SOME DIFFERENTIAL-GEOMETRIC ASPECTS OF THE LAGRANGE VARIATIONAL PROBLEM

BY

ROBERT HERMANN[1]

1. Introduction

Our principal aim is to present results relating the classical theory [2], [4] of the Lagrange problem to topics of current interest in differential geometry. We also sketch a reformulation of the classical theory in differential-geometric language, guided especially by E. Cartan's treatment of the ordinary variational problem in [7].

Let B be a manifold, let $T(B)$ be its tangent bundle, and let $M = T(B) \times R$ (R = real numbers). An *ordinary* variational problem is defined by a *Lagrangian*, i.e., by a real-valued function L on M, denoted by $L(v, t)$, $v \in T(B)$, $t \in R$. Such a function defines by integration a real-valued function $\mathbf{L}$ on the space of curves of B. The *extremal curves* are the solutions of the Euler equations, i.e., the "critical points" of the real-valued function $\mathbf{L}$ defines on the space P of curves of B joining two fixed points. However, in the regular cases the extremals can be defined without reference to coordinate systems as follows: There is a vector field on M such that the projection in B of its integral curves are the extremal curves. Further, there is a closed 2-differential-form on M which is an "invariant integral" of the vector field, i.e., the orbit space of M under the one-parameter group generated by the vector field locally has a symplectic manifold structure. In the nonregular cases, this 2-form is still defined, and the extremals in B are the projections of its characteristic curves.

A Lagrange problem can be formulated as follows. In addition to the Lagrangian L on M, a "constraint submanifold" $S \subset M$ is given: The *extremals* of the Lagrange problem are the "critical points" (in the sense of infinite-dimensional differential geometry) of the restriction of $\mathbf{L}$ to the space $P(S)$ of curves of B joining two fixed points whose tangent vector lies in S.

The *formal extremals* are the curves in B given by the classical Lagrange multiplier rule which, in the regular cases, can be described as follows: Let M' be the vector bundle of normal vectors to S, a manifold of the same dimension as M. There is a vector field on M', again having a closed 2-differential-form as an "invariant integral" whose integral curves projected in B are the formal extremals. In the nonregular cases, the 2-form on M' can still be defined, and the formal extremals are the projection in B of its characteristic

Received May 25, 1961; received in revised form August 27, 1961.

[1] Lincoln Laboratory, Massachusetts Institute of Technology, operated with support from the U. S. Army, Navy, and Air Force.

curves. (However, it is a nontrivial matter to prove that the extremals are also formal extremals [2].)

The *class* of an extremal is a very important invariant, measuring a type of degeneracy condition that does not occur in the theory of the ordinary problem. To describe it, suppose that $M' = S \times R^p$, where $p = \dim M - \dim S$, and λ_a $(1 \leqq a, b, \cdots \leqq p)$ are coordinates in R^p. There are 1-differential-forms θ_a and θ on S locally defined by the variational problem. A curve $\sigma(t)$, $0 \leqq t \leqq 1$, is a formal extremal if and only if its tangent vector curve $\tau(t)$ in S satisfies: There is a curve $\lambda_a(t)$ in R^p such that the curve $(\sigma(t), \lambda_a(t))$ in $M' = S \times R^p$ is a characteristic curve for the 2-form $d(\theta + \lambda_a \theta_a)$ on M' and $\theta_a(\tau'(t)) = 0$ ($\tau'(t)$ is the tangent vector to τ at t). These two conditions are equivalent to the conditions

1.1.(a) $\quad \theta_a(\tau'(t)) = 0, \quad$ and

$$\text{(b)} \quad \tau'(t) \,\lrcorner\, d\theta + (d\lambda_a(t)/dt)\theta_a + \lambda_a(t)(\tau'(t) \,\lrcorner\, d\theta_a) = 0.$$

(If ω is a p-differential-form on a space, v a tangent vector, $v \,\lrcorner\, \omega$ is the inner product or contraction of ω by v, a $(p - 1)$-covector of the space at the same point as v.) The curve $\lambda_a(t)$ is the *Lagrange multiplier curve* associated with σ. Notice that it is not necessarily unique: The difference $u_a(t)$ of two such multipliers satisfies

$$1.2. \qquad u_a(t)(\sigma'(t) \,\lrcorner\, d\theta_a) + (du_a/dt)\theta_a = 0.$$

This is a system of linear homogeneous differential equations for $u_a(t)$. The dimension of the set of solutions is called the *class* of the extremal σ. If it is zero, $P(S)$ may be thought of as being a "regular submanifold" of P locally about σ in the sense of the differential geometry of path spaces. On the other hand, Carathéodory has defined [5] an integer called the class attached to σ, whose vanishing ensures that all points close to $\sigma(1)$ can be reached as end-points of formal extremals starting at $\sigma(0)$. These two notions of class thus have different geometric meaning and are not necessarily equal. One of our main results (Section 8) gives a sufficient condition that both classes be zero. In general, these matters lie at the heart of what may be called the "differential geometry of submanifolds of path spaces" and, although local in nature, will be found useful in possible attempts to extend Morse theory to the Lagrange problem.

Another topic we shall treat is suggested by the fact that the "space" of all formal extremals of a Lagrange problem would be a symplectic manifold if it were a manifold at all; hence the usual correspondence between functions on the space of formal extremals and one-parameter groups of symplectic automorphisms makes some sort of sense. We shall apply this idea to prove the following result: Suppose X is a vector field on B such that the one-parameter pseudogroup of diffeomorphisms of B generated by X, when extended to $T(B) \times R = M$, leaves L and S invariant, i.e., maps extremal into extremals. Generalizing the same question dealt with earlier for Riemannian geometry [14], we ask for the points $b \in B$ such that the integral curve of X starting at b is itself a formal extremal. Modulo degeneracy problems, the answer in the case where L and S are time-independent is that they are the critical points of the function $b \to L(b, X(b))$ restricted to the submanifold $\{b \in B : X(b) \in S\}$ of B.

It appears that the most interesting special case of the Lagrange problem occurs when L defines a Riemannian metric on B and where the constraint manifold S is linear. We apply the general theory of the Lagrange problem to this situation. We shall state here our principal results on this topic.

A. *Let B be a Riemannian manifold, and let $T(B)$ be the tangent bundle to B. $T(B)$ has a natural Riemannian metric defined by the Levi-Civita parallelism. Consider the Lagrange Problem on $T(B)$ obtained by taking the Riemannian metric as Lagrangian, and the constrained curves as those which are horizontal, i.e., perpendicular to the fibres of $T(B)$. If $t \to v(t)$, $0 \leqq t \leqq 1$, is a minimizing curve for this Lagrange problem and $t \to \sigma(t)$ is its projection in B, then σ is the solution of the following variational problem on B:*

Given tangent vectors v_0, $v_1 \in T(B)$, with b_0, $b_1 \in B$ their projections in B, σ is the shortest curve joining b_0 to b_1 among the class of all curves joining b_0 to b_1 such that v_0 parallel-translates along the curve to v_1.

If σ and v define a formal extremal of this Lagrange problem, and σ is parameterized proportionally to arc-length, there is a vector field $\lambda(t)$ along σ such that

(a) $\nabla v = 0$, (b) $\nabla \lambda = \sigma'$,
(c) $(R(\sigma', u)(\lambda), v) = 0$ *for all vector fields $u : t \to u(t)$ along σ.*

∇v is the covariant derivative vector field of v along σ, σ' is the tangent vector field to σ, and $R(\ ,\ )(\ )$ is the Riemannian curvature tensor. (As a first observation, these equations are interesting because they involve the curvature tensor directly at the first order, instead of at the second order as do the equations of the geodesics.)

B. *With the notations of* A, *suppose that X is a Killing vector field on B, i.e., X generates a one-parameter pseudogroup of isometries of B. Let X' be the vector field on $T(B)$ that is the first-order prolongation of X. Then, X' is a Killing vector field with respect to the metric on $T(B)$, and preserves the extremals of the Lagrange problem on $T(B)$ defined by* A. *Given a $v \in T(B)$, with b its projection on B, the integral curve of X' starting at v is horizontal if and only if*

$$\nabla_v X = 0.$$

Any integral curve of X' is a geodesic of the metric on $T(B)$ if and only if it is horizontal and its projection into B is a geodesic. The integral curve of X' at a $v \in T(B)$, satisfying $\nabla_v X = 0$ is an extremal of the Lagrange variational *problem on* $T(B)$ *defined in* A *if and only if there is a tangent vector* λ *at* b *such that*

$$1.3. \qquad (a) \quad \nabla_\lambda X = 0, \qquad (b) \quad R(\lambda, v)(X(b)) = \nabla_{X(b)} X.$$

C. *Now let* B *be a Riemannian manifold, and let* $H : b \to H_b \subset B_b$ (B_b = *tangent space to* B *at* b) *be any field of tangent subspaces of constant dimension on* B, *i.e.,* H *defines a Pfaffian system on* B. *Consider the Lagrange problem of minimizing the length of curves which are integral curves of the Pfaffian system* H, *i.e., satisfy*

$$\sigma'(t) \in H_{\sigma(t)} .$$

Such a σ *parameterized proportionally to arc-length is a formal extremal if and only if there is a vector field* λ *along* σ *perpendicular to* H *with*

$$\nabla\sigma' + \nabla\lambda = Q(\sigma', \lambda),$$

where $Q(\ ,\ )$ *is a tensor field on* M *determined by* H. *If* B *is complete, these extremal curves* σ *can be extended indefinitely. There is a tensor field on* M *whose vanishing is necessary and sufficient that every geodesic of* M *that is tangent to* H *at one point be everywhere tangent. (This is a generalization of a result of B. Reinhart* [17] *in the case where* H *is the horizontal field associated with a foliation on* B.) *A real-valued function* S *on* B *is a solution of the Hamilton-Jacobi equation associated with the Lagrange problem if*

$$|\, \pi(\text{grad } S) \,| = F(S),$$

where π *is the projection of vector fields on* B *parallel to* H, *where* grad S *is the usual gradient vector field of a function defined by the Riemannian metric, where* $|\, X \,|$ *is the length function of the vector field* X, *and where* $F(\ )$ *is some function of one variable. The corresponding field of extremals of the Lagrange problem is formed by the integral curves of the vector field*

$$\pi(\text{grad } S).$$

2. Symplectic foliations

We continue the notations of [12] and [13], which we briefly summarize. All manifolds, maps, curves, tensor fields, etc. will be of differentiability class C^∞ unless mentioned otherwise. If M is a manifold (usually considered as connected and paracompact), M_x denotes the tangent space to M at a point x. If $\phi : M \to M'$ is a map of manifolds, $\phi_* : M_x \to M'_{\phi(x)}$, for $x \in M$, is the linear map ϕ induces on tangent vectors at x. In the special case $M = [a, b] = \{t \in R : a \leqq t \leqq b\}$, i.e., ϕ defines a curve in M', $\phi'(t)$ denotes the tangent vector to ϕ at t. Let $V(M)$ be the set of vector fields on M, considered as a module over the ring $C(M)$ of real-valued C^∞ functions on M and as a real Lie algebra with respect to the Jacobi bracket operation

$[X, Y]$. If X is a vector field and ω is a p-differential-form on M, denote by

(a) $X(\omega)$, the Lie derivative of ω by X, a p-form;
(b) $X \lrcorner\, \omega$, the inner product or contraction of ω by X, a $(p-1)$-form;
(c) $d\omega$, the exterior derivative of ω, a $(p+1)$-form.

Recall the following rules these operations obey (omitting the standard bilinearity conditions). In the following formulas, $X, Y, Z, \cdots$ denote vector fields, $\omega_1, \omega_2, \cdots$ denote differential forms.

2.1. $$X(\omega_1 \wedge \omega_2) = X(\omega_1) \wedge \omega_2 + \omega_1 \wedge X(\omega_2)$$
($\wedge$ denotes the exterior product of forms).

2.2. $$[X, Y](\omega) = X(Y(\omega)) - Y(X(\omega)).$$

2.3. $$X(Y \lrcorner\, \omega) = [X, Y] \lrcorner\, \omega + Y \lrcorner\, (X(\omega)).$$

2.4. $$X(\omega) = d(X \lrcorner\, \omega) + X \lrcorner\, d\omega.$$

2.5. $$X \lrcorner\, (\omega_1 \wedge \omega_2) = (X \lrcorner\, \omega_1) \wedge \omega_2 + (-1)^p \omega_1 \wedge (X \lrcorner\, \omega_2)$$
($p =$ degree ω_1).

Let H be a real subspace of $V(M)$. For $x \in M$, define

$$\mathbf{H}_x = \{v \in M_x : \exists X \in \mathbf{S} \text{ such that } X(x) = v\}.$$

A submanifold N of M is called an integral manifold of $\mathbf{H}$ providing $N_x \subset \mathbf{H}_x$ for all $x \in N$. (N_x is to be identified with a subspace of M_x.) In particular, a curve $\sigma : [a, b] \to M$ is an *integral curve* of $\mathbf{H}$ if $\sigma'(t) \in \mathbf{H}_{\sigma(t)}$ for $a \leqq t \leqq b$. For $x \in M$, let

$$L^x = \{y \in M : y \text{ can be joined to } x \text{ by an integral curve of } \mathbf{H}\}.$$

The following facts are proved in [13], and are standard if dim $\mathbf{H}_x$ is constant for $x \in M$.

2.6. *Suppose that*
(a) *For every integral curve* $\sigma : [a, b] \to M$ *of* $\mathbf{H}$, dim $\mathbf{H}_{\sigma(t)}$ *is independent of* t, $a \leqq t \leqq b$,
(b) $[\mathbf{H}, \mathbf{H}] \subset \mathbf{H}$.

Then for all $x \in M$, L^x *is a maximal connected integral submanifold of* $\mathbf{H}$ *with* $L^x_y = \mathbf{H}_y$ *for all* $y \in L^x$. $\mathbf{H}$ *is said to define a foliation on* M, *and* L^x *is the leaf of* $\mathbf{H}$ *through* x. *The foliation is said to be nonsingular if* dim $\mathbf{H}_x$ *is constant for* $x \in M$.

2.7. *Suppose that* $[\mathbf{H}, \mathbf{H}] \subset \mathbf{H}$ *and that* $\mathbf{H}$ *is locally finitely generated in the sense that each* $x \in M$ *has a neighborhood* U *with a finite-dimensional subspace* $\mathbf{H}_U$ *such that every element of* $\mathbf{H}$ *can, in* U, *be written as a linear combination of elements of* $\mathbf{H}_U$ *with coefficients from* $C(U)$. *Then* $\mathbf{H}$ *defines a foliation on* M, *i.e.*, 2.6(a) *is satisfied.*

From now on, we shall denote foliations by $\mathbf{F}$, and more general subspaces of $V(M)$ that are not necessarily subalgebras of $V(M)$ by $\mathbf{H}$. Let ω be a closed 2-differential-form on M. Let

$$\mathbf{F} = \{X \in V(M) : X \lrcorner\, \omega = 0\}.$$

Note that the following facts follow from 2.1–2.5.

2.8. *If* $X \in \mathbf{F}$, *then* $X(\omega) = d(X \lrcorner\, \omega) + X \lrcorner\, d\omega = 0$.

2.9. $[\mathbf{F}, \mathbf{F}] \subset \mathbf{F}$.

2.10. *If* $Y \in V(M)$ *satisfies* $Y(\omega) = f\omega$, *for* $f \in C(M)$, *then* $[Y, \mathbf{F}] \subset \mathbf{F}$. *If further* $Y(\omega) = 0$, *then* $d(Y \lrcorner\, \omega) = 0$.

In this section, we shall assume further that

2.11. $\dim \mathbf{F}_x$ *is constant for* $x \in M$, *i.e., the form* ω *has constant rank on* M.

