"smooth"; to be precise about this, we regard X as a map of U into $T.(U)$. Since $Xa \in T_a(U)$, we can write X as $a \rightsquigarrow (a, Xa)$.

<u>Definition.</u> A vector field is a smooth function $X\colon U \longrightarrow T.(U)$ such that the composite $\pi \circ X$ is the identity function: $\pi \circ X = 1_U$.

The last requirement means simply that

$$X : u \rightsquigarrow (u, \text{some tangent vector at } u).$$

Recall that the vector space $T^a U$ has the basis $d_a q^1, \ldots, d_a q^n$, where the q^i are coordinates in U. It seems natural to put coordinates on $T^{\cdot}(U)$ so that the point $(a, w) \in T^{\cdot}(U)$ is viewed as $(q^1, \ldots, q^n, h^1 d_a q^1, \ldots, h^n d_q q^n)$, where the q^i and the h^i are the coordinates of a and w, respectively. Thus $w = h^1 d_a q^1 + \ldots + h^n d_a q^n$. In a similar way, we put $2n$ coordinates on the tangent bundle. We have seen that, if f is a function on U, then

$$d_a f = \sum \left(\frac{\partial f}{\partial q^i}\right)_a d_a q^i$$

is a point of $T^a U$. Hence we can write

$$(*) \qquad df = \sum \left(\frac{\partial f}{\partial q^i}\right) dq^i$$

for the function which assigns to every point a of U the cotangent vector $d_a f$; it is thus the cotangent-space analogue of a vector field. Expressions like the above constantly pop up in physics; for example, if the three components of force in space are represented by the functions F_i of x, y, and z, the infinitesimal work is usually defined by

$$dW = F_1 dx + F_2 dy + F_3 dz .$$

When we interpret dW, dx, and so forth, not as infinitesimals, but as cotangent vector fields, this equation resembles equation (*). This will actually be the way we put the notion of "infinitesimal" on a sound mathematical basis.

§4 Vector Bundles

Take a long narrow strip of paper, draw a line down the middle lengthwise, and paste the ends together by twisting once to form a figure called the mobius strip. The line down the middle now becomes a circle, and the surface can be thought of as composed of vectors ("fibers") perpendicular to the circle and radiating from it. We are about to see how the mobius strip can be viewed as part of a vector bundle over the circle.

First we define the simpler notion of pre-bundle. Given two sets U and V, their cartesian product $U \times V$ is defined to be the set of all pairs (u, v) with $u \in U$, $v \in V$. Given sets G, E, U, and maps σ and π forming the following diagram

$$\begin{tikzcd} & E \arrow[d, "\pi"] \\ G \arrow[r, "\sigma"] & U \end{tikzcd}$$

we define the pullback $G \times_U E$ as the universal object making the square

$$\begin{tikzcd} G \times_U E \arrow[r] \arrow[d] & E \arrow[d, "\pi"] \\ G \arrow[r] & U \end{tikzcd}$$

commutative; (universal objects and commutative squares will be defined in due course). Specifically, the pullback can be defined as the following set:

$$G \times_U E = \{(e, g) \mid e \in E,\ g \in G,\ \pi e = \sigma g\}.$$

A pre-bundle, usually written $\begin{array}{c} B \\ \downarrow \pi \\ U \end{array}$, is a pair of sets B, U and a map between them such that B and U are open in some Euclidean space, π is a smooth map, and on each fiber $\pi^{-1}(u) = \{b \mid b \in B, \pi b = u\}$ there is a smooth addition and a scalar multiplication defined so that each fiber forms a vector space. By this we mean that

(a) if $b_1, b_2 \in B$ with $\pi(b_1) = \pi(b_2)$ (i.e., b_1 and b_2 are in the same fiber), and if k is a real number, then there are points $b_1 + b_2$ and $b_1 k$ in B with $\pi(b_1 + b_2) = \pi(b_1 k) = \pi(b_1)$,

(b) the resulting operations on each fiber satisy the vector space axioms, so that each fiber by itself becomes a real vector space,

(c) the maps $+ : B \times_U B \longrightarrow B$ and $\cdot : B \times \mathbb{R} \longrightarrow B$ are smooth.

To state property (c) we must make the additional postulate that $B \times_U B$, the pullback in the diagram

$$\begin{tikzcd} B \times_U B \arrow[r] \arrow[d] & B \arrow[d, "\pi"] \\ B \arrow[r, "\pi"] & U \end{tikzcd}$$

may be embedded in some Euclidean space as an open set. U is usually called the base space, B the total space, and π the projection of B on U. Notice that while U is an open subset of some vector space $\mathbb{R}^k$, U is not itself a vector space. Neither is B; instead, it is made up of a lot of vector spaces glued together like a bundle of sticks.

The easiest example of a pre-bundle is the product space $U \times V$, where U is an open set in $\mathbb{R}^n$ and V is a real vector space, $V = \mathbb{R}^m$ for some m. When we define $\pi(u,v) = u$, we see that the fiber $\pi^{-1}(a) = \{(a,v) \mid v \in V\}$ is just a copy of V, and so inherits the vector-space structure of V. One can check that the other requirements are satisfied. Such a pre-bundle is called a <u>local vector bundle</u>.

