MATH. SCAND. 42 (1978), 313–333

THE EULER DERIVATIVE

AN INTRINSIC APPROACH TO THE CALCULUS OF VARIATIONS

JAAK PEETRE

0. Introduction.

In this note we consider variational problems involving ν-jets of curves and we wish in particular to give an intrisic form of the Euler equations analogous to the one we obtained in [4] (cf. also [5], [6]) for $\nu = 1$. In [4] we also considered multiple integral problems but presently at least we refrain from that generalization.

A standard reference to the theory of jets is [2]. We also make extensive use of the fact that the bundle of ν-jets can be canonically mapped into the νth order osculating bundle (see e.g. [7]). Since the latter is a vector bundle we thereby achieve a certain *linearization* of the setup.

For an over all treatment of the Calculus of Variations from the classical standpoint see [9].

NOTATION.

$$\binom{\nu}{k} = \frac{\nu!}{k!(\nu-k)!} \quad (0 \leqq k \leqq \nu)$$

are the binomial coefficients; they are often interpreted as 0 if $k<0$ or $k>\nu$. l.c. = local coordinate.

1. Jets.

Let M be a C^∞ manifold, $n = \dim M$.

A *curve* in M is a C^∞ map $x \colon \mathsf{R} \to M$. Let $t \in \mathsf{R}$ and ν an integer >0 be given. Regard two curves x and y as equivalent if they have contact up to order ν at t, i.e. in terms of l.c.'s:

$$\frac{d^k x^i(t)}{dt^k} = \frac{d^k y^i(t)}{dt^k} \qquad (i=1,\ldots,n,\ 0 \leqq k \leqq \nu)\,.$$

Received June 6, 1977.

The ν-*jet* at t of x is the equivalence class to which x belongs. We denote it by $x^{(\nu)}(t)$. If $\nu=1$ we can identify it with a tangent vector $x'(t)$ of M, i.e. an element of the tangent bundle $T=TM$ of M. The set of all ν-jets at t of curves is noted by $J_t^\nu = J_t^\nu(\mathsf{R},M)$ and their union again by $\bar{J}^\nu = \bar{J}^\nu(\mathsf{R},M)$ $(=\bigcup_{t\in\mathsf{R}} J_t^\nu(\mathsf{R},M))$. Thus $J_t^1 = J_t^1(\mathsf{R},M)\approx T$. Some of our considerations are formally valid also for $\nu=0$ if we agree to interpret $J_t^0=J_t^0(\mathsf{R},M)$ as M. J_t^ν is a bundle over M, with projection $x^{(\nu)}(t)\mapsto x(t)$. $\bar{J}^\nu$ is a bundle over $\mathsf{R}\times M$ with projection $x^{(\nu)}(t)\mapsto (t,x(t))$. There is also a canonical map $J_t^\nu \to J_t^{\nu-1}$: $x^{(\nu)}(t)\mapsto x^{(\nu-1)}(t)$, (which means that J_t^ν can be considered as a bundle over $J_t^{\nu-1}$ too).

The additive group of reals R^+ has a canonical action R on J^ν which is defined as follows: Let $a\in\mathsf{R}$. If x is any curve in M define the curve x_a by $x_a(t)=x(t-a)$. Then $R(a)\colon J_t^\nu \to J_{t+a}^\nu\colon x^{(\nu)}(t)\mapsto x^{(\nu)}(t+a)$. Since this action obviously is free we can identify J_t^ν with J_{t+a}^ν and we have then $\bar{J}^\nu\approx\mathsf{R}\times J^\nu$ where we now write $J^\nu = J_0^\nu$.

Chanbe of l.c.'s in J_t: Consider a ν-jet γ at t and let x be any curve such that $\gamma = x^{(\nu)}(t)$. In a l.c. neighbourhood in M containing the point $x(t)$ of M γ is then entirely determined by the $(\nu+1)n$-tuple $(\gamma^{ik})=(\gamma^{i0},\gamma^{i1},\ldots,\gamma^{i\nu})$ where we have put

$$\gamma^{ik} = \frac{d^k x^i(t)}{dt^k} \qquad (i=1,\ldots,n;\ 0\leqq k\leqq \nu)\,.$$

In a different l.c. neighbourhood we have instead the $(\nu+1)n$-tuple $(\gamma^{i'k}) = (\gamma^{i'0},\gamma^{i'1},\ldots,\gamma^{i'\nu})$ where we have put

$$\gamma^{i'k} = \frac{d^k x^{i'}(t)}{dt^k} \qquad (i'=1,\ldots,n;\ 0\leqq k\leqq \nu)\,.$$

Let the change of l.c.'s (in the overlap of these l.c. neighbourhoods) be mediated by $x^{i'}=\varphi^{i'}(x^i)$ and put $\psi_i^{i'}(x^i)=\partial\varphi^{i'}(x^i)/\partial x^i$. Then we have

$$x^{i'}(t) = \varphi^{i'}(x^i(t)) \qquad (i'=1,\ldots,n)\,.$$

Repeated differentiation gives

$$\frac{dx^{i'}(t)}{dt} = \psi_i^{i'}(x^i(t))\frac{dx^i(t)}{dt}\,,$$

$$\frac{d^2x^{i'}(t)}{dt^2} = \psi_i^{i'}(x^i(t))\frac{d^2x^i(t)}{dt^2} + \frac{d}{dt}[\psi_i^{i'}(x^i(t))]\frac{dx^i(t)}{dt}$$

$$\ldots$$

$$\frac{d^\nu x^{i'}(t)}{dt^\nu} = \sum_{k=0}^{\nu-1}\binom{\nu-1}{k}\frac{d^{\nu-1-k}}{dt^{\nu-1-k}}[\psi_i^{i'}(x^i(t))]\frac{d^k x^i(t)}{dt^k}$$

$(i'=1,\ldots,n)$.

Here the coefficients

$$\frac{d^j}{dt^j}[\psi_i^{i'}(x^i(t))] \qquad (0 \leqq j \leqq \nu-1)$$

again can be expressed in terms of the derivatives $d^k x^i(t)/dt^k$ $(i=1,\ldots,n;\ 0 \leqq k \leqq \nu-1)$. Therefore we get an expression for $(\gamma^{i'k})$ in terms of (γ^{ik}). In particular this proves that J_t^ν is indeed a manifold.

A curve x in M is called *regular* at t if the 1-jet (i.e. the tangent) at t is different from 0. Similarly we say that a ν-jet at t is regular if its image under the canonical map $J_t^\nu \to J_t^{\nu-1}$ is different from 0. The set of regular ν-jets at t is denoted by $J_t^{\nu *}$. It is an open submanifold of J_t^ν. We also agree to write $J^{\nu *} = J_0^{\nu *}$.

2. Osculating bundles.

J^ν is *not* a vector bundle. However there exists a canonical map from J^ν into a certain vector bundle, namely the νth order osculating bundle (definition below).