$\mathbf{F}$ then defines a (nonsingular) foliation on M, the *characteristic foliation* of ω [9]. From 2.8, we see that ω is a base-like form for this foliation and defines a *symplectic manifold* structure on $M/\mathbf{F}$, the space of leaves of $\mathbf{F}$, if it is a manifold at all. (More precisely, suppose U is an open set of M having a map $\phi : U \to B$ that is a decomposition map for the foliation $\mathbf{F}$ restricted to U [12]. There is a closed 2-form ω_ϕ on B such that

$$\text{(a)}\quad \phi^*(\omega_\phi) = \omega, \qquad \text{(b)}\quad \omega_\phi \wedge \cdots \wedge \omega_\phi \ (p\text{-times}) \neq 0,$$

where $2p = \dim B = \dim M - \dim \mathbf{F}$.)

Suppose now that θ is a 1-form on M such that $d\theta = w$. Let U be an open set of M. Adopt the range of indices $1 \leqq i, j, \cdots \leqq n$ and the summation convention. Suppose (x_i, y_i, t) is a set of functions in U such that

$$\theta = y_i\, dx_i - H\, dt,$$

where H is a function in U that is functionally dependent on (x_i, y_i, t). θ will be said to be in *Hamiltonian form*. If $H =$ constant, and the (x_i, y_i, t) are functionally independent, θ will be said to be in *normal form*. Notice then that $X \in V(U)$ is a characteristic vector field for $d\theta$ if and only if

$$X(y_i)\, dx_i - X(x_i)\, dy_i - X(H)\, dt + X(t)\, dH = 0.$$

Consider a $(2n + 1)$-dimensional Euclidean space R^{2n+1}, with coordinates (u_i, v_i, s). Construct a mapping $\phi : U \to R^{2n+1}$ by the conditions:

$$u_i(\phi(m)) = x_i(m), \qquad v_i(\phi(m)) = y_i(m), \qquad s(\phi(m)) = t(m) \qquad \text{for } m \in U.$$

Suppose that $H'(u, v, s)$ is a function on R^{2n+1} such that

$$\phi^*(H') = H.$$

Then also $\phi^*(\theta') = \theta$, where θ' is the 1-form $v_i\,du_i - H'\,ds$ on R^{2m+1}. In particular, ϕ maps a characteristic curve of $d\theta = w$ into a characteristic curve of $d\theta'$. However, it is easy to see, by using 2.5, that the characteristic curves σ of $d\theta'$ have the following properties:

2.12.(a) *If σ has an everywhere nonzero tangent vector, its parameterization can be changed so that it is defined by giving u_i and v_i as functions of s.*

(b) *u_i and v_i as functions of s satisfy the Hamilton equations with Hamiltonian H':*

$$\frac{d}{ds}u_i = \frac{\partial H'}{\partial v_i}(u, v, s), \qquad \frac{d}{ds}v_i = -\frac{\partial H'}{\partial u_i}(u, v, s).$$

Conversely, every solution of these equations determines a characteristic curve of $d\theta'$. This provides the link with classical Hamilton-Jacobi theory. ϕ is usually called a *Legendre transformation*. However, following Cartan [7] we shall usually work with the more geometric differential-form formalism.

Suppose $Y \in V(M)$ satisfies $Y(\omega) = 0$, and M has zero first Betti-number. Since $d(Y \lrcorner\, \omega) = 0$, there is a function $f_Y \in C(M)$ with $df_Y = Y \lrcorner\, \omega$. A critical point x_0, of f_Y, i.e., a point at which $df_Y = 0$, is then a point for which $Y(x_0) \in \mathbf{F}_{x_0}$. Since $[Y, \mathbf{F}] \subset \mathbf{F}$, this means that the integral curve of Y starting at x_0 lies in the leaf of $\mathbf{F}$ through x_0, i.e., the integral curve is a characteristic curve of ω.

Let x_0 be a critical point of f_Y. If $X \in \mathbf{F}$, notice that then

2.13. $dX(f_Y) = [X, Y] \lrcorner\, \omega + Y \lrcorner\, X(\omega) = 0$; *hence the whole leaf of* $\mathbf{F}$ *through x_0 is a critical manifold for f_Y.*

The Hessian of f_Y at x_0 is a symmetric, bilinear form h on M_{x_0} defined as follows [1]: For $v_1, v_2 \in M_{x_0}$, let $X_1, X_2 \in V(M)$ be such that $X_1(x_0) = v_1$, $X_2(x_0) = v_2$. Then

2.14. $h(v_1, v_2) = X_1(X_2(f_Y))(x_0)$. *(It is easily seen that this is independent of the X_1, X_2 chosen.)*

We see that

2.15.(a) *For $v \in \mathbf{F}_{x_0}$, $h(v, M_{x_0}) = 0$.*

(b) *Let* $\mathbf{h}$ *(resp.* $\boldsymbol{\omega}$*) denote the bilinear form induced on $M_{x_0}/\mathbf{F}_{x_0}$ by h (resp. ω). Since $Y(x_0) \in \mathbf{F}_{x_0}$, Y induces a linear transformation $l_Y : M_{x_0}/\mathbf{F}_{x_0} \to M_{x_0}/\mathbf{F}_{x_0}$ by passing to the quotient from* Ad $Y : V(M) \to V(M)$ *defined by* Ad $Y(X) = [Y, X]$. *Then* $\mathbf{h}(\mathbf{v}_1, \mathbf{v}_2) = \boldsymbol{\omega}(\boldsymbol{l}_Y(\mathbf{v}_1), \mathbf{v}_2)$, *for $v_1, v_2 \in M_{x_0}$, where $\mathbf{v}_1$ denotes the image of v_1 in $M_{x_0}/\mathbf{F}_{x_0}$.*

The geometric interpretation of 2.15(b) is as follows: Suppose that $M/\mathbf{F}$ is a manifold and that $\phi : M \to M/\mathbf{F}$ is the decomposition map. There is a form $\boldsymbol{\omega}$ on $M/\mathbf{F}$ such that $\phi^*(\boldsymbol{\omega}) = \omega$. The tangent space to $M/\mathbf{F}$ at $\phi(x_0)$ is isomorphic under ϕ_* to $M_{x_0}/\mathbf{F}$. There is a vector field $\mathbf{Y}$ on $M/\mathbf{F}$ such that $\phi_*(Y(x)) = \mathbf{Y}(\phi(x))$ for all $x \in M$, and $\mathbf{Y}(\boldsymbol{\omega}) = 0$. $\mathbf{Y}(\phi(x_0)) = 0$, i.e., $\mathbf{Y}$

has a singular point at $\phi(x_0)$. f_Y can be chosen as $\phi^*(f_Y)$, with $df_{\mathbf{Y}} = \mathbf{Y} \lrcorner \boldsymbol{\omega}$ (assuming $M/\mathbf{F}$ also has zero first Betti-number); hence $\mathbf{Y}$ has a nondegenerate critical point at $\phi(x_0)$ if and only if $f_{\mathbf{Y}}$ has a nondegenerate point at $\phi(x_0)$, i.e., if $\mathbf{h}$ is nondegenerate on $M_{x_0}/\mathbf{F}_{x_0}$.

Example. To give a simple illustration of the possible applications of this machinery, recall the Poincaré-Birkhoff fixed-point theorem (Poincaré's "Last Geometric Theorem").

M is the region between two concentric circles in the plane with Cartesian coordinates (x, y). Let $\omega = dx \wedge dy$, the Euclidean volume element. Let $T : M \to M$ be a transformation such that

(a) T is extendable to the boundary: $C_0 \cup C_1 \to C_0 \cup C_1$, ($C_0$ (resp. C_1) is the outer (resp. inner) circle) and rotates them in opposite senses.

(b) $T^*(\omega) = \omega$, i.e., T is volume-preserving.

The theorem states that T has at least two fixed points.

It remains a challenge to imbed this famous theorem in general fixed-point theorems. We shall now show, however, that the infinitesimal version of the theorem can be proved very simply by the methods we have been using. Suppose then that X is a vector field on $M \cup C_0 \cup C_1$ that is (a) tangent to C_0 and C_1 and points in opposite directions on C_0 and C_1, with no zeros on either boundary, and (b) is infinitesimally volume-preserving, i.e., $X(\omega) = 0$.

We know then that $d(X \lrcorner \omega) = 0$. We do not know a priori that there is an $f \in C(M)$ such that $df = X \lrcorner \omega$, since M is not simply connected. However, this is where hypothesis (b) enters: Since $X \lrcorner \omega(X) = 0$, and X restricted to C_0 and C_1 is, up to a factor, the tangent vector, the line integral of $X \lrcorner \omega$ on the boundaries is zero; hence by de Rham's theorem such an f actually exists. We use (b) again to verify that either grad f or $-$grad f always points inward to the region. Standard Morse theory for manifolds with boundary now guarantees that f has at least two critical points, i.e., X has at least two zeros.

Return now to the general theory. If $Y(\omega) = 0$, then $X \lrcorner \omega$ is a base-like form for the foliation $\mathbf{F}$; hence, if $df_Y = Y \lrcorner \omega$, then f_Y is an invariant of the foliation $\mathbf{F}$, i.e., is constant along the leaves. If f_{Y_1} is so associated with another such vector field Y_1, then the Poisson bracket $\{f_Y, f_{Y_1}\}$ can be defined as in [7] as the $\mathbf{F}$-invariant function such that

2.16. $$\omega^{p-1} \wedge df_Y \wedge df_{Y_1} = \{f_Y, f_{Y_1}\}\omega^p, \qquad \text{where } 2p = \dim M - \dim \mathbf{F}.$$

Note that

2.17. $$d(Y(f_{Y_1})) = Y(df_{Y_1}) = Y(Y_1 \lrcorner \omega) = [Y, Y_1] \lrcorner \omega,$$

i.e., $Y(f_{Y_1})$ can be taken as $f_{[Y,Y_1]}$.

2.18. $$Y(f_{Y_1}) = p\{f_{Y_1}, f_Y\} = f_{[Y,Y_1]}.$$

Proof.

$$\begin{aligned} Y(f_{Y_1})\omega^p &= Y(f_{Y_1}\,\omega^p) = d(f_{Y_1}(Y \mathbin{\lrcorner} \omega^p)) \\ &= d(f_{Y_1}\cdot p\cdot(Y \mathbin{\lrcorner} \omega) \wedge \omega^{p-1}) \\ &= p\, df_{Y_1} \wedge (Y \mathbin{\lrcorner} \omega) \wedge \omega^{p-1} = p\, df_{Y_1} \wedge df_Y \wedge \omega^{p-1}, \qquad \text{Q.E.D.} \end{aligned}$$

(Note that a proof of the Jacobi identity for the Poisson bracket is easily obtained from 2.18.)

Now suppose that θ is a 1-form on M such that $d\theta = \omega$ and that $Y \in V(M)$ satisfies $Y(\theta) = 0$. Then,

$$d(\theta(Y)) = -Y\, d\theta, \qquad \text{i.e., } \theta(Y) \text{ can be taken as } -f_Y .$$

Note then that if Y_1 also satisfies $Y_1(\theta) = 0$, then

$$Y_1(f_Y) = -\theta([Y_1, Y]), \qquad \text{i.e., } Y_1(f_Y) \text{ is } f_{[Y_1, Y]} .$$

3. The Lagrange variational problem

Let B be a manifold, and let $T(B) = \bigcup_{b \in B} B_b$ be its tangent bundle. If (x_i) $(1 \leqq i, j, \cdots \leqq n$; summation convention) is a coordinate system for B, we can construct a coordinate system for $T(B)$, denoted by $(x_i, \dot{x}_i)$, so that (a) the (x_i) considered as functions on $T(B)$ are just the x_i on B lifted up to $T(B)$, and (b) $\dot{x}_i(v) = dx_i(v)$ for $v \in T(B)$. Let $X = a_i\, \partial/\partial x_i$ be a vector field on B. Define its first-order prolongation X', a vector field on $T(B)$, as

$$a_i \frac{\partial}{\partial x_i} + \frac{\partial a_i}{\partial x_j} \dot{x}_j \frac{\partial}{\partial \dot{x}_i} .$$

It is easy to verify that X' is independent of the coordinate system used to define it, and hence is globally defined on $T(B)$. This is so because X' has the following geometric interpretation: Consider the one-parameter pseudogroup generated by X. Each element prolongs to a transformation on tangent vectors; hence there is defined a one-parameter pseudogroup acting on $T(B)$. X' is just its infinitesimal generator.

Let $M = T(B) \times R$. Let t be the real function on M obtained by projecting on the second factor. Consider vector fields Z on $B \times R$ of the form:

3.1. $$Z = X^t + \partial/\partial t ,$$

where, for each t, X^t is a vector field on B.

These fields have the property that the pseudogroups they generate preserve the subclass of curves of $B \times R$ that are graphs of curves in B. We define the *first-order prolongation* of Z as the field $Z' = X^{t\prime} + \partial/\partial t$ on M.

A *Lagrangian* for B is a real-valued function L on M. It defines a real-valued function $\mathbf{L}$ on the space of curves of B as follows: If $\sigma : [a, b] \to B$ is a curve, then

$$\mathbf{L}(\sigma) = \int_a^b L(\sigma'(t), t)\, dt.$$

($\sigma'(t) \in B_{\sigma(t)}$ denotes the tangent vector to σ at t.) Note that a 1-differential-form on B defines a Lagrangian. If σ is a curve in B, define its *prolongation* to M as the curve $\sigma_* : t \to (\sigma'(t), t) \in M$.

LEMMA 3.1 (Cartan [7]). *There is a 1-form $\theta(L)$ on M such that $\theta(\mathbf{L})(\sigma_*) = \mathbf{L}(\sigma)$ for each curve σ in B. A curve σ in B is an extremal for L, i.e., satisfies the Euler equations, if and only if σ_* is a characteristic curve for $d\theta(L)$. In a coordinate system (x_i) for B, $\theta(L)$ has the form*

3.2. $$\theta(L) = \frac{\partial L}{\partial \dot{x}_i}\, dx_i - \left(\frac{\partial L}{\partial \dot{x}_i}\, \dot{x}_i - L\right) dt.$$

The following properties follow from 3.2:

3.3. $$\theta(fL_1 + gL_2) = f\theta(L_1) + g\theta(L_2)$$

for any two Lagrangians L_1 and L_2, any two functions on $B \times R$, f and g;

3.4 $$Z'(\theta(L)) = \theta(Z'(L))$$

for a vector field Z on $B \times R$ of the form 3.1.

Cartan actually defined $\theta(L)$ via 3.2. The proof that it is the same no matter what coordinate system is used, hence is globally defined on M, is straightforward and is left to the reader. 3.3 and 3.4 are similarly proved by a direct computation.

It should be noticed, following Cartan [7] again, that Lemma 3.1 provides a way of developing the basic facts of the ordinary problem in the calculus of variations in an elegant, completely global manner. For example, a function W on $B \times R$ is a solution of the *Hamilton-Jacobi* partial differential equation if there is a cross-section $\phi : B \times R \to M$ such that

3.5. $$dW = \phi^*(\theta(L)).$$

Following Carathéodory's ideas, such solutions of the Hamilton-Jacobi equation can be used, together with a suitable "Legendre condition", to prove that the extremals locally minimize $\mathbf{L}$. On the other hand, if one wants to write the extremals of L as solutions of the Hamilton ordinary differential equations 2.7, proceed as follows to define the Hamiltonian function H: It is any function of $2n + 1$ real variables $H(u_i, v_i, t)$ such that

3.6. $$H\left(x_i, \frac{\partial L}{\partial \dot{x}_i}, t\right) = \frac{\partial L}{\partial \dot{x}_j}\, \dot{x}_j - L.$$

As another application of the formalism, suppose that Z is a vector field on $B \times R$ satisfying 3.1 and such that $Z'(L) = 0$, i.e., the prolonged pseudo-group preserves L. By 3.4, $Z'(\theta(L)) = 0$, and the work of Section 2 can be applied: Define $f_Z = \theta(L)(Z')$. Let $(b_0, t_0) \in B \times R$. We see that the integral curve of Z starting at (b_0, t_0) is an extremal for L if and only if $(X^{t_0}(b_0), t_0)$ is a critical point for f_Z. Using 3.2 and a short calculation, we see that

3.7. *The integral curve of Z starting at (b_0, t_0) is an extremal for L if and only if (b_0, t_0) is a critical point for the function $(b, t) \to L(X^t(b))$ on $B \times R$.*

We shall not give the details of the calculation since we shall prove a more general result (Theorem 3.3) below. Note further that we have dealt with a special case of this result (i.e., when L defines a Riemannian metric for B) in [14], which in turn was a generalization of a theorem due to H. Busemann [3].

