved than it had been in the case of minimal curves, such as geodesics. Previously, the theory of sufficient conditions for a minimal curve had been adequately described by Legendre, who gave a sufficient condition for a weak minimum and Weierstrass, who gave a sufficient condition for a strong minimum. The Legendre condition was concerned with the character of the eigenvalues of the Hessian of the Lagrangian function for the action functional, while the Weierstrass condition involved introducing an "excess" function that allowed one to describe the difference between the values of the action functional on an extremal curve and a neighboring curve that did not have to be extremal. What Hadamard contributed was mostly a strengthening of the Legendre condition for a weak minimum.

---

[1] Some standard references on the calculus of variations that address the details of these conditions are Bliss [**1**] and Gelfand and Fomin [**2**].

It was the seminal paper [**5**] of Caratheodory that first addressed the issues brought up by Hadamard in a definitive way. In his formulation of the Weierstrass condition for the strong minimum, he introduced the concept of a *geodesic* field, whose existence is a stronger condition than the existence of a single extremal, and amounts to a foliation by extremal submanifolds. In the case of extremal submanifolds of dimension greater than one, he found that the existence of such a field was a necessary part of defining the Weierstrass excess function, as well as the embedding of the given extremal in such a field. In the process, he made essential use of the Hilbert independent integral, the Hamilton-Jacobi equation, and ultimately introduced the *Caratheodory complete figure* of a variational problem, which consisted of a pair of complementary foliations of space that were defined by the contact elements of the geodesic field and some transversal vector fields that appear in the process of its definition. This construction generalized the picture that emerges in geometrical optics, or the theory of first-order partial differential equations, in which there is a foliation of dimension one defined by the null geodesic congruence and a transversal foliation of codimension one that is defined by the level hypersurfaces of the eikonal function, which are then geodesically equidistant.

Later, Caratheodory presented his work on the calculus of variations in an influential book [**6**], in which he took the canonical – i.e., Hamiltonian – approach for the most part. The general theory of the Hamilton-Jacobi equation and its application to the problems of physics was also eventually documented in the definitive treatise of Rund [**7**].

Some time after Caratheodory's paper on multiple integrals in the calculus of variations, Hermann Weyl commented on Caratheodory's formulation of the problem, first in a brief note to the Physical Review [**8**], in which he applied the methodology to a critique of an early attempt that Max Born made at formulating quantum electrodynamics, and later in a more comprehensive article in the Annals of Mathematics [**9**]. Weyl characterized Caratheodory's approach as one in which the integrand of the action functional was defined by a certain determinant, while his own version of the theory was a simpler version of the theory that involved a trace. He also showed that, in effect, the trace theory was an infinitesimal approximation to the determinant theory. However, it was eventually recognized that the two theories produced different conclusions.

One of the parallel chains of development in this series of advances in the name of the Weierstrass problem was the fact that the mathematicians were gradually recognizing that the calculus of exterior differential forms seemed to represent a natural tool for applications to the calculus of variations. In 1921, Cartan published his lectures on integral invariants [**10**], in which he more specifically showed how exterior differential forms could illuminate the problems of the calculus of variations at a fundamental level. This calculus was also discussed by Goursat in a book [**11**] that he published in 1922 on the Pfaff problem, which grew out of his own work in partial differential equations. Later, de Donder developed this idea further in two books [**12**, **13**] on integral invariants and the invariant theory of the calculus of variations, respectively.

As one sees in the selected readings, Boerner [**14**], Géhéniau [**15**], Lepage [**16**], Debever [**17**], Hölder [**18**], Paquet [**19**], van Hove [**20**], and Dedecker [**21, 22**] carried on this tradition quite effectively; one also notices that the Belgian school of researchers was mostly developing the ideas of De Donder. As van Hove points out, the Lepage congruences define a concise way of relating that various definitions of geodesic fields as being equivalent in a reasonable sense; in particular, the differing approaches of De Donder-Weyl and Carathéodory are seen to be equivalent in the sense of Lepage congruence and can be converted to each other by means of a contact transformation, as shown by Hölder.

When one proceeds from the study of the exterior differential calculus to study of exterior differential *systems*, one arrives at one of the modern approaches to the calculus of variations (see, [**23**, **24**]). One of the directions that this took was an increasing emphasis on the formulation of Hamiltonian mechanic on symplectic manifolds, as well well as the Hamilton-Jacobi problem in mechanics. However, one should notice that some of the identifications that make this all quite natural in the case of point mechanics actually seem increasingly contrived and non-intuitive when one goes to variations of higher-dimensional objects, such as submanifolds and fields. Thus, it is probably best to keep an open mind about the relative advantages of the Lagrangian and Hamiltonian pictures when one is not always concerned exclusively with point mechanics, especially if one wishes to *weaken* the integrability assumptions that one makes about a physical system, not strengthen them. For instance, in real-world physical systems, there are such things as non-conservative forces and non-holonomic constraints to contend with.

An important aspect of Dedecker's formulation of the calculus of variations in [**21**] is his choice of *differentiable singular cubic chains* in a differentiable manifold M as the basic objects that one is varying. These topological building blocks are no loss in generality from the use of compact submanifolds, and are entirely natural when one is using exterior differential forms, since one usually defines the integration of differential forms over such chains at the elementary level. Hence, linear functionals on chains that are defined by the integration of differential forms can be immediately identified with real cochains, which is an elementary form of de Rham's theorem. A finite variation of a k-chain in a manifold M is then a differentiable homotopy of that chain, which can also be regarded as a differentiable singular cubic $k+1$-chain. Clearly, if one intends to investigate the topological aspects of the calculus of variations, this is the correct foundation. Indeed, one might confer a later monograph [**22**] by Dedecker, in which he expanded upon some of the themes of the paper included in this collection.

Some of the ancillary topics that were being developed elsewhere in mathematics along the way, but which did not get applied in the articles featured here, were the introduction of jets by Ehresmann [**25**] and the increasing recognition that the most fundamental and unavoidable kind of geometry that pertained to the study of jets and the calculus of variations was *contact geometry*. Indeed, this fact was already recognized by some figures of mathematical physics, such as Hamilton [**26**], Lie and Scheffer [**27**], Vessiot [**28**], and the aforementioned Hölder [**18**]. Hence, in our introductory remarks on a more modern formulation of the Weierstrass problem, we shall discuss some of these innovations as they relate to the readings.

What follows in this Introduction is then a summary of how one uses geodesic fields in the formulation of the Weierstrass problem when one includes some of the more modern considerations of jet manifolds and contact elements. Thus, we commence with a somewhat lengthy general discussion of these latter concepts.

1. Contact elements and jets. Suppose $\mathcal{O}$ is an orientable r-dimensional differentiable manifold with boundary, which will play the role of a parameter manifold for us. Indeed, we will usually regard $\mathcal{O}$ as a closed subset of $\mathbb{R}^r$, so one can also refer to the points by r-tuples $(t^1, \ldots, t^r)$ of coordinates t^a directly. For instance, if the central object of consideration is a curve then $\mathcal{O}$ will be a line segment, such as [0, 1], for the sake of specificity. If one is concerned with surfaces then $\mathcal{O}$ will be two-dimensional, such as a closed disc D^2.

Now, let M be an orientable m-dimensional differentiable manifold in which our objects are defined as submanifolds, where an r-dimensional submanifold in M is a differentiable map x: $\mathcal{O} \to M$. If t is a point in $\mathcal{O}$ and $x(t)$ is its image in M then the differential map $dx|_t$: $T_t\mathcal{O} \to T_{x(t)}M$ takes the r-dimensional vector space $T_t\mathcal{O}$ to an $r'(t)$-dimensional subspace of $T_{x(t)}M$, where $r'(t)$ is the rank of the linear map $dx|_t$. Hence, $r'(t) = r$ at t iff $dx|_t$ is a linear injection, so, if one desires that this should always be the case then one implicitly demands that x must be an *immersion*. Since immersions can still have self-intersections, which would make their images no longer represent differentiable manifolds, a stronger condition to impose is that x be an *embedding*, which means that when one gives $x(\mathcal{O})$ the subspace topology in M – viz., its open subsets are intersections of open subsets in the topology of M with $x(\mathcal{O})$ itself – the map x becomes a homeomorphism.

a. Contact elements. One can think of the $r'(t)$-dimensional linear subspace $dx|_t(T_t\mathcal{O})$ as the *contact element* to $x(t)$ that is defined by x at t. In the simplest case of curves in M this contact element will be the line in $T_{x(t)}M$ that is tangent to the curve $x(t)$ at each point. Note that this line field is generated by the velocity vector field dx/dt iff the velocity is non-vanishing for every t. One can also introduce the *projectivized tangent bundle* $PT(M)$, whose elements are lines through the origins in the tangent space to M, and think of a line field along the curve $x(t)$ as a special type of curve in $PT(M)$ itself.

More generally, the contact element $dx|_t(T_t\mathcal{O})$ will be an $r'(t)$-plane in $T_{x(t)}M$, and if x is an immersion then one can introduce the Grassmanian manifold $V^m_{r,x(t)}(M)$ of r-planes in $T_{x(t)}M$ and define a section of $V^m_r(M) \to M$ along $x(t)$ to be the association of each $t \in \mathcal{O}$ with the image vector subspace $dx|_t(T_t\mathcal{O})$ in $V^m_{r,x(t)}(M)$.

This is the route that Dedecker followed in his article [**21**] on the use of differential forms in the calculus of variations, but in the intervening years it has become more customary to go the route of jet manifolds in order to describe contact elements and contact geometry. Fortunately, the transition from Grassmanian manifolds of tangent subspaces to jets is quite immediate.

b. 1-jets. The *1-jet* $j^1_t x$ of a C^1 (i.e., continuously differentiable) function x: $\mathcal{O} \to M$ at $t \in \mathcal{O}$ is defined to be the set of all C^1 functions x': $U_t \to M$ that are defined in some neighborhood U_t of t and have the properties that their values $x'(t)$ all agree with $x(t)$ and the values of their differentials $dx'|_t$ all agree with that of $dx|_t$. If the coordinates of $\mathcal{O} \subset \mathbb{R}^r$ are t^a, $a = 1, \ldots, r$ and the coordinates of some neighborhood of $x(t)$ in M are x^i, $i = 1, \ldots, m$ then the points in the image of $\mathcal{O}$ that lie inside the coordinate neighborhood around $x(t)$ can be described by the system of m equations in r independent variables of the form $x^i = x^i(t^a)$, while the values of the differential map $dx|_t$ can be described by a system of rm equations in r independent variables of the form:

$$x_a^i = \frac{\partial x^i}{\partial t^a}(t) .$$

More generally, we can regard the $(r + m + rm)$-tuple of coordinates (t^a, x^i, x_a^i) as a coordinate representation of the 1-jets $j_t^1 x$ that are associated with the point $t \in \mathcal{O}$ described by the t^a and the point $x \in M$ that is described by x^i. However, the $(r + m + rm)$-tuple (t^a, x^i, x_a^i) is more general than the previous discussion of contact elements in several ways:

1. We are not requiring that the x^i be associated with t^a by way of a C^1 function $x: \mathcal{O} \to M$, and similarly x_a^i is not functionally related to either t^a or x^i.

2. Since there is no functional relationship between x_a^i and t^a, one cannot require that x_a^i take the form of a matrix of partial derivatives.

3. We have made no restriction on the rank of the matrix x_a^i; indeed, it could even be 0.

In order to address these situations as special cases of a more general construction, we first agree that the 1-jet $j_t^1 x$ does indeed represent an r-parameter family of contact elements of dimension $0 \leq r'(t) \leq r$ in the tangent space T_xM. The set $J^1(\mathcal{O}; M)$ of all such 1-jets can be given a topology and local coordinates systems of the form (t^a, x^i, x_a^i) that makes it a differentiable manifold of dimension $r+m+rm$. One calls $J^1(\mathcal{O}; M)$ the *manifold of 1-jets of C^1 functions from $\mathcal{O}$ into M.*

There are three canonically defined projections that are associated with the manifold $J^1(\mathcal{O}; M)$:

source projection:	$J^1(\mathcal{O}; M) \to \mathcal{O}$,	$j_t^1 x \mapsto t$,
target projection:	$J^1(\mathcal{O}; M) \to M$,	$j_t^1 x \mapsto x$,
contact projection:	$J^1(\mathcal{O}; M) \to \mathcal{O}\times M$	$j_t^1 x \mapsto (t, x)$.

None of these projections are fibrations, but they are submersions; i.e., their differentials have maximum rank in each case. As a consequence, the fiber over any point of the image manifold will be a submanifold of $J^1(\mathcal{O}; M)$, although one does not always have local triviality of some neighborhood of each fiber that would make the submersion a fibration. Hence, $J^1(\mathcal{O}; M)$ is referred to as a *fibered manifold* relative to any of these projections. Locally, the fibers are parameterized by (x^i, x^i_a), (t^a, x^i_a), and (x^i_a), respectively.

In case of the contact projection $J^1(\mathcal{O}; M) \to \mathcal{O}\times M$ one can see that the fibers are affine spaces that are modeled on the rm-dimensional vector space $L(r, m)$ of linear maps from $\mathbb{R}^r$ into $\mathbb{R}^m$. In fact, it will be crucial at a later point to understand that any element $j^1_t x$ can be regarded as a linear map from $T_t\mathcal{O}$ to T_xM, which defines a canonical identification of the fiber of the contact projection over (t, x) with the fiber of the projection of $T^*\mathcal{O} \otimes T(M) \to \mathcal{O}\times M$ over (t, x).

It will prove intuitively useful to represent the elements of $J^1(\mathcal{O}; M)$ schematically as in Fig. 1, in which the dimensions of the manifolds $\mathcal{O}$, M, and $L(r, m)$ have been reduced to one for ease of representation.

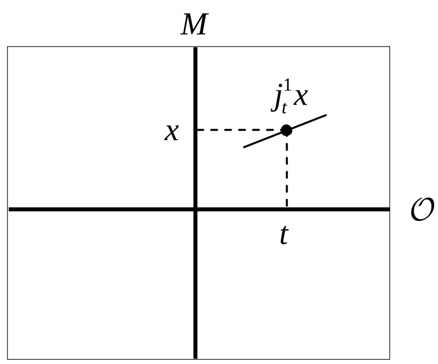

Figure 1. Schematic representation of an element $j^1_t x$ in $J^1(\mathcal{O}; M)$.

c. Sections of the projections. In order to account for the functional relationship between the t^a and both the x^i, as well as the x^i_a, we introduce the concept of a *section* of the source projection, which will be a C^1 map s: $\mathcal{O} \to J^1(\mathcal{O}; M)$, $t \mapsto s(t)$ such that the source projection takes each $s(t)$ back to t. When this is not the case, we shall call the section s a *singular section* of the source projection. Hence, in either case if the coordinates in a neighborhood of t are t^a then the coordinates of the points in the image of s will be:

$$s(t) = (t^a, x^i(t), x^i_a(t)).$$

In the particular case:

$$s(t) = (t^a, x^i(t), x^i_{,a}(t)),$$

where the $x^i_{,a}(t)$ are the partial derivatives $\partial x^i/\partial t^a(t)$, we shall refer to the section s as *integrable* and also regard it as the *1-jet prolongation* of the map x:

$$s = j_t^1 x .$$

This means that the integrability conditions for a general section take the local form:

$$x_a^i = \frac{\partial x^i}{\partial t^a} .$$

If one rewrites this in terms of coordinate differentials on $J^1(\mathcal{O}; M)$ as:

$$dx^i = x_a^i \, dt^a$$

then one sees that if one defines the set of m 1-forms on $J^1(\mathcal{O}; M)$ that take the local form[2]:

$$\omega^j = dx^i - x_a^i \, dt^a$$

then one can say that a section s: $\mathcal{O} \rightarrow J^1(\mathcal{O}; M)$ is integrable iff:

$$s^* \omega^j = 0 \qquad \text{for all } i = 1, \ldots, m.$$

As we shall see, these 1-forms ω^j play a recurring role in all of what follows, and their simultaneous vanishing at a point of $J^1(\mathcal{O}; M)$ defines a subspace in $T(J^1)$ of codimension m. The sub-bundle of $T(J^1)$ that consists of all such linear subspaces is referred to as the *contact structure* on $J^1(\mathcal{O}; M)$. If one regards the section s: $\mathcal{O} \rightarrow J^1(\mathcal{O}; M)$ as a submanifold of $J^1(\mathcal{O}; M)$ then the integrability of the section s means that the tangent spaces to the submanifold must be subspaces of the contact structure on $J^1(\mathcal{O}; M)$.

We illustrate the nature of general sections, singular sections, and integrable sections of the source projection schematically in Fig. 2.

[2] Although these 1-forms can be generalized to non-local expressions, since it will not be necessary for us to have that definition in what follows, we only refer the curious to the literature, such as Goldschmidt and Sternberg [**30**] or Saunders [**31**].

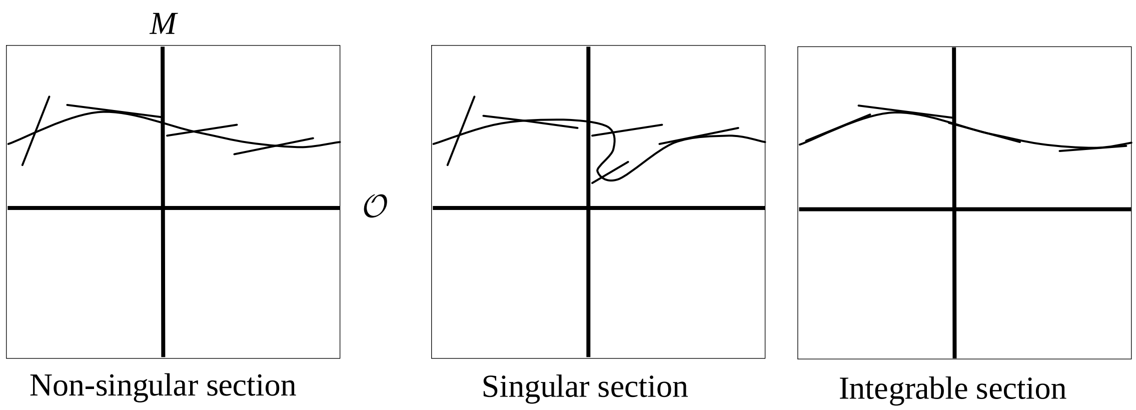

Figure 2. The basic types of sections of $J^1(\mathcal{O}; M) \to \mathcal{O}$.

In order to make sense of the Lepage congruences that are defined in [**16**], and which we will discuss in more detail later, one needs only to say that two k-forms α and α' in the exterior algebra $\Lambda^k(J^1)$ are *congruent modulo* ω^i iff their difference $\alpha - \alpha'$ is an element of the ideal $I\{\omega^1, \ldots, \omega^m\}$ in that exterior exterior algebra that is generated by the set $\{\omega^1, \ldots, \omega^m\}$. Its elements take the form of finite linear combinations of the form $A_\mu \wedge \omega^\mu$, where the A_μ, $\mu = 1, \ldots, N$ are k–1-forms on $J^1(\mathcal{O}; M)$. That is, one writes:

$$\alpha \equiv \alpha' \ (\mathrm{mod}\ \omega^i)$$

iff there exist $A_i \in \Lambda^{k-1}(J^1)$ such that:

$$\alpha - \alpha' = A_i \wedge \omega^i.$$

One sees that whenever two k-forms are congruent in the Lepage sense their pull-backs $s^*\alpha$ and $s^*\alpha'$ by any section s: $\mathcal{O} \to J^1(\mathcal{O}; M)$ will agree whenever s is integrable. In particular, as we will discuss later, Lepage concentrated on the congruence class of the fundamental p-form that defines the integrand of the action functional.

d. The integrability of the contact structure. Since the sub-bundle of $T(J^1)$ that is defined by its contact structure also represents a differential system on $T(J^1)$ of corank m the question naturally arises whether that differential system is completely integrable. That is: can it be foliated by integral submanifolds of codimension m?

The necessary and sufficient condition for this to be the case is given by Frobenius's theorem, which can be phrased in various forms. The one that is of immediate interest to us is based in the fact that since the sub-bundle in question is defined by the vanishing of the contact forms ω^i, $i = 1, \ldots, m$, it is the algebraic solution to the exterior differential system:

$$0 = \omega^i, \ i = 1, \ldots, m.$$

The form that Frobenius's theorem takes for such a system is to say that the system is completely integrable iff there are 1-forms η^i_j on $J^1(\mathcal{O}; M)$ such that:

$$d\omega^j = \eta^i_j \wedge \omega^j, \qquad i = 1, \dots, m.$$

Now if we define:

$$\Theta^i \equiv d\omega^j = dt^a \wedge dx^i_a$$

then the question at hand is whether Θ^i is expressible in the form $\eta^i_j \wedge \omega^j$ for suitable 1-forms η^i_j; i.e., can one solve:

$$dt^a \wedge dx^i_a = \eta^i_j \wedge (dx^j - x^i_a dt^a) = \eta^i_j \wedge dx^j - x^j_a \eta^i_j \wedge dt^a$$

for some set of η^i_j. However, since the left-hand side does not contain dx^i as an exterior factor, one must have the vanishing of $\eta^i_j \wedge dx^j$, which forces the forms η^i_j to take the form:

$$\eta^i_j = \eta^i\, dx^j.$$

Since such 1-forms give:

$$x^j_a \eta^i_j \wedge dt^a = x^i_a\, \eta^i\, dx^j \wedge dt^a,$$

and not something of the form $dt^a \wedge dx^i_a$, one must conclude that no such 1-forms exist, and the differential system is not completely integrable.

However, the fact that integral submanifolds of dimension $r + rm$ do not exist does not imply that integral submanifolds of lower dimension cannot exist. In particular, any integrable section $s: \mathcal{O} \to J^1(\mathcal{O}; M)$ defines an integral submanifold of dimension r.

e. Sections of the source projection. Any integrable section $s: \mathcal{O} \to J^1(\mathcal{O}; M)$ will also have the property that:

$$s^*\Theta^i = s^*(d\omega^j) = d(s^*\omega^j) = 0 \quad \text{for all } i,$$

since integrability makes $s^*\omega^j$ vanish.

More generally, when a submanifold $s: N \to J^1(\mathcal{O}; M)$, with N an n-dimensional manifold, has the property that

$$s^*\Theta^i = 0 \qquad \text{for all } i$$

one calls it an *isotropic submanifold*.

Such a submanifold is also an integral submanifold of the differential system on $J^1(\mathcal{O}; M)$ that is defined by the vanishing of all Θ^i, that is the sub-bundle of $T(J^1)$ that is the algebraic solution to the exterior differential equation:

$$\Theta^i = 0.$$

One sees that although the differential system that is defined by $\omega^i = 0$ does not have to be completely integrable, nonetheless, the system defined by $\Theta^i = 0$ does have to be completely integrable, since all of the Θ^i are closed.

The maximum dimension of an isotropic submanifold is dictated by the rank of the Θ^i; viz., maximum dimension of an tangent subspace to $J^1(\mathcal{O}; M)$ on which $\Theta^i(X, Y)$ for all i when one is given any two tangent vectors X, Y in that subspace. When an isotropic submanifold has this maximum dimension, we shall call it a *Legendrian submanifold*. In general, that maximum dimension is difficult to ascertain, although it is well-known in the case of one-dimensional parameter manifolds, as we shall discuss shortly.

If $s(u)$ has the local expression $s(u) = (t^a(u), x^i(u), x_a^i(u))$ then the pull-back of the Θ^i to N by way of s has the local expression:

$$s^*\Theta^i = \left(\frac{\partial t^a}{\partial u^\alpha}du^\alpha\right)\wedge\left(\frac{\partial x_a^i}{\partial u^\beta}du^\alpha\right) = \frac{1}{2}\left(\frac{\partial t^a}{\partial u^\alpha}\frac{\partial x_a^i}{\partial u^\beta} - \frac{\partial t^a}{\partial u^\beta}\frac{\partial x_a^i}{\partial u^\alpha}\right)du^\alpha\wedge du^\alpha .$$

This vanishes iff all of the Lagrange brackets:

$$[u^\alpha, u^\beta] = \frac{\partial t^a}{\partial u^\alpha}\frac{\partial x_a^i}{\partial u^\beta} - \frac{\partial t^a}{\partial u^\beta}\frac{\partial x_a^i}{\partial u^\alpha}$$

vanish. Hence, one can regard the vanishing of Lagrange brackets as an integrability condition for integral submanifolds of the exterior differential system defined by the vanishing of the Θ^i. It is for this reason that maximal isotropic manifolds, at least when one is concerned with 1-jets of curves, are also referred to as *Lagrangian submanifolds.*

In the case of a section s: $\mathcal{O} \to J^1(\mathcal{O}; M)$ of the source projection, when one expresses it in the local form $(t^a, x^i(t), x_a^i(t))$, one sees that the local form of the pulled-back 2-forms on $\mathcal{O}$ that are defined by $s^*\Theta^i$ is:

$$s^*\Theta^i = dt^a \wedge s^* dx_a^i = dt^a \wedge\left(\frac{\partial x_a^i}{\partial t^b}dt^b\right) = \tfrac{1}{2}(x_{a,b}^i - x_{b.a}^i)dt^a\wedge dt^b .$$

Hence, if one defines the m 1-forms on $\mathcal{O}$:

$$\alpha^i = x_a^i dt^a, \qquad i = 1, \ldots, m$$

then the isotropy of s implies the necessary condition that all of the α^i must be closed 1-forms.

By contrast, the requirement that s be integrable implies the necessary condition that all of the α^i be exact, hence, closed. That is, integrable sections define isotropic submanifolds.

f. Sections of target projection. A section s: $M \to J^1(\mathcal{O}; M)$ will have the local coordinate form:

$$s(x) = (t^a(x), x^i, x_a^i(x))$$

Now, in effect, for each $x \in M$ the section s associates the contact element $x_a^i(x)$ to the point $t^a(x)$ in $\mathcal{O}$.

Although the question of integrability of the section s is meaningless, nonetheless, one still derives some useful consequences from investigating the vanishing of $s^*\omega^j$ and $s^*\Theta^i$.

In the first case, one has $dt^a = (\partial t^a/\partial x^j)\, dx^j$, so:

$$s^*\omega^j = dt^i - x_a^i\left(\frac{\partial t^a}{\partial x^j}dx^j\right) = (\delta_j^i - x_a^i t_{,j}^a)dx^j .$$

Hence, the vanishing of $s^*\omega^j$ is locally equivalent to the condition:

$$\delta_j^i = x_a^i t_{,j}^a .$$

This means that the matrix $x_a^i(x)$ of the contact element at x must be a left-inverse to the matrix $t^a{}_{,j}(x)$, which is also a contact element at x, although the dimensions of the two elements are generally different. Since a left-inverse can only exist when the first map x is a linear injection and the second one t is a linear surjection, we see that in order for this condition to obtain one must have that $r \leq m$ and the rank of both maps is r.

One can then prove that the contact elements defined by both maps are transversal. Consider the situation in its most elementary form:

$$\mathbb{R}^r \xrightarrow{x} \mathbb{R}^m \xrightarrow{t} \mathbb{R}^r$$

The contact element that is defined by x is its image, while the contact element that is defined by t is its kernel. Since the composition $t \cdot x$ is non-zero everywhere except the origin, no non-zero element of the image of x can be in the kernel of t. Hence, the intersection of Im x with ker t must be 0, and, from the nullity-rank theorem, one must have:

$$\mathbb{R}^m = \mathrm{Im}\, x \oplus \ker t, \qquad \mathrm{Im}\, x \cap \ker t = 0.$$

Therefore, when $r \leq m$ and the rank of both maps is r the contact elements are transversal. This fact is at the root of the construction of the Caratheodory complete figure for more general extremal problems than geodesic curves.

As for the expression $s^*\Theta^i$, it takes the form:

$$s^*\Theta^i = \left(\frac{\partial t^a}{\partial x^j}dx^j\right)\wedge\left(\frac{\partial x_a^i}{\partial x^k}dx^k\right) = \frac{1}{2}\left(\frac{\partial t^a}{\partial x^j}\frac{\partial x_a^i}{\partial x^k} - \frac{\partial t^a}{\partial x^k}\frac{\partial x_a^i}{\partial x^j}\right)dx^j\wedge dx^k\,.$$

Hence, the vanishing of $s^*\Theta^i$ is locally equivalent to the condition:

$$0 = [x^j, x^k]^i = \frac{\partial t^a}{\partial x^j}\frac{\partial x_a^i}{\partial x^k} - \frac{\partial t^a}{\partial x^k}\frac{\partial x_a^i}{\partial x^j}\,, \qquad \text{all } i, j, k.$$

g. Sections of the contact projection. It is the sections z: $\mathcal{O}\times M \to J^1(M; \mathcal{O})$, $(t, x) \mapsto z(t, x)$ that will be the primary focus in the sequel, since geodesic fields represent special cases of such sections. Hence, we shall defer our discussion of them until a later point when we can discuss them in that context.

h. The dual jet manifold $J^1(M; \mathcal{O})$. One can just well consider 1-jets $j_x^1 t$ of C^1 maps t: $M \to \mathcal{O}$ and denote the resulting manifold of all such jets by $J^1(M; \mathcal{O})$. It will then have local coordinate systems of the form (x^i, t^a, t_i^a).

It is important to note that the fibers of the contact projection $J^1(M; \mathcal{O}) \to M\times\mathcal{O}$ over each (x, t) are affine spaces modeled on the vector space $L(\mathbb{R}^m, \mathbb{R}^r)$, which is canonically isomorphic to the dual of $L(\mathbb{R}^r, \mathbb{R}^m)$ by the map that take any linear map from $\mathbb{R}^m$ to $\mathbb{R}^r$ to its transpose map from $\mathbb{R}$ to $\mathbb{R}^m$. Indeed, since jet $j_x^1 t \in J^1(M; \mathcal{O})$ is also a linear map from T_xM to T_tO the fiber of the contact projection $J^1(M; \mathcal{O}) \to M\times\mathcal{O}$ over each (x, t) is also identified with the vector space $T_x^*M \otimes T_tM$, which maps isomorphically to the dual space $T_t^*M \otimes T_xM$ by transposition.

If x: $\mathcal{O} \to M$ is C^1 and so is t': $M \to \mathcal{O}$ then the composition $t' \cdot x$: $\mathcal{O} \to \mathcal{O}$ is also C^1. The differential map $d(t' \cdot x)|_t = dt'|_{x(t)} \cdot dx|_t$ then takes each $T_t\mathcal{O}$ to a subspace of $T_{t'}\mathcal{O}$, where $t' = t'(x(t))$. In order for the composition $t' \cdot x$ to be a diffeomorphism of $\mathcal{O}$ to itself t' would have to be a surjection and x would have to be an injection, so x would be effectively a section of t'. This would also make the differential maps invertible and the local condition for this would take the matrix form:

$$\det(t_i^a x_b^i) \neq 0.$$

We shall refer to the maps x and t' as *conjugate* whenever $t' \cdot x = I$, in which case:

$$t_i^a x_b^i = \delta_b^a\,.$$

Recall that we encountered this condition above as a transversality condition on two contact elements. Hence, one must have that $m \leq r$ in this case and the rank of both x and t' must be m.

The three canonical projections differ only in that the source and target manifolds have been exchanged. Hence, one can effectively exchange the t's with x's and the a, b, … indices with i, j, … indices. We therefore only summarize the form that the key equations of the previous section take.

For instance, since the integrability conditions for a section s: $M \to J^1(M; \mathcal{O})$, $x \mapsto (x^i, t^a(x), t_i^a(x))$ of the source projection now take the local form:

$$t_i^a = \frac{\partial t^a}{\partial x^i},$$

one sees that the corresponding 1-forms they define are:

$$\omega^a = dt^a - t_i^a dx^i, \qquad a = 1, \ldots, r.$$

The corresponding set of 2-forms is then:

$$\Theta^a = dx^i \wedge dt_i^a.$$

If a section s pulls all of the Θ^a down to 0 on M then locally one has:

$$s^*\Theta^a = dx^i \wedge s^* dt_i^a = \tfrac{1}{2}(t_{i,j}^a - t_{i.j}^a) dx^i \wedge dx^j, \qquad \text{all } a.$$

Hence, the local 1-forms on M $\tau^a = t_i^a dx^i$ make:

$$s^*\Theta^a = d\tau^a,$$

so $s^*\Theta^a$ vanishes iff τ^a is closed; once again, if s is integrable then τ^a must be exact.

A section s: $\mathcal{O} \to J^1(M; \mathcal{O})$ of the target projection has the local coordinate form:

$$s(t) = (x^i(t), t^a, t_i^a(t)).$$

Hence, the pull-down of ω^a takes the local form:

$$s^*\omega^a = dt^a - t_i^a \left(\frac{\partial x^i}{\partial t^b} dt^b\right) = (\delta_b^a - t_i^a x_{,b}^i) dt^b.$$

Its vanishing then gives the transversality condition on the relevant contact elements.

The pull-down of $s^*\Theta^a$ takes the local form:

$$s^*\Theta^a = \frac{1}{2}\left(\frac{\partial x^i}{\partial t^b}\frac{\partial t_i^a}{\partial t^c} - \frac{\partial x^i}{\partial t^c}\frac{\partial t_i^a}{\partial t^b}\right) dt^b \wedge dt^c .$$

The vanishing of $s^*\Theta^a$ is therefore locally equivalent to the condition:

$$0 = [t^b, t^c]^a = \frac{\partial x^i}{\partial t^b}\frac{\partial t_i^a}{\partial t^c} - \frac{\partial x^i}{\partial t^c}\frac{\partial t_i^a}{\partial t^b}, \qquad \text{all } a, b, c.$$

i. The case r = 1. Since the main purpose of the articles that follow is to extend the methods of the calculus of variations that were established in the case of extremal curves to extremal submanifolds of higher dimension, we should discuss the more elementary case in order to show that the general methods do indeed reduce to the more elementary ones when $r = 1$.

The first thing to address is the fact that when $r = 1$ the two jet manifolds $J^1(\mathbb{R}; M)$ and $J^1(M; \mathbb{R})$ take on simpler forms, namely, $J^1(\mathbb{R}; M)$ becomes $\mathbb{R}\times T(M)$ and $J^1(M; \mathbb{R})$ becomes $T^*M\times\mathbb{R}$. As we shall see, when the things that one defines on $J^1(\mathbb{R}; M)$ and $J^1(M; \mathbb{R})$, such as Lagrangians and Hamiltonians, are time-independent, it is customary to simply start with $T(M)$ and T^*M as the kinematical dynamical state spaces, respectively.

The canonical 1-forms on $\mathbb{R}\times T(M)$ become:

$$\omega^i = dx^i - v^i\, dt.$$

Hence, its exterior derivative is:

$$\Theta^i = d\omega^i = dt \wedge dv^i.$$

A section v: $\mathbb{R} \to \mathbb{R}\times T(M)$, $t \mapsto v(t)$:

$$v(t) = (t, x^i(t), v^i(t))$$

of the source projection is simply a vector field $v^i(t)$ along the curve $x^i(t)$, and an integrable section is a velocity vector field; i.e.:

$$v^i(t) = \frac{dx^i}{dt} .$$

When one pulls Θ^i down to $\mathbb{R}$ by way of a general section v, one gets:

$$v^*\Theta^i = \left(\frac{dx^i}{dt} - v^i\right) dt .$$

Hence, it vanishes iff $v(t)$ is a velocity vector field.

As for Θ^i, it pulls down to:

$$v^*\Theta^i = 0$$

since $\mathbb{R}$ is one-dimensional.

A section of the target projection v: $M \to \mathbb{R}\times T(M)$, $x \mapsto v(x)$, with:

$$v(x) = (t(x), x^i, v^i(x))$$

is a vector field on M, although the time function $t(x)$ is somewhat unconventional as an extra component. Its basic effect is to foliate M – or at least the open subset $U\subset M$ over which the local section v is defined – by simultaneity hypersurfaces relative to the time variable t.

The pull-down of ω^j by the section v is now:

$$v^*\omega^j = \left(\delta^i_j - v^i t_{,j}\right) dx^j .$$

It vanishes iff:

$$v^i\, t_{,j} = \delta^i_j .$$

However, this is impossible unless $m = 1$ since we are basically mapping $\mathbb{R}^m$ to $\mathbb{R}$ and then back to $\mathbb{R}^m$, which can be of rank one, at best, not rank m.

Similarly, the pull-down of Θ^i is:

$$v^*\Theta^i = \left(\frac{\partial t}{\partial x^j} dx^j\right)\wedge\left(\frac{\partial v^i}{\partial x^k} dx^k\right) = \frac{1}{2}\left(\frac{\partial t}{\partial x^j}\frac{\partial v^i}{\partial x^k} - \frac{\partial t}{\partial x^k}\frac{\partial v^i}{\partial x^j}\right) dx^j \wedge dx^k .$$

This vanishes iff:

$$[x^i, x^j] = 0,$$

with the previous notations.

When we consider the dual situation on $J(M; \mathbb{R}) = T^*M\times\mathbb{R}$, we see the familiar machinery of symplectic geometry emerge.

The canonical 1-form ω, which we now write as θ, for the sake of convention, takes the form:

$$\theta = dt - p_i\, dx^i,$$

which defines the usual contact structure on $T^*M\times\mathbb{R}$.

Its exterior derivative Θ is now:

$$\Omega = dx^i \wedge dp_i ,$$

which defines the canonical symplectic form on T^*M. That is, it is a closed 2-form that is non-degenerate, in the sense that the linear map from each tangent space $T_{(x,p)}T^*M$ to the corresponding dual cotangent space $T^*_{(x,p)}T^*M$ that takes the tangent vector $\mathbf{v} = v^i\, \partial/\partial x^i + v_i\, \partial/\partial p_i$ to the covector:

$$i_{\mathbf{v}}\Omega = - v_i\, dx^i + v^i\, dp_i$$

is a linear isomorphism.

Customarily, the contact element p_i that is associated with (t, x^i) is regarded as linear momentum, since it generally appears as the conjugate momentum to velocity in Hamiltonian mechanics. Of course, this essentially mixes kinematics with dynamics, and one finds that in geometrical optics, it is more conceptually consistent to first regard the frequency-wave number 1-form $k = k_\mu\, dx^\mu$ as the *kinematical* dual of velocity, while phase is the kinematical dual to time. However, for now, we shall simply revert to the dynamical notation.

A section $p: \mathbb{R} \to T^*M\times\mathbb{R}$, $t \mapsto p(t)$ of the source projection, with the local form:

$$p(t) = (t, x^i(t), p_i(t)),$$

is a covector field $p_i(t)$ along the curve $x^i(t)$.

When one pulls the canonical 1-form θ down to $\mathbb{R}$ by means of it, one gets:

$$p^*\theta = (1 - p_i\, v^i)\, dt,$$

which vanishes iff:

$$1 = p_i\, v^i.$$

This is simply the transversality condition that the tangent hyperplane annihilated by the covector $p_i\, dx^i$ cannot contain the velocity vector $v^i\, \partial/\partial x^i$. If one interprets the covector as linear momentum then this is also related to the requirement that the point mass that follows the curve in question must have non-vanishing kinetic energy, which takes the form $1/2\; p_i v^i$ classically.

If one pulls the canonical symplectic form Ω down to $\mathbb{R}$ then one gets zero again, since $\mathbb{R}$ is one-dimensional.

A section of the target projection $\pi: M \to T^*M\times\mathbb{R}$, $x \mapsto \pi(x)$, with the local form:

$$\pi(x) = (x^i, t(x), p_i(x)),$$

is a covector field $p_i(x)$ on M, together with a simultaneity foliation defined by the level hypersurfaces of $t(x)$.

The pull-down of θ to M by π is:

$$\pi^*\theta = \left(\frac{\partial t}{\partial x^i} - p_i\right) dx^i\,.$$

It vanishes iff the 1-form $p = p_i\, dx^i$ is exact:

$$p = dt.$$

The pull-down of Ω by π is:

$$\pi^*\Omega = \tfrac{1}{2}(p_{i,j} - p_{j,i})\, dx^i \wedge dx^j = dp.$$

It vanishes iff p is closed.

Hence, closed 1-forms on any differentiable manifold M give elementary examples of *Lagrangian submanifolds* of the symplectic manifold that is defined by T^*M, with its canonical 2-form Ω. That is, they are isotropic submanifolds of maximal dimension, namely m, and therefore maximal integral submanifolds of the differential system on T^*M that is defined by the exterior differential equation $\Omega = 0$.

In order to derive the usual Lagrange brackets of classical mechanics, one does not consider sections of any projection of $T^*M\times\mathbb{R}$, but simply submanifolds in it. Hence, let $f: N \to T^*M\times\mathbb{R}$, $u \mapsto f(u)$ be an n-dimensional submanifold in $T^*M\times\mathbb{R}$, with the local form:

$$f(u) = (x^i(u), t(u), p_i(u)).$$

Hence, it represents a covector field $p = p_i\, dx^i$ on the image of the submanifold, together with a simultaneity foliation of N by way of the function t.

The pull-back to N of θ by f takes the local form:

$$f^*\theta = dt - p_i\, dx^i = \left(\frac{\partial t}{\partial u^\alpha} - p_i \frac{\partial x^i}{\partial u^\alpha}\right) du^\alpha,$$

which vanishes iff:

$$\frac{\partial t}{\partial u^\alpha} = p_i \frac{\partial x^i}{\partial u^\alpha}.$$

The pull-back of Ω to N by f locally looks like:

$$f^*\Omega = \left(\frac{\partial x^i}{\partial u^\alpha} du^\alpha\right) \wedge \left(\frac{\partial p_i}{\partial u^\beta} du^\beta\right) = \frac{1}{2}\left(\frac{\partial x^i}{\partial u^\alpha}\frac{\partial p_i}{\partial u^\alpha} - \frac{\partial x^i}{\partial u^\beta}\frac{\partial p_i}{\partial u^\alpha}\right) du^\alpha \wedge du^\beta.$$

Hence, its vanishing is locally equivalent to the vanishing of the Lagrange brackets:

$$[u^\alpha, u^\beta] = \frac{\partial x^i}{\partial u^\alpha}\frac{\partial p_i}{\partial u^\beta} - \frac{\partial x^i}{\partial u^\beta}\frac{\partial p_i}{\partial u^\alpha}$$

for all $\alpha, \beta = 1, \dots, n$.

2. The action functional. It is conceptually valid, although not always computationally useful, to regard the calculus of variations as something like "the calculus of infinity variables." That is, if one were to make all of the analytical restrictions on the definitions that would make the space of objects in question – e.g., curves, surfaces, fields, etc. – into an infinite-dimensional differentiable manifold then the basic variational problem would take the form of a typical critical-point problem in differential calculus. That is: Given a differentiable function on the infinite-dimensional manifold, one would first form the differential of the function and look for the points of the manifold at which the differential vanished, at least when applied to the vectors in some specified subspace of the tangent space at the point in question.

This approach has been developed in the context of qualitative problems under the banner of "global analysis," which seems to have originated as "the calculus of variations in the large" in the work of Morse [**32**] on the role of topology in the structure of spaces of geodesic curves. However, even some of the classic references on Morse theory, such as Milnor [**33**], still emphasize that in order to do more tangible calculations one usually avoids the functional analytic details of setting up the infinite-dimensional manifold machinery. Some classic references in which more functional-analytic approach is taken are Ljusternik [**34**] and Morrey [**35**].

Hence, since the nature of the articles that follow is more concerned with the *calculus* of variations than the *analysis* of variations, we shall only refer to the infinite-dimensional manifold picture heuristically in order to understand the motivation for the elementary definitions. One also finds that for most of the standard problems and constructions of the calculus of variations it is sufficient to replace the infinite-dimensional manifold of all objects being varied with the finite-dimensional manifold of jets that pertain to them.

a. Action functional. The basic objects that we shall be varying are submanifolds x: $\mathcal{O} \to M$, $t \mapsto x(t)$, where $\mathcal{O}$ is a compact orientable r-dimensional differentiable manifold with boundary and M is an m-dimensional differentiable manifold. Furthermore, we shall usually assume that $r \leq m$. Hence, this class of objects includes the curve segments, bounded surfaces, and solid regions that are most commonly addressed in variational problems. Although we shall eventually discuss the use of differentiable singular k-chains in M as basic objects, which is advocated by Dedecker [**21, 22**], nonetheless, for the most elementary discussion it is sufficient to simply use the objects that we have chosen.

The "differentiable function" that one starts with is the *action functional* that associates an object x with a real number $S[x]$. The way that one gets around the infinite-dimensional details is to factor this functional through the finite-dimensional manifold $J^1(\mathcal{O}; M)$ by means of the 1-jet prolongation j^1x and define the action functional by means of an integral over the region in the parameter space $\mathcal{O}$ that x is defined over:

$$S[x] = \int_{\mathcal{O}} \mathcal{L}(j^1x)\mathcal{V} = \int_{\mathcal{O}} \mathcal{L}(t^a, x^a(t), x^i_{,a}(t))\, dt^1 \wedge \cdots \wedge dt^a ,$$

in which $\mathcal{L}$: $J^1(\mathcal{O}; M) \to \mathbb{R}$ is a C^1 function that one calls the *Lagrangian density* of the action functional and $\mathcal{V} \in \Lambda^r\mathcal{O}$ is the volume element that one has chosen for $\mathcal{O}$.

It is important to recognize that although $\mathcal{L}(j^1x)$ is a differentiable function on $\mathcal{O}$, nevertheless, it factors through a function on $J^1(\mathcal{O}; M)$ and a section of $J^1(\mathcal{O}; M) \to \mathcal{O}$, which will affect the differentiation.

An *extremal* of this action functional will then be a "point" x at which this "function" has an extremal value, such as a minimum or maximum. Since we are assuming "differentiability" of the "function," we can treat this as a problem in finding the "critical points" of the "differentiable function."

b. Equivalent Lagrangians. Since the action functional is defined by an integral over $\mathcal{O}$, its integrand is not unique. Indeed, when one specifies a fixed-boundary problem, the degree of ambiguity increases again.

Any two Lagrangian r-forms $\mathcal{L}\mathcal{V}$ and $\mathcal{L}^*\mathcal{V}$ on $J^1(\mathcal{O}; M)$ that pull down to the same r-form on $\mathcal{O}$ by means of any 1-jet prolongation j^1x: $\mathcal{O} \to J^1(\mathcal{O}; M)$ – i.e., any integrable section – will give the same value of the action functional $S[x]$ for the submanifold x: $\mathcal{O} \to M$. Since the canonical 1-forms ω^j all vanish when they are pulled down to $\mathcal{O}$ by means of an integrable section of this form, one then sees that any two Lagrangian r-forms that differ by a finite sum of exterior products of k-forms A_i on $J^1(\mathcal{O}; M)$ with the Θ^i :

$$\mathcal{L}^*\mathcal{V} - \mathcal{L}\mathcal{V} = A_i \wedge \omega^j$$

will give the same action function for integrable sections.

One can also say that replacing the $\mathcal{L}\mathcal{V}$ with:

$$\mathcal{L}^*\mathcal{V} = \mathcal{L}\mathcal{V} + A_i \wedge \omega^j$$

will not affect the action.

In particular, one can use $A_i = \Pi_i^a \,\#\partial_a$, which makes:

$$\begin{aligned}\mathcal{L}^*\mathcal{V} &= \mathcal{L}\mathcal{V} + \Pi_i^a \,\#\partial_a \wedge \omega^j \\ &= \mathcal{L}\mathcal{V} + \Pi_i^a \,\#\partial_a \wedge (dx^i - x_b^i\, dt^b) \\ &= (\mathcal{L} - \Pi_i^a\, x_a^i)\mathcal{V} + \Pi_i^a \,\#\partial_a \wedge dx^i .\end{aligned}$$

One recognizes the form of the Legendre transformation in the first term, and we shall return to this fact shortly.

A weaker equivalence condition on Lagrangians is to require that their pull-downs by any integrable section must differ by an exact r-form. Hence, since the exterior derivative operator commutes with pull-downs, one must have that:

$$\mathcal{L}^*\mathcal{V} - \mathcal{L}\mathcal{V} = dS$$

for some suitable r-form S on $J^1(\mathcal{O}; M)$.

This type of equivalence means that when one is considering a fixed-boundary problem the action functionals will agree everywhere except possibly the boundary $\partial\mathcal{O}$.

The case of interest in the sequel is when:

$$dS = \Pi_i^a \ \#\partial_a \wedge dx^i.$$

c. Variations of submanifolds. The first thing that one then does is to define the "differential" of the action functional, which is referred to as the *first variation functional* $\delta S|_x[\delta x]$. The "tangent vectors" δx that it acts on are vector fields over the submanifolds x, whose components then take the local form:

$$\delta x = \delta x^i(x(t)) \ \frac{\partial}{\partial x^i}.$$

Such a vector field δx can be regarded as the infinitesimal generator of a "differentiable" one-parameter family of *finite variations* of the initial submanifold x; i.e., a "differentiable curve" in our "differentiable manifold" of submanifolds. It is easiest to define this as a differentiable homotopy H: $[0, 1]\times\mathcal{O} \to M$, $(s, t) \mapsto H(s, t)$, that is, a differentiable map of that form with the property that $H(0, t) = x(t)$ and $H(1, t) = x'(t)$, where $x':\mathcal{O} \to M$ is some other submanifold in M. For the purposes of infinitesimal variations, its nature is irrelevant, since one usually considers only the vector field when $s = 0$, namely:

$$\delta x(x(t)) = \left.\frac{\partial H}{\partial s}\right|_{s=0} (x(t)).$$

However, the notion of finite variations is still unavoidable in what follows, since it is at the heart of the construction of the Weierstrass excess function that one uses in order to treat the sufficient condition for a strong minimum of the action functional. Hence, one might also define the vector field on the submanifold H: $[0, 1]\times\mathcal{O} \to M$:

$$\delta x(s, t) = \frac{\partial H}{\partial s}(s, x(t)).$$

We illustrate this situation in Fig. 3 schematically by representing the initial submanifold x and final submanifold x' as curves and the homotopy as a surface that is bounded by them. We emphasize that although when $r = 1$, one customarily treats only curve segments, for which ∂x has two components – viz., the initial and final point – nevertheless, when $r > 1$, the boundary of the submanifold might very well have just one component. For instance, in the isoperimetric problem, as well as the Plateau problem, one considers surfaces with a single boundary component.

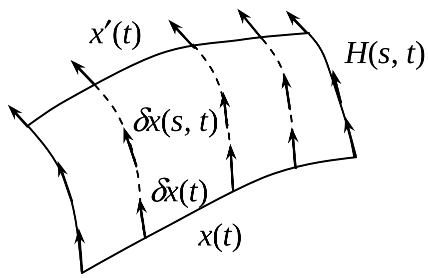

Figure 3. Finite and infinitesimal variations.

One of the aspects of finite variations that is clear in Fig. 3 is the fact that one can distinguish two basic types of finite variations:

1. Fixed-boundary variations, for which the lateral components of the boundary of $H([0, 1]\times\mathcal{O})$ contract to the boundary components $\partial x(t_0)$ and $\partial x(t_1)$.

2. Free-boundary variations, which is the general case.

The effect of fixing the boundary on the infinitesimal variation is to force $\delta x(t_0)$ and $\delta x(t_1)$ to vanish.

One then obtains the first-variation functional by means of the integral expression:

$$\delta S|_x[\delta x] = \int_{\mathcal{O}} \mathrm{L}_{\delta^1 x}(\mathcal{L}(j^i x)\mathcal{V}),$$

in which L represents the Lie derivative operator, which acts on the r-form $\mathcal{L}\mathcal{V}$ on $J^1(\mathcal{O}; M)$ in the manner that was described by Cartan:

$$\mathrm{L}_X(\mathcal{L}\mathcal{V}) = i_X d(\mathcal{L}\mathcal{V}) + d i_X(\mathcal{L}\mathcal{V}) = d\mathcal{L}(X)\mathcal{V} + d(\mathcal{L} i_X \mathcal{V}),$$

when X is a vector field on $J^1(\mathcal{O}; M)$.

In the case at hand, the vector field X takes the form of the *prolongation* $\delta^1 x$ of δx from a vector field on x to a vector field on j^1x. Once again, this is accomplished by differentiation and the local form of $\delta^1 x$ when $\delta x = \delta x^i(x(t))\,\partial_i$ is:

$$\delta^1 x = \delta x^i \frac{\partial}{\partial x^i} + \frac{d(\delta x^i)}{dt^a}\frac{\partial}{\partial x^i_a},$$

in which we are using the total derivative instead of the partial derivative.

Note that such a variation does not include a contribution from vectors tangent to $\mathcal{O}$, although when one considers symmetries of the action functional, one must indeed use variations that do include such contributions.

Since $d\mathcal{L}$ has the local form:

$$d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial t^a} dt^a + \frac{\partial \mathcal{L}}{\partial x^i} dx^i + \frac{\partial \mathcal{L}}{\partial x^i_a} dx^i_a ,$$

one can give $d\mathcal{L}(\delta^1 x)$ the local form:

$$d\mathcal{L}(\delta^1 x) = \frac{\partial \mathcal{L}}{\partial x^i} \delta x^i + \frac{\partial \mathcal{L}}{\partial x^i_a} \frac{d(\delta x^i)}{dt^a} = \frac{\delta \mathcal{L}}{\delta x^i} \delta x^i + \frac{d}{dt^a}\left(\frac{\partial \mathcal{L}}{\partial x^i_a} \delta x^i \right),$$

in which we have introduced the *variational derivative* of $\mathcal{L}$ with respect to x^i:

$$\frac{\delta \mathcal{L}}{\delta x^i} = \frac{\partial \mathcal{L}}{\partial x^i} - \frac{d}{dt^a} \frac{\partial \mathcal{L}}{\partial x^i_a} .$$

If one expands the total derivative then this takes the form:

$$\frac{\delta \mathcal{L}}{\delta x^i} = \frac{\partial \mathcal{L}}{\partial x^i} - \frac{\partial^2 \mathcal{L}}{\partial t^a \partial x^i_a} - x^j_a \frac{\partial^2 \mathcal{L}}{\partial x^j \partial x^i_a} - \frac{\partial x^j_b}{\partial t^a} \frac{\partial^2 \mathcal{L}}{\partial x^i_a \partial x^j_b} .$$

If we introduce the *generalized force* components F_i and the *conjugate momenta* 1-forms Π^a_i by way of:

$$F_i = \frac{\partial \mathcal{L}}{\partial x^i} , \qquad \Pi^a_i = \frac{\partial \mathcal{L}}{\partial x^i_a}$$

then we can also say:

$$\frac{\delta \mathcal{L}}{\delta x^i} = F_i - \frac{d\Pi^a_i}{dt^a} .$$

The momenta Π^a_i are conjugate to the generalized velocities x^i_a.

It is important to note that whether the Lagrangian density $\mathcal{L}$ is or is not a function of the t^a is irrelevant as long as one considers only variations δx that do not affect the points of $\mathcal{O}$, since the vanishing of the corresponding components of δx over $\mathcal{O}$ implies that the partial derivatives $\partial\mathcal{L}/\partial t^a$ do not appear in the final expression for the first variation $\delta S|_x[]$ when it is applied to such a δx. However, if one wishes to consider symmetries of $S[]$ then one must consider more general variations than the ones that give one the extremals themselves.

We can now express the first variation functional in the form:

$$\delta S|_x[\delta x] = \int_{\mathcal{O}} \left(\frac{\delta \mathcal{L}}{\delta x^i} \delta x^i \right) \mathcal{V} + \int_{\partial \mathcal{O}} (\Pi^a_i \delta x^i) \# \partial_i .$$

In the boundary integral, we have introduced the Poincaré duals of the coordinate vector fields:

$$\#\partial_i = i_{\partial i}\mathcal{V} = \frac{1}{(m-1)!}(-1)^i \varepsilon_{i_1\cdots i_m} dx^{i_1} \wedge \cdots \wedge \widehat{dx^i} \wedge \cdots \wedge dx^{i_m} ,$$

in which the caret signifies that the term has been suppressed from the exterior product. For instance, one has

$$\#\partial_1 = dx^2 \wedge \ldots \wedge dx^m, \qquad \#\partial_2 = -\, dx^1 \wedge dx^3 \wedge \ldots \wedge dx^m, \qquad \text{etc.}$$

It is interesting, though not directly relevant to the articles in this collection, that one can start the calculus of variations by defining a first variation functional, without the necessity of defining an action functional. This is not only analogous to the fact that not all 1-forms are exact – i.e., expressible in terms of a potential function – but actually makes it possible to treat non-conservative mechanical systems, along with ones that are subject to non-holonomic constraints, variationally, which is not usually possible in terms of action functionals. For a more detailed discussion of this, one can confer the author's papers [**36, 37**].

3. **Extremal submanifolds**. Now that we have a useful way of expressing the first variation functional for the purpose of local calculations, we can return to the basic critical point problem. Naively, one desires to find those submanifolds x for which the first variation functional $\delta S|_x[]$ vanishes.

However, this is usually overly general, since the specific problems usually involve first specifying whether one is varying the boundary ∂x. Hence, one then specifies that $\delta S|_x[\delta x]$ must vanish for some subspace of vector fields δx on x. If one considers a fixed-boundary problem then δx vanishes on the boundary points, and the boundary integral vanishes in the first variation functional. However, for a free-boundary problem, in order to make the boundary integral vanish one must use only variations that satisfy the *transversality* condition:

$$\Pi^a_i \delta x^i = 0, \qquad a = 1, \ldots, r$$

on ∂x.

a. Euler-Lagrange formulation. In either case, the only remaining contribution to the first variation is the first integral:

$$\delta S|_x[\delta x] = \int_{\mathcal{O}} \left(\frac{\delta \mathcal{L}}{\delta x^i} \delta x^i \right) \mathcal{V} ,$$

and if it is to vanish for all δx^i that make the boundary integral disappear then x must satisfy the *Euler-Lagrange equations:*

$$\frac{\delta \mathcal{L}}{\delta x^i} = 0,$$

which can also be given the form:

$$F_i = \frac{d\Pi_i^a}{dt^a},$$

which is essentially a balance law for the generalized momenta.

This system of partial differential equations for $x^i(t)$ is generally nonlinear and of second order when one considers only 1-jets. More generally, when one goes to k-jets, which corresponds to higher-order Lagrangians, as well, one obtains partial differential equations of order k+1.

In order to motivate the material of the next section, we point out that one can characterize an extremal manifold by the condition that:

$$j^1x^*[i_{\delta^1 x}d(\mathcal{L}\mathcal{V})] = 0$$

for every integrable vector field $\delta^1 x$ on $J^1(\mathcal{O}; M)$ that is vertical for the source projection.

b. Extremal submanifolds: Hamilton-Cartan formulation. Returning to the conjugate momenta Π_i^a, one sees that what they really represent are the components of a one-form on $J^1(\mathcal{O}; M)$, namely, the vertical part $d_V\mathcal{L}$ of the 1-form $d\mathcal{L}$ relative to the contact projection.

Recall that the vertical sub-bundle $V(J^1)$ of the tangent bundle $T(J^1)$ relative to that projection consists of all tangent vectors to the manifold $J^1(\mathcal{O}; M)$ that project to zero under the differential map to the projection $J^1(\mathcal{O}; M) \to \mathcal{O}\times M$. Vertical vectors will then be tangent to the fibers of the projection, so in terms of local coordinates, they will take the form:

$$V = X_a^i(t^b, x^j, x_b^j)\frac{\partial}{\partial x_a^i}.$$

Although one cannot canonically *project* an arbitrary vector field X on $J^1(M; \mathcal{O})$ onto its vertical part without defining a complementary "horizontal" sub-bundle to $V(J^1)$, nevertheless, one can *restrict* the 1-form $d\mathcal{L}$ to the vertical subspaces without making such a definition. Thus, one can regard the vertical differential $d_V\mathcal{L}$ as a section $d_V\mathcal{L}$: $J^1(M; \mathcal{O}) \to V^*(J^1)$ of the projection $V^*(J^1) \to J^1(M; \mathcal{O})$; hence, to every 1-jet $p\in J^1(M; \mathcal{O})$ one associates the 1-form $d_V\mathcal{L}|_p$ on the vertical vector space $V_p(J^1)$.

The conjugate momenta then take the local form:

$$\Pi(t, x, x_a^i) = \Pi_i^a(t^b, x^j, x_b^j)dx_a^i \equiv \frac{\partial\mathcal{L}}{\partial x_a^i}(t^b, x^j, x_b^j)dx_a^i.$$

When one considers the variational problem with $r > 1$, one first sees that the duality that one must now address is between $J^1(\mathcal{O}; M)$ and $J^1(M; \mathcal{O})$. However, one must also now consider sections π: $\mathcal{O} \to J^1(M; \mathcal{O})$ of the *target* projection $J^1(M; \mathcal{O}) \to \mathcal{O}$, which will then have the local form:

$$\pi(t) = (x^i(t), t^a, \pi_i^a(t)).$$

We now wish to define a Lagrangian density $\mathcal{L}^*$ on $J^1(M; \mathcal{O})$ such that one can define an action functional on submanifolds x: $\mathcal{O} \to M$ that agrees with an action functional that is defined by some Lagrangian density $\mathcal{L}$ on $J^1(\mathcal{O}; M)$.

However, since the notion of integrability does not mean anything for sections of the target projection, we can see that we also need to address what it means for a section, such as π, to correspond to some integrable section j^1x of $J^1(\mathcal{O}; M) \to \mathcal{O}$.

This is where the previous discussions of the canonical maps associated with the fibers of the two contact projections $J^1(\mathcal{O}; M) \to \mathcal{O}\times M$ and $J^1(M; \mathcal{O}) \to M\times\mathcal{O}$ proves to be most useful. As we said, the fiber of the former projection over (t, x) is canonically identified with $T_t^*\mathcal{O} \otimes T_x M$. Now, the vertical derivative $\Pi = d_V\mathcal{L}$ takes any point p of that fiber to the point $\Pi(p) = d_V\mathcal{L}|_p$ of the vector space $V_p^* J^1$.