Let first $T^\nu = T^\nu M$ be the *ν-th order iterated tangent bundle* of M (that is, $T^\nu M = T(T^{\nu-1}M)$ $(\nu>1)$, $T^1 M = TM = T$). If x is any curve in M then the assignment $t \mapsto x'(t)$ defines a curve x' in TM. This gives a canonical map

$$J^\nu(\mathrm{R}, M) \to TJ^{\nu-1}(\mathrm{R}, M)\ .$$

(For $\nu=1$ this can be interpreted as the canonical isomorphism $J'(\mathrm{R}, M) \approx TM$.) If we reiterate this construction we get the sequence

$$J^\nu(\mathrm{R}, M) \to TJ^{\nu-1}(\mathrm{R}, M) \to \ \ldots \ \to T^{\nu-1}J^1(\mathrm{R}, M) \to T^\nu M$$

and thus by composition the map

(2.1) $$J^\nu \to T^\nu\ .$$

Next let $T^{(\nu)} = T^{(\nu)}M$ be the *ν-th order osculating bundle* of M (see Pohl [7], Feldman [2]). An element D of $T^{(\nu)}$ (a νth order tangent vector) over a point x of M is essentially a νth order linear partial differential operator without constant term at x; in terms of l.c.'s:

$$D = \sum_{0<|I|\leqq \nu} D^I(\partial_I)_x \qquad (D^I \in \mathrm{R})$$

where I denotes a multi-index, $I = i_1 \ldots i_k$, and $|I|$ its order, $|I| = k$, and where we have put $\partial_I = \partial_{i_1} \ldots \partial_{i_k}$, $\partial_i = \partial/\partial x^i$. A section of $T^{(\nu)}$ is then nothing but a νth order linear partial differential operator without constant term on M, i.e. in terms of l.c.'s:

$$D = \sum_{0<|I|\leq \nu} D^I \partial_I \quad (D^I \text{ scalar functions in the l.c.'s}) .$$

We have the obvious canonical imbedding

$$T^{(\nu-1)}M \to T^{(\nu)}M .$$

We have also a canonical map

$$TT^{(\nu)}M \to T^{(\nu+1)}M$$

which is defined as follows. A tangent vector $\mathscr{X}$ of $T^{(\nu)}M$ at D is defined by a curve E in $T^{(\nu)}M$ with $D=E(0)$, $\mathscr{X}=E'(0)$. On the other hand each $E(t)$ can be interpreted as an element of $\mathscr{D}'(M)$ (the space of L. Schwartz distributions on M) and in the sense of the topology of that space the limit

$$\frac{dE(0)}{dt} = \lim_{t\to 0} \frac{E(t)-E(0)}{t}$$

too exists, defining a linear partial differential operator of order $\nu+1$ at x, thus an element of $T^{(\nu+1)}M$. In terms of l.c.'s: If

$$E(t) = \sum E^I(t)(\partial_I)_{y(t)}, \qquad D = \sum D^I(\partial_I)_x$$

where y is the curve in M over which E lies, the image of $\mathscr{X}$ is given by

$$\frac{dE(0)}{dt} = \sum XE^I(\partial_I)_x + \sum X^i D^I(\partial_i \partial_I)_x$$

where $X=\sum X^i(\partial_i)_x$ is the tangent of y at x, that is, $X=y'(0)$.

(A slightly different point of view: Let D be a section of $T^{(\nu)}$ and X a vector field on M. Denote by φ_t the flow corresponding to X (in general only locally defined). Then on one hand we have $(d/dt)D_{\varphi_t}|_{t=0}$, a section of $TT^{(\nu)}M$, and on the other hand XD, a $(\nu+1)$th order linear partial differential operator on M.)

We now get the sequence

$$T^\nu M \to T^{\nu-1}T^{(1)}M \to T^{\nu-2}T^{(2)}M \to$$
$$\to \dots \to TT^{(\nu-1)}M \to T^{(\nu)}M$$

and by composition the map

(2.2) $$T^\nu \to T^{(\nu)} .$$

Finally if we compose the maps (2.1) and (2.2) there results the map

(2.3) $$J^\nu \to T^{(\nu)} .$$

Note that the map (2.3) is not injective. However as we shall see below its restriction to $J^{\nu *}$ is everywhere regular (i.e. the indiced map on the tangent space is an isomorphism), thus in particular locally injective.

In what follows we usually without further comment identify a ν-jet with its image in $T^{(\nu)}$ under the map (2.3). As far as we are only concrned in local questions such an identification is certainly permissible.

Let γ be a ν-jet at 0 and pick up a curve x such that $\gamma = x^{(\nu)}(0)$. Then the corresponding νth order tangent vector D obtained by application of (2.3) is defined by $Df \doteq d^\nu f(t)/dt^\nu$.

Also if X is a vector field over M there corresponds to it by this procedure a particular section D of $T^{(\nu)}$, namely $D = X^\nu$. (Through every point x of M there passes an integral curve of X. Therefore a ν-jet is determined at each point x of M and we may apply (2.3).)

The previous claim that the restriction of the map (2.3) to $J^\nu{}^*_\cdot$ is everywhere regular will result from corollary 2.2 below to the following important

LEMMA 2.1. *Let γ be an element of J^ν over the point x of M and pick up a vector field X on M such that γ is the ν-jet at 0 of the integral curve of X through x, so that $X^\nu(x)$ is the image of γ under the map* (2.3). *Then for any other vector field Y on M the partial differential operator $[YX^\nu](x)$ at x which we regard as an element of $T^{(\nu+1)}$, belongs to the image of the map $TJ^{(\nu)} \to TT^{(\nu)}$, induced by* (2.3), *composed with the map $TT^{(\nu)} \to T^{(\nu+1)}$. The same is also true for the operator $YX^\nu(x)$.*

PROOF. Let ψ_t be the flow belonging to Y (i.e. $Y = \psi'_t|_{t=0}$). Set $X_t = \bar{\psi}_t \circ X \circ \psi_t^{-1}$ where $\bar{\psi}_t$ is the flow on TM induced by ψ_t. Then for each t $X_t^\nu(x)$ belongs to the image of J^ν under the map (2.3). Also $(d/dt)X_t^\nu|_{t=0} = [YX^\nu](x)$ (in the sense of distributions). This proves the first assertion of the lemma. The second one we have essentially already established.

COROLLARY 2.1. *Also all the partial differential operators $X^kY(x)$ where $0 \leqq k \leqq \nu$ belong to that image.*

PROOF. In view of the second part of lemma 2.1 it suffices to take $k < \nu$. Apply the first part with Y replaced by ηY, η any scalar function on M. We have the formula

$$
\begin{aligned}
(2.4)\qquad [\eta Y, X^\nu] &= \eta YX^\nu - X^\nu(\eta Y) \\
&= \eta YX^\nu - \sum_{k=0}^{\nu} \binom{\nu}{k} (X^{\nu-k}\eta) X^k Y \\
&= \eta[YX^\nu] - \sum_{k=0}^{\nu-1} \binom{\nu}{k} (X^{\nu-k}\eta) X^k Y .
\end{aligned}
$$

$[\eta Y, X^\nu](x)$ is thus always in the said image. Now choose η such that $X^{\nu-k}\eta(x) \neq 0$, $X^j\eta(x)=0$ if $j \neq \nu-k$. Then $[\eta Y, X^\nu](x)$ is proportional to $X^k Y(x)$ and the desired conclusion follows.

COROLLARY 2.2. *Assume that γ is regular. Then the image contains $(\nu+1)n$ linearly independent elements.*

PROOF. Since γ is regular we may choose l.c.'s such that $X=\partial_1$ $(=\partial/\partial x^1)$. Choose now $Y=\partial_i$ $(=\partial/\partial x^i)$. Then we have the $(\nu+1)n$ linearly independent operators $\partial_1^k \partial_i$ $(0 \leqq k \leqq \nu,\ i=1,\ldots,n)$.