We now turn to showing how this formalism can be extended to treat Lagrange problems. In addition to the Lagrangian L we are then given a submanifold $S \subset M$, the *constraint manifold.* Let $P(B, S)$ be the set of curves $\sigma : [a, b] \to B$ such that $(\sigma'(t), t) \in S$ for $a \leqq t \leqq b$. The *extremals* of the Lagrange problem are the critical points of $\mathbf{L}$ restricted to $P(B, S)$. We describe the extremals via the classical Lagrange multiplier device.[2] (The reader should keep in mind that it is a nontrivial problem to identify the extremals defined by means of Lagrange multipliers with those defined as the critical points of $\mathbf{L}$.) Suppose that $(x_i, \dot{x}_i, t)$ are the usual local coordinates for M, and $\phi_a(x, \dot{x}, t)$ are functions such that $\phi_a = 0$ defines S. $(1 \leqq i, j, \cdots \leqq n; 1 \leqq a, b, \cdots \leqq m = \dim M - \dim S$; summation conventions.) Following the classical idea, we are to introduce new auxillary functions of t, $\lambda_a(t)$, form the new Lagrangian $L'(x, \dot{x}, t) = L + \lambda_a(t)\phi_a(x, \dot{x}, t)$, and find the ordinary extremals of L' that satisfy the conditions of constraint. It is easy to see that this is equivalent to the following procedure: Introduce (λ_a) as coordinates of a new space R^m. Form $B \times R \times R^m$ and on it a new Lagrangian

$$L'(x, \dot{x}, \lambda, \dot{\lambda}, t) = L(x, \dot{x}, t) + \lambda_a \phi_a(x, \dot{x}, t).$$

Since L' does not depend on $\dot{\lambda}_a$, the associated 1-form $\theta(L')$ is equal to $\theta(L) + \lambda_a \theta(\phi_a)$, which can be considered as a form on $M \times R^m$. The characteristic curves of $d\theta(L')$ that lie on $S \times R^m$, when projected to B, are the extremals of the Lagrange problem. But of course, the characteristic curves of $d\theta(L')$ that lie on $S \times R^m$ are precisely the characteristic curves of $d\phi(L')$ restricted to $S \times R^m$, i.e., the extremals of the Lagrange problem are defined by a symplectic foliation on $S \times R^m$. As for writing the equations of the extremals in Hamiltonian form, the following facts follow from these observations:

A function $H(u_i, v_i, t)$ of $2n + 1$ real variables is a *Hamiltonian* for the Lagrange problem if

3.8.
$$H\left(x_i(p), \frac{\partial L}{\partial \dot{x}_i} + \lambda_a \frac{\partial \phi_a}{\partial \dot{x}_i}(p), t(p)\right) = \left(\frac{\partial L}{\partial \dot{x}_i} + \lambda_a \frac{\partial \phi_a}{\partial \dot{x}_i}\right)\dot{x}_i - (L + \lambda_a \phi_a)(p)$$

[2] From now on, by "extremal" we will always mean the curves satisfying the Lagrange multiplier rule. These curves were described as the "formal" extremals in the introduction.

for all $p \in S$, all $(\lambda_a) \in R^m$. Then, under the mapping defined by

$$u_i = x_i\,, \qquad v_i = \frac{\partial L}{\partial \dot{x}_i} + \lambda_a \frac{\partial \phi_a}{\partial \dot{x}_i}$$

(with the relations $\phi_a = 0$ holding), the extremal curves go over into solutions of the Hamilton equations with Hamiltonian H. Of course, this mapping is not necessarily onto the (u, v, t)-space, i.e., every solution of the Hamilton equations does not necessarily arise as the image of an extremal. However, a reasonable sufficient condition for every solution of the Hamilton equations to correspond to an extremal is that

3.9. *For x_i held constant and ϕ_a held zero, the Jacobian matrix of*

$$\frac{\partial L}{\partial \dot{x}_i} + \lambda_a \frac{\partial \phi_a}{\partial \dot{x}_i}$$

with respect to the other independent variables of $M \times R^m$ is nonzero, i.e., the variational problem is regular in the classical sense.

So far this is local, dependent on a choice of coordinate system and functions ϕ_a whose vanishing defines S. Suppose however, that the ϕ_a are functionally independent, i.e., the $d\phi_a$ are linearly independent. (It would suffice that this be true in a neighborhood of S.) $d\phi_a$ is zero on S_p, for $p \in S$, and hence passes to the quotient and defines a linear form on M_p/S_p. The forms $d\phi_a$ in fact form a basis for the dual space of M_p/S_p. Under change of coordinates in B, L remains invariant, while the $d\phi_a$ change, by the above remark, like the dual of the normal bundle of S. Since it is natural to require that $\theta(L')$ be invariant when restricted to S, the λ_a must change dually to the ϕ_a, i.e., λ_a are to be interpreted as the coordinates of M_p/S_p with respect to the basis of linear forms $d\phi_a$, and $\theta(L')$ is to be interpreted as a form on the normal vector bundle to S, which we denote by M'.

To verify this interpretation is now a straightforward matter. In the portion of M covered by the coordinate system $(x, \dot{x}, t)$, define the diffeomorphism of M' with $S \times R^m$ as described above, carry $\theta(L')$ back to M' via this map, and check by direct calculation (left to the reader) that $\theta(L')$ so defined on M' is independent of the coordinate system we have used. We can now state the result of these remarks:

THEOREM 3.2. *With the above notations, the extremals of the Lagrange problem in B (defined the Lagrange multiplier way) are the projection in B of the integral curves of the characteristic (symplectic) foliation of the 2-form $d\theta(L')$ on M', the normal vector bundle to the constraint manifold S.*

We have so far done nothing more than take the classical arguments, fit them into Cartan's approach to the calculus of variations, and then remark that in this form they admit an immediate global interpretation. We shall now use this point of view to generalize 3.7. Suppose that Z is a vector field on $B \times R$ of the form $X^t + \partial/\partial t$ and that $Z' = X^{t'} + \partial/\partial t$ satisfies

3.10. $$Z'(L) = 0 \quad \textit{and} \quad Z' \textit{ is tangent to } S.$$

The one-parameter pseudogroup generated by Z' then maps S into itself, hence also generates a one-parameter pseudogroup acting on the normal bundle, M', to S, whose infinitesimal generator we also denote by Z'. Suppose, for example, that

3.11. $$Z'(d\phi_a) = A_{ab}\, d\phi_b \quad \text{on } S.$$

To determine explicitly the action of Z' on M', it only remains to determine $Z'(\lambda_a)$, where the λ_a are considered as functions on M'. From 3.11 and our previous remarks about the dual nature of λ_a and $d\phi_a$, it is easy to see that

3.12. $$Z'(\lambda_a) = -A_{ba}\lambda_b .$$

From 3.10, 3.11, and 3.12, we have

3.13. $$Z'(\theta(L')) = 0 \quad \text{on } M'.$$

The considerations of Section 2 now apply, and we have

3.14. *The integral curve of Z' starting at $p \in M'$ is a characteristic curve of $d\theta(L')$ restricted to M' if and only if p is a critical point of $\theta(L')(Z')$ restricted to M'.*

Suppose, for example, that $X^t = a_i^t\, \partial/\partial x_i$. Then,

3.15. $$\theta(L')(Z') = \left(\frac{\partial L}{\partial \dot{x}_i} + \lambda_a \frac{\partial \phi_a}{\partial \dot{x}_i}\right)(a_i^t - \dot{x}_i) + L + (\lambda_a + u_a)\phi_a ,$$

and we find its critical points that satisfy the constraints $\phi_a = 0$. Notice however that we do not want to find all critical points, just those corresponding to orbits of Z that are extremals, i.e., those for which $a_i^t = \dot{x}_i$. Suppose $(x^0, \dot{x}^0, t_0, \lambda^0, u^0)$ is such a critical point, with $\dot{x}_i^0 = a_i^{t_0}(x^0)$. Differentiating 3.15 first with respect to $\dot{x}_i$ gives the condition:

3.16. $$u_a \frac{\partial \phi_a}{\partial \dot{x}_i} = 0 \qquad \text{at} \quad (x^0, a^{t_0}(x^0), t_0, \lambda_a^0, u_a^0).$$

Applying $\partial/\partial x_i$ and $\partial/\partial t$ and using 3.16 gives

3.17. $$\begin{gathered}\frac{\partial}{\partial x_i}(L(X^t) + (\lambda_a + u_a)(\phi_a(X^t))) = 0, \\ \frac{\partial}{\partial t}(L(X)^t) + (\lambda_a + u_a)(\phi_a(X^t))) = 0 \qquad \text{at} \quad (x^0, a^{t_0}, t_0, \lambda_a^0, u_a^0).\end{gathered}$$

Notice however, that 3.17 is the necessary and sufficient condition that (x^0, t_0) be a critical point of the function $(x, t) \rightarrow L(X^t(x))$ restricted to the subset $A = \{(b, t) \in B \times R : (X^t(b, t)) \in S\}$, provided that

3.18. *The system of Pfaffian forms $d(\phi_a(X^t(b)))$ has a constant dimension on A, and A is a submanifold of $B \times R$.*

If 3.18 is not satisfied, note at least that 3.16 implies that (x^0, t_0) is a critical point of $L(X^t)$ restricted to A. Summing up, we have proved

THEOREM 3.3. *With the above notations, if the integral curve of Z starting at $(b_0, t_0) \in A \subset B \times R$ is an extremal of the Lagrange variational problem, then (b_0, t_0) is a critical point of the function $(b, t) \to L(X^t(b), t)$ on A providing that A is a submanifold of $B \times R$. Conversely, if (b_0, t_0) is a critical point and if* 3.18 *is satisfied in a neighborhood of (b_0, t_0), then the integral curve of Z starting there is an extremal.*

In the general case, i.e., if 3.18 is not necessarily satisfied, let r be the dimension of the space of m-tuples (α_a) such that $\alpha_a\, d(\phi_a(X^t(b)))$, considered as a covector at (x^0, t_0), is zero. To identify r with a natural invariant of the Lagrange problem, proceed as follows: Let $\sigma : [a, b] \to B$ be any (formal) extremal of the Lagrange problem, i.e., $\sigma_* : [a, b] \to M$ is a characteristic curve of the 2-form $d(\theta(L) + \lambda_a(t)\theta(\phi_a))$, for *some* curve $\lambda_a(t)$ in R^m. Consider now the set of such curves $\lambda_a(t)$. The difference of two such, say a curve $\alpha_a(t)$ in R^m, satisfies a linear, homogeneous system of ordinary differential equations, namely:

3.19. $$\sigma'_* \lrcorner \left\{\frac{d}{dt}\alpha_a(t)\, dt \wedge \theta(\phi_a) + \alpha_a(t)\, d\theta(\phi_a)\right\} = 0.$$

The set of all solutions $\alpha_a(t)$ of 3.19 forms a vector space, which, under the classical "regularity conditions" for the constraints ϕ_a, is finite-dimensional. Its dimension is the *class* of the extremal σ. (This definition of the class is slightly different from that given by Carathéodory [5], but seems to be geometrically more natural, since it does not depend on the choice of a Hamiltonian function for the variational problem.)

Now suppose that $Z = X^t + \partial/\partial t$ is a vector field on $B \times R$ whose integral curve σ starting at $(x^0, t_0) \in B \times R$ is an extremal and such that Z' leaves invariant the Lagrangian and the constraints. It should be clear that the number r defined above is precisely the class of the extremal σ.

4. A Lagrange problem in Riemannian geometry

We suppose now that L defines a Riemannian metric on B; in local coordinates, $L = (g_{ij}\,\dot{x}_i\,\dot{x}_j)^{1/2}$. Suppose further that the constraints are linear, i.e., there is a field $b \to \mathbf{H}_b \subset B_b$ of tangent subspaces (of constant dimension) on B. $P(B, S)$ consists of those curves $\sigma : [a, b] \to B$ with $\sigma'(t) \in \mathbf{H}_{\sigma(t)}$, $a \leqq t \leqq b$, i.e., $P(B, S)$ consists of the integral curves of the Pfaffian system determined by $\mathbf{H}$. In local coordinates, there are 1-forms $w_a = A_{ai}\, dx_i$, $(1 \leqq a, b, \cdots \leqq m = \dim B - \dim \mathbf{H};\ 1 \leqq i, j, \cdots \leqq n;\ m + 1 \leqq u, v, \cdots \leqq n$; summation conventions) such that $w_a(\sigma'(t)) = 0$, $a \leqq t \leqq b$.

Introduce the Lagrange multipliers λ_a and the Lagrangian $L' = L + \lambda_a A_{ai} \dot{x}_i$. For $b \in B$, let $V_b = \mathbf{H}_b^\perp$, the orthogonal complement of $\mathbf{H}_b$ with respect to the positive definite quadratic form defined in B_b by L^2. (Think of vectors in $\mathbf{H}_b$ (resp. V_b) as *horizontal* (resp. *vertical*).) Since $w_a = 0$ on $\mathbf{H}_b$, the w_a restricted to V_b form a basis for the dual space of V_b. In accordance with the remark that the λ_a transform dually to the w_a, we see that (λ_a) are to be interpreted as the coordinates with respect to the w_a of a vector on V_b, i.e., the λ_a are to be interpreted as functions on the bundle of vertical vectors.

Suppose for the moment that (w_a) is part of a globally defined orthonormal moving frame (w_i) of 1-forms, i.e., $L^2(v) = w_i(v)w_i(v)$ for $v \in T(B)$. Let y_i be the functions on $T(B)$ such that $y_i(v) = w_i(v)$ for $v \in T(B)$.

(We simply relabel w_i by y_i, since we shall want to consider dy_i, a 1-form on $T(B)$, and want to avoid confusion with dw_i, a 2-form on B. Further, we want to consider the 1-forms $p^*(w_i)$, where $p : T(B) \to B$ is the projection map. Introducing the y_i allows us to simplify notation and denote these forms also by w_i.)

Then, (dy_i, w_i) forms a basis for the 1-forms of $T(B)$. L', considered as a function on $T(B) \times R^m$, takes the form $(y_i y_i)^{1/2} + \lambda_a y_a$. A short computation, left to the reader, shows that

4.1. $$\theta(L') = y_i w_i/(y_j y_j)^{1/2} + \lambda_a w_a .$$

(Since in this problem there is no explicit time dependence, we consider this as a form on $T(B) \times R^m$. The problem is then to find its characteristic curves lying in the constraint manifold.)

S, the constraint manifold, is defined by $y_a = 0$; hence

4.2. $\theta(L')$ *restricted to* $S \times R^m$ $= y_u w_u/(y_v y_v)^{1/2} + \lambda_a w_a$.

Introduce the components of the Levi-Civita connection with respect to the frame we are working with, i.e., the forms w_{ij} with $dw_i = w_{ij} \wedge w_j$ and $w_{ij} + w_{ji} = 0$. Let w_{ij} denote the same forms pulled up to $T(B)$.