Suprisingly enough, the tangent bundle to an open set U is a pre-bundle, (indeed, a local vector bundle) with U the base space, $T.(U)$ the total space, and $\pi: (a,v) \rightsquigarrow v$ the projection. Moreover, there is an isomorphism $T.(U) \xrightarrow{\cong} U \times \mathbb{R}^n$ obtained by mapping $(a,v) \rightsquigarrow (a,(v^1, \ldots, v^n))$, where v has coordinates v^i ; that is,

$$v = v^1\left(\frac{\partial}{\partial q^1}\right)_a + \ldots + v^n\left(\frac{\partial}{\partial q^n}\right)_a .$$

This isomorphism is a bundle map, in the sense that points lying in the fiber over a point a of U are mapped into points in the same fiber. More generally, the pair of smooth maps (H, h) in the diagram

$$\begin{array}{ccc} B & \xrightarrow{H} & \tilde{B} \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \tilde{\pi}} \\ U & \xrightarrow{h} & \tilde{U} \end{array}$$

is called the <u>bundle map</u> if $\tilde{\pi}H = h\pi$ (that is, if the diagram commutes), and if the map H is linear on each fiber. The first condition means that the fiber over a is mapped into the fiber over $h(a)$; the second can be stated

$\pi b_1 = \pi b_2 \Longrightarrow H(b_1 + b_2) = H(b_1) + H(b_2), \quad H(bk) = H(b)k.$

A <u>cross-section</u> of a pre-bundle is a smooth map $\chi : U \longrightarrow B$ such that $\pi\chi = 1_U$; in other words, $\chi(u)$ lies in the fiber over u . For example, we always have the zero section: $\chi(u) = 0$, all $u \in B$. In the case of a local vector bundle, for any constant vector v_o of V the map $u \rightsquigarrow (u, v_o)$ is a cross-section. In general, the set of all cross-sections of a bundle π, denoted by $\Gamma(\pi)$, has an additive structure $[(\chi_1 + \chi_2)(u) = \chi_1(u) + \chi_2(u)]$ as well as right multiplication not just by ordinary scalars but by real-valued smooth functions: if $f \in \mathcal{F}$, $(\chi f)(u) = \chi(u) \cdot f(u)$. Since χf and $\chi_1 + \chi_2$ are also smooth, we find that $\Gamma(\pi)$ forms a module over $\mathcal{F}$.

Now that we know that $T_\cdot(U)$ is a pre-bundle, we can redefine a vector field to be a cross-section of $T_\cdot(U)$. The inevitable dual object, a cross-section of $T^\cdot(U)$, will be called a <u>differential one-form</u>, or a <u>co-vector field.</u> Before considering these creatures, we observe that an ordinary vector field may act on functions much like a directional derivative; that is, given a function and a vector field, we define a new function whose value at every point is the derivative of the original function in the direction given by the vector field at that point. Formally, given $g \in \mathcal{F}$ and $X \in \Gamma(\pi)$, define

$$(Xg)(a) = \langle d_a g, Xa \rangle .$$

It is important not to confuse Xg, which is a function, with the vector field Xg, the product of the cross-section X of $T_\cdot(U)$ with the function g. The context should usually make clear which one is meant.

Since $T^\cdot(U)$ is a pre-bundle in exactly the same way that $T_\cdot(U)$ was, it makes sense to speak of a cross-section of $T^\cdot(U)$, and such a function ω will be called a <u>differential one-form</u> on U. For example, if f is a smooth function taking U into $\mathbb{R}$, the function df defined by $df(a) = (a, d_a f) \in T^\cdot(U)$

is a differential one-form. From the i^{th}-coordinate function q^i, which maps a point of U to its i^{th}-coordinate, we get a form dq^i, and we can express df for a general f in terms of this basis:

$$(1) \qquad df(a) = (q^1(a), \ldots, q^n(a), (\frac{\partial f}{\partial q^1})_a d_a q^1, \ldots, (\frac{\partial f}{\partial q^n})_a \, d_a q^n).$$

The differential one-forms are a module over $\mathfrak{F}$, just as the vector fields were; given a form ω and a function $g \in \mathfrak{F}$, we define $g\omega$ (a new form) by the equation $g\omega(a) = (a, g(a)\omega_a)$ if $\omega(a) = (a, \omega_a)$. Then formula (1) shows that

$$df = \sum (\frac{\partial f}{\partial q^i}) dq^i ,$$

which is now a rigorously based statement of the familiar law of partial differentiation.

It is easy to see that any differential one-form on U can be written as $\omega = \sum_{i=1}^{n} h^i dq^i$, where the h^i are n smooth functions on U. But it is an important fact that in general not every form ω is of the form df for some smooth f. This problem is related to the well-known condition for exactness of a first-order differential equation (cf. Apostol, vol. II, p. 239). Thus we can define a one-form ω to be <u>exact</u> if there is a smooth function f on U such that $\omega = df$. In a simply-connected two-dimensional region, like the interior of a circle, a necessary and sufficient condition for the one-form $\omega = h\,dx + k\,dy$ to be exact is that $\partial h/\partial y = \partial k/\partial x$. In general, however, this is only a necessary condition (cf. Mackey, p. 18, footnote), so we define a new expression, the differential of ω, as

$$d\omega = \sum_{\substack{i,j=1 \\ i<j}}^{n} \left(\frac{\partial h^i}{\partial q^j} - \frac{\partial h^j}{\partial q^i} \right) dq^i \, dq^j \qquad \text{if } \omega = \sum h^i dq^i$$

and we say that ω is closed if $d\omega = 0$. Thus any exact form is closed but not vice versa. The proverbial "alert reader" may have noticed that the above conclusion depends on the definition of exactly what $dq^i dq^j$ means; for the moment, however, we merely point out the analogy between the form of $d\omega$ and that of the "curl" of a vector field in three-space, which allows us to compare and contrast our condition for a form to be exact with the commonplace of physics that a field has a potential function if and only if its curl is zero. (It is worth pointing out that in n-space there are $\frac{n(n-1)}{2}$ different $dq^i dq^j$ for i and j ranging between 1 and n, $i<j$; when $n = 3$ it happens that $\frac{1}{2}n(n-1)$ is also 3. So while in the three-dimensional case the curl of a field is a vector of the same dimension as the field, this need not be true in general.)