Let G^ν be the group of ν-jets of diffeomorphisms φ of R such that $\varphi(0)=0$. G^ν is a Lie group and operates on the right on J^ν. We identify $T\mathsf{R}$ with $\mathsf{R} \times \mathsf{R}$. The map $\varphi \mapsto |\varphi'(0)|$ then induces a canonical homomorphism $G^\nu \to \mathsf{R}^*$: $g \mapsto |g|$ ("norm"). A scalar function f on J^ν is said to be *homogeneous* (of degree 1) if $f(\gamma g)=|g|\ f(\gamma)$ of all $\gamma \in J^\nu$ and all $g \in G^\nu$. Let X be a vector field on M and consider the corresponding section X^ν of $T^{(\nu)}$. Let λ be any non-vanishing scalar function on M. Then the vector fields X and λX have the same integral curves. I.e. if x is an integral curve of X then $x \circ \varphi$, where φ is a suitable diffeomorphism of R, is an integral curve of λX. It follows that if f is homogeneous then $f((\lambda X)^\nu)=\lambda f(X^\nu)$. Here we may clearly after a passage to the limit drop the assumption that λ is non-vanishing. Notice however that the homogeneity of f entails that f must be *singular* on the set $J^\nu \setminus J^{\nu *}$.

3. Integrals.

Let $\bar{f}$ be a fixed C^∞ function on $\bar{J}^\nu$. Then for every curve x on M and every compact intervall $[\alpha,\beta] \subset \mathsf{R}$, $x_{\alpha\beta}$ denoting the restriction ('arc') of x to $[\alpha,\beta]$, we can define the integral

$$\Phi = \Phi(x_{\alpha\beta}) = \int_\alpha^\beta \bar{f}(x^{(\nu)}(t))\,dt\,. \tag{3.1}$$

A convenient way of defining $\bar{f}$ is to give a one-parameter family f_t of C^∞ functions on J^ν (with C^∞ dependence on t) and use the identification of J_t^ν with $J^\nu=J_0^\nu$ via the action $R(t)$ (see Section 1). We can then put $\bar{f}(\gamma)=f_t((R(t))^{-1}\gamma)$ for $\gamma \in J_t^\nu$ so the integral becomes

$$\Phi = \Phi(x_{\alpha\beta}) = \int_\alpha^\beta f_t((R(t))^{-1}x^\nu(t))\,dt\,, \tag{3.1'}$$

which we abusively write as

$$\Phi = \Phi(x_{\alpha\beta}) = \int_\alpha^\beta f_t(x^{(\nu)}(t))\,dt\,. \tag{3.1''}$$

We distinguish two special cases.

1° The "time-independent" case, f_t is independent of t. Writing $f = f_0$ we thus have the integral

$$\Phi = \Phi(x_{\alpha\beta}) = \int_\alpha^\beta f(x^{(\nu)}(t))\,dt \tag{3.1'''}$$

2° The "parameter-invariant" case. The above f is homogeneous (see Section 1). Let φ be any diffeomorphism of R and consider the curve $y = x \circ \varphi$. We claim that the integral takes the same value if we replace x by y; that is, $\Phi(y_{\gamma\delta}) = \Phi(x_{\alpha\beta})$ where $\alpha = \varphi(\gamma)$, $\beta = \varphi(\delta)$. Indeed we have $y^{(\nu)}(t) = x^{(\nu)}(\varphi(t)) \circ \varphi^{(\nu)}(t)$ where $\varphi^{(\nu)}(t)$ is the ν-jet at t of φ. Therefore by homogeneity

$$f(y^{(\nu)}(t)) = f(x^{(\nu)}(\varphi(t))|\varphi'(t)|$$

and the classical formula for change of variable in an integral proves the point. Since f is singular on $J^\nu \setminus J^{\nu *}$ it is convenient in this case to restrict oneself to regular curves x.

The general case of the integral (3.1) can be formally reduced to the parameter-invariant case, by enlarging the number of variables. Consider the product manifold $\bar{M} = \mathrm{R} \times M$. Then given a one parameter family f_t of functions on $J^\nu = J^\nu M$ uniquely define a homogeneous function $\bar{f}$ on $J^\nu \bar{M}$ by requiring $f(\bar{x}^{(\nu)}(t)) = f_t(x^{(\nu)}(t))$ where x is any curve on M and $\bar{x}$ is the curve on $\bar{M}$ defined by $\bar{x}(t) = (t, x(t))$. Clearly we then have

$$\int_\alpha^\beta \bar{f}(\bar{x}^{(\nu)}(t))\,dt = \int_\alpha^\beta f_t(x^{(\nu)}(t))\,dt\,.$$

The main problem of the Calculus of Variations is to find conditions on a given curve x which secure that $\Phi(x_{\alpha\beta})$ is minimized (or maximized) within a set of arcs close to $x_{\alpha\beta}$ in the sense of a suitable topology.

By a *deformation* of a curve x we mean a one-parameter family of curves x_s (with C^∞ dependence on s) such that $x = x_0$. By the formula $Y = \partial x_s/\partial s|_{s=0}$ is defined a vector field along x. Then *variation* of Φ in the direction of Y is defined by the formula

$$\delta\Phi = s \cdot \frac{d\Phi(x_{s,\alpha\beta})}{ds}\bigg|_{s=0}\,.$$

The total variation is then a linear functional in $\delta x = sY$. A necessary condition for an extremum is this that $\delta\Phi = 0$ for all directions Y. In Section 8 we will work out this condition leading to the Euler equations in an intrinsic form. First we shall however develop the necessary machinery (Sections 4–7, Section 7 may be omitted at first reading).

4. Lie derivatives in general bundles.

Let B be any bundle over M with projection $\pi: B \to M$. Consider the category $\mathscr{C}_M$ whose objects are the non-empty open sets U of M and whose morphisms are diffeomorphisms φ from a non-empty U open subset of M onto a non-empty open subset V of M. Let there be given a functor from $\mathscr{C}_M$ into the corresponding category $\mathscr{C}_B$ formed with B which on the objects induces the mapping $U \mapsto \pi^{-1}(U)$. Assume also that π defines a natural transformation from that functor into the identity functor, meaning that we have the commutative diagram:

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\bar{\varphi}} & \pi^{-1}(V) \\ \pi \downarrow & & \downarrow \\ U & \xrightarrow{\varphi} & V \end{array}$$

$\bar{\varphi}$ is the map induced by φ. Then every vector field X on M induces a vector field $\bar{X}$ on B defined as follows. Let φ_t be the flow associated with X, that is,

$$X = \left.\frac{d\varphi_t}{dt}\right|_{t=0} .$$

(In general $\bar{\varphi}_t$ is only locally defined.) Then we set

$$\bar{X} = \left.\frac{d\bar{\varphi}_t}{dt}\right|_{t=0}$$

where $\bar{\varphi}_t$ is the flow induced by φ_t.

We can now define two particular Lie-derivatives.

First let f be any C^∞ scalar function on B. Then, X being as before a vector field on M, we set $L_X f = \bar{X} f$. This is again a scalar function on M.