Suppose that $\sigma : [0, a] \to B$ is an extremal curve for the Lagrange problem, i.e., $\sigma' : [0, a] \to T(B)$, the tangent vector curve of σ, is a characteristic curve of $d\theta(L')$. We can suppose that σ is parameterized by arc-length, i.e., σ' lies in the submanifold of S defined by $y_u y_u = 1$. So restricted, $d\theta(L')$ takes the form

$$dy_u \wedge w_u + y_u w_{ui} \wedge w_i + d\lambda_a \wedge w_a + \lambda_a w_{ai} \wedge w_i .$$

Set $y_u(t) = w_u(\sigma'(t)) = y_u(\sigma'(t))$. The condition that $(\sigma'(t), \lambda_a(t))$ be a characteristic curve is then, after a short computation,

4.3. $$\frac{d}{dt} y_u(t) w_u + y_u(t) w_{ui}(\sigma'(t)) w_i + \frac{d}{dt} \lambda_a(t) w_a + \lambda_a w_{ai}(\sigma'(t)) w_i - \lambda_a(t) w_{av} y_v = 0.$$

To put these conditions in an invariant form, recall the notion of the covariant derivative ∇v of a vector field v along a curve

$$w_i(\nabla v(t)) = \frac{d}{dt} w_i(v(t)) - w_{ij}(\sigma'(t))w_j(v(t)).$$

Let $\lambda : r \rightarrow \lambda(t) \in B_{\sigma(t)}$ be the vertical vector field along σ such that $w_a(\lambda(t)) = \lambda_a(t)$ and $w_u(\lambda(t)) = 0$. The first four terms of 4.3 then take the form

$$w_i(\nabla\sigma'(t) + \nabla\lambda(t))w_i .$$

The fifth term in 4.3 depends bilinearly on $\lambda(t)$ and $\sigma'(t)$; we write it as $-w_i(Q(\lambda(t), \sigma'(t)))w_i$, where $(h, v) \rightarrow Q(v, h) \in B_b$ is a tensor field on B defined for $h \in H_b$, $v \in V_b$. As definition,

4.4. $$w_i(Q(h, v))w_i = w_a(v)w_{au}\, w_u(h).$$

With this definition, the condition for an extremal takes the form

4.5. $$\nabla\sigma'(t) + \nabla\lambda(t) = Q(\sigma'(t), \lambda(t)) \qquad \text{for } \theta \leqq t \leqq a.$$

Since the condition expressed by 4.5 is independent of choice of moving frame, and the left-hand side of 4.5 is independent of this choice, the right-hand side, i.e., 4.4, must be also. Q is then a bona fide tensor field on all of B, even when B does not admit a global moving frame. This can also be verified directly, as follows: If (w_i') is another orthonormal moving frame, with $w_i = M_{ij}\, w_j'$, and with (w_{ij}') as connection forms, a short, well-known computation shows that

4.6. $$w_{ij} = dM_{ik}\, M_{jk} + M_{ik}\, w_{kl}'\, M_{jl} .$$

If $w_a' = 0$ also defines $\mathbf{H}$, then $M_{au} = 0 = M_{ua}$. Hence, from 4.6

4.7. $$w_{au} = M_{ab}\, w_{bv}'\, M_{uv} ,$$

i.e., the w_{au} transform tensorially. Summing up, we have proved

THEOREM 4.1. *With the above notations, the horizontal curves σ that occur as solutions of* 4.5, *parameterized by arc-length, are precisely the extremals of the Lagrange variational problem defined on integral curves of the Pfaffian system* $\mathbf{H}$ *by the condition that arc-length be stationary. If the metric on M is complete, then, given $b \in B$, $h \in \mathbf{H}_b$, $\lambda_0 \in V_b$, there are a unique integral curve σ of $\mathbf{H} : [0, \infty) \rightarrow B$, parameterized by arc-length, and a vertical vector field $\lambda(t)$, $0 \leqq t < \infty$, along σ such that both satisfy* 4.5, *with $\sigma'(0) = h$, $\lambda(0) = \lambda_0$. (If $\mathbf{H}$ is not complete, σ and λ both exist and are unique, but over a domain possibly smaller than* $[0, \infty)$, *depending on h, b, and λ_0.)*

Proof. Since we have seen that the extremals, i.e., solutions of 4.5, are integral curves of a vector field on the direct product on R^m and the sphere bundle of unit horizontal tangent vectors of B, all should be clear except perhaps the assertion that σ can be extended over $[0, \infty)$ in case B is complete. To prove this, of course we try to apply the standard analytic continuation arguments. Suppose then that σ and λ are defined over $[0, a)$, with $a < \infty$.

Looking at 4.3, we see that the differential equations for $\lambda(r)$ are linear, nonhomogeneous with continuous coefficients that have a limit as $t \to a$. It is then a standard fact in the theory of ordinary differential equations that $\lim_{t\to a} \lambda(t)$ exists, since $\lim_{t\to a} \sigma(t)$ exists by completeness, Q.E.D.

Suppose that $\mathbf{H} \subset V(B)$ continues to define a Pfaffian system $b \to \mathbf{H}_b$ on the Riemannian manifold B. We may ask: When does $\mathbf{H}$ have the property that an ordinary geodesic of B whose tangent vector at one point satisfies the constraints, i.e., lies in $\mathbf{H}$, must always lie in $\mathbf{H}$?

Let us look at this in greater generality, in terms of the notation of Section 3. Since there is no essential loss in generality, we shall only deal with homogeneous Lagrangians and constraints. Suppose then that L is a Lagrangian on B, i.e., a function on $T(B)$, and that ϕ_a, $1 \leqq a, b, \cdots \leqq m$, are constraint functions on $T(B)$. Let S be the subset $\phi_a = 0$ of $T(B)$. Let $\theta(L)$ be the 1-form on $T(B)$ defined by Lemma 3.1. Suppose $Z \in V(T(B))$ is a characteristic vector field of $d\theta(L)$; i.e., $Z \lrcorner\, d\theta(L) = 0$. Then

4.8. *If $Z(\phi_a)$ lies in the ideal of functions on $T(B)$ generated by the ϕ_a, the integral curves of Z starting at S lie in S. In particular, if this is true for each Z satisfying $Z \lrcorner\, d\theta(L) = 0$, then the extremals $\sigma : [a, b] \to B$ of L in B have the following property: If $\sigma'(t_0) \in S$ for some $t_0 \in [a, b]$, then $\sigma'(t) \in S$ for $a \leqq t \leqq b$.*

For each Lagrange variational problem, the question arises of interpreting the conditions involved in the hypothesis of 4.8. We proceed to deal with this question in the case mentioned above, and then go on to present a generalization in Section 5 to the case where $\mathbf{H}$ defines a Pfaffian system with singularities on B. Then, $L = ds$ defines a Riemannian metric on B. Suppose (ω_i) is an orthonormal basis for 1-forms on B, with $\mathbf{H}$ defined by $\omega_a = 0$.

$$(1 \leqq i, j, \cdots \leqq n; 1 \leqq a, b, \cdots \leqq m; m + 1 \leqq u, v, \cdots \leqq n.)$$

As before, no notational distinction is made between a form on B and its image in $T(B)$, and (y_i) are the functions on $T\ (B)$ such that $y_i(v) = \omega_i(v)$. Then,

$$\theta(L) = y_i\, \omega_i/(y_j\, y_j)^{1/2}.$$

S is defined by $y_a = 0$. Further, we can obviously restrict everything to $S(B)$, the unit sphere handle to B. Then $\theta(L) = y_i\, \omega_i$, $y_i\, dy_i = 0$, there is a unique $Z \in V(S(B))$ satisfying

4.9. (a) $\omega_i(Z) = y_i$, (b) $Z(y_j) + y_i\, \omega_{ij}(X)\omega_j = 0$

(ω_{ij} are the connection forms; i.e., $d\omega_i = \omega_{ij} \wedge \omega_j$, $\omega_{ij} + \omega_{ji} = 0$), and Z then satisfies $Z \lrcorner\, d\theta(L) = 0$, and is unique subject to 4.9(a). (Geometrically then, the integral curves of Z when projected down to B are in arc-length parameterization.) Suppose that

$$\omega_{ij} = \Gamma_{ijk}\, \omega_k\ .$$

Then, we have proved

PROPOSITION 4.2. *$Z(y_a)$ lies in the ideal generated by the y_a if and only if $\Gamma_{auv} + \Gamma_{avu} = 0$. If this condition is satisfied, a geodesic of B whose tangent vector lies in $\mathbf{H}$ at one point is an integral curve of $\mathbf{H}$.*

It is seen by using 4.7 that in any case the $\Gamma_{auv} + \Gamma_{avu}$ transform tensorially under change of frame. In the case the vector field $b \to V_b = \mathbf{H}_b^\perp$ is completely integrable, i.e., defines a foliation, this tensor field is zero if the metric is bundle-like in the sense of Reinhart [17] with respect to the foliation V. (This gives another proof of Reinhart's theorem that a geodesic perpendicular to one leaf of a bundle-like foliation is perpendicular to all.) Conversely, it is possible to show, by using Theorem 3.2 of [12], that the vanishing of this tensor field implies that the metric is bundle-like. (Notice that this proves that if the geodesics have this property, the metric is bundle-like.)

5. Application to the study of the geometry of the tangent bundle of a Riemannian manifold

Let B be a Riemannian manifold, and let $M = T(B)$ be its tangent bundle. Choose the range of indices and summation convention:

$$1 \leqq i, j, k, \cdots \leqq n = \dim B.$$

The fibre bundle $T(B) \to B$ has a connection, i.e., a field $v \to \mathbf{H}_v$ of tangent subspaces complementary to the fibres of the projection $T(B) \to B$. $\mathbf{H}$ is defined by the standard Levi-Civita affine connection on B. We shall recall the definition below. Let X be an infinitesimal isometry vector field on B, a Killing field, and let X' be its first-order prolongation to a vector field of $T(B)$. X' preserves $\mathbf{H}$; hence for a $v \in T(B)$ such that $X'(v) \in \mathbf{H}_v$ we have: The integral curve of X' starting at v is an integral curve of $\mathbf{H}$, i.e., is horizontal. If $v \in B_b$, for $b \in B$, and if $\sigma : [0, a] \to B$ is the integral curve of X starting at b, this means that the vector field $t \to (\operatorname{Exp} tX)_*(v)$ on σ is self-parallel. Such $v \in T(B)$ have an evident geometric importance.

Further, $T(B)$ has a natural Riemannian metric (to be described below); hence we can apply the machinery described above to find the orbits of X' that are either geodesics of $T(B)$ or extremals of the Lagrange variational problem defined by the metric and by the horizontal field $\mathbf{H}$. Our aim in this section is to carry out the details of these calculations and try to find the natural geometric interpretations of the results.

Since our work will be basically local, we can suppose that B has a global basis (w_i) of 1-forms that defines an orthonormal moving frame for the Riemannian metric, i.e.,

$$ds^2 = w_i \cdot w_i \qquad (\cdot = \text{symmetric product}).$$

Let (w_{ij}) be the 1-form components of the Levi-Civita connection, i.e.,

$$dw_i = w_{ij} \wedge w_j , \qquad w_{ij} + w_{ji} = 0.$$

As before, we introduce functions y_i on $T(B)$ such that

$$y_i(v) = w_i(v) \quad \text{for } v \in T(B).$$

This artifice enables us to make no notational distinction between a differential form on B and the same form pulled up to $T(B)$ via the projection map. Define

$$\theta_i = dy_i - w_{ij}\, y_j .$$

It is easy to see that the field **H** is defined by setting $\theta_i = 0$. Suppose X is a Killing field on B, i.e.,

$$X(w_{ij}) = a_{ij}\, w_j , \qquad \text{with } a_{ij} + a_{ji} = 0.$$

A standard calculation gives

$$X(w_{ij}) = da_{ij} + a_{ik}\, w_{kj} - w_{ik}\, a_{kj} .$$

By definition,

$$X'(w_i) = X(w_i) \quad \text{and} \quad X'(y_i) = a_{ij}\, y_j .$$

It is easy to calculate that

$$X'(\theta_i) = a_{ij}\, \theta_j .$$

The metric on $T(B)$ is

$$ds'^2 = w_i \cdot w_i + \theta_i \cdot \theta_i .$$

It is clear from these facts that X' is a Killing vector field for this metric and that the metric is bundle-like in the sense of Reinhart [17] with respect to the foliation of $T(B)$ defined by the fibres of the projection map $T(B) \rightarrow B$. It is also readily verified that the metric on $T(B)$ does not depend on the moving frame (w_i) used to define it, and hence is defined even if B cannot be covered by a global moving frame. Note that

5.1. $$\theta_i(X') = (a_{ij} - w_{ij}(X))y_j .$$

To calculate the right-hand side of 5.1, use

$$\begin{aligned} a_{ij}\, w_j &= X \lrcorner\, (w_{ij} \wedge w_j) + dX_i , \quad \text{where } X_i = w_i(X), \\ &= w_{ij}(X)w_j - w_{ij}\, X_j + dX_i \\ &= w_{ij}(X)w_j + X_{i;\,j}\, w_j , \end{aligned}$$

with $X_{i;\,j}$ the components with respect to the moving frame of the classical covariant derivative of the tensor field X.

In general, if T is a tensor field on B with components $T_{i_1 \cdots i_p}$, the components of the covariant derivative are denoted by $T_{i_1 \cdots i_p;\, j}$, i.e., for $v \in B_b$, $\nabla_v(T)$ is a tensor at b of the same algebraic type as T, with components $T_{i_1 \cdots i_p;\, j}\, w_j(v)$. (Since we are using orthonormal moving frames, it is unnecessary to distinguish between contravariant and covariant components.) For example,

$$T_{i;j}\,w_j = dT_i - w_{ij}\,T_j\,,$$

$$T_{i_1,i_2;j}\,w_j = dT_{i_1,i_2} - w_{i_1 j}\,T_{j i_2} + w_{i_2 j}\,T_{i_1,j}\,, \quad \text{etc.}$$

Note the following classical relation:

5.2. $T_{i;j;k} = T_{i;k;j} + \frac{1}{2} R_{ihjk}\,T_h$, *where* (R_{ijkh}) *are the components of the Riemann curvature tensor.*

Recall that, for $v, v_1, v_2 \in B_b$,

5.3. $$(w_j(R(v_1, v_2)(v))) = w_i(v)R_{ijkh}\,w_k(v_1)w_h(v_2) \quad \text{and}$$

$$dw_{ij} = w_{ik} \wedge w_{kj} + R_{ijkh}\,w_k \wedge w_h\,, \qquad R_{ijkh} + R_{ihjk} + R_{ikhj} = 0.$$

Note now that we can rewrite 5.1 as

5.4. $$\theta_i(X') = X_{i;j}\,y_j\,.$$

From this, we see immediately that

5.5. *The integral curve of X' at $v \in T(B)$ is horizontal if and only if $\nabla_v X = 0$.*

To put this result in more geometric language, suppose $v \in B_b$ and suppose $t \to \operatorname{Exp}(tX) : B \to B$ denotes the one-parameter group of isometries of B generated by X. Let σ be the curve in B, and let $t \to \sigma(t) = \operatorname{Exp}(tX)_*(v)$ be the vector field on σ resulting from translating v by the one-parameter group $\operatorname{Exp}(tX)$. 5.5, then, says that $v(t)$ is a self-parallel vector field along σ if and only if $\nabla_v X = 0$.

The necessary and sufficient condition that X be a Killing vector field is

5.6. $$X_{i;j} + X_{j;i} = 0.$$

Applying 5.3, and 5.6 several times, we have:

5.7. $$\begin{aligned} X_{i;j;k} &= X_h(R_{ihjk} - R_{khij} + R_{jhki}) \\ &= X_h(-R_{hjik} - R_{khij} + R_{jhki}) \\ &= X_h(R_{hkij} + R_{hjki} - R_{khij} + R_{jhki}) \qquad \text{by using 5.3} \\ &= 2X_h\,R_{hkij}\,. \end{aligned}$$

Let (X', X') be the square of the length of X' in the ds'^2-metric. Using 5.4, 5.6, and 5.7, we calculate

5.8. $$\begin{aligned} d(X', X') &= 2w_h(X_{i;h}\,X_i + X_{i;j;h}\,y_j\,X_{i;k}\,y_k) \\ &= 2w_h(-X_{h;i}\,X_i + 2X_l\,X_{i;k}\,y_j\,y_k\,R_{lhij}) \\ &\qquad + 2X_{i;j}\,y_j\,X_{i;k}\,dy_k\,. \end{aligned}$$

5.9. $$\begin{aligned} d(\theta_i(X')) &= X_{i;k}\,dy_k + X_{i;j;k}\,y_j\,w_k \\ &= X_{i;k}\,dy_k + 2X_h\,y_j\,w_k\,R_{hkij}\,. \end{aligned}$$

PROPOSITION 5.1. *With the above notations, the integral curve of X' at $v \in B_b$ is a geodesic of $T(B)$ if and only if the integral curve of X at b is a geodesic of B and $\nabla_v X = 0$, i.e., $X'(v)$ is horizontal.*

Proof. Suppose that the right-hand side of 5.8 is zero at v. In particular, $X_{i;j}(b)y_j(v)X_{i;k}(b) = 0$. The skew-symmetry of $X_{i;j}$ forces $X_{i;j}(b)y_j(v)$ to be zero. Applying this again in the right-hand side of 5.8 forces $X_{h;i}(b)X_i(b)$ to be zero. This means that the integral curve of X at b is a geodesic of B [14]. The converse is immediate.