If this map is invertible then we can think of it as a linear isomorphism of $T_t^*\mathcal{O} \otimes T_x M$ with $V_p^* J^1$. This implies a restriction in the choice of Lagrangian, since not all Lagrangians will satisfy that condition, but only the ones that satisfy the local condition that:

$$\det\left[\frac{\partial^2\mathcal{L}}{\partial x_a^i\,\partial x_b^j}\right] \neq 0.$$

If the Lagrangian permits this then since $V_p^* J^1$ would be isomorphic to the dual of $T_t^*\mathcal{O} \otimes T_x M$ – namely, $T_x^* M \otimes T_t\mathcal{O}$ – by transposition, one sees that such a Lagrangian allows one to define an diffeomorphism of λ: $J^1(\mathcal{O}; M) \to J^1(M; \mathcal{O})$ that preserves the contact projections and takes each $j_t^1 x \in J^1(\mathcal{O}; M)$ to the element of $J^1(M; \mathcal{O})$ that corresponds to $\Pi(j_t^1 x)$. This map takes the local form:

$$\lambda(t^a, x^i, x_a^i) = (t^a, x^i, \Pi_i^a(t^b, x^j, x_b^j)).$$

It is essential that this map be invertible, since the way that one defines a Lagrangian $\mathcal{L}^*\mathcal{V}$ on $J^1(M; \mathcal{O})$ that corresponds to the Lagrangian density $\mathcal{L}\mathcal{V}$ on $J^1(\mathcal{O}; M)$ is by pulling it back along λ^{-1}:

$$\mathcal{L}^* = \lambda^{-*}\mathcal{L} = \mathcal{L} \cdot \lambda^{-1},$$

which takes the local form:

$$\mathcal{L}^*(x^i, t^a, \pi_i^a) = \mathcal{L}(t^a, x^i, x_a^i(x^j, t^b, \pi_j^b)).$$

We further define the C^1 function $\mathcal{H}$: $J^1(M;\mathcal{O}) \to \mathbb{R}$, which one calls a *Hamiltonian density* by the *Legendre transform* of $\mathcal{L}$, which takes the form:

$$\mathcal{H}\mathcal{V} = \pi_i^a\, dx^i \wedge \#\partial_a - \mathcal{L}^*\mathcal{V}.$$

Hence, one can put the Lagrangian density on $J^1(M;\mathcal{O})$ into the local form:

$$\theta = \mathcal{L}^*(x^i, t^a, \pi_i^a)\mathcal{V} = \pi_i^a\, dx^i \wedge \#\partial_a - \mathcal{H}(x^i, t^a, \pi_i^a)\mathcal{V} = \theta^a \wedge \#\partial_a,$$

in which the 1-forms:

$$\theta^a = \pi_i^a\, dx^i - \mathcal{H}\, dt^a$$

will serve as the r-dimensional equivalents of the Poincaré-Cartan 1-form of point mechanics.

One then finds that the differential equations for the submanifold x can be obtained by making our condition for an extremal in terms of the section π: $\mathcal{O} \to J^1(M;\mathcal{O})$ that corresponds to j^1x. One requires that for all vector fields X on $J^1(M;\mathcal{O})$ that are vertical form the target projection, and will then have the local form:

$$X = X^i \frac{\partial}{\partial x^i} + X_i^a \frac{\partial}{\partial \pi_i^a},$$

one will have:

$$\pi^*(i_X d\theta) = 0.$$

Now:

$$\begin{aligned} d\theta &= d\pi_i^a \wedge dx^i \wedge \#\partial_a - d\mathcal{H} \wedge \mathcal{V} \\ &= d\pi_i^a \wedge dx^i \wedge \#\partial_a - \frac{\partial \mathcal{H}}{\partial x^i}\, dx^i \wedge \mathcal{V} - \frac{\partial \mathcal{H}}{\partial \pi_i^b}\, d\pi_i^b \wedge \mathcal{V}. \end{aligned}$$

Thus, we have:

$$i_X d\theta = -\left(d\pi_i^a + \frac{\partial H}{\partial x^i} dt^a\right) \wedge \#\partial_a + X_i^a \left(dx^i \wedge \#\partial_a - \frac{\partial H}{\partial \pi_i^a}\mathcal{V}\right).$$

In order to pull this down to $\mathcal{O}$ one substitutes:

$$d\pi_i^b = \frac{\partial \pi_i^b}{\partial t^c} dt^c.$$

One gets:

$$\pi^* i_X d\theta^a = \left[-X^i\left(\frac{\partial \pi_i^a}{\partial t^a} + \frac{d\mathcal{H}}{dx^i}\right) + X_i^a\left(\frac{\partial x^i}{\partial t^a} - \frac{d\mathcal{H}}{d\pi_i^a}\right)\right]\mathcal{V}.$$

If this vanishes for all X^i and X_i^a then one must have:

$$\frac{\partial x^i}{\partial t^a} = \frac{\partial \mathcal{H}}{\partial \pi_i^a}, \qquad \frac{\partial \pi_i^a}{\partial t^a} = -\frac{\partial \mathcal{H}}{\partial x^i},$$

which constitute a system of partial differential equations for x^i and π_i^a that generalize the canonical ordinary differential equations of Hamiltonian mechanics.

Note that actually the issues of integrability for both the section π and the vector field X did not figure in the derivation of these extremal equations.

c. Extremal curves. In the case of point mechanics, for which $r = 1$, and when one restricts oneself to time-invariant Lagrangians, one can replace $J^1(\mathcal{O}; M)$ with $T(M)$ and $J^1(M; \mathcal{O})$ with T^*M. In order to associate a Lagrangian $\mathcal{L}$ on $T(M)$ with a Hamiltonian H on T^*M, one must first specify an isomorphism of $T(M)$ with T^*M if one is to define the Legendre transformation. One starts by noting that the conjugate momentum p, as we have defined it, is a covector field on $T(M)$, not on M. If one has a vector field $\mathbf{v}$: $M \to T(M)$ then one can pull p down to a 1-form $\mathbf{v}^*p$ on M by means of $\mathbf{v}$. Its local expression is then:

$$\mathbf{v}^*p(x) = p_i(x^j, v^j(x))\, dx^i.$$

Hence, we can define a map π: $\mathfrak{X}(M) \to \Lambda^1 M$, $\mathbf{v} \mapsto \mathbf{v}^*p$ from vector fields on M to covector fields on M. By the inverse function theorem, a necessary and sufficient condition for its local invertibility is given by the invertibility of the matrix:

$$\gamma_{ij} = \frac{\partial p_i}{\partial v^j} = \frac{\partial^2 \mathcal{L}}{\partial v^i \partial v^j}.$$

In the case of Newtonian point mechanics in a Euclidian space, this matrix will take the form $m\delta_{ij}$ when $\mathcal{L}$ depends upon v^i only by way of the kinetic energy 1/2 $m\delta_{ij}\, v^i\, v^j$. Hence, it is conformal to the Euclidian metric. This latter restriction on $\mathcal{L}$ is also fundamental aspect of Finsler geometry, which uses the definition of $\mathcal{L}$ as the basis for the rest of the geometry, which has an unavoidably variational flavor to it. (See, for instance, Rund [**38**] or Bao, Chern, and Shen [**39**].)

Once one has an isomorphism of velocity vector fields with momentum covector fields, one can define a Hamiltonian function H on T^*M that corresponds to $\mathcal{L}$ by means of the Legendre transformation:

$$H(x^i, p_i) = p_i\, v^i(x, p) - \mathcal{L}(x^i, v^i(x, p)).$$

When $\mathbf{v}$ is integrable – so $v^i = dx^i/dt$ – if one multiplies both sides of this equation times the 1-form dt then one gets the one form:

$$H\,dt = p_i\,dx^i - \mathcal{L}\,dt,$$

which means the integrand of the action functional takes the form:

$$\mathcal{L}\,dt = p_i\,dx^i - H\,dt,$$

Hence, one can also express the action functional in terms of sections of $T^*M \to M$ by way of:

$$S[x] = \int_{[0,1]} [p_i(t)v^i(p(t)) - H(t, x^i(t), p_i(t))]\,dt\,.$$

The 1-form on $T^*M\times\mathbb{R}$:

$$\theta = p_i\,dx^i - H\,dt$$

represents the *Poincaré-Cartan 1-form*, since the first term was discussed by Poincaré in the context of integral invariants for autonomous mechanical systems, while Cartan [**10**] showed that one could formulate the least-action principle in a third formalism (besides the Euler-Lagrange and Hamiltonian formalisms) by demanding that the 1-form θ be a "relative integral invariant" of the motion; i.e., its exterior derivative:

$$d\theta = dp_i \wedge dx^i - dH \wedge dt$$

would be an "absolute integral invariant." The equations of motion then followed by considering the "characteristic system" of the exterior differential system on $T^*M\times\mathbb{R}$:

$$d\theta = 0.$$

This is obtained by solving the equation:

$$i_X d\theta = 0.$$

for the vector field X on $T^*M \times\mathbb{R}$ and then finding its integral curves.

One should note that this equation can also be expressed as:

$$i_X(dp_i \wedge dx^i) = i_X(dH \wedge dt) = i_X\,dH \wedge dt - X^t\,dH,$$

which can be put into the form:

$$i_X\Omega = i_X\,dH \wedge dt - X^t\,dH,$$

if one defines the canonical symplectic form on T^*M:

$$\Omega = d\theta = dp_i \wedge dx^i.$$

A typical vector field X on $T^*M\times\mathbb{R}$ has the local form:

$$X = X^i \frac{\partial}{\partial x^i} + X_i \frac{\partial}{\partial p_i} + X^t \frac{\partial}{\partial t}.$$

Now, since:

$$dH = \frac{\partial H}{\partial x^i} dx^i + \frac{\partial H}{\partial p_i} dp_i ,$$

we get:

$$d\theta = dp_i \wedge dx^i - \frac{\partial H}{\partial x^i} dx^i \wedge dt - \frac{\partial H}{\partial p_i} dp_i \wedge dt .$$

This makes:

$$i_X d\theta = -\left(\frac{\partial H}{\partial x^i} X^i + \frac{\partial H}{\partial p_i} X_i\right) dt + \left(X_i + \frac{\partial H}{\partial p_i} X^t\right) dx^i - \left(X^i - \frac{\partial H}{\partial x^i} X^t\right) dp_i .$$

If this vanishes then one must have the characteristic system:

$$\frac{\partial H}{\partial x^i} X^i + \frac{\partial H}{\partial p_i} X_i = 0, \qquad X^i = \frac{\partial H}{\partial p_i} X^t , \qquad X_i = -\frac{\partial H}{\partial x^i} X^t .$$

One immediately sees that if the last two sets of equations are valid then the first set is automatically satisfied. Hence, if one seeks the integral curves of the characteristic vector field X for the 2-form $d\theta$ then one obtains the system of ordinary differential equations:

$$\frac{dH}{dt} = 0, \qquad \frac{dx^i}{ds} = \frac{\partial H}{\partial p_i} X^t , \qquad \frac{dp_i}{ds} = -\frac{\partial H}{\partial x^i} X^t .$$

Except for the factor of X^t, these are essentially Hamilton's equations, combined with the requirement that the Hamiltonian be constant along the integral curves. This also represents a restriction on the possible choices of Hamiltonian.

In order to account for the factor of X^t, which is arbitrary, but non-zero, since the characteristic equations are homogeneous in X, one must regard it as the derivative dt/ds that is associated with a change of parameterization for the integral curves. One then can put the last two sets of equations into the form:

$$\frac{dx^i}{dt} = \frac{\partial H}{\partial p_i} , \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x^i} ,$$

which has the customary form.

The factor of X^t did not appear in the previous section since we were restricting ourselves to vector fields for which it vanished. Furthermore, if one treats only the time-invariant case, so H is a function T^*M, one will not have to consider it in that case, either.

One can specialize the above argument in that case by defining H as a differentiable function on T^*M and then defining the *Hamiltonian vector field* X_H that is associated with by:

$$i_{X_H} d\theta = - dH,$$

which is then the previous condition if one sets $X^t = 1$.

4. Geodesic fields. There is a fundamental difference between saying that a particular submanifold in a space, such as a curve segment or surface, is extremal for a particular action functional and saying that a region of that same space is *foliated* by extremal manifolds. For instance, there is a difference between finding a geodesic between two points in space and finding a geodesic congruence that foliates a region of that space with one-dimensional leaves that consist of geodesics. Hence, in the older literature one often finds the term "field of extremals" used to mean "foliation;" that is, an $(m - r)$-parameter family of r-dimensional extremals in an m-dimensional manifold M that partitions an m-dimensional region of M.

Basically, the key issue is one of integrability in reference to the action functional. In order for an exterior differential system on a manifold M to define a foliation, it must be completely integrable, in the Frobenius sense. Thus, if the exterior differential system is defined by $\omega = 0$, where ω is an r-form on M, then in order this system to be completely integrable, in the Frobenius sense, one must have that $\omega \wedge d\omega$ vanishes identically. It is, of course, sufficient that ω be closed, which is locally equivalent to the condition that it be exact.

In the study of geodesic fields, the r-form one is concerned with is obtained by first pulling down either the fundamental Lagrangian p-form $\Omega = \mathcal{L}\ \mathcal{V}$ on $J^1(\mathcal{O}; M)$, or one that is Lepage-congruent to it, to a r-form on $\mathcal{O} \times M$ by means of a field z of contact elements. Such a field is then geodesic, in various senses, when $[\Omega] = z^* \Omega$ is an exact r-form. One then finds oneself concerned with the Hamilton-Jacobi equation, originally in Lagrangian form, but also in Hamiltonian form, after a Legendre transformation.

a. Fields of contact elements. Previously, we did not address the issue of sections of the contact projection $J^1(\mathcal{O}; M) \to \mathcal{O}\times M$. That is not because it is insignificant, but just the opposite. The sections of that projection, which we shall call *fields of contact elements*, play the central role in the articles that follow since a geodesic field is a special type of field of contact elements. We caution the reader that in the literature the term *slope field* is used, to be consistent with the terminology used in the study of extremal curves, but since "slope" suggests that we are only using a one-dimensional parameter space $\mathcal{O}$ whose contact elements to M are tangent lines we shall modernize the terminology.

Hence, a field of contact elements is a section z: $\mathcal{O}\times M \to J^1(\mathcal{O}; M)$, $(t, x) \mapsto z(t, x)$, with the local form:

$$z(t, x) = (t^a, x^i, x^i_a(t, x)).$$

In the case of $r = 1$, one can regard a field of contact elements as a time-varying vector field on M. More generally, it is an r-parameter family of $r'(t, x)$-planes in T_xM, which will be r-planes if and only if one restricts the jets to 1-jets of immersions.

A submanifold $x: \mathcal{O} \to M$ is said to be *embedded* in the field of contact elements z iff the following diagram commutes:

$$\begin{tikzcd}
& J^1(\mathcal{O}; M) & \\
\mathcal{O} \arrow[ur, "j^1x"] \arrow[rr, "\text{graph } x"] & & \mathcal{O}\times M \arrow[ul, "z"']
\end{tikzcd}$$

That is:

$$j^1x = z \cdot (\text{graph } x).$$

In local form, one has:

$$(t^a, x^i(t), x^i_{,a}(t)) = (t^a, x^i(t), x^i_a(t, x)).$$

We illustrate the general case of a field of contact elements and the embedding of a submanifold schematically in Fig. 4.

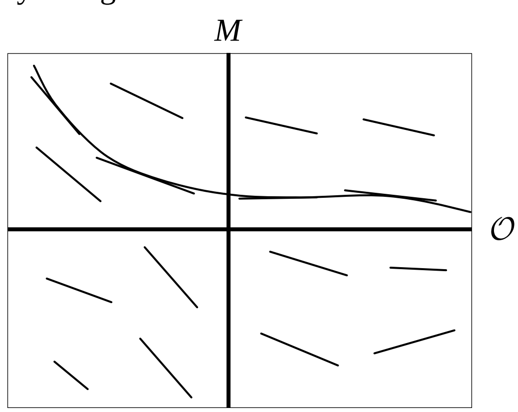

Figure 4. General field of contact elements with a submanifold embedded in it.

Now, the local condition on the fields is:

$$\frac{\partial x^i}{\partial t^a}(t) = x^i_a(t, x^i(t)),$$

so if one is given the field of contact elements, *a priori*, then the integrability condition for it to admit an embedded submanifold is obtained by differentiating both sides with respect to t^b and demanding the symmetry of the mixed partial derivatives. However, it is crucial to note that the differentiation on the right-hand side becomes a *total derivative* with respect to t^a:

$$\frac{dx_a^i}{dt^b} = \frac{\partial x_a^i}{\partial t^b} + \frac{\partial x^j}{\partial t^b}\frac{\partial x_a^i}{\partial x^j} = \frac{\partial x_a^i}{\partial t^b} + x_b^j \frac{\partial x_a^i}{\partial x^j}.$$

Note that although we started by taking a total derivative along a submanifold, nevertheless, the final expression does not depend upon a submanifold for its definition. Hence, the second equality is really a *replacement* of the one expression with the other.

The resulting integrability condition on the field of contact elements $x_a^i(t, x^i(t))$ is:

$$0 = \frac{\partial x_a^i}{\partial t^b} - \frac{\partial x_b^i}{\partial t^a} + x_b^j \frac{\partial x_a^i}{\partial x^j} - x_a^j \frac{\partial x_b^i}{\partial x^j}.$$

If one pulls down the canonical 1-forms ω^j to $\mathcal{O}\times M$ by means of z then the result is:

$$z^* \omega^j = dx^i - x_a^i(t, x)\, dt^a.$$

Since this can never vanish for any field of contact elements $x_a^i(t, x)$, one sees that fields of contact elements cannot be integral submanifolds of the exterior differential system $\Theta^i = 0$.

The pull-down of Θ^i takes the form:

$$z^*\Theta^i = dt^a \wedge \left(\frac{\partial x_a^i}{\partial t^b} dt^b + \frac{\partial x_a^i}{\partial x^j} dx^j \right) = \frac{1}{2}\left(\frac{\partial x_a^i}{\partial t^b} - \frac{\partial x_b^i}{\partial t^a} \right) dt^a \wedge dt^b + \frac{\partial x_a^i}{\partial x^j} dt^a \wedge dx^j.$$

In order for this to vanish, one must have:

$$\frac{\partial x_a^i}{\partial t^b} = \frac{\partial x_b^i}{\partial t^a}, \qquad \frac{\partial x_a^i}{\partial x^j} = 0.$$

Hence, such a field of contact elements will also satisfy the integrability condition given above, and a field of contact elements is isotropic only if it admits an embedded submanifold.

As we saw above, an extremal submanifold x: $\mathcal{O} \to M$ must satisfy the condition:

$$j^1 x^*[i_{\delta^1 x} d\theta] = 0$$

for all vector fields $\delta^1 x$ on $J^1(\mathcal{O}; M)$ that represent the prolongations of vector fields δx on x. In this expression:

$$\theta = \mathcal{L}\mathcal{V} + \Pi_i^a dx^i \wedge \#\partial_a = (\mathcal{L}\, dt^a + \Pi_i^a dx^i) \wedge \#\partial_a$$

is the Lagrangian form of the Poincaré-Cartan r-form.

One can also formulate the condition in terms of the Hamilton-Cartan form on $J^1(M; \mathcal{O})$:

$$\theta = \Pi^a_i dx^i \wedge \#\partial_a - \mathcal{H}\,\mathcal{V}.$$

which corresponds to the Lagrangian form under the Legendre transformation.

The condition is then that for an section π: $\mathcal{O} \to J^1(M; \mathcal{O})$ that corresponds to j^1x under Legendre transformation one must have that:

$$\pi^* i_X d\theta = 0$$

for all vector fields X on $J^1(M; \mathcal{O})$.

b. Extremal fields of contact elements. Now, suppose z: $\mathcal{O}\times M \to J^1(\mathcal{O}; M)$ is a field of contact elements, so a submanifold x: $\mathcal{O} \to M$ is embedded in z iff $j^1x = z \cdot x$. (Here, we abbreviate the notation by replacing the reference to the graph of x with simply a reference to the map x.) Once again, we let $\Omega = \mathcal{L}\,\mathcal{V}$ denote the fundamental r-form on $J^1(\mathcal{O}; M)$ that is defined by a choice of Lagrangian density $\mathcal{L}$ and let $\delta^1 x$ be the 1-jet prolongation of a vector field δx on x.

Hence, the integrand in the first variation functional $\delta S|_x[\delta x]$ initially takes the form:

$$(j^1x)^*[\, i_{\delta^1 x} d\Omega\,] = x^*(z^*\, i_{\delta^1 x} d\Omega\,).$$

We shall call a field of contact elements z *extremal* iff $z^*\Omega$ is a closed r-form on $\mathcal{O}\times M$; hence:

$$d(z^*\Omega) = z^* d\Omega = 0.$$

In the articles that follow, one often sees the notation $[\Omega] = z^*\Omega$ used, so the latter condition takes the form:

$$d[\Omega] = 0.$$

This clearly represents a stronger condition than the one that makes a submanifold extremal by itself, since one is now essentially defining a foliation of $\mathcal{O}\times M$ by extremal submanifolds that are integral submanifolds of the exterior differential system $[\Omega] = 0$ on $\mathcal{O}\times M$. Such a foliation is sometimes referred to as a *Mayer field* in the articles that follows, which is one example of using the term "field" to refer to foliations, as well as the differential forms that define them.

One sees that it will automatically follow from the previous condition on z that any submanifold that is embedded in an extremal field of contact elements must be an extremal submanifold, since one will have:

$$(j^1x)^* d\Omega = (z \cdot x)^* d\Omega = x^*(z^* d\Omega) = 0.$$

We can examine the local form of the differential equations for an extremal field of contact elements in either Lagrangian or Hamiltonian form. In Lagrangian form, we start with:

$$d\Omega =$$
$$\left(\frac{\partial\mathcal{L}}{\partial x^i}-\frac{\partial\Pi_i^a}{\partial t^a}\right)dx^i\wedge\mathcal{V}+\Pi_i^a dx_a^i\wedge\mathcal{V}-\tfrac{1}{2}(\Pi_{i,j}^a-\Pi_{j,i}^a)dx^i\wedge dx^j\wedge\#\partial_a+\frac{\partial\Pi_i^a}{\partial x_b^j}dx_b^j\wedge dx^i\wedge\#\partial_a\,.$$

In order to pull this down to $\mathcal{O}\times M$ by means of a Lagrangian field of contact elements z, one must make the replacement:

$$dx_a^i=\frac{\partial x_a^i}{\partial t^b}dt^b+\frac{\partial x_a^i}{\partial x^j}dx^j\,.$$

This leads to:

$$z^*d\Omega=\left(\frac{\delta\mathcal{L}}{\delta x^i}+\Pi_j^a\frac{\partial x_a^j}{\partial x^i}\right)dx^i\wedge\mathcal{V}-\frac{1}{2}\left(\frac{d\Pi_i^a}{dx^j}-\frac{d\Pi_j^a}{dx^i}\right)dx^i\wedge dx^j\wedge\#\partial_a\,,$$

in which we have introduced the total derivative of Π_i^a with respect to x^i:

$$\frac{d\Pi_i^a}{dx^j}=\frac{\partial\Pi_i^a}{\partial x^j}+\frac{\partial x_b^j}{\partial x^j}\frac{\partial\Pi_i^a}{\partial x_b^j}\,.$$

If $z^*d\Omega$ is to vanish identically then one must have:

$$\frac{\delta\mathcal{L}}{\delta x^i}+\Pi_j^a\frac{\partial x_a^j}{\partial x^i}=0,\qquad\frac{d\Pi_i^a}{dx^j}=\frac{d\Pi_j^a}{dx^i}\,.$$

When a submanifold is embedded in an extremal field of contact elements, the first set of equations reduces to the customary Euler-Lagrange equations, since the supplementary term will vanish. The second set of equations then represents an integrability condition on the conjugate momenta if they are to take the form:

$$\Pi_i^a=\frac{dS^a}{dx^i}\,.$$

In Hamiltonian form, one starts with Ω in the Hamilton-Cartan form then one has:

$$d\Omega=d\Pi_i^a\wedge dx^i\wedge\#\partial_a-\frac{\partial\mathcal{H}}{\partial x^i}dx^i\wedge\mathcal{V}-\frac{\partial\mathcal{H}}{\partial\Pi_i^a}d\Pi_i^a\wedge\mathcal{V}\,.$$

In order to pull it down by means of a Hamiltonian field of contact elements π: $\mathcal{O}\times M \to J^1(M; \mathcal{O})$, one must make the replacement:

$$d\Pi_i^a = \frac{\partial \Pi_i^a}{\partial t^b} dt^b + \frac{\partial \Pi_i^a}{\partial x^j} dx^j .$$

This makes:

$$\pi^* d\Omega = -\left(\frac{\partial \Pi_i^a}{\partial t^a} + \frac{\partial \mathcal{H}}{\partial x^i} + \frac{\partial \mathcal{H}}{\partial \Pi_i^b} \frac{\partial \Pi_i^b}{\partial x^j} \right) dx^i \wedge V - \frac{1}{2} (\Pi_{i,j}^a - \Pi_{j,i}^a) dx^i \wedge dx^j \wedge \# \partial_a .$$

If this vanishes then one must have the system of equations:

$$\frac{\partial \Pi_i^a}{\partial t^a} = -\frac{d\mathcal{H}}{dx^i}, \qquad \frac{\partial \Pi_i^a}{\partial x^j} = \frac{\partial \Pi_j^a}{\partial x^i} .$$

If x: $\mathcal{O} \to M$ is embedded in π then one will have:

$$x_a^i = \frac{\partial x^i}{\partial t^a} = \frac{\partial \mathcal{H}}{\partial \Pi_i^a} ,$$

which gives the set of canonical equations, while its substitution in the previous system of equations puts them into the form:

$$\frac{d\Pi_i^a}{dt^a} = -\frac{\partial \mathcal{H}}{\partial x^i} ,$$

which is then the other set of canonical equations.

c. Geodesic fields. In the previous subsection, we examined the nature of fields of contact elements that made the r-form Ω closed when one pulled it down to $\mathcal{O}\times M$. In order to define *geodesic fields*, as they are treated in the following papers, one must strengthen this to the requirement that the pull-down of Ω must be exact. That is, a Lagrangian field of contact elements z: $\mathcal{O}\times M \to J^1(\mathcal{O}; M)$ is a *geodesic field* if there exists an $(r - 1)$-form S on $\mathcal{O}\times M$ such that:

$$z^*\Omega = dS.$$

Obviously, the $(r - 1)$-form S is not unique, but defined only up to a closed $(r - 1)$-form.

Of course, since one is dealing with r-forms on the manifold $\mathcal{O}\times M$, if the de Rham cohomology in dimension r vanishes then all closed forms will be exact, anyway.

Part of the motivation for the requirement of exactness was based in the desire that the action functional be independent of the choice of submanifold x, but dependent only upon the choice of its boundary ∂x. From Stokes's theorem, this is precisely what exactness of an r-form gets one:

$$S[x] = \int_x (j^1x)^*\Omega = \int_x (j^1x)^* dS = \int_x d(j^1x^*S) = \int_{\partial x} (j^1x)^* S\,.$$

First, we observe that locally one has:

$$dS = d(S^a \wedge \#\partial_a) = \frac{\partial S^a}{\partial t^a}\mathcal{V} + \frac{\partial S^a}{\partial x^i}dx^i \wedge \#\partial_a\,.$$

One also has:

$$z^*\Omega = \mathcal{L}(t, x, z(t, x))\mathcal{V} + \Pi_i^a dx^i \wedge \#\partial_a\,.$$

If the two expressions are identical then:

$$\mathcal{L}(t, x, z(t, x)) = \frac{\partial S^a}{\partial t^a}, \qquad \Pi_i^a = \frac{\partial S^a}{\partial x^i}\,.$$

If x: $\mathcal{O} \to M$ is embedded in z then one must replace $s^a(t, x)$ with $s^a(t, x(t))$ and the partial derivatives with respect to t^a with total derivatives, so one obtains the condition on $\mathcal{L}$ that:

$$\mathcal{L}(z \cdot x) = \frac{dS^a}{dt^a},$$

which explains Weyl's terminology "Lagrangians of divergence type."

In Hamiltonian form, one uses:

$$z^*\Omega = [\pi_i^a(t,x)dx^i - \mathcal{H}(x, t, \pi(t, x))] \wedge \#\partial_a\,,$$

in which π: $\mathcal{O}\times M \to J^1(M; \mathcal{O})$ is a Hamiltonian field of contact elements.

Equating this with dS gives the system of partial differential equations for S^a:

$$\frac{\partial S^a}{\partial t^a} = -\mathcal{H}(x, t, \pi(t, x)), \qquad \pi_i^a = \frac{\partial S^a}{\partial x^i}\,.$$

These equations are clearly of a generalized Hamilton-Jacobi type, which becomes more evident when one embeds a submanifold x in the field of contact elements π, which involves replacing the partial derivatives with respect to t^a with total derivatives:

$$\frac{dS^a}{dt^a} = -H\left(x(t), t, \frac{\partial S^a}{\partial t^i}(t, x(t))\right), \qquad \pi_i^a = \frac{\partial S^a}{\partial x^i}\,.$$

If one solves an initial-value problem for the first equation for the S^a then the second equation serves to define the geodesic field by differentiation. In Weyl's article [**9**], he discusses the problem of solving the equation by Cauchy's method of characteristics, as well the method of majorants.

d. Lepage congruences, De Donder-Weyl, and Carathéodory fields. So far, we have been considering $\Omega = \mathcal{L}\,\mathcal{V}$ to be the fundamental r-form for any variational problem involving extremal r-dimensional submanifolds of an m-dimensional manifold M. However, since $s^*\omega^j = 0$ for any integrable section $s: \mathcal{O} \to J^1(\mathcal{O}; M)$, one sees that as long as s is integrable $s^*\Omega'$ will be the same for any other r-form Ω' on $J^1(\mathcal{O}; M)$ that is congruent to W modulo the ideal in the exterior algebra that is generated by the fundamental 1-forms ω^j. Similarly, if $d\Omega = 0$ then $d\Omega'$ will be congruent to 0 (mod ω^j).

The congruences that Lepage defined pertained to the definition of geodesic fields, in which one weakens the definition of the pull-down $[\Omega]$ and the condition on $d[\Omega]$ so they would be valid for all r-forms that are congruent to $\Omega = \mathcal{L}\,\mathcal{V}$ (mod ω^j):

$$\Omega \equiv \mathcal{L}\,\mathcal{V}\ (\mathrm{mod}\ \omega^j), \qquad\qquad d[\Omega] \equiv 0\ (\mathrm{mod}\ \omega^j).$$

Some of the authors that follow distinguish *stationary* fields from extremal fields by the requirement that when s satisfies these congruences in general it is stationary, and becomes extremal only when s is also integrable – i.e., $s = j^1x$ for some submanifold embedding $x: \mathcal{O} \to M$.

The Lepage congruences then help us clarify the distinction between the geodesic fields that were defined by De Donder-Weyl and the ones that were defined by Carathéodory. It basically comes down to the particular choice of representative for the Lepage congruence class of Ω. For a De Donder-Weyl field, the choice takes the Poincaré-Cartan form:

$$\Omega_0 = \mathcal{L}(j^1x)\mathcal{V} + (-1)^? \, \Pi_i^a \, \#\partial_a \wedge \omega^i .$$

For a Carathéodory field, one chooses the unique simple (i.e., decomposable) r-form in the congruence, which takes the form:

$$\Omega^* = \frac{1}{\mathcal{L}^{r-1}} \prod_{a=1}^{r} (\mathcal{L}dt^a + \Pi_i^a \omega^i) .$$

However, Hölder shows that the two choices are related by a contact transformation.

e. Caratheodory complete figure. If one has a Lagrangian field of contact elements $z: \mathcal{O}\times M \to J^1(\mathcal{O}; M)$ then its values in $J^1(\mathcal{O}; M)$ are contact elements, and thus define linear subspaces in the tangent spaces of $T(\mathcal{O}\times M)$. For the rest of this section, we assume that z is regular, in the sense that all of its values $z_a^i(t,x)$ are matrices of rank r. They then locally define r-frames in $T(\mathcal{O}\times M)$ by way of the r local vector fields:

$$\mathbf{z}_a(t, x) = z_a^i(t,x)\,\partial_i\,, \quad a = 1, \dots, r.$$

Since they are presumed to be linearly independent, they span an r-plane $[\mathbf{z}_a](t, x)$ in $T(\mathcal{O})$ for each (t, x). This r-plane can also be associated with the r-vector field:

$$\mathbf{Z} = \mathbf{z}_1 \wedge \dots \mathbf{z}_r\,.$$

The rank-r sub-bundle $E(\mathcal{O}\times M)$ of $T(\mathcal{O}\times M)$ that is defined by all of these r-planes then constitutes a differential system on $\mathcal{O}\times M$. By Frobenius, the necessary and sufficient condition for it to be completely integrable is that $[\mathbf{z}_a, \mathbf{z}_b]$ be a linear combination of the $\mathbf{z}$'s again. Since:

$$[\mathbf{z}_a, \mathbf{z}_b] = \left(z_a^i \frac{\partial z_b^j}{\partial x^i} - z_b^i \frac{\partial z_a^j}{\partial x^i} \right) \frac{\partial}{\partial x^j}\,,$$

the issue comes down to whether one can solve the defining equation for the $\mathbf{z}$'s in terms of the ∂_i's. However, there are r $\mathbf{z}$'s and m ∂_i's, so as long as $r < m$ the system is over-determined if one desires to solve for the ∂_i 's. Hence, a given set of $\mathbf{z}$'s will not always be associated be associated with some set of ∂_i's, but only the ones that lie in a certain r-dimensional linear subspace of $\mathbb{R}^m$. If it so happens that $r = m$ then the condition of possibility for the solution is the invertibility of the matrix z_a^i , but this also follows from the demand that z be regular.

Note that in the case of $r = 1$, the differential system is always integrable into a foliation of curves. When $z(t, x)$ is an extremal field of contact elements, it is a geodesic congruence. (Recall that any submanifold that is embedded in an extremal field will be extremal.)

More generally, in the event that the differential system $E(\mathcal{O}\times M)$ is completely integrable the manifold $\mathcal{O}\times M$ will be foliated by r-dimensional leaves, and if the field of contact elements z is extremal then the leaves will be extremal submanifolds of the action functional that one is dealing with.

When z also happens to be a geodesic field, so $z = dS^a \wedge \#\partial_a$, the r functions S^a, $a = 1, \dots, r$ local foliate $\mathcal{O}\times M$ with m-dimensional leaves by the level hypersurfaces of the map $\mathcal{O}\times M \to \mathbb{R}^r$, $(t, x) \mapsto S^a(t, x)$. Hence, the differential system $S(\mathcal{O}\times M)$ on $\mathcal{O}\times M$ that is defined by the tangent planes that are annihilated by all of the 1-forms dS^a is trivially completely integrable.

Since the dS^a's are assumed to be linearly independent they define a non-zero r-form:

$$\Sigma = dS^1 \wedge \ldots \wedge dS^r.$$

The r-planes of $E(\mathcal{O}\times M)$ and the m-planes of $S(\mathcal{O}\times M)$ are transversal iff $\Sigma(\mathbf{Z})$ is non-zero at every point of $\mathcal{O}\times M$. This implies the non-vanishing of the determinant of the matrix:

$$[S\cdot z]^a_b = dS^a(\mathbf{z}_b) = \frac{\partial S^a}{\partial x^i} z^i_b .$$

Such a pair of complementary foliations on $\mathcal{O}\times M$ – when it exists – or the pair of fields $\{S^a, z\}$ is said to be the *Caratheodory complete figure* for the geodesic field in question, and one sees in what follows that he introduced the concept in the last section of his article [**5**]. We illustrate the sort of situation in Fig. 5:

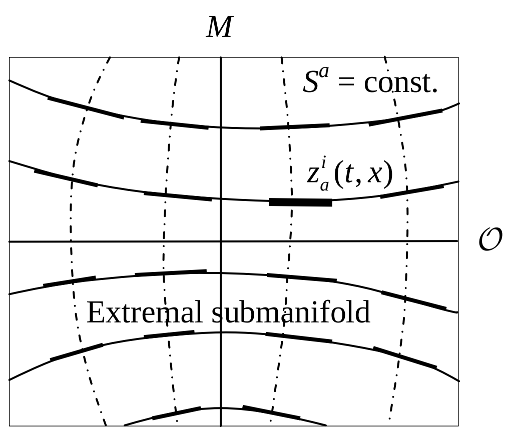

Figure 5. The Caratheodory complete figure for a geodesic field.

For $r = 1$, the Caratheodory complete figure is defined by a geodesic congruence and the transversal hypersurfaces are defined by the eikonal $S(t, x)$, to use the terminology of geometrical optics.

5. Sufficient conditions for a strong or weak local minimum. Now, let us return to the basic problem that was posed in the introductory remarks.

First, we should clarify the precise usage of the terms "strong" and "weak." Basically, they relate to two possible topologies for the set $C^1(\mathcal{O}, M)$ of all C^1 maps from $\mathcal{O}$ to M. We shall only sketch the essential elements, so for more rigor, one can consult Hirsch [**40**].

In the case of a strong local minimum, the topology is the C^0 – or *compact-open*-topology on $C^1(\mathcal{O}, M)$. If x: $\mathcal{O} \to M$ is a C^1 map then a neighborhood of x in the compact-open topology is defined by the set $N_x(K, V)$ of all C^1 maps y: $\mathcal{O} \to M$ that map a compact subset $K \subset \mathcal{O}$ into an open subset $V \subset M$ when $x(K) \subset V$; thus, the behavior of y outside of K is irrelevant. We can represent this situation schematically as in Fig. 6:

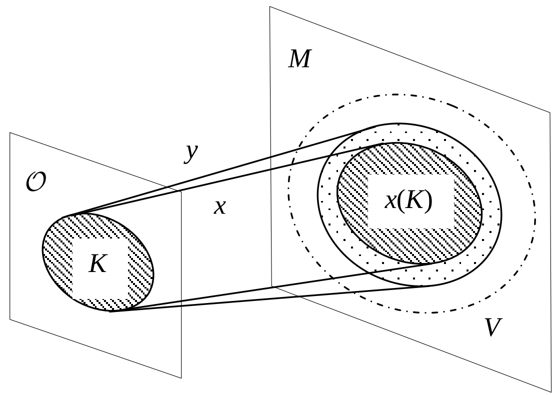

Figure 6. A neighborhood in the compact-open topology on $C^1(\mathcal{O}, M)$.

These neighborhoods then define a sub-basis for the open subsets of the C^0 topology. That is, any open subset is a union of some family of finite intersections of these neighborhoods.

The C^1 topology on $C^1(\mathcal{O}, M)$, which is a special case of the more general C^k topology introduced by Whitney, is finer that the C^0 topology, in the sense of having "more" open subsets; i.e., some of the open subsets of the C^1 topology are not open in the C^0 topology. This is because one further restricts the C^1 functions in the neighborhood above by the requirement that if $\varepsilon > 0$ and one replaces x and y by their local representatives x^i, y^i: $\mathbb{R}^r \to \mathbb{R}^m$ when one chooses coordinate charts (U, t^a) and (V, ξ^i), such that $K \subset U$ then in addition to the condition above, the two maps must also satisfy the constraint that:

$$\left\| \frac{\partial y^i}{\partial t^a}(t) - \frac{\partial x^i}{\partial t^a}(t) \right\| < \varepsilon$$

for every $t \in K$; one denotes the set of all such y by $N_x(K, V; \varepsilon)$.

These neighborhoods then constitute a basis for the open subsets of the C^1 topology. That is, any open subset of that topology can be expressed as a union of some family of these neighborhoods.

With these definitions, we clarify that an extremal submanifold x: $\mathcal{O} \to M$ is a *strong local minimum* for the action functional $S[x]$ iff:

$$S[x] \leq \mathrm{S}[x']$$

for all C^1 maps x': O $\to M$ that lie in a C^0 neighborhood of x and a *weak local minimum* iff this inequality is true for all maps in a C^1 neighborhood of x.

a. Hilbert independent integral. When a submanifold x is embedded in a geodesic field $x_a^i(t, x)$ the Lagrangian density must take the form:

$$\mathcal{L}(t^a, x^i, x_a^i(t, x(t))) = \frac{dS^a}{dt^a} = \frac{\partial S^a}{\partial t^a} + x_a^i \frac{\partial S^a}{\partial x^i} = \frac{\partial S^a}{\partial t^a} + x_a^i \frac{\partial \mathcal{L}}{\partial x_a^i}.$$

Hence, since the resulting action functional:

$$S[x] = \int_x dS = \int_{\partial x} S$$

is independent of the submanifold x – apart from its boundary – one defines the more general integral functional:

$$I[x] = \int_x dS(j^1 x) = \int_x \left(\frac{\partial S^a}{\partial t^a} + x_a^i \frac{\partial \mathcal{L}}{\partial x_a^i} \right) \mathcal{V} = \int_{\partial x} S(j^1 x)$$

the *Hilbert independent integral* that is defined by $\mathcal{L}$ and $x_a^i(t, x)$.

Since this functional is affected by only the boundary values of the r–1-form S, it will take the same value for any submanifold $\bar{x}$ that it does for the extremal submanifold x, and therefore all other submanifolds that have the same boundary as x will be extremals, as well. Hence, as an action functional it is quite ambiguous, but it can be added or subtracted from the action functional $S[x]$ without affecting the outcome.

One also notes that one can give it the form:

$$I[x] = \int_x \left(\frac{\partial S^a}{\partial t^a} + \Pi_i^a x_a^i \right) \mathcal{V}.$$

Hence, there is nothing to stop one from replacing the Lagrangian density $\mathcal{L}$ with:

$$\mathcal{L}^* = \mathcal{L} - \frac{\partial S^a}{\partial t^a} - \Pi_i^a x_a^i,$$

as this will change the value of the resulting action functional without changing the resulting extremals. In particular, one sees that:

$$S^*[x] = 0$$

when x is extremal.

b. Weierstrass excess function. When one has a geodesic field $x_a^i(t, x)$ for an action functional, one can use it to express the difference between its value $S[x]$ on an extremal submanifold x and its value $S[\bar{x}]$ another submanifold $\bar{x}$ in terms of the integral of a function $\mathcal{E}$: $J^1(\mathcal{O}; M) \to \mathbb{R}$ that vanishes for the extremal. Hence, $\mathcal{E}$ also defines a functional on submanifolds, just as action does, and it will have property that for the arbitrary submanifold $\bar{x}$ and an extremal submanifold x that have the same boundary, one will have:

$$S[\bar{x}] - S[x] = \int_{\bar{x}} \mathcal{E}(j^1\bar{x})\mathcal{V} .$$

The key to its construction is to see that although the two action functionals on left-hand side are defined over two different submanifolds, namely:

$$S[\bar{x}] = \int_{\bar{x}} \mathcal{L}(j^1\bar{x})\mathcal{V} , \qquad S[x] = \int_x \mathcal{L}(j^1x)\mathcal{V} ,$$

the *Weierstrass excess functional* is defined only on the submanifold $\bar{x}$.

Now, if x is embedded in a geodesic field $z(t, x)$ then $j^1x = z \cdot x$ and $S[x]$ will have the form:

$$S[x] = I[x] = \int_x dS(j^1x) .$$

Since the value of the functional $I[x]$ depends only upon the behavior of x on the boundary and $\bar{x}$ shares that boundary, one can infer that:

$$I[x] = I[\bar{x}] .$$

One then replaces $S[x]$ with:

$$I[\bar{x}] = \int_{\bar{x}} dS(j^1\bar{x}) ,$$

One further replaces $\mathcal{L}(p)$, where $p \in J^1(\mathcal{O}; M)$, with $\mathcal{L}^*$, as in the previous subsection, as we have a right to do, since the supplementary term will not affect the extremals, and this makes:

$$S[\bar{x}] - S[x] = \int_{\bar{x}} [\mathcal{L}(j^1\bar{x}) - \mathcal{L}(z(t, \bar{x}(t)) - \Pi_i^a(z(t, \bar{x}(t))(\bar{x}_{,a}^i(t) - x_a^i(t, \bar{x}(t))]\mathcal{V} ,$$

and we can then define the *Weierstrass excess function* to be:

$$\mathcal{E}(j^1\bar{x}) = \mathcal{L}(j^1\bar{x}) - \mathcal{L}(z(t, \bar{x}(t)) - \Pi_i^a(z(t, \bar{x}(t))(\bar{x}_{,a}^i(t) - x_a^i(t, \bar{x}(t)) .$$

The more traditional notation for this function is $\mathcal{E}(t^a,\ x^a, x_a^i, \bar{x}_a^i)$, although this somewhat obscures the fact that once the geodesic field x_a^i has been chosen, it becomes a function on $J^1(\mathcal{O};\ M)$ like any Lagrangian, as well as obscuring the functional dependencies of the last three sets of coordinates.

Clearly, $\mathcal{E}(j^1x) = 0$ when x is an extremal, but the primary reason that Weierstrass had for defining this function was to establish that a sufficient condition for that extremal x to be a strong local minimum for $S[x]$ was that one have:

$$\mathcal{E}(j^1\bar{x}) \geq 0$$

for all other submanifolds $\bar{x}$ that share a common boundary with x and are contained in a C^0 neighborhood of it.

We illustrate the situation that we have been discussing schematically in Fig. 7.

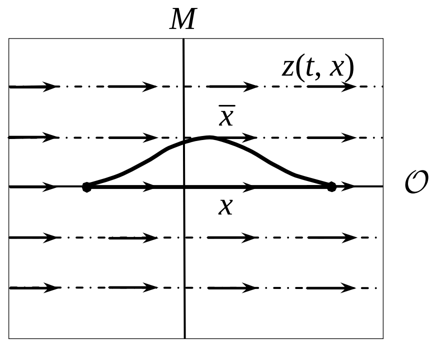

Figure 7. Constructions used in the Weierstrass excess function.

It is essential to understand that the existence of a geodesic field for the Lagrangian of the action function $S[x]$ is a necessary part of defining the excess function. Hence, it will be inapplicable to the consideration of more general Lagrangians that do not admit geodesic fields.

c. Legendre-Hadamard condition. A somewhat simpler sufficient condition for an extremal to be a local minimum was defined by Legendre in the case of extremal curves and discussed by Hadamard for higher-dimensional submanifolds.

It can be obtained by expanding the excess function in a Taylor series about each point of the extremal curve as a function of $\bar{x}_a^i - x_a^i$:

$$\mathcal{E}(j^1\bar{x}) = \mathcal{E}(j^1x) + \left.\frac{\partial \mathcal{E}}{\partial x_a^i}\right|_{j^1x} (\bar{x}_a^i - x_a^i) + \frac{1}{2}\left.\frac{\partial^2 \mathcal{E}}{\partial x_a^i \partial x_b^j}\right|_{j^1x} (\bar{x}_a^i - x_a^i)(\bar{x}_b^j - x_b^j) + \ldots$$

which then becomes:

$$\mathcal{E}(j^1\bar{x}) \approx \frac{1}{2}\left.\frac{\partial^2 \mathcal{E}}{\partial x_a^i \partial x_b^j}\right|_{j^1x} (\bar{x}_a^i - x_a^i)(\bar{x}_b^j - x_b^j),$$

up to higher-order terms in $\bar{x}_a^i - x_a^i$, since the first two terms vanish when evaluated on an extremal.

The *Legendre-Hadamard* sufficient condition that an extremal x be a weak local minimum is then the demand that the matrix:

$$\left.\frac{\partial^2 \mathcal{E}}{\partial x_a^i \partial x_b^j}\right|_{j^1x}$$

must be positive-definite at each point along x.

Appendix A.

Differentiable singular cubic chains [1].

Something that becomes gradually apparent in the work of Dedecker on the calculus of variations [**21**, **22**] is the fact that at one level of consideration it can be regarded as the study of differentiable homotopies. If one desires to pursue this algebraic topological aspect of variational problems then it becomes rapidly useful to make the objects that are being varied have an algebraic-topological character, as well.

In particular, instead of varying compact submanifolds $x: \mathcal{O} \to M$ in a differential manifold M, one can use more specific building blocks that lead one into singular homology and cohomology, and when one applies de Rham's theorem, to de Rham cohomology. Instead of a compact subset $\mathcal{O} \subset \mathbb{R}^r$, one first considers the standard r-cube $I^r = [0, 1]\times \ldots \times[0, 1]$. A *differentiable singular r-cube in M* is then a differentiable map $\sigma_r: I^r \to M$. In order to define differentiability when one is dealing with a piecewise linear manifolds, such as I^r, one must surround it with an open neighborhood, define a differentiable extension of x, and then restrict the extension to I^r; since differentiation is a local process, the choice of extension is irrelevant.

The reason that one refers to such an r-cube in M as singular is because unless one restricts the submanifold map to be an immersion or embedding the dimension of the image does not have to be r, and might very well be zero, as in the case of a constant map. If all one desires to examine is singular homology then this is no loss of generality since the effect of the degenerate r-cubes eventually disappears when one passes to homology, but if one also intends to consider geometrical issues then one usually must specify some regularity condition on the map. One also notes that the restriction from continuous maps of I^r into M to differentiable ones is not significant since every homotopy class $[\sigma_r]$ of continuous maps contains a differentiable element, which one proves by a smoothing construction.

The *boundary* $\partial \mathrm{I}^r$ of I^r is defined to be the sum of its "1-faces" minus the sum of its "0-faces," where the r 0-faces $I_a^r(0)$ of I^r take the form of all points of the form $I_1^r(0) = (0, t^2, \ldots, t^r), \ldots, I_a^r(0) = (t^1, \ldots, t^a, \ldots, t^r), \ldots, I_r^r(0) = (t^1, \ldots, t^{r-1}, 0)$, and the 1-faces $I_a^r(1)$, $a = 1, \ldots, r$ take the same form with the 0 replaced with a 1. Hence:

$$\partial \mathrm{I}^r = \sum_{a=1}^{r} [I_a^r(1) - I_a^r(0)] .$$

Since the r-faces of I^r are r–1-cubes in their own right, the process of taking the boundary can be applied to them as well. However, due to the sign alternation in the definition of the boundary one always has that:

$$\partial^2 = 0,$$

[1] For some standard mathematical references on singular homology and singular cohomology, one can confer Greenberg [**41**], Rotman [**42**], or Vick [**43**].

as one easily verifies in the case of a square.

A *differentiable singular cubic k-chain* c_k in M is composed of a "formal sum" of a finite number of k-cubes σ_i, i = 1, …, N in M whose coefficients come from some chosen ring R of coefficients:

$$c_k = \sum_{i=1}^{N} a_i \sigma_i \,.$$

Although it is possible to make the definition of a finite formal sum with coefficients in R more rigorous (one considers the "free R-module generated by the set of all k-cubes"), for the sake of computation it is entirely sufficient to deal with the formal sums naively and simply apply rules of computation to them.

A useful aspect of the use of cubic k-chains is the fact that a differentiable homotopy of a cubic k-chain becomes a cubic k+1-chain. Hence, if one regards finite variations of k-chains as differentiable homotopies then this also makes finite variations take the form of differentiable k+1-chains.

One extends the boundary operator from k-cubes to k-chains by linearity:

$$\partial c_k = \sum_{i=1}^{N} a_i \partial\sigma_i = \sum_{i=1}^{N} a_i \sigma_i(1) - \sum_{i=1}^{N} a_i \sigma_i(0)\,.$$

Something that is not entirely obvious at this point is that the *actual* definition in practice of this boundary operator for a given M does not follow automatically from the nature of the cubes in M, but must be introduced essentially "by hand" in order to account for the topology of M. Otherwise, the free R-module we have defined involves only the cardinality of M as a set, and completely ignores the details of its topology.

In order to pass to homology, one starts with the aforementioned free R-module $C_k(M; R)$ of finite formal sums of k-cubes with coefficients in R and then regards the boundary operator as a linear map ∂: $C_k(M; R) \to C_{k-1}(M; R)$ for each k from 0 to r. The image $B_{k-1}(M; R)$ of the boundary map is a submodule of $C_{k-1}(M; R)$ that one calls the module of *k*–1-*boundaries* in M, while its kernel $Z_k(M; R)$ is a submodule of $C_k(M; R)$ that one calls the module of *k-cycles* in M. The quotient module $H_k(M; R) = Z_k(M; R) / B_k(M; R)$ of all translates of $B_k(M; R)$ in $Z_k(M; R)$ is called the *(differentiable singular) homology module in dimension k.* Roughly speaking, its generators represent "k-dimensional holes" in M.

One can also think of elements in $H_k(M; R)$ as equivalence classes of k-cycles under the equivalence relation of homology. That is, two k-chains c_k and c'_k are *homologous* iff their difference is a boundary:

$$c'_k - c_k = \partial c_{k+1} \text{ (for some } c_{k+1} \in B_{k+1}(M; R)).$$

It is important to note that $H_k(M; R)$ will vanish for every $k > r$.

The dual notion to a differentiable singular cubic k-chain is that of a *differentiable singular cubic k-cochain*, which is simply a linear functional c^k: $C_k(M; R) \to R$, which makes it a finite formal sum of linear functionals on k-cubes with values in R. One denotes the corresponding free R-module of all k-cochains by $C^k(M; R)$.

Since k-cochains can be applied to k-chains, there is a natural bilinear pairing $C^k(M; R)\times C_k(M; R) \to R$, $(c^k, c_k) \mapsto \langle c^k, c_k\rangle$, where:

$$\langle c^k, c_k \rangle = c^k(c_k).$$

One then defines the *coboundary operator* $\delta: C^k(M; R) \to C^{k+1}(M; R)$ to be the adjoint to ∂ under this pairing:

$$\langle \delta c^k, c_{k+1}\rangle = \langle c^k, \partial c_{k+1}\rangle.$$

As we shall see shortly, this is really just an abstraction of Stokes's theorem for differential forms.

One defines analogous R-modules $Z^k(M; R)$, $B^k(M; R)$, and $H^k(M; R)$ that one calls the modules of *k-cocycles*, *k-coboundaries*, and the *cohomology module in dimension k*, respectively. The elements of the latter module are also called *k-cohomology classes* and two k-cochains are *cohomologous* iff their difference is a coboundary.

Although $C^k(M; R)$ is defined to be the dual R-module $\mathrm{Hom}(C_k(M; R); R)$, the same does not have to be true of $H^k(M; R)$; it can also include "torsion" factors; i.e., cyclic R-modules. However, if R is a principle ideal domain – such as, for instance, a field – there are no torsion factors, so, in particular $H^k(M; \mathbb{R}) = \mathrm{Hom}(H_k(M; \mathbb{R}); \mathbb{R})$.

In order to go from differentiable singular cubic cohomology to de Rham cohomology, one starts with the fact that any k-form α on M defines a linear functional on differentiable singular k-chains with values in $\mathbb{R}$ by integration:

$$\alpha[c_k] = \int_{c_k} \alpha .$$

Hence, one can regard α as a representative of a differentiable singular cubic k-cochain.

In fact, from Stokes's theorem, if α is a k–1-form then one has:

$$d\alpha[c_k] = \alpha[\partial c_k].$$

Hence, this, and the facts that d is a linear operator of degree +1 and $d^2 = 0$, one can treat k-forms as k-cochains in a different sort of cohomology that one calls de Rham cohomology after its inventor Georges de Rham [**44**]. The coboundary operator is d, which makes the k-cocycles take the form of closed k-forms, the k-coboundaries are exact k-forms, and the de Rham cohomology vector spaces are equivalence classes of closed forms that differ by exact forms. We say "vector spaces" in this case because the coefficient ring is the field $\mathbb{R}$, which makes the $\mathbb{R}$-modules into $\mathbb{R}$-vector spaces; hence, a set of generators for the de Rham cohomology vector space $H^k_{dR}(M)$ in dimension k is simply a basis for it as a vector space.

The theorem that de Rham had to prove was that the cohomology defined by the exterior differential forms on M was isomorphic to the singular cohomology $H^k(M; \mathbb{R})$

with values in $\mathbb{R}$. Hence, one sees that de Rham cohomology [1] represents something of an approximation to the topology of M, since manifolds that differ by torsion factors – e.g., an n-sphere and an n-dimensional projective space, will appear indistinguishable in the eyes of de Rham cohomology.

It is, however, a useful and powerful approximation, nonetheless.

The application of the foregoing discussion to the calculus of variations now becomes immediate when one restricts the objects being varied from submanifolds to chains, since we are defining the action functional to be essentially a differentiable singular r-cochain on M that is represented by the r-form $\mathcal{L}\mathcal{V}$ on M. Since the first variation functional takes the form of its coboundary, and does not generally vanish, except for extremal chains, one sees that in general the action functional is neither a cocycle nor a coboundary. However, the essence of Hamilton-Jacobi theory is that the functional becomes a cocycle for extremal chains and possibly a coboundary, as well.

Actually, the restriction from compact k-dimensional submanifolds with boundaries in M to differentiable singular cubic k-chains in M is no loss of generality, at least in the eyes of homotopy and homology, as Munkres [**47**] proved, in effect, that every compact submanifold is homotopically equivalent to such a chain.

[1] In addition to the reference by de Rham, one might also compare Bott and Tu [**45**], as well as Warner [**46**].

Appendix B

Characteristics of first-order partial differential equations.

Since the construction of a geodesic field often comes down to an initial-value problem for a first-order partial differential equation of Hamilton-Jacobi type, which can then be solved by the method of characteristics, it is worthwhile to point out that all of that is entirely natural within the context of jet manifolds and contact geometry. We then briefly summarize the essential points that relate to the present class of problems.

a. Differential equations and jet manifolds. When one defines a system of N first-order partial differential equations in the classical form:

$$F^{\nu}\left(t^a, x^i(t), \frac{\partial x^i}{\partial t^a}(t)\right) = 0, \qquad \nu = 1, \ldots, N$$

it becomes clear, from the foregoing discussions, that one can also put this system into the form of two systems:

$$F^{\nu}\left(t^a, x^i(t), x^i_a(t)\right) = 0, \qquad \nu = 1, \ldots, N,$$

$$x^i_a(t) = \frac{\partial x^i}{\partial t^a}(t), \qquad a = 1, \ldots, r, \quad i = 1, \ldots, m.$$

We now see that what the first one defines is a submanifold (or at least an algebraic subset) of $J^1(\mathcal{O}; M)$ by way of the zero-locus of a function F: $J^1(\mathcal{O}; M) \to \mathbb{R}^N$, while the second set is the integrability condition for a section of $J^1(\mathcal{O}; M) \to \mathcal{O}$. Thus, one could also express the system of equations as:

$$F^{\nu}(j^1 x) = 0.$$

Matters are simplest when one only has one function F to contend with. For instance, the Hamilton-Jacobi equation, in its homogeneous form, can be expressed in the form:

$$H\left(x^i, S(x), p_i(x)\right) = 0, \qquad p_i = \frac{\partial S}{\partial x^i},$$

as long as one includes the constraint that H is actually independent of S. Thus, if H is a real-valued function on $J^1(M; \mathbb{R})$ the Hamilton-Jacobi equation is defined by integrable sections s: $M \to J^1(M; \mathbb{R})$ that take their values in the zero locus of H.

b. Method of characteristics. As long as one is considering only one first-order partial differential equation whose solutions will take the form of differentiable functions

f on a manifold M, solving an initial-value problem (i.e., a *Cauchy* problem) can be reduced to a corresponding initial-value problem for a system of first-order ordinary differential equations that one calls the *characteristic equations* defined by the original PDE. Indeed, this construction is perfectly natural in the language of jet manifolds, so we now describe it that way.

One must be aware that since any partial differential equation of order higher than one can be converted into an equivalent system of first-order partial differential equations, the method of characteristics is no longer applicable to systems of more than one first-order PDE. In particular, the linear wave equation can be converted into a pair of first-order PDE's, but they cannot be solved directly using characteristics, only indirectly by introducing the geometrical optics approximation.

Say a differentiable function ϕ on a manifold M is a solution to a Cauchy problem for a first-order PDE:

$$F(x^\mu, \phi, p_\mu) = 0, \qquad p_\mu = \frac{\partial \phi}{\partial x^\mu}, \qquad \phi(0, x^i) = \phi_0,$$

in which $\mu = 0, 1, \ldots, m$, $i = 1, \ldots, m$.

The essence of the method of characteristics is to first use the contact geometry of $J^1(M; \mathbb{R})$ and the function F to define a global vector field on $J^1(M; \mathbb{R})$, which then allows one to treat each point of that manifold as potentially the initial point of a trajectory for that vector field.

The function F defines a 1-form by its differential:

$$dF = \frac{\partial F}{\partial x^\mu} dx^\mu + \frac{\partial F}{\partial \phi} d\phi + \frac{\partial F}{\partial p_\mu} dp_\mu \ .$$

The question is now how one might convert this covector field into a vector field of the form:

$$X_F = X^\mu \frac{\partial}{\partial x^\mu} + X^\phi \frac{\partial}{\partial \phi} + X_\mu \frac{\partial}{\partial p_\mu}.$$

Here is where we use the integrability assumption about p_μ . It amounts to the statements that at each point of $J^1(M; \mathbb{R})$, the tangent vector X_F is incident on the hyperplane defined by $\omega = 0$ and:

$$i_{X_F} \Theta = dF$$

in which the canonical 1-form ω and 2-form Θ on $J^1(M; \mathbb{R})$ are:

$$\omega = d\phi - p_\mu \, dx^\mu , \qquad \Theta = d\omega = dx^\mu \wedge dp_\mu ,$$

respectively.

From the first requirement on X_F, we find:

$$d\phi = p_\mu \, dx^\mu, \qquad X^\phi = p_\mu \; X^\mu.$$

From the second, when combined with first equation above, we find:

$$- X_\mu \, dx^\mu + X^\mu \, dp_\mu = (F_\mu + p_\mu \, F_\phi) \, dx^\mu + F^\mu \, dp_\mu \, .$$

This gives the following set of equations for the components of X_F :

$$X^\mu = F^\mu, \qquad X^\phi = p_\mu \; X^\mu, \qquad X_\mu = - (F_\mu + p_\mu \, F_\phi).$$

The vector field X_F, thus defined, represents the *characteristic vector field* defined by the first-order partial differential equation in question. Since any vector field on a manifold locally defines a system of first-order ordinary differential equations by assuming that it always gives the velocity vector field of a solution trajectory, we obtain the characteristic equations that are associated with the PDE:

$$\frac{dx^\mu}{dt} = \frac{\partial F}{\partial p_\mu}, \qquad \frac{d\phi}{dt} = p_\mu \, \frac{\partial F}{\partial p_\mu}, \qquad \frac{dp_\mu}{dt} = -\left(\frac{\partial F}{\partial x^\mu} + p_\mu \, \frac{\partial F}{\partial \phi} \right).$$

References

1. G. A. Bliss, *Lecture on the Calculus of Variations*, University of Chicago Press, Chicago, 1946.
2. I. M. Gelfand and S. V. Fomin, *Calculus of Variations*, Prentice-Hall, N. J., 1963.
3. A. Clebsch, "Über die zweite Variation vielfacher Integral," J. de Crelle, t. **56** (1859), 122-149.
4. J. Hadamard, "Sur une question de calcul des variations," Bull. Soc. Math. France **30** (1902), 253-256.
   "Sur quelques questions de calcul des variations," Bull. Soc. Math. de France **33** (1905), 73-80.
5. D. Hilbert, "Zur Variationsrechnung," Math. Ann. **69** (1906), 351-370; Gesammelte Abhandlungen, v. III, pp. …
6. C. Caratheodory, "Über die Variationsrechnung bei mehrfachen Integralen," Acta Szeged **4** (1929), 193-216; Gesammelte Mathematische Schriften, v. I, pp. 401-426.
7. C. Caratheodory, *Calculus of Variations and Partial Differential Equations of the First Order*, Chelsea, NY, 1982.
8. Rund, H., *The Hamilton-Jacobi Method in Mathematical Physics*, Krieger, NY, 1973.
9. H. Weyl, "Observations on Hilbert's independence Theorem and Born's quantization of field equations," Phys. Rev. **46** (1934), 505-508.
10. H. Weyl, "Geodesic Fields in the Calculus of Variations for Multiple Integrals," Ann. Math., **36** (1935), pp. 607-629.
11. Cartan, E., *Leçons sur les invariants intégraux*, Hermann, Paris, 1921.
12. E. Goursat, *Leçons sur les problème de Pfaff*, Hermann, Paris, 1922.
13. T. de Donder, *Théorie des invariants integraux*, Paris, 1927.
14. T. de Donder, *Théorie invariantive du calcul de variations*, 2nd ed., Gauthier-Villars, Paris, 1935.
15. H. Boerner, "Über die Extremalen und geodätischen Felder in der Variationsrechnung der mehrfachen Integrale," Math. Ann. **112** (1936), pp. 187-220.
    "Variationsrechnung aus dem Stokeschen Satz," Math. Zeit. **46** (1940), pp. 709-719.
    "Über die Legendrische Bedingung und die Feldtheorie in der Variationsrechnung der Mehrfachen Integrale," Math. Zeit **46** (1940), pp. 720-742.
16. J. Gehéniau, "Généralisation de la formule d'excès de Weierstrass déduite du théorème d'indépendence d'Hilbert-de Donder," Bull. Acad. Roy. Belg. **21** (1935), pp. 385 and 504.
17. T. Lepage., "Sur les champs géodésiques du calcul des variations," Bull. Acad. R. Belg., **22** (1936), pp. 716-729.
    T. Lepage., "Sur les champs géodésiques des intégrales multiples," Bull. Acad. R. Belg., **22** (1936), pp. 27-46.
    T. Lepage., "Champs stationnaires, champs géodésiques, et formes intégrables," Bull. Acad. R. Belg., **22** (1936), pp. 73-94.
18. R. Debever, "Les champs de Mayer dans le calcul des variations des integrals multiples," Bull. Roy. Acad. Belg. **23** (1937), pp. 809-815.
19. E. Hölder, Jber. d. Deutsche Math.-Verein **49** (1939), 162-178.
20. P. V. Paquet, "Les formes differentielles Ω_n dans le Calcul de Variations," Bull. Acad. Roy. Belg. **27** (1941), pp. 65-84.

21. L. van Hove, "Sur les champs de De Donder-Weyl et leur construction par la méthode des caractéristiques," Bull. Roy. Acad. Belg. **31** (1945), pp. 278-285.
"Sur les champs de Caratheodory et leur construction par la méthode des caractéristiques," ibid., pp. 625-638.
22. P. Dedecker, "Calcul de variations, différentielles, et champs géodésiques," *Géométrie différentielles,* Colloq. Int. de la Recherche Sci., Strasbourg (1953), pp. 17-34.
23. P. Dedecker, *Calcul des variations et topologie algébrique,* Mémoires de la Société Royale des Sciences de Liège, tome XIX, fasc. 1, 1957.
24. Griffiths, P., *Exterior Differential Systems and the Calculus of Variations*, Birkhäuser, Boston, 1983.
25. R. Bryant, P. Griffiths, and D. Grossman, *Exterior Differential Systems and Euler-Lagrange Partial Differential Equations,* Univ. of Chicago Press, Chicago, 2003.
26. C. Ehresmann, "Structures locales et structures infinitésimales," Comptes Rendus, t. **234** (1952), 487-489.
27. W. R. Hamilton, "Third Supplement to an Essay on the Theory of Systems of Rays," Trans. Roy. Irish Acad. **17** (1837), pp. 1-144.
28. S. Lie and G. Scheffer, *Berührungstransformationen*; reprinted by A. M. S. Chelsea Publishing, Providence, R. I., 2005.
29. E. Vessiot, "Sur la propagation par ondes sur le problème de Mayer," J. de Math. (6) **9** (1913), 39-76.
30. H. Goldschmidt and S. Sternberg, "Hamilton-Cartan Formalism in the Calculus of Variations," Ann. l'Inst. Fourier (Grenoble) 23 (1973), 203-267.
31. D. Saunders, *Geometry of Jet Bundles*, Cambridge University Press, Cambridge, 1989.
32. M. Morse, *The Calculus of Variations in the Large*, Amer. Math. Soc. Soc., Providence, R. I. 1934.
33. J. Milnor, *Morse Theory,* Princeton University Press, Princeton, 1963.
34. L. A. Ljusternik, *The Topology of Function Spaces and the Calculus of Variations in the Large,* Translations of the Amer. Math. Soc., v. **16**, Providence, R. I., 1966; translated from the Russian edition that was printed in Moscow in 1947.
35. C. B. Morrey, *Multiple Integrals in the Calculus of Variations,* Springer, Berlin, 1966.
36. D. H. Delphenich, "Integrability and the variational formulation of non-conservative mechanical systems," Ann. Phys. (Berlin) **18** (2009), 45-56.
37. D. H. Delphenich, "On the variational formulation of systems with non-holonomic constraints," Ann. Phys. (Berlin) **18** (2009), 649-670.
38. H. Rund, *The Differential Geometry of Finsler Spaces* (Springer, Berlin, 1959).
39. D. Bao, S.-S. Chern, and Z. Shen, *An Introduction to Riemann-Finsler Geometry* Springer, New York, 2000.
40. M. W. Hirsch, *Differential Topology*, Springer, Berlin, 1976.
41. M. Greenberg, *Lectures on Algebraic Topology*, Benjamin-Cummings, Boston, 1967.
42. J. Rotman, *An Introduction to Algebraic Topology,* Springer Verlag, Berlin, 1988.
43. J. Vick, *Homology Theory,* Academic Press, New York, 1973.
44. G. de Rham, *Differentiable Manifolds,* Springer, Berlin, 1984.
45. R. Bott and L. Tu, *Differential Forms in Algebraic Topology*, Springer, Berlin, 1982.

46. Warner, F., *Differentiable Manifolds and Lie Groups,* Scott Foresman, Glenview, IL, 1971.
47. J. Munkres, *Elementary Differential Topology,* Princeton University Press, NJ, 1966.

On the second variation of multiple integrals

(By *A. Clebsch*)

As is well-known, the examination of the criteria for the maximum and minimum of a simple integral has led to the consideration of the second variation, and this has given rise to the discovery of remarkable properties. It shows that one may, by partial integration, reduce it to a simpler form. This integration might not yield this result unless one had performed the integration of certain differential equations whose complicated character has long discouraged geometers from seeking this integration.

Jacobi, in the case of simple integrals of *one* dependent variable, discovered a connection between these equations and the ones that lead to the vanishing of the first variation, and insofar as he arrived at the integration of those transformation equations with great ease in this case, he arrived at an entirely new viewpoint of great significance by these investigations.

In an earlier paper that is included in the 55$^{\text{th}}$ volume, pp. 254, of this journal, I have proved that the *Jacobi* principles admit an application to all problems of the calculus of variations that depend on only simple integrals, and it is by means of the examination of the second variation that one is led back to the examination of values that can be assumed by a homogeneous function of second order, between whose arguments certain linear condition equations exist.

This advance, which is essentially necessitated by the method set down in the aforementioned paper, likewise leads to the conjecture that corresponding transformations will allow one to also pose similar problems in the calculus of variations that involve a greater number of independent variables. In fact, I have found the following theorem, whose development defines the content of the present paper:

The second variation of an arbitrary multiple integral will always lead back, by partial integration, to the integral of a homogeneous function of second order whose arguments correspond to the respective highest differential quotients of the variations of the dependent variables, while these arguments are likewise coupled to each other by a series of partial differential equations.

Here, as in the aforementioned paper, I will also first consider integrals that include only the first derivatives of the dependent functions, and arbitrarily many partial differential equations of the first order can exist between these functions themselves as condition equations. At the conclusion, I will briefly go on to the more general case, which can always be reduced, from the aforementioned.

§ 1.

We denote an arbitrary multiple integral by *V*, which will be a maximum or a minimum. *F* will denote the function under the integral sign, which contains the independent variables x^1, x^1, …, x^r and the dependent variables $y^{(1)}$, $y^{(2)}$, …, $y^{(n)}$, along with their first derivatives $\partial y / \partial x$, in such a way that:

$$V = \int^{(r)} F \, dx_1 \, dx_2 \ldots dx_r \,. \tag{1}$$

The functions y, which are determined in such a way that the first variation of V vanishes, may be coupled to each other by means of a series of partial differential equations of first order, which shall be denoted by:

$$\varphi_1 = 0, \qquad \varphi_2 = 0, \ldots, \qquad \varphi_r = 0. \tag{2}$$

Furthermore, we set:

$$\Omega = F + \lambda_1 \varphi_1 + \lambda_2 \varphi_2 + \ldots + \lambda_r \varphi_r \,, \tag{3}$$

where the λ mean certain multipliers, so one also has:

$$V = \int^{(r)} \Omega \; dx_1 \, dx_2 \ldots dx_r \,, \tag{4}$$

and the y, λ find their determination by means of equations (2) when they are linked with the following ones:

$$\begin{cases} \dfrac{\partial \Omega}{\partial y^{(1)}} = \sum_m \dfrac{\partial}{\partial x_m} \dfrac{\partial \Omega}{\partial \dfrac{\partial y^{(1)}}{\partial x_m}}, \\ \dfrac{\partial \Omega}{\partial y^{(2)}} = \sum_m \dfrac{\partial}{\partial x_m} \dfrac{\partial \Omega}{\partial \dfrac{\partial y^{(2)}}{\partial x_m}}, \\ \qquad \cdots \\ \dfrac{\partial \Omega}{\partial y^{(r)}} = \sum_m \dfrac{\partial}{\partial x_m} \dfrac{\partial \Omega}{\partial \dfrac{\partial y^{(r)}}{\partial x_m}}, \end{cases} \tag{5}$$

while equations (2) can also be represented in an analogous way by means of:

$$\frac{\partial \Omega}{\partial \lambda_1} = 0, \qquad \frac{\partial \Omega}{\partial \lambda_2} = 0, \ldots \quad , \frac{\partial \Omega}{\partial \lambda_\kappa} = 0. \tag{6}$$

The resolution of the question of whether some particular solution of these equations makes the integral V a maximum or a minimum depends upon the investigation of the second variation. If we let the $y^{(i)}$, λ_h in the expression (4) increase by small quantities $\varepsilon w^{(i)}$, $\varepsilon \mu_h$, and develop them in powers of ε then we obtain:

(7) $$V + dV + d^2V = \int^{(r)} (\Omega + \varepsilon\Omega_1 + \varepsilon^2\Omega_2 \dots)\ dx_1\, dx_2 \dots dx_r\,,$$

except for a piece that can be brought about by the variations on the boundary, and can only exist in integrals of lower order. The expression:

$$\int^{(r)} \Omega_1\, dx_1\, dx_2 \dots dx_r$$

is, due to equations (5), likewise soluble in integrals of lower order by partial integrations. However, the sign of the second variation, i.e., of:

(8) $$\delta^2 V = \varepsilon^2 \int^{(r)} \Omega_2\, dx_1\, dx_2 \dots dx_r\,,$$

is decisive for the existence of a maximum or minimum. This is the function that will now be examined more closely.

§ 2.

Ω_2 is a homogeneous function of second order of the $n(r + 1) + \kappa$ quantities w, $\partial w / \partial x$, μ, such that the last one enters in only in a linear way, and indeed, one can represent Ω_2 in terms of Ω_1 and Ω in such a way that:

(9) $$\begin{cases} \Omega = \sum_i w^{(i)} \dfrac{\partial\Omega}{\partial y^{(i)}} + \sum_i \sum_m \dfrac{\partial w^{(i)}}{\partial x_m} \dfrac{\partial\Omega}{\partial \dfrac{\partial y^{(i)}}{\partial x_m}} + \sum_h \mu_h \dfrac{\partial\Omega}{\partial\lambda_h}, \\[2ex] 2\Omega_2 = \sum_i w^{(i)} \dfrac{\partial\Omega_1}{\partial y^{(i)}} + \sum_i \sum_m \dfrac{\partial w^{(i)}}{\partial x_m} \dfrac{\partial\Omega_1}{\partial \dfrac{\partial y^{(i)}}{\partial x_m}} + \sum_h \mu_h \dfrac{\partial\Omega_1}{\partial\lambda_h}. \end{cases}$$

However, since the functions y are further coupled to each other by equations (2), which must also be fulfilled by the functions $y + \varepsilon w$, one obtains a series of condition equations for the w which, when one goes from the functions φ to the $\varphi + \varepsilon\psi$ by variation, can be represented by:

(10) $$\psi_1 = 0, \qquad \psi_2 = 0, \qquad \dots, \psi_\kappa = 0,$$

or also, which amounts to the same thing, by:

(11) $$\frac{\partial\Omega_1}{\partial\lambda_1} = 0, \qquad \frac{\partial\Omega_1}{\partial\lambda_2} = 0, \qquad \dots, \frac{\partial\Omega_1}{\partial\lambda_\kappa} = 0.$$

These equations show that the pieces of Ω_2 that are multiplied by μ vanish.

I now pose the problem of converting the expression $\delta^2 V$ by partial integration in such a way that in place of the function Ω, which contains $n(r + 1)$ arguments w, $\partial w / \partial x$, another one enters, in which only nr arguments appear. Ω_2 must then be decomposed into one piece that can be solved completely into an aggregate of integrals of lower order by partial integration and another one whose nr arguments must be represented as linear functions of the w and $\partial w / \partial x$.

However, the first part of the function Ω_2 must necessarily have the form:

$$\frac{\partial B}{\partial x_1} + \frac{\partial B}{\partial x_2} + \cdots + \frac{\partial B}{\partial x_r} = \Theta(w), \tag{12}$$

where the B mean homogeneous functions of second order in the w and $\partial w / \partial x$. However, if one imagines that the differentiations in Θ have been performed then the second derivatives of the w are not present in them. From this, it emerges that, firstly, the $\partial w / \partial x$ may enter into the B only in a linear way, and, secondly, that the coefficient of $w^{(i)} \cdot \partial w^{(h)} / \partial x_m$ in B_s and the coefficient of $w^{(i)} \cdot \partial w^{(h)} / \partial x_s$ in B_m must be equal and opposite. Thus, $\Omega(w)$ goes to a function of second order in the w and $\partial w / \partial x$ that has the property that it contains the second dimensions of the latter only in the combinations:

$$\frac{\partial w^{(i)}}{\partial x_n} \cdot \frac{\partial w^{(h)}}{\partial x_m} - \frac{\partial w^{(i)}}{\partial x_m} \cdot \frac{\partial w^{(h)}}{\partial x_n} .$$

The aggregate of these terms of higher order in Θ shall be denoted by $(\Theta\ (\partial w / \partial x))$.

Now, regardless of the form in which the nr new arguments can also be represented, one can consider the equations, with whose help they are composed from the w, $\partial w / \partial x$, and which can then always be solved for the nr quantities $\partial w / \partial x$, and the linear combinations of the new arguments that enter into the equations thus solved, to be the new arguments themselves. If we thus denote them by $W_m^{(i)}$ then they will have the form:

$$W_m^{(i)} = \frac{\partial w^{(i)}}{\partial x_m} + \alpha_m^{i,1} w^{(1)} + \alpha_m^{i,2} w^{(2)} + \cdots + \alpha_m^{i,n} w^{(n)} , \tag{13}$$

where the a still mean the coefficients to be determined, whose number adds up to $n^2 r$.

If one now, in analogy to the foregoing, denotes the aggregate of terms of higher order in Ω_2 by $(\Omega_2\ (\partial w / \partial x))$, then one will see, with no further assumptions, that the part of Ω_2 that remains after performing a partial integration can be nothing but:

$$(\Omega_2\ (W)) - \Theta(W)),$$

and that one must therefore have the equation:

$$\Omega_2 = (\Omega_2\ (W)) - \Theta(W)) + \Theta(w), \tag{14}$$

which infers the desired transformation in itself. The terms of higher order in it already agree; the coefficients of the $w^{(i)} \cdot w^{(h)}$ and $w^{(i)} \cdot \partial w^{(h)} / \partial x_m$ then give a series of equations for the determination of the *a* and *B*.

However, equation (14) is not necessarily an identity, but may become one only with the help of the condition equations (10) or (11) that link the *w* to each other. Thus, in order for equation (14) to become an identity one must add the sum of the expressions (11), when they are multiplied by linear factors of the *w* whose coefficients are arbitrary. However, we remark that in formula (9), Ω_2 includes the vanishing term:

$$\sum_h \mu_h \frac{\partial \Omega_1}{\partial \lambda_h},$$

which is an expression of just the form imagined, except that here the μ enter in place of linear functions of *w*. Therefore, if one lets this term in Ω_2 remain then one can employ it in order to make equation (14) into an identity, in such a way that one represents the μ as linear functions of the *w*, and the problem can be expressed thus:

Equation (14) *shall be fulfilled identically when the expressions* μ *are given in the form:*

(15) $$\mu_h = M_h^{(1)} w^{(1)} + M_h^{(2)} w^{(2)} + \cdots + M_h^{(n)} w^{(n)}.$$

§ 3.

However, the problem that was contained in § 2 is still not well-defined, as one can easily see.

It is also convenient from the outset to impose certain other demands upon the desired transformation. In fact, it is preferable that *all of the condition equations that exist between the W can be represented in terms of only them,* since otherwise it would be necessary for one to return to the *w*. The consideration of the second variation then has nothing whatsoever to do with the *w* any more, but only the *W* that seem to be coupled with each other by means of certain ordinary equations and differential equations.

This remark suffices to not only determine the problem posed, but also to make it soluble, which is, in no way, true in general for the indeterminate case.

The conditions that this yields are of two types: The first one arises from the fact that the functions *y* must be representable as linear functions of the *W* alone, or the fact that the coefficients of the *w* vanish when one eliminates the quantities $\partial w / \partial x$ by means of equations (13). For each *y*, one obtains *n* equations in this way.

The other type of condition equations arise from the fact that equations (13) should be replaced with certain partial differential equations that exist between the *W* without the help of the *w*. One defines them when one next eliminates the differential quotients of *w* from $W_m^{(i)}$ and $W_k^{(i)}$. This then yields:

$$\frac{\partial W_m^{(i)}}{\partial x_k}-\frac{\partial W_k^{(i)}}{\partial x_m}= \sum_h w^{(h)}\left(\frac{\partial \alpha_m^{i,h}}{\partial x_k}-\frac{\partial \alpha_k^{i,h}}{\partial x_m}\right)+\sum_h\left(\alpha_m^{i,h}\frac{\partial w^{(h)}}{\partial x_k}-\alpha_k^{i,h}\frac{\partial w^{(h)}}{\partial x_m}\right).$$

In the last term on the right-hand side, one can once more eliminate $\partial w^{(h)} / \partial x_k$ and $\partial w^{(h)} / \partial x_m$ with the help of equations (13), and it then becomes:

$$-\sum_h\sum_s w^{(s)}\left(\alpha_m^{i,h}\alpha_k^{h,s}-\alpha_k^{i,h}\alpha_m^{h,s}\right)+\sum_h\left(\alpha_m^{i,h}W_k^{(h)}-\alpha_k^{i,h}W_m^{(h)}\right).$$

If one is to then obtain partial differential equations between the W alone then it will be necessary that the coefficients of the w must vanish in this equation and that the α must therefore be linked to each other by the equations:

$$\frac{\partial \alpha_m^{i,s}}{\partial x_k}+\sum_h \alpha_k^{i,h}\alpha_m^{h,s} = \frac{\partial \alpha_k^{i,s}}{\partial x_m}+\sum_h \alpha_m^{i,h}\alpha_k^{h,s}, \tag{16}$$

while completely similar equations then exist between the W:

$$\frac{\partial W_m^{(i)}}{\partial x_k}+\sum_h \alpha_k^{i,h}W_m^{(h)} = \frac{\partial W_k^{(i)}}{\partial x_m}+\sum_h \alpha_m^{i,h}W_k^{(h)}. \tag{17}$$

The meaning of equations (18) shall now be examined and the general form that the α assume as a result will be deduced; however, it will thus be shown that for any system of W that satisfy the equations there is a system of w that satisfies equations (13) and *thus, in fact, can be regarded as completely equivalent to equations* (13) *and* (17).

§ 4.

Equation (16) represents a system of $n^2 \cdot (r \cdot r - 1)/2$ equations. If we now multiply each equation with any quantity u^s and take the sum over s then (16) gives the equation:

$$\sum_s \frac{\partial \alpha_m^{i,s}}{\partial x_k}u^s+\sum_h\sum_s \alpha_k^{i,h}\alpha_m^{h,s}u^s =\sum_s \frac{\partial \alpha_k^{i,s}}{\partial x_m}u^s+\sum_h\sum_s \alpha_m^{i,h}\alpha_k^{h,s}u^s .$$

If we add to the two sides the equal expressions:

$$\frac{\partial^2 u^i}{\partial x_k\,\partial x_m}+\sum_s \alpha_m^{i,s}\frac{\partial u^s}{\partial x_k}+\sum_h \alpha_k^{i,h}\frac{\partial u^h}{\partial x_m} = \frac{\partial^2 u^i}{\partial x_k\,\partial x_m}+\sum_s \alpha_k^{i,s}\frac{\partial u^s}{\partial x_m}+\sum_h \alpha_m^{i,h}\frac{\partial u^h}{\partial x_k},$$

and likewise set:

$$A_m^i = \frac{\partial u^i}{\partial x_m}+\alpha_m^{i,1}u^1+\alpha_m^{i,2}u^2+\ldots+\alpha_m^{i,n}u^n \tag{18}$$

then the equation above assumes the form:

$$(19) \qquad \frac{\partial A_m^i}{\partial x_k} + \sum_h \alpha_k^{i,h} A_m^h = \frac{\partial A_k^i}{\partial x_m} + \sum_h \alpha_m^{i,h} A_k^h .$$

It is now clear that, with no further assumptions, these equations can enter in place of equations (16) when one only introduces n different systems of u whose determinant is non-zero. If we denote the u that belong to one of these systems by:

$$u^{1,\sigma}, u^{2,\sigma}, \ldots, u^{n,\sigma}$$

then we must similarly distinguish n systems of A that are determined by the equation:

$$(20) \qquad A_m^{i,\sigma} = \frac{\partial u^{i,\sigma}}{\partial x_m} + \alpha_m^{i,1} u^{1,\sigma} + \alpha_m^{i,2} u^{2,\sigma} + \ldots + \alpha_m^{i,n} u^{n,\sigma} ,$$

and in place of equation (19) the following one comes into view:

$$(21) \qquad \frac{\partial A_m^{i,\sigma}}{\partial x_k} + \sum_h \alpha_k^{i,h} A_m^{h,\sigma} = \frac{\partial A_k^{i,\sigma}}{\partial x_m} + \sum_h \alpha_m^{i,h} A_k^{h,\sigma} .$$

However, the A indeed still contain the n^2 completely arbitrary quantities u. I can think of them as being determined such that all of the A_1 vanish; i.e., such that they represent n mutually independent systems of solutions of the simultaneous equations:

$$(22) \qquad 0 = \frac{\partial u^i}{\partial x_m} + \alpha_m^{i,1} u^1 + \alpha_m^{i,2} u^2 + \ldots + \alpha_m^{i,n} u^n .$$

Let us see what values the remaining functions A can therefore assume.

If we set $k = 1$ in (21) then we have for every value of m that is different from 1:

$$0 = \frac{\partial A_m^{i,\sigma}}{\partial x_m} + \alpha_m^{i,1} A_m^{1,\sigma} + \alpha_m^{i,2} A_m^{2,\sigma} + \ldots + \alpha_m^{i,n} A_m^{n,\sigma} ,$$

which, when combined with equation (22) shows, with no further assumptions, that the $A_m^{i,\sigma}$ can be nothing other than solutions of this system; i.e., linear functions of:

$$u^{i,1}, u^{i,2}, \ldots, u^{i,n},$$

or finally, that:

$$(23) \qquad A_m^{i,\sigma} = \beta_m^{i,\sigma} u^{i,1} + \beta_m^{2,\sigma} u^{i,2} + \ldots + \beta_m^{n,\sigma} u^{i,n} ,$$

where the β are independent of x_1 .

We now introduce these values into those of equations (21) that have either the value m or k. This gives:

$$(24) \qquad \sum_\rho \frac{\partial \beta_m^{i,s} u^{i,\sigma}}{\partial x_k} + \sum_h \sum_\rho \alpha_k^{i,h} \beta_m^{\rho,\sigma} u^{h,\rho} = \sum_\rho \frac{\partial \beta_k^{\rho,\sigma} u^{i,\rho}}{\partial x_m} + \sum_h \sum_\rho \alpha_m^{i,h} \beta_k^{\rho,\sigma} u^{h,\rho} .$$

For the sum $\sum_h \alpha_k^{i,h} u^{h,\rho} + \frac{\partial u^{i,\rho}}{\partial x_k}$, which is nothing but $A_k^{i,\rho}$, we can again set it to the value in (23), and thereby reduce the above equation to:

$$(25) \qquad \sum_\rho u^{i,\rho} \left(\frac{\partial \beta_m^{\rho,\sigma}}{\partial x_k} + \sum_\tau \beta_m^{\tau,\sigma} \beta_k^{\rho,\tau} \right) = \sum_\rho u^{i,\rho} \left(\frac{\partial \beta_k^{\rho,\sigma}}{\partial x_m} + \sum_\tau \alpha_k^{\tau,\sigma} \beta_m^{\rho,\tau} \right).$$

However, since the determinant of the u in these equations must not vanish, by assumption, one must fulfill the equations:

$$(26) \qquad \frac{\partial \beta_m^{\rho,\sigma}}{\partial x_k} + \sum_\tau \beta_m^{\tau,\sigma} \beta_k^{\rho,\tau} = \frac{\partial \beta_k^{\rho,\sigma}}{\partial x_m} + \sum_\tau \alpha_k^{\tau,\sigma} \beta_m^{\rho,\tau} ,$$

which must be completely free of the independent variable x_1, and thus differs from the system (16) only in that the latter involves the α.

One can now apply the same process to these β that just served for the representation of the α. We can thus multiply equations (26) by n mutually independent systems:

$$v^{1,\lambda}, v^{2,\lambda}, \ldots, v^{n,\lambda},$$

where the v, however, are independent of x_1, in order to obtain equations that are analogous to (21), and analogously introduce new functions B for the A such that:

$$(27) \qquad B_m^{i,\lambda} = \frac{\partial v^{i,\lambda}}{\partial x_m} + \beta_m^{i,1} v^{1,\lambda} + \beta_m^{i,2} v^{2,\lambda} + \cdots,$$

and when one imagines that the v are determined such that all of the B_2 vanish, one obtains, for $m = 3, 4, \ldots, n$, the equation:

$$(28) \qquad B_m^{i,\lambda} = \gamma_m^{1,\lambda} v^{i,1} + \gamma_m^{2,\lambda} v^{i,2} + \cdots,$$

where the γ are independent of x_1, x_2, and must satisfy a system of equations that is completely the same as the systems (16), (26), but includes either x_1 or x_2, as well.

Proceeding in this way, one gradually arrives at the complete representation of the α, which are expressed by:

n^2 functions $u^{i,\sigma}$ of $x_1, x_2, x_3, \dots, x_n$,
n^2 functions $v^{i,\sigma}$ of $x_2, x_3, \dots, x_n$,
n^2 functions $w^{i,\sigma}$ of $x_3, \dots, x_n$,

etc.

I would now like to show that, without compromising the generality of the solution, one can let all of these functions vanish up to the u.

Namely, if we multiply equation (23) by $v^{\sigma,\lambda}$ and sum over σ then, considering (27), one comes to:

$$\sum_\sigma A_m^{i,\sigma} v^{\sigma,\lambda} = u^{i,1} B_m^{1,\lambda} + u^{i,2} B_m^{2,\lambda} + \cdots - \left\{ u^{i,1} \frac{\partial v^{1,\lambda}}{\partial x_m} + u^{i,2} \frac{\partial v^{2,\lambda}}{\partial x_m} + \cdots \right\},$$

or, when one also appeals to (28):

$$(29) \quad \sum_\sigma A_m^{i,\sigma} v^{\sigma,\lambda} + \sum_\sigma u^{i,\sigma} \frac{\partial v^{\sigma,\lambda}}{\partial x_m} = \gamma_m^{1,\lambda} \sum_\sigma u^{i,\sigma} v^{\sigma,1} + \gamma_m^{2,\lambda} \sum_\sigma u^{i,\sigma} v^{\sigma,2} + \cdots \gamma_m^{n,\lambda} \sum_\sigma u^{i,\sigma} v^{\sigma,n},$$

in which the right-hand side vanishes when $m = 1$ or $m = 2$.

The left-hand side of this equation is nothing but the expression that $A_m^{i,\lambda}$ goes to as long as one everywhere uses the expression $\sum_\sigma u^{i,\sigma} v^{\sigma,\lambda}$ in place of $u^{i,\lambda}$. One then sees that everywhere the functions u, v enter only in the combinations:

$$\sum_\sigma u^{i,\sigma} v^{\sigma,\lambda},$$

and since the $u^{i,\sigma}$, which represent n independent solutions of equations (23), can possess no other property – a property that likewise corresponds to these combination – one can then obviously use these combinations immediately in place of the $u^{i,\sigma}$, and denote them by $u^{i,\sigma}$. However, equations (29) then show that, *with no loss of generality, the* A_1, *as well as the* A_2, *can be set equal to zero, and the remaining ones can be expressed by the equation:*

$$A_m^{i,\lambda} = \gamma_m^{1,\lambda} u^{i,1} + \gamma_m^{2,\lambda} u^{i,2} + \cdots + \gamma_m^{n,\lambda} u^{i,n},$$

where the γ *are independent of the* x_1, *as well as the* x_2.

One needs only to repeat the same argument in order to show that one can also let all of the A_3 vanish and one can then reduce the remaining ones to linear expressions in the u whose coefficients are also independent of x. One finally arrives at the conclusion that all of the A_m can be set equal to zero, and one has thus proved the theorem:

The most general values of the functions α *that satisfy the equations:*

$$\frac{\partial \alpha_m^{i,s}}{\partial x_k} + \sum_h \alpha_k^{i,h} \alpha_m^{h,s} = \frac{\partial \alpha_k^{i,s}}{\partial x_m} + \sum_h \alpha_m^{i,h} \alpha_k^{h,s}$$

are the ones that one arrives at upon determining the α *from the equations:*

$$(30) \qquad 0 = \frac{\partial u^{i,\sigma}}{\partial x_k} + \alpha_k^{i,1} u^{1,\sigma} + \alpha_k^{i,2} u^{2,\sigma} + \cdots + \alpha_k^{i,n} u^{n,\sigma},$$

where the u represent n^2 *completely arbitrary functions whose determinant does not vanish.*

§ 5.

One now easily connects the foregoing with the proof that the expression for W that is given by equations (13) also represents the most general functions that satisfy equations (17).

From the derivation of these equations, it next follows that they, in fact, must be fulfilled when one sets:

$$(31) \qquad \begin{cases} W_m^{(i)} = \dfrac{\partial w^{(i)}}{\partial x_m} + \alpha_m^{i,1} w^{(1)} + \alpha_m^{i,2} w^{(2)} + \cdots + \alpha_m^{i,n} w^{(n)} \\ \quad\;\; = \mathrm{W}_m^{(i)}, \end{cases}$$

where the w mean any arbitrary functions, and since the equations (17) are linear, they obviously always remain true when one introduces the difference $W_m^{(i)} - \mathrm{W}_m^{(i)}$, instead of $W_m^{(i)}$.

Now, it is undoubtedly always possible to determine the functions w in such a way that this difference $W_m^{(i)} - \mathrm{W}_m^{(i)}$ vanishes for $m = 1$, which also might be the actual values for the W. However, it then follows from equations (17) for $k = 1$ and an m that is different from 1 that:

$$\frac{\partial (W_m^{(i)} - \mathrm{W}_m^{(i)})}{\partial x_1} - \sum_h \alpha_1^{i,h} (W_m^{(i)} - \mathrm{W}_m^{(i)}) = 0;$$

i.e., $W_m^{(i)} - \mathrm{W}_m^{(i)}$ must be a linear function of the solutions of (22), or:

$$(32) \qquad W_m^{(i)} - \mathrm{W}_m^{(i)} = b_m^{(1)} u^{i,1} + b_m^{(2)} u^{i,2} + \cdots + b_m^{(n)} u^{i,n},$$

where the b_m are independent of x_1 .

Furthermore, we substitute these values into those of equations (17) for which either $m = 1$ or $k = 1$. We then first obtain:

$$\sum_\rho \frac{\partial (b_m^\rho u^{i,\rho})}{\partial x_k} + \sum_h \sum_\rho \alpha_k^{i,h} b_m^\rho u^{h,\rho} = \sum_\rho \frac{\partial (b_k^\rho u^{i,\rho})}{\partial x_m} + \sum_h \sum_\rho \alpha_k^{i,h} b_m^\rho u^{h,\rho} .$$

Here, however, the coefficients of b_m^ρ and b_k^ρ vanish as a result of equations (30), which define the α.

Thus, all that remains is:

$$\sum_\rho u^{i,\rho}\left(\frac{\partial b_m^\rho}{\partial x_k}-\frac{\partial b_k^\rho}{\partial x_m}\right)=0,$$

or, since the determinant of the *u* may not vanish:

$$\frac{\partial b_m^\rho}{\partial x_k}=\frac{\partial b_k^\rho}{\partial x_m},$$

or

(33) $$b_m^\rho=\frac{\partial c^\rho}{\partial x_m}.$$

However, when this is introduced into the right-hand side of equation (32), this takes the form:

$$\frac{\partial c^{(1)}}{\partial x_m}u^{i,1}+\frac{\partial c^{(2)}}{\partial x_m}u^{i,2}+\cdots+\frac{\partial c^{(n)}}{\partial x_m}u^{i,n},$$

or also:

$$\frac{\partial(c^{(1)}u^{i,1}+c^{(2)}u^{i,2}+\cdots+c^{(n)}u^{i,n})}{\partial x_m}-\left(c^{(1)}\frac{\partial u^{i,1}}{\partial x_m}+c^{(2)}\frac{\partial u^{i,2}}{\partial x_m}+\cdots+c^{(n)}\frac{\partial u^{i,n}}{\partial x_m}\right),$$

or finally, when one substitutes the values of the $\partial u / \partial x$ from equations (30):

$$\frac{\partial}{\partial x_m}\left(\sum_\rho c^\rho u^{i,\rho}\right)+\sum_h \alpha_m^{i,h}\left(\sum_\rho c^\rho u^{i,\rho}\right),$$

which is an expression of the form of the $\mathrm{W}_m^{(i)}$, except that the $\sum_\rho c^\rho u^{i,\rho}$ enter in place of the *w*, here.

We thus see from equation (32) that we arrive at the general expression for *W* when we substitute $w^{(i)}+\sum_\rho c^\rho u^{i,\rho}$ in place of $w^{(i)}$ in the expression for W, and since this expression is no more general than the completely arbitrary function $w^{(i)}$, one can, in turn, denote them by $w^{(i)}$, and thus arrive at the following theorem:

The most general values for the functions $W_m^{(i)}$ that satisfy the equations:

$$\frac{\partial W_m^{(i)}}{\partial x_k}+\sum_h \alpha_k^{i,h}W_m^{(h)}=\frac{\partial W_k^{(i)}}{\partial x_m}+\sum_h \alpha_m^{i,h}W_k^{(h)},$$

where the α are defined by the equations:

$$0 = \frac{\partial u^{i,\sigma}}{\partial x_k} + \alpha_k^{i,1} u^{1,\sigma} + \alpha_k^{i,2} u^{2,\sigma} + \cdots + \alpha_k^{i,n} u^{n,\sigma},$$

and the u represent completely arbitrary functions, are:

$$W_m^{(i)} = \frac{\partial w^{(i)}}{\partial x_k} + \alpha_k^{i,1} w^{(1)} + \alpha_k^{i,2} w^{(2)} + \cdots + \alpha_k^{i,n} w^{(n)},$$

where the w mean arbitrary functions.

§ 6.

After these preparations, we are finally in a position to completely formulate the problem that defined the actual objective of these investigations.

Convert the second variation:

$$\int^{(r)} \Omega_2 dx_1 dx_2 \cdots dx_r$$

by partial integration into the integral of a homogeneous function $W_m^{(i)}$ *with nr arguments that are linked with the previous arguments* $w^{(n)}$ *by the equations:*

$$W_m^{(i)} = \frac{\partial w^{(i)}}{\partial x_m} + \alpha_m^{i,1} w^{(1)} + \alpha_m^{i,2} w^{(2)} + \cdots,$$

while the coefficients α *are converted into* n^2 *functions u with the help of the equations:*

$$0 = \frac{\partial u^{i,\sigma}}{\partial x_m} + \alpha_m^{i,1} u^{1,\sigma} + \alpha_m^{i,2} u^{2,\sigma} + \cdots,$$

and furthermore, the κ *linear functions* $\partial\Omega / \partial\lambda$ *go to linear functions of the W.*

From (14) above, the transformation of the function Ω_2 is expressed by the equations:

(34) $$\Omega_2 = (\Omega_2 (W)) - (\Theta(W)) + \Theta(w),$$

in which the m are replaced with the expressions (15), and in which, ultimately, the coefficients of the $\frac{\partial w^{(i)}}{\partial x_k} \cdot \frac{\partial w^{(h)}}{\partial x_s}$ already agree.

If we denote, for the moment, the right-hand side of (34) by Φ then this yields linear equations for the w and $\partial w / \partial x$:

$$(35)\qquad \begin{cases} \dfrac{\partial \Omega_2}{\partial w^{(i)}} = \dfrac{\partial \Phi}{\partial w^{(i)}}, \\ \dfrac{\partial \Omega_2}{\partial \dfrac{\partial w^{(i)}}{\partial x_m}} = \dfrac{\partial \Phi}{\partial \dfrac{\partial w^{(i)}}{\partial x_m}}, \end{cases}$$

whose coefficients on both sides must agree. From the second of these equations, one deduces only $n^2 r$ equations, since the coefficients of the $\dfrac{\partial w^i}{\partial x_m}$ in them have already been made to coincide, and from the first equation, one then deduces $n(n + 1)/2$ new equations that originate in the coefficients of the products $w^{(i)} \cdot w^{(h)}$ in (34).

In any case, it must then suffice when one fulfills the first n equations (35) for n mutually independent systems of values of the w, and then show that the second equations can still be completely satisfied.

Instead of the first equations (35), one can, however, also appeal to the following ones, which are defined with the help of the second equation:

$$(36)\qquad \frac{\partial \Omega}{\partial w^{(i)}} - \sum_m \frac{\partial}{\partial x_m} \frac{\partial \Omega_2}{\partial \dfrac{\partial w^{(i)}}{\partial x_m}} = \frac{\partial \Phi}{\partial w^{(i)}} - \sum_m \frac{\partial}{\partial x_m} \frac{\partial \Phi}{\partial \dfrac{\partial w^{(i)}}{\partial x_m}}.$$

In this equation, I now successively introduce the various systems of u for the w and thus obtain n^2 equations from (36). Let us see what form the right-hand side takes as a result of this.

As a consequence of equations (13) and (30), the functions u have the property that when they replace the w, the W all vanish identically. The part of Φ that has the W for its arguments then vanishes completely from (36), and all that of Φ that remains is $\Theta(w)$. However, $\Theta(w)$ is a homogeneous function that admits the one-fold integration in all of its parts, so the equations:

$$(37)\qquad \frac{\partial \Theta}{\partial w^{(i)}} - \sum_m \frac{\partial}{\partial x_m} \frac{\partial \Theta}{\partial \dfrac{\partial w^{(i)}}{\partial x_m}} = 0$$

must be fulfilled identically, and the right-hand side of (36) vanishes completely. If one now lets Ω_2^σ denote the function that Ω_2 goes to when the $w^{(i)}$ in it are replaced with $u^{i,\sigma}$ then this shows that the $u^{i,\sigma}$ must then satisfy the equations:

$$(38)\qquad \frac{\partial \Omega_2^\sigma}{\partial u^{i,\sigma}} - \sum_m \frac{\partial}{\partial x_m} \frac{\partial \Omega_2^\sigma}{\partial \dfrac{\partial u^{i,\sigma}}{\partial x_m}} = 0.$$

Here, it is necessary to make a remark: The functions μ enter into Ω_2, which must be set equal to the expression (15) in order to make equation (34) an identity. Therefore, by the differentiations in (35), (36), (38), the μ must also be regarded as functions of the *w*.

In fact, however, the part of equations (38) that arises by a differentiation with respect to μ vanishes completely. The part of $\dfrac{\partial \Omega_2}{\partial w^{(i)}}$ in question, is, in fact:

$$\sum_h \frac{\partial \Omega_1}{\partial \lambda_h} \cdot \frac{\partial \mu_h}{\partial u^{(i)}} .$$

Now, since the expressions $\dfrac{\partial \Omega_1}{\partial \lambda_h}$ should likewise be regarded as linear functions of the *W*, they must obviously vanish identically when one goes from the $w^{(i)}$ to the $u^{i,\sigma}$; i.e., the *W* all vanish identically. Therefore, one can consider the μ to be completely independent of the *u* under the differentiation in (38), and they alone take on the system of values:

$$\mu_1^\sigma,\ \mu_2^\sigma,\ \ldots,\ \mu_\kappa^\sigma$$

under the introduction of the *u* into (15).

However, one must likewise fulfill the equations:

$$\text{(39)} \qquad \frac{\partial \Omega_1^\sigma}{\partial \lambda_1} = 0, \qquad \frac{\partial \Omega_1^\sigma}{\partial \lambda_2} = 0, \qquad \ldots, \frac{\partial \Omega_1^\sigma}{\partial \lambda_\kappa} = 0,$$

which are $n\kappa$ in number, and are thus the complete expression of the idea that the functions $\dfrac{\partial \Omega_1}{\partial \lambda}$ can be represented as linear combinations of the *W*, without recourse to the *w*.

Equations (38), (39) together now serve to determine the functions $u^{i,\sigma}$ and μ_h^σ, or, if one would like, the functions *u* and *M*, whose number is just as large, and indeed one has $n \cdot \kappa + n^2$ equations in just as many quantities. However, the systems of unknowns, divided by the various corresponding values of σ, are present in these equations, and always in the same way. One thus has the remarkable result, which represents the immediate extension of the argument that was introduced by *Jacobi:*

The different systems of $u^{(i)}$ and μ_h are just as numerous as the different solutions of the system of partial differential equations:

$$(40)\qquad \begin{cases} \dfrac{\partial\Omega_2}{\partial u^{(i)}} = \sum_m \dfrac{\partial}{\partial x_m} \dfrac{\partial\Omega_2}{\partial \dfrac{\partial u^{(i)}}{\partial x_m}}, \\[2ex] \dfrac{\partial\Omega_1}{\partial\lambda_h} = \dfrac{\partial\Omega_2}{\partial\mu_h} = 0 \end{cases}$$

that make the integral:

$$\delta^2 V = \varepsilon^2 \int^r \Omega_2\, dx_1\, dx_2 \dots dx_r$$

assume a maximum or minimum value.

§ 7.

One can now easily represent the solutions to equations (40) when the solutions of equations (5), (6) are assumed to be known.

Let $y^{(1)}, y^{(2)}, \dots, y^{(n)}, \lambda_1, \lambda_2, \dots, \lambda_n$ be the most general solutions of equations (5), (6). Let a be any arbitrary constant that enters into them, and let P an arbitrary function that enters into them. One then obviously obtains new solutions when one lets the constant a go to $a + \varepsilon\alpha$ and the function P go to $P + \varepsilon\Pi$, where α is a new arbitrary constant and Π is a new arbitrary function whose arguments coincide with those of P, and ε means a very small quantity. Now, under these operations, when one develops them in powers of ε, the y go to $y + \varepsilon u + \dots$, the λ go to $\lambda + \varepsilon\mu + \dots$, and one has the equations:

$$(41)\qquad \begin{cases} u^{(i)} = \sum \left(\alpha \dfrac{\partial y^{(i)}}{\partial a} + \Pi \dfrac{\partial y^{(i)}}{\partial P} \right), \\[2ex] \mu_h = \sum \left(\alpha \dfrac{\partial \lambda_h}{\partial a} + \Pi \dfrac{\partial \lambda_h}{\partial P} \right), \end{cases}$$

where the sums are extended over all of the arbitrary constants and functions that are included in the $y^{(i)}$, λ_h . Likewise, however, Ω goes to:

$$\Omega + \varepsilon\,\Omega_1 + \varepsilon^2\,\Omega_2 + \dots,$$

except that the u now enter into the functions Ω_1, Ω_2 in place of the w. Now, one obviously has:

$$\frac{\partial(\Omega + \varepsilon\Omega_1 + \varepsilon^2\Omega_2 + \cdots)}{\partial y^{(i)}} = \frac{1}{\varepsilon} \frac{\partial(\Omega + \varepsilon\Omega_1 + \varepsilon^2\Omega_2 + \cdots)}{\partial u^{(i)}},$$

$$\frac{\partial(\Omega + \varepsilon\Omega_1 + \varepsilon^2\Omega_2 + \cdots)}{\partial \dfrac{\partial y^{(i)}}{\partial x_m}} = \frac{1}{\varepsilon} \frac{\partial(\Omega + \varepsilon\Omega_1 + \varepsilon^2\Omega_2 + \cdots)}{\partial \dfrac{\partial u^{(i)}}{\partial x_m}},$$

$$\frac{\partial(\Omega+\varepsilon\Omega_1+\varepsilon^2\Omega_2+\cdots)}{\partial\lambda_h}=\frac{1}{\varepsilon}\frac{\partial(\Omega+\varepsilon\Omega_1+\varepsilon^2\Omega_2+\cdots)}{\partial\mu_h},$$

and from this, it likewise follows that:

$$\frac{\partial\Omega_1}{\partial y^{(i)}}=\frac{\partial\Omega_2}{\partial u^{(i)}},$$
$$\frac{\partial\Omega_1}{\partial\dfrac{\partial y^{(i)}}{\partial x_m}}=\frac{\partial\Omega_2}{\partial\dfrac{\partial u^{(i)}}{\partial x_m}},$$
$$\frac{\partial\Omega_1}{\partial\lambda_h}=\frac{\partial\Omega_2}{\partial\mu_h},$$

and since the equations (5), (6), when one lets $\Omega + \varepsilon\,\Omega_1 + \varepsilon^2\,\Omega_2 + \dots$ enter into it in place of Ω as coefficients of ε, as is allowed, give the equations:

$$\frac{\partial\Omega_1}{\partial y^{(i)}}=\sum_m\frac{\partial}{\partial x_m}\frac{\partial\Omega_1}{\partial\dfrac{\partial y^{(i)}}{\partial x_m}},$$
$$\frac{\partial\Omega_1}{\partial\lambda_h}=0,$$

one sees with the help of the equations above that the u and μ that were introduced here represent nothing but the general solutions of equations (40). *The different systems of* $u^{i,\sigma}$, μ_h^σ *must therefore be obtained from equations* (41) *when one only attributes different systems of values to the* α *and* Π; *i.e., one must have:*

$$(42)\qquad\begin{cases} u^{i,\sigma}=\sum\left(\alpha^\sigma\dfrac{\partial y^{(i)}}{\partial a}+\Pi^\sigma\dfrac{\partial y^{(i)}}{\partial P}\right), \\ \mu_h^\sigma=\sum\left(\alpha^\sigma\dfrac{\partial\lambda_h}{\partial a}+\Pi^\sigma\dfrac{\partial\lambda_h}{\partial P}\right), \end{cases}$$

from which, these quantities are completely related to known things.

§ 8.

All that still remains is the determination of the function Θ, and this comes about with the help of the second of equations (35).

We recall that each of the equations that represent it can lead to only n equations when one sets the coefficients of the w on both sides equal to each other, so in order to make this equation into an identity it obviously suffices that one adds that n different w fulfill independent systems of values. Let this system of values, in turn, be those of the u. Once again, all of the terms in the function Φ that are independent of the W then vanish, and the second of equations (35) goes to the following one:

$$(43) \qquad \frac{\partial \Omega_2^\sigma}{\partial \frac{\partial u^{i,\sigma}}{\partial x_m}} = \frac{\partial \Theta^\sigma}{\partial \frac{\partial u^{i,\sigma}}{\partial x_m}}.$$

We now go into the behavior of the function Θ more closely. In equation (12), $\Theta(w)$ was represented by the expression:

$$\frac{\partial B_1}{\partial x_1} + \frac{\partial B_2}{\partial x_2} + \cdots + \frac{\partial B_r}{\partial x_r},$$

where the B_h were homogeneous functions of second order of the w and $\partial w / \partial x$, such that the latter appeared only in a linear way and, in addition, the coefficient of $w^{(i)} \cdot \frac{\partial w^{(h)}}{\partial x_s}$ in B_k was equal and opposite to the coefficients of $w^{(i)} \cdot \frac{\partial w^{(h)}}{\partial x_k}$ in B_s. One can thus generally set:

$$(44) \qquad B_k = \sum_i \sum_h b_k^{i,h} w^{(i)} w^{(h)} + \sum_i \sum_h \sum_s a_{k,s}^{i,h} w^{(i)} \frac{\partial w^{(h)}}{\partial x_s},$$

where the coefficients b, a are subject to the conditions:

$$(45) \qquad b_k^{i,h} = b_k^{h,i}, \qquad a_{k,s}^{i,h} = -\, a_{k,s}^{h,i} = -\, a_{s,k}^{i,h} = a_{s,k}^{h,i}.$$

The number of the quantities to be determined is then:

$$\frac{n \cdot n+1}{2} r + \frac{n \cdot n-1}{2} \cdot \frac{r \cdot r-1}{2},$$

and since (43) delivers only $n^2 \cdot r$ equations then the problem is apparently indeterminate. It shows that it suffices to know a single solution and that, for the present purpose, all solutions deliver the same result.

When one substitutes the value in (44) for Θ, one obtains the following equation:

$$(45) \qquad \Theta = \sum_i \sum_h \sum_k \frac{\partial b_k^{i,h}}{\partial x_k} w^{(i)} w^{(h)}$$

$$+ \sum_i \sum_h \sum_s w^{(i)} \frac{\partial w^{(h)}}{\partial x_s} \left(2b_s^{i,h} + \sum_k \frac{\partial a_{k,s}^{i,h}}{\partial x_k} \right)$$

$$+ \sum_i \sum_h \sum_s \sum_k a_{k,s}^{i,h} \frac{\partial w^{(i)}}{\partial x_k} \frac{\partial w^{(h)}}{\partial x_s},$$

and from this, equation (43) goes to:

$$\text{(46)} \qquad \frac{\partial \Omega_2^\sigma}{\partial \dfrac{\partial u^{i,\sigma}}{\partial x_m}} = \sum_h u^{h,\sigma} \left(2b_m^{h,i} + \sum_k \frac{\partial a_{k,s}^{i,h}}{\partial x_k} \right) + 2\sum_h \sum_s a_{s,m}^{h,i} \frac{\partial u^{h,\sigma}}{\partial x_s}.$$

If we multiply this equation by $u^{i,\rho}$ and then take the sum over i then this gives:

$$\sum_i u^{i,\rho} \frac{\partial \Omega_2^\sigma}{\partial \dfrac{\partial u^{i,\sigma}}{\partial x_m}} = 2\sum_h \sum_i b_m^{h,i} u^{i,\rho} u^{h,\sigma}$$

$$+ \sum_h \sum_i u^{i,\rho} u^{h,\sigma} \left(\sum_k \frac{\partial a_{k,m}^{h,i}}{\partial x_k} \right)$$

$$+ 2\sum_h \sum_i \sum_s a_{s,m}^{h,i} u^{i,\rho} \frac{\partial u^{h,\sigma}}{\partial x_s}.$$

If we finally let $(\rho, \sigma)_m$ denote the expression:

$$\text{(47)} \qquad (\rho, \sigma)_m = \tfrac{1}{2} \sum_i \left(u^{i,\rho} \frac{\partial \Omega_2^\sigma}{\partial \dfrac{\partial u^{i,\sigma}}{\partial x_m}} - u^{i,\sigma} \frac{\partial \Omega_2^\rho}{\partial \dfrac{\partial u^{i,\rho}}{\partial x_m}} \right)$$

then this gives:

$$\text{(48)} \qquad (\rho, \sigma)_m = \sum_h \sum_i u^{h,\sigma} u^{i,\rho} \left(\sum_k \frac{\partial a_{k,m}^{h,i}}{\partial x_k} \right) + \sum_h \sum_i \sum_s a_{k,m}^{h,i} \left(u^{i,\rho} \frac{\partial u^{h,\sigma}}{\partial x_k} + u^{h,\sigma} \frac{\partial u^{i,\rho}}{\partial x_k} \right),$$

an equation that makes the functions b vanish identically. We remark that (48) represents $\frac{n \cdot n-1}{2} \cdot r$ equations, and if we think of the a as determined then what remain in the system (46) are $\frac{n \cdot n+1}{2} \cdot r$ equations that achieve the determination of the b completely. In order to perform this determination in a symmetric way, one can now, parallel to the expressions (47), introduce the notations:

(49) $$[\rho,\sigma]_m = 2\sum_i \left(u^{i,\rho} \frac{\partial \Omega_2^\sigma}{\partial \dfrac{\partial u^{i,\sigma}}{\partial x_m}} + u^{i,\sigma} \frac{\partial \Omega_2^\rho}{\partial \dfrac{\partial u^{i,\rho}}{\partial x_m}} \right),$$

which then yields:

(50) $$\sum_h \sum_i b_m^{h,i} u^{i,\rho} u^{h,\sigma} = [\rho,\sigma]_m + \sum_h \sum_i \sum_s a_{s,m}^{h,i} \left(u^{i,\rho} \frac{\partial u^{h,\sigma}}{\partial x_s} - u^{h,\sigma} \frac{\partial u^{i,\rho}}{\partial x_s} \right).$$

However, we shall now turn to the more precise consideration of equations (48).

§ 9.

Under the transformation of the function Ω_2, one comes down to just that part of it that remains under the r-fold integral in reduced form. However, this part is:

$$(\Omega_2(W) - (\Theta(W)),$$

and indeed, with the notation that we have just introduced:

(51) $$(\Theta(W)) = \sum_i \sum_h \sum_k \sum_s a_{k,s}^{i,h} W_k^i W_s^h .$$

However, the present problem has led to the fact that the a do not seem to be completely determined, but admit an infinite number of value systems. I would now like to show that *the various expressions that* $(\Theta(W))$ *can be equal to can differ only by functions that admit the one-fold integration in all of their parts, so for the present purpose they vanish from consideration completely.*

We let c denote the difference between two such corresponding distinct systems of associated a, and let H denote the difference between the functions $(\Theta(W))$ that arise in that way. I would like to prove that H always admits the one-fold integration in all of its parts.

From equation (48), it then follows that:

(52) $$0 = \sum_h \sum_i u^{h,\sigma} u^{i,\rho} \left(\sum_k \frac{\partial c_{k,m}^{h,i}}{\partial x_k} \right) + \sum_h \sum_i \sum_s c_{s,m}^{h,i} \left(u^{i,\rho} \frac{\partial u^{h,\sigma}}{\partial x_k} + u^{h,\sigma} \frac{\partial u^{i,\rho}}{\partial x_k} \right),$$

while from (51), it likewise emerges that:

(53) $$H = \sum_h \sum_i \sum_k \sum_s c_{s,m}^{h,i} W_k^i W_s^h .$$

In order for H to admit the one-fold integration, it is necessary and sufficient that it fulfill the equations:

(54) $$\frac{\partial H}{\partial w^{(p)}} = \sum_m \frac{\partial}{\partial x_m} \frac{\partial H}{\partial \frac{\partial w^{(p)}}{\partial x_m}} .$$

These equations are, however, as one easily shows, nothing but equations (52), from which the assertion is established. In fact, when we define equation (54), we next obtain, considering the value of W:

$$\sum_i \sum_h \sum_k \sum_s c_{k,s}^{i,h} \alpha_k^{i,p} W_s^{(h)} = \sum_m \frac{\partial}{\partial x_m} \left(\sum_h \sum_s c_{m,s}^{p.h} W_s^{(h)} \right)$$
$$= \sum_h \sum_s W_s^{(h)} \left(\sum_m \frac{\partial c_{m,s}^{p.h}}{\partial x_m} \right) + \sum_h \sum_s \sum_m c_{m,s}^{p.h} \frac{\partial W_s^{(h)}}{\partial x_m} .$$

If we now add equations (17), by which the W are defined, then this easily shows, when one always keeps in mind the equation:

$$c_{m,s}^{p,h} = - c_{m,s}^{h,p} = - c_{s,m}^{p,h} = c_{s,m}^{h,p} ,$$

that the second term on the right-hand side of this equation can be replaced by:

$$- \sum_h \sum_s \sum_m \sum_q c_{m,s}^{p,h} \alpha_m^{h,q} W_s^{(q)} ,$$

and in such a way that each equation goes to:

$$0 = \sum_h \sum_s W_s^{(h)} \left\{ \sum_m \frac{\partial c_{m,s}^{p,h}}{\partial x_m} - \sum_i \sum_k c_{k,s}^{i,h} \alpha_k^{i,p} - \sum_i \sum_k c_{k,s}^{p,i} \alpha_k^{i,h} \right\} ,$$

or that ultimately, since the w that are contained in W should remain completely arbitrary, the equations to be fulfilled must be:

$$0 = \sum_m \frac{\partial c_{m,s}^{p,h}}{\partial x_m} - \sum_k \sum_i (c_{k,s}^{i,h} \alpha_k^{i,p} + c_{k,s}^{p,i} \alpha_k^{i,h}) .$$

Instead of the $n^2 \cdot r$ different equations that this equation represents, one can also choose the $n^2 \cdot r$ equations that emerge from them when one multiplies them by $u^{p,\rho}\, u^{h,\sigma}$, and sums over h and p. However, this gives:

$$0 = \sum_\rho \sum_h \sum_m \frac{\partial c_{m,s}^{p,h}}{\partial x_k} u^{p,\rho} u^{h,\sigma}$$

$$-\sum_k\sum_i\left\{\sum_h c_{k,s}^{i,h}\left(\sum_\rho \alpha_k^{i,p} u^{p,\rho}\right)u^{h,\sigma}+\sum_p c_{k,s}^{p,i}\left(\sum_h \alpha_k^{i,h} u^{h,\sigma}\right)u^{p,\rho}\right\},$$

or, with the help of equations (30):

$$0 = \sum_k\sum_\rho\sum_m \frac{\partial c_{k,s}^{i,h}}{\partial x_k} u^{p,\rho}u^{h,\sigma}+\sum_k\sum_i\left\{\sum_h c_{k,s}^{i,h}u^{h,\sigma}\frac{\partial u^{i,\rho}}{\partial x_k}+\sum_p c_{k,s}^{p,i}u^{p,\rho}\frac{\partial u^{i,\rho}}{\partial x_k}\right\},$$

an equation that goes to (52) immediately with a different notation for the indices. The property of the function H that was alluded to is then proved, and likewise, the fact that one must only look for a single function Θ from amongst the infinitude of possibilities.

§ 10.

Finally, in order to actually perform the determination of the a, one can now give equations (48) the form:

$$(\rho,\ \sigma)_m = \sum_k\sum_h\sum_i \frac{\partial}{\partial x_k}(u^{h,\sigma}u^{i,\rho}a_{k,m}^{h,i}),$$

or when one sets:

(55) $$z_{k,m}^{\sigma,\rho} = \sum_h\sum_i\left(u^{h,\sigma}u^{i,\rho}a_{k,m}^{h,i}\right),$$

the form:

(56) $$(\rho,\ \sigma)_m = \frac{\partial z_{1,m}^{\rho,\sigma}}{\partial x_1}+\frac{\partial z_{2,m}^{\rho,\sigma}}{\partial x_2}+\cdots+\frac{\partial z_{r,m}^{\rho,\sigma}}{\partial x_r}.$$

The functions z, along with the a, then have the common property:

(57) $$z_{k,m}^{\sigma,\rho} = z_{m,k}^{\rho,\sigma} = -\ z_{k,m}^{\rho,\sigma} = -\ z_{m,k}^{\sigma,\rho},$$

and the a are determined from them easily by means of linear equations.