We previously considered the effect of a mapping $\varphi: U \longrightarrow \tilde{U}$ on cotangent vectors; for each point a of U, φ induces a linear map $\varphi^*: T^{\varphi(a)}\tilde{U} \longrightarrow T^a U$. It is easy to generalize this to the case of differential forms. Let $\tilde{\omega}$ be a one-form on $\tilde{U}$; we desire to construct a form $\varphi^*\tilde{\omega}$ on U. Given $a \in U$, we have $\varphi(a) \in \tilde{U}$. If $\tilde{\omega}(\varphi(a)) = (\varphi(a), \omega)$, $\omega \in T^{\varphi(a)}(\tilde{U})$, then define $(\varphi^*\tilde{\omega})(a) = (a, \varphi^*(\omega))$ where $\varphi^*(\omega) \in T^a(U)$, so $(a, \varphi^*(\omega)) \in T^{\cdot}(U)$. In a typical physical application, $\tilde{\omega}$ would represent the work for some displacement, and $\varphi^*\omega$ would be the same work function expressed in terms of a different generalized coordinate system, given by the change φ of coordinates from U to $\tilde{U}$.

In the case where $\omega = d\tilde{f}$, where $\tilde{f}$ is a real-valued smooth function on $\tilde{U}$, we can derive

$$\varphi^*(d\tilde{f}) = d(\tilde{f}\circ\varphi) = \sum_{j=1}^{n} \frac{\partial(\tilde{f}\circ\varphi)}{\partial q^j}\, dq^j \;.$$

In particular,

$$\varphi^*(d\tilde{q}^i) = \sum_{j=1}^{m} \frac{\partial(\tilde{q}^i\circ\varphi)}{\partial q^j}\, dq^j \;.$$

Here the coefficients of dq^j form the familiar Jacobian matrix $\dfrac{\partial(\tilde{q}^i\circ\varphi)}{\partial q^j}$.

Moreover, $\varphi^*(\tilde{f}\,\omega) = (\varphi^*\tilde{f})(\varphi^*\omega) = (\tilde{f}\circ\varphi)(\varphi^*\omega)$, where $\varphi^*\tilde{f} = \tilde{f}\circ\varphi$; and when we have defined the operator d on forms it will turn out that $d(\varphi^*\omega) = \varphi^*(d\omega)$. These equations can be interpreted as properties of the map $\varphi^*: \Gamma(T^{\cdot}\tilde{U}) \longrightarrow \Gamma(T^{\cdot}U)$, where $\Gamma(T^{\cdot}(U))$ is the $\mathcal{F}$-module of all differential one-forms on U. . Finally, if we have the diagram of maps $U_1 \xrightarrow{\varphi} U_2 \xrightarrow{\psi} U_3$ then $\Gamma(T^{\cdot}U_3) \xrightarrow{\psi^*} \Gamma(T^{\cdot}U_2) \xrightarrow{\varphi^*} \Gamma(T^{\cdot}U_1)$ and we have $(\psi\varphi)^*\omega = \varphi^*(\psi^*(\omega))$. We say that $(\;)^*$ is a "contravariant functor".

We can pull-back not only differential forms but also whole pre-bundles along the map φ. Specifically, given the situation
$$\begin{array}{ccc} & & \tilde{B} \\ & & \downarrow{\scriptstyle\pi} \\ U & \xrightarrow{\varphi} & \tilde{U} \end{array}$$
where $\tilde{B}$ is a pre-bundle, construct a new set

$$B = \{(a,\tilde{b}) \mid a \in U,\ \tilde{b}\in\tilde{B}\ ,\ \text{and}\ \varphi a = \pi\tilde{b}\}\ .$$

(This set B is exactly the "pullback" $U \times_{\tilde{U}} \tilde{B}$ discussed earlier.) If we define $\pi(a,\tilde{b}) = a \in U$, then it can be shown that B is a pre-bundle. Intuitively speaking, given any point a of U we have found the fiber over the image of a under φ and made that fiber the fiber over a in our new bundle. Notice that if $\tilde{B}$ is a tangent bundle then its pullback need not be; for example, if $\tilde{U}$ is some open set in the plane, and U is a line lying in $\tilde{U}$, the pullback B will have a two-dimensional vector space over each point of U, whereas the tangent bundle to a line is always one-dimensional.

§ 5. The Lagrange Equations

(As a reference for the following discussion of the Lagrange equations, cf. Goldstein, pp. 10-18, and Whittaker, pp. 30-35.)

We consider a system of N particles with coordinates $x^1, \ldots, x^{3N}$, moving subject to contraints, in such a way that the system can be described by n generalized coordinates $q^1, \ldots, q^n$. For example, the position of a sphere rolling on a plane is completely determined by x and y, the coordinates of the center of the sphere, and the three Eulerian angles θ, ψ, φ which determine how much the sphere rotates. We make the assumption that the constraints are holonomic; that is, we can make small displacements in each of the q^j independently. Given applied forces $F_i^{(a)}$ and constraint forces $F_i^{(c)}$ in each of the euclidean directions x^i, $i = 1, \ldots, 3N$, and under the assumption that the forces of constraint do no work, we will show that the system obeys the Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^j}\right) - \frac{\partial \mathcal{J}}{\partial q^j} = Q_j \quad , \quad j = 1, \ldots, n$$

where $\mathcal{J} = \mathcal{J}(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n)$ is the kinetic-energy function, and Q_j is called the "generalized force in direction j" as defined below.