Then, there are essentially no new geodesic integral curves of X'. However, the situation is considerably more interesting if we ask for the integral curves of X' that are extremals of the Lagrange problem defined by the Lagrangian ds'^2 and the constraints $\theta_i = 0$. According to previous results, to find these extremals, form the new function $(X', X')^{1/2} + \lambda_i \theta_i(X')$, with constants λ_i, and find all critical points such that 5.4 is zero. The orbits of these points will be extremals. By using 5.4, 5.8, and 5.9, these conditions on $v \in B_b$ become

5.10. (a) $\lambda_i X_{i;k}(b) = 0$,

(b) $\lambda_i X_h(b) R_{hkij}(b) y_j(v) = X_{k;i}(b)X_i(b)/(X(b), X(b))^{1/2}$.

We now put these conditions in invariant form. Define a vector $\lambda \in B_b$ by

$$\omega_i(\lambda) = \lambda_i \,|\, X(b) \,|.$$

Then 5.10(a) means $\nabla_\lambda X = 0$. 5.10(b) means

$$R(\lambda, v)(X(b)) = \nabla_{X(b)} X.$$

These two formulas prove formulas 1.3.

Finally, we turn to the following question: If $\sigma : [0, a] \to B$ is a curve parameterized by arc-length and $v : [0, a] \to T(B)$ is a self-parallel vector field, on σ, (i.e., v, considered as a curve in $T(B)$, is an integral curve of the Pfaffian system $\theta_i = 0$), what are the conditions that v be an extremal of the Lagrange problem defined by ds' and constraints $\theta_i = 0$? In principle, 4.5 is the condition, but we want to cast it solely in terms of the geometry of B. To do this, it is more convenient to return to the definition of the extremals as characteristic curves of the form defined by 4.2. This form, however, is to be considered as a form on the submanifold of $T(T(B)) \times R^m$ consisting of the $(u, \lambda) \in T(T(B)) \times R^m$ satisfying: $\theta_i(y) = 0$. The details of the computation are similar to those by which we reached 4.5. Note that

$$dw_i = w_{ij} \wedge w_j, \qquad d\theta_i = w_{ij} \wedge \theta_j - \tfrac{1}{2} R_{ijkl}\, y_j\, w_k \wedge w_l\,.$$

If y_i are the functions on $T(T(B))$ such that $y_i(u) = w_i(u)$, for $u \in T(T(B))$, then 4.2 takes the form:

$$y_i\, w_i/(y_i\, y_i)^{1/2} + \lambda_i\, \theta_i\,.$$

Taking the exterior derivative of this form, and expressing the fact that (a) $t \to v(t)$, considered as a curve in $T(B)$ is a characteristic curve, and (b) $(\sigma'(t), \sigma'(t)) = 1$, we have (details left to reader)

5.11. (a) $\nabla\lambda(t) = \sigma'(t)$,

(b) $(R(\sigma'(t), u(t))(\lambda(t)), v(t)) = 0$ *for all* $u(t) \in B_{\sigma(t)}$, *with* λ *a vector field on* σ *such that* $w_i(\lambda) = \lambda_i$.

Restated, $t \to v(t)$ is an extremal of the Lagrange problem if and only if there is a vector field λ along σ satisfying 5.11 (a) such that $v(t)$ satisfies 5.11 (b). Alternatively, one can regard 5.11 as a system of ordinary differential equations determining both λ and v.

Suppose conversely we ask for all λ satisfying 5.11 for a given v. If λ and λ_1 are solutions, with $\alpha = \lambda - \lambda_1$, we have

5.12. (a) $$\nabla\alpha = 0,$$

(b) $$0 = (R(\sigma'(t), u(t))(\alpha(t)), v(t)) \qquad \text{for all } u(t) \in B_{\sigma(t)}.$$

The solutions of 5.12 (a) and (b) form a finite-dimensional vector space. Its dimension is called the *class* of the extremal v, as defined in Section 3.

Suppose for example that B is a symmetric Riemannian space, i.e., the covariant derivative of the curvature tensor $R(\ ,\ )(\ )$ is zero, and that σ is a geodesic of B. It then follows that the class of v is the dimension of B minus the dimension of subspace $R(\sigma'(0), B_{\sigma(0)})(v(0))$ of $B_{\sigma(0)}$. (In the nonsymmetric cases, the covariant derivatives of R come in and enormously complicate the calculations.)

To put this into more familiar algebraic terms, suppose that $B = G/K$, where G is a connected semisimple Lie group, K is a compact symmetric subgroup [15]. Let $\mathbf{G}$ (resp. $\mathbf{K}$) be the Lie algebra of G (resp. K), and let $\mathbf{G} = \mathbf{K} + \mathbf{M}$ be the reduction of K in G, i.e., $[\mathbf{K}, \mathbf{M}] \subset \mathbf{M}$, $[\mathbf{M}, \mathbf{M}] \subset \mathbf{K}$. Suppose that we identify $\mathbf{M}$ with $B_{\sigma(0)}$ in the usual way [15], and that X (resp. Y) $\in \mathbf{M}$ corresponds to $\sigma'(0)$ (resp. $v(0)$) under this identification. Using the Cartan-Nomizu formula for the curvature tensor for a symmetric space [15], we see that $R(\sigma'(0), B_{\sigma(0)})(v(0))$ can be identified with $\operatorname{Ad} Y \operatorname{Ad} X(\mathbf{M})$, i.e., we have

PROPOSITION 5.2. *With the above notations, the class of the self-parallel vector field v is equal to the dimension of the kernel of the endomorphism* $\operatorname{Ad} Y \operatorname{Ad} X$ *acting on* $\mathbf{M}$. *In particular the class is never less than the rank of the symmetric space* G/K *(i.e., the dimension of a maximal abelian (Cartan) subalgebra of* $\mathbf{M}$*).*

We note further that

5.13. *The class of* σ *is equal to the rank of the symmetric space if* X *and* Y *are regular elements of the same Cartan subalgebra of* $\mathbf{M}$.

5.14. *The rank of the symmetric space is equal to the codimension of the orbit of greatest dimension of the holonomy group of* B *acting on* $B_{\sigma(0)}$.

Now, we turn to further work on the problem begun in Section 4. If $\mathbf{H}$ is a real subspace of $V(B)$ (possibly with singularities; i.e., dim $\mathbf{H}_x$ may vary for $x \in B$), when must a geodesic perpendicular to $\mathbf{H}$ at one point be perpendicular at all?

If Y is a vector field on B, let g_Y be the function on $S(B)$, the unit sphere bundle to B, such that

$$g_Y(v) = (Y(x), v) \qquad \text{for } v \in B_x, \quad x \in B.$$

If (ω_i) is an orthonormal basis of 1-forms on B, then

$$g_Y = \omega_i(Y) y_i, \qquad \text{with } y_i(v) = \omega_i(v) \quad \text{for } v \in S(B).$$

Let θ be the 1-form $y_i \omega_i$ on $S(B)$, and let Z be the vector field satisfying 4.9 whose integral curves projected down to B are the geodesics in arc-length parameterization. Suppose $Y_i = \omega_i(X)$, $dY_i = Y_{i;j} \omega_j + \omega_{ij} Y_j$, etc. Then,

$$\begin{aligned} Y(\omega_i) &= (Y_{i;j} + \omega_{ij}(Y))\omega_j, \\ Y'(y_i) &= (Y_{i;j} + \omega_{ij}(Y))y_j, \\ dg_Y &= (Y_{i;j}\omega_j + \omega_{ij} Y_j)y_i + Y_i\, dy_i, \\ Y' \lrcorner\, d\theta &= (Y_{i;j} + \omega_{ij}(Y))y_j \omega_j - Y_i\, dy_i + y_i \omega_{ij}(Y)\omega_j - y_i Y_j \omega_{ij} \\ &= Y_{i;j} y_j \omega_i - Y_i\, dy_i - y_i Y_j \omega_{ij}. \end{aligned}$$

5.15. (a) $$dg_Y + Y' \lrcorner\, d\theta = (Y_{i;j} + Y_j)y_i \omega_j, \quad \text{and hence}$$

(b) $$Z(g_Y) = (Y_{i;j} + Y_{j;j})y_i y_j.$$

Note that then

5.16. *If Y is a Killing vector field, i.e., $Y_{i;j} + Y_{j;i} = 0$, then $Z(g_Y) = 0$; i.e., a geodesic perpendicular to a Killing vector field at one point is perpendicular at all points* [2a].

In the general case we have

PROPOSITION 5.3. *Suppose that H is a linear space of vector fields on the Riemannian manifold B. If each geodesic of B that is perpendicular to $\mathbf{H}$ at one point is everywhere perpendicular, then H satisfies the following condition at each $x \in B$, for $Y \in \mathbf{H}$.*

5.17. *The quadratic form $v \to (\nabla_v, Y, v)$ on $\mathbf{H}_x^\perp$ is identically zero.*

Conversely, this condition is sufficient that $\mathbf{H}$ have this property, provided that the ideal generated by the functions $\{g_Y : Y \in H\}$ on $S(B)$ is prime.

6. Isoparametric problems

B and $M = T(B) \times R$ and a Lagrangian L on M are as before. In addition, we are given Lagrangians L_a, $1 \leqq a \leqq m$. The extremal curves of the isoparametric problem determined by $\{L; L_a\}$ are those curves $\sigma : [a, b] \to B$ for which

$$\mathbf{L}(\sigma) = \int_a^b L(\sigma'(t), t)\, dt$$

has an extremal value, subject to the constraints that the $\mathbf{L}_a(\sigma)$ have a fixed value. The Lagrange multiplier rule suggests introducing new constants (λ_a) (not, however, to be allowed to depend on t as in the Lagrange problem) and introducing a new Lagrangian, $L' = \lambda_a L_a$, with the associated 1-form $\theta(L')$ on M equal to $\theta(L) + \lambda_a\, \theta(L_a)$. The *formal extremals* are those curves σ in B that satisfy the constraints and such that σ_* is a characteristic curve of $d\theta(L')$ for some choice of the λ_a. For a given extremal σ, the set of all m-tuples (λ_a) such that σ_* is a characteristic curve of $d\theta(L')$ forms a subspace of R^m. Its dimension is called the *class* of the extremal σ.

Let $Z = X^t + \partial/\partial t$ be a vector field on $B \times R$, with X^t a one-parameter family of vector fields on B, such that $Z'(L) = Z'(L_a) = 0$, $1 \leqq a \leqq m$. The integral curve of Z at $(b_0, t_0) \in B \times R$ is then, by 3.7, a formal extremal for the isoparametric problem if and only if it is a critical point for the function $(b, t) \to L(X^t(b)) + \lambda_a L_a(X^t(b))$ for some choice of (λ_a). The *class* of this extremal orbit is then

$$m - (\text{dimension of 1-covectors at } (b_0, t_0) \text{ spanned by the } d(L_a(X^t))).$$

We shall restrict our discussion of the isoparametric problem to the classical isoparametric problem in the plane and a class of problems immediately generalizing it to other Riemannian manifolds. Suppose then first that $B = R^2$, the plane with coordinates (x, y), $m = 1$,

$$L = x\dot{y} - y\dot{x}, \qquad L_1 = (\dot{x}^2 + \dot{y}^2)^{1/2},$$
$$X = y\, \partial/\partial x - x\, \partial/\partial y,$$

the infinitesimal generator of the one-parameter group of rotations about the origin. Since there is no time dependence, we are led to finding the critical points of

$$L(X) + \lambda L_1(X) = -x^2 - \dot{y}^2 + \lambda(x^2 + y^2)^{1/2}, \quad \text{i.e.,} \quad \lambda = 2(x^2 + y^2).$$

Hence, for every $b \in B$, a λ which makes b a critical point can be found, and the integral curve of X starting at b, the circle about the origin, is an extremal. Obviously the only such extremal whose class is nonzero is the limiting case of the point circle.

To generalize, suppose that B is a Riemannian manifold with L_1 the Lagrangian determined by the Riemannian metric, and with L the Lagrangian on B determined by a vector field Y on B: $L(v) = (Y(b), v)$ for $v \in B_b$, $b \in B$. If X is a Killing vector field such that $[X, Y] = 0$, the critical points of $(Y, X) + \lambda(X, X)^{1/2}$, for some value of λ, are the origins of the integral curves of X that are extremals. A point $b \in B$ is such a critical point if and only if, for all $v \in B_b$,

$$0 = (\nabla_v Y, X) + (Y, \nabla_v X) + \lambda(\nabla_v X, X)/(X, X)^{1/2}$$
$$= (\nabla_v Y, X) - (\nabla_{\lambda(b)} X, v) - (\lambda/(X, X)^{1/2})(\nabla_{X(b)} X, v)$$

by using conditions 5.6 that X be a Killing field. An integral curve σ of X that is an extremal is then of nonzero class if and only if $\nabla_X X = 0$ along σ, i.e., if and only if σ is also a geodesic of B.

Suppose that Y is also a Killing vector field. Then, the condition becomes

6.1. $$\nabla_X Y + \nabla_Y X + (\lambda/(X, X)^{1/2}) \nabla_X X = 0 \quad \text{at } b.$$

But $\nabla_X Y = \nabla_Y X$ since $[X, Y] = 0$, i.e., 6.1 becomes

6.2. $$2 \nabla_Y X + (\lambda/(X, X)^{1/2}) \nabla_X X = 0 \quad \text{at } b.$$

In particular,

6.3. *If $Y = X$, the integral curve of X at B is always an extremal, since λ can be chosen as $-2(X, X)^{1/2}$.*

7. Remarks on the Lagrange multiplier rule and its relation to the Volterra calculus for path spaces

Our aim in this section is to show how the Lagrange multiplier rule for the Lagrange variational problem can be thought of as an application of Volterra's ideas on the differential geometry of function spaces [18]. First we recall how the Lagrange multiplier rule works for functions on finite-dimensional manifolds.

Let N be a manifold, and let f, f_a, $1 \leqq a \leqq m$, be C^∞ functions on N. Let S be a submanifold of N such that $f_a = 0$ on S. For any choice of constants λ_a, the critical points of $f + \lambda_a f_a$ that lie on S are critical points of f restricted to S. Conversely, if $S_x = \{v \in N_x : df_a(v) = 0\}$ and if $x \in S$ is a critical point of f restricted to S, then there are constants λ_a such that $f + \lambda_a f_a$ has a critical point at x. Note further, since $f_a = 0$ on S, that to look for critical points of $f + \lambda_a f_a$ with λ_a constant is the same as looking for them with λ_a variable, regarded as coordinates of an R^m, since $f_a \, d\lambda_a = 0$ on $S \times R^m$.

If $x \in N$ is a critical point of f subject to the constraints $f_a =$ constant, define the *class* of x as the dimension of the linear space of m-tuples (λ_a) such that $\lambda_a \, df_a = 0$ at x, i.e., the codimension of the space of covectors at x spanned by the df_a. Then the class zero case may be thought of as the *regular* or *nondegenerate* situation.

Now, the formal aspects of the calculus of variations are best understood by regarding the space of C^∞ curves on a manifold as an "infinite-dimensional manifold" [1]. The Lagrange multiplier device can then be regarded as a generalization of the Lagrange multiplier idea for finite-dimensional manifolds recalled above. However, there are often an infinite number of constraint functions f_a to consider, and summation over a must be replaced, at least formally, by an integration.

We present a brief outline of Volterra's ideas as applied to the calculus of variations, adopting of course an intrinsic, coordinate-free point of view. Let B be a manifold, with points b_0, b_1 fixed in B. Let P be the space of C^∞ curves on B going from b_0 to b_1. For convenience, we shall regard an element $\sigma \in P$ as a C^∞ map: $[0, 1] \to B$, with $\sigma(0) = b_0$, $\sigma(1) = b_1$.