Next let a be any C^∞ section of B. For each t $a_t = \bar{\varphi}_t^{-1} \circ a \circ \varphi_t$ is again a section of B. We therefore set

$$L_X a = \frac{d}{dt} a_t|_{t=0} .$$

This is now a section of TB. However in the case when B is a vector bundle we can identify it with a section of B again denoted by $L_X a$. Indeed for every element a of B with $x = \pi(a)$ there is a canonical isomorphism $b \mapsto b_a^V$ (the vertical lift) from the fiber of B over x to the tangent space of B at a, it is formally defined by

$$b_a^V = \left.\frac{d}{dt}(a+tb)\right|_{t=0} .$$

Thus elements of B and vertical tangent vectors of B can be identified. (It follows also that every section b of B can be lifted to a section b^V of TB.)

EXAMPLE 4.1. If $B=T^{(v)}M$ we have $L_X a=[Xa]$ where $[\cdot]$ denotes the commutator (Lie bracket).

We now wish, in the case of a vector bundle B, to relate $L_X f$ and $L_X a$.

First we define the *fiber derivatives* of a scalar function f on B. If a and b are any elements of B belonging to the same fiber we set

$$\dot{f}(b;\, a) = b_a^V f = \frac{d}{dt} f(a+tb)\bigg|_{t=0}$$

We can then announce the following result, the proof of which is immediate from definitions:

LEMMA 4.1. *Assume that B is a vector bundle over M. Let f be a scalar function on B, a a section of B and X a vector field over M. Then holds the formula*

$$L_X f(a) = X(f(a)) - \dot{f}(L_X a;\, a)\,.$$

5. The Lie derivative in the case of jets.

Let f be any C^∞ scalar function on $T^{(v)}$. Then in view of lemma 4.1 and ex. 4.1 holds (with due change of notation)

$$L_Y f(D) = Y(f(D)) - \dot{f}([YD], D) \tag{5.1}$$

where D is any section of $T^{(v)}$ and Y a vector field on M.

We now claim that the same formula (5.1) is also applicable if f is only defined on J^v (locally considered as a submanifold of $T^{(v)}$) and D is of the special form $D=X^v$, X a vector field on M, that is a section of J^v (locally regarded as a subbundle of $T^{(v)}$). To this end we locally extend f to a function $\tilde{f}$ defined on $T^{(v)}$. Then clearly we have $Y\tilde{f}(X^v)=Yf(X^v)$ and furthermore lemma 2.1 tells us that $\dot{\tilde{f}}([YX^v];\, X^v)$ too is independent of the particular choice of the extension. We can therefore unambiguously write $\dot{f}([YX^v];\, X^v)$ in place of $\dot{\tilde{f}}([YX^v];X^v)$. Therefore there results the formula

$$L_Y f(X^v) = Y(f(X^v)) - \dot{f}([YX^v];\, X^v) \tag{5.2}$$

where thus f is any scalar function on J^v and Y and X vector fields on M, X^v being interpreted as a section of J^v.

We wish also to see what happens if we replace Y by ηY, η a scalar function on M. To this end we replace Y by ηY in (5.2) and apply (2.4). We find:

$$L_{\eta Y} f(X^v) = \eta L_Y f(X^v) + \sum_{k=1}^{v} \binom{v}{k} (X_\eta^k) \dot{f}(X^{v-k}Y;\, X) \tag{5.3}$$

where the terms $\dot{f}(X^{v-k}Y;\, X^v)$ are defined in an analogous fashion as $\dot{f}([YX^v]:X^v)$.

6. The Euler derivative.

Let again f be any C^∞ function on J^ν and Y a vector field on M. We define the *Euler derivative* of f with respect to Y by the formula

$$E_Y f(X^\nu) = L_Y f(X^\nu) + \sum_{k=1}^{\nu} (-1)^k \binom{\nu}{k} X^k \dot{f}(X^{\nu-k}; X^\nu) \tag{6.1}$$

(In view of corollary 2.1 all the terms to the right in (6.1) involving $\dot{f}$ are certainly meaningful, with the same interpretation as in Section 6.) If we also introduce the *Noether derivative* of f with respect to Y by the formula

$$N_Y f(X^\nu) = \sum_{k=1}^{\nu} (-1)^{k-1} \binom{\nu}{k} X^{k-1} \dot{f}(X^{\nu-k}Y; X^\nu) \tag{6.2}$$

we can write (6.1) more concisely as

$$E_Y f(X^\nu) = L_Y f(X^\nu) + X(N_Y f(X^{\nu-1})) . \tag{6.1'}$$

Here is what this means for the first few values of ν:

$\nu=1$:
$$E_Y f(X) = L_Y f(X) - X\dot{f}(Y; X);$$
$$N_Y f(X) = \dot{f}(Y; X) \qquad \text{(see [4])}$$

$\nu=2$:
$$E_Y f(X^2) = L_Y f(X^2) - 2X\dot{f}(XY; X^2) + X^2\dot{f}(Y; X^2);$$
$$N_Y f(X) = 2f(XY; X^2) - X\dot{f}(Y; X^2)$$

$\nu=3$:
$$E_Y f(X^3) = L_Y f(X^3) - 3X\dot{f}(X^2Y; X^3) + $$
$$+ 3X^2\dot{f}(XY; X^3) - X^3\dot{f}(Y; X^3);$$
$$N_Y f(X^3) = 3\dot{f}(X^2Y; X^3) - 3X\dot{f}(XY, X^3) + X^2\dot{f}(Y; X^3) .$$

EXAMPLE 6.1. If $f(X^\nu) = Xg(X^{\nu-1})$ where g thus is a function on $J^{\nu-1}$ we have $E_Y f(X^\nu) = 0$. To see this we have to evaluate all the fiber derivatives involved. The method of proof of lemma 2.1 yields at once

$$\dot{f}([YX^\nu]; X^\nu) = [YX]g(X^{\nu-1}) + X\dot{g}([YX^{\nu-1}]; X^{\nu-1}) \tag{6.3}$$

from which again by the device used in the proof of corollary 2.1 readily follows

$$\binom{\nu}{k}\dot{f}(X^{\nu-k}Y; X^\nu) = \binom{\nu-1}{k-1} X\dot{g}(X^{\nu-k-1}Y; X^{\nu-1}) + \tag{6.4}$$
$$+ \binom{\nu-1}{k-2}\dot{g}(X^{\nu-k}Y; X^{\nu-1}) + [\text{if } k=1]\, Yg(X^{\nu-1}) -$$
$$- \dot{g}([YX^{\nu-1}]; X^{\nu-1}) \qquad (1 \leq k \leq \nu) .$$

Inserting (6.3) and (6.4) into (6.1) we readily get

$$\begin{aligned} E_Y f(X^\nu) = {} & YXg(X^{\nu-1}) - [YX]g(X^{\nu-1}) - \\ & - X\dot{g}([YX^{\nu-1}];\, X^{\nu-1}) + \\ & + \sum_{k=1}^{\nu-2} (-1)^k \binom{\nu-1}{k-1} X^{k+1}\dot{g}(X^{\nu-k-1}Y;\, X^{\nu-1}) + \\ & + \sum_{k=2}^{\nu-1} (-1)^k \binom{\nu-1}{k-2} X^k \dot{g}(X^{\nu-k}Y;\, X^{\nu-1}) - \\ & - YXg(X^{\nu-1}) + \dot{g}([YX^{\nu-1}];\, X^{\nu-1}) = 0 \end{aligned}$$

where we used $[YX] = YX - XY$.