Newton's laws tell us exactly how the system behaves when viewed in the framework of euclidean space, $\mathbb{R}^{3N}$; our problem is to describe it in the configuration space U whose coordinates are the q^j. Since we have an evident map $\varphi: U \longrightarrow \mathbb{R}^{3N}$ which merely changes the q^j-description of a state of the system to the x^i-coordinates of $\mathbb{R}^{3N}$, a natural step is to try to "pull back" the laws of motion in $\mathbb{R}^{3N}$ along φ. For example, if

$$W = \sum_i^{3N} (F_i^{(a)} + F_i^{(c)})\, dx^i$$

is the differential form corresponding to work in euclidean space, then $\varphi^* W$ gives the work in terms of the q^j. This is so because the change of the state of the system in time is described by a path in $\mathbb{R}^{3N}$; but since we consider only motion subject to the constraints, each such path is the image under φ of a path in U. Now the "self-adjointness" rule given above for pullbacks makes it clear that $W^U = \varphi^* W$ has the same effect on a path of U as does W on the corresponding path in $\mathbb{R}^{3N}$. Computing,

$$W^U = \sum_{i=1}^{3N} \varphi^*(F_i^{(a)} + F_i^{(c)}) \sum_{j=1}^{n} \frac{\partial x^i}{\partial q^j}\, dq^j$$

$$= \sum_{j=1}^{n} \Big(\sum_i F_i^{(a)} \frac{\partial x^i}{\partial q^j} \Big) dq^j + \sum_{j=1}^{n} \Big(\sum_i F_i^{(c)} \frac{\partial x^i}{\partial q^j} \Big) dq^j \quad ,$$

where we have identified F and $\varphi^* F = F \circ \varphi$ (that is, the force F as a function of the x^i and F viewed as a function of the q^i) as is the custom in physics. Since we have assumed that the $F_i^{(c)}$ do no work, we get

$$W^u = \sum_{j=1}^{n} Q_j\, dq^j \qquad \text{where} \quad Q_j = \sum_{i=1}^{3N} F_i^{(a)} \frac{\partial x^i}{\partial q^j}$$

are the generalized forces; "generalized" since they need not have the dimensions of force, although the product $Q_j q^j$ always has the dimensions of work.

The situation is now as follows

$$\begin{tikzcd}
T.(U) \arrow[r, "\Phi"] \arrow[d] & T.(\mathbb{R}^{3N}) \arrow[rr, "\mathcal{J}"] \arrow[d] \arrow[dr, shift left] \arrow[dr, shift right] & & \mathbb{R} \\
U \arrow[r, "\varphi"] \arrow[dr, shift left, "q^n"] \arrow[dr, shift right, "q^1"'] & \mathbb{R}^{3N} \arrow[dr, shift left, "x^{3N}"] \arrow[dr, shift right, "x^1"'] & (x^1, \dots, x^{3N}, \dot{x}^1, \dots \dot{x}^{3N}) & \\
 & & &
\end{tikzcd}$$

where the map Φ is defined by $\Phi(a, v) = (\varphi a, \varphi^* v)$. As above, we will follow the convention of writing $\dfrac{\partial \mathcal{J}}{\partial q^i}$ when we really mean $\dfrac{\partial(\mathcal{J} \circ \Phi)}{\partial q^i}$.

Then we have the following straightforward derivation of the Lagrange equations, starting from Newton's second law:

$$m_i \frac{d^2 x^i}{dt^2} = F_i^{(a)} + F_i^{(c)} \quad , \quad i = 1, \dots, 3N \,;$$

therefore

$$(1) \quad \sum_{j=1}^{n} \Big(\sum_{i=1}^{3N} m_i \frac{d^2 x^i}{dt^2} \frac{\partial x^i}{\partial q^j} \Big) dq^j = \sum_{i=1}^{3N} (F_i^{(a)} + F_i^{(c)}) \sum_{j=1}^{n} \frac{\partial x^i}{\partial q^j} dq^j = \sum_{j=1}^{n} Q_j dq^j .$$

Now $\dot{x} = \dfrac{dx}{dt} = \displaystyle\sum_{j=1}^{n} \dfrac{\partial x}{\partial q^j} \dot{q}^j$. Therefore $\dfrac{\partial \dot{x}}{\partial \dot{q}^j} = \dfrac{\partial x}{\partial q^j}$. Hence

$$\frac{d^2 x}{dt^2} \frac{\partial x}{\partial q^j} = \frac{d}{dt}(\dot{x}) \frac{\partial x}{\partial q^j} = \frac{d}{dt}\Big(\dot{x} \frac{\partial x}{\partial q^j}\Big) - \dot{x} \frac{d}{dt}\Big(\frac{\partial x}{\partial q^j}\Big) = \frac{d}{dt}\Big(\dot{x} \frac{\partial \dot{x}}{\partial \dot{q}^j}\Big) - \dot{x} \frac{\partial}{\partial q^j}(\dot{x})$$

$$= \frac{d}{dt}\Big(\frac{\partial}{\partial \dot{q}^j} \frac{\dot{x}^2}{2}\Big) - \frac{\partial}{\partial q^j}\Big(\frac{\dot{x}^2}{2}\Big) .$$

$$\text{By (1)}, \quad \sum_{j=1}^{n} Q_j dq^j = \sum_{j=1}^{n} \left[\frac{d}{dt}\frac{\partial}{\partial \dot{q}^i} \sum_i \frac{m_i \dot{x}_i^2}{2} - \frac{\partial}{\partial q^j} \sum_i \frac{m_i \dot{x}_i^2}{2}\right] dq^j$$

$$= \sum_{j=1}^{n} \left(\frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^j}\right) - \frac{\partial \mathcal{J}}{\partial q^j}\right) dq^j .$$

Since the dq^j are independent, we can equate the individual coefficients and get the desired Lagrange equations.

What is really going on here? A first clue is that our formula $\frac{dx}{dt} = \sum \frac{\partial x}{\partial q^j}\dot{q}^j$ is just a special case of the formula in coordinates for $T_a(\varphi)$, the map induced by φ on the tangent bundle. We have really "lifted" the path $c: I \longrightarrow U$, which describes the change of the system in time, to a path $\tilde{c}: I \longrightarrow T.(U)$, where $\tilde{c}(t) = (c(t), \dot{c}(t))$, and examined its form under the map Φ. More pertinent is the reason why we have organized our equations of motion in this form to begin with. Examining the derivation above shows that we expressed Newton's second law, $m_i \frac{d^2x^i}{dt^2} = F^i$, in terms of $\mathcal{J} = \sum \frac{1}{2} m_i (\dot{x}^i)^2$, getting $\frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{x}^i}\right) = \frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{x}^i}\right) - \frac{\partial \mathcal{J}}{\partial x^i} = F_i$, $i = 1, \ldots, 3N$, the Lagrange equations in euclidean space; we then pulled back by φ to get a system of equations <u>in the same form</u> with respect to the q^i. What matters here is that the Lagrange equations in fact remain invariant under pullback by <u>any</u> 1-1 map φ. This will be done in the next section.

§6. <u>Lifting Paths</u>

Let W be a differentiable manifold, and $c: I \longrightarrow W$ be a path in W. We define the <u>lifted path</u> $\tilde{c}: I \longrightarrow T.(W)$ in the tangent bundle $T.(W)$ by

$\tilde{c}(t) = (c(t), \tau_{c(t)}c)$. Let $\varphi: W \longrightarrow U$. Then $(T.\varphi)(\tilde{c}) = (\widetilde{\varphi c})$, since by definition $\widetilde{\varphi c}(t) = (\varphi c(t), \tau_{\varphi ct}(\varphi c))$ and $(T.\varphi)(\tilde{c}(t)) = (\varphi c(t), T_{ct}\varphi(\tau_{ct}c))$ by definition of $T.\varphi$. We indicate the relationships of the above functions by a commutative diagram

$$\begin{tikzcd}
& T.W \arrow[r, "T.\varphi"] \arrow[d, "\pi"] & T.U \arrow[d, "\pi"] \\
I \arrow[ur, "\tilde{c}"] \arrow[r, "c"'] \arrow[urr, bend left, "\widetilde{\varphi c}"] & W \arrow[r, "\varphi"'] & U
\end{tikzcd}$$

If we choose coordinates $r^1, \ldots, r^m: W \longrightarrow \mathbb{R}$ for W and $q^1, \ldots, q^n: U \longrightarrow \mathbb{R}$ for U, we have coordinates $\dot{r}^1, \ldots, \dot{r}^m; r^1\pi, \ldots, r^m\pi: T.W \longrightarrow \mathbb{R}$ for the tangent bundle of W and $\dot{q}^1, \ldots, \dot{q}^n; q^1\pi, \ldots, q^n\pi: T.U \longrightarrow \mathbb{R}$ as coordinates for $T.U$.

The Jacobian $J = J(\varphi)$ of φ is the $n \times m$ matrix defined by $J^i_j = \dfrac{\partial q^i\varphi}{\partial r^j}$ $\left(= \dfrac{\partial q^i}{\partial r^j}\right.$ for short), where $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The Jacobian of $T.\varphi$ is defined similarly. It is a $2n \times 2m$ matrix of the form shown below

$$\begin{array}{c|c|c|}
 & r^1 \ldots r^m & \dot{r}^1 \ldots \dot{r}^m \\ \hline
q^1 \ldots q^n & J(\varphi) & 0 \\ \hline
\dot{q}^1 \ldots \dot{q}^n & ? & J(\varphi) \\ \hline
\end{array}$$

The lower right block can be calculated as follows

$$\frac{dq^i}{dt} = \sum_{j=1}^{m} \frac{\partial q^i}{\partial r^j} \frac{dr^j}{dt} \; ; \quad \dot{q}^i = \sum_{j=1}^{m} \frac{\partial q^i}{\partial r^j} \dot{r}^j \, .$$

Therefore

$$\frac{\partial \dot{q}^i}{\partial \dot{r}^j} = \frac{\partial q^i}{\partial r^j} \quad \text{for all } i \text{ and } j \, .$$

Thus the equations for the transformation $T.\varphi$ in terms of coordinates are

$$\dot{q}^i = \sum_{j=1}^{m} \frac{\partial q^i}{\partial r^j} \dot{r}^j \quad , \quad q^i = \varphi(r^1, \ldots, r^m).$$

The invariant description is $T.\varphi(b, \tau_b c) = (\varphi b, \tau_{\varphi b}(\varphi c))$.

Given U, ω a 1-form on U, and $\mathcal{J}: T.U \to \mathbb{R}$ a smooth function, we define: a path c in U satisfies Lagrange's equation (with respect to $\mathcal{J}$ and ω) in the coordinates $q^1, \ldots, q^n$ of U where

$$\omega = \sum_{j=1}^{n} Q_j dq^j \quad , \quad \mathcal{J} = \mathcal{J}(q^1, \ldots, q^n; \dot{q}^1, \ldots, \dot{q}^n)$$

if

(1) $$\frac{d}{dt}\left[\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right)(\tilde{c})\right] - \left(\frac{\partial \mathcal{J}}{\partial q^i}\right)(\tilde{c}) = Q_i c \quad , \quad i = 1, \ldots, n .$$

The functions are indicated in the diagram:

$$\begin{tikzcd}
 & T.U \arrow[r, "\mathcal{J}"] \arrow[d] & \mathbb{R} \\
I \arrow[ur, "\tilde{c}"] \arrow[r, "c"] & U \arrow[r, "Q_i"] & \mathbb{R}
\end{tikzcd}$$

As a special case, suppose the forces are conservative. By definition, the 1-form ω (the work) is conservative if and only if there exists a smooth function $\mathcal{V}: U \to \mathbb{R}$ with $\omega = -d\mathcal{V}$, in coordinates

$\sum_{i=1}^{n} Q_i dq^i = - \sum_{i=1}^{n} \frac{\partial \mathcal{V}}{\partial q^i} dq^i$ so $Q_i = \frac{-\partial \mathcal{V}}{\partial q^i}$. Thus (1) becomes

$\frac{d}{dt}\left[\left(\frac{\partial L}{\partial \dot{q}^i}\right)\tilde{c}\right] - \left(\frac{\partial L}{\partial q^i}\right)\tilde{c} = 0$, where L is defined by

$$L = (\mathcal{J} - \mathcal{V} \circ \pi): T.U \to \mathbb{R}$$

We now state the theorem asserting that Lagrange's equation can be "pulled back" along a smooth map φ.

Theorem. Let $\varphi: W \longrightarrow U$ be a smooth map between open sets in euclidean spaces, while ω is a 1-form on U and $\mathcal{J}$ a smooth function on $T.U$, as above. If $c: I \longrightarrow W$ is a path in W such that the composite path φc satisfies Lagrange's equation (with respect to $\mathcal{J}$ and ω), then the path c in W satisfies Lagrange's equation (with respect to $\mathcal{J}(T.\varphi)$ and $\varphi^* \omega$).

Proof. If $\omega = \sum_i Q_i dq^i$, for Q_i smooth functions on U, the form $\varphi^* \omega$ is given in the coordinates r^j of W as

$$\varphi^* \omega = \sum_j \sum_{i=1}^{n} Q_i \frac{\partial q^i}{\partial r^j} dr^j .$$

This can be written as $\sum_j R_j dr^j$, where each

$$R_j = \sum_{i=1}^{n} Q_i \frac{\partial q^i}{\partial r^j} , \quad j = 1, \ldots, m .$$

We are given, on the path $\varphi \circ c$, the equations

$$\frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right) - \frac{\partial \mathcal{J}}{\partial q^i} = Q_i , \quad i = 1, \ldots, n,$$

(here $\mathcal{J}$ is short for $T \circ \varphi \circ c$, Q_i for $Q_i \circ \varphi \circ c$). Multiply the i^{th} equation by $\dfrac{\partial q^i}{\partial r^j}$ and add over i to get

$$\sum_i \frac{d}{dt}\left(\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\right) \frac{\partial q^i}{\partial r^j} - \sum_i \frac{\partial \mathcal{J}}{\partial q^i} \frac{\partial q^i}{\partial r^j} = \sum_i Q_i \frac{\partial q^i}{\partial r^j} = R_j$$

(here again everything is, through c, a function of t) By the rule for differentiating the product $\dfrac{\partial \mathcal{J}}{\partial \dot{q}^i} \dfrac{\partial q^i}{\partial r^j}$, this can be rewritten as

$$\sum_i \frac{d}{dt}\left[\frac{\partial \mathcal{J}}{\partial \dot{q}^i}\frac{\partial q^i}{\partial r^j}\right] - \sum_i \frac{\partial \mathcal{J}}{\partial \dot{q}^i}\frac{d}{dt}\left(\frac{\partial q^i}{\partial r^j}\right) - \sum_i \frac{\partial \mathcal{J}}{\partial q^i}\frac{\partial q^i}{\partial r^j} = R_j .$$

The second and third terms combine to give $\dfrac{\partial \mathcal{J}\varphi}{\partial r^j}$; as for the first term,

$$\frac{\partial \mathcal{J}\varphi}{\partial \dot{r}^j} = \sum_i \frac{\partial \mathcal{J}}{\partial \dot{q}^i}\frac{\partial \dot{q}^i}{\partial \dot{r}^j} + \sum_i \frac{\partial \mathcal{J}}{\partial q^i}\frac{\partial q^i}{\partial \dot{r}^j} ,$$

(here $\dfrac{\partial q^i}{\partial \dot{r}^j} = 0$ and $\dfrac{\partial \dot{q}^i}{\partial \dot{r}^j} = \dfrac{\partial q^i}{\partial r^j}$, as noted above). Hence the whole equation (really with the path c substituted in) becomes

$$\frac{d}{dt}\left[\frac{\partial \mathcal{J}\varphi}{\partial \dot{r}^j}\right] - \frac{\partial \mathcal{J}\varphi}{\partial r^j} = R_j$$

and this is Lagrange's equation on W. q.e.d.

As a corollary we can prove the invariance of Lagrange's equation under a change of coordinates (i.e., under a smooth map φ with a two-sided inverse).

Corollary. If $\varphi: W \longrightarrow U$ is a smooth map with a smooth inverse $\varphi^{-1}: U \longrightarrow W$, while $c: I \longrightarrow W$ is a path, then $\varphi \circ c$ satisfies Lagrange's equation with respect to a function L and a form ω if and only if c satisfies Lagrange's equation with respect to $(T.\varphi)^* L$ and $\varphi^*\omega$.

This invariance under any change of coordinates is the advantage of Lagrange's equation over the original Newtonian equations. In the next section we give another explanation of this invariance.

The conservative case is that in which the 1-form ω on U which represents the work is the differential

$$\omega = d\mathcal{V}$$

of a smooth function $\mathcal{V} : U \longrightarrow \mathbb{R}$ called the potential. The function $\mathcal{V}$ can be lifted to the tangent bundle as $\mathcal{V} \circ \pi : T.U \longrightarrow \mathbb{R}$, and we often write $\mathcal{V}$ for $\mathcal{V} \circ \pi$. Then the Lagrangian function L is by definition

$$L = \mathcal{J} - \mathcal{V} : T.U \longrightarrow \mathbb{R}$$

and Lagrange's equations clearly take the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = 0 \;, \; i = 1, \ldots, n.$$

Here L is of course short for $L \circ \tilde{c}$, c a path. Expanding the derivative, these equations are

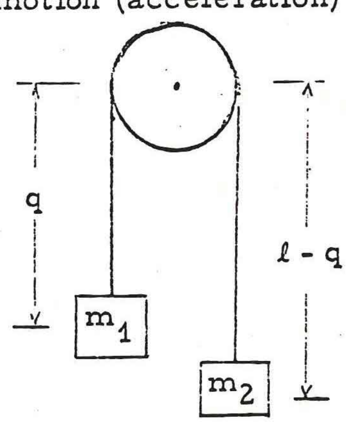
$$\sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}^i \partial \dot{q}^j} \ddot{q}^j + \sum_{j=1}^{n} \frac{\partial^2 L}{\partial \dot{q}^i \partial q^j} \dot{q}^j - \frac{\partial L}{\partial q^i} = 0,$$

a system of n second order differential equations.

We give an application of Lagrange's equation. Recall $L = (\mathcal{J} - \mathcal{V}) : T.U \longrightarrow \mathbb{R}$ and Lagrange's equation is

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}^i}\right] = \frac{\partial L}{\partial q^i} \qquad i = 1, \ldots, n.$$

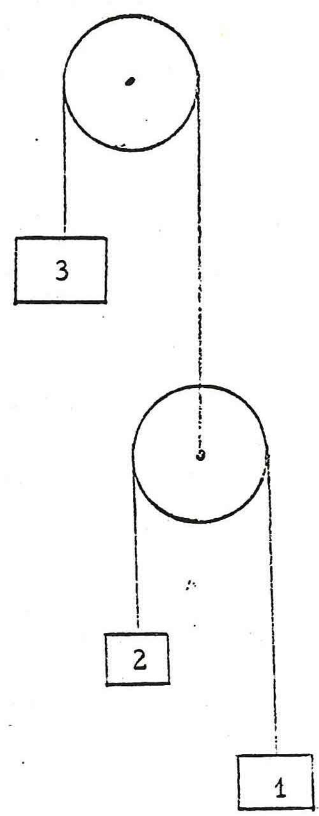
We will use this to derive the equation of motion (acceleration) for Atwood's machine:



a (massless) pulley supporting a (massless) chain with a weight m_1 a distance q below the center of the pulley and a weight m_2 a distance $(\ell - q)$ below the center of the pulley, where ℓ is fixed. As indicated, there is just one generalized coordinate q.

We calculate $\mathcal{J} = \frac{1}{2} m_1\dot{q}^2 + \frac{1}{2} m_2\dot{q}^2$ and $\mathcal{V} = -m_1gq - m_2g(\ell - q) =$ $(m_2 - m_1)gq +$ constant, therefore $L = \frac{1}{2}(m_1 + m_2)\dot{q}^2 - (m_2 - m_1)gq +$ constant, We differentiate $\frac{\partial L}{\partial \dot{q}} = (m_1 + m_2)\dot{q}$, $\frac{d}{dt}[\frac{\partial L}{\partial \dot{q}}] = (m_1 + m_2)\ddot{q}$ and $\frac{\partial L}{\partial q} = -(m_2 - m_1)g$. We equate $(m_1 + m_2)\ddot{q} = (m_1 - m_2)g$ and solve to get $\ddot{q} = ((m_1 - m_2)/(m_1 + m_2))g$, a second order differential equation which can readily be integrated.

We give a problem for mathematicians: Find the accelerations of the system shown below. The mass of the weights is indicated by the numbers labeling them.



§7 Hamilton's Principle.

The fact that a path $c: I \longrightarrow U$ satisfies Lagrange's equation is, we have seen, independent of the choice of coordinates in the configuration space U. This fact can also be explained by Hamilton's principle, which asserts that the solutions of Lagrange's equations are exactly the curves which "minimize" a certain integral formed from the Lagrangian function L. To cover the most general case, we will assume that L depends not only on position and velocity but also on time; that is, that L is a smooth function $L: T.U \times I \longrightarrow \mathbb{R}$. In coordinates, this means that $L(q^1, \ldots, q^n, \dot{q}^1, \ldots, \dot{q}^n, t) \in \mathbb{R}$.

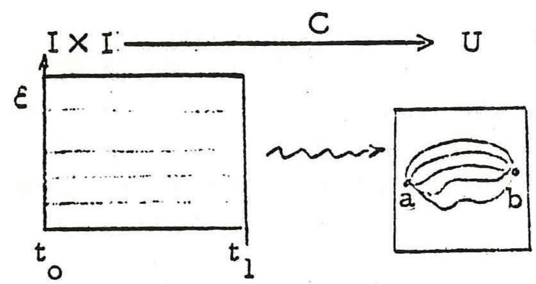
Given fixed points a and b in U, we consider paths c_o from a to b; that is, $c_o: I \longrightarrow U$ with $t_o, t_1 \in I$ and $c_o(t_o) = a$, $c_o(t_1) = b$. We can lift c_o to a path $\tilde{c}_o: I \longrightarrow T.U \times I$, the identity on I. We want to compare c_o with other smooth paths with the same endpoints a and b at the same time t_o and t_1. For any path c, let $J(c) = \int_{t_o}^{t_1} (L\tilde{c})dt$. We want to prove

<u>Hamilton's Principle:</u> If $c_o: I \longrightarrow U$ is a path from a to b in U, then c_o satisfies Lagranges's equations for L if and only if the corresponding integral $J(c)$ is "stationary" for $c = c_o$.

It remains to explain "stationary" for functions of paths like J. In the simpler case of a function f of a real number ε, we say that f is stationary at ε_o if $\frac{df}{d\varepsilon} = 0$ for $\varepsilon = \varepsilon_o$. Thus a function is stationary at a maximum, at a minimum, and at other points (horizontal inflections). Similarly, $J(c_o)$ will be stationary for c_o if c_o is a minimum $(J(c_o) \leq J(c))$ or a maximum or .... More exactly, J is stationary if the corresponding function of ε is stationary at c_o whenever c_o is embedded in a one-parameter family of paths. Such a family is described by a smooth function

$$C : I \times I \longrightarrow U,$$
$$(\varepsilon, t) \rightsquigarrow C(\varepsilon, t), \qquad \varepsilon \in I, \ t \in I$$

where $C(\varepsilon, t_o) = a$, $C(\varepsilon, t_1) = b$, and $C(0, t) = c_o(t)$. In other words, for each ε, $C(\varepsilon, -)$ is a path from a to b, while for $\varepsilon = 0$, $C(0, -)$ is the given path c_o. The situation is that of the following figure



By lifting each path, we get $\tilde{C} : I \times I \longrightarrow T.U \times I$. Then $J(\tilde{C}) = \int_{t_o}^{t_1} (L\tilde{C})dt$, is a function of ε. Calculate

$$\frac{dJ(\tilde{C})}{d} = \int_{t_o}^{t_1} \frac{\partial}{\partial \varepsilon}(L\tilde{C})\, dt = \int_{t_o}^{t_1} \left(\sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}^i} \frac{\partial \dot{q}^i}{\partial \varepsilon} + \sum_{i=1}^{n} \frac{\partial L}{\partial q^i} \frac{\partial q^i}{\partial \varepsilon}\right) dt.$$

Note that $\frac{\partial \dot{q}^i}{\partial \varepsilon} = \frac{d}{dt}\left(\frac{\partial q^i}{\partial \varepsilon}\right)$ and the integration-by-parts formula

$$\int_{t_o}^{t_1} U \frac{dV}{dt}\, dt = -\int_{t_o}^{t_1} \frac{dU}{dt} V\, dt + UV \Big|_{t_o}^{t_1} .$$

Therefore, $\frac{\partial J(\tilde{c})}{\partial \varepsilon} = \int \left[\sum [-\frac{d}{dt} (\frac{\partial L}{\partial \dot{q}^i})](\frac{dq^i}{d\varepsilon}) + \frac{\partial L}{\partial q^i} \frac{\partial q^i}{\partial \varepsilon} \right] dt + \underbrace{\sum \frac{dL}{\partial \dot{q}^i} \frac{\partial q^i}{\partial \varepsilon} \Big|_{t_o}^{t_1}}_{=0}$

(the last part equals zero because the paths have the same endpoints)

$$= \sum_{i=1}^{n} \int [-\frac{d}{dt} (\frac{\partial L}{\partial \dot{q}^i}) + \frac{\partial L}{\partial q^i}](\frac{\partial q^i}{\partial \varepsilon})\, dt .$$

Therefore, if the Lagrange equations hold along c_o, then $\frac{dJ(\tilde{c})}{d\varepsilon} = 0$.

Conversely, we show for a path c_o that if $J(\tilde{C})$ is stationary for every one parameter family C containing c_o as above, then c_o must satisfy Lagrange's equations. The proof will use the "variation" of the path given by a smooth $\eta : I \longrightarrow \mathbb{R}^n$ with $\eta(t_o) = \eta(t_1) = 0$. The "varied" path is the one parameter family C with

$$C(\varepsilon, t) = c_o(t) + \varepsilon\eta(t)$$

(assume that $\eta(t)$ is small, so that $C(\varepsilon, t)$ still lies in the open set U of $\mathbb{R}^n$). Then for the coordinates we have

$$q^i C(\varepsilon, t) = q^i \circ c_o + \varepsilon q^i \circ \eta \ , \quad \frac{\partial q^i}{\partial \varepsilon} = q^i \eta.$$

For this family the calculation above shows that

$$\frac{dJ(\tilde{C})}{d\varepsilon}\Big|_{\varepsilon = 0} = \sum_{i=1}^{n} \int_{t_o}^{t_1} [-\frac{d}{dt} (\frac{\partial L}{\partial \dot{q}^i} \tilde{c}_o) \quad \frac{\partial L}{\partial q^i} \tilde{c}_o]\, q^i \eta\, dt = 0.$$

This holds for all η; we wish to conclude that the expression in brackets is zero for each i, for this will give us Lagrange's equations. If we call this expression M, this will follow from the following lemma.