For $\sigma \in P$, P_σ, the "tangent space" to P at σ, will consist of the set of C^∞ vector fields $v : t \to v(t) \in B_{\sigma(t)}$, $0 \leqq t \leqq 1$, on σ which vanish at $t = 0$ and $t = 1$. Since vector fields on σ can be added pointwise, P_σ has a linear structure. A deformation σ_s, $0 \leqq s \leqq 1$, of σ (i.e., $\sigma_0 = \sigma$), $\sigma_s(0) = b_0$, $\sigma_s(1) = b_1$, will be thought of as a "curve" in P starting at σ. The "tangent vector" to the "curve" σ_s at $s = 0$ will be a vector field v on σ such that $v(t)$ is the tangent vector to the curve $s \to \sigma_s(t)$ at $s = 0$. In general, if $\delta(s, t)$, $0 \leqq s, t \leqq 1$, is a homotopy in B, $D_s\, \delta(s, t)$ (resp. $D_t\, \delta(s, t)$) will be the tangent vector to the curve $u \to \delta(u, t)$ (resp. $u \to \delta(s, u)$) at $u = s$ (resp. $u = t$). If $\delta(s, t) = \sigma_s(t)$, then v, the "tangent vector" to the curve in P at $s = 0$, will be the vector field $t \to v(t) = D_s\, \delta(0, t)$. If f is a real-valued function on P, the "differential" of f at $\sigma \in P$ will be (if it exists, of course), a linear function $df : P_\sigma \to R$, such that

$$7.1. \qquad \frac{d}{ds} f(\sigma_s)\bigg|_{s=0} = df(v).$$

For example, suppose that L is a (time-independent, homogeneous for simplicity) Lagrangian on B, i.e., L is a real-valued function on $T(B)$. Define a function $\mathbf{L}$ on P by

$$\mathbf{L}(\sigma) = \int_0^1 L(\sigma'(t))\, dt.$$

The classical "first variation" formula can be interpreted as giving just such a "differential" for the function $\mathbf{L}$ on P. Another example: Suppose in addition that S is a submanifold of $T(B)$, defined by setting functions φ_a, $1 \leqq a \leqq m$, on $T(B)$ equal to zero. We want to consider

$$P(S) = \{\sigma \in P : \sigma'(t) \in S \quad \text{for} \quad 0 \leqq t \leqq 1\}.$$

Looked at from the point of view of infinite-dimensional differential geometry, $P(S)$ is a "submanifold" of P, obtained by setting an infinite set of functions $F^{t,a}$, indexed by $[0, 1] \times \{1, \cdots, m\}$, equal to zero. For $t \in [0, 1]$, $F^{t,a}(\sigma) = \varphi_a(\sigma(t))$. It is then plausible to extend the Lagrange multiplier rule by introducing an infinite set of Lagrange multipliers $\lambda^{t,a}$, forming

$$\mathbf{L}' = \mathbf{L} + \textstyle\sum_{t,a} \lambda^{t,a} F^{t,a},$$

considered as a function on P. However, "summation" over t is, by the "Volterra Principle", to be replaced by integration over t. Rewriting $\lambda^{t,a}$

as $\lambda_a(t)$ leads us to form the new Lagrangian $L' = L + \lambda_a(t)\varphi_a$ as before, and to find the critical points of $\mathbf{L}$ that happen to lie on $P(S)$, which is of course the starting point of our work in Section 3.

Generalizing the definition of class for a finite-dimensional problem given above, we may define the class of an extremal $\sigma \in P(S)$ as follows: The set of m-tuple functions of t, $(\lambda_a(t))$ such that σ is an extremal in the usual sense of $L + \lambda_a(t)\varphi_a$ forms a linear space; its dimension is the class of σ. (The point is that the $\lambda_a(t)$ satisfying the condition usually satisfy a system of linear homogeneous ordinary differential equations, and hence form a vector space of finite dimension.)

We now list the results of the calculation of the differentials of various functions on P.

7.2. *Suppose that f is a function on $T(B)$, $t \in [0, 1]$, and F^t is the function on P such that $F^t(\sigma) = f(\sigma(t))$ for each $\sigma \in P$. Then,*

$$dF^t(v) = df(v'(t_0)) \quad \textit{for all } v \in P_\sigma .$$

(Since v can be considered as a curve in $T(B)$, v' is, as usual, the tangent-vector curve to it, a curve in $T(T(B))$.)

7.3. *Suppose that B is a Riemannian manifold, with metric ds^2, and*

(a) $L = |\, ds \,|$. *Then,*

$$d\mathbf{L}(v) = -\frac{1}{\mathbf{L}(\sigma)} \int_0^1 (\nabla\sigma'(t), v(t))\, dt,$$

provided that σ is parameterized according to arc-length, which we can suppose with no essential loss in generality. ∇ refers to covariant differentiation of a vector field along σ [12, p. 450].

(b) $L = ds^2$. *Then,*

$$d\mathbf{L}(v) = -2 \int_0^1 (\nabla\sigma'(t), v(t))\, dt.$$

(c) $L(v) = (X(b), v)$ *for each $v \in B_b$, where X is a vector field on B. Then,*

$$d\mathbf{L}(v) = \int_0^1 (\nabla_{v(t)} X, \sigma'(t)) - (\nabla_{\sigma'(t)} X, v(t))\, dt$$

which, providing that X is a Killing vector field, is equal to

$$-2 \int_0^1 (\nabla_{\sigma'(t)} X, v(t))\, dt.$$

Lemma 7.1. *Suppose w_a, $1 \leqq a \leqq m$, are 1-forms on a manifold B, considered as Lagrangians. For $t \in [0, 1]$, $1 \leqq a \leqq m$, construct the functions $F^{t,a}$ on P. Suppose $\sigma \in P$. Suppose that $\lambda_a(t)$, $0 \leqq t \leqq 1$, are functions such that formally*

$$\int_0^1 \lambda_a(t)\, dF^{t,a},$$

a linear form on P_σ, *is identically zero. Then,* $\lambda_a(t)$ *satisfy the differential equations*:

7.4. $$\left(\frac{d}{dt}\lambda_a(t)\right) w_a + \lambda_a(t)\sigma'(t) \lrcorner\, dw_a = 0 \quad \text{at } \sigma(t), \qquad \text{for } 0 \leqq t \leqq 1.$$

Then, if 7.4 has no solutions that are not identically zero, the functions $F^{t,a}$, $0 \leqq t \leqq 1$, on P may be thought of as "functionally independent", and the subset of P defined by $F_{a,t} = F_{a,t}(\sigma)$ may, at least locally about σ, be thought of as a "submanifold" of P.

Proof. It is easy to see that the problem is purely local. We can suppose then that B has a coordinate system of functions x_i $(1 \leqq i, j, \cdots \leqq n)$. Suppose that $w_a = A_{ai}\, dx_i$. Suppose $\delta(s, t)$, $0 \leqq s, t \leqq 1$, is a homotopy of curves of B with $\delta(0, t) = \sigma(t)$, $D_s\, \delta(0, t) = v(t)$, $\delta(s, 0) = b_0$, $\delta(s, 1) = b_1$, for $0 \leqq s \leqq 1$. Put $x_i(s, t) = x_i(\delta(s, t))$. Then,

$$dx_i(v(t)) = \frac{\partial}{\partial a} x_i(s, t)\bigg|_{s=0} = \frac{\partial A_{ai}}{\partial x_j}\frac{\partial x_j}{\partial s}\frac{\partial x_i}{\partial t} + A_{ai}\frac{\partial^2 x_i}{\partial t \partial s}\bigg|_{s=0}$$
$$= \frac{\partial A_{ai}}{\partial x_j}\frac{\partial x_j}{\partial s}\frac{\partial x_i}{\partial t} - \frac{\partial A_{ai}}{\partial x_j}\frac{\partial x_j}{\partial t}\frac{\partial x_i}{\partial s} + \frac{\partial}{\partial t}\left(A_{ai}\frac{\partial x_i}{\partial s}\right)\bigg|_{s=0}.$$

Hence,

$$0 = \int_0^1 \lambda_a(t)\, dF^{t,a}$$
$$= \int_0^1 \lambda_a(t)\, dw_a(v(t), \sigma'(t))\, dt + \int_0^1 \lambda_a(t) \frac{\partial}{\partial t}\left(A_{ai}\frac{\partial x_i}{\partial s}\right)\bigg|_{s=0} dt.$$

The second term on the right-hand side, after integration by parts and taking into account the boundary conditions $v(0) = 0 = v(1)$ is

$$-\int_0^1 \left(\frac{d}{dt}\lambda_a(t)\right) w_a(\sigma'(t))\, dt,$$

whence 7.4.

Let us now express 7.4 in a more intrinsic form. Suppose that $w_a = 0$ defines a nonsingular Pfaffian system on B, i.e., if $\mathbf{H}_b = \{v \in B_b : w_a(v) = 0\}$, then dim $\mathbf{H}_b$ is constant on B. We can suppose then that the w_a are everywhere linearly independent. Let $\mathbf{H} = \{X \in V(B) : w_a(X) = 0\}$. Suppose that σ is an integral curve of H and that X is an element of $\mathbf{H}$ such that $X(\sigma(t)) = \sigma'(t)$. Suppose w is a differential form on B such that $w(\mathbf{H}) = 0$ and $w = \lambda_a(t)w_a$ on σ. It is readily seen that 7.4 is equivalent to the condition:

7.5. $$X(w) = 0 \quad \text{on } \sigma.$$

Since $w(X) = 0$, we have

7.6. $$X(w) = X \lrcorner\, dw, \qquad \text{i.e., } 0 = dw(X(\sigma(t), Y(\sigma(t))$$
$$= -w([X, Y])(\sigma(t)) \qquad \text{for } 0 \leqq t \leqq 1, \quad \text{all } Y \in \mathbf{H}.$$

For X, $Y \in V(B)$, define $\operatorname{Ad}^j X(Y)$ inductively so that

$$\operatorname{Ad}^0 X(Y) = Y, \qquad \operatorname{Ad}^j X(Y) = [X, \operatorname{Ad}^{j-1} X(Y)].$$

Using 7.5 and 7.6, applying $\operatorname{Ad}^j X$ to 7.6, we have

7.7. $w(\operatorname{Ad}^j X(Y)) = 0$ *on* σ *for all integers* $j \geqq 0$, *all* $Y \in \mathbf{H}$.

This suggests the following definitions: Let $\mathbf{H} \subset V(B)$ define a Pfaffian system (possibly with singularities) on a manifold B, i.e., $\mathbf{H}$ is a $C(B)$-submodule of $V(B)$. For $b \in B$, let $\mathbf{H}_b = \{X(b) : X \in \mathbf{H}\}$. Let $X \in \mathbf{H}$. Define a submodule $D_j(\mathbf{H}, X)$, the j^{th} *partially derived system* of $\mathbf{H}$ by X, inductively as follows: $D_0(\mathbf{H}, X) = \mathbf{H}$; $D_j(\mathbf{H}, X)$ is the submodule of $V(B)$ spanned by $D_{j-1}(\mathbf{H}, X)$ and $[X, D_{j-1}(\mathbf{H}, X)]$. Finally, let

$$D(\mathbf{H}, X) = \cup_j D_j(\mathbf{H}, X).$$

We can also write $D(\mathbf{H}, X) = D_\infty(\mathbf{H}, X)$.

7.8. *If* $X(b) = 0$ *for some* $b \in B$, *then* $D_j(\mathbf{H}, X_b) = \mathbf{H}_b$, *for all* j.

Proof. Suppose $Y \in D_{j-1}(\mathbf{H}, X)$ and that X is of the form of X', $f \in C(B)$, $X' \in \mathbf{H}$, with $f(b) = 0$. (Since $X(b) = 0$, X can be written as the sum of fields of this form.) Then,

$$[X, Y] = -Y(f)X' + f[X', Y], \quad \text{i.e.,} \quad [X, Y](b) = -Y(f)(b)X'(b) \in \mathbf{H}_b .$$

Hence, we have proved 7.8 inductively. This result allows us to define $D_j(\mathbf{H}, v) \subset B_b$, $0 \leqq j \leqq \infty$, for $v \in \mathbf{H}_b$, as follows: Choose an $X \in \mathbf{H}$ such that $X(b) = v$ and define $D_j(\mathbf{H}, v) = D_j(\mathbf{H}, X)_b$.

Returning now to X and w as defined above, e.g. in 7.5, we see that 7.7 implies

7.9. $w(D(\mathbf{H}, \sigma'(t))) = 0$ *for* $0 \leqq t \leqq 1$. *In particular, if* $D(\mathbf{H}, \sigma'(t)) = B_{\sigma(t)}$ *for* $0 \leqq t \leqq 1$, *then* 7.4 *admits no nonzero solutions.*

We now want to show that if $\dim D(\mathbf{H}, \sigma'(t))$ is constant for $0 \leqq t \leqq 1$ and less than $\dim B$, then 7.4 admits nonzero solutions. This is the special case of the following result:

PROPOSITION 7.2. *Let* $\mathbf{H} \subset V(B)$ *define a Pfaffian system (possibly with singularities) on a manifold* B. *Let* S *be a submanifold of* B *such that* $\dim \mathbf{H}_b$ *is constant for* $b \in S$. *Let* X *be a vector field on* B *such that* (1) X *is tangent to* S, *and* (2) $[X, \mathbf{H}]_b \subset \mathbf{H}_b$ *for* $b \in S$. *Let* w_0 *be a* 1*-covector to* B *at* $b_0 \in S$ *such that* $w_0(\mathbf{H}_{b_0}) = 0$. *Suppose that* $X(b_0) = 0$. *Then, there is a neighborhood* U *of* b_0 *and a* 1*-form* w *on* U *such that* (a) $w = w_0$ *at* b_0, (b) $w(\mathbf{H}_b) = 0$ *for* $b \in U \cap S$, *and* (c) $X(w) = 0$ *on* $U \cap S$.

Proof. We can first assume that U is chosen so small that there are everywhere independent 1-forms w_a $(1 \leqq a, b, \cdots \leqq m)$ in U such that

$$w_a(v) = 0 \quad \text{if and only if} \quad v \in H_b \qquad \text{for } v \in B_b,\ b \in S \cap U.$$

Then, there are functions (A_{ab}) on $S \cap U$ such that

$$X(w_a) = A_{ab} w_b \quad \text{on } S \cap U.$$

We look for w of the form $A_a w_a$ for some functions (A_a) on S. The A_a must then satisfy

7.10. (a) $A_a(b_0)w_a = w_a$ at b_0, (b) $X(A_a) = A_{ab} A_b$.

Let $\sigma : [0, 1] \to S \cap U$ be an integral curve of X. Then, 7.10 (b) implies that

$$\frac{d}{dt} A_b(\sigma(t)) = A_{ab}(\sigma(t))A_b(\sigma(t)),$$

i.e., the A_a along σ are the solutions of a system of linear homogeneous differential equations. Then, to find locally solutions of 7.10 (b), we have only to choose any hypersurface S' of S transversal to the integral curves of X and solve these linear differential equations along the integral curves of X starting at points of S', Q.E.D.

We can now give the following heuristic answer to the problem with which we began. Let $\mathbf{H}$ be a nonsingular Pfaffian system on B. Let P be the space of C^∞ curves on B joining two fixed points b_0 and b_1. Let $P(\mathbf{H})$ be the subset of P consisting of the integral curves of H. Then, the *formal* condition that $P(H)$ be a "submanifold" in a neighborhood of $\sigma \in P(H)$ and that the Lagrange multiplier rule work is that

7.11. $$D(\mathbf{H}, \sigma'(t)) = B_{\sigma(t)}, \qquad \text{for } 0 \leqq t \leqq 1.$$

Example.

$$\omega_a = dx_a - (A_{ab} x_b + V_{au} y_u)\, dt = 0,$$
$$1 \leqq a, b, \cdots \leqq m; m + 1 \leqq u, v, \cdots \leqq n,$$

$A_{ab} = A_{ab}(t)$, $V_{au} = V_{au}(t)$ functions of t, determines a Pfaffian system in the space B of variables (X_a, y_u, t). The subspace $\mathbf{H} \subset V(B)$ annihilating the ω_a is spanned by

$$X_u = \partial/\partial y_u, \qquad Y = \partial/\partial t + (A_{ab} x_b + V_{au} y_u)\, \partial/\partial x_a.$$

Introduce a matrix-vector notation:

$A = (A_{ab})$, an $(m \times m)$-matrix function of t.