We begin our discussion of (6.1) by remarking that the right hand side of (6.1) at a given point x of M only involves the ν-jet of X^ν in the direction of X. Therefore (6.1) really defines a function on $J^{2\nu}$.

Next we investigate what happens if we replace Y by ηY, η a scalar function on M. We claim that we have

$$E_{\eta Y} f(X^\nu) = \eta E_Y f(X^\nu) . \tag{6.5}$$

Since obviously $E_Y f(X^\nu)$ for a fixed X is additive in Y it follows from (6.5) that making Y vary we get a differential form on M. We denote this differential form by $Ef(X^\nu)$ (the *Euler form*). If $Ef(X^\nu) = 0$ (the *Euler equations* in intrinsic form) we say that X is an *extremal field* (with respect to f). We say also that a curve x is an *extremal curve* (with respect to f) if $Ef(X^\nu) = 0$ for some vector field X tangent to x. By the above observation this definition is independent of which vector field we choose. In particular the integral curves of an extremal field are extremal. An extremal field is called *normalized* if $Xf(X^\nu) = 0$. This implies that $f(X^\nu)$ is constant along the integral curves.

In order to prove (6.5) we have to evaluate the second term to the right in (6.1) with Y replaced by ηY. (The first term was already considered in Section 5.) We get

$$\begin{aligned} & \sum_{k=1}^{\nu} (-1)^k \binom{\nu}{k} X^k \dot{f}\left(\sum_{j=0}^{\nu-k} \binom{\nu-k}{j} (X^j\eta) X^{\nu-k-j}Y;\, X^\nu\right) \\ = & \sum_{k=1}^{\nu} \sum_{j=0}^{\nu-k} (-1)^k \binom{\nu}{k}\binom{\nu-k}{j} \sum_{h=0}^{k} \binom{k}{h} (X^{j+k-h}\eta) X^h \dot{f}(X^{\nu-k-j}Y;\, X^\nu) \\ = & \sum_{r=1}^{\nu} \sum_{h=0}^{\nu} \sum_{k=h}^{r} (-1)^k \binom{\nu}{k}\binom{\nu-k}{r-k}\binom{k}{h} (X^{r-h}\eta) X^h \dot{f}(X^{\nu-r}Y;\, X^\nu) . \end{aligned}$$

Here the product of the three binomial coefficients can be rewritten as

$$\frac{1}{h!}\binom{\nu}{r}\binom{r}{k} k(k-1) \ldots (k-(h+1)) .$$

On the other hand by considering the Taylor expansion for $z=1$ of the function $(1-z)^r-1$ we readily get the identity

$$(6.6)\qquad \sum_{k=1}^{r}(-1)^k\binom{r}{k}k(k-1)\dots(k-(h+1)) = \begin{cases} -1 & h=0 \\ 0 & 0<h<r \\ (-1)^h h! & h=r \end{cases}$$

Thus the end result becomes

$$\eta\sum_{h=1}^{\nu}(-1)^h\binom{\nu}{k}X^h\dot{f}(X^{\nu-h}Y;\,X^\nu)-\sum_{r=1}^{\nu}\binom{\nu}{r}(X^r\eta)\dot{f}(X^{\nu-r}Y;\,X^\nu)\,.$$

If we compare this with (5.3) we readily get (6.5).

The transformation properties of $E_Y f(X^\nu)$ when X is varied will be studied in Section 7.

We conclude by the following formula for the Euler derivative of the pth power of f:

$$(6.7)\qquad E_Y f^p(X^\nu) = p(f(X^\nu))^{p-1}E_Y f(X^\nu)$$
$$+p\sum_{k=1}^{\nu}(-1)^k\binom{\nu}{k}\sum_{j=1}^{k}\binom{k}{j}X^j[(f(X^\nu))^{\nu-1}]X^{j-k}\dot{f}(X^{\nu-k}Y;\,X^\nu)\,.$$

The proof is immediate. From (6.7) follows in particular that if X is a normalized extremal field for f then it is also a normalized extremal field for any power of f.

7. The homogeneous case.

In this somewhat technical section we investigate the Euler derivative in the special case when f is homogeneous (see Section 2). This entails that we have

$$(7.1)\qquad f((\lambda X)^\nu) = \lambda f(X^\nu)$$

for every scalar function λ on M and every vector field X on M. f is singular on $J^\nu\setminus J^{\nu *}$. Therefore in order to be safe let us temporarily at least assume that neither λ nor X vanishes.

First we investigate what (7.1) means for the fiber derivative. Take λ of the form $\lambda=1+t\zeta$ where ζ is an arbitrary function on M. This gives the relation

$$(7.2)\qquad f(((1+t\zeta)X)^\nu) = (1+t\zeta)f(X^\nu)\,.$$

We have the expansion

$$((1+t\zeta)X)^\nu = X^\nu+t\sum_{h=0}^{\nu-1}X^h\cdot(\zeta X)\cdot X^{\nu-1-h}+\dots$$

$$= X^\nu + t \sum_{h=0}^{\nu-1} \sum_{j=0}^{\nu} \binom{h}{j} (X^j\zeta) X^{\nu-j} + \dots$$

where the dots ... indicate terms of higher order in t. Therfore differentitation of (7.2) gives

$$\sum_{h=0}^{\nu-1} \sum_{j=0}^{h} \binom{h}{j} (X^j\zeta) \dot{f}(X^{\nu-j}; X^\nu) = \zeta f(X^\nu).$$

Choose now ζ such that $X^{\nu-k}\zeta(x) \neq 0$, $X^j\zeta(x)=0$ $(j \neq \nu - k)$ for a given point x of M. This proves the formula

$$\begin{cases} \nu\dot{f}(X^\nu; X^\nu) = f(X) \\ \dot{f}(X^k; X^\nu) = 0 \qquad (0<k<\nu) \end{cases} \tag{7.3}$$

Next let Y be an auxiliary vector field on M. We claim that we have

$$\dot{f}([Y, (\lambda X)^\nu]; (\lambda X)^\nu) - Y\lambda f(X) = \lambda\dot{f}([YX^\nu]; X^\nu). \tag{7.4}$$

To prove (7.4) we consider the flow ψ_t belonging to Y and apply (7.4) with X replaced by $X_t = \bar{\psi}_t^{-1} \cdot X \cdot \psi_t$, $\bar{\psi}_t$ being the flow in TM induced by ψ_t. Then holds $X_t u = \psi_t^*(X(\psi_t^{-1}{}^* u))$ for any scalar function u on M. It follows that $X_t^\nu u = \psi_t^*(X^\nu(\psi_t^{-1}{}^* u))$ so that differentiation of the right hand side of (7.1) gives effectively the contribution $\lambda\dot{f}([YX^\nu]; X^\nu)$. On the other hand since

$$\lambda X_t = \bar{\psi}_t^{-1} \cdot ((\psi_t^{-1}{}^*\lambda)X) \cdot \psi_t$$

holds for any u

$$(\lambda X_t)^\nu u = \psi_t^*((\psi_t^{-1}{}^*\lambda)X)^\nu \psi_t^{-1}{}^* u.$$

Therefore differentiation of the left hand side gives

$$\dot{f}([Y, (\lambda X)^\nu]; (\lambda X)^\nu) - Y\lambda f(X)$$

where we also used (7.3) applied to λX. This establishes (7.4).

As an application of (7.4) we can now quickly settle the corresponding question for the Lie-derivative. We claim that

$$L_Y f((\lambda X)^\nu) = \lambda f(X^\nu). \tag{7.5}$$

Indeed (5.2) combined with (7.4) gives

$$\begin{aligned} L_Y f((\lambda X)^\nu) &= Y(\lambda f(X^\nu)) - \dot{f}([Y, (\lambda X)^\nu]; (\lambda X^\nu)) \\ &= \lambda Y(f(X^\nu)) + (Y\lambda) f(X^\nu) - \lambda\dot{f}([YX^\nu]; X^\nu) - Y\lambda f(X^\nu) \\ &= \lambda(Y(f(X^\nu)) - \dot{f}([YX^\nu]; X^\nu)) = \lambda L_Y f(X^\nu) \end{aligned}$$

proving (7.5).

To proceed further let us in (7.4) replace Y by ηY, η a scalar function. With (2.4) this gives

$$\eta\dot{f}([Y,(\lambda X)^\nu];(\lambda X)^\nu)-\sum_{k=0}^{\nu-1}\binom{\nu}{k}((\lambda X)^{\nu-k}\eta)\dot{f}((\lambda X)^k Y;\,(\lambda X)^\nu)-$$
$$-\eta Y\lambda f((\lambda X)^\nu) = \lambda\eta\dot{f}([YX^\nu];\,X^\nu)-\sum_{k=0}^{\nu-1}\binom{\nu}{k}\lambda(X^{\nu-k}\eta)\dot{f}(X^kY;\,X^\nu)\,.$$

At this juncture we involve the following combinatorial formula:

$$(\lambda X)^k = \sum_{j=1}^{k}\sum A_{e_1\ldots e_k}(X^{e_1}\lambda)\ldots(X^{e_k}\lambda)X^j \tag{7.6}$$

where the inner summation is extended over all k-sets of integers such that $e_1\geqq 0,\ldots,e_k\geqq 0$, $e_1+\ldots+e_k+j=k$, and the coefficients $A_{e_1\ldots e_k}$ are certain uniquely determined integers which can be found inductively. Applied in our case (7.6) gives

$$\sum_{j=0}^{k}\sum\binom{\nu}{j}A_{e_1\ldots e_{\nu-j}}(X^{e_1}\lambda)\ldots(X^{e_{\nu-j}}\lambda)\dot{f}((\lambda X)^jY,(\lambda X)^\nu) \tag{7.7}$$
$$= \lambda\binom{\nu}{k}\dot{f}(X^kY;\,X^\nu)\qquad(0\leqq k<\nu)$$

where we now sum over all $(\nu-j)$-sets of integers subject to $e_1\geqq 0,\ldots,e_{\nu-j}\geqq 0$, $e_1+\ldots+e_{\nu-j}+j=k$.

We will also need a dual form of (7.6), namely

$$(-1)^k(\lambda X)^k = \sum_{j=0}^{k}\sum(-1)^jA_{e_1\ldots e_k}\lambda X^j(X^{e_1}\lambda)\ldots(X^{e_k}\lambda)\,. \tag{7.8}$$

We are now in a position to prove the main result of this section which expresses thus the homogeneity of the Euler derivative:

PROPOSITION 7.1. *Assume that f is homogeneous. Then holds*

$$E_Yf((\lambda X)^\nu) = \lambda E_Yf(X^\nu) \tag{7.9}$$

where X and Y are vector fields and λ a scalar function on M.

PROOF. In view of (7.5) it suffices to consider the second term to the right in (6.1). Using (7.8) and (7.7) we now obtain

$$\sum_{k=0}^{\nu-1}(-1)^k\binom{\nu}{k}(\lambda X)^k\dot{f}((\lambda X)^{\nu-k}Y;\,(\lambda X)^\nu)$$

$$= \sum_{k=0}^{\nu-1} \sum (-1)^{k-e_1-\dots-e_k} \binom{\nu}{k} A_{e_1\dots e_k} \lambda X^{k-e_1-\dots-e_k}(X^{e_1}\lambda)\dots$$

$$\dots (X^{e_k}\lambda)\lambda^{-1}\dot{f}((\lambda X)^{\nu-k}Y;\ (\lambda X)^{\nu})$$

$$= \lambda \sum_{j=0}^{\nu-1} (-1)^j \binom{\nu}{j} X^j \dot{f}(X^{\nu-j}Y;\ X^{\nu})$$

which is all there is needed.

Open question. Is it possible to establish all this — notably proposition 7.1 — in a more direct and less computational manner? (We remark that (7.5) at least certainly can be obtained directly from the definition of the Lie derivative.)

8. Applications to the Calculus of Variations.

We begin by the time-invariant case (see Section 3). Thus we consider the integral (3.1‴) where f is any function on J^ν and x any curve on M. Then holds the following formula for the variation of Φ in the direction of Y:

$$\delta\Phi = s \int_\alpha^\beta E_Y f(x^{(\nu)})\,dt + s[N_Y f(x^{(\nu)}(t))]_\alpha^\beta\,. \tag{8.1}$$

In order to prove (8.1) we extend x' to a vector field X defined on M and likewise Y to a vector field on M still denoted by Y. (Initially x' and Y were vector fields along X only.) If X is regular, which we may as well assume, at least locally this is possible. We can even achieve that in a suitable l.c. neighbourhood holds $x(t)=(t,0,\dots,0)$, $X=\partial/\partial x^1$, $Y=\partial/\partial x^2$ so that the deformation x_s of x corresponding to Y can be taken to be $x_s(t)=(t,s,0,\dots,0)$. Then holds in particular $[Y,X^\nu]=0$. We therefore get by virtue of (6.1′) and (6.2)

$$\left.\frac{d\Phi(x_{s,\alpha\beta})}{ds}\right|_{s=0} = \int_\alpha^\beta Yf(X^\nu)\,dt$$

$$= \int_\alpha^\beta E_Y f(X^\nu)\,dt + \int_\alpha^\beta XN_Y f(X^\nu)\,dt$$

$$= \int_\alpha^\beta E_Y f(x^{(\nu)}(t))\,dt + [N_Y f(x^{(\nu)}(t))]_\alpha^\beta\,.$$

This establishes (8.2).

The general case, i.e. the integral (3.1″) where f_t is a one parameter family of functions on J^ν, can most simply be treated by reducing it to a homogeneous problem in the product space $\bar{M} = \mathbf{R} \times M$, by the process indicated in Section 3. Then will enter in place of $E_Y f(X^\nu)$ the expression (compare (6.1))

$$L_Y f_t(X^\nu) + \sum_{k=1}^{k} (-1)^k (-1)^k \binom{\nu}{k} \bar{X}^k \dot{f}_t(X^{\nu-k} Y;\ X^\nu)$$

where $\bar{X} = (0, X)$ thus is a vector field on $\bar{M}$. A similarly modified expression will appear in place of $N_Y f(X^\nu)$. We leave the details to the reader, if there is any left.

In what follows we will however concentrate on the time-independent case, with special attention to the parameter-invariant case (i.e. f homogeneous).

As a first consequence of (8.1) we notice that if x is an extremal curve (i.e. we have the *Euler equations* $Ef(X^{(\nu)}) = 0$, see Section 6) then $\delta\Phi = 0$ in all directions Y such that Y vanishes up to order $\nu - 1$ in the direction of x' at the points $x(\alpha)$ and $x(\beta)$. The converse is of course also true but we refrain from entering into the details.

Another immediate consequence of (8.1) which by the way also results directly from (6.1′) – is *Noether's theorem*: Assume that f is invariant for Y in the sense that $L_Y f = 0$. Then $N_Y f(X^{(\nu)}(t))$ is constant along any extremal curve.

We want also to give sufficient conditions for an extremum. Consider an extremal field X for f (i.e. $Ef(X^\nu) = 0$, see Section 6). Trying to imitate a classical procedure (Caratheodory's "Königsweg", see e.g. Rund [9]) we wish to replace our integrand f with another one f^* *equivalent* to f in the sense that the value of the integral $\Phi(x_{\alpha\beta})$ where x is any integral curve of X. Then X is extremal for f^* too. We require furthermore that $f^*((\tilde{X})^\nu) \geqq 0$ for any other vector field $\tilde{X}$ with strict inequality unless $\tilde{X}$ and X have contact up to order ν somewhere. Let us agree to say that f^* is *positive definite* with respect to X. (Also the difference $f - f^*$ should be of some simple type.) If all this is the case it then follows that we have a strong minimum: If $x_{\alpha\beta}$ is an integral curve x of X and $\tilde{x}$ is any other curve with

$$\tilde{x}^{(\nu-1)}(\alpha) = x^{(\nu-1)}(\alpha), \qquad \tilde{x}^{(\nu-1)}(\beta) = x^{(\nu-1)}(\beta)$$

then holds $\Phi(\tilde{x}_{\alpha\beta}) \geqq \Phi(x_{\alpha\beta})$ with equality only if $\tilde{x}_{\alpha\beta} = x_{\alpha\beta}$.

A natural choice for f^* seems to be

$$f^*(\tilde{X}^\nu) = f(\tilde{X}^\nu) - \tilde{X} g(\tilde{X}^{\nu-1}) \tag{8.2}$$

where g is a function on $J^{\nu-1}$. Clearly this f^* is always equivalent to t. Let us see what the positive definiteness of f^* implies. Firstly since $f^*(X^\nu) = 0$ we get

$$f(X^\nu) = X g(X^{\nu-1}) \tag{8.3}$$

(We notice in passing that (8.3) implies that X indeed has to be an extremal, by virtue of example 6.1.) Moreover $f^*(\tilde{X}^\nu)$ attains a minimum for $\tilde{X} = X$ so differentiation shows that (6.3) must hold which again entails (6.4). In the homogeneous case (8.3) is a consequence of (6.3) or (6.4).

In contrast to the classical case $v=1$, these conditions (8.3) and (6.3) or (6.4), do not suffice to determine g uniquely. We are therefore lead to try with the following more special form

$$g(X^{v-1}) = X^{v-1}u_0 + X^{v-2}u_1 + \ldots + u_{v-1} \tag{8.4}$$

where $u_0, u_1, \ldots, u_{v-1}$ are certain functions on M. (6.4) now gives

$$\binom{v}{k}\dot{f}(X^kY;\, X^v) = \sum_{h=0}^{k} \binom{v-h}{k-h} X^{k-h} Y u_h \,. \tag{8.5}$$

This is a recursion for determining Yu_h, and thus g, and leads to

$$Yu_k = \binom{v}{k} P_Y^k f(X^v) \qquad (0 \leqq k < v) \tag{8.6}$$

where we have put

$$P_Y^k f(X^v) = \sum_{h=0}^{k} (-1)^h \binom{k}{h} X^h \dot{f}(X^{k-h}Y;\, X^v) \qquad (0 \leqq k < v) \tag{8.7}$$

(Compare the definition of $N_Y f(X^v)$, formula (6.2)!) It may be easily verified that for any function η on M

$$P_{\eta Y}^k f(X^v) = \eta P_Y^k f(X^v) \qquad (0 \leqq k < v)\,. \tag{8.8}$$

Thus letting vary Y we see that (8.7) defines for fixed X certain differential forms denoted by $P^k f(X^v)$ $(0 \leqq k < v)$. It follows that the integrability conditions for the equations (8.6) take the simple form

$$dP^k f(X^v) = 0 \qquad (0 \leqq k < v)\,. \tag{8.9}$$

Finally we insert the expression for g thus obtained by (8.6) in the formula for f^* (8.2). We find the expression

$$\mathscr{E}(\tilde{X};\, X) = f(\tilde{X}^v) - \sum_{k=0}^{v-1} \binom{v}{k} \tilde{X}^{v-k-1} P_{\tilde{X}}^k f(X^v) \tag{8.10}$$

which may be considered as a generalization of the classical *Weierstrass excess function*, to which it reduces if $v=1$, at least in the homogeneous case: the non-homogeneous case is most conveniently treated by passage to the product manifold $\bar{M} = \mathsf{R} \times M$ (see Section 3).

REMARK 8.1. For $v=2$ a definition of an excess function was suggested by Rund [8] and (according to Rund [9]) a generalization to general v was given by Lister [3].

9. Connections in osculating bundles.

A (linear) *connection* in $T^{(\nu)} = T^{(\nu)}M$ (regarded as a vector bundle over M) determines a linear horisontal lift X^H to $TT^{(\nu)}$ of any vector field X on M such that for each section D of $T^{(\nu)}M$ the difference $XD - X^H(D)$, which thus is a vertical section of $TT^{(\nu)}M$, in addition is linear in D. We can therefore write

$$(\nabla_X D)^V = XD - X^H(D)$$

where thus $\nabla_X D$ is a section of $T^{(\nu)}$ (the *covariant derivative* of D with respect to X), V standing for the vertical lift (see Section 4). We have the standard rules:

$$\nabla_X(D_1+D_2) = \nabla_X D_1 + \nabla_X D_2, \quad \nabla_X(uD) = u\nabla_X D + XuD\,,$$

$$\nabla_{X_1+X_2} D = \nabla_{X_1} D + \nabla_{X_2} D,$$

$$\nabla_{uX} D = u\nabla_X D \qquad (u \text{ a scalar function on } M)\,.$$

If f is any C^∞ scalar function on $T^{(\nu)}$ we set $\nabla_X f = X^H f$. We then have the formula

$$\nabla_X f(D) = X(f(D)) - \dot{f}(\nabla_X D;\, D) \tag{9.1}$$

where D is any section of $T^{(\nu)}$ and X a vector field on M (compare [6] and lemma 4.1).

We say that our connection is *special* if furthermore holds

$$\nabla_X D = XD \tag{9.2}$$

for any section of $T^{(\nu-1)}$ regarded as a subbundle of $T^{(\nu)}$ compare [2]). In what follows we restrict our discussion to special connections.

For a special connection we define its *torsion* by the formula

$$T(Y;\, X^\nu) = \nabla_Y X^\nu - \nabla_X(X^{\nu-1}Y) - [YX^\nu] \tag{9.3}$$

where Y and X are vector fields on M. (If $\nu=1$ (9.3) reduces to the usual definition of torsion of a connection in $T=TM$.)

