A GEOMETRICAL THEORY OF MULTIPLE INTEGRAL PROBLEMS IN THE CALCULUS OF VARIATIONS

H. RUND

1. Introduction. Let X_n denote an n-dimensional differentiable manifold referred to local coordinates x^i. An m-dimensional subspace C_m $(m < n)$ of X_n can be represented parametrically in the form

$$x^i = x^i(t^\alpha) \qquad (i = 1, \ldots, n; \alpha = 1, \ldots, m), \tag{1.1}$$

where t^α denotes a system of independent parameters on C_m. It will be assumed throughout that the functions (1.1) under consideration are of class C^4, their first derivatives being denoted by

$$\dot{x}^i_\alpha = \partial x^i / \partial t^\alpha. \tag{1.2}$$

These define the elements of an $n \times m$ matrix, which is always supposed to be of rank m.

The following notation is adopted. Lower-case Latin indices run from 1 to n, while all Greek indices assume the values 1 to m; the summation convention is applied to both sets. If $F(t^\alpha, x^i)$ is a given function of class C^1 in all its arguments, we shall write

$$\frac{dF}{dt^\alpha} = \frac{\partial F}{\partial t^\alpha} + \frac{\partial F}{\partial x^i} \dot{x}^i_\alpha \tag{1.3}$$

for the derivatives of F on C_m.

Now let us suppose that we are given a function $L(x^i, \dot{x}^i_\alpha)$ of the $n + nm$ variables x^i, $\dot{x}^i_\alpha$, these being defined as functions of t^β on C_m, and let R_t denote a finite, simply connected region in the configuration space of the independent variables t^α. The fundamental m-fold integral

$$I = \int_{R_t} L(x^i, \dot{x}^i_\alpha)\, dt^1 \ldots dt^m \tag{1.4}$$

will, in general, depend on the choice of the subspace C_m by means of which the arguments (1.1) and (1.2) in the integrand are determined. Indeed, the basic problem of the calculus of variations associated with (1.4) can be formulated rather roughly as follows. Suppose that certain initial values $X^i(t^\alpha)$ are prescribed for all values of t^α on the boundary ∂R_t of R_t; from amongst

Received September 23, 1966.

the distinct sets of functions such as (1.1), of which all assume the same values $X^i(t^\alpha)$ for t^α on ∂R_t, one seeks that particular set which affords an extreme value (usually a minimum) to the integral (1.4). It is well known that in order that the functions $x^i(t^\alpha)$ assign such an extreme value to (1.4) it is necessary that they satisfy the following system of n Euler–Lagrange equations:

$$\frac{d}{dt^\alpha}\left(\frac{\partial L}{\partial \dot{x}^i_\alpha}\right) - \frac{\partial L}{\partial x^i} = 0. \tag{1.5}$$

In this article we shall be concerned with the geometrical theory of such problems in the sense that we shall regard the integral (1.4) as the m-dimensional "area" of the portion of the subspace C_m which is determined by R_t. It is natural to suppose that this area is positive and independent of the choice of the parameters t^α of C_m; consequently, it will be assumed throughout that $L > 0$ and that the integral (1.4) is parameter-invariant.

Two special cases of the general geometry which is thus defined are well known and have been studied exhaustively (**2**; **7**), namely, Finsler spaces (when $m = 1$) and Cartan spaces (when $m = n - 1$). In both of these a natural procedure for constructing a 2-index metric tensor in terms of the derivatives of L presents itself, but when $2 \leqslant m \leqslant n - 2$, the case of the so-called *areal spaces*, formidable new difficulties emerge for reasons that we shall outline very briefly. The condition that (1.4) be parameter-invariant permits the introduction of simple m-vectors such that L can be replaced by a function which is positively homogeneous of the first degree in the latter, which gives rise to direct analogies with Finsler geometry. But, unless $m = 1$ or $m = n - 1$, these m-vectors are not independent as a consequence of the so-called Plücker relations, and thus they cannot give rise to suitable derivatives of L. Still greater problems are encountered when one seeks a connection in areal spaces. However, since we shall not in any way be concerned directly with the method based on m-vectors, we shall not pursue these matters any further, reference being made to the very extensive literature, and, in particular, to the survey article of Kawaguchi (**5**), or to the introduction of a recent paper by Davies (**3**), in which a connection is found for a special class of areal spaces known as the submetric type.

An attempt is made here to develop *an entirely different theory of the geometry of spaces whose areal metric is defined by the integral* (1.4). The insistence on a 2-index metric tensor is abandoned altogether, and, instead, a 4-index tensor, which has been recently introduced elsewhere (**8**, p. 288), is regarded as the fundamental entity (although it must be emphasized that for $m = 1$ this tensor reduces to the usual metric tensor of Finsler geometry). As a result of the parameter-invariance of (1.4) the metric tensor and its derivatives satisfy certain identities, which are briefly described in §2. Furthermore, it is shown in §3 that this tensor can be used to construct quantities which generalize the Christoffel symbols of classical differential geometry, by means of which the covariant derivatives of certain vector fields may be formed. These are related to the expressions on the left-hand side of the Euler–Lagrange equations (1.5), and thus the contact with the calculus of variations as such is re-established. Section 4 is devoted to the derivation of suitable connection coefficients which give rise to covariant partial derivatives of arbitrary tensor fields. In §5 the corresponding theory of curvature is briefly sketched. Two distinct curvature tensors appear in the theory: the first emerges from the integrability conditions of the transformation law satisfied by the connection coefficients, while the second results from a consideration of the order in which repeated partial covariant differentiations are carried out. Each tensor plays a fundamental role in the theory (although they coincide when $m = 1$). Finally, some identities which these tensors satisfy are derived, including the counterpart of the classical Bianchi identities.

It should be emphasized that none of the problems of the earlier theories mentioned above are solved here, nor is an attempt made at this stage to relate their known invariants to those of the present theory. This article merely furnishes the beginnings of an alternative theory, and it is highly probable that important modifications will become desirable in the light of future developments.

2. Fundamental identities. The Lagrangian $L(x^i, \dot{x}^i_\alpha)$ of the fundamental integral (1.4) is supposed to satisfy certain conditions, some of which have been motivated in the Introduction. It will be evident immediately that these entail certain identities involving L and its derivatives on which the entire theory is vitally dependent. We shall therefore begin by listing our assumptions as follows:

1. The Lagrangian L is of class C^4 in all its arguments, and it is a scalar with respect to transformations of the local coordinates x^i of X_n.

2. The Lagrangian $L(x^j, \dot{x}^j_\alpha)$ is positive for all independent sets of arguments $\dot{x}^j_\alpha$.

3. The integral (1.4) is independent of the choice of the parameters t^α of the subspaces C_m.

4. The $nm \times nm$ determinant

$$D = \det\left[\frac{m}{2}\frac{\partial^2 L^{2/m}}{\partial \dot{x}^i_\alpha\, \partial \dot{x}^j_\beta}\right]$$

is non-vanishing for linearly independent $\dot{x}_\alpha{}^j$.

The reason for the last requirement will emerge presently. As regards condition 3, it is well known (**8**, p. 268) that it is equivalent to the relations

$$\frac{\partial L}{\partial \dot{x}^i_\alpha}\,\dot{x}^i_\beta = \delta^\alpha_\beta\, L, \tag{2.1}$$

which are necessary and sufficient to guarantee the parameter-invariance of the fundamental integral (1.4). Indeed, the assertion regarding the necessity of (2.1) is a special case of the following theorem, due to Douglas **(4)**, which is stated here in detail since we shall repeatedly have occasion to appeal to it: If $F^{\alpha_1\ldots\alpha_r}{}_{\beta_1\ldots\beta_s}(x^j, \dot{x}^j_\alpha)$ are functions of class C^1 which constitute the components of a tensor density of weight w, with co- and contra-variant valencies s and r with respect to transformations of the parameters t^α, then the derivatives of F with respect to $\dot{x}^i_\beta$ satisfy the following relations:

$$(2.2)\quad \frac{\partial F^{\alpha_1\ldots\alpha_r}{}_{\beta_1\ldots\beta_s}}{\partial \dot{x}^i_\mu}\dot{x}^i_\nu = w\delta^\mu_\nu F^{\alpha_1\ldots\alpha_r}{}_{\beta_1\ldots\beta_s} - \sum_{a=1}^{r}\delta^{\alpha_a}_\nu F^{\alpha_1\ldots\alpha_{a-1}\mu\alpha_{a+1}\ldots\alpha_r}{}_{\beta_1\ldots\beta_s} + \sum_{b=1}^{s}\delta^\mu_{\beta_b} F^{\alpha_1\ldots\alpha_r}{}_{\beta_1\ldots\beta_{b-1}\nu\beta_{b+1}\ldots\beta_s}.$$

Since condition 3 implies that L is a scalar density of unit weight with respect to transformations of the t^α, it is obvious how this result gives rise to (2.1).