$V_u = (V_{au})$, an $(m \times 1)$-vector function of t.

$\partial/\partial x = (\partial/\partial x_a)$, a $(1 \times m)$-vector operator function of t.

In terms of matrix multiplication,

$$Y = \frac{\partial}{\partial t} + \frac{\partial}{\partial x}(Ax + V_u y_a)$$

(u is, in this notation, considered to be merely a counting index and *not* a vector index). Introduce a "covariant differentiation" operation ∇ on matrix functions of t:

$$\nabla V = \frac{d}{dt} V - AV,$$

if $V = (V_i)$ is a 1-vector function of t. Computing, we have

$$[Y, X_a] = \frac{\partial}{\partial x} V_a = \frac{\partial}{\partial x} (\nabla^0 V_a),$$

$$[X_a, [Y, X_a]] = 0, \qquad [[Y, X_u], [Y, X_v]] = 0,$$

$$[Y, [Y, X_u]] = -\frac{\partial}{\partial x} (\nabla^1 V_u),$$

etc., i.e.,

$$\mathrm{Ad}^N(Y)(X_u) = -\frac{\partial}{\partial x} (\nabla^{N-1} V_u) \quad \text{and} \quad [\mathrm{Ad}^N(Y)(X_u), \mathrm{Ad}^M(Y)(X_u)] = 0$$

if N and M are > 0. Combining the calculations with W. L. Chow's accessibility theorem [11], we have

PROPOSITION 7.3. *Suppose that the dimension of the space of* $(m \times 1)$-*vectors spanned by the* $V_u(t)$, $\nabla V_u(t)$, $\nabla^2 V_u(t)$, $\cdots$ *is independent of* t; *call it* p. *Then, every point of* (x, u, t)-*space can be reached by starting from a given point along piecewise differentiable integral curves of* H *if and only if* $p = m$.

Now suppose $Z = z_u X_u + z_0 Y \in \mathbf{H}$. We want to calculate dim $D(\mathbf{H}, Z)$. In carrying out these calculations, we might as well suppose that z_u and z_0 are constants. Then,

$$[Z, X_u] = z_0[Y, X_u] = -z_0 \frac{\partial}{\partial x} (\nabla^0 V_u),$$

$$[Z, Y] = z_u[X_u, Y] = z_u \frac{\partial}{\partial x} (\nabla^0 V_u);$$

hence,

(i) If $z_0 \neq 0$, $D(\mathbf{H}, Z) = D(\mathbf{H})$.

(ii) If $z_0 = 0$, dim $D(\mathbf{H}, Z) -$ dim $\mathbf{H} = 1$, unless $z_u V_u = 0$, in which case $D(\mathbf{H}, Z) = \mathbf{H}$.

If $D(\mathbf{H}) = V(B)$, then case (i) implies 7.11 is satisfied. Notice also that the integral curves of fields Z satisfying (i) (with z_0 not necessarily constant) are precisely those which can be reparameterized by t so as to be a solution of the system:

7.12. $$\frac{dx_a}{dt} = A_{ab}\, x_b(t) + V_{au}\, y_u(t), \qquad \text{for some choice of } y_u(t).$$

Returning to the general case where $\mathbf{H}$ is a Pfaffian system on a manifold B, we now investigate another infinitesimal property of $P(\mathbf{H})$. First, we must make precise what is meant by the tangent space $P(\mathbf{H})_\sigma$ at a "point" $\sigma \in P(\mathbf{H})$.

$$1 \leqq a, b, \cdots \leqq m; \quad m + 1 \leqq u, v, \cdots \leqq n.$$

Suppose (ω_i), $1 \leqq i, j, \cdots \leqq n$, is a basis of 1-forms such that $\omega_a = 0$ determines $\mathbf{H}$. Suppose that

$$d\omega_i = c_{jki}\,\omega_j \wedge \omega_k .$$

Let $\sigma : [0, 1] \to B$ be an integral curve of $\mathbf{H}$ with $\sigma(0) = x_0$, and $\sigma^*(\omega_a) = 0$. Change the definition of $P(\mathbf{H})$ slightly: It is now the set of integral curves of $\mathbf{H}$ starting at x_0, with the end-point $\sigma(1)$ free to wander. To define $P(\mathbf{H})_\sigma$, proceed as follows: Let σ_s, $0 \leqq s \leqq 1$, be a curve in $P(\mathbf{H})$, with $\delta(s, t) = \sigma_s(t)$. The corresponding infinitesimal deformation $v(x) = D_s\,\delta(0, t)$ then satisfies a certain set of *linear* differential equations, the "linear variational equations" of the Pfaffian system $\omega_a = 0$ about σ; $P(\mathbf{H})_\sigma$ is then the set of vector fields satisfying these equations. (It is not a priori obvious that *every* solution arises in this way from a curve in $P(\mathbf{H})$. This question in turn is related to the "submanifold" structure of $P(\mathbf{H})$ but will not be pursued here.) Now, let $\phi : P(\mathbf{H}) \to B$ be the mapping $\sigma \to \sigma(1)$. The "differential" at σ, $\phi_* : P(\mathbf{H})_\sigma \to B_{\sigma(1)}$ is then the mapping which assigns to each $v \in P(\mathbf{H})_\sigma$ the vector $v(1)$ at $\sigma(1)$. ϕ is said to be of *maximal rank about* σ if ϕ_* is onto.

To derive the linear variational equations, suppose then that $\delta(s, t)$ is a homotopy, with $\delta(0, t) = \sigma(t)$. We have

$$\frac{\partial}{\partial s}\,(\omega_i(D_t\,\delta)) - \frac{\partial}{\partial t}\,(\omega_i(D_s\,\delta)) = c_{jki}(\delta(s, t))\omega_j(D_t\,\delta)\omega_k(D_s\,\delta).$$

Hence

7.13. $$\frac{d}{dt}\,\omega_a(v(t)) = c_{jua}(\sigma(t))\omega_j(v(t))\omega_u(\sigma'(t)), \quad \text{or}$$

7.14. $$d\omega_a(v(t), \sigma'(t)) - \frac{d}{dt}\,\omega_a(v(t)) = 0.$$

$P(H)_\sigma$ is then the set of vector fields v along σ satisfying 7.14, and the boundary condition $v(0) = 0$.

THEOREM 7.4. *ϕ is of maximal rank about σ if 7.11 is satisfied.*

Proof. Given $v^1 \in B_{\sigma(1)}$, we must prove there is a vector field $v(t)$ satisfying 7.13 with $v(1) = v^1$. Now 7.13 is a differential equation of type 7.12, with $\omega_a(v(t))$ identified with $x_a(t)$, $\omega_u(v(t))$ identified with $y_u(t)$, $A_{ab}(t)$ identified with $c_{bua}(\sigma(t))\omega_u(\sigma'(t))$, and with $v_{au}(t)$ identified with $c_{uva}(\sigma(t))\omega_v(\sigma'(t))$. Then, the proof follows from Proposition 7.3 if we can identify the codimension of the space of $(m \times 1)$-vectors spanned by

V_u, ∇V_u, $\cdots$ with the codimension of $D(\mathbf{H}, \sigma'(t))$ in $B_{\sigma'(t)}$. This follows from the following argument.

Let (E_i) be a basis of vector fields dual to the ω_i, i.e., $\omega_i(E_j) = \delta_{ij}$. Then, (E_u) span $\mathbf{H}$. Let $E = f_u E_u$ be a field such that σ is an integral curve of E, i.e.,

$$f_u(\sigma(t)) = \omega_u(\sigma'(t)).$$

Then,

$$[E, E_v] \equiv f_u c_{uva} E_a \pmod{\mathbf{H}},$$

i.e., $[E, E_v]$ on σ can, mod $\mathbf{H}$, be identified with $V_{au}(\sigma(t))E_a(\sigma(t))$. Notice further that $[E, [E, E_v]]$ on σ can, mod $\mathbf{H}$, be identified with $(\nabla V_u)_a E_a(\sigma(t))$, etc., Q.E.D.

8. The connection with Carathéodory's definition of the class of an extremal

Let $K(x_i, y_i)$ be a function of $2n$ real variables (x_i, y_i), $(1 \leqq i, j, \cdots \leqq n$; summation convention). Let M be the space of these $2n$ variables, and let σ be the curve determined by functions $x_i(t)$, $y_i(t)$, $0 \leqq t \leqq 1$, that are solutions of the Hamilton equations:

8.1. $$\frac{d}{dt} x_i(t) = \frac{\partial K}{\partial y_i}(x, y), \qquad \frac{d}{dt} y_i(t) = \frac{\partial K}{\partial x_i}(x, y).$$

The linear variational equations based on σ are

8.2. $$\begin{aligned} \frac{d}{dt} X_i(t) &= \frac{\partial^2 K}{\partial y_i\, \partial x_j}(\sigma(t))X_j(t) + \frac{\partial^2 K}{\partial y_i\, \partial y_j}(\sigma(t))Y_j(t), \\ \frac{d}{dt} Y_i(t) &= -\frac{\partial^2 K}{\partial x_i\, \partial x_j}(\sigma(t))X_j(t) - \frac{\partial^2 K}{\partial x_i\, \partial y_j}(\sigma(t))Y_j(t). \end{aligned}$$

Consider the solutions of 8.2 for which the $X_i(t)$ are identically zero, i.e., the solutions of

8.3. $$\frac{d}{dt} Y_i(t) = -\frac{\partial^2 K}{\partial x_i\, \partial y_j}(\sigma(t))Y_j(t), \qquad \frac{\partial^2 K}{\partial y_i\, \partial y_j}(\sigma(t))Y_j(t) = 0.$$

The solutions of 8.3, considered as vector-valued functions of t, form a vector space. Its dimension is the *class* of the extremal σ in the sense of Carathéodory [5], C-class for short.

The C-class of σ has the following geometric meaning: Under certain additional assumptions [5], if the class is zero, all points of x-space sufficiently close to $x_i(1)$ can be joined to $x_i(0)$ by a curve which is the projection on x-space of a solution of 8.1.

Let $\omega_a = a_{ai}\, dx_i$ be everywhere independent 1-forms such that $\omega_a = 0$ defines a nonsingular Pfaffian system $\mathbf{H}$, $1 \leqq a, b, \cdots \leqq m$, on B, the space of the variables (x_i). Suppose that

$$a_{ai}(x)K_{n+i}(x, y) = 0,$$

i.e., the solutions of the Hamilton equations 8.1 when projected down to x-space are integral curves of **H**. Further, suppose that rank $(K_{n+i,n+j}) = n - m$, i.e., if $\lambda_i K_{n+i,n+j}(x^0, y^0) = 0$ at a point (x^0, y^0), then the vector $\lambda = (\lambda_i)$ is a linear combination of the vectors $(a_{\alpha i}(x^0))$.

$$(K_i = \partial K/\partial x_i, \qquad K_{n+i} = \partial K/\partial y_i, \qquad K_{n+i,j} = \partial^2 K/\partial y_i\, \partial x_j, \qquad K_{n+i,n+j} = \partial^2 K/\partial y_i\, \partial y_j, \quad \text{etc.})$$

Let $\sigma(t) = (x(t), y(t))$ be a solution of 8.1, and let $\sigma_B(t) = (x(t))$ be the projection onto x-space. Let $Y_i(t)$ be a solution of 8.3. Let ω be a 1-form such that $\omega(\mathbf{H}) = 0$ and $\omega(\sigma_B(t)) = Y_i(t)\, dx_i$, and let X be a vector field such that $X(\sigma_B(t)) = \sigma_B'(t)$. If $\omega = Y_i\, dx_i$ and $X = X_i\, \partial/\partial x_i$,

$$X(\omega) = X(Y_i)\, dx_i + Y_i\, d(X_i).$$

We can suppose that there are functions $y_i(x)$ such that

$$X_i(x) = K_{n+i}(x, y(x)), \qquad y(x(t)) = y(t).$$

(We use the fact that rank $K_{n+i,n+j} = n - m$.) Then,

$$dX_i = K_{n+i,j}(x, y(x))\, dx_j + K_{n+i,n+j}\, dy_j(x).$$

Hence, by using 8.3,

8.4. $X(\omega)(\sigma_B(t)) = (dY_i(t)/dt)\, dx_i + Y_i K_{n+i,j}(\sigma(t))\, dx_j$, *i.e.*, $X(\omega) = 0$ *at every point of* $\sigma_B(t)$.

We are now back to the situation dealt with in Section 7. Applying the results and notation developed there, we have

THEOREM 8.1. *The C-class of the extremal σ is no greater than*

$$\min_{0 \leqq t \leqq 1} (\dim B - \dim D(\mathbf{H}, \sigma_B'(t))).$$

In particular, if $D(\mathbf{H}, \sigma_B'(t)) = B_{\sigma(t)}$ for $0 \leqq t \leqq 1$, the C-class is zero. Then, if in addition the Legendre condition that the quadratic form $(K_{n+i,n+j})$ be positive semidefinite is satisfied, and if σ_B is sufficiently small, every point of x-space sufficiently near to $\sigma_B(1)$ can be reached as the end-point of an extremal, i.e., as the projection of a solution of 8.1, *starting at $\sigma_B(0)$.*

This result has the following intuitive meaning: If $D(\mathbf{H}) = V(B)$ by Chow's accessibility theorem [11], every point of B can be reached as the end-point of an integral curve of **H** starting at a fixed point $x_0 = \sigma_B(0)$, i.e., if $P(\mathbf{H}) \to B$ is the space of integral curves of **H** starting at x_0, then $\phi(P(\mathbf{H})) = B$. As we have seen,

$$D(\mathbf{H}, \sigma_B'(t)) = B_{\sigma(t)}, \qquad 0 \leqq t \leqq 1,$$

is the condition that $\phi_* : P(\mathbf{H})_{\sigma_B}$, the differential of ϕ at σ_B, be onto $B_{\sigma_B(1)}$. Since $P(\mathbf{H})$ is some sort of infinite-dimensional manifold, the implicit function theorem cannot be used a priori to prove that ϕ maps a neighborhood of σ_B in $P(\mathbf{H})$ onto a neighborhood of $\sigma_B(1)$. However, if σ_B is an extremal, Theorem 8.1 gives conditions that this be so, and even that points near $\sigma_B(1)$ can be reached on extremals that are near to σ_B. This stability phenomenon might be important in applied problems.

9. Appendix. A proof of the local minimizing property of the formal extremals

Despite the fact that it forms one of the most elegant parts of the classical theory, Carathéodory's proof of the local minimizing property of the extremals of a Lagrange variational problem [4] does not seem to be well known. Of all the classical methods, it is best suited to the needs of applied mathematics, where "nonclassical" problems involving inequality constraints arise [16]. We present a treatment, based on Carathéodory's ideas, for the case where the Lagrangian and constraint functions are homogeneous and time-independent. (The general case can be reduced to this one.)

Let B be a manifold of dimension n ($1 \leqq i, j, \cdots \leqq n$). Let $T(B)$ be the tangent bundle to B. Let L (resp. φ_a, $1 \leqq a, b, \cdots \leqq m$) be real-valued functions on $T(B)$ such that

9.1. $L(rv) = rL(v)$ *(resp.* $\varphi_a(rv) = r\varphi_a(v)$*) for each* $r > 0$, $v \in T(B)$.