EXAMPLE 9.1. The connections called dissections by Feldman [2] are special in our sense and moreover fulfill $T(Y;\, X^\nu)=0$. Indeed (9.3) can also be written as

$$T(Y; X^\nu) = K(Y,X)X^{\nu-1} \tag{9.3'}$$

where K is the curvature of our connection. Notice also that (9.2) implies

$$K(YX)X^k = 0 \qquad (0<k<\nu)\,. \tag{9.2'}$$

Let us investigate as usual the transformation properties of $T(Y; X^\nu)$. Let η be any scalar function on M. We claim that

(9.4) $$T(\eta Y; X^{\nu}) = \eta T(Y; X^{\nu})$$

so that making η vary we obtain for a fixed X a form on M.

To prove (9.4) we consider in turn the various terms entering to the right in (9.3). The first one causes no trouble:

(9.5) $$\nabla_{\eta Y} X^{\nu} = \eta \nabla_Y X^{\nu} .$$

Next we notice that

$$X^{\nu-1}(\eta Y) = \sum_{k=0}^{\nu-1} \binom{\nu-1}{k} (X^k \eta) X^{\nu-k-1} Y .$$

Using (9.2) and the Pascal triangle this gives

(9.6) $$\begin{aligned} \nabla_X X^{\nu-1}(\eta Y) &= \sum_{k=0}^{\nu-1} \binom{\nu-1}{k} (X^{k+1}\eta) X^{\nu-k-1} Y + \\ &\quad + \sum_{k=1}^{\nu-1} \binom{\nu-1}{k} X^k \eta X^{\nu-k} Y + \eta \nabla_X (X^{\nu-1} Y) \\ &= \sum_{k=1}^{\nu} \binom{\nu}{k} (X^k \eta) X^{\nu-k} Y + \eta \nabla_X (X^{\nu-1} Y) \end{aligned}$$

(9.5) and (9.6) together with (2.4 readily lead to (9.4). (Another proof of (9.4) follows from (9.3′).)

From (9.2′) and (9.3′) follows likewise that for any scalar function λ on M holds

(9.7) $$T(Y; (\lambda X)^{\nu}) = \lambda^{\nu} T(Y; X^{\nu})$$

but we omit the proof.

We now come to the main result of this section. Let f be a scalar function on J^{ν}. Assume that f can be (locally) extended to a function on $T^{(\nu)}$. We still denote the extension by f and set for a fixed vector field X on M

(9.8) $$\mu(D) = \dot{f}(D; X^{\nu})$$

D being any section of $T^{(\nu)}$. Let Y be an auxiliary vector field on M. Then holds the following rather bizarre formula (cf. [4] for $\nu = 1$)

(9.9) $$E_Y f(X^{\nu}) = \nabla_Y f(X^{\nu}) + \mu(T(Y; X^{\nu})) - \nabla_X^{\nu} \mu(Y) .$$

In order to prove (9.9) we consider the various terms contributing to the right member of (6.1).

First (5.2), (9.8), (9.1) give

(9.10) $$\begin{aligned} L_Y f(X^{\nu}) &= Y(f(X^{\nu})) - \mu([YX^{\nu}]) \\ &= \nabla_Y f(X^{\nu}) + \mu(\nabla_Y X^{\nu} - [Y, X^{\nu}]) . \end{aligned}$$

Next again by (9.1), applied to μ and reiterated, and (9.2)

$$X^k\mu(X^{\nu-k}Y) = \sum_{j=0}^{k} \binom{k}{j} \nabla_X^j \mu(\nabla_X^{k-j}(X^{\nu-k}Y))$$
$$= \sum_{k=1}^{k} \binom{k}{j} \nabla_X^j (X^{\nu-j}Y) + \mu(\nabla_X X^{\nu-1}Y) .$$

Taking the sum we get

$$(9.11) \qquad \sum_{k=1}^{\nu} (-1)^k \binom{\nu}{k} X^k \tilde{f}(X^{\nu-k}Y;\, X^\nu)$$
$$= \sum_{j=1}^{\nu} \sum_{k=j}^{\nu} (-1)^k \binom{\nu}{k}\binom{k}{j} \nabla_X^j \mu(X^{\nu-j}Y) +$$
$$+ \sum_{k=1}^{\nu} (-1)^\nu \binom{\nu}{k} \mu(\nabla_X X^{\nu-1}Y) = \nabla_X^\nu \mu(Y) - \mu(\nabla_X X^{\nu-1}Y)$$

where we use the formula

$$\sum_{k=j}^{\nu} (-1)^k \binom{\nu}{k}\binom{k}{j} = \begin{cases} -1 & j=0 \\ 0 & 0<j<\nu \\ (-1)^\nu & j=\nu \end{cases}$$

which is just (6.4) rewritten. (9.8) now follows from (9.10) and (9.11) if we also take account of (9.4).

Notice that from (9.9) combined with (9.5) gives a new proof of (6.2). In the same way using also (9.7) we can probably get a new proof of (7.9) (homogeneous case).

Open problem. We do not know if there exists a special connection which leaves invariant J^ν (regarded as a submanifold of $T^{(\nu)}$). If this were the case all the terms to the right in (9.9) would have a meaning independent of the extension of f.

Note (nov. 77). In a paper entitled "Further comments on the Euler derivative" (technical report, Lund, 1977: 7), not intended for publication in the present form, some concrete illustrations to problems pertaining to Differential Geometry and Physics are given. The author will send copies on request. I have also become aware that the definition of the Euler derivative given in the present paper, viz. (6.1), is equivalent to the one already used by Tulczyjew (see e.g. Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 23 (1975), 964–969, 24 (1976), 1089–1096). I now plan a long paper where I will generalize all this adopting the point of view of A-points of A. Weil (see Colloques Internationaux du C.N.R.S., Géométrie Différentielle, Strasbourg, 1953).

REFERENCES

1. C. Ehresmann, *Les prolongements d'une variété différentiable*, Atti del quarto Congresso dell'Unione Matematica Italiana tenuto in Taormina nei giorgni 25–31 ottobre 1951, vol. II, Cremonese, Roma, 1953.
2. E. A. Feldman, *The geometry of immersions*. I, Trans. Amer. Math. Soc. 20 (1965), 185–224.
3. T. C. Lister, *The theory of problems in the calculus of variations whose Lagrangians involve higher order derivatives*, Colloquium on the calculus of variations, University of South Africa, Pretoria, 1963.
4. J. Peetre, *A non-linear exterior derivative and the foundations of the Calculus of Variations*, Technical report, Lund, 1976.
5. J. Peetre, *On É. Cartan's approach to Finsler spaces*, Technical report, Lund, 1977.
6. J. Peetre, *Dissections, connections, and sprays*, Technical report, Lund, 1977. Submitted to Math. Ann.
7. W. F. Pohl, *Connections in the differential geometry of higher order*, Topology 1 (1962), 169–211.
8. H. Rund, *The theory of problems in the calculus of variations whose Lagrangian function involves second-order derivatives: a new approach*, Ann. Mat. Pura Appl. 55 (1961), 77–104.
9. H. Rund, *The Hamilton–Jacobi theory in the calculus of variations. Its role in mathematics and physics*, Van Nostrand, London, Toronto, New York, Princeton, 1966.

TEKNISKA HÖGSKOLAN I LUND
FACK
S-22007 LUND 7
SWEDEN