Equation (56) is connected with an important property of the functions $(\rho,\ \sigma)_m$. Namely, if we differentiate this equation with respect to x_m and take the sum over m then we obtain:

(58) $$\frac{\partial(\rho,\sigma)_1}{\partial x_1}+\frac{\partial(\rho,\sigma)_2}{\partial x_2}+\cdots+\frac{\partial(\rho,\sigma)_r}{\partial x_r} = 0.$$

This is, in fact, a fundamental equation for the functions $(\rho,\ \sigma)_m$. It can be derived directly from the equations that the u satisfy. To that end, we need only to multiply the equations:

$$\frac{\partial\Omega_2^\sigma}{\partial u^{i,\sigma}} = \sum_m \frac{\partial}{\partial x_m}\frac{\partial\Omega_2^\sigma}{\partial\dfrac{\partial u^{i,\sigma}}{\partial x_m}},$$

$$\frac{\partial\Omega_2^\rho}{\partial u^{i,\rho}} = \sum_m \frac{\partial}{\partial x_m}\frac{\partial\Omega_2^\rho}{\partial\dfrac{\partial u^{i,\rho}}{\partial x_m}},$$

by $u^{i,\rho}$ and $u^{i,\sigma}$, respectively, and then add them, and take the sum over i. This gives:

$$\sum_i\left(u^{i,\rho}\frac{\partial\Omega_2^\sigma}{\partial u^{i,\sigma}} - u^{i,\sigma}\frac{\partial\Omega_2^\rho}{\partial u^{i,\rho}}\right)$$

$$= \sum_i\sum_m\left(u^{i,\rho}\frac{\partial}{\partial x_m}\frac{\partial\Omega_2^\sigma}{\partial\dfrac{\partial u^{i,\sigma}}{\partial x_m}} - u^{i,\sigma}\frac{\partial}{\partial x_m}\frac{\partial\Omega_2^\rho}{\partial\dfrac{\partial u^{i,\rho}}{\partial x_m}}\right),$$

and when one now adds the expression:

$$\sum_i\sum_m\left(\frac{\partial u^{i,\rho}}{\partial x_m}\frac{\partial\Omega_2^\sigma}{\partial\dfrac{\partial u^{i,\sigma}}{\partial x_m}} - \frac{\partial u^{i,\sigma}}{\partial x_m}\frac{\partial\Omega_2^\rho}{\partial\dfrac{\partial u^{i,\rho}}{\partial x_m}}\right)$$

to both sides, one obtains identically zero on the left-hand side, from known properties of homogeneous functions of second order, and with consideration to the fact that $\dfrac{\partial\Omega_2}{\partial\mu_k}$ vanishes, while the right-hand side likewise yields the equation to be proved:

$$0 = \frac{\partial(\rho,\sigma)_1}{\partial x_1} + \frac{\partial(\rho,\sigma)_2}{\partial x_2} + \cdots + \frac{\partial(\rho,\sigma)_r}{\partial x_r}.$$

This equation may also be expressed by saying that *the functions:*

$$(\rho, \sigma)_1,\ (\rho, \sigma)_2,\quad \ldots, (\rho, \sigma)_r$$

can always be represented as sub-determinants of one and the same functional determinant, and this gives rise to a remarkable representation of the functions z. *If we imagine that the functions of this determinant have been found, i.e., that the partial differential equation:*

(59) $$0 = (\rho,\sigma)_1 \frac{\partial \varphi^{\rho,\sigma}}{\partial x_1} + (\rho,\sigma)_2 \frac{\partial \varphi^{\rho,\sigma}}{\partial x_2} + \cdots + (\rho,\sigma)_r \frac{\partial \varphi^{\rho,\sigma}}{\partial x_r},$$

whose multiplier is unity, has been integrated completely, and if we denote the solutions to it by:

$$\varphi^{\rho,\sigma},\ \varphi_1^{\rho,\sigma},\ \ldots,\ \varphi_{r-2}^{\rho,\sigma},$$

and the partial functional determinants that are defined from them, which were equal to the functions (ρ, σ), *by* $\Delta_1^{\rho,\sigma}$, $\Delta_2^{\rho,\sigma}$, …, $\Delta_r^{\rho,\sigma}$, *then one obtains solutions to equation* (56) *when one sets:*

(60) $$z_{h,m}^{\rho,\sigma} = \varphi^{\rho,\sigma} \left(\frac{\partial \Delta_m}{\partial \dfrac{\partial \varphi}{\partial x_h}} \right)^{\rho,\sigma} = -\, \varphi^{\rho,\sigma} \left(\frac{\partial \Delta_h}{\partial \dfrac{\partial \varphi}{\partial x_m}} \right)^{\rho,\sigma}.$$

From this, it then follows from known properties of functional determinants that:

$$\frac{\partial z_{1,m}^{\rho,\sigma}}{\partial x_1} + \frac{\partial z_{2,m}^{\rho,\sigma}}{\partial x_2} + \cdots = \sum_h \frac{\partial \varphi^{\rho,\sigma}}{\partial x_h} \left(\frac{\partial \Delta_m}{\partial \dfrac{\partial \varphi}{\partial x_h}} \right)^{\rho,\sigma} = \Delta_m^{\rho,\sigma},$$

or $(\rho, \sigma)_m$, which was to be proved.

However, since the functions z here take the form of the second differential quotients of a functional determinant that is comprised of elements that one may, in general, assume to be known, there is another path that leads to the immediate representation of certain solutions of (56). *Namely, one can set:*

(61) $$z_{h,m}^{\rho,\sigma} = \frac{1}{r} \left\{ \int (\rho,\sigma)_m \, dx_h - \int (\rho,\sigma)_h \, dx_m \right\}.$$

If one then differentiates with respect to x_h and sums then this gives, upon consideration of equations (58):

$$\sum_h \frac{\partial z_{h,m}^{\rho,\sigma}}{\partial x_h} = (\rho, \sigma)_m,$$

as it should be.

Equation (61) thus contains the solution to the present problem. I thus point out the remarkable fact that *in the entire course of the investigation, no condition was imposed upon the character of the solutions* $u^{i,\sigma}$ *of the system* (40). It is known that such conditions enter in when only *one* independent variable is present, and indeed, it is the condition that the constant values that the functions (ρ, σ) can assume must all be equal to zero. Such a condition then appears in no other cases. In fact, the second of equations (35) leads to $n^2 r$ equations in:

$$\frac{n\cdot n+1}{2}r+\frac{n\cdot n-1}{2}\cdot\frac{r\cdot r-1}{2}$$

undetermined functions. The difference between the two numbers is $\frac{n\cdot n-1}{2}\cdot\frac{r\cdot r-3}{2}$. One thus has, when $r = 3$, just as many functions to be determined as equations that are present, and too many when $r > 3$. However, for $r = 1$ and $r = 2$ there are $\frac{n\cdot n-1}{2}$ more equations than functions, and (56) then actually leads to the $\frac{n\cdot n-1}{2}$ condition equations:

$$(\rho,\ \sigma) = 0.$$

For $r = 2$, however, these $\frac{n\cdot n-1}{2}$ conditions are nothing but equations (58), which are fulfilled identically. Indeed, this situation also diminishes the number of equations by $\frac{n\cdot n-1}{2}$ for $r = 3$ such that the problem will be truly undetermined for $r = 3$. Furthermore, this shows that $r = 2$ yields the only case in which the problem is determinate, without leading to conditions on the u, and in this case, one has:

$$z_{1,2}^{\rho,\sigma} = \tfrac{1}{2}\left\{\int(\rho,\sigma)_2\,dx_1 - \int(\rho,\sigma)_1\,dx_2\right\},$$

or also:

$$z_{1,2}^{\rho,\sigma} = \int((\rho,\sigma)_2\,dx_1 - (\rho,\sigma)_1\,dx_2),$$

where a complete differential is found under the integral sign.

§ 11.

The result of the present developments may be summarized in the following theorem:

Let $u^{1,\sigma}$, $u^{2,\sigma}$, …, $u^{n,\sigma}$ *be a system of functions that are composed from the solutions to equations* (5), (6) *in such a way that:*

(42) $$u^{i,\sigma} = \sum \alpha^\sigma \frac{\partial y^{(i)}}{\partial a} + \sum \Pi^\sigma \frac{\partial y^{(i)}}{\partial P},$$

while U denotes their determinant and $U^{i,\sigma}$ *is equal to* $\frac{\partial U}{\partial u^{i,\sigma}}$. *By partial integration, the integral:*

$$\int^{(r)} \Omega_2(w)\,dx_1 dx_2 \cdots dx_r$$

reduces to:

$$\int^{(r)} ((\Omega_2(W) - (\Theta(W)))\, dx_1 dx_2 \cdots dx_r \;,$$

where the arguments $W_k^{(i)}$ *have the form:*

(13.30)
$$W_k^{(i)} = \frac{\partial w^{(i)}}{\partial x_k} - \sum_\sigma \sum_h w^{(h)} \frac{\partial u^{h,\sigma}}{\partial x_k} \cdot \frac{U^{i,\sigma}}{U} \,,$$

or, when they are coupled to each other by the equations:

$$\frac{\partial W_m^{(i)}}{\partial x_k} - \sum_\sigma \sum_h \frac{\partial u^{h,\sigma}}{\partial x_k} \cdot \frac{U^{i,\sigma}}{U} \cdot W_m^{(h)} = \frac{\partial W_k^{(i)}}{\partial x_m} - \sum_\sigma \sum_h \frac{\partial u^{h,\sigma}}{\partial x_m} \cdot \frac{U^{i,\sigma}}{U} \cdot W_k^{(h)} \,,$$

and where the coefficient $a_{k,n}^{h,i}$, *with which* $W_k^{(h)} \cdot W_m^{(i)}$ *is multiplied in* $(\Theta(W))$, *has the form:*

$$\begin{aligned} a_{k,n}^{h,i} &= \sum_\rho \sum_\sigma \frac{U^{h,\sigma} \cdot U^{i,\rho}}{rU^2} \left(\int (\rho, \sigma)_m \, dx_h - \int (\rho, \sigma)_h \, dx_m \right) \\ &= \frac{1}{2rU} \sum_\rho \sum_\sigma \frac{\partial^2 U}{\partial u^{h,\sigma} \partial u^{i,\rho}} \left(\int (\rho, \sigma)_m \, dx_h - \int (\rho, \sigma)_h \, dx_m \right). \end{aligned}$$

Likewise, the equations:

(10) $\psi_1 = 0, \qquad \psi_2 = 0, \ldots, \; \psi_\kappa = 0,$

go to linear condition equations between the $W_k^{(i)}$.

From the above, the form of the latter is easy to give, as well as that of the integrals of lower order that were excluded from the partial integration.

I couple this with a remark that concerns the more general problem in which higher differential quotients also enter under the integral sign. One can then introduce all of the lower differential quotients as new independent variables, and define their character by the condition equations that one appends. However, we remark that in the aforementioned theorem $W_k^{(i)}$ has the form:

$$U \cdot W_k^{(i)} = \begin{vmatrix} \frac{\partial w^{(i)}}{\partial x_k} & \frac{\partial u^{i,1}}{\partial x_k} & \frac{\partial u^{i,2}}{\partial x_k} & \cdots & \frac{\partial u^{i,n}}{\partial x_k} \\ w^{(1)} & u^{1,1} & u^{1,2} & \cdots & u^{1,n} \\ w^{(2)} & u^{2,1} & u^{2,2} & \cdots & u^{2,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w^{(n)} & u^{n,1} & u^{n,2} & \cdots & u^{n,n} \end{vmatrix} .$$

If $w^{(i)}$ does not correspond to one of the highest-order differential quotients that are now introduced as new variables then amongst the sequences:

$$w^{(i)}, u^{h,1}, u^{h,2}, \ldots, u^{h,n},$$

there is always one of them that is equal to the sequence:

$$\frac{\partial w^{(i)}}{\partial x_k}, \frac{\partial u^{i,1}}{\partial x_k}, \frac{\partial u^{i,2}}{\partial x_k}, \ldots, \frac{\partial u^{i,n}}{\partial x_k}$$

precisely. $W_k^{(i)}$ will then always be vanishing, and one thus obtains the following theorem:

After an application of the present transformations under the integral sign, an integral that contains differential quotients of arbitrarily higher order in the dependent variables exhibits a homogeneous function of second order whose arguments correspond to the respective highest-order differential quotients of the dependent variables.

If the differential quotients go up to the degrees p_1, p_2, …, p_n, respectively, then the number of arguments in the reduced function of second order is:

$$\sum_h \frac{r \cdot r+1 \cdot r+2 \cdots r+p_h-1}{1 \cdot 2 \cdot 3 \cdots p_h}.$$

Berlin, 12 June 1858.

On a question in the calculus of variations

By J. HADAMARD

At the present time, one knows that conditions have been rigorously established for the maximum and minimum of a simple integral that contains an arbitrary number of unknown functions, or again a multiple integral that contains only one of these functions ([1]). In the first case, if one limits oneself, for simplicity, to the study of an integral:

$$\int f(x, y_1, y_2, \cdots, y_m, y_1', y_2', \cdots, y_m') dx ,$$

where the functions $y_1, y_2, \ldots, y_m$ are not subject to any given relation and appear to the first order, the Legendre condition for the weak minimum, which is generally equivalent ([2]) to that of Weierstrass, is that the quadratic form:

$$\sum_{i=1}^{m} \sum_{k=1}^{m} \frac{\partial^2 f}{\partial y_i' \partial y_k'} u_i u_k$$

must be positive definite. In the second case, the integral being:

$$\iint \cdots \iint f(y, x_1, x_2, \cdots, x_n, p_1, p_2, \cdots, p_n) dx_1 dx_2 \cdots dx_n$$

and $p_1, p_2, \ldots, p_n$ denoting the partial derivatives of the unknown function y with respect to the independent variables $x_1, x_2, \ldots, x_n$, the condition will be on the form:

$$\sum_{i,k} \frac{\partial^2 f}{\partial p_i \partial p_k} u_i u_k ,$$

which must be definite.

The most general case, in which both several independent variables and several unknown functions are involved, has been, to the contrary, almost universally neglected although one must regard it as a natural and immediate generalization of the preceding cases.

From that analogy, one must, upon being given an integral:

(1) $$\iint \cdots \int f(y_1, y_2, \cdots, y_m, x_1, \cdots, x_n, p_1', \cdots, p_n^m) \, dx_1 dx_2 \cdots dx_n ,$$

([1]) Here, I am making an abstraction of the cases in which some of the inequalities that must be satisfied in order for there to be a maximum or a minimum are replaced by the corresponding inequalities, cases that are yet to be elucidated and will probably never be complete (*see* HEDRICK, *Bulletin of the American Math. Society*, 2nd series, v. IX, 1902). These cases are exceptional, due to the fact that the difficulties that are at issue in the text are presented in a completely general case.

([2]) The exceptions are valid only for the limiting cases to which we alluded in the preceding footnote.

where the quantity under the ∬ …∫ depends on m unknown functions y, n independent variables x, and mn derivatives:

$$p_k^i = \frac{\partial y_i}{\partial x_k} \qquad \begin{pmatrix} i = 1, 2, \cdots, m \\ k = 1, 2, \cdots, n \end{pmatrix}$$

of the former ones with respect to the latter, consider a necessary condition for the minimum to be that the quadratic form:

(2) $$\sum_{i,k,i',k'} \frac{\partial^2 f}{\partial p_k^i \, \partial p_{k'}^{i'}} u_k^i u_{k'}^{i'}$$

in the mn indeterminates $u_1^1, \dots, u_n^m$ must be positive definite.

It is not pointless to remark that it is nothing of the sort, and that the case of $m > 1$, $n > 1$ presents a difficulty that is peculiar to it. This remark is not, moreover, completely new: It results from the transformations that were performed by Clebsch in his memoir *Über die zweite Variation vielfacher Integral* ([1]). These transformations, which generalize the ones that have been performed by Jacobi on the second variation of simple integrals, consist uniquely in the addition of terms that integration by parts allows to disappear when one transports them to the frontier. Clebsch showed that by a convenient introduction of terms of this type one may supplement the form (2) with particular linear combinations of the $\frac{m(m-1)}{2}\frac{n(n-1)}{2}$ expressions:

(3) $$u_k^i u_{k'}^{i'} - u_{k'}^i u_k^{i'} \qquad \begin{pmatrix} i, i' = 1, 2, \cdots, m \\ k, k' = 1, 2, \cdots, n \\ i \neq i', \quad k \neq k' \end{pmatrix}.$$

It is therefore absolutely necessary only that the form (2) be definite: *It suffices that it become that way when one combines it in an arbitrary manner with the forms* (3).

Furthermore, we see that this condition is sufficient to render the second variation essentially positive. However, it is easy to deduce from it the desired sufficient conditions for a minimum.

Indeed, suppose that the form (2) is defined by the addition of one or more terms of the form:

$$\lambda(u_k^i u_{k'}^{i'} - u_{k'}^i u_k^{i'}),$$

where λ is a quantity, either constant or variable, that one may always assume to be expressed as a function of x. We may, with Clebsch, consider these terms to be pre-existing in the form in question, with the condition that one adds to f corresponding terms:

([1]) *Journal de Crelle*, t. 56, 1859, pp. 122-149.

$$\frac{\partial}{\partial x_k}(\lambda y_i p_{k'}^{i'}) - \frac{\partial}{\partial x_{k'}}(\lambda y_i p_k^{i'})$$
$$= \lambda(p_k^i p_{k'}^{i'} - p_{k'}^i p_k^{i'}) + y_i\left(p_{k'}^{i'}\frac{\partial \lambda}{\partial x_k} - p_k^{i'}\frac{\partial \lambda}{\partial x_{k'}}\right),$$

which naturally do not change the question that was posed, and consequently also remain without influence on the other elements of the solution – I would like to say, on the Lagrange equation and the Jacobi condition. Once the expression for f has been thus transformed, everything proceeds as for $m = 1$.

It remains to know whether the condition, thus modified, is necessary.

Now, one may easily establish ([1]) the following necessary condition: *The form* (2) *must be essentially positive for all the values of u that annul the expressions* (3).

However, it is not obvious that this condition is equivalent to the former one. Indeed, if one considers p arbitrary quadratic forms $\varphi_1, \varphi_2, \ldots, \varphi_p$, and a $(p + 1)^{\text{th}}$ form ψ, and if one knows that it is positive for all values of the variables that annul the former then it results only from this that ψ must be a linear combination of the forms φ and a definite form. At the very most, one may be more affirmative on this point if, instead of arbitrary forms φ, one considers the particular forms (3).

Meanwhile, the deduction in question is legitimate when the number p is equal to unity. Since this is true for $m = 2$, $n = 2$, the question is solved in this case.

On the contrary, it seems to call for further research when $m, n > 2$.

([1]) This proof, which rests, moreover, on considerations that are completely distinct from the integration by parts that must be utilized for the converse proof, appear in my *Leçons sur la propagation des ondes et les equations de l'Hydrodynamique*, which will be published shortly.

On some questions in the calculus of variations

By J. HADAMARD

In a previous Communication ([1]), I gave a necessary condition (corresponding to the Legendre condition or that of Weierstrass) for the minimum of an n-fold integral in which m unknown functions appear. I added that the method of Clebsch ([2]) furnishes a condition that is equivalent to the preceding one for $m = n = 2$ (the equivalence being doubtful for higher values of m and n), and which is capable of playing the role of the Legendre-Weierstrass condition as a sufficient condition.

This latter point is not exact, and, as one sees, the question is yet to be elucidated, except for the simplest case of $m = n = 2$.

Let z_1, z_2 be unknown functions of the variables x, y and let p_1, q_1, p_2, q_2 be their partial derivatives. Any condition that is analogous to that of Legendre or Weierstrass for the weak minimum (for us upon considering this case) must involve a certain form F that is quadratic in p_1, q_1, p_2, q_2 .

I. The *necessary* condition that we have previously obtained is that F must be essentially positive for all the (non-null) values of the variables p_1, q_1, p_2, q_2 that satisfy the relation:

$$p_1 q_2 - q_1 p_2 = 0. \tag{1}$$

II. The method *Clebsch* gives the *sufficient* condition that the quadratic form:

$$\Phi + \lambda(p_1 q_2 - q_1 p_2) \tag{2}$$

(where λ is a function of x and y, but not the p, q) must be positive definite. To this condition one must, of course, add a *Jacobi condition*.

In the present case, this consists in the existence of two solutions (ζ_1, ζ_2), (τ_1, τ_2) of the *equations of variation*, such that the determinant:

$$\Delta = \zeta_1 \tau_2 - \zeta_2 \tau_1,$$

is not annulled in the domain of integration.

If λ may be chosen arbitrarily then the condition that $\Phi + \lambda(p_1 q_2 - q_1 p_2)$ may be rendered definite by a choice of λ will be equivalent to the previously stated necessary condition.

However, this is not the case. Just as it results from the analysis of Clebsch, the values of λ are determined by those of the solutions ζ, τ, or at least, once these solutions are chosen, λ contains only an arbitrary constant C.

It therefore does not suffice that for each system of values of x and y there exist values of λ that render the form (2) positive definite. Let:

([1]) This *Bulletin*, t. XXX, 1902, pp. 253.
([2]) *Journal de Crelle*, t. 56, 1859.

$$\lambda_1 \leq \lambda \leq \lambda_2 \tag{3}$$

be the interval that comprises these values of λ. It is further necessary that one must determine the constant C in such a manner that for all pairs of values of x, y contained in the domain of integration the inequality (3) must be verified; in other words, that the minimum of the values of C deduced (at each point) from the relation $\lambda = \lambda_2$ are not inferior to the maximum of the value of C that is deduced from $\lambda = \lambda_1$.

Likewise, if one takes, in all possible ways, the solutions ζ, τ of the equation of variations then the function λ may not be taken at will. It will satisfy a system of partial differential equations S (most likely complicated) resulting from the elimination of ζ_1, ζ_2; τ_1, τ_2 between the equations of variation (which are four in number for the two systems) and the two relations that define λ, the six equations thus written reducing, moreover, to five, thanks to the fact that the system of equations of variation is identical to its adjoint.

One will then be led to the following question:

Does there exist a solution to the system S satisfying the inequalities (3) *in the entire domain of integration?*

This problem belonging to the same category of questions as the ones that we encountered a moment ago, it offers a simple example, a category that is probably quite worthy of attention.

Of course, once one has such a solution λ, one must calculate the corresponding solutions ζ, τ and verify the Jacobi condition $\Delta \gtrless 0$.

III. The method of *Hilbert* (which is not fundamentally distinct from that of Clebsch) leads to results that are entirely similar.

In order for a function of x, z, z_1, z_2, p_1, p_2, q_1, q_2, when integrated over x and y, to give a result that depends upon only the contour, it is necessary and sufficient that it have the form:

$$\varphi = A + B_1 p_1 + B_2 p_2 + C_1 q_1 + C_2 q_2 + D(p_1 q_2 - q_1 p_2)$$

(A, B, … functions of x, y, z_1, z_2), a form that generally refers, as one sees, to a term that is nonlinear with respect to the first derivatives.

In order to follow the path pointed out by Hilbert, we first take:

$$\varphi = f(x, y, z_1, \varpi_1, \chi_1, \varpi_2, \chi_2) + (p_1 - \varpi_1) f_{\varpi_1} + (p_2 - \varpi_2) f_{\varpi_2} + (q_1 - \chi_1) f_{\chi_1} + (q_2 - \chi_2) f_{\chi_2}, \tag{4}$$

where f is the given function under the $\int\int$ sign, in such a manner that:

$$\iint (x, y, z_1, z_2, p_1, p_2, q_1, q_2)\, dx\, dy$$

is the integral whose extremum one seeks and ϖ_1, ϖ_2, χ_1, χ_2 are defined in the following manner:

One considers a family of extremals that depend upon two arbitrary constants a, b:

(5) $$z_1 = \Psi_1(x, y, a, b), \qquad z_2 = \Psi_2(x, y, a, b),$$

these equations being supposed soluble with respect to a, b, in such a way that the determinant:

(6) $$\frac{\mathrm{D}(\Psi_1, \Psi_2)}{\mathrm{D}(a, b)}$$

is not annulled at any point of the domain of integration.

Geometrically speaking, the functions z_1, z_2 of x and y represent a doubly-extended multiplicity traced out in the space of four dimensions, or, more conveniently, a pair of surfaces in ordinary space, with the convention that one considers any two points to be an indissoluble entity when, taken on these two surfaces respectively, they have the same projection on the xy-plane ([1]). These are two points thus situated on the same parallel to the z-axis that we call a *pair* of points. The condition that was imposed on the extremals (5) amounts to saying that one may make them pass through any pair of given points (that are sufficiently close to the extremal that one considers).

Having said this, ϖ_1, χ_1, ϖ_2, χ_2 are the partial derivatives of z_1, z_2 that are deduced from equations (5). They are thus functions of x, y, a, b, and consequently (by virtue of the previously-postulated solubility) functions of x, y, z_1, z_2, in such a way that expression (4) is a function of x, y, z_1, z_2, p_1, q_1, p_2, q_2.

The calculations involve two types of derivatives, namely:

1. The derivatives of a function of x, y, z_1, z_2, p_1, q_1, p_2, q_2, when these eight quantities are considered to be independent variables; we denote them by indices. For example, Φ_{p_1} denotes the derivatives of Φ with respect to p_1. This is what we did in formula (4), where f_{ϖ_1} represents the value of f_{p_1} when one replaces p_1, q_1, p_2, q_2 by ϖ_1, ϖ_2, χ_1, χ_2 in it.

2. The derivatives of a function of x, y, z_1, z_2 with respect to these quantities considered as independent variables; we denote them by the symbol δ. For a function Φ of x, y, z_1, z_2, ϖ_1, χ_1, ϖ_2, χ_2 one has:

(7) $$\frac{\delta\Phi}{\delta z_i} = \Phi_{z_i} + \Phi_{\varpi_1}\frac{\delta\varpi_1}{\delta z_i} + \Phi_{\varpi_2}\frac{\delta\varpi_2}{\delta z_i} + \Phi_{\chi_1}\frac{\delta\chi_1}{\delta z_i} + \Phi_{\chi_1}\frac{\delta\chi_2}{\delta z_i} \qquad (i = 1, 2).$$

3. The derivatives are taken along the pair of surfaces (which are arbitrary, moreover) whose tangent planes have the angular coefficients p_1, q_1, p_2, q_2 . We denote them by the symbol ∂; one has (for a function of x, y, z_1, z_2):

([1]) In particular, these surfaces are limited by two given contours C_1, C_2 that are both subsets of the same cylinder parallel to the z-axis.

$$(8)\qquad \begin{cases} \dfrac{\partial}{\partial x} = \dfrac{\delta}{\delta x} + p_1 \dfrac{\delta}{\delta z_1} + p_2 \dfrac{\delta}{\delta z_2}, \\ \dfrac{\partial}{\partial y} = \dfrac{\delta}{\delta y} + q_1 \dfrac{\delta}{\delta z_1} + q_2 \dfrac{\delta}{\delta z_2}. \end{cases}$$

4. The derivatives are taken along the extremals (4) upon considering, as a consequence, z_1, z_2 to be functions of x, y and a, b to be constants. These derivatives, which will be denoted by the symbol $\mathfrak{d}$, are coupled to the derivatives δ by the relation:

$$(9)\qquad \begin{cases} \dfrac{\mathfrak{d}}{\mathfrak{d} x} = \dfrac{\delta}{\delta x} + \varpi_1 \dfrac{\delta}{\delta z_1} + \varpi_2 \dfrac{\delta}{\delta z_2}, \\ \dfrac{\mathfrak{d}}{\mathfrak{d} y} = \dfrac{\delta}{\delta y} + \chi_1 \dfrac{\delta}{\delta z_1} + \chi_2 \dfrac{\delta}{\delta z_2}. \end{cases}$$

Having said this, the conditions for the integral $\int\int \varphi \;dx\,dy$, φ being the expression (4), to depend upon only the contour, namely:

$$\frac{\delta\varphi}{\delta z_1} - \frac{\partial}{\partial x}\varphi_{p_1} - \frac{\partial}{\partial y}\varphi_{q_1} = 0, \qquad \frac{\delta\varphi}{\delta z_2} - \frac{\partial}{\partial x}\varphi_{p_2} - \frac{\partial}{\partial y}\varphi_{q_2} = 0$$

(equalities that must be valid for any p, q), reduce, upon taking into account formulas (8) and the fact that the surfaces (5) are extremals, to:

$$(10)\qquad \begin{cases} \dfrac{\delta}{\delta z_2} f_{\varpi_1} - \dfrac{\delta}{\delta z_1} f_{\varpi_2} = 0, \\ \dfrac{\delta}{\delta z_2} f_{\chi_1} - \dfrac{\delta}{\delta z_1} f_{\chi_2} = 0. \end{cases}$$

They will therefore not be satisfied in general.

However, if (as we have the right to do) we add to f the two integrable terms:

$$\begin{aligned} \frac{\partial}{\partial y}(\mu\, p_1 z_2) - \frac{\partial}{\partial x}(\mu\, q_1 z_2) &= \frac{\partial\mu}{\partial y} p_1 z_2 - \frac{\partial\mu}{\partial x}\, q_1 z_2 + \mu(p_1 q_2 - q_1 p_2) \\ &= \frac{\delta\mu}{\delta y} p_1 z_2 - \frac{\delta\mu}{\delta x}\, q_1 z_2 + \mu(p_1 q_2 - q_1 p_2) \end{aligned}$$

(m being a function of x, y, z_1, z_2, and λ denoting the combination:

$$\lambda = \mu + z_2 \frac{\delta\mu}{\delta z_2}\Big),$$

or furthermore, if, without changing f, we add to the expression (4), the term:

$$-\lambda[(p_1-\varpi_1)(q_2-\chi_2)-(p_2-\varpi_2)(q_1-\chi_1)],$$

then equations (10) are replaced by:

$$(10')\qquad \begin{cases} \dfrac{\delta}{\delta z_2} f_{\varpi_1} - \dfrac{\delta}{\delta z_1} f_{\varpi_2} + \lambda\left(\dfrac{\delta\chi_1}{\delta z_1}+\dfrac{\delta\chi_2}{\delta z_2}\right)+\dfrac{\partial\lambda}{\partial y}=0, \\ \dfrac{\delta}{\delta z_2} f_{\chi_1} - \dfrac{\delta}{\delta z_1} f_{\chi_2} + \lambda\left(\dfrac{\delta\varpi_1}{\delta z_1}+\dfrac{\delta\varpi_2}{\delta z_2}\right)+\dfrac{\partial\lambda}{\partial x}=0, \end{cases}$$

and, to satisfy the conditions of the problem, we only have to determine λ by these latter equations.

However, these are nothing but the equations that were posed by Clebsch. In order to convert them to the Clebsch form, it suffices to transform the derivatives δ upon replacing the independent variables z_1, z_2 by the variables a, b. The derivatives of one of them with respect to the other are the quantities that were previously denoted by ζ_1, ζ_2, τ_1, τ_2, and the functional determinant (6) is what we have called Δ.

Consequently, as in the Clebsch method, the equations (10′) form a completely integrable system and have, as a consequence, a solution λ that depends upon an arbitrary constant.

We are thus led to exactly the same point as in the preceding method, and we have to study a system that is analogous to S (but notably more complicated and more difficult to form, since the equations will no longer be linear) obtained by eliminating z_1, z_2 between equations (10′) and the ones that express that the family (5) is composed of extremals. We must express the notion that this system admits a solution satisfying the inequalities (3).

IV. In reality, one must presume that the discussion that we just had is not necessary.

Indeed, if we consider, no longer the method of Hilbert, but the method of *Weierstrass* in its original form then we arrive at some conclusions in a notably different form.

Indeed, let (s_1, s_2) be the pair of surfaces that constitute the extremal studied, while (S_1, S_2) is another pair that one may compare them with, and which is limited by the same contour (C_1, C_2). On (S_1, S_2), trace out a pair of variable contours (γ_1, γ_2) (γ_1 and γ_2 being, of course, situated on the same cylinder parallel to Ox). Make a pair of surfaces (Σ_1, Σ_2) that constitute an *extremal* pass through (γ_1, γ_2). If this latter construction is always possible then it will suffice to vary the contours (γ_1, γ_2) from a pair of points to the position (C_1, C_2) in order to apply the reasoning of Weierstrass.

Now, the condition that is analogous to that of Legendre to which one thus arrives (for a weak minimum) is the *necessary* condition that was recalled to begin with.

However, upon proceeding thus the difficulty appears in the Jacobi condition. Instead of simply supposing the existence of a two-parameter family of extremals one must express that one may construct the extremal pair (Σ_1, Σ_2), i.e., solve a Dirichlet problem in two unknowns under the most general and difficult conditions, since γ_1 and γ_2 are arbitrary.

The present state of science does not, for example, permit one to deduce from this method the existence of a minimum for a very small integration domain. On the contrary, it results from the method of Clebsch-Hilbert since, by reason of the arbitrary constants that appear in l, one may always suppose that this quantity satisfies the inequalities (3) in the environment of an arbitrary given point.

I will conclude by indicating some points for which, in the lectures taught at the Collège de France for two years, I have yet to complete the results acquired from the calculus of variations.

The first concerns the *isoperimetric problem*, in which one seeks the extremum of a certain integral I_0, knowing the values of one or more other given integrals and certain accessory conditions. The fundamental result (which reduces that extremum to a free minimum by the intervention of a multiplier) has yet to be extended to the case where, among the accessory conditions the conditions of inequality appear. It still remains true in this case; however, the proof must appeal to some considerations that are noticeably different from the classical considerations.

On the other hand, the question of knowing whether the *Weierstrass construction* is possible in this same isoperimetric problem, may, in many cases, be considered to be solved if the arc considered satisfies the sufficient conditions for a minimum (for example) for the integral $I_0 + lI_1$ (I_1 being the given integral and l, the multiplier), this minimum being considered to be free. Indeed, one confirms that the extremities remain fixed. This remark will be very useful, for example, in the proof of the existence of a minimum in a sufficiently small region.

Finally, I will further note a very great simplification that one may achieve in the proof of the theorem of Osgrod, from which one may assign a lower limit on the difference that exists between the minimum integral and a varied integral. A procedure that is completely similar to the one that was employed by Kneser in the context of stability of equilibrium for a massive string permits one to pass to the inequalities that were employed in the different direct proofs that given up to now. Unfortunately, this method is no longer applicable to multiple integrals, for which, what is more, the question is always much more delicate.

I will content myself with merely mentioning these various points, which will be treated in more detail in a later work.

On the calculus of variations [1]

By David HILBERT

Necessity of the existence of the Lagrange differential equations

The question of the necessity of the Lagrange criterion – i.e., the existence of the differential equations that are implied by the vanishing of first variation – has been treated by A. MAYER [2] and A. KNESER [3]. Here, I would like to present a rigorous and likewise very simple method that leads to the desired proof of the Lagrangian criterion.

For the sake of brevity, I will always assume in the following communication that the given functions and differential relations are analytic, from which the analytical character of the solutions that come to be employed will be assured.

We further choose the more convenient representation – without diminishing the generality of the method – in the case of three desired functions $y(x)$, $z(x)$, $s(x)$ of the independent variable x; between them and their first derivatives with respect to x:

$$\frac{dy}{dx} = y'(x), \qquad \frac{dz}{dx} = z'(x), \qquad \frac{ds}{dx} = s'(x)$$

let there be given two conditions of the form:

$$\left.\begin{aligned} f(y', z', s', y, z, s; x) = 0, \\ g(y', z', s', y, z, s; x) = 0. \end{aligned}\right\} \tag{1}$$

From this, one arrives at the following theorem to be proved:

It might be that $y(x)$, $z(x)$, $s(x)$ are three particular functions that satisfy the conditions (1) and possess the following properties: for all values of x that lie between $x = a_1$ and $x = a_2$, one has:

$$\begin{vmatrix} \dfrac{\partial f}{\partial y'} & \dfrac{\partial f}{\partial z'} \\ \dfrac{\partial g}{\partial y'} & \dfrac{\partial g}{\partial z'} \end{vmatrix} \neq 0; \tag{2}$$

if we choose any other three functions $Y(x)$, $Z(x)$, $S(x)$ that likewise satisfy the conditions (1), for which one has:

---

[1] Essentially unchanged from the version that was published in the Göttinger Nachrichten 1905, pp. 159-180.

[2] Math. Ann. v. 26 and Leipziger Berichte 1895; in the latter note, A. MAYER has extended his foundation of the Lagrange differential equations to the most general problem.

[3] Lehrbuch der Variationsrechnung § 56-58, Braunschweig 1900; the problem was likewise posed in full generality in this work.

$$
\begin{array}{ll}
Y(a_1) = y(a_1), & \\
Z(a_1) = z(a_1), & Z(a_2) = z(a_2), \\
S(a_1) = s(a_1), & S(a_2) = s(a_2),
\end{array}
$$

then one always lets – assuming that the functions $Y(x)$, $Z(x)$, $S(x)$, along with their derivatives (any particular functions $y(x)$, $z(x)$, $s(x)$ and their derivatives, resp.) differ sufficiently little:

$$Y(a_2) \geq y(a_2). \tag{3}$$

If this minimality requirement is fulfilled then there are necessarily two functions $\lambda(x)$, $\mu(x)$ that do not both vanish identically for all x, and which together with the functions $y(x)$, $z(x)$, $s(x)$ fulfill the Lagrange differential equations that arise from the annulling of the first variation of the integral:

$$\int_{a_1}^{a_2} \{\lambda f(y', z', s', y, z, s; x) + \mu g(y', z', s', y, z, s; x)\} dx,$$

namely:

$$\frac{d}{dx}\frac{\partial(\lambda f + \mu g)}{\partial y'} - \frac{\partial(\lambda f + \mu g)}{\partial y} = 0, \tag{4}$$

$$\frac{d}{dx}\frac{\partial(\lambda f + \mu g)}{\partial z'} - \frac{\partial(\lambda f + \mu g)}{\partial z} = 0, \tag{5}$$

$$\frac{d}{dx}\frac{\partial(\lambda f + \mu g)}{\partial s'} - \frac{\partial(\lambda f + \mu g)}{\partial s} = 0. \tag{6}$$

In order to carry out the proof of this theorem, we take any two well-defined functions $\sigma_1(x)$, $\sigma_2(x)$ that vanish for $x = a_1$ and $x = a_2$, and substitute for y, z, s in (1):

$$
\begin{aligned}
Y &= Y(x, \varepsilon_1, \varepsilon_2), \\
Z &= Z(x, \varepsilon_1, \varepsilon_2), \\
S &= s(x) + \varepsilon_1 \sigma_1(x) + \varepsilon_2\, \sigma_2(x),
\end{aligned}
$$

resp., where ε_1, ε_2 mean two parameters. We regard the resulting equations:

$$\left.\begin{aligned} f(Y', Z', S', Y, Z, S; x) = 0, \\ g(Y', Z', S', Y, Z, S; x) = 0 \end{aligned}\right\} \tag{7}$$

as a system of two differential equations for the determination of the two functions Y, Z. As the theory of differential equations teaches [1], due to assumption (2), for sufficiently small values of ε_1, ε_2 there is certainly a system of two functions:

$$Y(x, \varepsilon_1, \varepsilon_2) \qquad \text{and} \qquad Z(x, \varepsilon_1, \varepsilon_2),$$

[1] Cf., É. Picard: Traité d'Analyse, t. III, ch. VIII.

that fulfill both equations identically in x, ε_1, ε_2 and go to $y(x)$, $z(x)$ for $\varepsilon_1 = 0$, $\varepsilon_2 = 0$, and assume the values $y(a_1)$ and $z(a_2)$ for $x = a_1$ and arbitrary ε_1, ε_2, moreover.

Since our minimality requirement (3) demands that $Y(x, \varepsilon_1, \varepsilon_2)$ must have a minimum as a function of ε_1, ε_2 at $\varepsilon_1 = 0$, $\varepsilon_2 = 0$, while the equation:

$$Z(x, \varepsilon_1, \varepsilon_2) = z(a_2)$$

exists between ε_1, ε_2, the theory of relative minima for functions of two variables teaches that there must necessarily be two constants l, m that are not both null, and for which one has:

$$\left.\begin{aligned} \left[\frac{\partial(l\,Y(a_2,\varepsilon_1,\varepsilon_2)+mZ(a_2,\varepsilon_1,\varepsilon_2))}{\partial\varepsilon_1}\right]_0 &= 0, \\ \left[\frac{\partial(l\,Y(a_2,\varepsilon_1,\varepsilon_2)+mZ(a_2,\varepsilon_1,\varepsilon_2))}{\partial\varepsilon_2}\right]_0 &= 0, \end{aligned}\right\} \tag{8}$$

in which the index 0 means, in both cases, that both parameters ε_1, ε_2 are set to zero.

Moreover, we determine – as is certainly possible according to (2) – two functions $\lambda(x)$, $\mu(x)$ of the variable x that satisfy the differential equations (4), (5), which are both linear and homogeneous for them, and for which at the place where $x = a_2$ the boundary conditions:

$$\left.\begin{aligned} \left[\frac{\partial(\lambda l+\mu g)}{\partial y'}\right]_{x=a_2} &= l, \\ \left[\frac{\partial(\lambda l+\mu g)}{\partial z'}\right]_{x=a_2} &= m \end{aligned}\right\} \tag{9}$$

are valid. Since l, m are not both null, both of the functions $\lambda(x)$, $\mu(x)$ thus determined also vanish, but certainly not identically.

By differentiation of the equations (7) with respect to ε_1, ε_2, and subsequent annulling of these two parameters, we obtain the equations:

$$\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0\frac{\partial f}{\partial y'}+\left[\frac{\partial Y}{\partial\varepsilon_1}\right]_0\frac{\partial f}{\partial y}+\left[\frac{\partial Z'}{\partial\varepsilon_1}\right]_0\frac{\partial f}{\partial z'}+\left[\frac{\partial Z}{\partial\varepsilon_1}\right]_0\frac{\partial f}{\partial z}+\sigma_1'\frac{\partial f}{\partial s'}+\sigma_1\frac{\partial f}{\partial s}=0,$$

$$\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0\frac{\partial g}{\partial y'}+\left[\frac{\partial Y}{\partial\varepsilon_1}\right]_0\frac{\partial g}{\partial y}+\left[\frac{\partial Z'}{\partial\varepsilon_1}\right]_0\frac{\partial g}{\partial z'}+\left[\frac{\partial Z}{\partial\varepsilon_1}\right]_0\frac{\partial g}{\partial z}+\sigma_1'\frac{\partial g}{\partial s'}+\sigma_1\frac{\partial g}{\partial s}=0,$$

$$\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0\frac{\partial f}{\partial y'}+\left[\frac{\partial Y}{\partial\varepsilon_2}\right]_0\frac{\partial f}{\partial y}+\left[\frac{\partial Z'}{\partial\varepsilon_2}\right]_0\frac{\partial f}{\partial z'}+\left[\frac{\partial Z}{\partial\varepsilon_2}\right]_0\frac{\partial f}{\partial z}+\sigma_2'\frac{\partial f}{\partial s'}+\sigma_2\frac{\partial f}{\partial s}=0,$$

$$\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0\frac{\partial g}{\partial y'}+\left[\frac{\partial Y}{\partial\varepsilon_2}\right]_0\frac{\partial g}{\partial y}+\left[\frac{\partial Z'}{\partial\varepsilon_2}\right]_0\frac{\partial g}{\partial z'}+\left[\frac{\partial Z}{\partial\varepsilon_2}\right]_0\frac{\partial g}{\partial z}+\sigma_2'\frac{\partial g}{\partial s'}+\sigma_2\frac{\partial g}{\partial s}=0,$$

in which the index 0 always means that both parameters ε_1, ε_2 are set to zero. Starting from these equations, on the one hand, one multiplies the first and second ones by λ and μ, resp., adds the resulting equations, and integrates between the limits $x = a_1$ and $x = a_2$; on the other hand, one multiplies the third and fourth ones by λ and μ, resp., adds the resulting equations, and integrates between the limits $x = a_1$ and $x = a_2$. From this, we obtain:

$$\left.\begin{aligned}
&\int_{a_1}^{a_2}\left\{\frac{\partial(\lambda f+\mu g)}{\partial y'}\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial y}\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial z'}\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0\right.\\
&\qquad\left.+\frac{\partial(\lambda f+\mu g)}{\partial z}\left[\frac{\partial Y'}{\partial\varepsilon_1}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial s'}\sigma_1'+\frac{\partial(\lambda f+\mu g)}{\partial s}\sigma_1\right\}dx \quad = 0,\\
&\int_{a_1}^{a_2}\left\{\frac{\partial(\lambda f+\mu g)}{\partial y'}\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial y}\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial z'}\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0\right.\\
&\qquad\left.+\frac{\partial(\lambda f+\mu g)}{\partial z}\left[\frac{\partial Y'}{\partial\varepsilon_2}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial s'}\sigma_2'+\frac{\partial(\lambda f+\mu g)}{\partial s}\sigma_2\right\}dx \quad = 0.
\end{aligned}\right\} \tag{10}$$

We now have, on the one hand, the conditions defined above:

$$Y(a_1, \varepsilon_1, \varepsilon_2) = y(a_1), \qquad Z(a_1, \varepsilon_1, \varepsilon_2) = z(a_1),$$

and therefore at the location $x = a_1$:

$$\left[\frac{\partial Y}{\partial\varepsilon_1}\right]_0 = 0, \qquad \left[\frac{\partial Z}{\partial\varepsilon_1}\right]_0 = 0,$$

$$\left[\frac{\partial Y}{\partial\varepsilon_2}\right]_0 = 0, \qquad \left[\frac{\partial Z}{\partial\varepsilon_2}\right]_0 = 0;$$

on the other hand, we deduce from equations (8) and (9) at the location $x = a_2$:

$$\frac{\partial(\lambda f+\mu g)}{\partial y'}\left[\frac{\partial Y}{\partial\varepsilon_1}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial z'}\left[\frac{\partial Z}{\partial\varepsilon_1}\right]_0 = 0,$$

$$\frac{\partial(\lambda f+\mu g)}{\partial y'}\left[\frac{\partial Y}{\partial\varepsilon_2}\right]_0+\frac{\partial(\lambda f+\mu g)}{\partial z'}\left[\frac{\partial Z}{\partial\varepsilon_2}\right]_0 = 0.$$

Keeping this in mind, it follows from (10), using (4), (5), and by means of the formula for the integration of a product (partial integration) that we have the equations:

$$\int_{a_1}^{a_2}\left\{\frac{\partial(\lambda f+\mu g)}{\partial y'}\sigma_1'+\frac{\partial(\lambda f+\mu g)}{\partial z'}\sigma_1\right\}dx = 0,$$
$$\int_{a_1}^{a_2}\left\{\frac{\partial(\lambda f+\mu g)}{\partial s'}\sigma_2'+\frac{\partial(\lambda f+\mu g)}{\partial s}\sigma_2\right\}dx = 0.$$

We set, to abbreviate:

$$(\lambda\,\mu,\,\sigma) = \int_{a_1}^{a_2}\left\{\frac{\partial(\lambda f+\mu g)}{\partial s'}\sigma'+\frac{\partial(\lambda f+\mu g)}{\partial s}\sigma\right\}dx,$$

so we can express the result obtained as follows: *For any two functions* σ_1, σ_2 *that vanish at* $x = a_1$ *and* $x = a_2$ *there is always one system of solutions* λ, μ *of the differential equations* (4), (5) *that does not vanish identically and is such, that:*

$$(\lambda\,\mu,\,\sigma_1) = 0 \qquad \textit{and} \qquad (\lambda\,\mu,\,\sigma_2) = 0.$$

If we now assume that there is a function σ_3 for this system of solutions *l*, *m* such that the inequality:

$$(\lambda\,\mu,\,\sigma_3) \neq 0 \tag{11}$$

is valid then we define any system of solutions λ', μ' of the differential equations (4), (5) that does not vanish identically, such that one has:

$$(\lambda'\,\mu',\,\sigma_3) = 0. \tag{12}$$

If we assume, moreover, that there is a function σ_4 for which the inequality:

$$(\lambda'\,\mu',\,\sigma_4) \neq 0 \tag{13}$$

is valid then we can apply our previous result to the functions σ_3, σ_4 and recognize from this the existence of a system of solutions λ'', μ'' of (4), (5), such that the equations:

$$(\lambda''\,\mu'',\,\sigma_3) = 0, \tag{14}$$
$$(\lambda''\,\mu'',\,\sigma_4) = 0, \tag{15}$$

are valid. Since λ, μ; λ', μ'; λ'', μ'' are solutions of a system of two homogeneous, linear, first-order differential equations, there must exist homogeneous linear relations between two of them that take the form:

$$a\,\lambda + a'\,\lambda' + a''\,\lambda'' = 0,$$
$$a\,\mu + a'\,\mu' + a''\,\mu'' = 0,$$

where a, a', a'' mean constants that are not all zero. However, from (11), (12), (14), it would then follow necessarily that $a = 0$ and then, from (13), (15), it would follow that a'

= 0, which is not possible, since one indeed has $a'' \neq 0$, moreover, and the system of solutions λ'', μ'' does not vanish identically in x.

Our assumptions are therefore inapplicable, and we conclude from this that either λ, μ or λ', μ' is a system of solutions of (4), (5) such that the aforementioned integral relation:

$$(\lambda\,\mu,\,\sigma) = 0 \qquad\qquad ((\lambda'\,\mu',\,\sigma) = 0, \text{ resp.})$$

is valid for any function σ. The application of the product integration (partial integration) to this relation then shows that, equation (6) must necessarily be valid for the system of solutions λ, μ (λ' μ', resp.), and with that, our desired proof is brought to completion.

The independence theorem and the Jacobi-Hamilton theory of the associated integration problem

In my talk [1] "Mathematical Problems," I put forth the following method for extending the necessary and sufficient conditions in the calculus of variation:

It treated the simplest problem in the calculus of variations, namely, the problem of finding a function y of the variables x such that the integral:

$$J = \int_a^b F(y', y; x)\,dx\,, \qquad\qquad \left[y' = \frac{dy}{dx}\right]$$

attains a minimum value in comparison to the values that the integral assumes when we replace $y(x)$ with other functions of x that have the same given initial and final values.

We now consider the integral:

$$J^* = \int_a^b \{F + (y' - p)F_p\}\,dx$$
$$\left[F = F(p, y; x), \qquad F_p = \frac{\partial F(p, y; x)}{\partial p}\right],$$

and we ask how the p in it is to be regarded as a function of x, y in order that the value of this integral J^* be independent of the path of integration in the xy-plane – i.e., of the choice of function y of the variable x. The answer is: One takes any one-parameter family of integral curves of the Lagrange differential equation:

$$\frac{d\,\dfrac{\partial F}{\partial y'}}{dx} - \frac{\partial F}{\partial y} = 0, \qquad\qquad [F = F(y', y; x)]$$

[1] Presented to the International Congress of Mathematicians in Paris 1900; this volume, paper no. 17.

and determines at each point x, y, the value of the derivative y' of the curve of the family that goes through this point. The value of this derivative y' is a function $p(x, y)$ with the desired property.

From this "independence theorem," what immediately follows is not only the known criterion for the attainment of the minimum, but also all essential facts from the Jacobi-Hamilton theory of the associated integration problem.

A. MAYER [1] has proved the corresponding theorem for the case of several functions and exhibited its connection with the Jacobi-Hamilton theory. In the following, I would like to show that the independence theorem is capable of a general conception, and also without any application of calculation, and can be proved very simply by returning to the aforementioned special case that was treated in my talk.

For the sake of simplicity, I start with only two functions $y(x)$, $z(x)$. The variational problem consists in choosing them in such a way that the integral:

$$J = \int_a^b F(y', z', y, z; x)\, dx\,, \qquad \left[y' = \frac{dy}{dx}, \quad z' = \frac{dz}{dx} \right]$$

attains a minimum value compared to the values that the integral assumes when one replaces the $y(x)$, $z(x)$ with other functions of x that have the same given initial and final values.

We now consider the integral:

$$J^* = \int_a^b \{F + (y' - p)F_p + (z' - q)F_q\}\, dx$$

$$\left[F = F(p, q, y, z; x), \quad F_p = \frac{\partial F(p, q, y, z; x)}{\partial p}, \quad F_q = \frac{\partial F(p, q, y, z; x)}{\partial q} \right],$$

and ask how the p, q, in it can be regarded as functions of x, y, z in order that the value of this integral J^* *be independent of the path of integration in xyz-space – i.e., of the choice of functions* $y(x)$, $z(x)$.

In order to respond to this question, we choose an arbitrary surface $T(x, y, z) = 0$ and think of the same functions p, q as being determined in such a way that the integral J^* attains a value that is independent of the choice of curve when we take it over a curve lying in $T = 0$ that goes between two points of $T = 0$. Thus, we construct the integral curve of the Lagrange equations:

$$\frac{d\dfrac{\partial F}{\partial y'}}{dx} - \frac{\partial F}{\partial y} = 0, \qquad\qquad [F = F(y', z', y, z; x)],$$

[1] Math. Ann., v. 58, pp. 235.

$$\frac{d\frac{\partial F}{\partial z'}}{dx} - \frac{\partial F}{\partial z} = 0$$

through each point P of the surface $T = 0$ in xyz-space, for which one has:

$$y' = p, \qquad z' = q$$

at each point P, such that a two-parameter family of integral curves that fills up a spatial field comes about in this way. We now think of each point x, y, z of this field as determining the integral curve of the family that goes through it. *The value of the derivatives* y', z' *at each point* x, y, z *are then functions* $p(x, y, z)$, $q(x, y, z)$ *with the desired property.*

In order to prove this assertion, we connect a certain point A of the surface $T = 0$ with an arbitrary point Q of the spatial field by means of a path w; we think of an integral curve of our two-parameter family as going through each point of this path w. The one-parameter family of integral curves that thus arises will be represented by the equations:

$$\left.\begin{aligned} y &= \psi(x, \alpha), \\ z &= \chi(x, \alpha). \end{aligned}\right\} \tag{17}$$

Those points of the surface $T = 0$ from which these integral curves (17) begin define a path w_T on the surface $T = 0$ that leads from the point A to that point P of $T = 0$ that is the starting point of the integral curve of the family that goes through Q.

A surface will be generated by a one-parameter family of curves (17) whose equation:

$$z = f(x, y) \tag{18}$$

is obtained when one eliminates the parameter α from the two equations (17).

If we now introduce the function $f(x, y)$ into F in place of z, and set:

$$F\left(y', \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y', y,\ f(x, y); x\right) = \Phi(y', y; x)$$

then for any curve that lies on the surface (18) one has:

$$\int_a^b F(y', z', y, z; x)\, dx = \int_a^b \Phi(y', y, x)\, dx,$$

and thus vanishes in the xy-plane for curves of the family:

$$y = y(x, \alpha) \tag{19}$$

as well as the first variation of the integral:

$$\int_a^b \Phi(y', y, x)\, dx\,; \tag{20}$$

i.e., the family of curves (19) in the *xy*-plane is a family of integral curves of the Lagrange differential equations, which implies the vanishing of the first variation of the integral (20). From the validity of the independence theorem for *one* function *y* it follows that the integral:

$$\int_a^b \{\Phi + (y' - p)\Phi_p\}\, dx\,, \qquad [\Phi = \Phi(p, y; x)] \tag{21}$$

possesses a value that is independent of the choice of function *y*.

Since:

$$z' = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y'\,,$$
$$q = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} p\,,$$

one will have, however:

$$\frac{\partial f}{\partial y}(y' - p) = z' - q,$$

and as a result, we have:

$$\begin{aligned}\Phi(p, y; x) + (y' - p)\Phi_p &= F(p, q, y, z; x) + (y' - p)\left(F_p + F_q \frac{\partial f}{\partial y}\right)\\ &= F(p, q, y, z; x) + (y' - p)\, F_p + (z' - q)\, F_q\,.\end{aligned}$$

The independence of the integral (21) that we just proved also brings with it the fact that our original integral:

$$J^* = \int_a^b \{F + (y' - p)F_p + (z' - q)F_q\}\, dx$$

preserves its value when we choose our integration path to be, not *w*, but another path from *A* to *Q* that lies on the surface (18) – say, a curve that consists of the path w_T and the integral curve of the family (17) that runs from *P* to *Q*. This fact may be expressed by considering that equations (16) are valid on the path segment *PQ*, so one has the equation:

$$\begin{aligned}&\int_{(w)} \{F + (y' - p)F_p + (z' - q)F_q\}\, dx\\ &\qquad = \int_{(w_T)} \{F + (y' - p)F_p + (z' - q)F_q\}\, dx + \int_P^Q F\, dx\,.\end{aligned} \tag{22}$$

If we let $\bar{w}$ denote any other path in our spatial *pq*-field that leads from *A* to *Q* and let $\bar{w}_T$ denote the corresponding path from *A* to *P* on the surface *T* = 0 then, for the stated reasons, the equation:

$$\int_{(\bar{w})}\{F+(y'-p)F_p+(z'-q)F_q\}\,dx$$
$$=\int_{(\bar{w}_T)}\{F+(y'-p)F_p+(z'-q)F_q\}\,dx+\int_P^Q F\,dx \tag{23}$$

follows, and since, from our assumptions, the first integrals on the right-hand sides of (22) and (23) have equal values – since w_T and $\bar{w}_T$ lie in $T = 0$ – it then follows that the integrals in the left-hand sides of (22) and (23) are also equal to each other, from which, our independence theorem is proved.

The simplest type of functions p, q on surface $T = 0$ that one can choose that are consistent with our demand consists of ones that are determined by the equations:

$$F - pF_p - qF_q : F_p : F_q = \frac{\partial T}{\partial x} : \frac{\partial T}{\partial y} : \frac{\partial T}{\partial z}; \tag{24}$$

the integrand of the integral J^* then vanishes for any path that lies on $T = 0$, and this integral thus has the value zero on $T = 0$, independently of the path.

In particular, one can replace the surface $T = 0$ with a point; all of the integral curves of the Lagrange differential equations that go through this point then define a two-parameter family of curves that one has employed in the construction of the spatial pq-field.

Since the integral J^* will be independent of the path, it represents a function of the variables of the upper limit – i.e., a function of the endpoint x, y, z in the spatial pq-field; we set:

$$J(x, y, z) = \int_A^{x,y,z}\{F+(y'-p)F_p+(z'-q)F_q\}\,dx\,. \tag{25}$$

This function obviously satisfies the equations:

$$\frac{\partial J}{\partial x} = F - pF_p - qF_q\,,$$
$$\frac{\partial J}{\partial y} = F_p\,,$$
$$\frac{\partial J}{\partial z} = F_q\,.$$

If we eliminate the quantities p, q from them then what results is the first-order "Jacobi-Hamilton partial differential equation" for the function $J(x, y, z)$. In particular, if the values of p, q on $T = 0$ are determined by the construction of the spatial pq-field in such a way that the integrand of the integral J^* vanishes – i.e., such that (24) is true – then $J(x, y, z)$ is the solution of a Jacobi-Hamilton differential equation that vanishes on $T = 0$.

If we think of the surface $T = 0$ as belonging to a two-parameter family of surfaces and denote the parameters of this family by a, b then the function p, q of the spatial field, and therefore the function $J(x, y, z)$ will also be independent of these parameters. The differentiation of the equation (25) with respect to the parameters a, b gives:

$$\frac{\partial J}{\partial a} = \int_A^{x,y,z} \left\{ (y' - p)\frac{\partial F_p}{\partial a} + (z' - q)\frac{\partial F_q}{\partial a} \right\} dx\,,$$

$$\frac{\partial J}{\partial b} = \int_A^{x,y,z} \left\{ (y' - p)\frac{\partial F_p}{\partial b} + (z' - q)\frac{\partial F_q}{\partial b} \right\} dx\,,$$

and since, from (16), the integrands of the integrals in the right-hand sides obviously vanish while progressing along an integral curve, these integrals represent functions of *x*, *y*, *z* that assume the same value on any individual integral curve; i.e., the equations:

$$\frac{\partial J}{\partial a} = c,$$

$$\frac{\partial J}{\partial b} = d$$

are – when *c*, *d*, as well as *a*, *b*, mean integration constants – nothing but integrals of the Lagrange differential equations.

This way of looking at things suffices to show how the essential theorems of the Jacobi-Hamilton theory arise immediately from the independence theorem.

Adaptation of the method of independent integrals to double integrals

When one merely treats the question of the conditions for the minimum of an integral, one does not require the given construction of a spatial *pq*-field; it usually suffices to construct a one-parameter family of integral curves (17) of the Lagrange differential equations in such a way that the surface thus generated includes the varied curve *w*. The application of the independence theorem for *one* function in the aforementioned way then leads to the conclusion.

This remark is of use when one wishes to adapt the method of the independent integral to the problem of finding the minimum of a double integral that includes several functions of several variables.

In order to treat such a problem, we let *z*, *t* denote two functions of two variables *x*, *y* and seek to determine these functions in such a way that the double integral:

$$J = \int_{(\Omega)} F(z_x, z_y, t_x, t_y, z; z, t; x, y)\, d\omega\,,$$

$$\left[z_x = \frac{\partial z}{\partial x}, \quad z_y = \frac{\partial z}{\partial y}, \quad t_x = \frac{\partial t}{\partial x}, \quad t_y = \frac{\partial t}{\partial y} \right],$$

which is taken over a given domain Ω in the *xy*-plane, assumes a minimum value when compared to those values that the integral assumes when we replace *z*, *t* with other functions $\bar{z}$, $\bar{t}$ that possess the same prescribed values as z, t on the boundary S of the domain Ω. The Lagrange equations, which are given by the vanishing of the first variation, read:

$$\frac{d}{dx}\frac{\partial F}{\partial z_x}+\frac{d}{dy}\frac{\partial F}{\partial z_y}-\frac{\partial F}{\partial z}=0,$$
$$\frac{d}{dx}\frac{\partial F}{\partial t_x}+\frac{d}{dy}\frac{\partial F}{\partial t_y}-\frac{\partial F}{\partial t}=0$$

in this case.

Furthermore, we start with a certain solution z, t of the Lagrange differential equations, and let $\bar{z}$, $\bar{t}$ be any varied system of functions that satisfy the same boundary conditions as z, t. We then determine a function $S(x, y)$ of the variables x, y such that the equation $S(x, y) = 0$ represents the boundary curve of Ω in the xy-plane, while $S(x, y) = 1$ will be fulfilled only by the coordinates of a single point inside of Ω; finally, the equation $S(x, y) = \alpha$, when α runs through the value between 0 and 1, shall represent a family of curves that fill up the interior of Ω simply without gaps. Thus, we determine those functions:

$$\left.\begin{aligned} z &= \psi(x, y, \alpha), \\ t &= \chi(x, y, \alpha) \end{aligned}\right\} \tag{26}$$

that satisfy the Lagrange differential equations and possess the same prescribed values on the curve $S(x, y) = \alpha$ as the varied system of functions $\bar{z}(x, y)$, $\bar{t}(x, y)$, and are such that for $\alpha = 0$ the functions (26) go over to the original solution z, t. These functions (26) then onviously define a one-parameter family of solution systems for the Lagrange equations for which the equations:

$$\bar{z}(x, y) = \psi(x, y, S(x, y)),$$
$$\bar{t}(x, y) = \chi(x, y, S(x, y)),$$

are satisfied identically.

If we interpret the basic solution z, t of the Lagrange equations as a two-dimensional surface in the $xyzt$-space, and likewise the arbitrarily varied system of functions $\bar{z}$, $\bar{t}$, then the two-dimensional integral surfaces of the one-parameter family (26) generate a three-dimensional subspace of this $xyzt$-space whose equation is obtained by eliminating the α in (26); let the equation of this three-dimensional space take the form:

$$t = f(x, y, z).$$

We assume that the one-parameter family (26) fills up this three-dimensional space simply and without gaps.

If we replace the t in F with the function $f(x, y, z)$ and set:

$$F\left(z_x, z_y, \frac{\partial f}{\partial x}+\frac{\partial f}{\partial z}z_x, \frac{\partial f}{\partial y}+\frac{\partial f}{\partial z}z_y, z, f(x, y, z); x, y\right) = \Phi(z_x, z_y, z; x, y)$$

then it is only necessary for us to apply the independence theorem that I proved in the cited talk for *one* unknown function and the argument that is linked with it to the integral:

$$\int_{(\Omega)} \Phi(z_x, z_y, z; x, y)\, d\omega,$$

in order to recognize that the integral J, under the assumption of a positive E-function for the system of functions $z(x, y)$, $t(x, y)$ before us, actually assumes a minimum value. The appearance of the minimum is thus bound by the following two requirements:

1. Constructibility of the family (26). This requirement is certainly fulfilled when the Lagrange partial differential equations always possess a system of solutions z, t that possess arbitrarily prescribed values on any closed curve K that lies inside of Ω, while they are regular functions of x, y on K.

2. Simple and gapless covering of the three-dimensional space by the family (26). This requirement is certainly fulfilled when each system of solutions z, t of the Lagrange equations is *uniquely* determined by its boundary values on any arbitrary closed curve K that lies inside of Ω.

We can briefly summarize the result as follows:

Our criterion for the attainment of the minimum requires that the boundary-value problem for the Lagrange differential equations relative to any closed curve K that lies inside of Ω be uniquely soluble for arbitrary boundary values. Our argument shows that this criterion is certainly sufficient.

If, in particular, the given function F under the integral sign in the problem treated only happens to be of second degree in z_x, z_y, t_x, t_y, z, t then the Lagrange differential equations will be linear in these quantities, and in this case the boundary-value problem that is required for the application of our criterion is treated completely with the help of my theory of integral equations.

In order to develop the reasoning that comes about in the application in this case more closely, we define that *homogeneous*, linear system of differential equations that arises from the Lagrange equations by dropping the terms that are free of z, t; we would like to refer to this system of equations as the "Jacobi equations." It is now immediately obvious that the boundary-value problem for a curve K only allows several systems of solutions when the Jacobi equations possess a system of solutions z, t that are zero on a curve K, but not everywhere inside the domain that is bounded by K. The theory of integral equations now shows that the latter case is likewise the only one in which the boundary-value problem for the curve K will *not* be soluble for the curve K for certain prescribed boundary values.

In the case of a quadratic F, our criterion for the attainment of the minimum thus comes down to the demand that the Jacobian equations allow no system of solutions z, t, other than zero that are null on boundary S or a closed curve that lies inside of Ω. (The fulfillment of the criterion is also necessary in this case.)

In the general case when the given function F under the integral sign is not quadratic, in particular, but arbitrary, in the functions z, t to be determined and their derivatives, we have to apply the aforementioned criterion on the second variation of the integral J and thus arrive at a criterion is completely analogous to the well-known Jacobi criterion in the case of *one* independent variable or *one* function of several variable to be determined, and which will thus be briefly referred to here as the Jacobi criterion.

Minimum of the sum of a double integral and a simple boundary integral

We finally treat the problem of determining the function z of the variables x, y in such a way that a double integral that is extended over a given domain Ω in the xy-plane and augmented by an integral that is extended over a part S_1 of the boundary of Ω, namely the integral sum:

$$J = \int_{(\Omega)} F(z_x, z_y, z; x, y)\, d\omega + \int_{(S_1)} f(z_s, z; s)\, ds$$

$$\left[z_x = \frac{\partial z}{\partial x}, \quad z_y = \frac{\partial z}{\partial y}, \quad z_s = \frac{\partial z}{\partial s} \right],$$

attains a minimum value, while z shall have prescribed values on the remaining part S_2 of the boundary; thus, F, f are given functions of their arguments and s means the arc length of the boundary curve S of Ω, as measured from a fixed point in the positive sense of the circuit.

The vanishing of the first variation requires that the desired function z, as a function of x, y in the interior of Ω, must fulfill the partial differential equation:

$$\frac{d}{dx}\frac{\partial F}{\partial z_x} + \frac{d}{dy}\frac{\partial F}{\partial z_y} - \frac{\partial F}{\partial z} = 0, \tag{27}$$

while the differential relation:

$$\left(\frac{\partial F}{\partial z_y}\right)_{S_1} \frac{dx}{ds} - \left(\frac{\partial F}{\partial z_x}\right)_{S_1} \frac{dy}{ds} + \frac{d}{ds}\frac{\partial f}{\partial z_s} - \frac{\partial f}{\partial z} = 0 \tag{28}$$

must be valid on the boundary S_1; in it, we understand dx/ds, dy/ds to mean the derivatives of the functions $x(s)$, $y(s)$ that the boundary curve S_1 defines.

We now consider the integral sum:

$$J^* = \int_{(\Omega)} \{F + (z_x - p)F_p + (z_y - q)F_q\}\, d\omega + \int_{(S_1)} \{f + (z_s - \pi)\} ds$$

$$\left[F = F(p, q, z; x, y), \quad F_p = \frac{\partial F}{\partial p}, \quad F_q = \frac{\partial F}{\partial q}, \quad f = f(\pi, z; s), \quad f_\pi = \frac{\partial f}{\partial \pi} \right],$$

and would like to seek to determine the p, q in them as functions of x, y, z and p as a function of s, z in such a way that the value of this integral sum is independent of the surface $z = z(x, y)$ over Ω – i.e., of the choice of function z – when they have only prescribed boundary values in S_2 . The integral sum J^* has the form:

$$\int_{(\Omega)} \{Az_x + Bz_y - C\}\, d\omega + \int_{(S_1)} \{az_s - b\}\, ds\,,$$

where A, B, C represent functions of x, y, z and a, b are functions of s, z. This integral sum is, as one easily recognizes, independent of the surface $z = z(x, y)$ in the desired sense, when the differential equation:

$$\frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z} = 0 \tag{29}$$

is fulfilled identically for x, y, z inside of the xyz-space that projects onto the domain Ω, and the differential equation:

$$(B)_{S_1} \frac{dx}{ds} - (A)_{S_1} \frac{dy}{ds} + \frac{\partial a}{\partial s} + \frac{\partial b}{\partial z} = 0 \tag{30}$$

is fulfilled identically in s, z on the sz-cylinder that projects into the boundary curve S_1 . When we replace A, B, C, a, b with their values:

$$\left.\begin{aligned} A &= F_p, \\ B &= F_q, \\ C &= pF_p + qF_q - F, \\ a &= f_\pi, \\ b &= \pi f_\pi - f, \end{aligned}\right\} \tag{31}$$

the two equations (29), (30) represent partial differential equations for the functions p, q, π.

We now determine a one-parameter family of functions:

$$z = y(x, y, s) \tag{32}$$

that satisfy the Lagrange equations (27), (28), and set:

$$z = \psi(x(s), y(s), \alpha) = \psi(s, \alpha) \tag{33}$$

on the boundary; we assume that this one-parameter family fills up the spatial field in a single-valued and gapless manner. Thus, we compute α as a function of x, y, z from (32) and α as a function of s, z from (33), and form the expressions:

$$p(x, y, z) = \left[\frac{\partial \psi(x, y, \alpha)}{\partial x}\right]_{\alpha=\alpha(x,y,z)},$$
$$q(x, y, z) = \left[\frac{\partial \psi(x, y, \alpha)}{\partial y}\right]_{\alpha=\alpha(x,y,z)},$$
$$\pi(s, z) = \left[\frac{\partial \psi(s, \alpha)}{\partial s}\right]_{\alpha=\alpha(s,z)}.$$

The functions p, q of x, y, z and π of s, z that thus result have the expected property.

In fact, that the functions p, q satisfy the equation (29), follows easily from our consideration of the equation:

$$\frac{\partial p}{\partial y} + q\frac{\partial p}{\partial z} = \frac{\partial q}{\partial x} + p\frac{\partial q}{\partial z}$$

if we think that $\psi(x, y, \alpha)$ shall fulfill the Lagrange equation identically for all values of x, y, α. In order to prove the existence of (30), we substitute:

$$z_x = p,$$
$$z_y = q,$$
$$z_s = p,$$
$$\frac{d^2 z}{ds^2} = \frac{\partial \pi}{\partial s} + \pi\frac{\partial \pi}{\partial z}$$

in the Lagrange equation (28), which is satisfied identically in s, α, and it goes over to the equation:

$$(F_q)_{S_1}\frac{dx}{ds} - (F_p)_{S_1}\frac{dy}{ds} + \frac{\partial^2 f}{\partial \pi^2}\left(\frac{\partial \pi}{\partial s} + \pi\frac{\partial \pi}{\partial z}\right) + \frac{\partial^2 f}{\partial \pi\,\partial z}\pi + \frac{\partial^2 f}{\partial \pi\,\partial s} - \frac{\partial f}{\partial z} = 0,$$

which is valid identically for all s, z. We obtain precisely the same equation when we substitute the expressions (31) in formula (30). With that, the proof of the independence theorem for the present problem is completed.

From the independence theorem, it follows, as before, that:

$$E(z_x, z_y, p, q) \equiv F(z_x, z_y) - F(p, q) - (z_x - p)\, F_p - (z_y - q)\, F_q > 0,$$
$$E(z_s, \pi) \equiv f(z_s) - f(\pi) - (z_s - p)\, f_\pi > 0,$$

such that in the present problem *two Weierstrass E-functions come into consideration: one for the interior and one for the boundary* S_1 .

On the other hand, in order for a one-parameter family (32) to exist that generates a simple, gapless, spatial field in the desired way, we pose the requirement that *any solution z of the Lagrange equations* (27), (28) *must be uniquely determined by its boundary values on any arbitrary curve K that is either closed or begins and ends in* S_1 *and lies inside of* Ω. *Our argument then shows that this criterion is certainly sufficient.*

In particular, if the given functions F, f under the integral sign in the problem being treated happen to be of only second degree in z_x, $_y$, z (z_s , s, resp.) then the Lagrange differential equations become linear. If we then define the *homogeneous*, linear, differential equations that come about by dropping the terms that are free of z from the Lagrange equations and refer to them as the Jacobi equations then it is immediately clear that the boundary-value problem for a curve K admits several solutions only when the Jacobi equations possess a solution z that is null on K, but not everywhere inside of the domain that is bounded by K (K and S_1, resp.).

Thus, in the case of quadratic F, f our criterion for the attainment of the minimum leads to the requirement that the Jacobi equations allow no solution z besides zero that is zero on the boundary S_2 or on a curve K inside of Ω that is either closed or begins and ends inside of S_1 .

In the general case, if the given functions F, f are not quadratic, in particular, but depend arbitrarily upon the function z to be determined and its derivatives then we must apply the aforementioned criterion to the second variation of the integral sum J, and thus arrive at a criterion that is precisely analogous to the well-known Jacobi criterion, and will thus be briefly referred as such here.

When the problem is posed of making the double integral:

$$\int_{(\Omega)} F(z_x, z_y, z; x, y)\, d\omega$$

attain a minimum, while the boundary values for the desired function z shall satisfy the supplementary condition:

$$f(z_s, z; s) = 0,$$

then we can immediately apply the formulas and reasoning of the aforementioned problem; it is only necessary to append the equation $f = 0$ and replace $f(s)$ with $\lambda(s)\, f$ in formulas everywhere, where the Lagrange factor $\lambda(s)$ is to be regarded as a yet-to-be-determined function of s.

General rule for the treatment of variational problems and the statement of a new criterion

In conclusion, let me state a general rule for the treatment of variational problems in which the values of the functions to be determined are prescribed everywhere on the boundary that is an abstraction of the cases that were dealt with above.

First, one obtains the Lagrange equations L of the variational problem by annulling the first variation. Then let a system Z of such solutions of these differential equations L be known that fulfills given conditions B of the variational problem that relate to the interior, as well as the boundary.

When the Weierstrass E-function for our system of solutions Z happens to be positive, we refer to the system of solutions Z as having a *positive-definite character.*

We now fix our attention on any part T of the domain of integration, and denote the boundary of this sub-domain T, as long as it belongs to the boundary of the original domain of integration, by S_T, as long as it lies in the interior of the original domain of integration; thus, the new boundary that results is s_T . Let the conditions B be valid for the first boundary S_T, as well as for the interior of T, as demanded by the present variational problem; for s_T , we prescribe that the values of the functions of the system of solutions Z on it be the boundary values. The system of conditions that comes about for the sub-domain T will be denoted by B_T .

Thus, when no system of solutions of the Lagrange equations L exists that satisfies the conditions B other than the system of solutions Z, and when, furthermore, no system of solutions of the Lagrange equations L exists for each sub-domain T that fulfills the conditions B_T other than the system of solution Z inside of T, the system of solutions Z is said to have an *intrinsically unique character.*

For the system of solutions Z a certain minimum occurs when it has a positive-definite and intrinsically unique character.

As one sees, in the statement thus expressed, a new requirement enters in along with the Weierstrass requirement of the definite character of the solution Z, namely, the requirement of the intrinsically unique character of the solution Z. The latter requirement now has the same relationship to the Jacobi criterion – as it was formulated in the calculus of variations up to now – as the Weierstrass criterion does to the Legendre one, when one regards the Weierstrass criterion as the necessary correct strengthening of the Legendre criterion to arbitrary variations. In fact, just as one will obtain the Legendre criterion by an application of the second variation to the Weierstrass one, so will the Jacobi criterion arise from the one that I presented here (the requirement of the intrinsically unique character of the solution Z) by an application of the second variation. Namely, if we define, by an easily recognizable analogy, the *homogeneous, linear* Jacobi equations [L] from the Lagrange equations L, and likewise, the *homogeneous, linear* associated conditions [B] from the given conditions B then our criterion comes down to the requirement that this homogeneous, linear system of equations and conditions may possess no solution besides zero – and indeed not for any sub-domain T, either – when one also prescribes the boundary value of zero on the new boundary s_T of this sub-domain that then comes about. However, the criterion that I posed is – in analogy with the Weierstrass criterion – also valid as a sufficient criterion without restriction when arbitrary variations come into consideration, not merely ones in a sufficiently close neighborhood; this is then applicable, by way of example, when the decision about the minimum must be made for a curve between two conjugate points, where the Jacobi criterion breaks down.

Whether the criterion that I posed also suffices for boundary values that are not given as fixed – how it is to be modified in that case, resp. – requires an examination in particular cases.

On the calculus of variations for multiple integrals

By C. Caratheodory

Introduction

1. The *Weierstrass* theory of the calculus of variations and the *Jacobi-Hamilton* theory that he employed were completely established by him in two extreme cases, namely, when one has a *simple* integral and n independent functions to vary and when one has a μ-fold integral with *one* function to be varied. By contrast, the general problem of the form:

$$\int\cdots\int f\left(x_1,\ldots,x_n;t_1,\ldots,t_\mu;\frac{\partial x_1}{\partial t_1},\ldots,\frac{\partial x_n}{\partial t_\mu}\right)dt_1\cdots dt_\mu \tag{1.1}$$

is actually never well-posed, when one considers some brief remarks that *Hadamard* foresaw when he remarked on some peculiarities of this problem.[1] In the following pages, I will thoroughly set down the first attempts at a treatment of this problem that seem indispensible to me. My investigations in this regard already go back several years and have also been published piecemeal.[2]

Upon studying the important work of *Haar* on the adjoint variational problem [3] I then remarked that my old computations could be written down in a much more symmetric fashion by a minor modification in the notation. On this basis, the entire system of formulas was derived once again. The first chapter, which is devoted to the derivation of some purely formal identities, thus includes simply the results of my prior work in a new form. However, by means of the newer notation, as well as some advice of Dr. *T. Radó*, I hope that the representation has become more transparent. The second chapter is devoted to the Weierstrass theory for the problem (1.1), which I had previously only inadequately touched upon. The *E*-function that belongs to this problem will be presented here for this first time in canonical coordinates, as well as in the usual coordinates. The same is true for the Legendre condition, as well as the differential equations that the "geodetic fields" must satisfy. Finally, it will be shown that when a geodetic field intersects a surface transversally then this surface must necessarily be a solution of the Euler-Lagrange equations.

The opposite problem of constructing "geodetic fields", i.e., ones through which a *complete figure* of our variational problem can be constructed, can still not yet be solved.

---

[1] *J. Hadamard*, Sur quelques questions de calcul de variations. Bull. Soc. Math. de France **33** (1905), 73-80.

[2] *C. Caratheodory*, Über die kanonischen Veränderlichen in der Variationsrechnung der mehrfachen Integrale. Math. Annalen **85** (1922), 78-88; Über ein Reziprozitätsgesetz der verallgemeinerten Legendreschen Transformation. Math. Annalen **86** (1922), 272-275. [In this work, see v. I, pp. 383-395 and pp. 396-400.]

[3] *A. Haar*. Über adjungierte Variationsprobleme und adjungierte Extremalflächen. Math. Annalen **100** (1928), 481-502.

Chapter I. Formal identities

2. Elementary examples of birational involutory contact transformations. Since, in the sequel, we will occupy ourselves with a contact transformation that is birational and involutory, it is of interest to recall that also any transformations that possess these properties have played a prominent role in the calculus of variations for some time now:

We let the symbols:

(2.1) $$f,\ \varphi,\ p_i,\ \pi_i \qquad (i = 1, 2, \ldots, n)$$

denote a number of quantities, between which (with the usual suppression of the summation sign) the relation:

(2.2) $$f + \varphi = p_i\, \pi_i$$

must exist. We introduce a second sequence of $2n + 2$ quantities, F, Φ, P_i, Π_i, through which the following equations:

(2.3) $$F = \varphi, \qquad \Phi = f, \qquad P_i = \pi_i, \qquad \Pi_i = p_i\,.$$

This transformation is nothing but the *Legendre transformation*; It possesses the following properties:

a) It is *birational* and *involutory*. This means: One solves equations (2.3) for the small symbols by simply exchanging the large symbols with the small ones. Thus, it follows from (2.2) and (2.3) that the relation:

(2.4) $$F + \Pi = P_i\, \Pi_i$$

must exist.

b) It is a *contact transformation*. In fact: If f, φ, p_i, π_i are functions of arbitrary parameters then there always exists the relation:

(2.5) $$dF - \Pi_i\, dP_i = -\,(df - \pi_i\, dp_i).$$

3. The Legendre transformation is naturally not the only transformation that has the properties *a*) and *b*) of § 2. An entirely trivial transformation that obeys them is, e.g., the following one:

(3.1) $$F = -\,f_i\,, \qquad \Phi = -\,\phi_i\,, \qquad P_i = -\,\pi_i\,.$$

4. As a third example, we consider the *generalized inversion* that is defined by the following relations:

(4.1) $$F = \frac{1}{f}, \qquad \Phi = \frac{1}{\varphi}, \qquad P_i = \frac{p_i}{f}, \qquad \Pi_i = \frac{\pi_i}{\varphi}\,.$$

The transformation (4.1) is obviously involutory and birational. In addition, one verifies that the relation (2.4) is a consequence of (2.2), as well as the fact that one is concerned with a contact transformation, with the help of the immediate relations:

(4.2) $$F + \Phi - P_i \Pi_i = \frac{1}{f\varphi}(f + \varphi - p_i \pi_i)$$

(4.3) $$dF - \Pi_i \, dP_i = \frac{1}{f\varphi}(df - \pi_i \, dp_i).$$

One remarks, moreover, that not only (4.1), but also (2.2) must be used to establish (4.3).

5. The transformation that *A. Haar* used in the citation in our footnote 3 is a simple combination of the previous ones when sets:

(5.1) $$F = -\frac{1}{\varphi}, \qquad \Phi = -\frac{1}{f}, \qquad P_i = -\frac{\pi_i}{\varphi}, \qquad \Pi_i = \frac{p_i}{f}.$$

6. A very interesting, but somewhat complicated, birational and involutory contact transformation was used by *T. Levi-Civita* in the regularization of the three-body problem with great success.[1] It consists in the following: If one introduces the notation:

(6.1) $$a = p_i \, p_i, \qquad b = p_i \, \pi_i, \qquad c = \pi_i \, \pi_i,$$

(6.2) $$F = f, \qquad \Phi = \varphi - 2b, \qquad P_i = \frac{p_i}{a}, \qquad \Pi_i = a\pi_i - 2bp_i,$$

(6.3) $$A = P_i P_i, \qquad B = P_i \, \Pi_i, \qquad C = \Pi_i \, \Pi_i$$

then one obtains by completely elementary calculations:

(6.4) $$Aa = 1, \qquad B + b = 0, \qquad AC = ac.$$

From this, one easily verifies properties *a*) and *b*) of § 2.

7. The canonical transformations of the calculus of variations. The main subject of our investigation is a birational, involutory contact transformation, which, when combined with generalized inversion of § 4, gives the generalized Legendre transformation of my older work. The latter possess the advantage that the small symbols can be switched with the large ones in all formulas, but they also have one small disadvantage, that they do not go over to the ordinary Legendre transformation in the limiting cases ($n = 1$ or $\mu = 1$), but the transformations that *Haar* employed.

From now one, we will use the Latin symbols $i, j, k, \ldots$, which run from 1 to n, along with the Greek indices $\alpha, \beta, \gamma, \rho, \sigma, \ldots$, which are to be taken from 1 to μ; e.g., the symbol $p_{i\alpha}$ then represents a matrix of n rows and μ columns.

8. We consider the variables:

[1] *T. Levi-Civita*, Sur la regularization du problème des trois corps. Acta mathematica **42** (1920), 99-144.

(8.1) $$f,\ \varphi,\ p_{i\alpha},\ \pi_{i\alpha},$$
between which the relation:
(8.2) $$f + \varphi = p_{i\alpha}\,\pi_{i\alpha},$$
must exist.

Furthermore, we introduce the symbol:

(8.3) $$a_{\alpha\beta} = \delta_{\alpha\beta}\,f - p_{i\alpha}\,\pi_{i\beta},$$

where, as usual, $\delta_{\alpha\beta}$ shall represent the number one or zero, depending on whether $\alpha = \beta$ or $\alpha \neq \beta$.

To abbreviate, we set the determinant $|\,a_{\alpha\beta}\,|$ equal to a and denote the algebraic complement of $a_{\alpha\beta}$ in this determinant by $\bar{a}_{\alpha\beta}$. One thus has:

(8.4) $$a = |\,a_{\alpha\beta}\,|,$$

(8.5) $$\delta_{\alpha\beta}\,a = a_{\alpha\rho}\,\bar{a}_{\beta\rho} = a_{\sigma\beta}\,\bar{a}_{\sigma\beta}.$$

9. Now, we introduce a new sequence of $2(n\mu + 1)$ variables:

(9.1) $$F,\ \Phi,\ P_{i\alpha},\ \Pi_{i\alpha},$$

which are defined by the following equations:

(9.2) $$\frac{F}{f} = \frac{\Phi}{\varphi} = \frac{f^{\mu-2}}{a},$$

(9.3) $$P_{i\alpha} = \frac{1}{a}\,\pi_{i\rho}\,\bar{a}_{\alpha\rho},$$

(9.4) $$\Pi_{i\alpha} = \frac{f^{\mu-2}}{a}\,p_{i\sigma}\,a_{\alpha\sigma}.$$

It is very remarkable that one can represent the original variables (8.1) as rational functions of the quantities (9.1) by successive solutions of the linear system of equations that was given in (9.2) to (9.4).

10. We first derive some identities that follow from the previous relations. If one replaces the summation symbols α in (9.3) by σ and contract this equation with $a_{\sigma\alpha}$ then it follows, from considering (8.5), that:

(10.1) $$p_{i\alpha} = P_{i\sigma}\,a_{\sigma\alpha}.$$

In an entirely similar manner, one obtains from (9.4):

(10.2) $$p_{i\alpha} = f^{2-\mu}\,\Pi_{i\sigma}\,\bar{a}_{\rho\alpha}\,.$$

Third, it follows from (9.3), when one considers (8.3) that:

$$\begin{aligned} P_{i\alpha}\,p_{i\beta} &= \frac{1}{a}\pi_{i\rho}\,p_{i\beta}\bar{a}_{\alpha\rho} \\ &= \frac{1}{a}\bar{a}_{\alpha\rho}\,(\delta_{\beta\rho}\,f - a_{\beta\rho})\,, \end{aligned}$$

and thus it follows from (8.5) and (9.2) that:

(10.3) $$\bar{a}_{\alpha\beta} = \frac{f^{\mu-2}}{F}\,(\delta_{\alpha\beta} + P_{i\alpha}\,p_{i\beta}).$$

In order to ultimately present the latter of the relations that we have considered here, we state the following equation with the help of (9.3) and (9.4):

$$P_{i\alpha}\,\Pi_{i\beta} = \frac{f^{\mu-2}}{a^2}\,\pi_{i\rho}\,p_{i\sigma}\bar{a}_{\alpha\rho}\,a_{\beta\sigma}\,.$$

It then follows from (8.3) and (8.5) that:

$$P_{i\alpha}\,\Pi_{i\beta} = \frac{f^{\mu-2}}{a^2}\,(\delta_{\sigma\rho} - a_{\sigma\rho})\,\bar{a}_{\alpha\rho}\,a_{\beta\sigma} = \delta_{\alpha\beta}\,\frac{f^{\mu-1}}{a} - \frac{f^{\mu-2}}{a}\,a_{\beta\alpha}$$

or, from (9.2):

(10.4) $$P_{i\alpha}\,\Pi_{i\beta} = \delta_{\alpha\beta}\,F - \frac{F}{f}\,a_{\beta\alpha}\,.$$

By (8.3), the relation (10.4) can also be symmetrically written as:

(10.5) $$\frac{1}{F}\,P_{i\alpha}\,\Pi_{i\beta} = \frac{1}{f}\;p_{i\alpha}\,\pi_{i\beta}.$$

11. From (9.2) and (8.2), we find that:

$$F + \Phi = \frac{F}{f}\,(f + \varphi) = \frac{F}{f}\,p_{i\alpha}\,\pi_{i\alpha},$$

or, from (10.5):

(11.1) $$F + \Phi = P_{i\alpha}\,\Pi_{i\alpha}.$$

12. We now introduce a notation that is analogous to (8.3):

(12.1) $$A_{\alpha\beta} = \delta_{\alpha\beta} F - P_{i\alpha}\Pi_{i\beta};$$

it then follows from (10.4) that:

(12.1) $$\frac{A_{\alpha\beta}}{F} = \frac{a_{\beta\alpha}}{f},$$

and thus, when one also recalls our previous notation for the large symbols:

(12.3) $$\frac{A}{F^{\mu}} = \frac{a}{f^{\mu}},$$

(12.4) $$\frac{\bar{A}_{\alpha\beta}}{F^{\mu-1}} = \frac{\bar{a}_{\beta\alpha}}{f^{\mu-1}}.$$

The comparison of (9.2) with (12.3) now gives:

(12.5) $$\frac{f}{F} = \frac{\varphi}{\Phi} = \frac{F^{\mu-2}}{A}.$$

Furthermore, it follows from (10.2), when one uses (12.4) and then (12.5) that:

(12.6) $$p_{i\alpha} = \frac{1}{A}\Pi_{i\rho}\bar{A}_{\alpha\rho};$$

and likewise one obtains from (10.1), (12.2), and (12.3):

(12.7) $$\pi_{i\alpha} = \frac{F^{\mu-1}}{A}\, P_{i\sigma}A_{\alpha\sigma}.$$

If one compares (8.2) with (11.1), and then (9.2), (9.3), and (9.4), resp., with (12.5), (12.6), and (12.7), resp. then one sees that one can switch the large symbols with the small ones in these and therefore also all of the remaining equations.

Our transformation is thus *birational* and *involutory*.

13. Introduction of f, F, $p_{i\alpha}$, $P_{i\alpha}$ as variables. Up till now, we have alternately based our calculations on the system of quantities (8.1) and (9.1). For many purposes, it is more convenient to develop formulas in which the quantities:

(13.1) $$f,\ F,\ p_{i\alpha},\ P_{i\alpha},$$

appear as basic variables.

Thus, we set:

(13.2) $$g_{\alpha\beta} = \delta_{\alpha\beta} + P_{i\alpha}p_{i\beta},$$

such that, from (10.3), one has:

$$g_{\alpha\beta} = \frac{F}{f^{\mu-2}} \bar{a}_{\alpha\beta} . \tag{13.3}$$

In order to compute the value g of the determinant $|g_{\alpha\beta}|$, we remark that $|\bar{a}_{\alpha\beta}| = a^{\mu-1}$; it then follows from (13.3) and (9.2) that:

$$g = Ff, \tag{13.4}$$

and (13.3) can then be written:

$$g_{\alpha\beta} = \frac{g}{f^{\mu-1}} \bar{a}_{\alpha\beta} . \tag{13.5}$$

From this latter equation, we gather that:

$$g_{\rho\sigma}\bar{g}_{\rho\beta}a_{\alpha\sigma} = \frac{g}{f^{\mu-1}} \bar{a}_{\rho\sigma}\bar{g}_{\rho\beta}a_{\alpha\sigma}$$

or:

$$a_{\alpha\beta} = \frac{a}{f^{\mu-1}} \bar{g}_{\alpha\beta} ; \tag{13.6}$$

hence, from (9.2):

$$\bar{g}_{\alpha\beta} = Fa_{\alpha\beta} . \tag{13.7}$$

Due to (10.1), it now follows that:

$$F\pi_{i\alpha} = P_{i\sigma}\bar{g}_{\sigma\alpha} , \tag{13.8}$$

and, from (9.4):

$$f\,\Pi_{i\alpha} = p_{i\sigma}\bar{g}_{\alpha\sigma} . \tag{13.9}$$

Ultimately, when one solves the latter two equations for $p_{i\alpha}$ and $P_{i\alpha}$ they give the relations:

$$F\,p_{i\alpha} = \Pi_{i\rho}\,g_{\rho\alpha} , \tag{13.10}$$

$$f\,P_{i\alpha} = \pi_{i\rho}\,g_{\alpha\rho} . \tag{13.11}$$

14. The property of contact transformation. We now assume that the quantities that we are considering depend upon arbitrary parameters and then form the total differential of (13.4) with respect to these parameters. In this way, we obtain the relation:

$$F\,df + f\,dF = dg; \tag{14.1}$$

however, it is now well-known that:

$$dg = \bar{g}_{\alpha\beta}dg_{\alpha\beta}$$

and, from (13.2):

$$dg_{\alpha\beta} = P_{i\alpha}\, dp_{i\beta} + p_{i\beta}\, dP_{i\alpha}\,.$$

From the latter two equations one then has, when one recalls (13.8) and (13.9):

$$dg = F\, p_{i\beta}\, dp_{i\beta} + f\, \Pi_{i\alpha}\, dP_{i\alpha}\,. \tag{14.2}$$

The comparison of (14.2) with (14.1) ultimately leads to the relation:

$$F(df - p_{i\beta}\, dp_{i\beta}) + f(dF - \Pi_{i\alpha}\, dP_{i\alpha}) = 0, \tag{14.3}$$

from which it follows that our transformation is a *contact transformation.*

15. Reciprocity. In a previous work [1] I made the remark, which can, moreover, be immediately confirmed, that the determinant a, as we showed in § **8**, can also be written as a $(\mu + n)$-rowed determinant in the following way:

$$a = \begin{vmatrix} \delta_{ij} & \pi_{i\beta} \\ p_{j\alpha} & \delta_{\alpha\beta} f \end{vmatrix}; \tag{15.1}$$

in this formula, the rows are denoted by i and α, and the columns by j and β. In the same way, one sees, when one introduces a new system of variables by the equations:

$$b_{ij} = \delta_{ij}\, f - p_{i\rho}\, \pi_{j\rho}\,, \tag{15.2}$$

that the determinant b of b_{ij} can be written:

$$b = \begin{vmatrix} \delta_{ij} f & \pi_{i\beta} \\ p_{j\alpha} & \delta_{\alpha\beta} \end{vmatrix}. \tag{15.3}$$

Comparing (15.1) and (15.3) then leads to the relation:

$$f^n\, a = f^\mu\, b, \tag{15.4}$$

from which we deduce, with the help of (9.2):*

$$\frac{F}{f} = \frac{\Phi}{\varphi} = \frac{f^{n-2}}{b}\,. \tag{15.5}$$

Furthermore, it follows from (15.2) that:

$$b_{si}\, P_{s\alpha} = f\, P_{i\alpha} - p_{s\rho}\, \pi_{i\rho}\, P_{s\alpha}\,;$$

[1] Cf., footnote 2 [on page 1 of this article].

* [A printing error in (15.5) in the original has been corrected.]

from (13.2), one can write this as:

$$b_{si}\,P_{s\alpha} = f\,P_{i\alpha} - \pi_{i\rho}\,(g_{\alpha\rho} - \delta_{\alpha\rho});$$

or, taking (13.11) into account:

(15.6) $$\pi_{i\alpha} = b_{si}\,P_{s\alpha}\,.$$

Similarly, we deduce from (15.2) that:

$$\begin{aligned} b_{si}\,p_{s\alpha} &= f\,p_{i\alpha} - p_{i\rho}\,\pi_{i\sigma}\,p_{i\alpha} \\ &= p_{i\alpha}\,(\delta_{\alpha\rho}\,f - p_{i\alpha}\,\pi_{i\sigma}), \end{aligned}$$

or, from (8.3):**

(15.7) $$b_{it}\,p_{t\alpha} = p_{i\rho}\,a_{\alpha\rho}\,.$$

From (9.4), one thus obtains, when one observes (15.4):

(15.8) $$\Pi_{i\alpha} = \frac{f^{n-2}}{b}\,p_{t\alpha}\,b_{it}\,.$$

Finally, it follows by solving (15.6) that:

(15.9) $$P_{i\alpha} = \frac{1}{b}\,\pi_{r\alpha}\,\bar{b}_{ir}\,.$$

16. The similarity between formulas (15.2), (15.5), (15.9), and (15.8) and (8.3), (9.2), (9.3), and (9.4) shows that [in the same sense as on pp. 397 of this volume]* in all of our equations one can switch the Latin indices with the Greek ones when one simply replaces $a_{\alpha\beta}$ with b_{ij} .

17. Introduction of the parameters $S_{\alpha\beta}$, $S_{\alpha i}$, and $c_{\alpha\beta}$. For the treatment of our variational problem, it is necessary to introduce new parameters and to examine their connection with the previous notation.

To that end, we consider three matrices:

(17.1) $$S_{\alpha\beta},\ S_{\alpha i}\ ,\ c_{\alpha\beta}\,,$$

which shall be linked with the previous quantities by the relations:

(17.2) $$c_{\alpha\beta} = S_{\alpha\beta} + S_{\alpha i}\,p_{i\beta}\,,$$

(17.3) $$S_{\alpha i} = p_{i\rho}\,S_{\alpha\rho}\,,$$

** [A printing error in the original was corrected here.]

* [The comment in square brackets was incorrect in the original.]

(17.4) $$\frac{1}{F} = |S_{\alpha\beta}|.$$

By replacing (17.3) in (17.2) one now obtains:[**]

$$\begin{aligned} c_{\alpha\beta} &= S_{\alpha\beta} + S_{\alpha\rho} P_{i\rho} p_{i\beta}, \\ &= S_{\alpha\rho} (\delta_{\rho\beta} + P_{i\rho} p_{i\beta}), \end{aligned}$$

or, from (13.2):

(17.5) $$c_{\alpha\beta} = S_{\alpha\rho} g_{\rho\beta}.$$

From the laws of multiplication for determinants, it now follows, when one observe (13.4), that $c = Ff \,|\, S_{\alpha\rho}|$, or, from (17.4):

(17.6) $$c = f.$$

Furthermore, one deduces from (17.5) that:

$$c_{\lambda\rho} \bar{c}_{\lambda\beta} \bar{g}_{\alpha\sigma} = S_{\lambda\sigma} g_{\rho\sigma} \bar{g}_{\alpha\sigma} \bar{c}_{\lambda\beta},$$

and from this it follows, from (13.4) and (17.6) that:

(17.7) $$\bar{g}_{\alpha\beta} = F\, S_{\lambda\alpha}\, \bar{c}_{\lambda\beta}.$$

From this, it follows, using (13.8), that:

(17.8) $$\pi_{i\alpha} = S_{\lambda i} \bar{c}_{\lambda\alpha}.$$

18. We would now like to show that when one assumes (17.2) equations (17.6) and (17.8) are equivalent to equations (17.3) and (17.4). We thus now come to the equations:

(18.1) $$c_{\alpha\beta} = S_{\alpha\beta} + S_{\alpha i}\, p_{i\beta},$$

(18.2) $$\pi_{i\alpha} = S_{\rho i} \bar{c}_{\rho\alpha},$$

(18.3) $$c = f,$$

and would like to derive (17.3) and (17.4). First, it follows from (18.2) that:

$$\pi_{i\sigma}\, c_{\alpha\sigma} = S_{\rho i} \bar{c}_{\rho\sigma} c_{\alpha\sigma},$$

hence, upon considering (18.3):

(18.4) $$f\, S_{\alpha i} = \pi_{i\sigma}\, c_{\alpha\sigma}.$$

[**] [In the following equation, a printing error in the original was corrected.]

This, when substituted in (18.1), gives:

$$f\,c_{\alpha\beta} = f\,S_{\alpha\beta} + c_{\alpha\sigma}\,p_{i\beta}\,\pi_{i\sigma}\,,$$

from which, from (8.3), it follows that:

$$f\,S_{\alpha\beta} = c_{\alpha\sigma}\,a_{\beta\sigma}\,. \tag{18.5}$$

From (18.5), it next follows that:

$$f\,P_{i\rho}\,S_{\alpha\rho} = c_{\alpha\sigma}\,P_{i\rho}\,a_{\rho\sigma}\,.$$

The right-hand side of the latter equation is, from (10.1), equal to $c_{\alpha\sigma}\,\pi_{i\sigma}$, and with the help of (18.4) one ultimately obtains:

$$S_{\alpha i} = P_{i\sigma}\,S_{\alpha\sigma}; \tag{18.6}$$

i.e., relation (17.3), as we wished to prove. Equation (17.4) is likewise a consequence of (18.5) when one observes (18.3) and (9.2); it then becomes:

$$f^{\mu}\,|\,S_{\alpha\beta}\,| = ac = f\,\frac{f^{\mu-1}}{F}$$

or:

$$F\,|\,S_{\alpha\beta}\,| = 1. \tag{18.7}$$

Chapter II. The variational problem.

19. Definition of the geodetic fields. In an $(n + \mu)$-dimensional space whose coordinates are $x_1, \ldots, x_n, t_1, \ldots, t_\mu$, or, with the previous notation, x_i, t_α, we consider a μ-parameter family of n-dimensional surfaces. A family of this sort can be represented by μ equations of the form:

$$S_\alpha(x_i;\,t_\beta) = \lambda_\alpha\,. \tag{19.1}$$

Furthermore, a μ-dimensional manifold will be defined by the equations:

$$x_i = \xi_i(t_\alpha) \qquad (i = 1, 2, \ldots, n), \tag{19.2}$$

which intersect the family (19.1). This is the case when and only when a one-to-one map of a region G_t in the μ-dimensional space of t_α onto a region G_λ of the μ-dimensional parameter space of the λ_α is generated by the system of equations:

$$S_\alpha(\xi_i(t_\gamma);\,t_\beta) = \lambda_\alpha\,. \tag{19.3}$$

In this, however, the functional determinant, in particular:

(19.4) $$\Delta = \left|\frac{\partial S_\alpha(\xi_i; t_\beta)}{\partial t_\beta}\right|$$

must be non-zero in G_t .

If one sets, to abbreviate:

(19.5) $$S_{\alpha i} = \frac{\partial S}{\partial x_i}, \qquad S_{\alpha\beta} = \frac{\partial S_\alpha}{\partial x_\beta},$$

(19.6) $$p_{i\alpha} = \frac{\partial \xi_i}{\partial t_\alpha},$$

(19.7) $$c_{\alpha\beta} = S_{\alpha\beta} + S_{\alpha i}\, p_{i\beta},$$

then (19.4) takes the form:

(19.8) $$\Delta = |\, c_{\alpha\beta} \,| = c.$$

We further remark that the integral:

(19.9) $$\int \underset{G_t}{\cdots} \int \Delta\, dt_1 \cdots dt_\mu$$

represents the volume of the region G_λ in the parameter space, onto which the region G_t is mapped by the relation (19.3).

However, this volume depends only[*] *upon the form of the boundary of* G_λ .

If one thus considers a second μ-dimensional surface:

(19.10) $$x_i = \xi_i(t_\alpha)$$

and a region $\bar{G}_t$ that is mapped onto *the same* region G_λ that we have just considered through this new surface then the integral:

(19.11) $$\int \underset{\bar{G}_t}{\cdots} \int \bar{\Delta}\, dt_1 \cdots dt_\mu ,$$

which will be mapped in a manner that is completely analogous to (19.9), will possess the same value as (19.9).

In particular, if a manifold that also lies on (19.10) is taken from surface (19.2) through the boundary of the region G_t then one must compute the integrals (19.9) and (19.11) for the same region G_t ; i.e., one must set $\bar{G}_t = G_t$.

20. The coordinates of a μ-dimensional surface element of the $(n + \mu)$-dimensional space shall now be represented by the $n + \mu + n\mu$ quantities:

[*] [In the original, the word was "nun" here. In the Gesammelte Mathematische Schriften, the word was "uur," when it should have been "nur;" i.e., there was a printing error in the correction to the printing error!]

(20.1) $$x_i,\ t_\alpha,\ p_{i\alpha}.$$

We now consider a *positive* function:

(20.2) $$f(x_i,\ t_\alpha,\ p_{i\alpha})$$

of these quantities and form the expression:

(20.3) $$\frac{f(x_i, t_\alpha, p_{i\alpha})}{\Delta(x_i, t_\alpha, p_{i\alpha})},$$

in which Δ shall have the same meaning as in (19.8). We now hold the $(x_i,\ t_\alpha)$ fixed in (20.3) and seek to determine the $p_{i\alpha}$ in such a way that:

(20.4) $$\frac{f}{\Delta} = \text{minimum};$$

we say of a surface element (20.1) for which the condition (20.4) is satisfied that it intersects the family of surfaces (19.1) *transversally*.

We now assume that in a certain $(n+\mu)$-dimensional region of the space of $(x_i,\ t_\alpha)$ we can determine functions:

(20.5) $$p_{i\alpha} = p_{i\alpha}(x_j,\ t_\beta),$$

which generate nothing but surface elements that will intersect our family (19.1) transversally.

If we now substitute the values (20.5) of the $p_{i\alpha}$ in $f(x_i,\ t_\alpha,\ p_{i\alpha})$ and $\Delta(x_i,\ t_\alpha,\ p_{i\alpha})$, and if one then has the validity of the equation:

(20.6) $$f = \Delta$$

at any point of the region in question then we would like to say that the family (19.1) forms a *geodetic field* (associated with f). A necessary condition for the validity of (20.4) will be given by the equations:

$$\frac{\partial}{\partial p_{i\alpha}}\left(\frac{f}{\Delta}\right) = 0,$$

which can, due to (20.6), be written in the following manner:

(20.7) $$f_{p_{i\alpha}} = \frac{\partial \Delta}{\partial p_{i\alpha}}.$$

The equations (20.6) and (20.7) comprise the fundamental relations through which a geodetic field is defined.

21. Solution of the variational problem. If one has constructed, by whatever means, a geodesic field that intersects a manifold (19.2) transversally, moreover, then it always constitutes a solution of the variational problem that is associated with the integral:

$$\int \cdots \int f(x_i, t_\alpha, p_{i\alpha})\, dt_1 \ldots dt_\mu . \tag{21.1}$$

Namely, we consider a piece of (19.2) that projects onto a region G_t of the space of t_α and a corresponding piece of (19.10) that projects onto $\bar{G}_t$, in which the relations between G_t and $\bar{G}_t$ that were specified at the end of § **19** shall be valid. Then, from the results of § **19**, it is linked with (20.6) that:

$$\int_{G_t} f\, dt_1 \cdots dt_\mu = \int_{G_t} \Delta\, dt_1 \cdots dt_\mu = \int_{\bar{G}_t} \bar{\Delta}\, dt_1 \cdots dt_\mu . \tag{21.2}$$

If one then denotes the value of f on the surface (19.10) by $\bar{f}$ then, from (21.2), one has:

$$\int_{\bar{G}_t} \bar{f}\, dt_1 \cdots dt_\mu - \int_{G_t} f\, dt_1 \cdots dt_\mu = \int_{\bar{G}_t} (\bar{f} - \bar{\Delta})\, dt_1 \cdots dt_\mu . \tag{21.3}$$

One now remarks that from the fact that $f > 0$ and also from (20.6) it follows that $\Delta > 0$. For a weak variation of our original surface piece one therefore also has $\bar{\Delta} > 0$. From this, it follows, upon considering (20.4) and (20.6) that:

$$\bar{f} - \bar{\Delta} = \bar{\Delta}\left(\frac{\bar{f}}{\bar{\Delta}} - \frac{f}{\Delta}\right) > 0, \tag{21.4}$$

by which our main assertion is proved.

22. Introduction of canonical variables. The further treatment of our problem will be simplified considerably if we now introduce the canonical quantities F, $P_{i\alpha}$, $\Pi_{i\alpha}$ that we examined in the first chapter. In fact, from (19.7) and (19.8), one has:

$$\frac{\partial \Delta}{\partial p_{i\alpha}} = S_{\lambda i} \bar{c}_{\lambda\alpha}, \tag{22.1}$$

and comparison of this formula with (17.8) and (20.7) shows that we must set:

$$p_{i\alpha} = f_{p_{i\alpha}} . \tag{22.2}$$

Thus, from §§ **8** and **9**, one can now compute the $a_{\alpha\beta}$, F, Φ, $P_{i\alpha}$, $\Pi_{i\alpha}$ as functions of x_i, t_α, $p_{i\alpha}$ by rational operations. Likewise, one can compute the determinant a and, in particular, verify that it does not vanish. In case it vanishes identically, the function f, upon which our variational problem depends, is not useful for our theory.

However, our goal is to take x_i, t_α, $P_{i\alpha}$ as the independent variables, and we must therefore present the condition that must be verified in order to express the $P_{i\alpha}$ in terms of these quantities. Thus, the best equation to make use of is (10.1), an equation that, from (8.3), can be written most effectively in the following form:

$$M_{i\alpha} = p_{i\alpha} - P_{i\sigma}(\delta_{\sigma\alpha} f - p_{k\sigma}\,\pi_{k\alpha}) = 0. \tag{22.3}$$

This latter system of equation shall thus be soluble in terms of the $P_{i\alpha}$, and we must therefore demand that the functional determinant must satisfy:

$$\left|\frac{\partial M_{i\alpha}}{\partial p_{j\beta}}\right| \neq 0. \tag{22.4}$$

If one sets, to abbreviate:

$$\frac{\partial^2 f}{\partial p_{i\alpha}\,\partial p_{j\beta}} = \pi_{i\alpha,\,j\beta}\,, \tag{22.5}$$

then it follows from (22.3) that:

$$\frac{\partial M_{i\alpha}}{\partial p_{j\beta}} = \pi_{i\alpha,\,j\beta} - P_{i\alpha}\,\pi_{j\beta} + P_{i\beta}\,\pi_{j\alpha} + P_{i\sigma}\,p_{k\sigma}\,\pi_{k\alpha,\,j\beta}. \tag{22.6}$$

From our assumptions, it now follows that $b \neq 0$; therefore, one can replace the condition (22.4) with the non-vanishing of a determinant whose elements are:

$$N_{i\alpha,\,j\beta} = b_{ri}\,\delta_{\sigma\alpha}\frac{\partial M_{r\sigma}}{\partial p_{j\beta}} = b_{ri}\frac{\partial M_{r\alpha}}{\partial p_{j\beta}}. \tag{22.7}$$

From (22.6) and (22.7), it now follows, with the help of (15.6), that:

$$N_{i\alpha,\,j\beta} = b_{ri}\;\pi_{r\alpha,\,j\beta} - \pi_{i\alpha}\,\pi_{j\beta} + \pi_{i\beta}\,\pi_{j\alpha} + \pi_{i\sigma}\,p_{k\sigma}\,\pi_{k\alpha,\,j\beta}; \tag{22.8}$$

now, since one has $b_{ri}\;\pi_{r\alpha,\,j\beta} = b_{ki}\;\pi_{k\alpha,\,j\beta}$ and $b_{ri} + \pi_{i\sigma,}\,p_{k\sigma,} = \delta_{ik}\,f$, one can also write (22.8) in the form:

$$N_{i\alpha,\,j\beta} = f\,\pi_{i\alpha,\,j\beta} - \pi_{i\alpha}\,\pi_{j\beta} + \pi_{i\beta}\,\pi_{j\alpha}; \tag{22.9}$$

The introduction of the $P_{i\alpha}$ as independent variables is then always possible, as long as the determinant satisfies:

$$\left| f\frac{\partial^2 f}{\partial p_{i\alpha}\,\partial p_{j\beta}} + \frac{\partial f}{\partial p_{i\beta}}\frac{\partial f}{\partial p_{j\alpha}} - \frac{\partial f}{\partial p_{i\alpha}}\frac{\partial f}{\partial p_{j\beta}}\right| \neq 0. \tag{22.10}$$

23. Once we have represented the $P_{j\beta}$ as functions of x_i, t_α, $P_{i\alpha}$, we can, from chapter I, determine the remaining quantities – hence, in particular, F, Φ, $\Pi_{j\beta}$ – as functions of these variables.

If one now substitutes these functions in (14.3) then it immediately follows that one has the relations:

$$F f_{x_i} = - f\, F_{x_i}\,, \qquad F\, f_{t_\alpha} = - f\, F_{t_\alpha}\,, \tag{23.1}$$

$$\Pi_{i\alpha} = \frac{\partial F}{\partial P_{i\alpha}}\,, \tag{23.2}$$

which we shall use later.

24. The *E*-function. We are now in a position to compute the Weierstrass excess function that belongs to the integral (1.1) for any geodetic field.

Thus, the equations (20.6) and (20.7) shall be valid; however, if one considers (19.8), (22.2), and (22.1) then one sees that equations (18.2) and (18.3) must also be valid, and, from § **18**, equivalent to (17.3) and (17.4).

If we thus substitute the right-hand side of (17.3) for the $S_{\alpha i}$ in:

$$\bar{\Delta} = \left| S_{\alpha\beta} + S_{\alpha i}\, \bar{p}_{i\beta} \right| \tag{24.1}$$

then it follows from the multiplication rules for determinants, when one observes (17.4), that:

$$\bar{\Delta} = \frac{1}{F} \left| \delta_{\alpha\beta} + P_{\alpha i}\, \bar{p}_{i\beta} \right|. \tag{24.2}$$

Now, we define new quantities $h_{i\beta}$ by the equations:

$$\bar{p}_{i\beta} = p_{i\beta} + f \cdot h_{i\beta}\,, \tag{24.3}$$

and obtain from (24.2), upon consideration of (13.2):

$$\bar{\Delta} = \frac{1}{F} \left| g_{\alpha\beta} + f\, P_{i\alpha}\, h_{i\beta} \right|. \tag{24.4}$$

If we now remark that from (13.4) and (13.8) it follows that:

$$(g_{\rho\beta} + f\, P_{i\rho}\, h_{i\beta})\, \bar{g}_{\sigma\beta} = F\, f(\delta_{\alpha\beta} + \pi_{i\alpha}\, h_{i\beta}) \tag{24.5}$$

then, from (24.4), we obtain:

$$\bar{\Delta} \left| \bar{g}_{\alpha\beta} \right| = F^{\mu-1} f^{\mu} \left| \delta_{\alpha\beta} + \pi_{i\alpha}\, h_{i\beta} \right|.$$

Now, since, from (13.4), one has $\left| \bar{g}_{\alpha\beta} \right| = F^{\mu-1} f^{\mu}$, one finally obtains:

(24.6) $$\bar{\Delta} = f\,|\delta_{\alpha\beta} + \pi_{i\alpha}\,h_{i\beta}\,|.$$

We now compute $h_{i\beta}$ from (24.3) and remark that, from § **21**, one must take:

$$E = \bar{f} - \bar{\Delta}\,.$$

Ultimately, one thus has:

(24.7) $$\boxed{E = \bar{f} - \frac{1}{f^{\mu-1}}\,|\;\delta_{\alpha\beta}\,f + \pi_{i\alpha}(\,\bar{p}_{i\beta} - p_{i\beta})\;|\,.}$$

This is a formula for the E-function that goes over to the usual one when $\mu = 1$; it is also noteworthy that the E-function depends only upon the surface elements $p_{i\beta}$, $\bar{p}_{i\beta}$ here, but not, however, on the geodetic field.

25. The Legendre condition. We develop the determinant (24.6) in powers of $h_{i\beta}$ and determine and quadratic terms of this development. Thus, we introduce the notation:

(25.1) $$m_{\alpha\beta} = \delta_{\alpha\beta} + \pi_{i\alpha}\,h_{i\beta}\,,$$

from which, by abbreviations that are similar to the ones in § **8**, it follows that:

$$\delta_{\alpha\beta}\,m = m_{\rho\beta}\bar{m}_{\rho\alpha}$$

and, by differentiation:

$$\delta_{\alpha\beta}\,dm = m_{\rho\beta}d\bar{m}_{\rho\alpha} + \bar{m}_{\rho\alpha}dm_{\rho\beta}\,.$$

We contract this equation with $\bar{m}_{\sigma\beta}$ and obtain, when we replace the summation symbol β with λ:

$$m\,d\bar{m}_{\sigma\alpha} = \bar{m}_{\sigma\alpha}dm - \bar{m}_{\sigma\alpha}\bar{m}_{\sigma\lambda}dm_{\rho\lambda}\,.$$

Now, it is well-known that $dm = \bar{m}_{\rho\lambda}dm_{\rho\lambda}$, and we ultimately have that:

(25.2) $$m\,d\bar{m}_{\sigma\alpha} = (\bar{m}_{\sigma\alpha}\bar{m}_{\rho\lambda} - \bar{m}_{\rho\alpha}\bar{m}_{\sigma\lambda})\,dm_{\rho\lambda}\,.$$

It then follows from (24.6) that:

(25.3) $$\frac{\partial\bar{\Delta}}{\partial h_{i\alpha}} = f\,\bar{m}_{\sigma\alpha}\pi_{i\sigma}$$

and from (25.2):

(25.4) $$m\frac{\partial\bar{m}_{\sigma\alpha}}{\partial h_{j\beta}} = (\bar{m}_{\sigma\alpha}\bar{m}_{\rho\beta} - \bar{m}_{\rho\alpha}\bar{m}_{\sigma\beta})\,\pi_{j\rho}\,.$$

Hence, from (25.3):

(25.5) $$\frac{\partial^2\bar{\Delta}}{\partial h_{i\alpha}\,\partial h_{j\beta}} = \frac{f}{m}(\bar{m}_{\sigma\alpha}\bar{m}_{\rho\beta} - \bar{m}_{\rho\alpha}\bar{m}_{\sigma\beta})\,\pi_{i\sigma}\pi_{j\sigma}\,.$$

For $h_{i\beta} = 0$, we now have $\bar{m}_{\alpha\beta} = \delta_{\alpha\beta}$, and from (25.3) and (25.5), it thus follows that:

(25.6) $$\left.\frac{\partial\bar{\Delta}}{\partial h_{i\alpha}}\right|_0 = f\,\pi_{i\alpha}\,,$$

(25.7) $$\frac{\partial^2\bar{\Delta}}{\partial h_{i\alpha}\,\partial h_{j\beta}} = f(\pi_{i\alpha}\,\pi_{j\beta} - \pi_{i\beta}\,\pi_{j\alpha}).$$

Thus, if we now develop the E-function (24.7) in powers of ($\bar{p}_{i\beta} - p_{i\beta}$) then the constant drops away along with the linear terms. The quadratic terms in the development define a quadratic form, which reads as follows:

(25.8) $$\boxed{2Q = (\bar{p}_{i\alpha} - p_{i\alpha})(\bar{p}_{j\beta} - p_{j\beta})\left\{\frac{\partial^2 f}{\partial p_{i\alpha}\,\partial p_{j\beta}} - \frac{1}{f}\left(\frac{\partial f}{\partial p_{i\alpha}}\frac{\partial f}{\partial p_{j\beta}} - \frac{\partial f}{\partial p_{j\alpha}}\frac{\partial f}{\partial p_{i\beta}}\right)\right\}.}$$

The *Legendre condition* of our problem consists in the requirement that the quadratic form (25.8) must be positive definite. One should observe that the determinant of this quadratic form agrees [up to a positive factor]* with the expression (22.10); whenever the Legendre condition is satisfied, one also has the possibility of introducing canonical coordinates.

Finally, we remark that the first derivatives of f with respect to $p_{i\alpha}$ are also present in the Legendre condition.

26. The *E*-function in canonical coordinates. For the case in which one presents the variational problem in canonical coordinates from the outset by means of the function $F(x_i,\ t_\alpha,\ P_{i\alpha})$, it is useful to have an expression for the E-function in which these coordinates alone appear. Thus, one sets:

(24.1) $$P_{i\alpha} = \bar{P}_{i\alpha} + \bar{F}\,k_{i\alpha}$$

in (24.2) and transforms the expression (24.2) in a completely similar manner to what we did in § **24**. One ultimately finds that:

(26.2) $$\boxed{\frac{F}{\bar{f}}E = F - \frac{1}{\bar{F}^{\mu-1}}\left|\delta_{\alpha\beta}\bar{F} + \bar{\Pi}_{i\beta}(P_{i\alpha} - \bar{P}_{i\alpha})\right|.}$$

* [The comment in square brackets was absent in the original.]

If one computes the Legendre condition from (26.2) then one finds a formula that is completely analogous to the relation (25.8).

Finally, one remarks that as a result of reciprocity (§ **15**) the E-function can also be represented by n-rowed determinants in the original coordinates, as well as in the canonical ones.

27. The differential equations of geodetic fields. From §§ **20** and **22**, one obtains a geodetic field when one simultaneously satisfies the equations:

$$f = \Delta = c, \qquad f_{p_{i\alpha}} = p_{i\alpha} = S_{\lambda i}\bar{c}_{\lambda\alpha}, \tag{27.1}$$

with the notations of § **19**. From § **18**, however, this system of equations is completely equivalent to the following one:

$$S_{\alpha i} = P_{i\rho}\, S_{\alpha\rho}, \tag{27.2}$$

$$F \cdot |\, S_{\alpha\beta}\,| = 1. \tag{27.3}$$

If one now computes the function $F(x_i, t_\alpha, P_{i\alpha})$ then one can find a geodetic field in the following way: One determines the $P_{i\alpha}$ as rational functions of the first partial derivatives of the $S_\alpha(x_i, t_\beta)$ from equations (27.2) and substitutes the values thus found in (27.3). One then obtains *one* first order partial differential equation for μ functions S_α, $(\mu - 1)$ of which can therefore be chosen completely at will.

28. By means of an arbitrary given geodetic field, with the help of (27.2) and (27.3), the $P_{i\alpha}$ and F – and therefore, by applying the formulas of chapter I, all remaining quantities – will be determined as functions of (x_i, t_α); i.e., as functions of position in $(n + \mu)$-dimensional space.

However, conversely one can also give the $P_{i\alpha}$ as such functions of position to begin with and ask what the necessary and sufficient conditions are in order for one to find functions $S_\alpha(x_i, t_\beta)$ for which the relations (27.2) and (27.3) are valid.

We introduce the linear operator:

$$L_i = \frac{\partial}{\partial x_i} - P_{i\rho}\frac{\partial}{\partial t_\rho}. \tag{28.1}$$

Equations (27.2) then say that the system of differential equations:

$$L_i\, S_\alpha = 0 \tag{28.2}$$

for the μ independent functions S_α must be valid, and thus one must have a *Jacobi system*. The necessary and sufficient condition for this is well-known to be the vanishing of the bracket expressions $(L_iL_j - L_jL_i)S$; this is equivalent to the following relations:

$$L_j\, P_{i\rho} - L_i\, L_{j\rho} = 0. \tag{28.3}$$

If we then set, to abbreviate:

(28.4) $$\boxed{[ij\rho] = \frac{\partial P_{i\rho}}{\partial x_j} - \frac{\partial P_{j\rho}}{\partial x_i} - \left(P_{j\sigma} \frac{\partial P_{i\rho}}{\partial t_\sigma} - P_{i\sigma} \frac{\partial P_{j\rho}}{\partial t_\sigma} \right),}$$

then we must write:

(28.5) $$[ij\rho] = 0.$$

29. Let the conditions (28.5) all be verified. Between two systems S_α and T_α of any μ independent solutions of the Jacobi system (28.2) there always exists the relation:

(29.1) $$\left|S_{\alpha\beta}\right| \frac{\partial(T_1,\ldots,T_\mu)}{\partial(S_1,\ldots,S_\mu)} = |\, T_{\alpha\beta}\,|,$$

which represents a well-known property of the functional determinant. If one then gives the T_α then equation (27.3) is soluble when and only when one can determine the S_α as functions of the T_β such that the equation:

(29.2) $$\frac{\partial(T_1,\ldots,T_\mu)}{\partial(S_1,\ldots,S_\mu)} = (F)\,|\, T_{\alpha\beta}\,|$$

is satisfied. Thus, (F) means any function in the $(n + \mu)$ variables (x_i, t_α) that one obtains when one expresses the $P_{i\alpha}$ in $F(x_i, t_\alpha, P_{i\alpha})$ as functions of (x_i, t_β). Relation (29.2) can, however, be satisfied when and only when the right-hand side of this equation is itself a function of the T_β; i.e., when it satisfies the Jacobi system (28.2). Equation (27.3) is then equivalent to the system:

(29.3) $$L_i((F) + (F)\,\overline{T}_{\rho\sigma} L_i \frac{\partial T_\rho}{\partial t_\sigma} = 0.$$

Now, one has:

(29.5) $$L_i \frac{\partial T_\rho}{\partial t_\sigma} = \frac{\partial^2 T_\rho}{\partial x_i\, \partial t_\sigma} - P_{i\lambda} \frac{\partial^2 T_{\rho\sigma}}{\partial t_\sigma\, \partial t_\lambda}.$$

On the other hand, because T_ρ is, by assumption, a solution of (28.2) one has:

$$\frac{\partial^2 T_\rho}{\partial x_i\, \partial t_\sigma} = \frac{\partial}{\partial t_\sigma}\left(\frac{\partial T_\rho}{\partial x_i}\right) = \frac{\partial}{\partial t_\sigma}\left(P_{i\lambda} \frac{\partial T_\rho}{\partial t_\lambda}\right);$$

when this is substituted in (29.5) it gives the equation:

(29.6) $$L_i \frac{\partial T_\rho}{\partial t_\sigma} = \frac{\partial P_{i\lambda}}{\partial t_\sigma} T_{\rho\lambda}.$$

We substitute this value in (29.4) and obtain, after dividing by $| T_{\alpha\beta} |$, the condition that we wished to present:

(29.7) $$L_i(F) + (F)\frac{\partial P_{i\sigma}}{\partial t_\sigma} = 0.$$

30. We now set, to abbreviate:

(30.1) $$[i] = L_i(F) + (F)\frac{\partial P_{i\sigma}}{\partial t_\sigma} + \Pi_{j\rho}[ij\rho],$$

and develop $L_i(F)$, while taking (23.2) into account; we obtain:

$$[i] = \frac{\partial F}{\partial x_i} + \Pi_{j\sigma}\frac{\partial P_{j\rho}}{\partial x_i} - P_{i\sigma}\left(\frac{\partial F}{\partial t_\sigma} + \Pi_{j\rho}\frac{\partial P_{j\rho}}{\partial t_\sigma}\right) + F\frac{\partial P_{i\sigma}}{\partial t_\sigma} + \Pi_{j\sigma}[ij\rho].$$

From this, it follows, when one uses (28.4), that:

(30.2) $$[i] = \frac{\partial F}{\partial x_i} - P_{i\sigma}\frac{\partial F}{\partial t_\sigma} + \Pi_{j\sigma}\left(\frac{\partial P_{i\rho}}{\partial x_j} - P_{j\rho}\frac{\partial P_{i\rho}}{\partial t_\sigma}\right) + F\frac{\partial P_{i\sigma}}{\partial t_\sigma},$$

an equation that can be written, with our previous notation:

(30.3) $$\boxed{[i] = \frac{\partial F}{\partial x_i} - P_{i\sigma}\frac{\partial F}{\partial t_\sigma} + \Pi_{j\sigma}\frac{\partial P_{i\rho}}{\partial x_j} + A_{\sigma\rho}\frac{\partial P_{i\rho}}{\partial t_\sigma}.}$$

The necessary and sufficient conditions for the existence of a geodetic field ultimately take the form:

(30.4) $$[ij\rho] = 0, \qquad [i] = 0.$$

31. The Euler equations. It is actually not difficult to prove directly that any μ-dimensional manifold that intersects a geodetic field transversally must satisfy the Euler differential equations:

(31.1) $$\frac{d}{dt_\alpha} f_{p_{i\alpha}} - f_{x_i} = 0;$$

however, it is much more interesting and instructive to present a general identity from which this requirement will be deduced immediately.

To that end, we give the $p_{i\alpha}$ as completely arbitrary functions of (x_j, t_β) and likewise compute the remaining quantities $f(x_i, t_\alpha, p_{i\alpha})$, $\pi_{i\alpha} = f_{p_{i\alpha}}$, etc., as functions of position with the help of our previous formulas.

Furthermore, we introduce the notation $d\psi / dt_\alpha$ for any function $\psi(x_i, t_\alpha)$ that enters into (31.1), in particular, and is defined by the relation:

$$\frac{d\psi}{dt_\alpha} = \frac{\partial \psi}{\partial t_\alpha} + \frac{\partial \psi}{\partial x_i} p_{i\alpha}. \tag{31.2}$$

With these preparations, we consider the relation (13.8):

$$F p_{i\alpha} = P_{i\sigma} \bar{g}_{\sigma\alpha},$$

and deduce from this equation, by differentiation, that:

$$F\, dp_{i\alpha} = -p_{i\alpha}\, dF + \bar{g}_{\sigma\alpha} dP_{i\sigma} + P_{i\sigma} d\bar{g}_{\sigma\alpha}. \tag{31.3}$$

From a formula that is derived from equation (25.2), one now has:

$$g\, d\bar{g}_{\sigma\alpha} = (\bar{g}_{\sigma\alpha}\bar{g}_{\rho\lambda} - \bar{g}_{\rho\alpha}\bar{g}_{\sigma\lambda})\, dg_{\rho\lambda};$$

hence, from (13.4), (13.8), (13.9), and (13.2):

$$F\ f\ P_{i\alpha} d\bar{g}_{\sigma\alpha} = F(\pi_{i\alpha}\bar{g}_{\rho\lambda} - \pi_{i\lambda}\bar{g}_{\rho\alpha})(p_{j\lambda} dP_{j\rho} + P_{j\rho} dp_{j\lambda}),$$

$$f\ P_{i\sigma} d\bar{g}_{\sigma\alpha} = (f\ \pi_{i\alpha}\, \Pi_{j\rho} - \bar{g}_{\rho\alpha}\, p_{j\lambda}\, \pi_{i\lambda})\, dP_{j\rho} + F(\pi_{i\alpha}\, \pi_{j\lambda} - \pi_{i\lambda}\, \pi_{j\alpha})\, dp_{j\lambda}. \tag{31.4}$$

By substituting this relation in (31.3) we obtain:

$$F\, f\, d\pi_{i\alpha} = -f\ \pi_{i\alpha}\, dF + (f\ \pi_{i\alpha}\, \Pi_{j\rho} + \bar{g}_{\rho\alpha}\, b_{ji})\, dP_{j\rho} + F(\pi_{i\alpha}\, \pi_{j\lambda} - \pi_{i\lambda}\, \pi_{j\alpha})\, dp_{j\lambda}. \tag{31.5}$$

We now introduce the notation:

$$\Omega_i = Ff\left(\frac{d\pi_{i\alpha}}{dt_\alpha} - f_{x_i}\right) - F\pi_{i\alpha}\pi_{j\lambda}\left(\frac{dp_{j\lambda}}{dt_\alpha} - \frac{dp_{j\alpha}}{dt_\lambda}\right), \tag{31.6}$$

and obtain from (31.5) that:

$$\Omega_i = -F\ f\ f_{x_i} - f\pi_{i\alpha}\frac{d(F)}{dt_\alpha} + (f\pi_{i\alpha}\Pi_{j\rho} + \bar{g}_{\rho\alpha} b_{ji})\frac{dP_{j\rho}}{dt_\alpha}. \tag{31.7}$$

Now, from (23.1), (23.2), and (31.2), one has:

$$-F\ f\ f_{x_i} = f^2 F_{x_i} = f^2\frac{\partial(F)}{\partial x_i} - f^2 \Pi_{j\rho}\frac{\partial P_{j\rho}}{\partial x_i},$$

$$\frac{d(F)}{dt_\alpha} = \frac{\partial(F)}{\partial x_h} p_{h\alpha} + \frac{\partial(F)}{\partial t_\alpha},$$
$$\frac{dP_{j\rho}}{dt_\alpha} = \frac{\partial P_{j\rho}}{\partial x_h} p_{h\alpha} + \frac{\partial P_{j\rho}}{\partial t_\alpha}.$$

When all of this is substituted in (31.7), it gives, with some simplifications:

(31.8) $\Omega_i =$

$$f b_{ki} \frac{\partial(F)}{\partial x_k} + f(\Pi_{k\rho} b_{ji} - \Pi_{j\rho} b_{ki}) \frac{\partial P_{j\rho}}{\partial x_k} - f \pi_{i\alpha} \frac{\partial(F)}{\partial t_\alpha} + (f \pi_{i\alpha} \Pi_{j\rho} + \bar{g}_{\rho\alpha} b_{ji}) \frac{\partial P_{j\rho}}{\partial t_\alpha}.$$

With the use of (15.6) and (28.1) this can be written:

(31.9) $$\Omega_i = f\, b_{ki}\, L_k(F) - f \Pi_{j\rho} b_{ki} \left(\frac{\partial P_{j\rho}}{\partial x_k} - \frac{\partial P_{k\rho}}{\partial x_j} \right) + (f \pi_{i\alpha} \Pi_{j\rho} + \bar{g}_{\rho\alpha} b_{ji}) \frac{\partial P_{j\rho}}{\partial t_\alpha}.$$

However, from (28.4) and (30.1), one has:

$$\frac{\partial P_{j\rho}}{\partial x_k} - \frac{\partial P_{k\rho}}{\partial x_j} = [jk\rho] + P_{k\sigma} \frac{\partial P_{j\rho}}{\partial t_\sigma} - P_{j\sigma} \frac{\partial P_{k\rho}}{\partial t_\sigma},$$
$$L_k(F) = [k] + \Pi_{j\rho}\, [jk\rho] - (F) \frac{\partial P_{k\sigma}}{\partial t_\sigma}.$$

If one substitutes these quantities in (31.9) and then remarks that, from (15.6), (13.7), (12.2), and (12.1), one has:

$$b_{ki}\, P_{k\sigma} = \pi_{i\sigma},$$
$$\bar{g}_{\rho\sigma} = F\, a_{\rho\sigma} = f\, A_{\sigma\rho} = f(\delta_{\rho\sigma} F - P_{j\sigma}\, \Pi_{j\rho})^*,$$

then nearly all of the terms vanish and what remains is:

(31.10) $$\Omega_i = f\, b_{ki} [k].$$

One obtains the identity that we wish to present by equating (31.6) and (31.10); it reads:**

(31.11) $$\boxed{\frac{d\pi_{i\alpha}}{dt_\alpha} - f_{x_i} = \frac{b_{ki}}{F} [k] - \frac{\pi_{i\alpha} \pi_{j\beta}}{f} \left(\frac{dp_{j\alpha}}{dt_\beta} - \frac{dp_{j\beta}}{dt_\alpha} \right).}$$

* [A printing error in this equation in the original was corrected.]
** [A printing error in (31.11) in the original was corrected.]

32. Now, if, as we have assumed, the functions of position $p_{i\alpha}(x_j, t_\beta)$, in particular, belong to a geodetic field that intersects a μ-dimensional manifold transversally then one has at any point of this manifold:

$$[k] = 0, \qquad \frac{dp_{j\alpha}}{dt_\beta} = \frac{dp_{j\beta}}{dt_\alpha}.$$

The left-hand side of (31.11) must then vanish on this manifold, and this is an integral of the Euler equation (31.1).

We would like call the geodetic field a *distinguished field* when an extremal can be found through any point of this field that intersects this field transversally. The extremals in such a case then define a field in their own right and the figure that is defined by these extremals and the manifolds $S_\alpha = \lambda_\alpha$ is called a *complete figure* of the variational problem.

33. We consider an arbitrary family of extremals that simply cover a region of $(n + \mu)$-dimensional space. The left-hand side and the last term in the identity (31.11) must then vanish, from which it follows that all $[k] = 0$.

However, in order for the extremals of a field to generate a complete figure of the variational problem one must further have that all $[jk\rho] = 0$, which is already well-known to not always be the case when $\mu = 1$.

Observations on Hilbert's independence theorem and Born's quantization of field equations

Hermann Weyl, *Institute for Advanced Study, Princeton, New Jersey*

(Received July 6, 1934)

Born recently proved a quantization of the field equations which is based upon Hilbert's independence theorem of the calculus of variations.[1] My intention here is to give, in the first, purely mathematical, Part A, a formulation as simple and explicit as possible of the independence theorem. The agreement between the principle of variation and the independence theorem, complete in the case of one independent variable and one unknown function, fails in two respects in the case of several variables and functions; the independence theorem specializes the extremal vector field, on the one hand, and it discards the assumption of integrability, on the other hand.[2] In Part B, I first suggest a modification of Born's scheme, without which it would be in disagreement with ordinary quantum mechanics, even in the one-dimensional case. After the modification, a comparison of Heisenberg-Pauli's quantization becomes possible under the simplest circumstances. Born's scheme proves to be too narrow. Finally, I raise the principal objection that the quantum-mechanical equation should not be of the form: four-dimensional divergence of ψ equals $H\psi$ with a scalar operator of action H, but that it should rather consist of four components stating that the differentiation of ψ with respect to four space-time coordinates is performed by means of the operators: energy and momentum.[*]

A. HILBERT'S INDEPENDENCE THEOREM FOR SEVERAL ARGUMENTS

§ 1. The problem of variation. The problem of the calculus of variations in r independent variables $t^1 \dots t^r$ consists in determining ν functions, or a "surface:"

$$z^\alpha = z^\alpha(t^1 \dots t^r), \qquad (\alpha = 1, 2, \dots, \nu) \tag{1}$$

such that the variation:

$$\delta \int L(t^i, z^\alpha, z_i^\alpha)\, dt^1 \cdots dt^r = 0 \tag{2}$$

[1] Proc. Roy. Soc. A**143**, 410 (1934).

[2] These are known facts. Born himself refers to: Prange, Thesis, Göttingen 1915, but the theory was developed before Prange, and in a more general and suitable form, by Volterra (1890), Fréchet (1905), and de Donder. Cf., de Donder, Mém. Acad. Roy. de Belgique, ser. 2, III (1911); *Théorie invariative du calcul des variations*, Paris, 1930, Chaps. VII and VIII.

[*] The abstract will also serve as a summary and introduction to the paper. – EDITOR.

for arbitrary variations $\delta z^\alpha(t)$ which vanish at the border of the domain of integration. L is here a given function of the arguments t^i, z^α, z_i^α; one has to substitute the functions (1) for z^α and the derivatives dz^α / dt^i for z_i^α.

§ 2. Surface field and vector field. A family of ∞^r surfaces $z^\alpha = z^\alpha(t^1 \dots t^r)$ simply covering a piece Ω of the $(r+\nu)$ dimensional space of coordinates (t^i, z^α) may be called a *surface field* in Ω. At every point (t, z) of Ω we have the "gradient vector:"

$$dz^\alpha / dt^i = z_i^\alpha(t, z) \tag{3}$$

of the surface passing through (t, z). Conversely, if one is given the *vector field* $z_i^\alpha(t, z)$ arbitrarily, one can find a corresponding field of surfaces provided Eqs. (3) are completely integrable. As one readily sees, the necessary and sufficient conditions of integrability are the relations:

$$\left(\frac{\partial z_k^\alpha}{\partial t^i} - \frac{\partial z_i^\alpha}{\partial t^k}\right) + \left(\frac{\partial z_k^\alpha}{\partial z^\beta} z_i^\beta - \frac{\partial z_i^\alpha}{\partial z^\beta} z_k^\beta\right) = 0. \tag{4}$$

(Always sum over two-fold occurring indices!) A vector field satisfying these equations may be called *integrable*.

§ 3. Three stages of independent variables. We distinguish three standpoints:

(1) t^i, z^α, z_i^α are taken as independent variables, as for instance, in the function L. The derivatives with respect to these variables are distinguished by an attached index.

(2) By using a given vector field, the z_i^α are replaced by functions of the t^i and z^α. The partial derivatives with respect to the arguments t^i and z^α are then denoted by $\partial/\partial t^i$, $\partial/\partial z^\alpha$.

(3) The subscription:

$$z^\alpha = z^\alpha(t^1, \dots, t^r) \qquad\qquad [\, z_i^\alpha = dz^\alpha / dt^i]$$

changes functions which appeared in the second (or the first) standpoint, into functions of the t alone. The derivatives with respect to the t's are denoted by d/dt^i.

We have already complied with these conventions in paragraphs 1 and 2.

§ 4. Extremal vector field. The Lagrangian equations of the problems of variation (2) are (standpoint 3):

$$dL_{z_i^\alpha} / dt^i = L_{z^\alpha}. \tag{5}$$

A solution of these equations may be designated as an *extremal surface.* We start with a *field of extremal surfaces.* Such a field is, according to (5), characterized by the equations (standpoint 2):

$$\frac{\partial L_{z_i^\alpha}}{\partial t^i} + \frac{\partial L_{z_i^\alpha}}{\partial z^\beta} z_i^\beta = L_{z^\alpha} . \tag{6}$$

A vector field $z_i^\alpha(t, z)$ satisfying (6) is called an *extremal vector field* whether it is integrable or not.

§ 5. Legendre transformation. We introduce (standpoint 1) the momenta:

$$p_\alpha^i = L_{z_i^\alpha} \tag{7}$$

and:

$$p = L - p_\alpha^i z_i^\alpha . \tag{8}$$

From the total differential:

$$\delta L = L_{t^i} \delta t^i + L_{z^\alpha} \delta z^\alpha + L_{z_i^\alpha} \delta z_i^\alpha ,$$

there follows:

$$\delta p = L_{t^i} \delta t^i + L_{z^\alpha} \delta z^\alpha - z_i^\alpha \delta p_\alpha^i .$$

It is therefore natural to assume that p is given as a function of t^i, z^α, p_α^i:

$$p = H(t^i, z^\alpha, p_\alpha^i).$$

We then have $z_i^\alpha = -H_{p_\alpha^i}$ as the converse of the relation (7).

We now should write (standpoint 2) p_α^i, instead of $L_{z_i^\alpha}$, on the left side of (6). In order to determine the right side, one has to differentiate Eq. (8) or:

$$p = L - p_\beta^i z_i^\beta$$

with respect to z^α:

$$\frac{\partial p}{\partial z^\alpha} = L_{z^\alpha} + L_{z_i^\beta} \frac{\partial z_i^\beta}{\partial z^\alpha} - p_\beta^i \frac{\partial z_i^\beta}{\partial z^\alpha} - \frac{\partial p_\beta^i}{\partial z^\alpha} z_i^\beta .$$

Here, the second and third terms cancel each other, and consequently the equations characteristic of an extremal field read as follows:

$$\frac{\partial p_\alpha^i}{\partial t^i} + \left(\frac{\partial p_\alpha^i}{\partial z^\beta} - \frac{\partial p_\beta^i}{\partial z^\alpha} \right) z_i^\beta = \frac{\partial p}{\partial z^\alpha} ; \tag{9}$$

$$p = H(t^i, z^\alpha, p_\alpha^i), \qquad z_i^\alpha = -H_{p_\alpha^i} . \tag{10}$$

§ 6. Special extremal vector field and Jacobi-Hamilton equation. Eq. (9) is satisfied in particular if:

$$\frac{\partial p_\alpha^i}{\partial z^\beta} - \frac{\partial p_\beta^i}{\partial z^\alpha} = 0, \frac{\partial p_\alpha^i}{\partial t^i} = \frac{\partial p}{\partial z^\alpha}\,. \tag{11}$$

We then speak of a *special extremal vector field*.

One makes good these equations by putting:

$$p_\alpha^i = \partial s^i \,/\, \partial z^\alpha,$$

and by assuming that the new unknown quantities s^i fulfill the equation:

$$\partial s^i \,/\, \partial t^i = p.$$

In this way, determination of a special extremal vector field is reduced to the integration of the *Jacobi-Hamilton equation:*

$$(\mathrm{div}\, s =) \qquad \partial s^i \,/\, \partial t^i = H(t^i, z^\alpha, \partial s^i \,/\, \partial z^\alpha). \tag{12}$$

If one wants the vector field $z_i^\alpha = -H_{p_\alpha^i}$ to be integrable one has to satisfy further Eqs. (4).

§ 7. Invariance. Flux. The t^i, as well as the z^α, may be subjected to an arbitrary transformation among themselves. Let us treat z_i^α as a vector contravariant in the Greek, covariant in the Latin indices, L as a scalar density with respect to the t^i and as a scalar in z^α; p_α^i as a contravariant vector density in i¸ as a covariant vector in α, and s^i as a contravariant vector density with respect to the t^i (as a scalar with respect to z^α). Under such circumstances, all our equations remain unaltered by the transformation. Hence, the *flux* of s^i through an arbitrary $(r-1)$-dimensional "cross section" Λ of the r-dimensional t-space:

$$S = \int_\Lambda \begin{vmatrix} s^1 & \cdots & s^r \\ dt^1 & \cdots & dt^r \\ . & \cdots & . \\ \delta t^1 & \cdots & \delta t^r \end{vmatrix} \tag{13}$$

has an invariant significance. In forming (13), we operate on a surface Σ: $z^\alpha = z^\alpha(t^1 \ldots t^r)$ (standpoint 3), and Λ is to be considered as an $(r-1)$-dimensional "line" on the surface Σ.

§ 8. The independent integral. Following Gauss's theorem, one can change the flux (13) through a *closed* Λ into an integral extending over the piece of Σ bounded by Λ; its integrand:

$$\frac{ds^i}{dt^i} = \frac{\partial s^i}{\partial t^i} + \frac{\partial s^i}{\partial z^\alpha}\frac{dz^\alpha}{dt^i} = p + p_\alpha^i \frac{dz^\alpha}{dt^i}$$

contains only the original quantities p_α^i and p instead of s. This is Hilbert's "*independent integral*," for it does not change its value if one changes in an arbitrary manner the piece of the surface Σ bounded by Λ in the $(r+1)$-dimensional (t^i, z^α)-space, provided the boundary line Λ is preserved.

§ 9. Case of no forces. If L depends only on the third group of variables z_i^α, and if H consequently depends only on the p_α^i then:

$$p_\alpha^i = \text{const.}, \qquad p = H(p_\alpha^i) = \text{const.}$$

yields a special integrable extremal field. The corresponding s^i is:

$$s^i = p_\alpha^i z^\alpha + (1/r) p t^i \; . \tag{14}$$

This particular solution is, of course, not endowed any longer with the general invariance as described in section 7.

B. CRITICAL REMARKS CONCERNING BORN'S PROPOSAl OF A QUANTIZATION OF ELECTROMAGNETIC FIELD EQUATIONS

§ 10. Born's procedure. Professor Born propounds the following procedure for the transition to quantum physics. One first forms the flux S, (13) of the vector density (14) through an arbitrary closed line Λ and takes $\psi = e^{iS}$ ("plane wave"); one then builds up wave packets or a general ψ by forming linear superpositions of plane waves that correspond to several values of the constants p_α^i. Each such ψ is a function of the following arguments:

$$V = \int_\Sigma dt^1 \cdots dt^r \; , \qquad \int_\Sigma \frac{dz^\alpha}{dt^i} dt^1 \cdots dt^r \; . \tag{15}$$

Here, Σ denotes the domain in our r-dimensional t-space surrounded by the line Λ. $|\psi|^2$ should be interpreted as the probability that the integrals (15) assume given values in a domain Σ of given volume V. All domains of the t-space here, be it noticed, if they only have the same volume, are thrown into the same pot without regard to their shape and situation! This sounds queer enough, and, as a matter of fact, Born's interpretation does not coincide with the usual well-proved interpretation of quantum mechanics even in the one-dimensional case where t = time is the only variable. For there $|\psi|^2$ is the probability that the quantities z^α assume given values at the instant t, whereas Born is urged to look upon it as a probability of transition, namely the probability that the quantities z^α experience given changes $\Delta z^\alpha = \int (dz^\alpha / dt)\, dt$ in a time internal of given length Δt – irrespective of the temporal localization of the interval Δt. One obviously has to tear asunder the closed "null-dimensional line," which bounds the one-dimensional time interval and consists of two time points, into its initial and end points. We are able to imitate this procedure in r dimensions by determining the flux S through a *cross section* Λ of the t-space instead of a *closed* Λ. Let us think of the whole t-space as dissolved into a simply infinite sequence of such cross sections. In the physical applications, r is equal to 4 and t^1, t^2, t^3; $t^0 = t$ are the 4 space and time coordinates. *After choosing the planes of simultaneity* t = const. as our cross sections Λ, Born's procedure becomes somewhat comparable to the Heisenberg-Pauli quantization.

§ 11. Comparison of the Born and Heisenberg-Pauli process in the simplest case. The comparison can actually be carried out for the particular case of an L depending only on the temporal derivatives $z_0^\alpha = dz^\alpha / dt$. In this case, the fundamental ψ – the plane wave – becomes, according to Born:

$$\exp\left[i \iiint \{ \tfrac{1}{4} t H(p_\alpha^0) + p_\alpha^0 z^\alpha \} dt^1 dt^2 dt^3 \right].$$

Here, the p_α^0 are constants; the z^α are arbitrary functions of t^1, t^2, t^3. The Heisenberg-Pauli procedure yields the same result, with the difference, however, that this time the p_α^0 are arbitrary functions of the space coordinates t^1, t^2, t^3; the probability refers to the question as to which values the physical quantities z^α assume *at all possible space points.* This more general formulation, obviously not required by the nature of the problem, is not entirely beyond the scope of Born's quantization. For in the present circumstances the following p's:

$$p_\alpha^0 = \text{arbitrary functions of } t^1, t^2, t^3, \quad [\, p_\alpha^i = 0 \ (i = 1, 2, 3)]$$

furnish a special extremal vector field. But it is not integrable! Thus, one is led to *renounce the assumption of integrability.*

§ 12. Objections and hopes. Nevertheless, I am unable to see how, by means of an analogous extension of Born's scheme, the general case could be brought into agreement with the fundamental physical experience, for the characteristic commutation rules of coordinates and corresponding momenta, q and p, are missing in Born's theory, owing to the fact that he subsumes the field equations under the mechanical "problem without forces," but these commutation rules seem to be essential for the possibility of considering the electromagnetic ether as a superposition of oscillators (photons). On the other hand, I am fairly sure that the scheme of quantum physics should not be obtained from the one equation (12) in the form div = H by means of Schrödinger's quantum-mechanical transmutation, but that it should consist, rather, of four equations:

$$d/dt^i = T_i\,,$$

in which the four operators T_i represent the energy and the three components of momentum. The recipe for forming the T_i is rather complicated in the Heisenberg-Pauli theory, and the fact that they form a covariant 4-vector, in the sense of relativity theory, needs a special proof. One may, perhaps, expect that a way similar to that followed by Born will lead to an essentially simpler formulation, and perhaps a modification of this prescription, so as to put the relativistic invariance in evidence from the beginning.

Geodesic fields in the calculus of variations for multiple integrals

BY HERMANN WEYL

(Received April 15, 1935)

Introduction

Carathéodory recently drew my attention to an "independent integral" in the calculus of variation for several variables exhibited by him in an important paper in 1929,[1] and he asked me about its relation to a different independent integral I made use of in a brief exposition of the same subject in the Physical Review, 1934.[2] The present note was drafted to meet Carathéodory's question (§ 11). To facilitate comparison, I first serve my own dish again in Carathéodory style (trace theory, Part I), and then expound the essentials of his theory (Part 2); the link between them thereby becomes fairly obvious. In Part 3, I consider the approximation known as the second variation. Thus, the whole formal apparatus of the calculus of variations – Lagrange's equation, Legendre's, Jacobi's, Weierstrass's conditions, and Hilbert's independent integral – will be found in these three Parts, packed together in a nutshell, as it were. Chapter 4 solves the problem of embedding a given extremal in a geodesic slope field – this notion taken in the sense of the trace theory.[3] The reader who does not care for technical details, but wants the lucid simplicity of the general foundations not to be marred by toilsome existential considerations is warned to ignore this last Part.

Part 1. The linear trace theory.

§ 1. The problem of variation. ν functions of r variables t^i:

$$z^\alpha = z^\alpha(t^1 \dots t^r), \qquad (t^1 \dots t^r) \text{ in } G, \qquad (\alpha = 1, \dots, \nu) \tag{1}$$

describe an r-dimensional "surface" Σ in the $(r + \nu)$-dimensional t-z-space covering a given region G of the t-space. We consider only surfaces Σ lying in a certain domain Ω of the t-z-space which have their boundary in common; i.e., the values of the functions (1) at the boundary of G are prescribed once for all.

The situation in the calculus of variations with r independent variables t^i ($i = 1, \dots, r$) is this: A function L of the variables:

$$t^i, z^\alpha, z_i^\alpha \qquad (\alpha = 1, \dots, \nu;\ i = 1, \dots, r)$$

[1] Acta litt. ac. scient. universe. Hungaricae, Szeged, Sect. Math. **4** (1929), 193.

[2] Physical Review **46** (1934), 505.

[3] Prof. Carathéodory advises me that Mr. Boerner did the same for his more sophisticated theory.

is given. By an appropriate choice of Σ, one tries to *minimize the integral:*

(2) $$J = J(\Sigma) = \int_G L(t^i, z^\alpha(t), dz^\alpha / dt^i)\, dt^1 \cdots dt^r \, .$$

§ 2. Three stages of independent variables. *ν-r* functions $z_i^\alpha(t, z)$ in Ω define what we call a *slope field* $\mathfrak{F}$ in Ω. The surface Σ, (1) is *embedded* in the slope field $\mathfrak{F}$ if:

$$dz^\alpha / dt^i = z_i^\alpha(t^k, z^\beta(t))$$

holds.

We distinguish three standpoints concerning the arguments in our functions:

(1) t^i, z^α, z_i^α are taken as independent variables, as for instance, in the function *L*. The derivatives with respect to these variables are marked by attaching the respective variable as an index.

(2) By using a given slope field $\mathfrak{F}$, the z_i^α are replaced by functions of the t^i and z^a. The partial derivatives with respect to the arguments t^i and z^α are then denoted by $\partial/\partial t^i$, $\partial/\partial z^\alpha$.

(3) The substitution:

$$z^\alpha = z^\alpha(t^1, \ldots, t^r) \qquad [\, z_i^\alpha = dz^\alpha / dt^i]$$

referring to a given surface Σ changes functions which appeared in the second (or the first) standpoint into functions of the *t* alone. Their derivation with respect to t^i is denoted by d/dt^i.

In keeping with these conventions and the further one that one always has to sum over two-fold occurring indices, the vanishing of the first variation: $\delta J = 0$ is expressed by *Euler's equations:*

(3) $$dL_{z_i^\alpha} / dt^i - L_{z^\alpha} = 0.$$

Σ is called an *extremal* when satisfying these relations. The arguments in L_{z^α} and $L_{z_i^\alpha}$ are t^i, z^α, dz^α / dt^i.

§ 3. Lagrangian of the divergence type. One may form *L* for which the integral (2) is independent of Σ by the following method: Let:

(4) $$s^i(t, z) \qquad (i = 1, \ldots, r)$$

be given functions in Ω. After substituting the functions (1) for the arguments *z*, we consider the divergence:

$$\frac{ds^i}{dt^i} = \frac{\partial s^i}{\partial t^i} + \frac{\partial s^i}{\partial z^\alpha} \cdot \frac{dz^\alpha}{dt^i} \, .$$

Its integral is the flux of the vector field $s^i(t, z(t))$ through the boundary of G, and therefore depends on the values of $z^\alpha(t)$ at the border of G only. Hence:

$$D(t^i, z^\alpha, z_i^\alpha) = \frac{\partial s^i}{\partial t^i} + \frac{\partial s^i}{\partial z^\alpha} \cdot z_i^\alpha . \tag{5}$$

All surfaces Σ are extremals of this Lagrangian, which is linear in z_i^α.

§ 4. Geodesic field and independent integral. Let L be a given Lagrangian, and let us now suppose we succeeded in determining our functions (4) and the slope field $z_i^\alpha(t, z)$ such that:

$$L = D, \qquad L_{z_i^\alpha} = D_{z_i^\alpha} \qquad \text{for} \qquad z_i^\alpha = z_i^\alpha(t, z). \tag{6}$$

A slope field of this kind may be called *geodesic*. We notice, in passing, that $D_{z_i^\alpha} = \partial s^i / \partial z^\alpha$ does not contain the variables z_i^α. For the "momenta" $L_{z_i^\alpha}$, we often use the abbreviations p_α^i.

As:

$$D(t^i, z^\alpha, dz^\alpha / dt^i) = D(t^i, z^\alpha, z_i^\alpha) + \frac{\partial s^i}{\partial z^\alpha}\left(\frac{dz^\alpha}{dt^i} - z_i^\alpha\right),$$

its integral (*the independent integral*) under these circumstances changes into:

$$W = W(\Sigma) = \int_\Sigma \{L + p_\alpha^i(\dot{z}_i^\alpha - z_i^\alpha)\} dt . \tag{7}$$

A surface integral like:

$$\int_\Sigma F(t^i, z^\alpha, z_i^\alpha, \dot{z}_i^\alpha)\, dt$$

is always to be interpreted as meaning:

$$\int_G F(t^i, z^\alpha, z_i^\alpha(t, z(t), dz^\alpha / dt^i)) \cdot dt^1 \cdots dt^r .$$

The arguments of the functions L and p_α^i in (7) are t^i, z^α, z_i^α.

A surface Σ embedded in our geodesic slope field is, of necessity, an extremal for the Lagrangian L. Indeed, on account of:

$$\frac{\partial L}{\partial z^\alpha} = L_{z^\alpha} + L_{z_i^\beta} \frac{\partial z_i^\beta}{\partial z^\alpha} \quad \text{or} \quad L_{z^\alpha} = \frac{\partial L}{\partial z^\alpha} - L_{z_i^\beta} \frac{\partial z_i^\beta}{\partial z^\alpha}$$

one can supplement equations (6) by:

$$L_{z^\alpha} = D_{z^\alpha} \qquad \text{for} \qquad z_i^\alpha = z_i^\alpha\,(t, z).$$

Hence, the identity:

$$dD_{z^\alpha} / dt^i - D_{z_i^\alpha} = 0,$$

which is satisfied for *every* surface, leads, for a Σ embedded in our geodesic field, to (3).

In the case $r = 1$, $\nu = 1$, the independent integral (7) was first propounded by Hilbert.

§ 5. Legendre transformation. The equations (6) with the definition (5) of D are equivalent to:

(8) $$\frac{\partial s^i}{\partial z^\alpha} = p_\alpha^i, \qquad \frac{\partial s^i}{\partial t^i} = L - p_\alpha^i z_i^\alpha .$$

We therefore have to introduce into the function:

$$H = L - p_\alpha^i z_i^\alpha$$

the momenta:

(9) $$p_\alpha^i = L_{z_i^\alpha},$$

instead of the z_i^α, as independent variables:

(10) $$H = H(t^i, z^\alpha, p_\alpha^i)$$

(*Legendre's transformation*). The total differential:

$$\delta L = L_{t^i}\delta t^i + L_{z^\alpha}\delta z^\alpha + p_\alpha^i \delta z_i^\alpha$$

leads at once to:

(11) $$dH = L_{t^i}\delta t^i + L_{z^\alpha}\delta z^\alpha - z_i^\alpha \delta p_\alpha^i ;$$

thus, one gets:

$$z_i^\alpha = -H_{p_\alpha^i}$$

as the converse of the equations (9). *In order to construct a geodesic field, one has to solve the one Jacobi-Hamilton equation:*

(12) $$\frac{\partial s^i}{\partial t^i} = H(t^i, z^\alpha, \partial s^i / \partial z^\alpha);$$

the formula:

$$z_i^\alpha = -H_{p_\alpha^i}(t, z, \partial s / \partial z)$$

then furnishes the geodesic field.

By the way, equation (12) can be formulated in such a manner that it does not involve any derivatives with respect to the t's; the integral of $H(t^i, z^\alpha, \partial s^i / \partial z^\alpha)$ over an arbitrary part V of the region G in t-space is equal to the flux of the vector-field s^i through the boundary of V.[1]

§ 6. Weierstrass's formula. A surface Σ embedded in our geodesic field $\mathfrak{F}$ is extremal, and the integral $W(\Sigma)$, (7) coincides with $J(\Sigma)$ for this surface.

Let us then suppose we have an extremal Σ_0:

$$z^\alpha = \dot{z}^\alpha(t^1 \cdots t^r), \qquad (t^1 \ldots t^r) \text{ in } G, \tag{13}$$

lying in a region Ω of t-z-space and embedded in a geodesic field $\mathfrak{F}$ that covers Ω. We compare Σ_0 with other surfaces Σ, (1) in Ω of the same boundary. Under the notations:

$$J(\Sigma) = J, \qquad J(\Sigma_0) = J_0\,; \qquad W(\Sigma) = W, \qquad W(\Sigma_0) - W_0\,; \quad \Delta J = J - J_0$$

we have:

$$\Delta J = \Delta(J - W) = (J - W) - (J_0 - W_0),$$

because of the independence of W, and furthermore $J_0 = W_0$, because of the embedding of Σ_0 in $\mathfrak{F}$. In this simple fashion we arrive at *Weierstrass's formula:*

$$\Delta J = J - W = \int_\Sigma E(t^i, z^\alpha; z_i^\alpha, \dot{z}_i^\alpha)\, dt, \tag{14}$$

$$E(t^i, z^\alpha; z_i^\alpha, \dot{z}_i^\alpha) = [L(t^i, z^\alpha, \dot{z}_i^\alpha) - L(t^i, z^\alpha, z_i^\alpha)] - L_{z_i^\alpha}(\dot{z}_i^\alpha - z_i^\alpha), \tag{15}$$

the clue to which is the fact that the difference $J - J_0$ is expressed by a single integral *extending over* Σ; Σ_0 has been mysteriously juggled out. In (15), $L_{z_i^\alpha}$ depends on the arguments t^i, z^α, z_i^α.

One may say that the method consists in replacing L by $L - D$, subtracting a suitable D of the type (5) from L; this process does not change the extremals of L. The "suitable" choice of D is effected by solving the Jacobi-Hamilton equation (12).

Sufficient for a ("strong") minimum is the positive-definite character of Weierstrass's E-function:

$$E(t^i, z^\alpha, z_i^\alpha, \dot{z}_i^\alpha) \geq 0\,. \tag{16}$$

Here, the $\dot{z}_i^\alpha$ range independently over all values from $-\infty$ to $+\infty$; $z_i^\alpha = z_i^\alpha(t, z)$ are the slope functions of the embedding geodesic field, and the point (t, z) varies in a region Ω surrounding the extremal Σ_0 in the t-z-space. *The existence of such a field is an integral part of Weierstrass's criterion.*

[1] With such limitations as to the spread of V, of course, as are necessary for the statement to make sense: $V = V_z$ has to be such for a given point (z) that all points (t, z) lie in Ω when (t) lies in V.

§ 7. Invariance. The t^i may be subjected to an arbitrary transformation among themselves. We might even replace the region G of the t-space by an arbitrary r-dimensional manifold G, only parts of which can be referred to coordinates $t^1, \ldots, t^r$. The r quantities z_i^α ($i = 1, \ldots, r$) are to be treated as components of a covariant vector (with respect to the Latin indices, matched with the variables t^i). The Lagrangian L is to be transformed as a scalar density (of weight 1); i.e., it is to be multiplied with the absolute value of the functional determinant of the transformation of the t. The integral $J(\Sigma)$ then has an invariant significance – even when the whole G is not coverable by a single coordinate system t. Covariance and contravariance are designated by the position of the indices in the usual way. Some of the quantities – in particular, L, s^i, p_α^i, H, E – are densities in the sense just described; I would have denoted them by German letters, in accordance with the usage in my book "Raum, Zeit, Materie," had I not to reckon with the Anglo-Saxon aversion to these types.

It is conceptually simpler to take as the realm of integration G the whole manifold, not a finite portion (= compact subset) thereof. We then must replace the boundary condition for Σ by the requirement that S coincides with the standard extremal Σ_0 outside a sufficiently large *finite* portion of G (depending on Σ). Under these circumstances, the difference ΔJ, as its integrand vanishes outside that finite region, has a meaning (though not the integral $J(\Sigma)$ itself.).

At a higher standpoint of invariance, the dependent variables z^α may be included in the transformations. But, in contrast to the t^i, they should not be looked upon as a separate set in the row of $r + \nu$ variables $t^1, \ldots, t^r, z^1, \ldots, z^\nu$; we have the case of "reduction," not of "decomposition." The situation prevailing can be described in this way: An $(r + \nu)$-dimensional manifold Ω is mapped upon the r-dimensional manifold G; this mapping – called the projection – is given once for all. Thus, G may be considered as the manifold arising from Ω by identifying points ω in Ω with the same projection t. The coordinates $t^1, \ldots, t^r, z^1, \ldots, z^\nu$ covering a part of Ω are subject to the restriction that the coordinates $t^1, \ldots, t^r$ have the same values at points ω with the same projection, but all transformations in agreement with this requirement are admissible. Σ is a mapping of G in Ω: $t \to \omega$, such that the image ω has t as its projection. The behavior of all our quantities could be easily discussed under this wider aspect of invariance, but I do not wish to dwell upon it here.

Part 2. Carathéodory's determinant theory and its relation to the trace theory.

§ 8. Lagrangian of the determinant type. Carathéodory uses a different independent integral. He, too, starts with r functions:

$$S^i(t, z), \tag{17}$$

from which he forms, with reference to a given surface Σ, (1), instead of the divergence (5), the functional determinant:

(18) $$\left|\frac{dS^i(t, z(t))}{dt^k}\right|, \qquad \frac{dS^i}{dt^k} = \frac{\partial S^i}{\partial t^k} + \frac{\partial S^i}{\partial z^\alpha} \cdot \frac{dz^\alpha}{dt^k}.$$

Its integral over G is independent of Σ, as long as the boundary of Σ is preserved, for it gives the volume in the λ-space upon which the region G in the t-space is mapped by the transformation:

$$S^i(t, z(t)) = \lambda^i.$$

In accordance with the formation (18), we now take:

(19) $$D(z_i^\alpha) = \left|\frac{\partial S^i}{\partial t^k} + \frac{\partial S^i}{\partial z^\alpha} \cdot z_k^\alpha\right|.$$

This Lagrangian, too, has the property of possessing all surfaces as its extremals. Since D is not linear in the arguments z_i^α – the only one we put in evidence – one needs a little algebraic computation to compare $D(z_i^\alpha)$ for two sets of values z_i^α: $D(z_i^\alpha) = D$ and $D(\dot{z}_i^\alpha)$.

§ 9. An algebraic identity. $D = D(z_i^\alpha)$ is the determinant of certain quantities of the form:

$$S_k^i = s_k^i + \sigma_\alpha^i z_k^\alpha.$$

The element $s_k^i + \sigma_\alpha^i \dot{z}_k^\alpha$ of the second determinant $D(\dot{z}_i^\alpha)$ can be written as:

$$S_k^i + \sigma_\alpha^i u_k^\alpha,$$

the u_i^α being the differences $\dot{z}_i^\alpha - z_i^\alpha$. Application of the multiplication theorem of determinants readily leads to the formula:

$$|S_k^i + \sigma_\alpha^i u_k^\alpha| = |S_k^i| \cdot |\delta_k^i + \pi_\alpha^i u_k^\alpha|,$$

where the π_α^i are determined by the equations:

(20) $$S_r^i \pi_\alpha^r = \sigma_\alpha^i.$$

I maintain that:

(21) $$\pi_\alpha^i = D_{z_r^\alpha} / D.$$

Indeed, let $\|T_k^i\|$ be the inverse matrix of $\|S_k^i\|$. The general formula:

$$dD / D = T_r^k dS_k^r,$$

when applied to derivatives with respect to z_i^α, yields:

$$D_{z_k^\alpha} / D = T_r^k \sigma_\alpha^r ,$$

and this shows exactly that (21) are the solutions of the equations (20). Hence, the following identity obtains:

$$D(\dot z_i^\alpha) = D(z_i^\alpha) \cdot \left| \delta_k^i + (D_{z_i^\alpha} / D)(\dot z_k^\alpha - z_k^\alpha) \right| . \tag{22}$$

§ 10. Geodesic field and independent integral once more. When the functions:

$$S^i(t, z), \qquad z_i^\alpha(t, z)$$

are such that:

$$L = D, \qquad L_{z_i^\alpha} = D_{z_i^\alpha} \qquad \text{for} \qquad z_i^\alpha = z_i^\alpha(t, z),$$

Carathéodory calls the slope field $z_i^\alpha(t, z)$ *geodesic*. In a geodesic field, (22) changes into:

$$D(\dot z_i^\alpha) = L \cdot \left| \delta_k^i + (L_{z_i^\alpha} / L)(\dot z_k^\alpha - z_k^\alpha) \right| .$$

The arguments of L and $L_{z_i^\alpha} = p_\alpha^i$ are here: t^i, z^α, $z_i^\alpha(t, z)$. The *independent integral* takes on the form:

$$W(\Sigma) = \int_\Sigma L \cdot \left| \delta_k^i + (L_{z_i^\alpha} / L)(\dot z_k^\alpha - z_k^\alpha) \right| \cdot dt .$$

All further developments follow the same line as Part 1.

The differential equations, though, that are imposed on S^i by the requirement that the slope field $z_i^\alpha(t, z)$ be geodesic are essentially more complicated. The role of the Hamiltonian H in Part 1 is taken over by the determinant:

$$L \cdot \left| \delta_k^i - \frac{1}{L} p_\alpha^i z_k^\alpha \right| .$$

The theory will work only if this function, as well as L, is of constant sign in the region to be considered.

§ 11. Mutual relationship of the two independent integrals. The relation between the two competing theories of Parts 1 and 2, which serve the same end, is now fairly obvious; they do not differ in the case of only one variable t. In the general case, the extremals for the Lagrangian L are the same as for $L^* = 1 + \varepsilon L$, ε being an arbitrary constant. Notwithstanding, Carathéodory's theory is not linear with respect to L, but applying it to $1 + \varepsilon L$, instead of L, and then letting ε tend to zero, we fall back on the linearity of Part 1. One has to choose Carathéodory's functions:

$$S^i(t, z) = t^i + \varepsilon \cdot s^i(t, z).$$

Neglecting quantities that tend to zero with ε more strongly than ε itself, one then gets:

$$\left|\frac{dS^i}{dt^k}\right| = 1 + \varepsilon \cdot \frac{ds^i}{dt^i},$$

or Carathéodory's D^*, (19) becomes $= 1 + \varepsilon D$, where D has the significance (5) of Part 1. One may therefore describe Carathéodory's theory as a *finite determinant theory*, and the simpler one of Part 1 as the corresponding *infinitesimal trace theory*.

The Carathéodory theory is invariant when the S^i are considered as scalars not affected by the transformations of t. It appears unsatisfactory that the transition here sketched, by introducing the density 1 relative to the coordinates t^i, breaks the invariant character. This, however, is related to the existence of a distinguished system of coordinates t^i in the determinant theory, consisting of the functions $S^i(t, \dot{z}(t))$. This remark reveals, at the same time, that, in contrast to the trace theory, it is not capable of being carried through without singularities on a manifold G that cannot be covered by a single coordinate system t.

§ 12. Special extremal slope fields. Returning, for the rest of the paper, to the theory of Part 1, we keep to the definitions and notations explained there. In my article in the Physical Review, I viewed the problem from a slightly different angle. One is accustomed, in the classical case of one variable t and one unknown z, to perform the embedding by means of a *field of extremals*. I therefore started with a field of extremal surfaces simply covering Ω, and I introduced the gradient:

$$dz^\alpha / dt^i = z_i^\alpha(t, z) \tag{23}$$

of the field surface passing through (t, z). Such a gradient field of extremals is, according to (3), characterized by the relations:

$$\left(\frac{\partial L_{z_i^\alpha}}{\partial t^i} + \frac{\partial L_{z_i^\alpha}}{\partial z^\beta} \cdot z_i^\beta\right) - L_{z^\alpha} = 0. \tag{24}$$

Conversely, if one is given the slope field $z_i^\alpha(t, z)$ arbitrarily, one can find a corresponding field of surfaces provided equations (23) are *completely integrable*, the conditions of integrability being:

$$\left(\frac{\partial z_k^\alpha}{\partial t^i} - \frac{\partial z_i^\alpha}{\partial t^k}\right) + \left(\frac{\partial z_k^\alpha}{\partial z^\beta} \cdot z_i^\beta - \frac{\partial z_i^\alpha}{\partial z^\beta} \cdot z_k^\beta\right) = 0.$$

I proposed to call a slope field $z_i^\alpha(t,z)$ satisfying the equations (24) an *extremal slope field* whether it be integrable or not.

With respect to a Lagrangian D of the special form (5), not only is every surface an extremal, but every slope field is an extremal field. This is an immediate consequence of the fact that D is linear in z_i^α, as we shall see at once. Therefore, our geodesic field must needs be an extremal field for L. On account of $p_\alpha^i = \partial s^i / \partial z^\alpha$, it satisfies the conditions:

$$\frac{\partial p_\alpha^i}{\partial z^\beta} - \frac{\partial p_\beta^i}{\partial z^\alpha} = 0. \tag{25}$$

For this reason, I conceived the geodesic fields in the Physical review as "*special extremal slope fields,*" and thus the essential modification imposed upon the classical concept of an extremal field appeared as dropping off integrability and replacing it by the new conditions (25). For Carathéodory's D, however, it is not true at all that every slope field is extremal – notwithstanding the fact that all surfaces are extremals of D. This robs the notion of a special extremal field of its primary importance for our present purpose.

In order to justify our assertion that the left side of (24) vanishes identically for $L = D$, (5), one merely needs to observe that it does not contain the derivatives of $z_i^\alpha(t,z)$, since $D_{z_i^\alpha} = \partial s^i / \partial z^\alpha$ does not contain the variables z_i^α. A surface $z^\alpha(t)$ may be chosen such that $z^\alpha(t)$, dz^α / dt^i have arbitrarily given values at one specific point t. Hence, our statement is evident from the fact that every surface is extremal for D. He who is not afraid of a simple calculation could verify the averred identical vanishing at once.

Part 3. Second variation

§ 13. Legendre's quadratic form. Let us consider the Weierstrass E-function for definite values of t^i, z^α, z_i^α, and expand it into a power series in terms of the variables $u_i^\alpha = \dot{z}_i^\alpha - z_i^\alpha$. The expression (15) shows that the constant and linear terms are missing, and the development starts with the quadratic terms:

$$\tfrac{1}{2} L_{z_i^\alpha z_k^\beta} u_\alpha^i u_\beta^k = \tfrac{1}{2} F(t, z^\alpha, z_i^\alpha \mid u)\,. \tag{26}$$

It should not go unnoticed that the discriminant of this quadratic form in the u's is that determinant whose non-vanishing makes possible the solving of the equations (9) for z_i^α. Our form F, when taken on Σ_0, i.e., for:

$$z^\alpha = \dot{z}^\alpha(t)\,, \qquad z_i^\alpha = d\dot{z}^\alpha / dt^i\,,$$

may be designated by $F_0(t \mid u)$. The positive definite character of the quadratic form $F_0(t, u)$ – for every (t) in G – is, as is seen from this whole development, a sufficient condition for a "*weak minimum*" (*Legendre's condition*).

Whereas Weierstrass's condition refers explicitly to an embedding geodesic field, Legendre's condition does not. Does it therefore guarantee a weak minimum without assuming the existence of an embedding geodesic field? No, that is exactly where Legendre was wrong. But, only the *approximate* geodesic field (Jacobi's condition) enters into the proof of Legendre's criterion – approximate to the same degree as (26) approximates the E-function. Legendre's stunt of subtracting a divergence:

$$\frac{d}{dt^i}(s^i_{\alpha\beta}\delta z^\alpha \delta z^\beta)$$

from the integrand of the second variation $\delta^2 J$ is exactly the same procedure for that infinitesimal variation as the Weierstrass-Hilbert-Carathéodory method of subtracting a D from L with respect to the finite "variation" ΔJ.

§ 14. Trivial preparations for solving the problem of embedding. This coincidence will become clearer when we now attack the problem of embedding a given extremal:

$$\Sigma_0 : z^\alpha = \dot{z}^\alpha(t^1 \cdots t^r)$$

in a geodesic slope field. We have to construct a solution s^i of (12) such that $\partial s^i / \partial z^\alpha$ reduces to $\dot{p}^i_\alpha(t)$ for $z^\alpha = \dot{z}^\alpha(t)$. Here, let $\dot{p}^i_\alpha(t)$ be the value of $p^i_\alpha = L_{z^\alpha_i}$ for $z^\alpha = \dot{z}^\alpha(t)$, $z^\alpha_i = d\dot{z}^\alpha / dt^i$, so that we have, conversely:

$$d\dot{z}^\alpha / dt^i = -H_{p^i_\alpha}(t, \dot{z}(t), \dot{p}(t)).$$

Σ_0 being an extremal, the equation:

$$d\dot{p}^i_\alpha / dt^i = H_{z^\alpha}(t, \dot{z}(t), \dot{p}(t))$$

obtains [observe that $H_{z^\alpha} = L_{z^\alpha}$, because of (11)]. We rid ourselves of the constant and linear terms in s^i and H in the following simple way:

Writing $\dot{z}^\alpha(t) + z^\alpha$, $\dot{p}^i_\alpha(t) + p^i_\alpha$, instead of z^α and p^i_α, we put:

$$s^i(t, \dot{z}(t) + z) - s^i(t, \dot{z}(t)) = \dot{p}^i_\alpha(t) z^\alpha + \sigma^i(t, z),$$

$$H(t, \dot{z}(t) + z, \dot{p}(t) + p) - H(t, \dot{z}(t), \dot{p}(t)) = \frac{d\dot{p}^i_\alpha}{dt^i} z^\alpha - \frac{d\dot{z}^\alpha}{dt^i} p^i_\alpha + H^*(t, z, p).$$

The differential equation (12) now changes into:

$$\frac{\partial \sigma^i}{\partial t^i} = H^*\left(t^i, z^\alpha, \frac{\partial \sigma^i}{\partial z^\alpha}\right),$$

and the initial conditions:

$$\frac{\partial s^i}{\partial z^\alpha} = \dot{p}^i_\alpha(t) \quad \text{for } z = \dot{z}(t)$$

into:

$$\frac{\partial \sigma^i}{\partial z^\alpha} = 0 \qquad \text{for} \qquad z^1 = \ldots = z^r = 0.$$

The Taylor expansions of $\sigma^i(t, z)$ and $H^*(t, z, p)$ in terms of z or z, p, respectively, contain no constant and linear terms. Restoring our original notations s and H, instead of σ and H^*, we thus have shown that we may put, without any loss of generality: $\dot{z}^\alpha(t) = 0$, $\dot{p}^i_\alpha(t) = 0$.

§ 15. First approximation: Legendre's differential equations. When limiting H to its quadratic term:

(27) $$H_2 = \tfrac{1}{2} A_{\alpha\beta} z^\alpha z^\beta + A^\alpha_{i\beta} p^i_\alpha z^\beta + \tfrac{1}{2} A^{\alpha\beta}_{ik} p^i_\alpha p^k_\beta,$$

the quadratic part of s^i:

(28) $$\tfrac{1}{2} s^i_{\alpha\beta} z^\alpha z^\beta$$

provides an exact solution of the Jacobi-Hamilton differential equation. The coefficients A and $s^i_{\alpha\beta}$ are functions of t only and are, of course, written in symmetrical fashion:

$$A_{\alpha\beta} = A_{\beta\alpha}, \qquad A^{\alpha\beta}_{ik} = A^{\beta\alpha}_{ki}, \qquad s^i_{\alpha\beta} = s^i_{\beta\alpha}.$$

(12) yields the following system of differential equations for the unknown $s^i_{\alpha\beta}$**:**

(29) $$\frac{ds^i_{\alpha\beta}}{dt^i} = A_{\alpha\beta} + A^\rho_{i\beta} s^i_{\rho\alpha} + A^{\rho\sigma}_{ik} s^i_{\rho\alpha} s^k_{\sigma\beta}.$$

The transformation character is indicated again by the position of the indices. It should be added that the three A's on the right side are densities of weight +1, 0, −1, respectively. A solution of these Legendre equations furnishes what may properly be called an approximate geodesic field. (Legendre's method, as he applied it to the second variation, would lead exactly to the same result.)

Whereas Legendre's condition is only a part of the much stronger Weierstrass condition, it is to be guessed that the existence of a geodesic field, in the approximate sense of the "second variation," implies its existence in the exact sense. Our conjecture will be proved in the last Chapter. The result is twofold:

1) The embedding of Σ_0 by a geodesic slope field is always *locally* possible. This sufficies for answering all questions about *local minima* (when only surfaces Σ are admitted to competition that differ from Σ_0 in a small enough neighborhood of a point t).

2) The embedding goes through, even in the large, for the whole extremal Σ_0, *provided the first approximation – viz., the solution of Legendre's equations – can be effected.*

§ 16. Appendix: Necessary Local Conditions. Let us consider the extremal $\Sigma_0 : z^\alpha = 0$ in the neighborhood of a given point $t^i = t_0^i$, and denote the E-function at that point of Σ_0 – namely, $E(t_0, 0; 0, u_i^\alpha)$ – by $E_0(u_i^\alpha)$. One gets a necessary local condition for a strong minimum by putting a little cone-shaped hood on Σ_0. Its basis may be defined by $f(\tau^1 \dots \tau^r) \leq 1$ in terms of the relative coordinates $\tau^i : t^i = t_0^i + \varepsilon\tau^i$; here, ε is a positive constant doomed to approach zero and f is a ray function, i.e., a positive homogeneous function of degree 1:

$$f(\lambda\tau^1, ..., \lambda\tau^r) = \lambda \cdot f(\tau^1, ..., \tau^r) \qquad (\text{for } \lambda \geq 0);$$

$$f(\tau^1, ..., \tau^r) > 0 \qquad \text{except for } (\tau^1, ..., \tau^r) = (0, \dots, 0).$$

In terms of further arbitrary constants v^α, the varied surface Σ itself – the "hood" – is described by:

$$z^\alpha = \varepsilon v^\alpha \{1 - f(\tau^1, ..., \tau^r)\} \text{ for } f(\tau^1, ..., \tau^r) \leq 1,$$

$$= 0 \text{ outside this region.}$$

The inequality $\Delta J \geq 0$ with the expression (14) for ΔJ and with $\varepsilon \to 0$ leads to:

[1] $$\underset{f(\tau)\leq 1}{\mathfrak{M}} \{E_0(u_i^\alpha = v^\alpha f_i(\tau))\} \geq 0 .$$

$f_i(\tau)$ are the derivatives $df / d\tau^i$, $\mathfrak{M}$ is the integral extending over the domain $f(\tau^1, ..., \tau^r) \leq 1$ in τ-space, which that should now be looked upon as the affine "tangent space" of the r-dimensional manifold G in (t_0); the left side of [1] is invariant in this sense. As the $f_i(\tau)$ – the components of the normal vector – are homogeneous of order zero, the integral may equally well be interpreted as an average over the "sphere" of all directions in τ-space.

One can show, by specializing the function f in an appropriate manner, that not only the integral [1] but every element of it must be ≥ 0. We choose a positive constant k and put:

[2] $$\begin{aligned} f(\tau^1\ \tau^2 \dots \tau^r) &= \max(|\tau^1|, k|\tau^2|, \dots, k|\tau^r|) \text{ for } \tau^1 \geq 0, \\ &= \max(k|\tau^1|, k|\tau^2|, \dots, k|\tau^r|) \text{ for } \tau^1 \leq 0. \end{aligned}$$

Afterwards, we let k in [1] tend to zero. The volume of the negative half $\tau^1 \leq 0$ of the region $f(\tau) \leq 1$ equals $2^{r-1}/k^r$, whereas the volume of the positive part $\tau^1 \geq 0$ equals $2^{r-1}/k^{r-1}$. Let us write, for a moment:

[3] $$(1, 0, \dots, 0) = (u_1, u_2, \dots, u_r).$$

f_i is of order k in the negative half, whereas it differs from u_i by quantities of the same order in the positive half of our region. Considering the fact that $E(u_i^\alpha)$ for arguments u_i^α of the order of magnitude of k is $= O(k^2) = o(k)$ one finds for the left side of [1], after multiplication by $(k/2)^{r-1}$, an expression:

$$E_0(v^\alpha u_i) + \frac{1}{k}o(k),$$

and consequently one arrives with $k \to 0$ at:

$$E_0(v^\alpha u_i) \geq 0. \tag{[4]}$$

The particular covariant vector [3] may be here replaced by an arbitrary one. The result, formerly obtained in a slightly different manner by McShane,[1] is the following:

Necessary local condition for a strong minimum: Unless [4] *holds for arbitrary values* v^α, u_i *at any point* (t_0) *of* G, *the surface* Σ_0 *cannot have the minimizing property.*

An immediate consequence is the similar:

Necessary local condition for a weak minimum: The quadratic form $F_0(t \mid u_i^\alpha)$ *must be* ≥ 0 *for such values of the variables* u_i^α *that nullify all the quadratic forms* $u_i^\alpha u_k^\beta - u_k^\alpha u_i^\beta$.

In the general case $r > 1$, $\nu > 1$, there yawns a wide gap between the necessary and sufficient conditions; unfortunately, it seems not likely that one will be able to set up a more complete set of local necessary conditions that are comparable in simplicity to McShane's inequalities [4].

Part 4. Construction of Geodesic Fields

§ 17. Cylindrical domains and fields. For the purpose of the local problem, G can be assumed to be a cube. We shall solve the problem in the large for *cylindrical regions* G, i.e., for regions G which are the product of an $(r-1)$-dimensional manifold G^* and the open one-dimensional continuum – such that the points P of G appear as pairs (P^*, t) consisting of an arbitrary point P^* of G^* and an arbitrary number t. G^* may be referred (locally) to coordinates $t^2, \ldots, t^r$, and t may be used as the coordinate t^1. Since the Hamilton-Jacobi equation (12) – preferably in its undifferentiated form as stated at the end of § 5 – is invariant under topological transformations, our method yields a solution for all manifolds topologically equivalent to a cylinder. The complete intrinsic topological characterization of the 'cylinders" is not yet known, but we certainly get a

[1] Annals of Math. **32** (1931), 578.

fairly general picture of the situation in the large, even though we have to make this restriction of a topological nature. Its necessity shows, however, that our mode of approach is not quite adequate. Every "cell," as for instance, a convex region in ordinary $(t^1, \ldots, t^r)$-space, is, of course, a cylinder.

We start out to construct in our cylindrical manifold G a solution s^i for which all components $s^2, \ldots, s^r$ except s^1 vanish identically. Writing t, s, instead of t^1 and s^1, and dropping the upper index 1 where it appears with a similar meaning, we reduce (12) to the partial differential equation with only one unknown s:

$$\frac{\partial s}{\partial t} = H(t, z^\alpha, p_\alpha), \qquad p_\alpha = \frac{\partial s}{\partial z^\alpha} \qquad (-\infty < t < +\infty) \tag{30}$$

The coordinates $t^2, \ldots, t^r$ play now merely the role of accessory parameters. We have:

$$H = 0, \qquad H_{z^\alpha} = 0, \qquad H_{p_\alpha} = 0 \quad \text{for} \qquad z = 0, \quad p = 0 \tag{31}$$

(i.e., for $z^1 = \ldots = z^\nu = 0$, $p_1 = \ldots = p_\nu = 0$), and our aim is to find a solution $s(t, z^\alpha)$ making:

$$s = 0, \qquad \frac{\partial s}{\partial z^\alpha} = 0 \qquad \text{for } z = 0. \tag{32}$$

One can get at the partial differential equation (30) with two different tools: either with the *theory of characteristics*, or, following Cauchy, by *power series and their dominants*. Let us first go the former way.

§ 18. The characteristic equations. The differential equations for the characteristics of (30) read as follows:

$$\begin{cases} \dfrac{dz^\alpha}{dt} = -H_{p_\alpha}(t, z^\beta, p_\beta), \\[2mm] \dfrac{dp_\alpha}{dt} = H_{z^\alpha}(t, z^\beta, p_\beta). \end{cases} \tag{33}$$

When one is called upon to determine that solution $s(t, z^\alpha)$ of (30) which satisfies the initial conditions (32), one has to has to proceed in the following manner: One integrates (33):

$$z^\alpha = \zeta^\alpha(t;\ z_0^\beta), \qquad p_\alpha = \pi_\alpha(t, z_0^\beta), \tag{34}$$

with the initial values:

$$\zeta^\alpha(0;\ z_0^\beta) = z_0^\alpha, \qquad \pi_\alpha(0, z_0^\beta) = 0,$$

and the further equation:

$$\frac{ds}{dt} = -\sum_\alpha p_\alpha \cdot H_{p_\alpha}$$

by quadrature:

$$s = \sigma(t;\ z_0^\alpha) = \int_0^1 \left(\pi_\alpha \frac{d\zeta^\alpha}{dt} \right) dt\,. \tag{35}$$

One then must express the initial values z_0^α by means of the s^α themselves in solving the equations:

$$z^\alpha = \zeta^\alpha(t;\ z_0^\beta), \tag{36}$$

and in doing so one changes the quantities π_α, (34), and σ, (35), into functions p_α and s of (t, z^α). They satisfy all the relations (30).

The solution of the ordinary differential equations (33) is possible in the neighborhood of $t = 0$ for sufficiently small initial values z_0^α. Furthermore, the desired inversion of the functions (36) near $t = 0$, $z^\alpha = 0$ is possible since the functional determinant:

$$\left| \frac{\partial \zeta^\alpha}{\partial z_0^\beta} \right| \text{ equals 1} \qquad \text{for} \qquad t = 0. \tag{37}$$

This remark settles the *local* question.

§ 19. The characteristics in the large. *The first step goes through in the large, too.* That is to say: to a finite interval $-a \le t \le a$ arbitrarily given, one may assign a positive constant ε such that (33) is solvable throughout that whole interval, provided all the initial values z_0^α are of modulus less than ε. Let us briefly repeat the well-know proof.

Our differential equations (33) are of the type:

$$\frac{dx_i}{dt} = f_i(t;\ x_1,\ \ldots,\ x_n) \qquad (i = 1,\ \ldots,\ n)$$

where $f_i(t;\ 0,\ \ldots,0) = 0$. Combined with the initial conditions $x_i = x_i^0$ for $t = 0$, one replaces them by the integral equations:

$$x_i(t) = x_i^0 + \int_0^t f_i(t; x(t))\, dt$$

and determines successive approximations $x^{(0)}$, x', x'', … recursively according to:

$$x^{(h+1)}(t) = x^0 + \int_0^t f(t; x^{(h)}(t))\, dt \qquad [x^{(0)}(t) = x^0]. \tag{38}$$

Using the abbreviation $|\ x\ |$ for the largest of the n moduli $|\ x_1\ |$, …, $|\ x_n\ |$, and supposing the functions f_i to satisfy the Lipschitz inequality:

$$|\ f(t; x) - f(t; y)\ | \le M\ |\ x - y\ | \qquad (-\ a \le t \le a),$$

as long as $| x | \leq A$, $| y | \leq A$, one sees from (38) that the sequence of the successive approximation is majorized by the partial sums of the series:

$$\varepsilon \sum_{h=0}^{\infty} \frac{1}{h!}(Mt)^h = \varepsilon \cdot e^{Mt} \qquad (t \geq 0),$$

and that one is allowed to go one step further in this development as long as the preceding approximations keep within the range $| x | \leq A$. It is supposed that the initial values x^0 satisfy the inequality $| x^0 | \leq \varepsilon$. The first step is all right because the integrand in:

$$x' - x^0 = \int_0^t f(t; x^0)\, dt$$

can be replaced by the difference $f(t; x^0) - f(t; 0)$ of modulus less than εM. Hence, the whole estimation is legitimate, and the approximations converge to a solution x for which:

$$| x(t) | \leq \varepsilon \cdot e^{M|t|} \qquad (-a \leq t \leq a),$$

when ε is taken as $A \cdot e^{-Ma}$.

Notwithstanding the solubility of the characteristic equations (33) thus proved, the construction in the large of the embedding geodesic field might fail in the second step, because the functional determinant:

$$\left| \frac{\partial \zeta^\alpha}{\partial z_0^\beta} \right| \qquad \text{for} \qquad z_0^1 = \ldots = z_0^\nu = 0 \tag{39}$$

becomes zero for some value of t (Jacobi's "conjugate point"). Therefore, one has the necessity of *requiring Legendre's equations* (29) *to have a solution* $s_{\alpha\beta}^i$ *throughout the whole domain G.*

§ 20. Determination of the geodesic field by means of characteristics. *But, this admitted*, one is able to overcome the obstacle just mentioned. We split off the quadratic part:

$$s_2^i(t^k, z^\alpha) = \frac{1}{2} \sum_{\alpha,\beta} s_{\alpha\beta}^i(t^k) z^\alpha z^\beta$$

as formed by the given solution $s_{\alpha\beta}^i(t^k)$ of Legendre's equations as our first approximation, and thus put:

$$s^i(t^k, z^\alpha) = s_2^i(t^k, z^\alpha) + \bar{s}^i(t^k, z^\alpha),$$

$$H(t^i, z^\alpha, \partial s_2^i / \partial z^\alpha + \bar{p}_\alpha^i) - H_2(t^i, z^\alpha, \partial s_2^i / \partial z^\alpha) = \bar{H}(t^i, z^\alpha, \bar{p}_\alpha^i).$$

The equation (12) remains valid for the "corrections" $\bar{s}$ and $\bar{H}$:

$$\frac{\partial \bar{s}^i}{\partial t^i} = \bar{H}\left(t, z, \frac{\partial \bar{s}}{\partial z}\right),$$

but the situation is improved, insofar as the quadratic part $\bar{H}_2$ of $\bar{H}(z, \bar{p})$ contains no terms $z^\alpha z^\beta$ (only products $\bar{p}^i_\alpha z^\beta$, $\bar{p}^i_\alpha \bar{p}^k_\beta$). It is material that we start with *any* given solution of Legendre's equations without introducing the "cylindrical" specialization $s^2 = \ldots = s^\nu = 0$ for the $s^i_{\alpha\beta}(t^k)$. *The corrections $\bar{s}^i$, though, shall be determined in the cylindrical manner again:* $\bar{s}^2 = \ldots = \bar{s}^r = 0$. Thus, after returning to the old notations s, p, H, instead of $\bar{s}$, $\bar{p}$, $\bar{H}$, all previous relations are preserved, but we have won the further condition:

(40) $$H_{z^\alpha, z^\beta} = 0 \qquad \text{for} \qquad z = 0, \quad p = 0.$$

We treat equation (12) with the *new* Hamiltonian H by the method of characteristics again, *and now prove the non-vanishing of the determinant* (39).

For this purpose, we must consider the derivatives:

$$\zeta^\alpha_\beta = \frac{\partial \zeta^\alpha}{\partial z^\beta_0}, \qquad \pi_{\alpha\beta} = \frac{\partial \pi_\alpha}{\partial z^\beta_0} \qquad \text{for} \qquad z^1_0 = \ldots = z^\nu_0 = 0.$$

If $C^\alpha_\beta(t)$ denotes the second derivative:

$$H_{p_\alpha, z^\beta} \text{ for } \qquad z = 0, \quad p = 0,$$

one deduces, by differentiating the second line of equations (33) with respect to z^β_0 and taking into account the fact (40):

$$\frac{d\pi_{\alpha\beta}}{dt} = C^\gamma_\alpha(t)\, \pi_{\gamma\beta}(t).$$

Since $\pi_{\alpha\beta} = 0$ for $t = 0$, this leads at once to the result that $\pi_{\alpha\beta}(t) = 0$ for all values of t. In view of this situation, the first line (33) gives rise to the relations:

$$\frac{d\zeta^\alpha_\beta}{dt} = -\, C^\alpha_\gamma(t)\, \zeta^\gamma_\beta(t).$$

Hence, the determinant Δ of the ζ^α_β fulfills the simple equation:

(41) $$\frac{d\Delta}{dt} + c(t)\,\Delta = 0,$$

where $c(t)$ is the trace of the matrix $\| C^{\alpha}_{\beta}(t) \|$. The initial value of $\Delta(t)$ for $t = 0$ is 1; hence, from (41):

$$\Delta(t) = e^{-\int_0^t c(\tau)\,d\tau}.$$

This shows that $\Delta(t)$ is *positive throughout the whole interval* $-a \leq t \leq a$, and it even gives a fixed positive lower limit: $\Delta \geq e^{-ca}$, c being an upper bound to $c(t)$ in that interval. One easily infers now that a certain neighborhood $\mathfrak{N}_0$ of $z_0 = 0$ in a z_0-space is put into one-to-one correspondence with a neighborhood $\mathfrak{N}_t$ of $z = 0$ in z-space by means of the relations (36) for every fixed t in the interval $-a \leq t \leq a$.

Thus, one succeeds in building up the correction s^i that is to be added to Legendre's approximation s_2^i in order to get an *exact* geodesic field.

§ 21. The method of power series. One can hardly avoid a feeling of discontinuity regarding this whole process of solving the Jacobi-Hamilton equation – an equation that served as a tool for the theory of extremals – by means of its characteristics, which are something much akin to, but not quite identical with, the extremals. Furthermore, one ought to understand better why everything goes smoothly once the existence of the first approximation is granted.

Anyhow, I thought it worthwhile to carry through also the second, more direct, method: the application of *power series* whose convergence has to be secured through simple dominant series. Here, the reason becomes perspicuous: *the subsequent approximations depend on linear equations only*, whereas Legendre's equations for the first approximation are of the quadratic Ricatti type.

For our present purpose, one must assume at the outset that H is *analytic* in z and p, and is thus given as a power series in terms of all these variables z^{α} and p^i_{α}. The expansion begins with the quadratic terms H_2 only. Starting with a given solution $s^i_{\alpha\beta}(t^k)$ of Legendre's equations, we make use of the same trick as in § 20, and thus are able to assume H_2 to contain no products $z^{\alpha} z^{\beta}$. Let us subtract from H the part bilinear in z and p:

$$H = C^{\alpha}_{i\beta}(t)\, p^i_{\alpha} z^{\beta} + H^*, \tag{42}$$

and put the first term on the left side of our equation (12). Our solution s^i should be a power series in z, the terms of which we arrange by increasing order:

$$s^i(t, z) = s^i_3 + s^i_4 + \ldots$$

$$s^i_{\alpha}(t, z) = \sum \frac{n!}{n_1! \cdots n_{\nu}!} s^i(n_1 \cdots n_{\nu}; t)(z^1)^{n_1} \cdots (z^{\nu})^{n_{\nu}} \qquad (n_1 + \ldots + n_{\nu} = n)$$

is the totality of all terms of order n. The lowest order occurring is 3. The coefficients of the n^{th} approximation s^i_n have to satisfy equations of the type:

$$\frac{ds^i(n_1 \cdots n_\nu; t)}{dt^i} - \sum_\beta C^\alpha_{i\beta}(t)\, n_\beta \cdot s^i(\ldots\, n_{\alpha+1} \,\ldots\, n_{\beta-1} \,\ldots;\, t) = F(n_1 \,\ldots\, n_\nu;\, t). \tag{43}$$

The s^i in the second term on the left side contain the same indices $n_1 \ldots n_\nu$ as the first term if $\beta = \alpha$; the same holds for $\beta \neq \alpha$, except that n_α is increased and n_β diminished by 1. The right-hand side becomes a known function after *the preceding approximations of order lower than n have been computed.* This was the reason for our shoving over the first part of H in (42) to the left side of our equations.

§ 22. Solving and majorizing the differential equations for the approximations. At this stage, we introduce again our assumption of the cylinder-like topological nature of G, enabling us to put $s^2 = \ldots = s^r = 0$ and to forget about the variables $t^2, \ldots, t^r$. (43) are changed into ordinary differential equations:

$$\frac{ds(n_1 \cdots n_\nu; t)}{dt} - \sum_\beta C^\alpha_\beta(t)\, n_\beta \cdot s(\ldots\, n_{\alpha+1} \,\ldots\, n_{\beta-1} \,\ldots;\, t) = F(n_1 \,\ldots\, n_\nu;\, t), \tag{44}$$

which we want to solve under the initial conditions:

$$s(n_1 \ldots n_\nu;\, t) = 0 \qquad \text{for} \qquad t = 0.$$

The coefficients $C^\alpha_\beta(t)$ are the same as in § 20.

One knows how the solution is effected explicitly by an infinite series. One first combines the differential equations with the initial conditions into an integral equation:

$$s(t) = \int_0^t F(t)\,dt + \int_0^t C(t)\,s(t)\,dt\,.$$

s stands here for all those $s(n_1 \ldots n_\nu;\, t)$ for which $n_1 + \ldots + n_\nu$ has the prescribed value $n \geq 3$, arranged in a single column; F has the same significance, while $C(t)$ is the matrix of the linear transformation occurring in (44):

$$s(n_1 \ldots n_\nu) \to \sum_\beta C^\alpha_\beta(t)\, n_\beta \cdot s(\ldots\, n_\alpha + 1 \,\ldots\, n_\beta - 1 \,\ldots).$$

The solving series:

$$s(t) = s^{(0)}(t) + s^{(1)}(t) + s^{(2)}(t) + \ldots$$

is computed by successive integrations according to:

$$s^{(0)}(t) = \int_0^t F(\tau)\,d\tau\,, \qquad\qquad s^{(h+1)}(t) = \int_0^t C(\tau)\,s^{(h)}(\tau)\,d\tau\,.$$

This was mentioned for the purpose of deducing from it the *majorizing property:* if

(45) $$| C^{\alpha}_{\beta}(t) | \leq \Gamma^{\alpha}_{\beta}(t), \qquad | F(n_1 \ldots n_\nu ; t) | \leq \Phi(n_1 \ldots n_\nu ; t)$$

then the corresponding solution σ of the equations with Γ and Φ, instead of C and F, dominates s:

$$| s(n_1 \ldots n_\nu ; t) | \leq \sigma(n_1 \ldots n_\nu ; t).$$

Let us assume, in particular, that we are in possession of upper bounds:

(46) $$| C^{\alpha}_{\beta}(t) | \leq \Gamma, \qquad | F(n_1 \ldots n_\nu ; t) | \leq A_n \cdot e^{(n-2)At},$$

involving certain constants Γ, A, A_n, and valid throughout the interval $0 \leq t \leq a$. It is essential that neither Γ nor A depend on n. A bound like Γ can be assigned *a priori*, whereas the proper choice of A and A_n is to be kept open for later decision. With these dominants, (46), instead of (45), all the elements of our column $s(n_1 \ldots n_\nu; t)$ become equal. $\sigma_n(t)$ and the majorizing system (44) reduces to the simple equation:

$$\frac{d\sigma_n}{dt} - n\Gamma \cdot \sigma_n = A_n \cdot e^{(n-2)At},$$

with the solution:

$$\sigma_n = \frac{A_n}{(n-2)A - n\Gamma} \{e^{(n-2)At} - e^{n\Gamma t}\}.$$

If:

$$\frac{1}{3}A - \Gamma = B$$

is positive then the denominator $(n-2)A - n\Gamma$ will be $\geq nB > 0$ for $n \geq 3$. Thus, one is led to the estimation:

(47) $$| s(n_1 \ldots n_\nu ; t) | \leq \frac{1}{nB} \cdot A_n\, e^{(n-2)At}.$$

Consequently:

$$s(t, z^\alpha) \text{ is dominated by } \sum_{n=3}^{\infty} \frac{A_n}{nB} \cdot z^n\, e^{(n-2)At},$$

$$p_\alpha = \frac{\partial s}{\partial z^\alpha} \text{ is dominated by } \sum_{n=3}^{\infty} \frac{A_n}{B} \cdot z^{n-1}\, e^{(n-2)At},$$

$$(z = z^1 + \ldots + z^r).$$

§ 23. Recursive formula for the upper bounds. In order to determine an upper bound of the desired form (46) for $F(n_1 \ldots n_\nu; t)$, we first have to majorize the given Hamiltonian $H(t, z^\alpha, p_\alpha)$. Such a dominant may obviously be chosen in the form:

$$\frac{M(z+p)^2}{1-R(z+p)} - Mz^2,$$

since the products $z^\alpha z^\beta$ in the quadratic term H_2 of H are missing. p stands for $p_1 + \ldots + p_\nu$, just as z stands for $z^1 + \ldots + z^\nu$. The factors R and M are constants valid throughout the whole interval $-a \le t \le a$. $\Gamma = 2M$ is then a proper upper bound for the $C^\alpha_\beta(t)$, and H^* is dominated by:

(48) $$\frac{M(z+p)^2}{1-R(z+p)} - M(z^2 + 2zp).$$

We now replace p by its dominant as given at the end of the last section; that is, by $z \cdot f(\zeta)$, where:

(49) $$f(\zeta) = \frac{\nu}{B} \cdot \sum_{n=3}^{\infty} A_n \zeta^{n-2}$$

depends only on the combined argument $\zeta = z \cdot e^{At}$. The dominant (48) is still enlarged when one replaces R in the denominator by $R \cdot e^{At}$; it then takes on the form:

(50) $$\frac{Mz^2(1+f(\zeta))^2}{1-R\zeta(1+f(\zeta))} - Mz^2(1 + 2f(\zeta)),$$

The coefficient of z^n herein is an upper bound for $F(n_1 \ldots n_\nu; t)$ *provided that the inequalities* (47) *prevail for all orders less than n.* Because (50) equals z^2 times a function of $\zeta = z \cdot e^{At}$, that upper bound is precisely of the form (46). In this way, we have arrived at a proof of (47). The factors A_n are determined by the following recurrent equation for the generating function (49):

$$\frac{Mz^2(1+f(\zeta))^2}{1-R\zeta(1+f(\zeta))} - Mz^2(1 + 2f(\zeta)) = \sum_{n=3}^{\infty} A_n z^n e^{(n-2)At}$$

or

(51) $$\frac{(1+f)^2}{1-R\zeta(1+f)} - (1 + 2f) = \frac{B}{\nu M} \cdot f.$$

§ 24. The auxiliary quadratic equation. Final conclusions. The recurrent computation of the coefficients A_n of $f(\zeta)$ guarantees that they are positive, whereas the solution of the quadratic equation (51) for f will show that the series (49) is convergent in a circle round the origin. *This settles the convergence for our successive approximations.*

But let us be a little more explicit! On putting:

$$R\zeta = u, \qquad 1 + f = \varphi,$$

our equation becomes:

$$\frac{\varphi^2}{1-\varphi u} = \left(2+\frac{B}{\nu M}\right)\varphi - \left(1+\frac{B}{\nu M}\right).$$

Hence, we choose a constant $\alpha > 2$, take $\beta = \alpha - 1$, and consider the equation for φ:

$$\varphi^2 = (\alpha\varphi - \beta)(1 - \varphi u).$$

If

(52) $$\varphi = a_0 + a_1 u + a_2 u^2 + \ldots$$

is its solution with the initial coefficient $a_0 = 1$ then the inequalities (47) will hold with:

$$B = \nu M(\beta - 1), \qquad \frac{1}{3}A = M[\nu(\beta - 1) + 2],$$

$$A_n / B = a_{n-2} R^{n-2}.$$

We find:

(53) $$\varphi = \frac{(\alpha+\beta u) - \sqrt{(\alpha+\beta u)^2 - 4\beta(1+\alpha u)}}{2(1+\alpha u)}.$$

The square root must be taken with the minus sign at $u = 0$ in order to have the expansion (52) of φ start with the term 1. The quadric under the square root:

$$(\beta u - \alpha)^2 - 4\beta = \beta^2(u - u_1)(u - u_2)$$

has two positive roots:

$$u_1, u_2 = (a\ \frac{1}{\beta}\left(\alpha \mp 2\sqrt{\beta}\right).$$

Cauchy's integral formula gives the following expression for a_n:

$$a_n = \frac{1}{2\pi i}\int_k \frac{\varphi(u)\,du}{u^{n+1}}.$$

The integral extends over a small circle k about the origin. The function $\varphi(u)$ is regular in the complex u-plane to be slit along the line $u_1 \leq u \leq u_2$. It has no pole, since the numerator in (53) vanishes for $u = -1/\alpha$, where $1 + \alpha u = 0$, and it is finite at infinity. For negative real values of u the square root in (53) is positive, so that the value of φ at infinity equals β/α. Thus, K may be replaced, for $n \geq 1$, by a path closely surrounding the incision. One adds together in the usual manner the contributions from the opposite points on the two borders of the slit, and thus arrives at the formula:

(54) $$a_n = \frac{\beta}{2\pi}\cdot\int_{u_1}^{u_2} \frac{\sqrt{(u_2 - u)(u - u_1)}}{(1+\alpha u)\,u^{n+1}}\,du \qquad (n \geq 1)$$

which proves anew the positiveness of a_n . In the case $n = 0$ one has to add to the path around the slit an infinitely large circle K whose contribution will be:

$$\frac{\beta}{\alpha}\cdot\frac{1}{2\pi i}\int_K \frac{du}{u} = \frac{\beta}{\alpha}.$$

Since $a_0 = 1$ and $1 - (\beta/\alpha) = 1/\alpha$ one finds here:

$$\frac{1}{\alpha} = \frac{\beta}{2\pi}\cdot\int_{u_1}^{u_2} \frac{\sqrt{(u_2 - u)(u - u_1)}}{(1+\alpha u)u}\,du\,. \tag{55}$$

(55) yields the following bound for (54):

$$a_n \leq \frac{1}{\alpha}\cdot\frac{1}{u_1^n} = \frac{1}{\alpha}\left(\frac{\beta}{\alpha - 2\sqrt{\beta}}\right)^n.$$

Let us put $\beta = \gamma^2$, $\alpha = \gamma^2 + 1$, and replace by $\frac{1}{3}A$ in the final result. *We then find $p_a = \partial s / \partial z^\alpha$ to be dominated by:*

$$\frac{z}{1+\gamma^2}\cdot\sum_{n=1}^{\infty}\left\{\left(\frac{\gamma}{\gamma - 1}\right)^2 Rze^{3A|t|}\right\}^n,$$

where:

$$z = z^1 + \ldots + z^\nu, \qquad A = M[\nu(\gamma^2 - 1) + 2].$$

The number $\gamma > 1$ may be chosen at random, whereas M and R are fixed by the nature of the Hamiltonian $H(t^i, z^\alpha, p^i_\alpha)$ and the solution $s^i_{\alpha\beta}(t)$ of Legendre's equations. A reasonable choice for γ would be $\gamma = 2$.

THE INSTITUTE FOR ADVANCED STUDY
PRINCETON, NEW JERSEY

On extremals and geodesic fields in the calculus of variations for multiple integrals

By

Hermann Boerner in Munich

Introduction

The calculus of variations for multiple integrals – viz., several independent variables and several desired functions – exhibits some characteristic difficulties, as long as one wishes to proceed from the derivation of the Euler differential equations to the statement of necessary and sufficient conditions. For that reason, from the earliest times onward only a few first steps towards the presentation of a "Legendre condition" were suggested [1]). Caratheodory wrote the first comprehensive work [2]). His methods in the calculus of variations [3]), which consequently made use of the connection between the Hilbert "independent integrals" and the Hamilton-Jacobi partial differential equation, have also proved fruitful in overcoming the initial difficulties in precisely this case. In between these two approaches, there is the notion of a "geodesic field," which one obtains as one does in the special case of ray optics as the wave surfaces (i.e., the eikonal) of the rays. Carathéodory's "Legendre condition" and "Weierstrass $\mathcal{E}$-function" have appeared to be something other than one suspected up to now. As a particular lemma for the mastery of geodesic fields he devised a sort of generalized Legendre transformation that reduced to something different from the ordinary one in the case of line integrals.

The algebraic and analytical properties of this Legendre transformation, which has very little to do with the calculus of variations, took up the most space in Carathéodory's treatment. Thus, a representation would be welcome in which the variational problem appears at the outset and remains in the foreground; it will be given in the first chapter of the following work, which will likewise serve for its generalization in the second. It also shows that one may make many things much simpler when one has it at one's disposal. Thus, I will first derive the $\mathcal{E}$-function and the Legendre condition with few calculations, and then the write down the comprehensive system of formulas for the Legendre transformation that I will need for later purposes. In the beginning, I choose the path that Carathéodory followed in his book [4]), which truly represents the most elegant, and

---

[1]) Clebsch, Crelles Journ. **56** (1859), pp. 122-148; Hadamard, Bull. Soc. Math. de France **30** (1902), pp. 253-256; **33** (1905), pp. 73-80; McShane, Ann. Math. **32** (1931), pp. 578. – Prange treated, *inter alia*, the case of two independent and two dependent variables in his Diss. (Gött. 1915); he employed a Legendre transformation as a means of integrating the Euler differential equations.

[2]) Carathéodory, Acta Szeged **4** (1929), pp. 193-216.

[3]) Carathéodory, Variationsrechnung und partielle Differentialgleichungen erster Ordnung, Leipzig and Berlin 1935.

[4]) pp. 197, et seq. He gave it for the first time in his lectures on geometrical optics in Summer 1934.

likewise simplest, path to all of the basic formulas of the calculus of variations. Here, one may convince oneself that one brings a difficult problem to its solution in an exceptionally quick and compelling way by this means.

However, this theory is still incomplete in one essential point: One has shown that any surface that is intersected transversally by a geodesic field [1]) is a solution of the variational problem when the Legendre and Weierstrass conditions are satisfied, moreover, and it has been proved that any such surface is an "extremal," i.e., it satisfies the Euler differential equations. However, in order to prove that these three conditions – viz., those of Euler, Legendre, and Weierstrass – when formulated in a particular way, are *necessary* (*sufficient*, resp.), one must be able to "embed" each (sufficiently small) piece of an extremal in a geodesic field, i.e., to find a geodesic field that intersects it transversally. The second chapter of this work is dedicated to the proof of this embedding theorem.

In the case of a single desired function there is nothing to prove: One needs only to construct more extremals in the neighborhood of the given extremal; here, any extremal field is a geodesic field. The theorem is also very easy to prove in the case of a single independent variable; in that case, any geodesic field is an extremal field, and the extremals are the characteristics of the Hamilton-Jacobi equation.

Things are different in the general case. However, one also arrives at the proof here with the help of the theory of characteristics. Namely, the geodesic fields will be still be obtained as the solutions of a single partial differential equation, in such a way that one can choose the system of functions that one uses to treat it with here up to an arbitrary function. When one does this in a particular way – as I show – then each characteristic curve of the partial differential equation that contacts the extremal surface lies completely within it. One thus needs only to choose the Cauchy initial values suitably in order for all of the characteristics that begin in the extremals to lie completely within them, and then they will, in fact, be transversally intersected by the geodesic field.

In general, the geodesic field will intersect no other extremals transversally, and is therefore not indeed an extremal field. However, that is also completely unnecessary, and precisely this circumstance pushes the use of the notion of a "geodesic field" into a brighter light. In order to be able to write down the Weierstrass formula, one needs, in fact, precisely a geodesic field and nothing more [2]).

In the third chapter, I give the beginnings of a theory of discontinuous solutions for multiple integrals. Here, as well, the latter viewpoint plays a decisive role, namely, that one does not necessarily have to construct a *field* of discontinuous solutions. The geodesic fields in canonical variables prove to be, moreover, the most convenient for the presentation of the generalized Erdmann equations.

---

[1]) For the definition of this notion, cf. sec. 2 of this work.

[2]) Recently, Weyl [Phys. Rev. **46** (1934), pp. 5050; Ann. Math. **36** (1935), pp. 607] has made a new attempt, and likewise, with some modification, an extremal goes through a geodesic field that likewise obeys the Hamilton-Jacobi method, and seems much simpler than what is presented here. Weyl's formulas are all linear, as in the simple problems, whereas for us determinants of linear expressions always appear. However, it also happens that the Weyl theory is not in a position to answer all of the questions that one can pose in the calculus of variations. Namely, *transversality* can only be defined by nonlinear formulas in the general problems, and therefore it follows that in the Weyl theory, simply stated, it is impossible to compare surfaces that do not possess the same boundary.

First Chapter

The geodesic field and the Legendre transformation

1. n functions:

(1.1) $$x_i(t_\alpha)$$

of μ variables [1]) define a μ-dimensional manifold in the $n + \mu$-dimensional space of the x_i , t_α.

Their derivatives, which we briefly denote by:

(1.2) $$\frac{\partial x_i}{\partial t_\alpha} = p_{i\alpha} ,$$

shall be piecewise continuous and differentiable. If $\Psi(x_i , t_\alpha)$ is an arbitrary function in space then we can replace the x_i with the functions (1.1), in particular; i.e., we consider the function to be defined on our surface, in particular. We then denote its derivatives on the surface by the plain d [2]); one thus has:

(1.3) $$\frac{d\Psi}{dt_\alpha} = \frac{\partial \Psi}{\partial t_\alpha} + p_{i\alpha}\frac{\partial \Psi}{\partial x_i}$$ [3]).

We will employ this differential operator when the $p_{i\alpha}$ are defined, not on a surface (1.1), but, for example, as functions in space. In particular, if these functions belong to a family of surfaces (1.1) that cover a piece of space simply then they satisfy the condition:

(1.4) $$\frac{dp_{i\alpha}}{dt_\beta} = \frac{dp_{i\beta}}{dt_\alpha}$$

everywhere.

Let $f(x_i, t_\alpha, p_{i\alpha})$ be a *positive* function that is analytic in its $n + \mu + n\mu$ arguments. We denote its derivatives by indices and introduce the abbreviated notation $\pi_{i\alpha} = f_{p_{i\alpha}}$.

If G_1 is a region in the space of the μ variables t_α in which the functions (1.1) are defined then one can consider the integral [4]):

(1.5) $$\int_{G_1} f(x_i, t_\alpha, p_{i\alpha})\, dt .$$

---

[1]) Latin indices always range through the numbers from 1 to n; Greek indices range from 1 to μ .

[2]) Total derivatives of a variable first appear in the second chapter, and there they will be denoted by a dot, so there should be no danger of confusion.

[3]) Any doubled index in a term is to be summed over.

[4]) We will consider only μ-fold integrals and thus we will briefly write $\int \dots dt$, instead of $\int \dots \int \dots dt_1 \dots dt_\alpha$.

The variational problem reads: How must the functions (1.1) *be chosen in order for the integral* (1.5) *to possess a smaller value than the same integral when it is taken over any other surface* $x_i'(t_\alpha)$? For this, the class of surfaces to be compared must be made somewhat more precise. In general, we will assume that the boundary of the surface is given, i.e., that we always integrate over the same region G_1 and the functions $x_i'(t_\alpha)$ shall agree with (1.1) on the boundary of this region.

2. In order to answer this question, with Carathéodory [1]), we embark upon the following path:

A surface (1.1) is apparently a solution of the variational problem when the following is true: In a region of space that contains the surface there exist functions $p_{i\alpha}(x_i, t_\alpha)$ that satisfy (1.2) on the surface – on the contrary, (1.4) need not be valid outside the surface, at all – and for these functions one has $f(x_i, t_\alpha, p_{i\alpha}) = 0$ everywhere; on the other hand, $f(x_i, t_\alpha, p'_{i\alpha}) > 0$ when $p'_{i\alpha} \neq p_{i\alpha}$. Our function f does not possess his property – indeed, we have assumed that it is always positive. However, we can seek to construct an *equivalent* problem that does possess it.

Two problems are called *equivalent* when any solution of the one is likewise a solution of the other. This is especially the case when the integrals differ from each other by a "path-independent" integral; i.e., one that possesses the same value on two surfaces that agree on the boundary. One arrives at such an independent integral in the following way:

We introduce μ functions $S_\alpha(x_i, t_\beta)$ that shall possess continuous derivatives up to second order. The equations:

$$S_\alpha(x_i, t_\beta) = \lambda_\alpha \tag{2.1}$$

represent a μ-parameter family of n-dimensional surfaces. If one replaces the x_i with functions of t_α then equations (2.1) define a one-to-one map from a region G_t to a region G_λ in the space of the t_α in the case that the surface $x_i(t_\alpha)$ goes through the manifolds (2.1) without touching them. In this case, the functional determinant:

$$\Delta = \frac{dS_\alpha}{dt_\beta} = | S_{\alpha\beta} + S_{i\alpha}\, p_{i\beta} | \qquad {}^{2})$$

is non-zero. In the event that it is positive, the integral:

$$\int_{G_t} \Delta\, dt = \int_{G_\lambda} d\lambda$$

represents the volume of the region. If one therefore considers any other functions $x_i'(t_\alpha)$ then one has:

$$\int_{G_t'} \Delta' dt = \int_{G_t} \Delta\, dt\,, \tag{2.2}$$

---

[1]) Variationsrechnung und partielle Differentialgleichungen erster Ordnung, pp. 197, et seq.

[2]) Here, as in the sequel, we have set $\partial S_\alpha / \partial x_i = S_{\alpha i}$ and $\partial S_\alpha / \partial t_\beta = S_{\alpha\beta}$.

as long as the region G_t' is mapped to the same region G_λ as G_t. Furthermore, this is certainly the case when the functions $x_i'(t_\alpha)$ agree with the $x_i(t_\alpha)$ on the boundary of G_t, and one takes $G_t' = G_t$.

We thus obtain an equivalent problem when we replace the function f with $f - \Delta$, and a solution of the variational problem in the previously suggested sense, when the family (2.1) possesses the following property: At every point the minimum of the function $f - \Delta$ is zero under variation of the $p_{i\alpha}$. A family (2.1) with this property is called a *geodesic field.* To a geodesic field there then belongs a system of functions $p_{i\alpha}(x_i, t_\alpha)$ that satisfy the equations:

(2.3) $$\pi_{i\alpha} = \frac{\partial \Delta}{\partial p_{i\alpha}}$$

and:

(2.4) $$f = \Delta.$$

Moreover, one says that the surface element that is defined by these $p_{i\alpha}$ will *transversally* intersect each surface of the family (2.1) that goes through these points.

Now, if these functions $p_{i\alpha}$ satisfy equations (1.2) on a surface (1.1) then this surface is, in fact, a solution of the variational problem. Then, if $x_i'(t_\alpha)$ is a comparison surface with the same boundary then due to (2.4) and (2.2) one has:

(2.5) $$\int_{G_t} f'\,dt - \int_{G_t} f\,dt = \int_{G_t} (f' - \Delta')\,dt,$$

and this is positive.

(2.5) is the Weierstrass formula, and the function $f' - \Delta'$, which defines our equivalent problem, is nothing but the Weierstrass $\mathcal{E}$-function for our problem. From (2.4) and (2.3), it follows that it vanishes and is stationary for $p_{i\alpha}' = p_{i\alpha}$. That one is really dealing with a *minimum* when the $\mathcal{E}$-function is therefore *positive* for $p_{i\alpha}' = p_{i\alpha}$ must then be introduced as a special condition. However, if this is not satisfied then our surface, which intersects the geodesic field transversally, actually provides a "strong" minimum for the integral (1.5); for the comparison surfaces, one needs to assume nothing more than that they lie in the field.

We would now like to derive another expression for this $\mathcal{E}$-function. We will see that one can define this function without the use of a geodesic field and express it in terms of only the functions $p_{i\alpha}$ and $p_{i\alpha}'$.

3. In order to do this, we summarize some formulas from the theory of determinants that we will need incessantly in the sequel.

If $\psi_{\alpha\beta}$ are the elements of a non-singular matrix with μ rows and μ columns then we denote its determinant by ψ, and the algebraic complement of $\psi_{\alpha\beta}$ in this determinant by $\bar{\Psi}_{\alpha\beta}$. One then has $\psi \neq 0$ and:

(3.1) $$\psi_{\alpha\beta}\bar{\Psi}_{\beta\gamma} = \psi_{\rho\alpha}\bar{\Psi}_{\rho\beta} = \delta_{\alpha\beta}\,\psi,$$

and thus, from the multiplication theorem for determinants:

(3.2) $$\left|\bar{\psi}_{\alpha\beta}\right| = \psi^{\mu-1}.$$

If the $\psi_{\alpha\beta}$ are functions of any other variables then one has:

(3.3) $$d\psi = \bar{\psi}_{\alpha\beta}\, d\psi_{\alpha\beta}\,.$$

Occasionally, we will also need the derivatives of the $\bar{\psi}_{\alpha\beta}$. From (3.1) and (3.3), it follows that:

$$\delta_{i\alpha}\bar{\psi}_{\rho\sigma}d\psi_{\rho\sigma} = \delta_{i\alpha}\, d\psi = \psi_{\lambda\mu}d\bar{\psi}_{\alpha\mu} + \bar{\psi}_{\alpha\mu}d\psi_{\lambda\mu}\,,$$

and when one contracts this with $\bar{\psi}_{\alpha\beta}$ (i.e., multiplies with $\bar{\psi}_{\alpha\beta}$ and sums over λ) one obtains, with a simple conversion:

(3.4) $$d\bar{\psi}_{\alpha\beta} = \frac{1}{\psi}(\bar{\psi}_{\alpha\beta}\bar{\psi}_{\lambda\mu} - \bar{\psi}_{\alpha\mu}\bar{\psi}_{\lambda\beta})d\psi_{\lambda\mu}\,.$$

What appears here as the coefficient of $d\psi_{\alpha\beta}$ on the right-hand side is likewise naturally the algebraic complement of $\psi_{\lambda\mu}$ in the sub-determinant $\bar{\psi}_{\alpha\beta}$.

4. Now, we would like to calculate the quantity Δ' as a function of the $p_{i\alpha}$ and $p'_{i\alpha}$. To that end, we denote the elements of the determinant Δ by $c_{\alpha\beta}$. Then, from (2.3), we calculate:

(4.1) $$\pi_{i\beta} = S_{\rho i}\bar{c}_{\rho\beta}\,.$$

Furthermore, from (2.4), it follows that:

$$\delta_{\alpha\beta} f = c_{\rho\alpha}\bar{c}_{\rho\beta} = S_{\rho\alpha}\bar{c}_{\rho\beta} + S_{\rho i}\, p_{i\alpha}\bar{c}_{\rho\beta} = S_{\rho\mu}\bar{c}_{\rho\beta} + p_{i\alpha}\pi_{i\beta}\,.$$

If we now introduce new quantities $a_{\alpha\beta}$ by way of:

(4.2) $$a_{\alpha\beta} = \delta_{\alpha\beta} f - p_{i\alpha}\,\pi_{i\alpha}$$

then we have:

(4.3) $$a_{\alpha\beta} = S_{\rho\mu}\bar{c}_{\rho\beta}\,.$$

Now, we contract the elements of Δ' with $\bar{c}_{\rho\beta}$. Due to (4.3) and (4.1), one obtains:

$$(S_{\rho\alpha} + S_{\rho i}\, p'_{i\alpha})\bar{c}_{\rho\beta} = a_{\alpha\beta} + p'_{i\alpha}\, p_{i\beta} = \delta_{\alpha\beta} f + (\, p'_{i\alpha} - p_{i\alpha})\, p_{i\beta}\,.$$

For the determinant, this means, from (3.2) and (2.4):

(4.4)
$$\Delta' = \frac{1}{f^{\mu-1}} \,|\, \delta_{\alpha\beta} f + (p'_{i\alpha} - p_{i\alpha})\, p_{i\beta} \,| \,.$$

The $\mathcal{E}$-function is, as saw, the quantity $f' - \Delta'$. We thus have to set:

(4.5)
$$\mathcal{E}(x_i, t_\alpha, p_{i\alpha}, p'_{i\alpha}) = f' - \frac{1}{f^{\mu-1}} \,|\, \delta_{\alpha\beta} f + (p'_{i\alpha} - p_{i\alpha})\, p_{i\beta} \,| \,.$$

5. We are now also in a position to give the Legendre condition for our problem. Thus, we develop the $\mathcal{E}$-function in powers of $p'_{i\alpha} - p_{i\alpha}$. To that end, we calculate the derivatives of Δ' with respect to $p'_{i\alpha}$ at the location $p_{i\alpha}$, or, what amounts to the same thing, the derivatives of Δ with respect to $p_{i\alpha}$ with the use of (2.3) and (2.4) *afterwards.*

We have already calculated the first derivative; it is:

$$\left.\frac{\partial \Delta'}{\partial p'_{i\alpha}}\right|_{p'_{i\alpha} = p_{i\alpha}} = \pi_{i\alpha} \,.$$

In order to obtain the second derivatives, we must differentiate $S_{\lambda i} \bar{c}_{\lambda\alpha}$ with respect $p_{i\beta}$. With the use of (3.4), we obtain:

$$\frac{\partial^2 \Delta}{\partial p_{i\alpha} \partial p_{j\beta}} = \frac{S_{\lambda i} S_{\mu j}}{\Delta} (\bar{c}_{\lambda\alpha} \bar{c}_{\mu\beta} - \bar{c}_{\lambda\beta} \bar{c}_{\mu\alpha})$$

and due to (4.1) and (2.4):

$$\left.\frac{\partial^2 \Delta'}{\partial p'_{i\alpha} \partial p'_{j\beta}}\right|_{p'_{i\alpha} = p_{i\alpha}} = \frac{1}{f} (\pi_{i\alpha}\, \pi_{j\beta} - \pi_{i\beta}\, \pi_{j\alpha}) \,.$$

If one thus considers the development of $\mathcal{E}$ then, as one sees, no terms of null or first order appear; the terms of second order define a quadratic form with the coefficients:

(5.1)
$$q_{i\alpha, j\beta} = f_{p_{i\alpha} p_{j\beta}} - \frac{1}{f} (f_{p_{i\alpha}} f_{p_{j\beta}} - f_{p_{i\beta}} f_{p_{j\alpha}})$$

in the $n \cdot \mu$ variables $(p'_{i\alpha} - p_{i\alpha})$. The *Legendre condition* consists in the requirement that this form shall be positive definite. Surface elements that satisfy this condition are called *regular*. If it is satisfied then the $\mathcal{E}$-function is certainly positive when $p'_{i\alpha}$ differs slightly from $p_{i\alpha}$. It therefore guarantees the existence of a "weak" minimum. In fact, one can indeed set:

(5.2)
$$\mathcal{E} = \bar{q}_{i\alpha, j\beta} (p'_{i\alpha} - p_{i\alpha})(p'_{j\beta} - p_{j\beta}) \,,$$

where one has to define the coefficients of the quadratic form for a value $p_{i\alpha} + \theta(p'_{i\alpha} - p_{i\alpha})$ on the connecting line from $p'_{i\alpha}$ to $p_{i\alpha}$.

6. We have seen that one can express the quantities that are important for the variational problem in terms of only $p_{i\alpha}$ and f without the use of the functions S_α . Furthermore, in fact, a certain arbitrariness indeed underlies the choice of these functions. Two geodesic fields are already regarded as identical when only their surface elements agree at each point. That prompts us to introduce a system of values that are connected to precisely these surface elements in a one-to-one manner, and we will see very soon that these quantities yield the must useful *canonical variables* for our variational problem.

We think of an n-dimensional surface of the family (2.1) that goes through a point as being given by functions $t_\alpha(x_i)$ in the neighborhood of this point and set:

$$\frac{\partial t_\alpha}{\partial x_i} = - P_{i\alpha} .$$

We would now like to consider those families of surfaces for which this is possible; thus the determinant of $S_{\alpha\beta}$ must be non-zero for given functions S_α, and therefore we calculate the functions $P_{i\alpha}$ from the equations:

(6.1) $$S_{i\alpha} = P_{i\rho}\, S_{\alpha\rho} .$$

We next introduce the quantity F by means of:

(6.2) $$F \cdot | S_{\alpha\beta} | = 1 .$$

Analogous to the differential operator (1.3), we can now introduce the symbol:

(6.3) $$\frac{d}{dx_i} = \frac{\partial}{\partial x_i} - P_{i\rho} \frac{\partial}{\partial t_\rho} ;$$

(6.1) then simply reads $dS_\alpha / dx_i = 0$.

As differential quotients, the $P_{i\alpha}$ must satisfy a system of differential conditions, namely:

(6.4) $$[i\, j\alpha] = 0 \qquad (i, j = 1, \ldots, n;\ \alpha = 1, \ldots, \mu),$$

where we have set:

(6.5) $$\frac{dP_{i\alpha}}{dx_j} - \frac{dP_{j\alpha}}{dx_i} = [i\, j\alpha] .$$

A further condition shall likewise be derived next. If one totally differentiates $S_{\alpha\beta}$ with respect to x_i then one obtains, with the use of (6.1):

$$\frac{dS_{\alpha\beta}}{dx_i} = S_{\alpha\rho}\frac{\partial P_{i\rho}}{\partial t_\beta} \qquad [1]).$$

When one thus differentiates the determinant $| S_{\alpha\beta} | = 1/F$ from (3.3) one obtains (cf., (3.1)):

$$\frac{d\frac{1}{F}}{dx_i} = \frac{1}{F}\frac{\partial P_{i\rho}}{\partial t_\rho}$$

or:

$$\frac{dF}{dx_i} + F\frac{\partial P_{i\rho}}{\partial t_\rho} = 0 \,. \tag{6.6}$$

From the theory of partial differential equations of first order it follows that (6.4) and (6.6) are the only conditions to which one must subject the functions $P_{i\alpha}$ and F; thus, there are functions S_α that satisfy (6.1) and (6.2).

All of this is valid for families of n-dimensional surfaces. Now, we again consider a geodesic field, in particular, and look for the relations that exist between μ-dimensional surface elements $p_{i\alpha}$ that will make it intersect them transversally. Our notation is completely symmetric: The μ-dimensional surface element will be spanned by the μ vectors:

$$(p_{i\beta},\ \delta_{\alpha\beta}), \qquad \beta = 1, \ldots, \mu,$$

and the n-dimensional element that is perpendicular to it by the n vectors:

$$(-\delta_{ij},\ p_{i\beta}), \qquad i = 1, \ldots, n;$$

by contrast, the element that is *transversal* to it is spanned by the n vectors:

$$(-\delta_{ij},\ P_{i\alpha}), \qquad i = 1, \ldots, n\,.$$

For any simple problem, where the notions of "orthogonal" and "transversal" agree, one has $P_{i\alpha} = p_{i\alpha}$ [2]).

However, in the general case, it is also easy to calculate $P_{i\alpha}$ and F from $p_{i\alpha}$ and f with the help of equations (2.3) and (2.4) for geodesic fields. From (4.3), it follows with the use of formula (3.2):

$$a = | S_{\alpha\beta} |\, f^{\mu-1}\,.$$

Hence:

$$F = \frac{f^{\mu-1}}{a}, \tag{6.7}$$

---

[1]) One observes that one has $\partial S_{\alpha i} / \partial t_\beta = \partial S_{\alpha\beta} / \partial x_i$, $\partial S_{\alpha\rho} / \partial t_\beta = \partial S_{\alpha\beta} / \partial t_\beta$.

[2]) For the problem of the shortest arclength or the smallest surface area, one has indeed $F = f$.

in the case $a \neq 0$, which we will assume from here on, and when one substitutes (6.1) in (4.1) one obtains:

(6.8) $$\pi_{i\beta} = S_{\rho\alpha} P_{i\rho} \, \bar{c}_{\rho\beta} = P_{i\rho} \, a_{\alpha\beta} \, .$$

If one contracts that with $\bar{a}_{\alpha\beta} / a$ then one obtains:

(6.9) $$P_{i\alpha} = \frac{\bar{a}_{\alpha\beta}}{a} \pi_{i\beta} \, .$$

In the future, we will define $P_{i\alpha}$ and F by means of formulas (6.7) and (6.9). Equations (6.1) and (6.2) are then completely equivalent to (2.3) and (2.4) and, at that point, these equations will serve as the definition of geodesic fields. Our "generalized Legendre transformation" is, however, still not finished: We must see whether we can solve (6.9) for the $p_{i\alpha}$, and thus calculate F as a function of the $P_{i\alpha}$, namely, as a "Hamiltonian function."

7. Before we do this, we briefly develop the algebraic relations between the lower-case and upper-case notations. The symmetry of our transformation will clearly emerge, and we will later need the majority of these formulas, anyway.

We write all of the formulas in the form where the elements of the determinants appear *without* overbars: By contracting components, one can succeed in solving each formula in terms of the variables that appear on the right.

As a starting point, we recall the previously-developed formulas:

(7.a1) $$a_{\alpha\beta} = \delta_{\alpha\beta} f - p_{i\alpha} \, \pi_{i\beta} \, ,$$

(7.a2) $$\pi_{i\alpha} = P_{i\rho} \, a_{\rho i\alpha} \, ,$$

(7.a3) $$F = \frac{f^{\mu-1}}{a} \, .$$

We introduce the matrix:

(7.g1) $$g_{\alpha\beta} = \delta_{\alpha\beta} f + P_{i\alpha} \, p_{i\beta} \, ,$$

with whose application many other particularly simple things may be written down. One immediately arrives at the simple connection between $g_{\alpha\beta}$ and $a_{\alpha\beta}$ when one subtitutes (6.9) in (7.g1) and observes that $p_{i\alpha} \, \pi_{i\beta} = \delta_{\alpha\beta} f - a_{\alpha\beta}$. One finds that:

$$g_{\alpha\beta} = \frac{f}{a} \bar{a}_{\alpha\beta} \, ,$$

and from this, it follows that:

(7.g2) $$g_{\alpha\beta} \, a_{\beta\rho} = \delta_{\alpha\beta} f,$$

and conversely:

$$\bar{g}_{\alpha\beta} = a_{\alpha\beta} F \, .$$

For the determinant, (7.g2) then yields, because of (7.a3):

(7.g3) $$g = f\,F\,.$$

From (7.a2), it follows, when one contracts with $g_{\alpha\beta}$, and due to (7.g2), that:

(7.g4) $$f\,P_{i\alpha} = \pi_{i\rho}\,g_{\alpha\rho}\ .$$

One can also switch the roles of the Greek and Latin indices [1]). We introduce the matrix:

(7.h1) $$h_{ij} = \delta_{ij} + P_{i\alpha}\,p_{i\alpha}\,.$$

When one substitutes $a_{\rho\alpha}$ into (7.a2), one sees immediately that:

(7.h4) $$f\,P_{i\alpha} = \pi_{r\alpha}\,h_{ir}\ .$$

In exactly the same way, it follows conversely from (7.g4), when one introduces:

(7.b1) $$b_{ij} = \delta_{ij}\,f - p_{i\alpha}\,\pi_{j\alpha}\,,$$

that:

(7.b2) $$\pi_{i\alpha} = P_{r\alpha}\,b_{ri}\,.$$

(7.h4) and (7.b2) are constructed in precisely the same way as (7.g4) and (7.a2). Therefore, one must also have:

(7.h2) $$h_{ir}\,b_{jr} = \delta_{ij}\,f\,.$$

Now, the determinants h and b are still missing. However, one immediately sees that $h = g$. Namely, both of them are equal to the $n + \mu$-rowed determinant:

$$\begin{vmatrix} \delta_{ij} & -p_{i\beta} \\ p_{j\alpha} & \delta_{\alpha\beta} \end{vmatrix}$$ [2]).

Thus, one has:

(7.h3) $$h = f\,F\,,$$

and therefore it follows from (7.h2) that:

$$f\,F\,b = f^{\,n},$$

hence:

(7.h3) $$F = \frac{f^{\,n-1}}{f}$$ [3]).

---

[1]) Cf., C. Carathéodory, Math. Ann. **86** (1922), pp. 272.

[2]) Cf., the previous citation. Usually, it follows from the non-vanishing of this determinant that the mutually transversal surface elements $p_{i\alpha}$ and $P_{i\alpha}$ that were considered in the previous section do not touch each other.

[3]) If $a \neq 0$ then one also has $b \neq 0$; this is indeed the case when one can also show directly that $f^n\,a = f^\mu\,b$.

The system of formulas is, however, still incomplete, as long as the analogous one relating $\Pi_{i\alpha}$ to $p_{i\alpha}$ is missing. We must introduce these quantities in such a way that when f is a function of $p_{i\alpha}$ and F is a function of $P_{i\alpha}$ the relations $p_{i\alpha} = f_{p_{i\alpha}}$ and $\Pi_{i\alpha} = F_{P_{i\alpha}}$ are mutually implicit. To that end, we next assume that all of our quantities depend upon arbitrary parameters, and construct the differential of equation (7.g3): $dg = F\,df + f\,dF$. From formula (3.3), one obtains, due to $\bar{g}_{\alpha\beta} = F\,a_{\alpha\beta}$:

$$dg = F\,a_{\alpha\beta}\,dg_{\alpha\beta} = F\,a_{\alpha\beta}\,(P_{i\alpha}\,dp_{i\alpha} + p_{i\alpha}\,dP_{i\alpha})$$

and finally, due to (7.a2) and (7.a3):

$$F(df - \pi_{i\alpha}\,dp_{i\alpha}) + f\left(dF - \frac{f^{\mu-2}}{a}\cdot a_{\alpha\beta}\,p_{i\beta}\,dP_{i\alpha}\right) = 0\ .$$

We must therefore set:

(7.a4) $$\Pi_{i\alpha} = \frac{f^{\mu-2}}{a}\,p_{i\rho}a_{\alpha\rho}\,,$$

and then one has:

(7.*) $$F(df - \pi_{i\alpha}\,dp_{i\alpha}) + f(dF - \Pi_{i\alpha}\,dP_{i\alpha}) = 0.$$

If one regards f as a function of x_i, t_α, and $p_{i\alpha}$, and F as a function of x_i, t_α, *and* $P_{i\alpha}$, and if $\pi_{i\alpha} = f_{p_{i\alpha}}$ and $\Pi_{i\alpha} = F_{P_{i\alpha}}$ then (7.*) implies the following important formulas:

(7.**) $$F\ f_{x_i} = -f\ F_{x_i}\,, \qquad F\ f_{t_\alpha} = -f\ F_{t_\alpha}\ .$$

However, we must still complete the system of algebraic formulas. Had we calculated dh instead of dg above, then we would have found, in place of (7.a4):

(7.b4) $$\Pi_{i\alpha} = \frac{F^{n-1}}{b}\,p_{r\alpha}b_{ir}\ .$$

From (7.a4) and (7.b4), it follows upon contracting with $g_{\alpha\beta}$ (h_{ij}, resp.):

(7.g5) $$F\,p_{i\alpha} = \Pi_{i\rho}\,g_{\rho\alpha}$$

and:

(7.h5) $$F\,p_{i\alpha} = \Pi_{r\alpha}\,h_{ri}\ .$$

Now, the solution of the system of formulas (7.a2-4) or (7.b2-4) for the upper case symbols, without taking into account the remark, yields that the formulas (7.g3-5) or (7.h3-5), from which one can derive everything else, are symmetric in the lower case and upper case symbols. In fact, (7.g3) and (7.h3) remain unchanged when one switches the lower case and upper case symbols; (7.g4) and (7.g5) will be switched, as well as (7.h4) and (7.h5). Clearly, we thus still need to introduce the matrices:

(7.A1) $$A_{\alpha\beta} = \delta_{\alpha\beta} F - P_{i\alpha}\Pi_{i\beta}$$
and:
(7.B1) $$B_{ij} = \delta_{ij} F - P_{i\alpha}\Pi_{j\alpha},$$

and then we simply rewrite the previous formulas with the exchange of lower and upper case quantities:

(7.A2) $$\Pi_{i\alpha} = p_{i\rho} A_{\rho\alpha},$$ (7.B2) $$\Pi_{i\alpha} = p_{r\alpha} B_{ri},$$

(7.A3) $$f = \frac{F^{\mu-1}}{A},$$ (7.B3) $$f = \frac{F^{n-1}}{B},$$

(7.A4) $$\pi_{i\alpha} = \frac{F^{\mu-2}}{A} P_{i\rho} A_{\alpha\rho},$$ (7.B4) $$\pi_{i\alpha} = \frac{F^{\mu-2}}{B} P_{r\alpha} B_{ir}.$$

From $\bar{g}_{\alpha\beta} = F\, a_{\alpha\beta}$ it finally follows that:

(7.***) $$F\, a_{\alpha\beta} = f\, A_{\beta\alpha}.$$

Finally, we introduce the quantities φ and Φ by way of:

(7.φ) $$f + \varphi = p_{i\alpha}\Pi_{i\alpha}$$
and:
(7.Φ) $$F + \Phi = P_{i\alpha}\Pi_{i\alpha}.$$

If one rewrites (7.***) in detail and identifies a and b then one obtains:

$$F\, p_{i\alpha}\,\pi_{i\alpha} = f\, P_{i\alpha}\,\Pi_{i\alpha}$$

or:

$$F(f + \varphi) = f(F + \Phi).$$

One thus has:

(7.****) $$\frac{\Phi}{\varphi} = \frac{F}{f}.$$

8. As a first application, we write the $\mathcal{E}$-function in canonical variables. If one substitutes (6.1) in:

$$\Delta' = |\, S_{\alpha\beta} + S_{\alpha i}\; p'_{i\beta}\,|$$

and observes (6.2), then one obtains:

(8.1) $$F\,\Delta' = |\, \delta_{\alpha\beta} + P_{i\alpha}\; p'_{i\beta}\,|.$$

Here, the right-hand side remains invariant when one simultaneously exchanges the primed and unprimed notations, along with the lower case and upper case ones. Thus, instead of (4.4), one can also write:

$$\Delta' = \frac{f'}{F\,F^{\mu-1}}\,|\,\delta_{\alpha\beta}F' + (P_{i\alpha} - P'_{i\alpha})\Pi'_{i\beta}\,|,$$

and obtain for the $\mathcal{E}$-function, in place of (4.5):

$$\frac{F}{f'}\mathcal{E} = F - \frac{f'}{F^{\mu-1}} \,|\, \delta_{\alpha\beta} F' + (P_{i\alpha} - P'_{i\alpha})\Pi'_{i\beta} \,|\, . \tag{8.2}$$

If one develops this in powers of $P_{i\alpha} - P'_{i\alpha}$ then one obtains for the second order terms a quadratic form with the coefficients:

$$Q_{i\alpha,\, i\beta} = F_{P_{i\alpha}P_{j\beta}} - \frac{1}{F}(\Pi_{i\alpha}\Pi_{j\beta} - \Pi_{i\beta}\Pi_{j\alpha}) , \tag{8.3}$$

which are to be constructed at the location $P'_{i\alpha}$, here.

9. However, we have anticipated this. Indeed, we still do not know whether one really can solve (6.9) in terms of $p_{i\alpha}$, and, in that way, calculate F as a function of $P_{i\alpha}$.

Instead of (6.9), we start with formula (7.a2), which has the same effect, and take the differential of this equation in general, which we will need another time in another context later on. We obtain:

$$a_{\alpha\rho}\, dP_{i\alpha} = d\pi_{i\rho} - P_{i\alpha}\, da_{\beta\rho} = d\pi_{i\rho} - P_{i\rho}\, df + P_{i\beta}\, \pi_{j\rho}\, dp_{j\beta} + P_{i\beta}\, p_{k\rho}\, d\pi_{k\rho} .$$

From (7.h1) and (7.h4), one can thus also write:

$$a_{\alpha\rho}\, dP_{i\alpha} = h_{ik}\{d\pi_{k\rho} - \frac{1}{f}(\pi_{k\rho}\, df - \pi_{j\rho}\, \pi_{k\beta}\, d\pi_{j\beta})\}. \tag{9.1}$$

We would like to differentiate this with respect to $p_{i\beta}$; hence, one has $df = \pi_{j\beta}\, dp_{j\beta}$. We next contract with $g_{\alpha\beta}$ and obtain, due to (7.g2) and (5.1):

$$\frac{\partial P_{i\alpha}}{\partial p_{j\beta}} = \frac{h_{ik}\, g_{\alpha\rho}}{f} q_{k\rho,\, j\beta} .$$

The determinant of this $\mu \cdot n$-rowed matrix is:

$$\left| \frac{\partial P_{i\alpha}}{\partial p_{j\beta}} \right| = F^{n\,\mu} \,|\, q_{i\alpha,\, j\beta} \,| \quad ^{1}). \tag{9.2}$$

This is certainly non-zero when the Legendre condition is satisfied. In this case, there therefore always exists the possibility of introducing canonical variables.

Had we started from (7.A2) instead of (7.a2) then we would have found:

---

[1]) One sees that the determinant of the $\mu + n$-rowed matrix $h_{ij}\, g_{\alpha\beta}$ has the value $h^{\mu}\, g^{n}$ as follows: We have, as we easily see, $|\, h_{ij}\, \delta_{\alpha\beta}\, | = h^{\mu}$ and $|\, \delta_{ij}\, g_{\alpha\beta} \,| = g^{n}$, and one has $h_{ij}\, g_{\alpha\beta} = h_{ir}\, \delta_{\alpha\rho}\, \delta_{rj}\, g_{\alpha\beta}$.

$$\frac{\partial p_{i\alpha}}{\partial P_{j\beta}} = \frac{h_{ki} g_{\rho\alpha}}{F} Q_{k\rho, j\beta} .$$

The matrix that appears on the right-hand side is thus the reciprocal of the one that we previously wrote down, and for the determinant, one has:

$$1 = f^{n\mu} F^{n\mu} \mid q_{i\alpha,j\beta} \mid \mid Q_{i\alpha,j\beta} \mid .$$

10. The method of Lagrange provides the Euler equations for the extremals of our variational problem:

(10.1) $$\frac{d\pi_{i\rho}}{dt_\rho} - f_{x_i} = 0 .$$

One then easily calculates that a surface that intersects a geodesic field satisfies these equations [1]). Under general assumptions, one has an identity that makes understanding the connection between fields of extremals and geodesic fields clear, and which we would now like to derive.

We assume that we are given the quantities $p_{i\alpha}$, $f(x_i, t_\alpha, p_{i\alpha})$, and $P_{i\alpha}$, $F(x_i, t_\alpha, P_{i\alpha})$ as functions of space, and that the relations that were written down in section 7 are valid between them. We then use equation (9.1) in order to obtain a relation in which the quantities $d\pi_{i\rho} / dt_\rho$ appear:

$$a_{\alpha\rho} \frac{dP_{i\alpha}}{dt_\rho} = h_{ik} \left\{ \frac{d\pi_{k\rho}}{dt_\rho} - \frac{1}{f} \left(\pi_{k\rho} \frac{df}{dt_\beta} - \pi_{j\rho} \pi_{k\beta} \frac{dp_{j\beta}}{dt_\rho} \right) \right\} .$$

Here, we calculate df / dt_ρ, using (7.h4), and obtain:

$$a_{\alpha\rho} \frac{dP_{i\alpha}}{dt_\rho} + P_{i\rho} (f_{t_\rho} + p_{k\rho} f_{x_k}) = h_{ik} \left\{ \frac{d\pi_{k\rho}}{dt_\rho} - \frac{1}{f} \left(\pi_{k\rho} \pi_{j\beta} - \pi_{j\rho} \pi_{k\beta} \right) \frac{dp_{j\beta}}{dt_\rho} \right\} .$$

However, one has $P_{i\rho}\, p_{k\rho} = h_{ik} - \delta_{ik}$. Thus:

$$a_{\alpha\rho} \left(\frac{dP_{i\alpha}}{dt_\rho} + p_{j\rho} \frac{dP_{i\alpha}}{dx_i} \right) + P_{i\rho} f_{t_\rho} - f_{x_k}$$

---

[1]) One has $\frac{d\pi_{i\rho}}{dt_\rho} = \frac{d}{dt_\rho} (S_{\lambda i} \bar{c}_{\lambda\rho})$. Since $c_{\lambda\beta} = \frac{dS_\lambda}{dt_\rho}$, the divergence $\frac{d\bar{c}_\lambda}{dt_\rho}$ vanishes; hence one has $\frac{d\pi_{i\rho}}{dt_\rho} = \bar{c}_{\lambda\rho} \frac{dS_{\lambda i}}{dt_\rho}$. However, if one has $\frac{dS_{\lambda i}}{dt_\rho} = \frac{\partial c_\lambda}{\partial x_i} - S_{\lambda j} \frac{\partial p_j}{\partial x_i}$ then $\frac{d\pi_{i\rho}}{dt_\rho} = \frac{\partial \Delta}{\partial x_i} - \pi_{j\rho} \frac{\partial p_{j\rho}}{\partial x_i} = \frac{\partial f}{\partial x_i} - \pi_{j\rho} \frac{\partial p_{j\rho}}{\partial x_i} = f_{x_i}$.

$$= h_{ik} \left\{ \frac{d\pi_{k\rho}}{dt_\rho} - f_{x_k} + \frac{\pi_{k\rho}\pi_{j\beta}}{f} \left(\frac{dp_{j\rho}}{dt_\sigma} - \frac{dp_{j\sigma}}{dt_\rho} \right) \right\}.$$

(In the last term on the right, we implicitly made a renaming of the summation indices.) Now, we introduce the upper case notation everywhere on the left, in which we use (7.***), (7.a4) with (7.a3), and (7.**). The left-hand side becomes:

$$\frac{f}{F} \left\{ A_{\rho\alpha} \frac{\partial P_{i\alpha}}{\partial t_\rho} + \Pi_{j\alpha} \frac{\partial P_{i\alpha}}{\partial x_j} + F_{x_i} - P_{i\rho} F_{t_\rho} \right\}$$

$$= \frac{f}{F} \left\{ F \frac{\partial P_{i\alpha}}{\partial t_\alpha} + \Pi_{j\alpha} \left(\frac{\partial P_{i\alpha}}{\partial x_j} - P_{j\rho} \frac{\partial P_{i\alpha}}{\partial t_\rho} \right) + F_{x_i} - P_{i\rho} F_{t_\rho} \right\},$$

and with the notations (6.3) and (6.5):

$$\frac{f}{F} \left\{ \frac{dF}{dx_i} + F \frac{\partial P_{i\alpha}}{\partial t_\alpha} + \Pi_{j\alpha} \left(\frac{\partial P_{i\alpha}}{\partial x_j} - \frac{\partial P_{j\alpha}}{\partial x_i} \right) \right\} = \frac{f}{F} \left\{ \frac{dF}{dx_i} + F \frac{\partial P_{i\alpha}}{\partial t_\alpha} + \Pi_{j\alpha} [i\ j\ \alpha] \right\}.$$

With:

(10.2) $$[i] = \frac{dF}{dx_i} + F \frac{\partial P_{i\alpha}}{\partial t_\alpha} + \Pi_{j\alpha} [i\ j\ \alpha]$$

we obtain the desired identity when we contract (7.h2) with b_{ij} :

(10.3) $$\left\{ \frac{d\pi_{k\rho}}{dt_\rho} - f_{x_k} \right\} + \frac{\pi_{k\rho}\pi_{j\alpha}}{f} \left(\frac{dp_{j\rho}}{dt_\sigma} - \frac{dp_{j\sigma}}{dt_\rho} \right) = \frac{b_{ik}}{F} [i].$$

In a geodesic field (6.4) and (6.6) are true, hence, $[i] = 0$. If the geodesic field intersects a surface transversally then this is true on the surface (1.4), and thus it follows from (10.3) that such a surface satisfies the Euler equations (10.1).

Conversely, if one has field of extremals, i.e., of solutions to the Euler equations, then it follows from (10.3) that $[i] = 0$, however, in order for a geodesic field to be present one must have that all $[i\ j\ \alpha] = 0$, from which (6.6) follows. In this case, one has what Carathéodory called a *complete figure* and what the Americans call a *Mayer field* (in the case of a simple integral).

There is no method of embedding a given extremal in a field of extremals that defines a complete figure. We have, however, already seen that this is completely unnecessary. One needs no field of extremals whatsoever to apply the Weierstrass theory. In order to construct a geodesic field, it is completely sufficient that the given extremals (and possibly no others) intersect transversally. In the next chapter we will show how one can make this happen; thus, it will likewise be proved that the extremals are solutions of the variational problem.

11. Now, a word about transversality. For a simple problem in the calculus of variations transversality ordinarily enters in connection with the boundary conditions, and indeed only when the boundary of the given curve or surface is not assumed to be fixed. Also, for the aforementioned general problem one easily arrives at such transversality conditions by the theory of geodesic fields. One needs only to show that the integral $\int \Delta' dt$ has the same value for all "permissible" comparison surfaces, or, what amounts to the same thing, that all such surfaces map the region G_t that is determined by its boundary to the same region G_λ [1]). In case the boundary is assumed to be fixed, this is, as we have seen, unavoidably the case. We assume, as an example, that a part of the – usually assumed fixed – boundary of a ν-dimensional manifold $\mathfrak{H}$ ($\nu \geq \mu$) can move freely! A surface $S_\alpha = \lambda_\alpha$ goes through each boundary point. In order for our condition to be satisfied, all of the boundary points of all of the permissible comparison surfaces must necessarily lie on these same surface $S_\alpha = \lambda_\alpha$. Therefore, the manifold $\mathfrak{H}$ must be contained in the $n + \mu - 1$-dimensional manifold that is defined by these surfaces.

On thus obtains the following transversality condition: If a part of the boundary of the desired surface moves freely on a n-dimensional manifold ($\nu \geq \mu$) then any solution of the problem on this manifold must intersect it transversally. Thus, an $m + \mu - 1$-dimensional surface element ($1 \leq m \leq n$) that has $\mu - 1$ directions in common with the extremal is called transversal to it when it contains precisely m linearly independent transversal directions (i.e., ones that are contained in the previously-defined n-dimensional transversal surface element).

In particular, an $n + \mu - 1$-dimensional that is represented by the equation $S(x_i, t_\alpha) =$ const is called transversal to an extremal when the equations:

$$S_{x_i} = S_{t_\rho} P_{i\rho} \tag{11.1}$$

are satisfied.

Second Chapter

Construction of a geodesic field that intersects a given extremal transversally

12. Carathéodory has remarked that in order for one to obtain any geodesic field one has to solve just one partial differential equation of the first order. The equations that one has to satisfy indeed read like:

$$S_{\alpha i} = P_{i\rho} S_{\alpha\rho} , \tag{12.1}$$

$$[S_{\alpha\beta} | \cdot F(x_i, t_\alpha, P_{i\alpha}) = 1. \tag{12.2}$$

One merely needs to solve (12.1) for $P_{i\alpha}$ and substitute the result in (12.2). One can therefore choose $\mu - 1$ of the functions S_α arbitrarily, and then (12.2) is a partial differential equation of first order for the μ^{th} function.

In order to examine this closer, we introduce some new notations. We denote an index that ranges through the numbers from 2 to μ with a primed Greek notation: α', β', etc. The index 1 that will frequently appear is always a Greek one: $\alpha = 1$.

---

[1]) Cf., section 2.

We choose the functions $S_{\alpha'}$ arbitrarily and denote the derivatives of the desired function $S_1 = \sigma(x_i, t_\alpha)$ by:

$$S_{11} = \sigma_{11}, \qquad S_{1\alpha} = \sigma_\alpha .$$

All of the elements in the matrix $S_{\alpha\beta}$, except for the ones in the first row – hence, the sub-determinant $\bar{S}_{1\alpha}$ in particular, as well – are now known functions of x_i and t_α. Since we would like to satisfy (12.2) and since one has:

$$(12.3) \qquad | S_{\alpha\beta} | = \sigma_\alpha \bar{S}_{1\alpha} ,$$

at least one of these sub-determinants must be everywhere non-zero. We assume that in the region that we consider one has:

$$(12.4) \qquad \bar{S}_{11} \neq 0 .$$

By solving (12.1) for $P_{i\alpha}$ and introducing the known functions $S_{\alpha'1}$, $S_{\alpha'\alpha}$, one obtains the $P_{i\alpha}$ as functions of x_i, t_α, σ_i, σ_α, which we will denote by:

$$(12.5) \qquad (P_{i\alpha}) = P_{i\alpha}(x_i , t_\alpha, \sigma_i , \sigma_\alpha) .$$

Correspondingly, let $(F) = F(x_i , t_\alpha, (P_{i\alpha}))$ and $(\Pi_{i\alpha}) = F_{P_{i\alpha}}(x_i, t_\alpha, (P_{i\alpha}))$.

Finally, we set:

$$(12.6) \qquad \sigma_\alpha \cdot \bar{S}_{1\alpha} \cdot (F) - 1 = M(x_i , t_\alpha, \sigma_i , \sigma_\alpha),$$

and then our partial differential equation (12.2) becomes:

$$(12.7) \qquad M(x_i , t_\alpha, \sigma_i , \sigma_\alpha) = 0 .$$

13. Recall the following facts from the theory of characteristics for first order partial differential equations: One can give the function σ on an $n + \mu - 1$-dimensional manifold (“hypersurface”):

$$(13.1) \qquad t_1 = \tau(x_i, t_{\alpha'})$$

arbitrarily:

$$(13.2) \qquad \sigma(x_i, \tau, t_{\alpha'}) = \Sigma(x_i, t_{\alpha'}).$$

From (13.2), it follows by differentiation, when we denote the derivatives of τ and Σ simply by the symbols i and α':

$$(13.3) \qquad \begin{aligned} \sigma_i &= \Sigma_i = \sigma_1 \tau_i , \\ \sigma_{\alpha'} &= \Sigma_{\alpha'} = \sigma_1 \tau_{\alpha'} . \end{aligned}$$

If one substitutes this in M then (12.7) can serve for the calculation of σ_1 in the event that the derivative with respect to σ_1 is non-zero:

$$M_{\sigma_1} - M_{\sigma_{\alpha'}} \tau_{\alpha'} - M_{\sigma_i} \tau_i \neq 0. \tag{13.4}$$

One then calculates σ_i and $\sigma_{\alpha'}$ from (13.3). With the initial values for for σ_i and $\sigma_{\alpha'}$ thus obtained, we integrate the equations of the characteristics:

$$\dot{t}_\alpha = M_{\sigma_\alpha}, \qquad \dot{x}_i = M_{\sigma_1}, \tag{13.5}$$

$$\dot{\sigma}_\alpha = -M_{t_\alpha}, \qquad \dot{\sigma}_i = -M_{x_i} \tag{13.6}$$ [1]).

By this, one obtains σ_i and $\sigma_{\alpha'}$ as functions in space and from them, by a quadrature, the desired function $\sigma = S_1$ and thereby, the geodesic field.

Instead of (13.4), one can now also write:

$$\dot{t}_1 - \dot{t}_{\alpha'} \tau_{\alpha'} - \dot{x}_i \tau_i \neq 0,$$

i.e., the characteristic curves shall not touch the surface (13.1). It will therefore cover a certain region of space simply.

One can read the proof that the functions σ_i, σ_α that we found satisfy the required integrability conditions in Carathéodory [2]).

14. Our aim is to show that these methods provide a geodesic field that intersects a given extremal transversally, in the case where one is given the functions $S_{\alpha'}$ and the initial values for S_1 (on a hypersurface that the extremal goes through) in a certain way.

For this, we must examine the characteristic equations closer. We begin with (13.5). From (12.6), these equations read:

$$\begin{aligned} \dot{t}_\alpha &= (F)\bar{S}_{1\alpha} + \sigma_\gamma \bar{S}_{1\gamma} (\Pi_{j\beta}) \frac{\partial (P_{i\alpha})}{\partial \sigma_\alpha}, \\ \dot{x}_1 &= \sigma_\gamma \bar{S}_{1\gamma} (\Pi_{j\beta}) \frac{\partial (P_{i\alpha})}{\partial \sigma_i}. \end{aligned} \tag{14.1}$$

The functions $(P_{j\alpha})$ are obtained by solving the equations:

$$\sigma_j = P_{j\rho}\, \sigma_\rho, \tag{14.2}$$

$$S_{\alpha' j} = P_{j\rho}\, S_{\alpha'\rho}, \tag{14.3}$$

for $P_{j\alpha}$. We differentiate these equations with respect to σ_α:

$$-P_{j\alpha} = \frac{\partial (P_{j\rho})}{\partial \sigma_\alpha} \sigma_\rho,$$

---

[1]) The dot means the derivative with respect to a parameter τ by which we represent the characteristics.

[2]) Variationsrechnung und partielle Differentialgleichungen erster Ordnung, Chapter 3.

$$0 = \frac{\partial(P_j)}{\partial \sigma_\alpha} S_{\alpha'\rho},$$

and obtain, by solving for the differential quotients:

$$\sigma_\gamma \bar{S}_{1\gamma} \frac{\partial(P_{j\beta})}{\partial \sigma_\alpha} = -(P_{j\alpha})\bar{S}_{1\beta}. \tag{14.4}$$

In the same way, one obtains:

$$\begin{aligned} \dot{t}_\alpha &= (F)\bar{S}_{1\alpha} - (P_{j\alpha})(\Pi_{j\beta})\bar{S}_{1\beta} = (A_{\alpha\beta})\bar{S}_{1\beta}, \\ \dot{x}_t &= (\Pi_{i\beta})\bar{S}_{1\beta}. \end{aligned} \tag{14.6}$$

15. For the following section we give σ_1 and σ_α to be arbitrary functions in space, between which the equation (12.7) exists. Likewise, functions in space originate from (12.5) that we denote by $P_{i\alpha}^*$, F^*, etc.

One can also first give the functions $P_{i\alpha}^*$. They must satisfy only the equations (14.3):

$$S_{\alpha' i} = P_{i\rho}^* S_{\alpha'\rho}. \tag{15.1}$$

There are still $\mu - 1$ of the σ that can be chosen entirely arbitrarily; e.g., $\sigma_{\alpha'}$. One always obtains σ_1 uniquely on account of (12.4) from:

$$\sigma_\gamma \bar{S}_{1\gamma} \cdot F^* = 1 \tag{15.2}$$

(viz., (12.7)), and then σ_i from:

$$\sigma_i = P_{i\rho}^* \sigma_\rho \tag{15.3}$$

(viz., (14.2)). The relation (15.2) allows us to replace $\sigma_\gamma \bar{S}_{1\gamma}$ with $1/F^*$.

The equations:

$$\begin{aligned} \dot{t}_\alpha &= A_{\alpha\beta}^* \bar{S}_{i\beta}, \\ \dot{x}_i &= \Pi_{i\beta}^* \bar{S}_{1\beta} \end{aligned} \tag{15.4}$$

(viz., (14.6)) define a particular family of curves in space. It is noteworthy that this family, when one proceeds in the manner that was previously described, does not depend upon the choice of the functions $\sigma_{\alpha'}$, at all. It is easy to clarify the geometric meaning of the curves. With the use of (14.3) and (7.A.1), one computes:

$$S_{\alpha' i}\dot{x}_i + S_{\alpha'\alpha}\dot{t}_\alpha = 0, \tag{15.5}$$

and when one lets $p^*_{i\alpha}$ denote the μ-dimensional surface element that is associated with $P^*_{i\alpha}$ by means of the Legendre transformation one finds, due to (7.A2):

(15.6) $$\dot{x}_i = p^*_{i\alpha}\dot{t}_\alpha .$$

The curves under consideration thus lie on the surfaces $S_{\alpha'} = \lambda_{\alpha'}$ and contact the surface elements $p^*_{i\alpha}$. In the case where $p^*_{i\alpha}$ belongs to a family of μ-dimensional surfaces that intersect the $(n + 1)$-dimensional surfaces $S_{\alpha'} = \lambda_{\alpha'}$, they are simply the intersection curves of these two families of surfaces.

16. Under the assumptions that were introduced in the last section, we now come to grips with the other half (13.6) of the characteristic equations. Due to (14.4) and (14.5), one has:

$$\begin{aligned}\frac{\partial P^*_{j\beta}}{\partial t_\alpha} &= \frac{\partial(P_{j\beta})}{\partial t_\alpha} + \frac{\partial(P_{j\beta})}{\partial \sigma_i}\frac{\partial \sigma_i}{\partial t_\alpha} + \frac{\partial(P_{j\beta})}{\partial \sigma_\rho}\frac{\partial \sigma_\rho}{\partial t_\alpha} \\ &= \frac{\partial(P_{j\beta})}{\partial t_\alpha} + F^*\overline{S}_{1\beta}\left(\frac{\partial \sigma_j}{\partial t_\alpha} - P^*_{j\rho}\frac{\partial \sigma_\rho}{\partial t_\alpha}\right)\end{aligned}$$

(one always observes (15.2)!), or, when one takes into consideration equation (15.3), differentiated with respect to t_α:

(16.1) $$\frac{\partial(P_{j\beta})}{\partial t_\alpha} = \frac{\partial P^*_{j\beta}}{\partial t_\alpha} - F^*\overline{S}_{1\beta}\sigma_\rho\frac{\partial P^*_{j\rho}}{\partial t_\alpha} .$$

One obtains precisely:

$$\frac{\partial(P_{j\beta})}{\partial x_i} = \frac{\partial P^*_{j\beta}}{\partial x_i} - F^*\overline{S}_{1\beta}\sigma_\rho\frac{\partial P^*_{j\rho}}{\partial x_i} ,$$

and therefore also, with the operator (6.3), in which we take $P_{i\rho} = P^*_{i\beta}$:

(16.2) $$\frac{d(P_{j\beta})}{dx_i} = \frac{dP^*_{j\beta}}{dx_i} - F^*\overline{S}_{1\beta}\sigma_\rho\frac{dP^*_{j\rho}}{dx_i} .$$

With the help of (16.1), we calculate $M^*_{t_\alpha}$ (i.e., the derivative of the function (12.6) with respect to t_α, in which we then substitute the given functions σ_i and σ_α) and obtain:

(16.3) $$\begin{aligned}M^*_{t_\alpha} &= F^*\sigma_\beta\frac{\partial \overline{S}_{i\beta}}{\partial t_\alpha} + \frac{1}{F^*}F^*_{t_\alpha} + \frac{\Pi^*_{j\beta}}{F^*}\left(\frac{\partial P^*_{j\beta}}{\partial t_\alpha} - F^*\overline{S}_{i\beta}\frac{\partial P^*_{j\rho}}{\partial t_\alpha}\right) \\ &= F^*\sigma_\beta\frac{\partial \overline{S}_{i\beta}}{\partial t_\alpha} + \frac{1}{F^*}F^*_{t_\alpha} - \dot{x}_j\sigma_\beta\frac{\partial P^*_{j\beta}}{\partial t_\alpha};\end{aligned}$$

here, we have used (15.4). In the same way, one computes, with (16.2):

(16.4) $$M^*_{x_i} - P^*_{i\rho} M^*_{t_\rho} = F^* \sigma_\beta \frac{d\bar{S}_{1\beta}}{dx_i} + \frac{1}{F^*}\frac{dF^*}{dx_i} - \dot{x}_j \sigma_\beta \frac{dP^*_{j\beta}}{dx_i}.$$

Now, we must carry out the differentiation along the curves (15.4). If $\Psi(x_i, t_\alpha)$ is any function in space then, from (15.4), one has:

$$\dot{\Psi} = \frac{\partial\Psi}{\partial x_i}\dot{x}_i + \frac{\partial\Psi}{\partial t_\alpha}\dot{t}_\alpha = \frac{\partial\Psi}{\partial x_i}\dot{x}_i + \frac{\partial\Psi}{\partial t_\alpha}F^* S_{1\alpha} - \frac{\partial\Psi}{\partial t_\alpha}P^*_{i\alpha}\dot{x}_i$$

or:

(16.5) $$\dot{\Psi} = \dot{x}_i \frac{d\Psi}{dx_i} + F^* \bar{S}_{1\alpha}\frac{d\Psi}{dt_\alpha}.$$

A further auxiliary formula gives the derivatives of $\bar{S}_{1\beta}$. With the help of (15.1), one obtains:

$$\frac{dS_{\alpha'\alpha}}{dx_i} = \frac{\partial S_{\alpha'\rho}}{\partial x_i} - P^*_{i\sigma}\frac{\partial S_{\alpha'\rho}}{\partial t_\sigma} = \frac{\partial S_{\alpha' i}}{\partial t_\rho} - P^*_{i\sigma}\frac{\partial S_{\alpha'\sigma}}{\partial t_\rho} = S_{\alpha'\sigma}\frac{\partial P^*_{i\sigma}}{\partial t_\rho},$$

and therefore the formula (3.4) gives us:

(16.6) $$\frac{d\bar{S}_{1\beta}}{dx_i} = \bar{S}_{1\beta}\frac{\partial P^*_{i\rho}}{\partial t_\rho} - \bar{S}_{1\rho}\frac{\partial P^*_{i\beta}}{\partial t_\rho}.$$

in which, since $\alpha = 1$, we may use a primed index for λ.

We go on to the differentiation of equations (15.2) and (15.3) along the curves (15.4), which we write as:

(16.7) $$\dot{\sigma}_\alpha \bar{S}_{1\alpha} = -\frac{\dot{F}^*}{F^{*2}} - \sigma_\beta \dot{\bar{S}}_{1\beta}$$

and:

(16.8) $$\dot{\sigma}_i - P^*_{i\rho}\dot{\sigma}_\rho = \sigma_\beta \dot{P}^*_{i\beta}.$$

From (16.3) and (16.7), we calculate the quantities $\bar{S}_{1\alpha}(\dot{\sigma}_\alpha + M^*_{t_\alpha})$ and then use formula (16.5), first for $\Psi \equiv F^*$, and then for $\Psi \equiv \bar{S}_{1\beta}$, and then formula (16.6). Most of them go away, and what remains is:

(16.9) $$\bar{S}_{1\alpha}(\dot{\sigma}_\alpha + M^*_{t_\alpha}) = -\frac{\dot{x}_j}{F^{*2}}\left(\frac{dF^*}{dx_j} + F^*\frac{\partial P^*_{j\alpha}}{\partial t_\alpha}\right).$$

Likewise, we calculate the expression $\dot{\sigma}_i + M^*_{x_i} - P^*_{i\rho}(\dot{\sigma}_\rho + M^*_{t_\rho})$ from (16.4) and (16.8), where we again use (16.6) and (16.5) for $\Psi = P^*_{i\beta}$. One obtains:

$$\begin{aligned}\dot{\sigma}_i &+ M^*_{x_i} - P^*_{i\rho}(\dot{\sigma}_\rho + M^*_{t_\rho}) \\ &= \frac{1}{F^*}\left(\frac{dF^*}{dx_i} + F^*\frac{\partial P^*_{i\rho}}{\partial t_\rho}\right) + \dot{x}_j\sigma_\beta\left(\frac{dP^*_{i\beta}}{dx_j} - \frac{\partial P^*_{j\rho}}{\partial x_i}\right) \\ &= \frac{1}{F^*}\left(\frac{dF^*}{dx_i} + F^*\frac{\partial P^*_{i\rho}}{\partial t_\rho}\right) + \Pi^*_{j\alpha}S_{1\alpha}\sigma_\beta[i\,j\,\beta]^*.\end{aligned} \tag{16.10}$$

In order to further put this into another form, we remark that due to (15.1) a close connection exists between $\bar{S}_{1\alpha}$ and $[i\,j\alpha]^*$. Namely, from (15.1), it follows that:

$$\frac{dS_{\alpha' i}}{dx_j} - \frac{dS_{\alpha' j}}{dx_i} = P^*_{i\rho}\frac{dS_{\alpha'\rho}}{dx_j} - P^*_{j\rho}\frac{dS_{\alpha'\rho}}{dx_i} + [i\,j\,\rho]^* S_{\alpha'\rho}.$$

One only needs to write out the differentiations and use the facts that $S_{\alpha' i}$ and $S_{\alpha'\rho}$ are the derivatives of $S_{\alpha'}$ with respect to x_i and t_ρ ; everything else goes away, and what remains is:

$$[i\,j\,\rho]^* S_{\alpha'\rho} = 0.$$

These $\mu - 1$ equations for $[i\ j\ \rho]^*$ are the same ones that the quantities $\bar{S}_{1\rho}$ satisfy. From this, it follows that for any pair of numbers i, j the μ numbers $[i\ j\ \rho]^*$ are proportional to $\bar{S}_{1\rho}$, which one can write as:

$$[i\,j\,\alpha]^*\bar{S}_{1\beta} = [i\,j\,\beta]^*\bar{S}_{1\alpha}. \tag{16.11}$$

One thus ultimately finds from (16.10), with the use of the abbreviation (10.2):

$$\dot{\sigma}_i + M^*_{x_i} - P^*_{i\rho}(\dot{\sigma}_\rho + M^*_{t_\rho}) = \frac{[i]^*}{F^*}. \tag{16.12}$$

17. Let there be given an extremal, i.e., a surface (1.1) that satisfies the Euler differential equations (10.1). The $p_{i\alpha}$ for the surface follow from (1.2). Let the determinant $|\,q_{i\alpha,\,j\beta}\,|$ (cf., (9.2)) be non-zero, such that one can calculate $P_{i\alpha}$ and the function $F(x_i, t_\alpha, P_{i\alpha})$ on the surface from (6.9) and (6.7). We shall show that one can choose the functions $S_{\alpha'}$ and the functions σ_1, σ_α that were introduced in section 15, in such a way that one has $P^*_{i\alpha} = P_{i\alpha}$ on the extremal.

The $S_{\alpha'}$ obviously need to satisfy, in addition to the inequality (12.4), only equations (14.3) on the extremal; i.e., the surface $S_{\alpha'} = \lambda_{\alpha'}$ must intersect the extremal transversally. From this, one can conclude that it is always possible to give such functions [1]). Now, in order to also find suitable functions σ we again first give ourselves the $P^*_{i\alpha}$. These functions must satisfy (15.1) everywhere and agree with the $P_{i\alpha}$ on the extremals (which indeed satisfy the same equation). The $\sigma_{\alpha'}$ remain completely arbitrary, as before, and the remaining σ will be calculated from them, as in the previous section.

Once we have determined all the functions in this way, we consider the μ – 1-parameter family of curves, in which the surfaces $S_{\alpha'} = \lambda_{\alpha'}$ go through our extremal. Due to (15.5) and (15.6), these curves belong to the family of curves that we considered in the previous section, and can therefore be represented by functions:

(17.1) $$x_i(\tau, u_1, \ldots, u_{\mu-1}), t_\alpha(\tau, u_1, \ldots, u_{\mu-1})$$

that satisfy equations (15.4). Now, we determine the $\mu - 1$ arbitrary functions $\sigma_{\alpha'}$ in such a way that they satisfy the $\mu - 1$ differential equations:

(17.2) $$\dot{\sigma}_{\alpha'} + M^{\dagger}_{t_{\alpha'}} = 0 ;$$

---

[1]) We show this by actually writing down such functions. Through each point $\overset{0}{t}_\alpha$, $\overset{0}{x}_i = x_i(\overset{0}{t}_\alpha)$ of the extremal, we pass an *n*-dimensional plane that is transversal to it; it will be given by the equations:

(1) $$t_\alpha = t_\alpha(x_i, \overset{0}{t}_\alpha) = \overset{0}{t}_\alpha - P_{i\alpha}(x_i - \overset{0}{x}_i).$$

A certain neighborhood of the extremal will covered by these planes simply. If the derivative of the functions (1) with respect to $\overset{0}{t}_\beta$ is $\delta_{\alpha\beta} + P_{i\alpha}\, p_{i\beta} = g_{\alpha\beta}$ the functional determinant g is, from section 6, non-zero. For that reason, one can solve equations (1) for $\overset{0}{t}_\alpha$ and obtain certain functions $\overset{0}{t}_\alpha(x_i, t_\beta)$.

Now, we choose $\mu - 1$ arbitrary functions $s_{\alpha'}(\overset{0}{t}_\alpha)$, and determine the functions $S_{\alpha'}$ in such a way that they have the following values:

(2) $$S_{\alpha'}(x_i(\overset{0}{t}_\alpha), \overset{0}{t}_\alpha) = s_{\alpha'}(\overset{0}{t}_\alpha),$$

on the extremal, and are constant on any plane; i.e., we set:

$$S_{\alpha'}(x_i, t_\alpha) = s_{\alpha'}(\overset{0}{t}_\alpha(x_i, t_\beta)).$$

These functions satisfy (14.3) on the extremals. Therefore, by differentiating (2), with the use of the prior abbreviations, one obtains the relation:

$$S_{\alpha'\rho}\, g_{\rho\beta} = c_{\alpha'\beta}$$

(here, the $c_{\alpha'\beta}$ are the derivatives of $s_{\alpha'}$), which allows us to calculate the derivatives of the S_α on the extremals from those of the $s_{\alpha'}$. For $\bar{S}_{11}$, one finds:

$$\bar{S}_{11} = \frac{1}{g}\, g_{1\alpha}\, \bar{c}_{1\alpha}.$$

The $s_{\alpha'}$ must therefore be chosen in such a way that this is non-zero; (12.4) is then satisfied on the extremal and thus in a neighborhood of it.

here, the new notation means that in M_{t_α} one expresses σ_1 and σ_i in terms of $\sigma_{\alpha'}$ by means of (15.2) and (15.3), and shall thus substitute the functions (17.1). The values that the $\sigma_{\alpha'}$ assume outside of the extremal are immaterial.

We likewise write the system of functions σ_i, $\sigma_{\alpha'}$ thus obtained in the form:

(17.3) $$\sigma_i\,(\tau, u_1, \ldots, u_{\mu-1}), \qquad \sigma_\alpha\,(\tau, u_1, \ldots, u_{\mu-1}).$$

These functions, together with (17.1), *define a* ($\mu - 1$)*-parameter family of characteristics of the differential equation* $M = 0$.

In fact, one has $(P_{i\alpha}) = P^*_{i\alpha} = P_{i\alpha}$ on the extremal. Thus, the functions (17.1) satisfy not only (15.4) (as we said), but also (17.3), together with (14.6), i.e., (13.5). Furthermore, due to (17.2), one has:

(17.4) $$\dot{\sigma}_{\alpha'} + M^*_{t_{\alpha'}} = 0.$$

Now, we consider the identity (16.9). Due to (10.3), all $[j]^* = 0$ on the extremal (here, we use the fact that it can be treated around an extremal), which one can write (cf., (10.2)):

$$\frac{dF^*}{dx_j} + F^*\frac{\partial P^*_{j\alpha}}{\partial t_\alpha} = -\,\Pi^*_{k\beta}[j\,k\,\beta]^*.$$

Thus, for the right-hand side of (16.9) we can write:

$$\frac{\dot{x}_j}{F^{*2}}\Pi^*_{k\beta}[j\,k\,\beta]^*,$$

or also, due to (16.11) and because $\bar{S}_{11} \neq 0$:

$$\frac{\dot{x}_j}{F^{*2}}\frac{\bar{S}_{1\beta}}{\bar{S}_{11}}\Pi^*_{k\beta}[j\,k\,1]^* = \frac{1}{F^{*2}\bar{S}_{11}}\,\dot{x}_j\dot{x}_k[j\,k\,1]^*.$$

This is, however, null because the coefficient of $[j\ k\ 1]^*$ is symmetric in j and k, but $[j\ k\ 1]^*$ itself is anti-symmetric. The left-hand side of (16.9) thus vanishes on the extremal and from this it follows, from (17.4):

(17.5) $$\dot{\sigma}_1 + M^*_{t_1} = 0.$$

Now, we consider the identity (16.12). Here, the right-hand side also vanishes on the extremal, and since all $\dot{\sigma}_\alpha + M^*_{t_\alpha} = 0$ there already, one must also have:

(17.6) $$\dot{\sigma}_1 + M^*_{x_1} = 0\,.$$

Since $(P_{i\alpha}) = P^*_{i\alpha}$, we can write (17.4) to (17.6) in just the same way without stars. In fact, all of the equations (13.6) are also satisfied, which was to be proved.

18. A characteristic is uniquely determined by the values of σ_i and σ_α at one of its points. Under the assumption that the $S_{\alpha'}$ satisfy (14.3) on the extremals, we have thus proved the following:

At a point of the extremal one determines the σ in such a way that $(P_{i\alpha}) = P_{i\alpha}$ [1]), and integrates the characteristic differential equations with these initial values. *All of the characteristics thus determined remain on the extremal and one has* $(P_{i\alpha}) = P_{i\alpha}$ *along the entire curve.*

We would like to see whether one can choose the initial values that were spoken of in section 13 in such a way that all characteristics that begin on the extremal have these properties.

We deduce nothing else from the hypersurface (13.1) – it is called $\mathfrak{F}$ – except for the fact that it goes through the extremal and the curves $S_{\alpha'} = \lambda_{\alpha'}$ on it (which will indeed be characteristics). The intersection is a $(\mu - 1)$-dimensional manifold $\bar{\mathfrak{F}}$:

(18.1) $$t_1 = \tilde{\tau}(t_{\alpha'}), \qquad x_i = \xi_i(t_{\alpha'}) .$$

This manifold thus lies on $\mathfrak{F}$:

(18.2) $$\tau(\xi_i(t_{\alpha'}), t_\beta) = \tilde{\tau}(t_{\alpha'}),$$

and on the extremal:

(18.3) $$x_i\,(\tilde{\tau}(t_{\alpha'}), t_\beta) = \xi_i(t_{\alpha'});$$

here, the functions (1.1) appear on the left-hand side.

On this manifold $\bar{\mathfrak{F}}$, we must likewise satisfy, along with (13.3), the equations:

(18.4) $$\sigma_\alpha \bar{S}_{1\alpha} = \frac{1}{F}$$

and:

(18.5) $$\sigma_i = P_{i\rho}\,\sigma_\rho$$

((12.7) is then also satisfied naturally). We can eliminate all σ from these $\mu + 2n$ equations, in total, between which there exist n relations between the Σ. The simplest expressions that one defines from (13.3) are the linear combinations $\Sigma_i - P_{i\alpha'}\,\Sigma_{\alpha'}$ and $\Sigma_{\alpha'}\bar{S}_{1\alpha'}$; by the use of (18.4) and (18.5), one obtains:

$$\Sigma_i - P_{i\alpha'}\,\Sigma_{\alpha'} = \sigma_i - P_{i\alpha'}\,\sigma_{\alpha'} + \sigma_1(\tau_i - P_{i\alpha'}\,\tau_{\alpha'}) = \sigma_1(P_{i1} + \tau_i - P_{i\alpha'}\,\tau_{\alpha'}),$$

---

[1]) In this, the $\sigma_{\alpha'}$ are still completely arbitrary at this point; in fact, we still do not have the integration of (17.2) for the initial values at our disposal.

$$\Sigma_{\alpha'}\bar{S}_{1\alpha'} = \sigma_{\alpha'}\bar{S}_{1\alpha'} + \sigma_1\tau_{\alpha'}\bar{S}_{1\alpha'} = \frac{1}{F} - \sigma_1(\bar{S}_{11} - \tau_{\alpha'}\bar{S}_{1\alpha'}),$$

and one then eliminates σ_1 from them [1]):

$$\left(\frac{1}{F} - \Sigma_{\alpha'}\bar{S}_{1\alpha'}\right)(P_{i1} + \tau_i - P_{i\beta'}\,\tau_{\beta'}) - (\Sigma_1 - P_{i\alpha'}\,\Sigma_{\alpha'})\ (\bar{S}_{11} - \tau_{\beta'}\bar{S}_{1\beta'}) = 0\,. \tag{18.6}$$

I assert: When the initial values Σ on the manifold $\bar{\mathfrak{F}}$ satisfy the n relations (18.6), there is always *one* system of solutions σ_i, σ_α of the equations (13.3), (12.7) that belongs to $P_{i\alpha}$, i.e., that satisfies (18.4) and (18.5).

This is almost self-explanatory. Since, from (18.4), we do not need to worry about (12.7), we have to concern ourselves – in contradiction to the general case in section 13 – with *linear* equations, and clearly need to verify that their determinant does not vanish. We would like to make it into something else, and in place of the $(\mu + n)$-rowed determinant, consider a $(\mu - 1)$-rowed one. Namely, we remark that one can already calculate all of the σ_i and σ_α on this manifold from their initial values on $\bar{\mathfrak{F}}$:

$$\sigma(\xi_i, \tilde{\tau}, t_{\alpha'}) = \Sigma(\xi_i,\, t_{\alpha'}) = \tilde{\Sigma}_{\alpha'}\,. \tag{18.7}$$

In fact, we know indeed that on it only the $\sigma_{\alpha'}$ are freely at one's disposal. We differentiate (18.7) with respect to the $t_{\alpha'}$:

$$\sigma_i\xi_{i\alpha'} + \sigma_1\tilde{\tau}_{\alpha'} + \sigma_{\alpha'} = \tilde{\Sigma}_{\alpha'}\,,$$

and substitute the values $p_{i1}\tilde{\tau}_{\alpha'} + p_{i\alpha'}$ that follow from (18.3) for the $\xi_{i\alpha'}$ in this equation, and for σ_i and σ_α, the expressions that are calculated from (18.4) and (18.5) in terms of $\sigma_{\alpha'}$. By the use of the abbreviation (7.g1), we obtain, from an easy conversion:

$$\frac{1}{\bar{S}_{11}}\left\{\bar{S}_{11}(g_{\beta'\alpha} + g_{\beta'1}\tilde{\tau}_{\alpha'} - \bar{S}_{1\beta'}(g_{1\alpha'} + g_{11}\tilde{\tau}_{\alpha'})\right\}\sigma_{\beta'} = \tilde{\Sigma}_{\alpha'} - \frac{1}{F\bar{S}_{11}}(g_{1\alpha'} + g_{11}\tilde{\tau}_{\alpha'})\,. \tag{18.6}$$

One can calculate $\sigma_{\alpha'}$ from these $\mu - 1$ equations in the event that their determinant is non-vanishing. In order to establish this, we contract its matrix with $S_{\rho\beta'}$ and use the relations $S_{1\beta}S_{\rho'\beta} = 0$ and $S_{\rho'\beta}\ g_{\beta\alpha'} = c_{\rho'\alpha'}$, in which the second one follows immediately from the first one when one uses (14.3) (cf., also rem. [23]). Contraction yields the matrix:

---

[1]) One easily convinces oneself which relations must appear in place of (18.6) when $\bar{S}_{11} - \tau_{\alpha'}\bar{S}_{1\alpha'}$ or all of the $P_{i1} + \tau_i - P_{i\alpha'}\,\tau_{\alpha'}$ vanish. One can conclude that one or the other condition enters in, as a geometric argument shows. Cf., below.

$$c_{\rho'\alpha'} + c_{\rho'1}\tilde{\tau}_{\alpha'} = \frac{dS_{\rho'}}{dt_{\alpha'}} + \frac{dS_{\rho'}}{dt_1}\tilde{\tau}_{\alpha'}.$$

Its determinant is nothing but the functional determinant of the $S_{\alpha'}$ with respect to the $t_{\beta'}$ on the surface (18.1), which is non-zero by our assumptions. The same is true for the determinant $|S_{\rho'\beta'}| = \bar{S}_{11}$, and thus one can calculate the desired determinant from both of them, and this is likewise non-vanishing.

One can choose the initial values $\tilde{\Sigma}$ on $\bar{\mathfrak{F}}$ arbitrarily. From (18.7), it follows that:

$$\Sigma_i\ \xi_{i\alpha'} + \Sigma_{\alpha'} = \tilde{\Sigma}_{\alpha'},$$

and these equations, together with (18.6), define a system of $u + \mu - 1$ equations for Σ_i and $\Sigma_{\alpha'}$. These equations will be satisfied for the values that follow from (13.3) once one has calculated σ from (18.8), (18.4), and (18.5).

The geometric meaning of the equations (18.6) is that of guaranteeing that the surface S_1 = const. intersects the extremal at the points in question. This is easy to see when one takes $\mathfrak{F}$ to be such a surface S_1 = const. and asks what sort of restriction one must then subject the orientation of $\mathfrak{F}$ to. One would thus like to choose Σ = const., so that all Σ_i and $\Sigma_{\alpha'}$ shall vanish, and it would thus follows from (18.6):

$$P_{i1} + \tau_i - P_{\beta'}\tau_{\beta'} = 0.$$

$\tau_i - 1, \tau_{\beta'}$ are, however, the components of the normal to $\mathfrak{F}$, so these equations then say nothing more than the fact that this hypersurface must be transversal to the extremal, as we expected (cf., 11.1).

We can summarize the result whose proof sections 14-18 were dedicated to in the following way:

In order for a given extremal to be embedded in a geodesic field, one must give functions $S_{\alpha'}(x_i, t_\alpha)$ such that the surface $S_{\alpha'} = \lambda_{\alpha'}$ is transversal to the extremal. One then gives the initial values for S_1 on a hypersurface that goes through the extremal, and on it, the curves $S_{\alpha'} = \lambda_{\alpha'}$, such that they satisfy the equations (18.6) *on the extremal, and integrate the partial differential equation* (12.7) *with these initial values. The geodesic field that is thus obtained intersects the extremal transversally.*

Third Chapter

Discontinuous solutions

19. Let two functions $F(x_i, t_\alpha, P_{i\alpha})$ and $F'(x_i, t_\alpha, P'_{i\alpha})$ be given. We would like to construct a geodesic field, on the one side of which, a hypersurface $\mathfrak{F}$ is given by the μ +

$n - 1$ parameters u_i , $u_{\alpha'}$, that belong to F and, on the other side, to F', in such a way that the functions S_α are continuous. That is, on $\mathfrak{F}$ the following equations shall be valid:

(19.1) $$S_\alpha = S'_\alpha ,$$

(19.2) $$S_{\alpha i} = P_{i\beta} S_{\alpha\beta} , \qquad S'_{\alpha 1} = P'_{i\beta} S'_{\alpha\beta} ,$$

(19.3) $$F \mid S_{\alpha\beta} \mid = 1, \qquad F' \mid S'_{\alpha\beta} \mid = 1.$$

From (19.1), it follows by differentiation with respect to the parameters that:

$$S_{\alpha j} \frac{\partial x_j}{\partial u_i} + S_{\alpha\beta} \frac{\partial t_\beta}{\partial u_i} = S'_{\alpha j} \frac{\partial x_j}{\partial u_i} + S'_{\alpha\beta} \frac{\partial t_\beta}{\partial u_i} ,$$
$$S_{\alpha j} \frac{\partial x_j}{\partial u_{\alpha'}} + S_{\alpha\beta} \frac{\partial t_\beta}{\partial u_{\alpha'}} = S'_{\alpha j} \frac{\partial x_j}{\partial u_{\alpha'}} + S'_{\alpha\beta} \frac{\partial t_\beta}{\partial u_{\alpha'}} .$$

That means that the μ vectors with the components:

$$S'_{\alpha j} - S_{\alpha j} , \qquad S'_{\alpha\beta} - S_{\alpha\beta} ,$$

all point in the direction of the normal to $\mathfrak{F}$, whose components we denote by ν_i, ν_β. That is, there are μ functions ρ_α on the hypersurface such that:

(19.4) $$\begin{aligned} S'_{\alpha j} - S_{\alpha j} &= \rho_\alpha \nu_j , \\ S'_{\alpha\beta} - S_{\alpha\beta} &= \rho_\alpha \nu_\beta . \end{aligned}$$

20. Now, we assume that we are given the quantities $P_{i\alpha}$ and $P'_{i\alpha}$ on a (μ – 1)-dimensional manifold:

(20.1) $$t_1 = \tau(t_{\alpha'}), \qquad x_i = \xi_i (t_{\alpha'}).$$

It shall be possible to construct a geodesic field through each hypersurface $\mathfrak{F}$ that goes through (20.1), such that the equations (19.1) to (19.4) are satisfied and whose surface elements agree with the given ones on (20.1). What conditions must these quantities therefore satisfy?

From (19.2), it follows by subtraction that:

(20.2) $$S'_{\alpha i} - S_{\alpha i} - P'_{i\beta} (S'_{\alpha\beta} - S_{\alpha\beta}) - (P'_{i\beta} - P_{i\beta}) S_{\alpha\beta} = 0 .$$

One now chooses the hypersurface $\mathfrak{F}$ especially such that at a point of (20.1), one has [1]):

(20.3) $$v_i - P'_{i\beta} v_\beta = 0$$

It then follows from (19.4) and (20.2) for this point that:

$$(P'_{i\beta} - P_{i\beta}) S_{\alpha\beta} = 0$$

and therefore, since $| S_{\alpha\beta} | \neq 0$:

(20.4) $$P'_{i\alpha} = P_{i\alpha}.$$

Instead of (20.2), one now has at the point in question:

$$S'_{\alpha i} - S_{\alpha i} - P'_{i\beta}(S'_{\alpha\beta} - S_{\alpha\beta}) = 0.$$

However, one can choose the hypersurface such that (20.3) *is not* valid. Due to (19.4), one can then satisfy the latter equation only with $S'_{\alpha i} = S_{i\alpha}$ and $S'_{\alpha\beta} = S_{\alpha\beta}$, and it then follows from (19.3) that:

(20.5) $$F' = F.$$

The given quantities must then satisfy equations (20.4), (20.5).

21. Now, we consider an extremal that possesses a "kink" along a manifold (20.1). Its surface element is represented by $p_{i\alpha}$ on one side of (20.1) and on the other, by $p'_{i\alpha}$, such that on the kink one has $p'_{i\alpha} \neq p_{i\alpha}$. One is then dealing with a special case of the previously treated problem, and therefore equations (20.4) and (20.5) must be satisfied on the kink.

From (20.4), it follows that one cannot operate on both sides of the kink with one and the same Legendre transformation (otherwise, it would then follow from $P'_{i\alpha} = P_{i\alpha}$ that $p'_{i\alpha} = p_{i\alpha}$ [2]). Here, one therefore also has to deal with two *different* Hamiltonian functions, and (20.5) is not a consequence of (20.4).

When written out in detail, the *generalized Erdmann corner conditions* read like:

(21.1) $$\frac{\bar{a}_{\alpha\beta}}{a} \pi_{i\beta} = \frac{\bar{a}'_{\alpha\beta}}{a'} \pi'_{i\beta},$$

[1]) Thus, - cf., (11.1) – it is "transversal to $p'_{i\alpha}$" in the event that the "lower case" quantities are defined.

[2]) We will likewise see that $\mathcal{E}(x_i, t_\alpha, p_{i\alpha}, p'_{i\alpha}) = 0$. Now, if the $\mathcal{E}$-function is usually positive, and one writes it in the form (5.2) then one sees that the quadratic form is singular in the point of the connecting line from $p_{i\alpha}$ to $p'_{i\alpha}$ that is denoted by the circumflex, and thus the determinant (9.2) vanishes.

(21.2) $$\frac{f^{\mu-1}}{a} = \frac{f'^{\mu-1}}{a'} .$$

22. One arrives at a system of equations that is equivalent to the Erdmann equations in the following way: We consider an extremal, and around it, a neighborhood in which it is "strong," i.e., where $\mathcal{E} > 0$ for $p'_{i\alpha} \neq p_{i\alpha}$. If we fix our attention on a particular boundary point of this neighborhood then the following is obviously valid there: We have $\mathcal{E} \geq 0$; however, there is at least one system of values $p'_{i\alpha} \neq p_{i\alpha}$ for which one has:

(22.1) $$\mathcal{E}(x_i , t_\alpha , p_{i\alpha}, p'_{i\alpha}) = 0 .$$

A surface element with this property may be called "semi-strong." We would still like to deduce for the semi-strong surface elements that there is a neighborhood of the system of values $p'_{i\alpha} \neq p_{i\alpha}$ in which the $\mathcal{E}$-function does not vanish for $p'_{i\alpha} \neq p_{i\alpha}$. In other words: Semi-strong surface elements shall be *regular* (sec. 5). In particular, the determinant is then $| q_{i\alpha,\, j\beta} | \neq 0$.

If $p_{i\alpha}$ is a semi-strong surface element then there is a system of values $p'_{i\alpha} \neq p_{i\alpha}$ that satisfy equation (22.1) and therefore, the $\mu\, n$ equations:

(22.2) $$\frac{\partial \mathcal{E}}{\partial p'_{i\alpha}} = 0 .$$

If we set:

$$q_{\alpha\beta} = \delta_{\alpha\beta} + \frac{p'_{i\alpha} - p_{i\alpha}}{f} \pi_{i\beta}$$

then, from (4.5), we can write:

(22.3) $$f' = q\, f$$

and:

(22.4) $$\pi'_{i\alpha} = \bar{q}_{\alpha\rho} \pi_{i\rho}$$

for (22.1) and (22.2).

We show that equations (21.1) and (21.2) follow from (22.5) and (22.4), and conversely; thus, the Erdmann equations have a solution on semi-strong surface elements. In fact, it follows from (22.4) that:

$$a'_{\alpha\beta} = \delta_{\alpha\beta} f' - p'_{i\alpha} \pi'_{i\beta} = \delta_{\alpha\beta} f' - p'_{i\alpha} \bar{q}_{\beta\rho} \pi_{i\rho} ,$$

and from this, one calculates, due to (22.3), and because:

$$f\, q_{\alpha\beta} = a_{\alpha\beta} + p'_{i\alpha}\, \pi_{i\beta} ,$$

that one has:

(22.5) $$a'_{\alpha\rho}\, q_{\rho\beta} = q_{\rho\beta}\, f' - p'_{i\alpha}\, q\, \pi_{i\beta} = \frac{f'}{f} q_{\alpha\beta} .$$

If one defines the determinant here then one finds that due to (22.3) one has:

$$\frac{f'}{f}a' = \frac{f'^{\mu}}{f^{\mu}}a\,,$$

hence, (21.2). However, if one contracts the rows of (22.5) with $\bar{a}_{\alpha\gamma}/a$, $\bar{q}_{\alpha\beta}/q$, and $\bar{a}'_{\sigma\rho}/a'$ then one obtains, due to (22.3):

$$\frac{\bar{a}_{\sigma\gamma}}{a} = \frac{\bar{a}_{\sigma\rho}}{a'}\bar{q}_{\rho\gamma}\,,$$

and when one multiplies this by $\pi_{i\beta}$ and observes (22.4) then one obtains (21.1).

In order to prove the converse, we use (8.1) and derive from this formula that:

$$\mathcal{E} = f' - \frac{1}{F}\,|\,\delta_{\alpha\beta} + P_{i\alpha}\;p'_{i\alpha}\,|. \tag{22.6}$$

When (21.1) and (21.2) are valid – i.e., when (20.4) and (20.5) are true – then the elements that enter into the determinant (22.6) are equal to $g'_{\alpha\beta}$, and then, on account of (7.g3), one has:

$$\mathcal{E} = f' - \frac{g'}{F'} = f - \frac{f'F'}{F'} = 0.$$

Furthermore, when one differentiates (22.6) with respect to $p'_{i\alpha}$, and then sets $P_{i\alpha} = P'_{i\alpha}$, $F = F'$, one obtains:

$$\frac{\partial\mathcal{E}}{\partial p'_{i\alpha}} = \pi'_{i\alpha} - \frac{1}{F'}\bar{g}'_{\rho\alpha}P'_{i\rho} = \frac{\bar{g}'_{\rho\alpha}}{F'\,f'}(\pi'_{i\beta}g'_{\rho\beta} - f'\,P'_{i\rho}) = 0,$$

due to (7.g4).

23. Two surface elements at a point (x_i^0, t_α^0), $p_{i\alpha}$, and $p'_{i\alpha}$ that satisfy the Erdmann equations with each other, satisfy, as we just saw, equations (22.1) and (22.2). Since the Erdmann equations are symmetric in the primed and unprimed variables we must also have the validity of the equations that arise from (22.1) and (22.2) by exchanging primed and unprimed; we can also see this directly. Namely, we remember that in section 2 of chapter I we introduced the quantities:

$$\tilde{f} - \tilde{\Delta} = \mathcal{E}(x_i^0, t_\alpha^0, p_{i\alpha}, \tilde{p}_{i\alpha})\,,$$

for the $\mathcal{E}$-function, which possess an extremum with the value 0 for $\tilde{p}_{i\alpha} = p_{i\alpha}$, in any case. From (22.1), (22.2) it follows that an extremum 0 also exists at the location $\tilde{p}_{i\alpha} = p'_{i\alpha}$; the considerations of the second section then show that one can also set:

$$\tilde{f} - \tilde{\Delta} = \mathcal{E}(x_i^0, t_\alpha^0, p'_{i\alpha}, \tilde{p}_{i\alpha}).$$

Hence, for all $\tilde{p}_{i\alpha}$ one has:

(23.1) $$\mathcal{E}(x_i^0, t_\alpha^0, p_{i\alpha}, \tilde{p}_{i\alpha}) = \mathcal{E}(x_i^0, t_\alpha^0, p'_{i\alpha}, \tilde{p}_{i\alpha}).$$

From this, it follows with no further conditions: If there are two surface elements at a point that satisfy the Erdmann conditions with each other and one of them is semi-strong then the other one is always semi-strong, as well.

A further consequence of (23.1) affects the second derivatives of the $\mathcal{E}$-function, which we will likewise need. Namely, one must have:

(23.2) $$\frac{\partial^2 \mathcal{E}(x_i^0, t_\alpha^0, p_{i\alpha}, p'_{i\alpha})}{\partial p'_{i\alpha} \partial p'_{j\beta}} = q'_{i\alpha, j\beta},$$

i.e., it is equal to the numbers that were defined by (5.1), which were constructed for the $p'_{i\alpha}$. In particular, the $\mu\, n$-rowed determinant of these quantities is non-zero since we have assumed the regularity of semi-strong surface elements.

24. We would now like to present the condition for an extremal to be isolated on a manifold where the Erdmann equations possess a solution. We thus consider the Erdmann equations on an extremal, i.e., we express x_i and $p_{i\alpha}$ in terms of t_α. We can then perhaps write these $\mu\, n + 1$ equations as:

$$\Psi(t_\rho, p'_{i\alpha}) = 0, \qquad \Psi_{j\beta}(t_\rho, p'_{i\alpha}) = 0.$$

When these equations possess a solution at a point (t_α) then – and only then – this point obviously belongs to a well-defined μ – 1-dimensional manifold of points with the same property when the matrix:

(24.1) $$\begin{pmatrix} \dfrac{\partial \Psi}{\partial p'_{i\alpha}} & \dfrac{\partial \Psi_{j\beta}}{\partial p'_{i\alpha}} \\ \dfrac{\partial \Psi}{\partial t_\rho} & \dfrac{\partial \Psi_{j\beta}}{\partial t_\rho} \end{pmatrix}$$

of $\mu\, n + 1$ rows and $\mu\, n + \mu$ columns has the rank $\mu\, n + 1$.

This condition can now be written in a particularly simply manner when one starts with the Erdmann equations in the form (22.1), (22.2). Namely, at the point in question the matrix (24.1) looks like:

$$\begin{pmatrix} 0 & \dfrac{\partial^2 \mathcal{E}}{\partial p'_{i\alpha}\,\partial p'_{j\beta}} \\ \dfrac{\partial \mathcal{E}}{\partial t_\rho} & \dfrac{\partial^2 \mathcal{E}}{\partial t_\rho\,\partial p'_{j\beta}} \end{pmatrix} \quad {}^{1}).$$

If the surface elements in question are semi-strong then, from the previous section, the determinant $\left|\dfrac{\partial^2 \mathcal{E}}{\partial p'_{i\alpha}\,\partial p'_{j\beta}}\right|$ does not vanish. Thus, the matrix (24.1) has rank $\mu\, n + 1$ when and only when the numbers $\partial\mathcal{E} / \partial t_\rho$ do not all vanish; i.e., when the $\mathcal{E}$-function possesses a gradient at the points in question.

The condition (24.1) for a semi-strong surface element to be an extremal will thus mean the same thing as saying that this surface element belongs to a μ – 1-dimensional manifold of surface elements of the extremal such that on it a "strong" neighborhood of one of them splits into "weak" ones.

Along the boundary of a strong neighborhood, which may be represented in the form (20.1), a solution of the Erdmann equations whose surface elements $p'_{i\alpha}$ are semi-strong is thus determined, and thus one obtains no broken extremals in this way. Therefore, if it is entirely possible to construct an extremal with the initial values $p'_{i\alpha}$ then it must satisfy the condition that we would now like to present, and which is not satisfied, in general.

25. Namely, if $p_{i\alpha}$ and $p'_{i\alpha}$ are functions on (20.1), and they belong to a μ-dimensional surface that this manifold contains then they satisfy the equations:

(25.1) $$\frac{\partial \xi_i}{\partial t'_\alpha} = p_{i1}\frac{\partial \tau}{\partial t_{\alpha'}} + p_{i\alpha'},$$

(25.2) $$\frac{\partial \xi_i}{\partial t_{\alpha'}} = p'_{i1}\frac{\partial \tau}{\partial t_{\alpha'}} + p'_{i\alpha'}.$$

This is what one calls a strip condition in the case of $\mu = 2$. If one starts with an extremal then (25.1) is satisfied, but (25.2) is not, in general.

On the other hand, there are always [2]) "kinked strips," i.e., manifolds (20.1) and functions $p_{i\alpha}$ and $p'_{i\alpha}$ on them that satisfy the Erdmann equations, along with (25.1) and (25.2). If one is given (20.1) then one must determine $2\mu\, n$ functions that satisfy:

$$\mu\, n + 1 + 2\,(\mu - 1) = (n + 2)\,\mu - 1$$

---

[1]) The derivative of $\mathcal{E}$ with respect to t_ρ is naturally understood to mean that one also considers the dependency of the x_i and $p_{i\alpha}$ on t_ρ.

[2]) Assuming that the variational problem is not regular for all surface elements (in which case, there are only strong surface elements).

equations. Only for $\mu = n = 1$ are there exactly as many equations as the number of desired quantities. Thus, only in the plane does there exist the characteristic property of the ordinary theory of discontinuous extremals that one can uniquely determine a field of "corners" in a neighborhood of a corner. If one lets μ increase then one obtains more equations than unknowns; e.g., one will find that on surfaces in space ($\mu = 2$, $n = 1$) the curves along which a kink can be found cannot be specified arbitrarily. By contrast, if $n > 1$ then one always has more unknowns than equations.

26. If one has found a discontinuous extremal that satisfies the Erdmann equations with the condition (24.1) and it is strong on both sides of the kink then a sufficiently small piece of this surface provides a strong minimum, in fact. In order to obtain a geodesic field that intersects the surface transversally one must only carry out the construction of the previous chapter on both sides of a hypersurface $\mathfrak{F}$ that contains the kink. One chooses the functions $S_{\alpha'}$ in such a way that they remain continuous with all of their derivatives. The initial values for σ_i , σ_α will then be the same on both sides, and therefore the $P_{i\alpha}$ are continuous over the entire geodesic field.

(Received on 2 July 1935.)

____