The metric tensor is now defined as

$$(2.3)\quad g^{\alpha\beta}_{ij}(x^h, \dot{x}^h_\epsilon) = \frac{m}{2}\frac{\partial^2\{L(x^h, \dot{x}_\epsilon{}^h)\}^{2/m}}{\partial\dot{x}^i_\alpha\,\partial\dot{x}^j_\beta},$$

the motivation for this definition being based (partly) on the fact that (2.1) gives rise to the following identity:

$$(2.4)\quad L(x^h, \dot{x}^h_\epsilon) = [m^{-1}g^{\alpha\beta}_{ij}(x^h, \dot{x}^h_\epsilon)\dot{x}^i_\alpha\,\dot{x}^j_\beta]^{m/2},$$

from which it is evident that, if L is interpreted as a measure of the area dA of an element of an m-dimensional subspace spanned by $\dot{x}_\alpha{}^j$ at the point x^j of X_n in the sense that

$$(2.5)\quad dA = L(x^j, \dot{x}^j_\alpha)dt^1\ldots dt^m,$$

then the tensor (2.3) can be regarded as a suitable areal metric tensor **(8**, p. 289**)**.

In analogy with the usual notation of Finsler geometry, we shall now write

$$(2.6)\quad C^{\alpha\beta\gamma}_{ijk} = \frac{1}{2}\frac{\partial g^{\alpha\beta}_{ij}}{\partial\dot{x}^k_\gamma},$$

where it is to be noted that $g^{\alpha\beta}_{ij}$, $C^{\alpha\beta\gamma}_{ijk}$ are symmetric in *pairs* of indices such as (α, i), (β, j), as is immediately evident from (2.3). The identity (2.2) is now applied to the tensor density (2.3), which yields

$$(2.7)\quad C^{\alpha\beta\gamma}_{ijk}\dot{x}^i_\epsilon = \frac{1}{m}\delta^\alpha_\epsilon g^{\beta\gamma}_{jk} - \tfrac{1}{2}\delta^\beta_\epsilon g^{\alpha\gamma}_{jk} - \tfrac{1}{2}\delta^\gamma_\epsilon g^{\beta\alpha}_{jk},$$

as may be verified also directly by repeated differentiation of (2.1). Contracting over α, ϵ, one obtains the fundamentally important identity

$$(2.8)\quad C^{\alpha\beta\gamma}_{ijk}\dot{x}^i_\alpha = 0,$$

with similar relations involving the index pairs (β, j), (γ, k).

From (2.7), several other useful results can also be deduced. If (2.7) is multiplied by $\dot{x}^j_\beta$, the identities (2.8) being taken into account, it is found that

(2.9) $$g^{\alpha\gamma}_{ik}\dot{x}^i_\epsilon = \frac{2}{m}\delta^\alpha_\epsilon g^{\gamma\beta}_{kj}\dot{x}^j_\beta - \delta^\gamma_\epsilon g^{\alpha\beta}_{kj}\dot{x}^j_\beta,$$

in contrast to

(2.10) $$g^{\alpha\gamma}_{ik}\dot{x}^i_\alpha = L^{(2/m-1)}\frac{\partial L}{\partial \dot{x}^k_\gamma},$$

which is a direct result of the definition (2.3). Furthermore, if we contract over β and ϵ in (2.7), we obtain

(2.11) $$C^{\alpha\beta\gamma}_{ijk}\dot{x}^i_\beta = \left(\frac{1}{m} - \frac{m}{2}\right)g^{\alpha\gamma}_{jk} - \tfrac{1}{2}g^{\gamma\alpha}_{jk},$$

and hence, again by virtue of (2.8),

(2.12) $$g^{\alpha\beta}_{ij}\dot{x}^j_\alpha = \left(\frac{2}{m} - m\right)g^{\beta\alpha}_{ij}\dot{x}^j_\alpha.$$

This identity gives rise to a result of basic importance. It is obvious from the definition (2.3) that in general

(2.13) $$g^{\alpha\beta}_{ij} \neq g^{\beta\alpha}_{ij}.$$

This is the source of many of the analytical difficulties which will be encountered later. One would therefore be inclined to consider a subclass of Lagrangians for which (2.13) could be replaced by an equality. However, if the latter were to hold, this would be compatible with (2.12) if and only if $(2/m - m) = 1$, which is feasible only for $m = 1$. *Thus the validity of the inequality* (2.13) *must be accepted except for single integral problems* (for which it is meaningless).

Also, when (2.9) is multiplied by $\dot{x}^k_\gamma$, we find, using (2.4), that

(2.14) $$g^{\alpha\beta}_{ij}\dot{x}^i_\epsilon\dot{x}^j_\beta = m^{-1}\delta^\alpha_\epsilon g^{\gamma\beta}_{kj}\dot{x}^k_\gamma\dot{x}^j_\beta = \delta^\alpha_\epsilon L^{2/m},$$

while (2.7), taken in conjunction with (2.12), similarly yields the identity

(2.15) $$C^{\alpha\beta\gamma}_{ijk}\dot{x}^i_\beta\dot{x}^j_\alpha = 0.$$

Finally, we briefly indicate how the inverse of the tensor (2.3) may be obtained. Defining the conjugate momenta by means of (2.10) as

(2.16) $$y^\alpha_i = g^{\alpha\beta}_{ij}\dot{x}^j_\beta = \frac{\partial L^{2/m}}{\partial \dot{x}^i_\alpha},$$

it follows that we may express the $\dot{x}^i_\alpha$ as functions of (x^j, y^β_j), for by virtue of our condition 4 the equations (2.16) can be solved for the $\dot{x}^i_\alpha$, which yields a relation of the form

(2.17) $$\dot{x}^i_\alpha = \psi^i_\alpha(x^j, y^\beta_j).$$

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

It has been shown by Martin **(6)** that the function defined by the substitution of (2.17) in L, namely

$$H(x^i, y^\beta_j) = L\{x^j, \psi^i_\alpha(x^j, y^\beta_j)\}, \tag{2.18}$$

can be regarded as a suitable Hamiltonian function for our given problem in the calculus of variations, and, in particular, that the quantities defined by

$$g^{ij}_{\alpha\beta} = \frac{m}{2}\frac{\partial^2\{H(x^h, y^h_\epsilon)\}^{2/m}}{\partial y^\alpha_i\,\partial y^\beta_j} \tag{2.19}$$

satisfy the relations

$$g^{ij}_{\alpha\beta}\, g^{\alpha\gamma}_{ik} = \delta^j_k\, \delta^\gamma_\beta. \tag{2.20}$$

Clearly, this method of obtaining an inverse of (2.3) depends significantly on our condition 4, which is obviously related to some form of generalized Legendre condition. In this connection it is necessary that we refer to the corresponding condition in the remarkable theory of multiple integrals due to Carathéodory **(1)**, whose canonical formalism depends on the analogous requirement that the determinant Δ of the quantities defined by

$$Q^{\alpha\beta}_{ij} = L\frac{\partial^2 L}{\partial \dot{x}^i_\alpha\,\partial \dot{x}^j_\beta} - \frac{\partial L}{\partial \dot{x}^i_\alpha}\frac{\partial L}{\partial \dot{x}^j_\beta} + \frac{\partial L}{\partial \dot{x}^i_\beta}\frac{\partial L}{\partial \dot{x}^j_\alpha} \tag{2.21}$$

should be non-vanishing **(1, § 22)**. However, by differentiation of (2.1) with respect to $\dot{x}^h_\epsilon$ we find that

$$\frac{\partial^2 L}{\partial \dot{x}^i_\alpha\,\partial \dot{x}^j_\beta}\dot{x}^i_\epsilon = \delta^\alpha_\epsilon\frac{\partial L}{\partial \dot{x}^j_\beta} - \delta^\beta_\epsilon\frac{\partial L}{\partial \dot{x}^j_\alpha}, \tag{2.22}$$

which, together with (2.1), indicates that

$$Q^{\alpha\beta}_{ij}\,\dot{x}^i_\epsilon = 0 \tag{2.23}$$

identically. This is in direct contrast to the condition $\Delta \neq 0$, and indicates that the maximal rank of the matrices $(Q^{\alpha\beta}_{ij})$, each matrix with (α, β) fixed, is $n - m$. Needless to say, the theory of Carathéodory does not presuppose the parameter-invariance of (1.4), from which (2.23) has been deduced.

3. The generalized Christoffel symbols. In analogy with the known geometries associated with a metric tensor, our primary objective is the construction of a set of connection parameters by means of which covariant derivatives of tensors of arbitrary rank may be defined. This aim will be realized partly in the present section by means of a direct investigation of the transformation properties of the derivatives of the tensor (2.3).

Initially, we shall suppose that we are given an m-dimensional subspace C_m (or, if necessary, a family of such subspaces). Under a twice differentiable transformation

$$\bar{x}^i = \bar{x}^i(x^j) \tag{3.1}$$

of the local coordinates of X_n, the quantities (1.2) transform as follows:

$$\dot{x}_\alpha^i = B_h^i \dot{\bar{x}}_\alpha^h, \tag{3.2}$$

where we have written

$$B_h^i = \partial x^i / \partial \bar{x}^h. \tag{3.3}$$

It will be assumed throughout that the transformation (3.1) is non-singular. We shall also put

$$B_h{}^i{}_l = \frac{\partial^2 x^i}{\partial \bar{x}^h \, \partial \bar{x}^l}, \tag{3.4}$$

and for future reference we note that the derivatives of $\dot{x}_\alpha^i$, to be denoted by $\dot{x}_\alpha{}^i{}_\beta$, satisfy the following transformation law:

$$\dot{x}_\alpha{}^i{}_\beta = B_h^i \dot{\bar{x}}_\alpha{}^h{}_\beta + B_h{}^i{}_k \dot{\bar{x}}_\alpha^h \dot{\bar{x}}_\beta^k. \tag{3.5}$$

Let us denote by $X_\epsilon^i(x^j)$ a set of m linearly independent differentiable vector fields ($\epsilon = 1, \ldots, m$) tangent to C_m, so that

$$X_\epsilon^i = U_\epsilon^\alpha \dot{x}_\alpha^i, \tag{3.6}$$

where the U_ϵ^α represent the projections of the X_ϵ^i onto the $\dot{x}_\alpha^i$. The values of U_ϵ^α are independent of the choice of the coordinate system in X_n. Under the transformation (3.1) we then have, by (3.2) and (3.6):

$$X_\alpha^i = U_\alpha^\beta B_h^i \dot{\bar{x}}_\beta^h = B_h^i \bar{X}_\alpha^h, \tag{3.7}$$

showing that the X_α^i obey the same transformation law as the $\dot{x}_\alpha^i$. Furthermore, the transform of the derivative of X_α^i with respect to t^β is given by

$$\frac{\partial X_\alpha^i}{\partial t^\beta} = B_h^i \frac{\partial \bar{X}_\alpha^h}{\partial t^\beta} + B_h{}^i{}_j \bar{X}_\alpha^h \dot{\bar{x}}_\beta^j. \tag{3.8}$$

Clearly, these quantities do not represent components of a tensor field, and, consequently, we must now endeavour to construct the corresponding covariant derivatives.

Since the given Lagrangian L is supposed to be a scalar with respect to the transformation (3.1), it follows directly from (3.2) that the quantities (2.3) represent components of a covariant tensor of rank 2 (as already anticipated by our nomenclature and notation). We therefore have

$$\bar{g}_{ts}^{\alpha\beta}(\bar{x}^r, \dot{\bar{x}}_\gamma^r) = g_{kj}^{\alpha\beta}(x^l, \dot{x}_\epsilon^l) B_t^k B_s^j. \tag{3.9}$$

This relation is now differentiated with respect to $\bar{x}^r$, being noted that, as a result of (3.2),

$$\partial \dot{x}_\lambda^l / \partial \bar{x}^r = B_v{}^l{}_r \dot{\bar{x}}_\lambda^v. \tag{3.10}$$

In terms of the notation (2.6) we thus obtain:

$$\frac{\partial \bar{g}_{ts}^{\alpha\beta}}{\partial \bar{x}^r} = \frac{\partial g_{kj}^{\alpha\beta}}{\partial x^i} B_r^i B_s^j B_t^k + g_{kj}^{\alpha\beta}(B_t{}^k{}_r B_s^j + B_t^k B_s{}^j{}_r) + 2C_{kjl}^{\alpha\beta\lambda}(B_v{}^l{}_r \dot{\bar{x}}_\lambda^v) B_t^k B_s^j. \tag{3.11}$$

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

Here we replace the indices t, s, r by r, t, s, respectively, and in the relation thus obtained the indices s and t are interchanged. The latter is then subtracted from the sum of the former and (3.11). After some simplification, the result of these operations is expressed in the following form:

$$(3.12)\qquad \bar{\gamma}^{(\alpha\beta)}_{rs,t} = \gamma^{(\alpha\beta)}_{ij,k} B^i_r B^j_s B^k_t + \tfrac{1}{2}(g^{\alpha\beta}_{kj} + g^{\alpha\beta}_{jk}) B^k_t B_s{}^j{}_r \\ + C^{\alpha\beta\lambda}_{kjl}[B_v{}^l{}_r B^k_t B^j_s + B_v{}^l{}_s B^k_r B^j_t - B_v{}^l{}_t B^k_r B^j_s]\dot{\bar{x}}^v_\lambda,$$

where we have written

$$(3.13)\qquad \gamma^{(\alpha\beta)}_{ij,k} = \frac{1}{2}\left(\frac{\partial g^{\alpha\beta}_{kj}}{\partial x^i} + \frac{\partial g^{\alpha\beta}_{ik}}{\partial x^j} - \frac{\partial g^{\alpha\beta}_{ij}}{\partial x^k}\right),$$

these quantities representing the *generalized Christoffel symbols* of the first kind (**8**, p. 291, equation (5.11)). The relation (3.12) thus represents the transformation law of these symbols.

Let us multiply (3.12) by $\dot{\bar{x}}^r_\alpha$, noting (3.2) and the identity (2.8), to obtain

$$(3.14)\qquad \bar{\gamma}^{(\alpha\beta)}_{rs,t} \dot{\bar{x}}^r_\alpha = \gamma^{(\alpha\beta)}_{ij,k} \dot{x}^i_\alpha B^j_s B^k_t + G^{\alpha\beta}_{kj} B^k_t B_s{}^j{}_r \dot{\bar{x}}^r_\alpha + C^{\alpha\beta\lambda}_{kjl}(B_v{}^l{}_r \dot{\bar{x}}^v_\lambda \dot{\bar{x}}^r_\alpha) B^k_t B^j_s,$$

where, for the sake of brevity, we have written

$$(3.15)\qquad G^{\alpha\beta}_{kj} = \tfrac{1}{2}(g^{\alpha\beta}_{kj} + g^{\alpha\beta}_{jk}).$$

We shall now eliminate the second derivatives $B_s{}^j{}_r$ from (3.14). By means of (3.5) the last term on the right-hand side of (3.14) can be expressed as

$$(3.16)\qquad C^{\alpha\beta\lambda}_{kjl}(\dot{x}_\lambda{}^l{}_\alpha - B^l_r \dot{\bar{x}}_\lambda{}^r{}_\alpha) B^k_t B^j_s = B^k_t B^j_s C^{\alpha\beta\lambda}_{kjl} \dot{x}_\lambda{}^l{}_\alpha - \bar{C}^{\alpha\beta\lambda}_{tsr} \dot{\bar{x}}_\lambda{}^r{}_\alpha,$$

where we have used the fact that the quantities (2.6) represent the components of a tensor of covariant valency 3 by virtue of (3.2). When (3.16) is substituted in (3.14), the latter assumes the following form:

$$(3.17)\qquad \bar{\gamma}^{(\alpha\beta)}_{rs,t} \dot{\bar{x}}^r_\alpha + \bar{C}^{\alpha\beta\lambda}_{tsr} \dot{\bar{x}}_\lambda{}^r{}_\alpha = G^{\alpha\beta}_{kj} B^k_t (B_s{}^j{}_r \dot{\bar{x}}^r_\alpha) + B^k_t B^j_s \{\gamma^{(\alpha\beta)}_{ij,k} \dot{x}^i_\alpha + C^{\alpha\beta\lambda}_{kjl} \dot{x}_\lambda{}^l{}_\alpha\}.$$

This is now multiplied by $\bar{X}^s_\epsilon$, equation (3.6) being taken into account, which yields

$$(3.18)\qquad (\bar{\gamma}^{(\alpha\beta)}_{rs,t} \dot{\bar{x}}^r_\alpha + \bar{C}^{\alpha\beta\lambda}_{tsr} \dot{\bar{x}}_\lambda{}^r{}_\alpha) \bar{X}^s_\epsilon = G^{\alpha\beta}_{kj} B^k_t (B_s{}^j{}_r \bar{X}^s_\epsilon \dot{\bar{x}}^r_\alpha) \\ + B^k_t (\gamma^{(\alpha\beta)}_{ij,k} \dot{x}^i_\alpha + C^{\alpha\beta\lambda}_{kjl} \dot{x}_\lambda{}^l{}_\alpha) X^j_\epsilon.$$

Using (3.8), the first term on the right-hand side can be written in the form

$$(3.19)\qquad G^{\alpha\beta}_{kj} B^k_t \left(\frac{\partial X^j_\epsilon}{\partial t^\alpha} - B^j_r \frac{\partial \bar{X}^r_\epsilon}{\partial t^\alpha}\right) = B^k_t G^{\alpha\beta}_{kj} \frac{\partial X^j_\epsilon}{\partial t^\alpha} - \bar{G}^{\alpha\beta}_{tr} \frac{\partial \bar{X}^r_\epsilon}{\partial t^\alpha},$$

where we have used the fact that the quantities defined in (3.15) are components of a tensor.

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

When (3.19) is substituted in (3.18), we obtain a relation of the type

$$\bar{V}_{\epsilon}{}^{\beta}{}_{t} = B^{k}_{t}\, V_{\epsilon}{}^{\beta}{}_{k}, \tag{3.20}$$

showing that the quantities defined by

$$V_{\epsilon}{}^{\beta}{}_{k} = G^{\alpha\beta}_{kj}\frac{\partial X^{j}_{\epsilon}}{\partial t^{\alpha}} + (\gamma^{(\alpha\beta)}_{ij,k}\,\dot{x}^{i}_{\alpha} + C^{\alpha\beta\lambda}_{kjl}\,\dot{x}_{\lambda}{}^{l}{}_{\alpha})X^{j}_{\epsilon} \tag{3.21}$$

are the components of a covariant vector. Clearly, the right-hand side of (3.21) contains the required covariant derivative of X^{j}_{ϵ}.

A special case of fundamental importance is obtained when we put $X^{j}_{\epsilon} = \dot{x}^{j}_{\epsilon}$ in (3.21), after which we contract over ϵ and β. The first term on the right-hand side then becomes

$$G^{\alpha\beta}_{kj}\,\dot{x}_{\beta}{}^{j}{}_{\alpha} = g^{\alpha\beta}_{kj}\,\dot{x}_{\alpha}{}^{j}{}_{\beta} \tag{3.22}$$

as a result of the symmetry of $\dot{x}_{\alpha}{}^{j}{}_{\beta}$ in α, β and (3.15), while the last term vanishes identically by virtue of (2.8). Denoting the results of these operations by V_k, we have thus established that the quantities

$$V_{k} = g^{\alpha\beta}_{kj}\frac{\partial^{2}x^{j}}{\partial t^{\alpha}\partial t^{\beta}} + \gamma^{(\alpha\beta)}_{ij,k}\,\dot{x}^{i}_{\alpha}\,\dot{x}^{j}_{\beta} \tag{3.23}$$

are components of a covariant vector, which does, in fact, represent a generalized divergence. This vector is of particular importance, for, as has been shown elsewhere (**8**, p. 292, equation (5.125)), it is identically related to the expression on the left-hand side of the Euler–Lagrange equation (1.5) as follows:

$$\frac{d}{dt^{\alpha}}\left(\frac{\partial L}{\partial \dot{x}^{k}_{\alpha}}\right) - \frac{\partial L}{\partial x^{k}} = L^{(1-2/m)}\,V_{k} + \frac{(m-2)}{mL}\,\frac{\partial L}{\partial \dot{x}^{k}_{\alpha}}\,\frac{dL}{dt^{\alpha}}, \tag{3.24}$$

and it can easily be shown (*ibid.*) that the condition $V_k = 0$ implies that (1.5) is satisfied. We shall call those subspaces C_m of V_n for which $V_k = 0$ *geodesic subspaces: clearly, these are also extremal subspaces in the sense that they satisfy the Euler–Lagrange equation of our variational problem.*

In order to be able to solve (3.21) for the derivatives of X^{j}_{ϵ} we must now assume that the quantities (3.15) possess an inverse in analogy with that of (2.3) in the sense of (2.20); i.e., we suppose that there exist quantities $G^{kh}_{\beta\gamma}$ such that

$$G^{\beta\alpha}_{kj}\,G^{kh}_{\beta\gamma} = \delta^{\alpha}_{\gamma}\,\delta^{h}_{j}. \tag{3.25}$$

When (3.21) is multiplied by $G^{hk}_{\beta\gamma}$ we thus obtain the following contravariant vector:

$$\frac{\delta X^{h}_{\epsilon}}{\delta t^{\gamma}} \equiv G^{hk}_{\beta\gamma}\,V_{\epsilon}{}^{\beta}{}_{k} = \frac{\partial X^{h}_{\epsilon}}{\partial t^{\gamma}} + (G^{kh}_{\beta\gamma}\,\gamma^{(\alpha\beta)}_{ij,k}\,\dot{x}^{i}_{\alpha} + G^{kh}_{\beta\gamma}\,C^{\alpha\beta\lambda}_{kjl}\,\dot{x}_{\lambda}{}^{l}{}_{\alpha})X^{j}_{\epsilon}. \tag{3.26}$$

Clearly this covariant derivative depends on the given subspace C_m in that the quantities $\dot{x}_{\lambda}{}^{l}{}_{\alpha}$ appear on the right-hand side. This may, in fact, be useful for certain special problems. For a general theory, however, it is desirable to avoid this type of dependence. We shall show in the next section how this can be done. To this end we note already at this stage that differentiation of (2.8) yields the identity

(3.27) $$C^{\delta\beta\lambda}_{kjl} = -C^{\delta\beta\lambda\alpha}_{kjli}\,\dot{x}^i_\alpha,$$

where we have put

(3.28) $$C^{\delta\beta\lambda\alpha}_{kjli} = \partial C^{\delta\beta\lambda}_{kjl}/\partial \dot{x}^i_\alpha.$$

It therefore follows that (3.26) can be written in the following form:

(3.29) $$\frac{\delta X^h_\epsilon}{\delta t^\gamma} = \frac{\partial X^h_\epsilon}{\partial t^\gamma} + G^{kh}_{\beta\gamma}(\gamma^{(\alpha\beta)}_{ij,k} - C^{\delta\beta\lambda\alpha}_{kjli}\,\dot{x}_\delta{}^l{}_\lambda)\dot{x}^i_\alpha X^j_\epsilon.$$

4. The connection coefficients. When we identify the arbitrary vector field X^h_ϵ, which is defined over the subspace C_m, with the field $\dot{x}^h_\epsilon$ in (3.26), we obtain a contravariant vector field, which is denoted as follows:

(4.1) $$\xi_\epsilon{}^k{}_\gamma = \dot{x}_\epsilon{}^k{}_\gamma + (G^{kh}_{\beta\gamma}\,\gamma^{(\alpha\beta)}_{ij,h}\,\dot{x}^i_\alpha + G^{kh}_{\beta\gamma}\,C^{\alpha\beta\lambda}_{hjl}\,\dot{x}_\lambda{}^l{}_\alpha)\dot{x}^j_\epsilon.$$

Clearly, these quantities are defined entirely by C_m and can, in fact, be used to describe important properties of the latter. In order to simplify (4.1), we apply (2.7) to the last term on the right-hand side, which, because of the symmetry of $\dot{x}_\lambda{}^l{}_\alpha$ in λ, α, can thus be written as

$$G^{kh}_{\beta\gamma}\left[\frac{1}{m}\,\delta^\beta_\epsilon\,g^{\alpha\lambda}_{hl}\,\dot{x}_\alpha{}^l{}_\lambda - \tfrac{1}{2}(g^{\beta\alpha}_{hl} + g^{\alpha\beta}_{hl})\dot{x}_\alpha{}^l{}_\epsilon\right],$$

and this, if we use (3.15) and (3.25), is seen to be equivalent to

$$\frac{1}{m}\,G^{kh}_{\epsilon\gamma}\,G^{\alpha\lambda}_{hl}\,\dot{x}_\alpha{}^l{}_\lambda - \dot{x}_\epsilon{}^k{}_\gamma.$$

Thus (4.1) becomes

(4.2) $$\xi_\epsilon{}^k{}_\gamma = \frac{1}{m}\,G^{kh}_{\gamma\epsilon}\,G^{\alpha\beta}_{hl}\,\dot{x}_\alpha{}^l{}_\beta + G^{kh}_{\gamma\beta}\,\gamma^{(\alpha\beta)}_{ij,h}\,\dot{x}^i_\alpha\,\dot{x}^j_\epsilon.$$

When this is multiplied by $G^{\gamma\lambda}_{ki}$, the identity (3.25) being taken into account, it is found that

$$\xi_\epsilon{}^k{}_\gamma\,G^{\gamma\lambda}_{ki} = \frac{1}{m}\,\delta^\lambda_\epsilon\,g^{\alpha\beta}_{il}\,\dot{x}_\alpha{}^l{}_\beta + \gamma^{(\alpha\lambda)}_{lj,i}\,\dot{x}^l_\alpha\,\dot{x}^j_\epsilon,$$

or, if we contract over λ, ϵ and apply the definition (3.25),

(4.3) $$V_i = G^{\gamma\epsilon}_{ki}\,\xi_\epsilon{}^k{}_\gamma = \tfrac{1}{2}G^{\gamma\epsilon}_{ki}(\xi_\epsilon{}^k{}_\gamma + \xi_\gamma{}^k{}_\epsilon).$$

In analogy with the well-known concept of "autoparallel" curves of metric differential geometry, this result suggests the following definition. An m-dimensional subspace C_m of X_n is said to be *autoparallel at a point* $P(x^j)$ if at P the conditions

(4.4) $$\xi_\alpha{}^j{}_\beta + \xi_\beta{}^j{}_\alpha = 0$$

are satisfied. Our definition is formulated in this particular manner for the following reasons. First, with a view to our present needs, we are interested in the validity of (4.4) merely at a single point, so that we may ignore the integrability conditions which would be relevant if one were to require that (4.4) be valid over a finite region of C_m (although such subspaces, if they exist, will be called autoparallel surfaces). Secondly, it is obvious from (4.2) that the $\xi_\epsilon{}^k{}_\gamma$ are not, in general, symmetric, although it is clear from (4.3) that only their symmetric parts play a significant role in relation to the problem in the calculus of variations under consideration. Thus, a symmetric formulation is chosen for (4.4).

It follows from (4.4) and (4.3) *that an autoparallel subspace* (when it exists) *is a geodesic subspace, and hence also an extremal surface of the fundamental integral* (1.4). It need hardly be added that the converse need not hold in general.

If (4.2) is substituted in (4.4), we see that the latter represents a set of $\frac{1}{2}nm(m+1)$ non-homogeneous linear equations in the $\frac{1}{2}nm(m+1)$ quantities $\dot{x}_\alpha{}^i{}_\beta$. We now make the additional assumption that this set of linear equations can be solved for the latter. (This obviously entails the requirement that a certain determinant is non-zero at the point under consideration: for $m = 1$ the corresponding matrix is simply the unit $(n \times n)$ matrix.) From the form of these equations it is obvious that the solution must be of the form

$$\dot{x}_\alpha{}^i{}_\beta = f_\alpha{}^i{}_\beta(x^j, \dot{x}_\epsilon^j). \tag{4.5}$$

Furthermore, since (4.1) represents a vector field, condition (4.4) is invariant under the coordinate transformation (3.1), and the same must apply, of course, to the relation (4.5); hence, the $f_\alpha{}^i{}_\beta$ transform under (3.1) in precisely the same way as the $\dot{x}_\alpha{}^i{}_\beta$, namely, as follows:

$$f_\alpha{}^i{}_\beta = B_h^i \bar{f}_\alpha{}^h{}_\beta + B_h{}^i{}_k \dot{\bar{x}}_\alpha^h \dot{\bar{x}}_\beta^k. \tag{4.6}$$

We now return to the last term on the right-hand side of (3.14), but instead of using (3.5) to decompose it, we use (4.6) for this process. The latter is again of the form (3.16), except for the replacement of $\dot{x}_\alpha{}^i{}_\beta$ by $f_\alpha{}^i{}_\beta$; otherwise, we continue as before, and instead of (3.29) we obtain the *covariant derivative*

$$\frac{DX_\epsilon^h}{Dt^\gamma} = \frac{\partial X_\epsilon^h}{\partial t^\gamma} + P_i{}^h{}_{j\gamma}^\alpha \dot{x}_\alpha^i X_\epsilon^j, \tag{4.7}$$

where we have put

$$P_i{}^h{}_j{}^\alpha_\gamma = G_{\beta\gamma}^{kh}(\gamma_{ij,k}^{(\alpha\beta)} - C_{ijkl}^{\alpha\beta\delta\lambda} f_\delta{}^l{}_\lambda). \tag{4.8}$$

Here it is to be noted that

$$P_i{}^h{}_j{}^\alpha_\gamma = P_i{}^h{}_j{}^\alpha_\gamma(x^k, \dot{x}_\epsilon^k), \tag{4.9}$$

in other words, these coefficients are independent of the second derivatives $\dot{x}_\alpha{}^j{}_\beta$; but, of course, the covariant derivative (4.7) differs from the vector

(3.29) since the latter refers to a subspace C_m for which (4.4) will not necessarily be valid.

The transition from (3.29) to (4.7) can be interpreted geometrically as follows. At the point $P(x^j)$ of the subspace C_m at which the vector (3.29) is constructed, one now considers another subspace $C_m{}^*$, which is tangent to C_m at P in the sense that the derivatives $\dot{x}^i_\alpha$ which define the tangent planes at P are common to both, where, however, it is assumed that $C_m{}^*$ is autoparallel at P. The covariant derivative (4.7) bears precisely the same relationship to $C_m{}^*$ that (3.29) bears to C_m.

Once this transition has been achieved, we can adopt an altogether more general point of view. Instead of basing our considerations on a given subspace C_m (or a family of such subspaces), we now merely consider sets of m linearly independent vectors $\dot{x}^i_\epsilon$ at each point P of X_n: the question as to whether or not such sets define integrable subspaces C_m ceases to be relevant to our subsequent analysis. (Of course, from the point of view of the calculus of variations this question is not without relevance: with regard to these matters we refer to some remarks made by Weyl (**9**; **10**).) As before, however, the functions to be considered in the following are to depend on the $n + nm$ independent variables $(x^j, \dot{x}^j_\epsilon)$, being required that the $\dot{x}^i_\alpha$ transform according to (3.2) under (3.1), and as covariant vectors under transformations of the independent variables t^α. Precisely the same behaviour is prescribed for a set of components $X^i_\alpha(x^j, \dot{x}^i_\epsilon)$, whose partial covariant derivatives will now be obtained.

Since (4.7) represents a vector, the transformation law of the coefficient of X^j_ϵ in (4.7) can be derived directly, namely, by substitution from (3.8) in the relation expressing the vector transformation property of (4.7). This gives

$$P_s{}^r{}_t{}^\alpha_\gamma \dot{x}^s_\alpha X^t_\epsilon = B^r_i \bar{P}_j{}^i{}_h{}^\nu_\gamma \dot{\bar{x}}^j_\nu \bar{X}^h_\epsilon - B_h{}^r{}_l \bar{X}^h_\epsilon \dot{\bar{x}}^l_\gamma. \tag{4.10}$$

Since the $\bar{X}^h_\epsilon$ appearing here are essentially arbitrary, we infer by means of (3.7) that

$$B^s_j P_t{}^r{}_s{}^\alpha_\gamma \dot{x}^t_\alpha = B^r_i \bar{P}_h{}^i{}_j{}^\nu_\gamma \dot{\bar{x}}^h_\nu - B_j{}^r{}_l \dot{\bar{x}}^l_\gamma. \tag{4.11}$$

This relation is now differentiated with respect to $\dot{\bar{x}}^k_\beta$. In the course of this process use is made of (3.2), according to which

$$\partial \dot{x}^t_\alpha / \partial \dot{\bar{x}}^k_\beta = B^t_k \delta^\beta_\alpha. \tag{4.12}$$

After a little simplification we thus obtain

$$B^s_j B^t_k \left[\frac{\partial P_q{}^r{}_s{}^\alpha_\gamma}{\partial \dot{x}^t_\beta} \dot{x}^q_\alpha + P_t{}^r{}_s{}^\beta_\gamma \right] = B^r_i \left[\frac{\partial \bar{P}_h{}^i{}_j{}^\alpha_\gamma}{\partial \dot{\bar{x}}^k_\beta} \dot{\bar{x}}^h_\alpha + \bar{P}_k{}^i{}_j{}^\beta_\gamma \right] - B_j{}^r{}_k \delta^\beta_\gamma. \tag{4.13}$$

This immediately suggests that we should consider the quantities defined by

$$P^*{}_k{}^i{}_j{}^\beta_\gamma = P_k{}^i{}_j{}^\beta_\gamma + \frac{\partial P_h{}^i{}_j{}^\alpha_\gamma}{\partial \dot{x}^k_\beta} \dot{x}^h_\alpha, \tag{4.14}$$

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

whose transformation law is expressed explicitly by (4.13). In the latter we now contract over the indices β and γ, at the same time putting

$$\Gamma_{k\ j}^{\ i} = m^{-1} P^{*\ i\ \beta}_{k\ j\ \beta}, \tag{4.15}$$

which yields

$$B_{j\ k}^{\ r} = B_i^r \bar{\Gamma}_{k\ j}^{\ i} - \Gamma_{t\ s}^{\ r} B_k^t B_j^s. \tag{4.16}$$

But this is the usual transformation law which any set of connection coefficients must satisfy. We shall therefore regard *the expressions* (4.15) *as the components of our connection*, where it is to be noted that these are functions of $(x^j, \dot{x}_\alpha^j)$.

We shall now use these quantities to construct partial covariant derivatives of a given vector field $X_\epsilon^r(x^j, \dot{x}_\lambda^j)$ for which we have

$$X_\epsilon^r(x^h, \dot{x}_\lambda^h) = B_i^r \bar{X}_\epsilon^i(\bar{x}^l, \dot{\bar{x}}_\lambda^l). \tag{4.17}$$

Differentiation of this with respect to $\dot{x}_\lambda^h$ gives

$$\frac{\partial X_\epsilon^r}{\partial \dot{x}_\lambda^h} = B_i^r \frac{\partial \bar{X}_\epsilon^i}{\partial \dot{\bar{x}}_\lambda^l} A_h^l, \tag{4.18}$$

where we have put

$$A_h^l = \partial \bar{x}^l / \partial x^h. \tag{4.19}$$

In analogy with (3.10) we also have that

$$\partial \dot{\bar{x}}_\lambda^l / \partial x^j = A_{k\ j}^{\ l} \dot{x}_\lambda^k, \tag{4.20}$$

where

$$A_{k\ j}^{\ l} = \frac{\partial^2 \bar{x}^l}{\partial x^k\, \partial x^j}, \tag{4.21}$$

and, since (4.19) is the inverse of (3.3), so that

$$A_{k\ j}^{\ l} = -A_h^l B_{p\ q}^{\ h} A_k^p A_j^q, \tag{4.22}$$

it follows that (4.20) can be expressed in the form

$$\partial \dot{\bar{x}}_\lambda^l / \partial x^j = -B_{p\ q}^{\ h} \dot{\bar{x}}_\lambda^p A_j^q A_h^l. \tag{4.23}$$

Let us now differentiate (4.17) with respect to x^j, obtaining

$$\frac{\partial X_\epsilon^r}{\partial x^j} = B_{i\ h}^{\ r} A_j^h \bar{X}_\epsilon^i + B_i^r \left[\frac{\partial \bar{X}_\epsilon^i}{\partial \bar{x}^h} A_j^h + \frac{\partial \bar{X}_\epsilon^i}{\partial \dot{\bar{x}}_\lambda^l} \frac{\partial \dot{\bar{x}}_\lambda^l}{\partial x^j} \right]. \tag{4.24}$$

The last term on the right-hand side of this equation can be reduced as follows. We substitute from (4.23), after which the $B_{p\ q}^{\ h}$ are replaced according to (4.16). This gives rise to the following expression:

$$-B_i^r \frac{\partial \bar{X}_\epsilon^i}{\partial \dot{\bar{x}}_\lambda^l} A_h^l A_j^q (B_s^h \bar{\Gamma}_{p\ q}^{\ s} - \Gamma_{u\ v}^{\ h} B_p^u B_q^v) \dot{\bar{x}}_\lambda^p, \tag{4.25}$$

and by means of (4.18) this is easily reduced to

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

$$(4.26)\qquad -B^r_i A^h_j \frac{\partial \bar{X}^i_\epsilon}{\partial \dot{\bar{x}}^v_\lambda} \bar{\Gamma}_p{}^v{}_h \dot{\bar{x}}^p_\lambda + \frac{\partial X^r_\epsilon}{\partial \dot{x}^h_\lambda} \Gamma_u{}^h{}_j \dot{x}^u_\lambda.$$

Thus (4.24) can be written in the form

$$(4.27)\qquad \frac{\partial X^r_\epsilon}{\partial x^j} - \frac{\partial X^r_\epsilon}{\partial \dot{x}^h_\lambda} \Gamma_u{}^h{}_j \dot{x}^u_\lambda = B_i{}^r{}_h A^h_j \bar{X}^i_\epsilon + B^r_i A^h_j \left(\frac{\partial \bar{X}^i_\epsilon}{\partial \bar{x}^h} - \frac{\partial \bar{X}^i_\epsilon}{\partial \dot{\bar{x}}^v_\lambda} \bar{\Gamma}_p{}^v{}_h \dot{\bar{x}}^p_\lambda \right).$$

By means of (4.16) the first term on the right-hand side of this relation can be expressed as

$$(4.28)\qquad A^h_j B^r_i \bar{\Gamma}_h{}^i{}_l \bar{X}^l_\epsilon - \Gamma_j{}^r{}_p X^p_\epsilon,$$

and when this is substituted in (4.27) we obtain a transformation law according to which the quantities defined by

$$(4.29)\qquad X^i_{\epsilon|j} = \frac{\partial X^i_\epsilon}{\partial x^j} - \frac{\partial X^i_\epsilon}{\partial \dot{x}^l_\lambda} \Gamma_p{}^l{}_j \dot{x}^p_\lambda + \Gamma_j{}^i{}_l X^l_\epsilon$$

emerge as components of a mixed tensor. This, then, represents *the covariant partial derivatives of the quantities $X^i_\epsilon(x^j, \dot{x}^j_\lambda)$ with respect to x^j.*

We shall conclude this section with a few general remarks concerning the process of covariant differentiation as exemplified by (4.29). First, it is easily seen that the definition of covariant derivative can be extended as usual to tensors of arbitrary rank, and that the usual sum and product rules hold, although the covariant derivative of a scalar ϕ is not, in general, its ordinary partial derivative (unless ϕ is independent of the $\dot{x}^j_\alpha$).

Secondly, the relation between (4.7) and (4.29) may be established as follows. According to (4.8) the $P_i{}^h{}_j{}^\alpha_\gamma$ are the components of an absolute mixed tensor of co- and contra-variant valencies unity under transformations of the t^α. According to (2.2) we therefore have

$$(4.30)\qquad \frac{\partial P_h{}^i{}_j{}^\alpha_\beta}{\partial \dot{x}^k_\epsilon} \dot{x}^k_\gamma = \delta^\epsilon_\beta P_h{}^i{}_j{}^\alpha_\gamma - \delta^\alpha_\gamma P_h{}^i{}_j{}^\epsilon_\beta,$$

and hence, contracting over ϵ, β and using (4.14) and (4.15), we find that

$$(4.31)\qquad \Gamma_k{}^i{}_j \dot{x}^k_\gamma = m^{-1}\left(P_k{}^i{}_j{}^\beta_\beta + \frac{\partial P_h{}^i{}_j{}^\alpha_\beta}{\partial \dot{x}^k_\beta} \dot{x}^h_\alpha \right) \dot{x}^k_\gamma$$

$$= m^{-1}(P_k{}^i{}_j{}^\beta_\beta \dot{x}^k_\gamma + m P_h{}^i{}_j{}^\alpha_\gamma \dot{x}^h_\alpha - P_h{}^i{}_j{}^\beta_\beta \dot{x}^h_\gamma) = P_h{}^i{}_j{}^\alpha_\gamma \dot{x}^h_\alpha,$$

and, in particular,

$$(4.32)\qquad \Gamma_k{}^i{}_j \dot{x}^k_\gamma \dot{x}^j_\epsilon = P_h{}^i{}_j{}^\alpha_\gamma \dot{x}^h_\alpha \dot{x}^j_\epsilon.$$

Thus, with $X^h_\epsilon = \dot{x}^h_\epsilon$ in (4.7), we have

$$(4.33)\qquad D\dot{x}^h_\epsilon / Dt^\gamma = \dot{x}_\epsilon{}^h{}_\gamma + \Gamma_k{}^h{}_j \dot{x}^k_\gamma \dot{x}^j_\epsilon.$$

However, for an arbitrary vector field $X^h(x^j, \dot{x}^j_\lambda)$ the partial derivative in (4.7) has to be interpreted as follows:

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

$$(4.34)\qquad \frac{\partial X^h_\epsilon}{\partial x^j}\dot{x}^j_\gamma + \frac{\partial X^h_\epsilon}{\partial \dot{x}^l_\lambda}\dot{x}_\lambda{}^l{}_\gamma.$$

Thus, if (4.29) is multiplied by $\dot{x}^j_\gamma$, the relations (4.31) and (4.33) being taken into account, we obtain by subtraction from (4.7):

$$(4.35)\qquad \frac{DX^h_\epsilon}{Dt^\gamma} - X^h_{\epsilon|j}\dot{x}^j_\gamma = \frac{\partial X^h_\epsilon}{\partial \dot{x}^l_\lambda}\cdot\frac{D\dot{x}^l_\gamma}{Dt^\lambda}.$$

This is the relation we have been seeking.

Thirdly, it should be remarked that (4.29) is not the only possible definition of covariant derivative. Since $\Gamma_j{}^i{}_l$ is not, in general, symmetric in j and l, an alternative tensorial expression results from (4.29) by the replacement of the last term on the right-hand side by $\Gamma_l{}^i{}_j X^l_\epsilon$. This procedure would, in fact, entail many advantages, but a reduction corresponding to (4.31) would be more cumbersome since (2.2) could not be invoked directly.

Finally, it may be asked whether the connection constructed in this section admits a counterpart of Ricci's lemma, i.e., whether the covariant derivative of the metric tensor vanishes identically. A little reflection shows that the answer to this must, in general, be in the negative, for it is not difficult to deduce that when $m = 1$ the connection coefficients (4.15) reduce to the parameters defined by Berwald for Finsler spaces (**7**, p. 79), and it is known that these give rise to a non-metric connection. Furthermore, a fairly long but straightforward calculation indicates that for $m > 1$ there cannot in general exist a *symmetric* connection for which the covariant derivatives of the metric tensor vanish. Whether or not it is possible to construct a non-symmetric connection for which the latter requirement is satisfied is still an open question.

5. The curvature tensors. The existence of connection coefficients satisfying the transformation law (4.16) naturally suggests the existence of curvature tensors with a corresponding theory of curvature. The present section is devoted to a brief discussion of these concepts. It will be seen that two distinct curvature tensors emerge from the theory mainly as a result of the lack of symmetry of our connection. To some extent this could possibly be avoided by suitably modifying our definition of covariant derivatives (for instance, by simply formulating all definitions in terms of the symmetric parts of the connection coefficients), but such a procedure does not seem to be very natural and many rather novel features inherent in the general theory would thus be obscured.

The first curvature tensor emerges directly from the integrability conditions associated with the transformation law (4.16) which we shall now investigate. To this end we differentiate (4.16) with respect to $\bar{x}^r$. In the course of this process we note that, in view of (3.10), we must write

$$(5.1)\qquad \frac{\partial \Gamma_l{}^i{}_h}{\partial \bar{x}^r} B^l_s B^h_t = \left(\frac{\partial \Gamma_l{}^i{}_h}{\partial x^k} B^k_r + \frac{\partial \Gamma_l{}^i{}_h}{\partial \dot{x}^p_\alpha} B_q{}^p{}_r \dot{\bar{x}}^q_\alpha\right) B^l_s B^h_t,$$

in which the term $B_q{}^p{}_r$ on the right-hand side is to be replaced according to (4.16), so that this expression is equal to

$$(5.2) \qquad B^k_r B_s{}^l B^h_t \left(\frac{\partial \Gamma_l{}^i{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^i{}_h}{\partial \dot{x}^p_\alpha} \Gamma_j{}^p{}_k \dot{x}^j_\alpha \right) + B^i_u \frac{\partial \bar{\Gamma}_s{}^u{}_t}{\partial \dot{\bar{x}}^v_\alpha} \bar{\Gamma}_q{}^v{}_r \dot{\bar{x}}^q_\alpha,$$

where, in the last step, we have made use of the fact that $\partial \Gamma_l{}^i{}_h / \partial \dot{x}^p_\alpha$ is a tensor of rank 4. Thus the derivative $B_s{}^i{}_{tr}$ of $B_s{}^i{}_t$ with respect to $\bar{x}^r$ can be expressed as

$$(5.3) \qquad B_s{}^i{}_{tr} = B^i_u \left(\frac{\partial \bar{\Gamma}_s{}^u{}_t}{\partial \bar{x}^r} - \frac{\partial \bar{\Gamma}_s{}^u{}_t}{\partial \dot{\bar{x}}^v_\alpha} \bar{\Gamma}_q{}^v{}_r \dot{\bar{x}}^q_\alpha \right) - B^l_s B^h_t B^k_r \left(\frac{\partial \Gamma_l{}^i{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^i{}_h}{\partial \dot{x}^p_\alpha} \Gamma_j{}^p{}_k \dot{x}^j_\alpha \right) + \bar{\Gamma}_s{}^u{}_t B_u{}^i{}_r - \Gamma_l{}^i{}_h (B^h_t B_s{}^l{}_r + B^l_s B_t{}^h{}_r).$$

In this relation we replace $B_u{}^i{}_r$, $B_s{}^l{}_r$ according to (4.16). After some simplification we thus obtain

$$(5.4) \qquad B_s{}^i{}_{tr} = B^i_u \left(\frac{\partial \bar{\Gamma}_s{}^u{}_t}{\partial \bar{x}^r} - \frac{\partial \bar{\Gamma}_s{}^u{}_t}{\partial \dot{\bar{x}}^v_\alpha} \bar{\Gamma}_q{}^v{}_r \dot{\bar{x}}^q_\alpha + \bar{\Gamma}_v{}^u{}_r \bar{\Gamma}_s{}^v{}_t \right) - B^l_s B^h_t B^k_r \left(\frac{\partial \Gamma_l{}^i{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^i{}_h}{\partial \dot{x}^p_\alpha} \Gamma_j{}^p{}_k \dot{x}^j_\alpha - \Gamma_j{}^i{}_h \Gamma_l{}^j{}_k \right) - \Gamma_p{}^i{}_q B^p_u (\bar{\Gamma}_s{}^u{}_t B^q_r + \bar{\Gamma}_s{}^u{}_r B^q_t) - \Gamma_l{}^i{}_h B^l_s B_t{}^h{}_r.$$

The required integrability condition is tantamount to the stipulation that the right-hand side be symmetric in t and r. Noting that the last three terms on this side are symmetric in these indices, we see that this symmetry condition is equivalent to the following:

$$(5.5) \qquad B^i_u \bar{K}_s{}^u{}_{tr} - B^l_s B^h_t B^k_r K_l{}^i{}_{hk} = 0,$$

where we have written

$$(5.6) \qquad K_l{}^i{}_{hk} = \left(\frac{\partial \Gamma_l{}^i{}_h}{\partial x^k} - \frac{\partial \Gamma_l{}^i{}_h}{\partial \dot{x}^p_\alpha} \Gamma_j{}^p{}_k \dot{x}^j_\alpha \right) - \left(\frac{\partial \Gamma_l{}^i{}_k}{\partial x^h} - \frac{\partial \Gamma_l{}^i{}_k}{\partial \dot{x}^p_\alpha} \Gamma_j{}^p{}_h \dot{x}^j_\alpha \right) + \Gamma_j{}^i{}_k \Gamma_l{}^j{}_h - \Gamma_j{}^i{}_h \Gamma_l{}^j{}_k.$$

The relation (5.5) states that the quantities (5.6) are the components of a tensor of contravariant valency 1 and covariant valency 3: this tensor is our *first curvature tensor.*

The other curvature tensor is obtained from the commutation rules satisfied by our covariant derivatives. Let us consider the covariant derivative of $X^i_{\epsilon|k}$ with respect to x^h. By definition, this is

$$(5.7) \qquad X^i_{\epsilon|k|h} = \frac{\partial X^i_{\epsilon|k}}{\partial x^h} - \frac{\partial X^i_{\epsilon|k}}{\partial \dot{x}^l_\beta} \Gamma_q{}^l{}_h \dot{x}^q_\beta + \Gamma_h{}^i{}_j X^j_{\epsilon|k} - \Gamma_h{}^j{}_k X^i_{\epsilon|j}.$$

We substitute in this expression from (4.29) and carry out the operations indicated in (5.7). After some rearrangement we obtain the following expression:

$$(5.8)\qquad \begin{aligned} X^i_{\epsilon|k|h} = {} & X^j_\epsilon\left[\frac{\partial\Gamma_k{}^i{}_j}{\partial x^h} - \frac{\partial\Gamma_k{}^i{}_j}{\partial \dot{x}^l_\alpha}\Gamma_p{}^l{}_h\dot{x}^p_\alpha + \Gamma_h{}^i{}_p\Gamma_k{}^p{}_j\right] \\ & - \frac{\partial X^i_\epsilon}{\partial \dot{x}^l_\alpha}\left[\frac{\partial\Gamma_p{}^l{}_k}{\partial x^h} - \frac{\partial\Gamma_p{}^l{}_k}{\partial \dot{x}^q_\alpha}\Gamma_j{}^q{}_h\dot{x}^j_\alpha - \Gamma_j{}^l{}_k\Gamma_p{}^j{}_h\right]\dot{x}^p_\alpha \\ & - \Gamma_h{}^j{}_k X^i_{\epsilon|j} + \frac{\partial^2 X^i_\epsilon}{\partial x^h\partial x^k} - \left(\frac{\partial^2 X^i_\epsilon}{\partial x^h\partial \dot{x}^j_\alpha}\Gamma_p{}^j{}_k + \frac{\partial^2 X^i_\epsilon}{\partial x^k\partial \dot{x}^j_\alpha}\Gamma_p{}^j{}_h\right)\dot{x}^p_\alpha \\ & + \frac{\partial^2 X^i_\epsilon}{\partial \dot{x}^j_\alpha\partial \dot{x}^l_\beta}\Gamma_p{}^j{}_k\Gamma_q{}^l{}_h\dot{x}^p_\alpha\dot{x}^q_\beta + \left(\Gamma_k{}^i{}_j\frac{\partial X^j_\epsilon}{\partial x^h} + \Gamma_h{}^i{}_j\frac{\partial X^j_\epsilon}{\partial x^k}\right) \\ & - \frac{\partial X^j_\epsilon}{\partial \dot{x}^l_\alpha}(\Gamma_k{}^i{}_j\Gamma_p{}^l{}_h + \Gamma_h{}^i{}_j\Gamma_p{}^l{}_k)\dot{x}^p_\alpha. \end{aligned}$$

In this relation we interchange the indices h, k, and subtract the result from (5.8), noting that the sum of the last eight terms on the right-hand side of (5.8) is symmetric in these indices. We thus obtain the following commutation rules:

$$(5.9)\qquad X^i_{\epsilon|k|h} - X^i_{\epsilon|h|k} = X^j_\epsilon K^*{}_j{}^i{}_{kh} - \frac{\partial X^i_\epsilon}{\partial \dot{x}^l_\alpha}K_p{}^l{}_{kh}\dot{x}^p_\alpha + X^i_{\epsilon|j}T_k{}^j{}_h,$$

where

$$(5.10)\qquad T_k{}^j{}_h = \Gamma_k{}^j{}_h - \Gamma_h{}^j{}_k$$

is the *torsion tensor* associated with our connection, and where the *second curvature tensor* is defined by

$$(5.11)\qquad \begin{aligned} K^*{}_j{}^i{}_{kh} = {} & \left(\frac{\partial\Gamma_k{}^i{}_j}{\partial x^h} - \frac{\partial\Gamma_k{}^i{}_j}{\partial \dot{x}^l_\alpha}\Gamma_p{}^l{}_h\dot{x}^p_\alpha\right) \\ & - \left(\frac{\partial\Gamma_h{}^i{}_j}{\partial x^k} - \frac{\partial\Gamma_h{}^i{}_j}{\partial \dot{x}^l_\alpha}\Gamma_p{}^l{}_k\dot{x}^p_\alpha\right) + \Gamma_h{}^i{}_p\Gamma_k{}^p{}_j - \Gamma_k{}^i{}_p\Gamma_h{}^p{}_j. \end{aligned}$$

The tensor appearing in the second term on the right-hand side of (5.9) is, of course, our first curvature tensor.

Clearly the two curvature tensors would coincide if the underlying connection were symmetric. However, a straightforward calculation shows that they are, in general, related according to the following identity:

$$(5.12)\qquad \begin{aligned} K^*{}_j{}^i{}_{kh} = {} & K_j{}^i{}_{kh} + T_k{}^i{}_{j|h} - T_h{}^i{}_{j|k} + T_l{}^i{}_j T_h{}^l{}_k \\ & + (T_k{}^i{}_l T_h{}^l{}_j - T_h{}^i{}_l T_k{}^l{}_j). \end{aligned}$$

In conclusion we shall briefly derive the identities satisfied by our curvature tensors. First, it is easily verified by direct calculation that

$$(5.13)\qquad K^*{}_j{}^i{}_{hk} + T_j{}^i{}_{h|k} + T_l{}^i{}_j T_h{}^l{}_k + \text{cycl.}(j,h,k) = 0,$$

https://doi.org/10.4153/CJM-1968-062-1 Published online by Cambridge University Press

where cycl.(j, h, k) denotes the sum of two sets of three terms each obtained by replacing the indices (j, h, k) by first (h, k, j) and then by (k, j, h) in the expression on the left-hand side of (5.13). If (5.12) is substituted in (5.13), it is found that

$$K_j{}^i{}_{hk} + T_h{}^i{}_{j|k} + T_j{}^i{}_l T_h{}^l{}_k + \text{cycl.}(j, h, k) = 0. \tag{5.14}$$

In order to obtain the counterparts of the Bianchi identities, we consider m linearly independent covariant vector fields $Y_i^\epsilon(x^j)$, which are functions of the positional coordinates only. (Since $\partial T^{i\cdots}{}_{j\ldots}/\partial \dot{x}_\alpha^k$ is always tensorial if $T^{i\cdots}{}_{j\ldots}(x^l, \dot{x}_\epsilon^l)$ is a tensor, this is an invariant and consequently justifiable requirement which is imposed merely to simplify the analysis below.) As in (5.9) we have

$$Y^\epsilon_{i|j|k} - Y^\epsilon_{i|k|j} = -Y^\epsilon_r K^*{}_i{}^r{}_{jk} + Y^\epsilon_{i|r} T_j{}^r{}_k. \tag{5.15}$$

In the relation obtained by taking the covariant derivative of (5.15) with respect to x^h the indices j, h, k are cyclically interchanged, and the resulting three equations are added. After some rearrangement one thus obtains

$$\begin{aligned} Y^\epsilon_{i|j|k|h} - Y^\epsilon_{i|j|h|k} + \text{cycl.}(j, h, k) = \\ -Y^\epsilon_r K^*{}_i{}^r{}_{jk|h} + Y^\epsilon_{i|r} T_j{}^r{}_{k|h} - Y^\epsilon_{r|j} K^*{}_i{}^r{}_{kh} \\ + Y^\epsilon_{i|r|h} T_j{}^r{}_k + \text{cycl.}(j, h, k). \end{aligned} \tag{5.16}$$

The commutation rule corresponding to $Y^\epsilon_{i|j}$ is now applied to the left-hand side of (5.16), which gives rise to the following expression:

$$-Y^\epsilon_{i|r} K^*{}_j{}^r{}_{kh} - Y^\epsilon_{r|j} K^*{}_i{}^r{}_{kh} + Y^\epsilon_{i|j|r} T_k{}^r{}_h - \frac{\partial Y^\epsilon_{i|j}}{\partial \dot{x}^l_\alpha} K_p{}^l{}_{kh} \dot{x}^p_\alpha + \text{cycl.}(j, h, k).$$

When this is substituted in (5.16), the latter can be simplified to read as follows:

$$\begin{aligned} Y^\epsilon_r K^*{}_i{}^r{}_{jk|h} + \text{cycl.}(j, h, k) = \\ Y^\epsilon_{i|r} (K^*{}_j{}^r{}_{kh} + T_j{}^r{}_{k|h}) + \frac{\partial Y^\epsilon_{i|j}}{\partial \dot{x}^l_\alpha} K_p{}^l{}_{kh} \dot{x}^p_\alpha \\ + T_j{}^r{}_k (Y^\epsilon_{i|r|h} - Y^\epsilon_{i|h|r}) + \text{cycl.}(j, h, k). \end{aligned} \tag{5.17}$$

Again, it follows almost directly from (5.15) that

$$\begin{aligned} T_j{}^r{}_k (Y^\epsilon_{i|r|h} - Y^\epsilon_{i|h|r}) + \text{cycl.}(j, h, k) = \\ Y^\epsilon_{i|r} T_l{}^r{}_h T_j{}^l{}_k - Y^\epsilon_r K^*{}_i{}^r{}_{lh} T_j{}^l{}_k + \text{cycl.}(j, h, k), \end{aligned} \tag{5.18}$$

while according to the definition of $Y^\epsilon_{i|j}$ we have

$$\frac{\partial Y^\epsilon_{i|j}}{\partial \dot{x}^l_\alpha} = - Y^\epsilon_r \frac{\partial \Gamma_j{}^r{}_i}{\partial \dot{x}^l_\alpha}. \tag{5.19}$$

When (5.18) is substituted in (5.17), we see that the coefficient of $Y^{\epsilon}_{i|r}$ on the right-hand side vanishes identically by virtue of (5.13). Also, because of (5.19), the remaining terms all involve Y^{ϵ}_{r}. Since the latter are essentially arbitrary, it follows that their coefficients must vanish, which is equivalent to the relation

$$K^*{}_i{}^r{}_{jk|h} + \frac{\partial \Gamma_j{}^r{}_i}{\partial \dot{x}^l_\alpha} K_p{}^l{}_{kh}\, \dot{x}^p_\alpha + T_j{}^l{}_k K^*{}_i{}^r{}_{lh} + \text{cycl.}(j, h, k) = 0. \tag{5.20}$$

These, then, are the *generalized Bianchi identities* we have been seeking. It is easily verified that for $m = 1$ they reduce to the Bianchi identities of Finsler geometry with respect to Berwald's connection (**7**, p. 128), in which case the tensor (5.10) vanishes.

These results suggest the investigation of several problems usually associated with the theory of curvature. We hope to return to some of these matters later.

References

1. C. Carathéodory, *Über die Variationsrechnung bei mehrfachen Integralen*, Acta Sci. Math. (Szeged), *4* (1929), 193–216.

2. E. Cartan, *Les espaces métriques fondés sur la notion d'aire*, Actualités scientifiques, *72* (Hermann, Paris, 1933).

3. E. T. Davies, *Areal spaces*, Ann. Mat. Pura Appl. (4), *55* (1961), 63–76.

4. J. Douglas, *Systems of K-dimensional manifolds in N-dimensional space*, Math. Ann., *105* (1931), 707–733.

5. A. Kawaguchi, *Theory of areal spaces*, Mat. e Rend. Appl. (5), *12* (1953), 373–386.

6. D. H. Martin, *The local geometry and extremal surfaces of areal spaces*, Ph.D. thesis, University of South Africa (Pretoria, 1965).

7. H. Rund, *The differential geometry of Finsler spaces* (Springer, Berlin-Göttingen–Heidelberg, 1959).

8. ——— *The Hamilton-Jacobi theory in the calculus of variations* (Van Nostrand, London and New York, 1966).

9. H. Weyl, *Observations on Hilbert's independence theorem and Born's quantization of field equations*, Phys. Rev. (2), *46* (1934), 505–508.

10. ——— *Geodesic fields in the calculus of variations for multiple integrals*, Ann. of Math. (2), *36* (1935), 607–629.

University of the Witwatersrand,
Johannesburg, South Africa