For a curve $\sigma : [a, b] \to B_1$, define

$$\mathbf{L}(\sigma) = \int_a^b L(\sigma'(t))\, dt.$$

Because of 9.1, it is independent of the parameterization of σ. For $x \in B$, let L^x (resp. φ_a^x) be the function on B_x resulting from restricting L (resp. φ_a) to B_x. Introduce local coordinates (x_i) for B, and the corresponding local coordinate $(x_i, \dot{x}_i)$ for $T(B)$. Let $L(x_i, \dot{x}_i)$ and $\varphi_a(x_i, \dot{x}_i)$ be the expressions for the corresponding functions in local coordinates. Let

$$L_i(x, \dot{x}) = \frac{\partial L}{\partial \dot{x}_i}(x, \dot{x}), \qquad \varphi_{ai} = \frac{\partial \varphi_a}{\partial \dot{x}_i}, \qquad L_{ij} = \frac{\partial^2 L}{\partial \dot{x}_i\, \partial \dot{x}_j}, \quad \text{etc.}$$

By Euler's relations for homogeneous functions (i.e., by suitably differentiating 9.1), we have

9.2. (a) $L_i \dot{x}_i = 0$,
(b) $L_i(x, r\dot{x}) = L_i(x, \dot{x})$ for $r > 0$,
(c) $L_{ij} \dot{x}_j = 0$.

(Similar relations hold for the φ_a, since they too are homogeneous.) If we construct the 1-form $\theta(L) = L_i\, dx_i - (L_i \dot{x}_i - L)\, dt$ on $T(B) \times R$, notice, by 9.2 (a) that $\theta(L) = L_i\, dx_i$, i.e., $\theta(L)$ is a form on $T(B)$ alone. We shall so consider it from now on and shall forget about explicit time dependence. Similarly, $\theta(\varphi_a) = \varphi_{a,i}\, dx_i$ are 1-forms on $T(B)$.

Let $S \subset T(B)$ be the constraint manifold defined by the φ_a, i.e.,

$$S = \{v \in T(B) : \varphi_a(v) = 0\}.$$

For $x \in B$, let $S^x = S \cap B_x$. We shall suppose that the following classical conditions are verified.

9.3. *The matrix φ_{ai} has constant rank m. In particular, S (resp. S^x, for $x \in B$) is an imbedded submanifold of $T(B)$ (resp. B_x).*

We shall suppose, as explained in Section 3, that M', the normal bundle of S, is $S \times R^m$, with (λ_a) the coordinates of R^m, i.e., M' is the subset of $R^n \times R^n \times R^m$ defined by $\varphi_a = 0$. Consider the form $\theta(L, S)$ on M':

$$\theta(L, S) = \theta(L) + \lambda_a\, \theta(\varphi_a) \qquad \text{restricted to } S \times R^m.$$

A real-valued function W on B is a *solution* of the *Hamilton-Jacobi equation* if

9.4. *There is a cross-section $f : B \to M' = S \times R^m$ and a function $F(\ )$ of one variable such that*

$$f^*(\theta(L, S)) = d(F(W)).$$

To see how such conditions on W arise in a natural geometric way, consider a cross-section $g : B \to S$ and a real-valued function W on B such that

9.5. (a) $dW(g(x)) = 1$ *for all* $x \in B$.

(b) *For each $x \in B$, $g(x)$ is a critical point of L^x restricted to*

$$S^x \cap \{v \in B_x : dW(v) = 1\}.$$

(c) *For all $x \in B$, dW and φ_a, considered as functions on B_x, are functionally independent in a neighborhood of S^x.*

Then, according to the Lagrange multiplier rule for ordinary functions, for each $x \in B$ there is a system of numbers $(\lambda(x), \lambda_a(x))$ such that $g(x)$ is a critical point of the function $L^x + \lambda(x)(dW(x) - 1) + \lambda_a(x)\varphi_a^x$ on B^x ($dW(x)$ denotes dW restricted to B_x). Further, the $\lambda(x)$ and $\lambda_a(x)$ are, by 9.5 (c), uniquely determined and hence are differentiable functions of x, by the implicit function theorem.

In local coordinates, put $g_i(x) = \dot{x}_i(g(x)) = dx_i(g(x))$ and $W_i = \partial W/\partial x_i$. Then, the conditions stated take the form:

9.6. $L_i(x, g(x)) + \lambda(x)W_i(x) + \lambda_a(x)\varphi_{ai}(x, g(x)) = 0 \qquad$ for each $x \in B$.

Let $f : B \to S \times R^m = M'$ be the cross-section such that

$$f(x) = (g(x), \lambda_a(x)) \qquad \text{for } x \in B.$$

Then, we see that 9.6 is equivalent to the condition:

9.7. $$f^*(\theta(L, S)) = -\lambda(x)\, dW.$$

Further, multiplying each term of 9.6 by $g_i(x)$ and summing, we have

$$0 = L_i(x, g(x))g_i(x) + \lambda(x)W_i(x)g_i(x) + \lambda_a(x)\varphi_{ai}(x, g(x))g_i(x)$$
$$= L(x, g(x)) + \lambda(x), \qquad \text{by 9.2 and 9.5(a).}$$

Hence, we have

9.8. *$g^*(L)$ is a constant on the level surfaces of W if and only if W is a solution of the Hamilton-Jacobi equation.*

THEOREM 9.1. *Suppose $f : B \to M' = S \times R^m$ and $g : B \to S$ are cross-sections such that $f(x) = (g(x), \lambda_a(x))$ for all $x \in B$. If W is a corresponding solution of the Hamilton-Jacobi equation such that* 9.5 (a) *is satisfied, then* 9.5 (b) *is satisfied. Suppose that the further conditions, strengthening* 9.5 (b), *are satisfied*:

9.9. *For each $x \in B$, $L(g(x)) < L(v)$ for all $v \in S^x$ such that $dW(v) = 1$ and $v \neq g(x)$.*

9.10. *The surfaces W = constant are connected.*

9.11. *$L(v) > 0$ if $v \neq 0$, $v \in S$.*

Let $\sigma : [a, b] \to B$ be a curve such that $\sigma'(t) = g(\sigma(t))$ for $a \leqq t \leqq b$, i.e., σ is an integral curve for the vector field determined by g. If $\sigma_1 : [a_1, b_1] \to B$ is a curve with $W(\sigma_1(a_1)) = W(\sigma(a))$, $W(\sigma_1(b_1)) = W(\sigma(b))$ and with $\sigma_1'(t) \in S$ for $a_1 \leqq t \leqq b$, then $\mathbf{L}(\sigma_1) > \mathbf{L}(\sigma)$ unless σ_1 is also an integral curve of g (with a possible change of parameterization), in which case $\mathbf{L}(\sigma_1) = \mathbf{L}(\sigma)$.

We may remark that it is, in the classical theory, customary to assure 9.9 by assuming a "Legendre condition" on the second partial derivatives L_{ij} and φ_{aij}.

Proof. The first part is easily proved by reversing the reasoning which led to 9.7. To prove the second part, note first that we are free to change the parameterization of σ so that $W(\sigma(t)) = t$, $a \leqq t \leqq b$.

Case 1. $dW(\sigma_1(t))/dt > 0$ for $a_1 \leqq t \leqq b_1$. We can then, by an allowed change of parameterization, suppose also that $W(\sigma_1(t)) = t$. By 9.7, 9.8, 9.9, and 9.10, for $x \in B$

$$L(g(x)) = \min \{L(v) : v \in S^y, y \in B, dW(v) = 1, W(y) = W(x)\}.$$

Then, $L(\sigma_1'(t)) \leqq L(\sigma'(t))$ for $W(a) \leqq t \leqq W(b)$, and equality holds if and only if $\sigma_1'(t) = g(\sigma_1(t))$. Then, $L(\sigma_1) \geqq L(\sigma)$, and equality holds, by continuity of σ' and σ_1' if and only if σ_1 is an integral curve of g.

Turning to the general case, notice that $[a_1, b_1]$ can be broken into subintervals in which $dW(\sigma_1(t))/dt$ is (a) positive or (b) nonpositive. Since $W(\sigma_1(t))$ must go from $W(\sigma(a))$ to $W(\sigma(b))$ as t goes from a_1 to b_1, and $L(\sigma_1)$ over an interval satisfying (a) is $\geqq L(\sigma)$ over that interval, and $L(\sigma_1) > 0$ over any intervals satisfying (b), the reasoning by which one obtains $L(\sigma_1) \geqq L(\sigma)$ is an obvious consequence of the fact that $L(\sigma_3) = L(\sigma_1) + L(\sigma_2)$ if a curve σ_3 is obtained by putting end-to-end curves σ_1 and σ_2.

COROLLARY 9.2. $\mathbf{L}(\sigma_1) > \mathbf{L}(\sigma)$ *for all other curves* σ_1 *satisfying the constraints joining* $\sigma(a)$ *to* $\sigma(b)$, *if conditions* 9.9, 9.10, *and* 9.11 *are satisfied.*

This does not exhaust the classical results. The next problem: Given a sufficiently small formal extremal σ of the Lagrange problem, i.e., the projection into B of a characteristic curve of $d\theta(L, S)$, can we find cross-sections $f : U \to M'$, $g : U \to S$, with $f(x) = (g(x), \lambda_a(x))$, defined in a sufficiently small neighborhood U of σ such that σ is an integral curve of g and there is a function W in U with $f^*(\theta(L, S)) = dW$?

This problem can be handled by standard Hamilton-Jacobi theory, as Carathéodory does [4], by showing the equivalence, via a "Legendre transformation", with a Cauchy problem for a first-order partial differential equation. However, it is also possible to deal with this "Cauchy problem" directly by using Cartan's theory of exterior differential systems. We shall *sketch* such a treatment here:

First, eliminate consideration of W by noticing that we are looking for a cross-section $f : U \to M'$ such that $f^*(d\theta(L, S)) = 0$. Consider $f(U)$ as a submanifold N of dimension n of M'. Notice that it is an integral manifold of the 2-form $d\theta$ ($= d\theta(L, S)$), i.e., $d\theta$ restricted to it is zero. The *characteristic vector fields* X (on M') of the differential ideal generated by $d\theta$ are then the $X \in V(M')$ such that $X \lrcorner d\theta = 0$, i.e., the vector fields defining what we called the characteristic foliation F of $d\theta$ in Section 2. The following property of the $X \in F$ is well known:

If $N_1 \subset M'$ is an integral submanifold of $d\theta$ to which X is nowhere tangent, there is an integral submanifold of $d\theta$ of one higher dimension containing N_1, obtained by finding the integral curves of X whose origin lies on N_1.

Then, we may hope to find integral manifolds of $d\theta$ of dimension n by finding integral manifolds of lower dimension which are not tangent to the characteristic foliation of $d\theta$.

DEFINITION. If A is a submanifold of B and $\tau : A \to M'$ is a cross-section map, then τ is *transversal to* A (with respect to the Lagrange variational problem) if $\tau^*(\theta) = 0$. Then, $\tau(A)$ is a submanifold of M' of the same dimension as A which is a fortiori an integral submanifold of $d\theta$.

Under suitable regularity conditions, it is easy to see that (a) $\tau(A)$ is not tangent to F, and (b) the existence of such τ's, given A with $\dim A < \dim B$, can be proved by invoking the implicit function theorem only. Rather than work out these points in the general case, it is more instructive to look at the special case considered in Section 4.

Suppose then that L defines a Riemannian metric on B and that the constraints are linear. As before, let (w_i) be an orthonormal moving frame of 1-forms on B, pulled up and considered as 1-forms on $T(B)$, and let (y_i) be the functions on $T(B)$ such that $w_i(v) = y_i(v)$ for $v \in T(B)$. Then, $L = (y_i y_i)^{1/2}$. Further, let the (w_i) be chosen so that $y_a = 0$ defines the constraint submanifold S of $T(B)$. $(1 \leqq a, b, \cdots \leqq m; m + 1 \leqq u, v, \cdots \leqq n.)$ The 1-form $\theta(L, S)$ $(= \theta)$ on $S \times R^m$ is then

$$\theta = y_u w_u/(y_i y_i)^{1/2} + \lambda_a w_a .$$

Let A be a submanifold of B, and let $T : A \to S \times R^m$ be a cross-section defined by giving y_u and λ_a as functions of $x \in A$, say $y_u(x)$ and $\lambda_a(x)$. The condition for transversality is then

9.12. $$y_u(x)w_u/(y_i(x)y_i(x))^{1/2} + \lambda_a(x)w_a = 0$$

for $x \in A$, (w_i) restricted to A.

For $x \in B$, consider H_x (resp. V_x), the set of horizontal (resp. vertical) vectors that satisfy $w_a = 0$ (resp. $w_u = 0$). Let h (resp. λ) $: A \to T(B)$ be the horizontal (resp. vertical) tangent vector field on A such that

$$w_u(h(x)) = y_u(x) \quad (\text{resp. } w_a(\lambda(x)) = \lambda_a(x)).$$

Then, 9.12 is equivalent to

9.13. $$(h(x), v)/| h(x) | + (\lambda(x), v) = 0 \quad \text{for all } x \in A, v \in A_x .$$

This suggests making the following definition:

DEFINITION. Let $H : x \to H_x$ be a field of tangent space subspaces of constant dimension on a Riemannian manifold B. For $x \in B$, let $P : B_x \to H_x$ (resp. $P^{\perp} : B_x \to H_x^{\perp}$) be the corresponding projections. A tangent vector $v \in B_x$ is said to be *transversal* to a $v_1 \in B_x$ if

9.14. $$(P(v), v_1)/| P(v) | + (P^{\perp}(v), v_1) = 0.$$

(Notice that this is not necessarily a transitive relation, i.e., v_1 is not necessarily transversal to v, unless dim $H = 0$, when the condition reduces to the standard notion of *perpendicular* vectors. This is, of course, typical of variational problems more general than those provided by Riemannian metrics.)

BIBLIOGRAPHY

1. W. AMBROSE, *The Cartan structural equations in classical Riemannian geometry*, J. Indian Math. Soc. (N.S.), vol. 24 (1960), pp. 23–76.
2. G. A. BLISS, *The problem of Lagrange in the calculus of variations*, Amer. J. Math., vol. 52 (1930), pp. 673–744.
2a. R. BOTT AND H. SAMELSON, *Applications of the theory of Morse to symmetric spaces*, Amer. J. Math., vol. 80 (1958), pp. 964–1029.
3. H. BUSEMANN, *The geometry of geodesics*, New York, Academic Press, 1955.
4. C. CARATHÉODORY, *Die Methode der geodätischen Äquidistanten und das Problem*

von Lagrange, Acta Math., vol. 47 (1926), pp. 199–236, or Gesammelte Mathematische Schriften, Bd. I, pp. 212–248.
5. ———, *Über die Einteilung der Variationsprobleme von Lagrange nach Klassen*, Comment. Math. Helv., vol. 5 (1933), pp. 1–19, or Gesammelte Mathematische Schriften, Bd. I, pp. 270–288.
6. ———, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, Leipzig und Berlin, B. G. Teubner, 1935.
7. E. CARTAN, *Leçons sur les invariants intégraux*, Paris, Hermann, 1922.
8. ———, *Leçons sur la géométrie des espaces de Riemann*, 2nd ed., Paris, Gauthier-Villars, 1946.
9. ———, *Les systèmes differentiels extérieures et leurs applications géométriques*, Paris, Hermann, 1945.
10. C. CHEVALLEY, *Theory of Lie groups I*, Princeton, Princeton University Press, 1946.
11. W. L. CHOW, *Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung*, Math. Ann., vol. 117 (1940/1941), pp. 98–105.
12. R. HERMANN, *On the differential geometry of foliations*, Ann. of Math. (2), vol. 72 (1960), pp. 445–457.
13. ———, *The differential geometry of foliations, II*, J. Math. Mech., vol. 11 (1962), pp. 303–315.
14. ———, *Totally geodesic orbits of groups of isometries*, Nederl. Akad. Wetensch. Proc. Ser. A, vol. 65 (1962), pp. 291–298.
15. K. NOMIZU, *Invariant affine connections on homogeneous spaces*, Amer. J. Math., vol. 76 (1954), pp. 33–65.
16. L. S. PONTRJAGIN, *Optimal control processes*, Uspehi Mat. Nauk, vol. 14 (1959), pp. 3–20 (in Russian).
17. B. L. REINHART, *Foliated manifolds with bundle-like metrics*, Ann. of Math. (2), vol. 69 (1959). pp. 119–132.
18. V. VOLTERRA, *Theory of functionals*, New York, Dover, 1959.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASSACHUSETTS