MATHEMATICAL PRELIMINARIES

The purpose of these preliminaries is to introduce the notation, terminology, and principal mathematical results to be used in the course of the lectures. These will allow an easy, uninterrupted development of the physical theory. The results presented here may be found in greater detail in the following sources:

Geometric Integration Theory
H. Whitney
Princeton University Press (1957)

Ricci-Calculus
J. A. Schouten
Springer-Verlag (1954)

Finite Dimensional Vector Spaces
P. Halmos

Tensor Fields
J. L. Ericksen
Appendix, Classical Field Theories
Handbuch der Physik, III/1 (1960)

1. COTENSORS

Let V^n denote an n-dimensional real vector space. Elements of V^n will be denoted by boldface, lower case Latin letters, **v**, **u**, ... and will be called vectors. A real valued multilinear (i.e., linear in each argument) function of r vectors is called an r-cotensor. Thus, denoting the real numbers by R, an r-cotensor is a multilinear mapping

$$\boldsymbol{\alpha}: V^{rn} \to R,$$

where

$$V^{rn} = V^n \times V^n \times \dots \times V^n$$

is the r-fold Cartesian product of V^n. Cotensors will be denoted by lower case, Greek boldface letters.

The sum of any two cotensors and the product of a cotensor by a real number are defined by the relations

$$\begin{aligned} (\boldsymbol{\alpha} + \boldsymbol{\beta})(\mathbf{v},\mathbf{u}) &= \boldsymbol{\alpha}(\mathbf{v},\mathbf{u}) + \boldsymbol{\beta}(\mathbf{v},\mathbf{u}) \\ (\lambda\boldsymbol{\alpha})(\mathbf{u}) &= \lambda\,\boldsymbol{\alpha}(\mathbf{u}), \qquad \lambda \in R, \quad \mathbf{u} \in V^n. \end{aligned} \tag{1.1}$$

As in (1.1), we always denote real numbers by lightface, lower case Greek or Latin letters. With these definitions, the set of all r-cotensors is a certain linear space which we denote by V_{n^r}, and $\dim(V_{n^r}) = n^r$. We call V^n the carrier space of V_{n^r}. For $r = 1$, we write V_n and call the elements of V_n, covectors. The space of covectors is called the conjugate of V^n. More generally, the space of real valued linear functions of the elements of any linear space L is called the conjugate space L^c.

Let $\underset{\sim}{e}_i$, $i = 1, 2, \ldots, n$ denote a basis (linearly independent set of vectors) in the carrier space V^n. Then an arbitrary vector $v \in V^n$ has the representation

$$\underset{\sim}{v} = v^i \underset{\sim}{e}_i, \quad v^i \in R, \quad i = 1, 2, \ldots, n \qquad (1.2)$$

where the components v^i of $\underset{\sim}{v}$ with respect to the basis e_i are uniquely determined by $\underset{\sim}{v}$ and $\underset{\sim}{e}_i$. As in (1.2) we use the summation convention wherein, if the same letter appears in a given term of an expression in both a superior and inferior position (not at the same level, however), summation over the corresponding index set is implied without writing the summation sign. When the index set is not clear from the context, the summation sign will be used.

It is evident that an r-cotensor $\underset{\sim}{\alpha}$ is uniquely determined by its set of values

$$\alpha_{i_1 i_2 \ldots i_r} = \alpha_{(i)} = \alpha(\underset{\sim}{e}_{i_1}, \underset{\sim}{e}_{i_2}, \ldots, \underset{\sim}{e}_{i_r}). \tag{1.3}$$

For $r = 1$, the set of covectors $\underset{\sim}{\varepsilon}^i$, $i = 1, 2, \ldots, n$ defined by

$$\underset{\sim}{\varepsilon}^i(\underset{\sim}{e}_j) = \delta^i_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \tag{1.4}$$

are linearly independent and constitute a basis in the space V_n of covectors. The sets $\underset{\sim}{e}_i$ and $\underset{\sim}{\varepsilon}^j$ so related are called reciprocal bases for V^n and the conjugate space V_n.

Every covector α has the representation

$$\alpha = \alpha_i \underset{\sim}{\varepsilon}^i, \tag{1.5}$$

and if v is represented as in (1.2), then

$$\underset{\sim}{\alpha}(\underset{\sim}{v}) = v^i \alpha_i. \tag{1.6}$$

More generally, now, for $r > 1$, the n^r r-cotensors defined by

$$\underset{\sim}{\varepsilon}^{i_1} \otimes \underset{\sim}{\varepsilon}^{i_2} \otimes \ldots \otimes \underset{\sim}{\varepsilon}^{i_r}(\underset{\sim}{e}_{j_1}, \underset{\sim}{e}_{j_2}, \ldots, \underset{\sim}{e}_{j_r}) = \delta^{i_1}_{j_1} \delta^{i_2}_{j_2} \ldots \delta^{i_r}_{j_r} \tag{1.7}$$

are linearly independent and constitute a basis in V_{n^r} called the tensor basis of V_{n^r} corresponding to the basis $\underset{\sim}{e}_i$ in the carrier space V^n of V_{n^r}. Every r-cotensor $\underset{\sim}{\alpha}$ has the representation

$$\underset{\sim}{\alpha} = \alpha_{i_1 i_2 \ldots i_r} \underset{\sim}{\varepsilon}^{i_1} \otimes \underset{\sim}{\varepsilon}^{i_2} \otimes \ldots \otimes \underset{\sim}{\varepsilon}^{i_r} \qquad (1.8)$$

where the tensor components $\alpha_{(i)}$ of $\underset{\sim}{\alpha}$ are given by (1. 3).

The tensor product of an r-cotensor $\underset{\sim}{\alpha}$ and an s-cotensor $\underset{\sim}{\beta}$ is the $(r + s)$-cotensor $\underset{\sim}{\alpha} \otimes \underset{\sim}{\beta}$ defined by

$$(\underset{\sim}{\alpha} \otimes \underset{\sim}{\beta})(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r, \underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_s) =$$

$$\underset{\sim}{\alpha}(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r)\, \underset{\sim}{\beta}(\underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_s). \qquad (1.9)$$

The tensor components of $\underset{\sim}{\alpha} \otimes \underset{\sim}{\beta}$ are given in terms of the tensor components of $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ by

$$(\underset{\sim}{\alpha} \otimes \underset{\sim}{\beta})_{i_1 i_2 \ldots i_r j_1 j_2 \ldots j_2} = \alpha_{i_1 i_2 \ldots i_r} \beta_{j_1 j_2 \ldots j_s}. \qquad (1.10)$$

2. TENSORS AND MIXED TENSORS

A multilinear real valued function

$$T: V_{rn} \to R,$$

where

$$V_{rn} = V_n \times V_n \times \dots V_n$$

is the r-fold Cartesian product of V_n, is called an r-tensor.* All that has been said in §1 can now be repeated with the roles of V^n and V_n interchanged and with obvious changes in the terminology. In particular, every r-tensor T has the representation

$$T = T^{i_1 i_2 \dots i_r} \underset{\sim}{e}_{i_1} \otimes \underset{\sim}{e}_{i_2} \otimes \dots \otimes \underset{\sim}{e}_{i_r}, \qquad (2.1)$$

where the tensor components $T^{(i)}$ of $\underset{\sim}{T}$ are given by

$$T^{i_1 i_2 \dots i_r} = \underset{\sim}{T}\left(\underset{\sim}{\varepsilon}^{i_1}, \underset{\sim}{\varepsilon}^{i_2}, \dots, \underset{\sim}{\varepsilon}^{i_r}\right). \qquad (2.2)$$

The tensor product $\underset{\sim}{T} \otimes \underset{\sim}{S}$ of an r-tensor and an s-tensor is defined in obvious analogy to the tensor product of r-cotensors.

The set of all r-tensors with addition and multiplication by

---

*In tensor analysis, it is common to call an r-tensor a tensor of rank r. But we shall use the term rank of a tensor in an entirely different sense below; hence, we avoid the common terminology here.

scalars defined by

$$(\underset{\sim}{T}+\underset{\sim}{S})(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r) = \underset{\sim}{T}(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r) + \underset{\sim}{S}(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r)$$

$$(\lambda \underset{\sim}{T})(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r) = \lambda \underset{\sim}{T}(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2, \ldots, \underset{\sim}{\alpha}_r) \qquad (2.3)$$

is a linear space V^{n^r} of dimension n^r.

By definition, $(V^n)^c = V_n$; i.e., the space of covectors is the conjugate of the space of vectors. The conjugate $(V_n)^c$ of the space of covectors is, by definition, what we have called the space of 1-tensors V^{n^1}. Now V^n, the carrier space, has the same dimension as V^{n^1}; hence, they are isomorphic. The <u>natural isomorphism</u> ϕ: $V^n \rightarrow V^{n^1}$ is defined by

$$\underset{\sim}{\phi}(\underset{\sim}{v})(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}(\underset{\sim}{v}) = \alpha_i v^i. \qquad (2.4)$$

It is customary to denote $\underset{\sim}{\phi}(\underset{\sim}{v})$ and $\underset{\sim}{v}$ by the same letter and not to distinguish between vectors and 1-tensors whose carrier space V^n is the corresponding space of vectors. We adopt this convention here, but the natural isomorphism and all these agreements should be kept in mind. More generally now, the conjugate space $(V_{n^r})^c$ of the space of r-cotensors has the same dimension as the space V^{n^r} of r-tensors and, hence, they are isomorphic, but distinct. The <u>natural isomorphism</u>

$\underset{\sim}{\Phi}: V^{n^r} \to (W_{n^r})^c$ is defined as follows. Call an r-tensor $\underset{\sim}{T}$ simple if it is the tensor product of r-vectors; $\underset{\sim}{T} = \underset{\sim}{v}_1 \otimes \underset{\sim}{v}_2 \otimes \ldots \otimes \underset{\sim}{v}_r$. Define $\underset{\sim}{\alpha}(\underset{\sim}{T})$ for every simple r-covector by

$$\underset{\sim}{\alpha}(\underset{\sim}{T}) = \underset{\sim}{\alpha}(\underset{\sim}{v}_1 \otimes \underset{\sim}{v}_2 \otimes \ldots \otimes \underset{\sim}{v}_r) = \alpha(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r) \quad (2.5)$$

and set

$$\underset{\sim}{\alpha}(\lambda \underset{\sim}{T} + \eta \underset{\sim}{S}) = \lambda \underset{\sim}{\alpha}(\underset{\sim}{T}) + \eta \underset{\sim}{\alpha}(\underset{\sim}{S}) \quad (2.6)$$

if $\underset{\sim}{T}$ and $\underset{\sim}{S}$ are simple, but $T+S$ not necessarily simple. But every r-tensor $\underset{\sim}{T}$ is the sum of a finite number of simple r-tensors (cf., the representation (2.1)). Hence, (2.5) and (2.6) define $\underset{\sim}{\alpha}(\underset{\sim}{T})$ for arbitrary $\underset{\sim}{T} \in V^{n^r}$. The natural isomorphism between V^{n^r} and $(V_{n^r})^c$ is then defined by

$$\underset{\sim}{\Phi}(\underset{\sim}{T})(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}(\underset{\sim}{T}). \quad (2.7)$$

Here, as in the case of vectors and 1-tensors, it is conventional not to distinguish between "co-cotensors" $\underset{\sim}{\Phi}(\underset{\sim}{T})$ and the tensor $\underset{\sim}{T}$ related by the natural isomorphism established by (2.7). Thus, $\underset{\sim}{\Phi}(\underset{\sim}{T})$ and $\underset{\sim}{T}$ are denoted by the common symbol $\underset{\sim}{T}$, and

$$\underset{\sim}{T}(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}(\underset{\sim}{T}) = \alpha_{i_1 i_2 \ldots i_r} T^{i_1 i_2 \ldots i_r}, \quad (2.8)$$

where $\alpha_{(i)}$ and $T^{(i)}$ are the components of α and T with respect to arbitrary tensor bases in V_{n^r} and V^{n^r}.

More generally, now, a real valued multilinear function

$$M: V_1^{n_1} \times V_2^{n_2} \times \ldots \times V_r^{n_r} \to R \qquad (2.9)$$

of r vector arguments drawn from an arbitrary collection of r vector spaces is called a <u>mixed</u> r-tensor unless $V_1 = V_2 = \ldots V_r$.

It suffices to illustrate the general case by the case $r = 2$. Then

$$M: V^n \times U^m \to R,$$

say. Let e_i and E_α, $i = 1, 2, \ldots, n$, $\alpha = 1, 2, \ldots, m$ be bases in V^n and U^m, respectively. Then, M is uniquely determined by its components

$$M_{i\alpha} = M(e_i, E_\alpha). \qquad (2.10)$$

Let $M + N$ and λM be defined by

$$(M + N)(v, u) = M(v, u) + N(v, u),$$
$$(\lambda M)(v, u) = M(v, u). \qquad (2.11)$$

Then, the set of all such mixed tensors is a linear space W_{mn}

of dimension mn. A basis in W_{mn} consists in the mn elements defined by

$$(\underset{\sim}{\varepsilon}^i \otimes \underset{\sim}{\mathcal{E}}^\alpha)(\underset{\sim}{e}_j, \underset{\sim}{e}_\beta) = \delta^i_j \delta^\alpha_b, \qquad (2.12)$$

and an arbitrary $M \in W_{mn}$ has the representation

$$\underset{\sim}{M} = M_{i\alpha} \underset{\sim}{\varepsilon}^i \otimes \underset{\sim}{\mathcal{E}}^\alpha, \qquad (2.13)$$

where the components of $\underset{\sim}{M}$, the $M_{i\alpha}$, are given by (2.10).

Consider the special case of the above where $V^n = U_m$, the conjugate space of the second argument U^m. Then,

$$\underset{\sim}{M}: U_m \times U^m \to R$$

and we set

$$M^\alpha{}_\beta = M(\underset{\sim}{\mathcal{E}}^\alpha, \underset{\sim}{E}_\beta) \qquad (2.14)$$

where the $\underset{\sim}{\mathcal{E}}^\alpha$ and $\underset{\sim}{E}_\beta$ are reciprocal, $\underset{\sim}{E}_\alpha(\underset{\sim}{\mathcal{E}}^\beta) = \delta^\beta_\alpha$. Thus,

$$\underset{\sim}{M} = M^\alpha{}_\beta \underset{\sim}{E}_\alpha \otimes \underset{\sim}{\mathcal{E}}^\beta. \qquad (2.15)$$

Every such mixed tensor $\underset{\sim}{M}$ determines a unique linear transformation

$$\underset{\sim}{M}^*: U^m \to U^m$$

defined as follows:

$$\underset{\sim}{M}^*(\underset{\sim}{u})(\underset{\sim}{\alpha}) = \underset{\sim}{M}(\underset{\sim}{\alpha}, \underset{\sim}{u})$$

where $\underset{\sim}{M}^*(\underset{\sim}{u}) \in (U_m)^c$ is that unique element of U^m determined by the natural isomorphism $\underset{\sim}{\phi}$ (cf., 2.4). With respect to tensor bases in U_m and U^m

$$M^{*\alpha}{}_{\beta} = M^{\alpha}{}_{\beta},$$

where

$$M^{*\alpha}{}_{\beta} = \underset{\sim}{\mathcal{E}}^{\alpha}(\underset{\sim}{M}^*(\underset{\sim}{E}_{\beta}))$$

are the components of $\underset{\sim}{M}^*$. In a similar way, the mixed tensor $\underset{\sim}{M}$ determines a linear transformation $\underset{\sim}{M}^{**}: U_m \to U_m$ defined by $\underset{\sim}{M}^{**}(\underset{\sim}{\alpha})(\underset{\sim}{v}) = \underset{\sim}{M}(\underset{\sim}{\alpha}, \underset{\sim}{v})$ and the tensor components of $\underset{\sim}{M}^{**}$ defined by $M^{**\alpha}{}_{\beta} = \underset{\sim}{E}_{\beta}(\underset{\sim}{M}^{**}(\underset{\sim}{\mathcal{E}}^{\alpha}))$ are also equal, respectively, to the components $M^{\alpha}{}_{\beta}$ of $\underset{\sim}{M}$. These definitions justify and are consistent with the usual rules and conventions of tensor algebra. There, it is traditional not to distinguish between $\underset{\sim}{M}$, $\underset{\sim}{M}^*$, and $\underset{\sim}{M}^{**}$ and to denote all three by the common symbol $\underset{\sim}{M}$. In the following we shall use absolute notations or the kernel index notations of tensor algebra interchangeably according to whichever seems most efficient and expressive in a given context. By the components of tensors, cotensors, or mixed tensors we shall always mean <u>tensor components</u> as these have been defined above. In the physical applications, a mixed tensor, r-tensor, or r-cotensor is generally introduced into the physical theory with a specific logical meaning; e. g., as a 1-tensor and not a vector, or as a linear transformation of some vector space, and not as a mixed tensor, but then the natural isomorphism established above is used freely to define other operations in which the mixed tensor, r-tensor, or r-cotensor plays a different logical role.

3. THE SYMMETRY PARTS OF TENSORS AND COTENSORS

Let $\underset{\sim}{L}$ be a nonsingular linear transformation

$$\underset{\sim}{L}: V^n \rightarrow V^n \tag{3.1}$$

of an n-dimensional vector space V^n. Then $\underset{\sim}{L}$ induces a linear tensor transformation

$$\underset{\sim}{L}_r: V_{n^r} \rightarrow V_{n^r} \tag{3.2}$$

in the space V_{n^r} of r-cotensors having V^n as carrier space. The transformation $\underset{\sim}{L}_r$ is defined by

$$(\underset{\sim}{L}_r \alpha)(\underset{\sim}{L}v_1, \underset{\sim}{L}v_2, \ldots, \underset{\sim}{L}v_r) = \underset{\sim}{\alpha}(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r). \tag{3.3}$$

If

$$\underset{\sim}{L}_r(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}, \tag{3.4}$$

then $\underset{\sim}{\alpha}$ is said to be <u>invariant under the tensor transformation</u> $\underset{\sim}{L}_r$. More generally, let $U \subset V_{n^r}$ be a proper subspace of V_{n^r}. Then if

$$\underset{\sim}{L}_r(U) \subset U \tag{3.5}$$

for every $\underset{\sim}{L}_r$ in some set $\{\underset{\sim}{L}_r\}$ of tensor transformations, then U is an <u>invariant subspace</u> under the set of transformations $\{\underset{\sim}{L}_r\}$.

Consider the symmetric (permutation) group S^r on the first r integers and let $\underset{\sim}{\Pi} = \begin{pmatrix} 1 & 2 & 3 & \ldots & r \\ \pi_1 & \pi_2 & \pi_3 & \ldots & \pi_r \end{pmatrix} \in S^r$. Then, in terms of an arbitrary r-cotensor $\underset{\sim}{\alpha}$ we define the r-cotensor $\Pi(\underset{\sim}{\alpha})$ by

$$\underset{\sim}{\Pi}\underset{\sim}{\alpha}(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r) = \underset{\sim}{\alpha}(\underset{\sim}{v}_{\pi_1}, \underset{\sim}{v}_{\pi_2}, \ldots, \underset{\sim}{v}_{\pi_r}). \quad (3.6)$$

The r-cotensor $\underset{\sim}{\Pi}\underset{\sim}{\alpha}$ is called an <u>isomer</u> of $\underset{\sim}{\alpha}$. Defining

$$\underset{\sim}{\Pi}(\underset{\sim}{\alpha}+\underset{\sim}{\beta}) = \underset{\sim}{\Pi}(\underset{\sim}{\alpha}) + \underset{\sim}{\Pi}(\underset{\sim}{\beta}), \quad \underset{\sim}{\Pi}(\lambda\underset{\sim}{\alpha}) = \lambda\underset{\sim}{\Pi}(\underset{\sim}{\alpha}), \quad (3.7)$$

every permutation $\underset{\sim}{\Pi} \in S^r$ determines a linear transformation

$$\underset{\sim}{\Pi}: V_{n^r} \rightarrow V_{n^r}. \quad (3.8)$$

It is easy to see that every linear transformation $\underset{\sim}{\Pi}$ defined in this way commutes with every tensor transformation $\underset{\sim}{L}_r$:

$$\underset{\sim}{\Pi}\underset{\sim}{L}_r(\underset{\sim}{\alpha}) = \underset{\sim}{L}_r\underset{\sim}{\Pi}(\underset{\sim}{\alpha}). \quad (3.9)$$

More generally, now, consider the enveloping algebra $A^{r!}$ of the symmetric group S^r. Define the transformations $\underset{\sim}{\Pi}+\underset{\sim}{\Omega}$, and $\lambda\underset{\sim}{\Pi}$ in V_{n^r} by

$$(\underset{\sim}{\Pi}+\underset{\sim}{\Omega})\underset{\sim}{\alpha} = \underset{\sim}{\Pi}\underset{\sim}{\alpha} + \underset{\sim}{\Omega}\beta, \quad (\lambda\,\underset{\sim}{\Pi})\underset{\sim}{\alpha} = \lambda\underset{\sim}{\Pi}\underset{\sim}{\alpha}. \quad (3.10)$$

Then, each element of $A^{r!}$, say,

$$\underset{\sim}{A} = \sum_{s=1}^{r!} a^s \underset{\sim}{\Pi}_s, \quad \underset{\sim}{\Pi}_s \in S^r \tag{3.11}$$

determines a linear transformation of V_{n^r}, defined by the above and

$$\underset{\sim}{A}(\underset{\sim}{\alpha}) = \sum_{s=1}^{r!} a^s \underset{\sim}{\Pi}_s(\underset{\sim}{\alpha}). \tag{3.12}$$

The algebra $A^{r!}$ possesses a resolution of its identity element $\underset{\sim}{I}$,

$$\underset{\sim}{I} = \underset{\sim}{J}_1 + \underset{\sim}{J}_2 + \dots + \underset{\sim}{J}_p, \tag{3.13}$$

where each $\underset{\sim}{J}_k$, $k = 1, 2, \dots, p$ is idempotent and irreducible and such that

$$\underset{\sim}{J}_k^2 = \underset{\sim}{J}_k, \quad \underset{\sim}{J}_k\underset{\sim}{J}_h = 0 = \underset{\sim}{J}_h\underset{\sim}{J}_k, \quad h \neq k. \tag{3.14}$$

Irreducible means that there exists no decomposition of any $\underset{\sim}{J}_h$ into a sum $\underset{\sim}{J}_h = \underset{\sim}{A} + \underset{\sim}{B}$ such that $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are idempotent and $\underset{\sim}{A}\underset{\sim}{B} = \underset{\sim}{B}\underset{\sim}{A} = 0$. From the existence of the resolution of the identity (3.13) it follows that every r-cotensor α can be resolved into symmetry parts $\underset{\sim}{\alpha}_k = \underset{\sim}{J}_k(\underset{\sim}{\alpha})$ as follows:

$$\underset{\sim}{\alpha} = \underset{\sim}{I}\underset{\sim}{\alpha} = \sum_{k=1}^{p} \underset{\sim}{J}_k\underset{\sim}{\alpha} = \sum_{k=1}^{p} \underset{\sim}{\alpha}_k, \tag{3.15}$$

and one has

$$\underset{\sim}{J}_k(\underset{\sim}{\alpha}_h) = \delta_{kh}\,\underset{\sim}{\alpha}_h \,. \tag{3.16}$$

If $\underset{\sim}{J}_k\underset{\sim}{\alpha} = \underset{\sim}{\alpha}$, α is said to have symmetry $\{k\}$. The sum $\underset{\sim}{\alpha}+\underset{\sim}{\beta}$ and product $\lambda\underset{\sim}{\alpha}$ of r-cotensors of symmetry $\{k\}$ are r-cotensors of the same symmetry class $\{k\}$. Hence, V_{n^r} is resolved as follows:

$$V_{n^r} = V^1_{n^r} \oplus V^2_{n^r} \oplus \cdots \oplus V^p_{n^r} \tag{3.17}$$

into a direct sum of subspaces of r-cotensors of given symmetry. (Some of these subspaces may be empty; i.e., may have dimension 0.)

It follows from the commutativity property (3.9) that the subspaces $V^k_{n^r}$ of r-cotensors of given symmetry are invariant subspaces of every tensor transformation $\underset{\sim}{L}_r$ of V_{n^r}:

$$\underset{\sim}{L}_r(V^k_{n^r}) = \underset{\sim}{L}_r\underset{\sim}{J}_k(V_{n^r}) = \underset{\sim}{J}_k\underset{\sim}{L}_r(V_{n^r}) \subset \underset{\sim}{J}_k(V_{n^r}),$$

or

$$\underset{\sim}{L}_r(V^k_{n^r}) \subset V^k_{n^r} \,.$$

<u>Theorem:</u> (Weyl) The resolution (3.17) of V_{n^r} into a direct

sum of invariant subspaces of r-cotensors of given symmetry is a maximal decomposition of V_{n^r} under the set of all tensor transformations $\{\underset{\sim}{L}_r\}$ induced in V_{n^r} by the set of all nonsingular linear transformations $\{\underset{\sim}{L}\}$ of V^n. In other words, if $\{\underset{\sim}{L}\}$ is the set of all nonsingular linear transformations of V^n and $\{\underset{\sim}{L}_r\}$ the corresponding set of tensor transformations induced in V_{n^r}, then no proper subspace of any $V^k_{n^r}$, $k = 1, 2, \ldots, p$ is invariant under the set $\{\underset{\sim}{L}_r\}$.

If $r > 1$, the number $p(r)$ of symmetry classes $\{k\}$ is always ≥ 2. Amongst these p symmetry classes for every value of $r > 1$ is the class of symmetric r-cotensors for which

$$\underset{\sim}{\Pi}\underset{\sim}{\alpha} = \underset{\sim}{\alpha}, \qquad \text{for every permutation } \underset{\sim}{\Pi} \in S^r$$

and the class of antisymmetric r-cotensors for which

$$\underset{\sim}{\Pi}\underset{\sim}{\alpha} = \underset{\sim}{\alpha} \qquad \text{for every even permutation } \underset{\sim}{\Pi},$$

$$\underset{\sim}{\Pi}\underset{\sim}{\alpha} = -\underset{\sim}{\alpha} \qquad \text{for every odd permutation } \underset{\sim}{\Pi}.$$

For brevity, antisymmetric r-cotensors are called r-covectors. The subspaces of symmetric and antisymmetric r-cotensors are denoted by the special symbols $V_{(n^r)}$ and $V_{[n^r]}$, respectively. For $r = 2$, the resolution (3.17) reduces to the familiar decomposition

$$V_{n^2} = V_{(n^2)} \oplus V_{[n^2]} \tag{3.18}$$

of 2-cotensors into their symmetric and antisymmetric parts. But for $r > 2$,

$$V_{n^r} = V_{(n^r)} \oplus V_{[n^r]} \oplus U,$$

where, in general, U is not empty.

If $\alpha_{i_1 i_2 \ldots i_r}$ are the components of an r-cotensor with respect to some basis e^i, we denote the components of the symmetric and antisymmetric parts of $\underset{\sim}{\alpha}$ by $\alpha_{(i_1 i_2 \ldots i_r)}$ and $\alpha_{i_1 i_2 \ldots i_r}$, respectively.

$$\alpha_{(i_1 i_2 \ldots i_r)} = (1/r!) \sum_{\Pi \in S^r} \alpha_{i_{\pi_1} i_{\pi_2} \ldots i_{\pi_r}}$$

$$\alpha_{i_1 i_2 \ldots i_r} = (1/r!) \sum_{\Pi \in S^r} (-)^{\sigma_\pi} \alpha_{i_{\pi_1} i_{\pi_2} \ldots i_{\pi_r}},$$

where $\sigma_\pi = 0, 1$ for even and odd permutations, respectively.

All that has been said above for r-cotensors holds with minor changes for r-tensors, perhaps with one exception. If L is a linear transformation of the carrier space V^n, then the linear tensor transformation of the space of r-tensors V^{n^r}

induced by L is defined by

$$\underset{\sim}{L}^r(\underset{\sim}{T}) = T^{i_1 i_2 \cdots i_r}(\underset{\sim}{L}\,\underset{\sim}{e}_{i_1}) \otimes (\underset{\sim}{L}\,\underset{\sim}{e}_{i_2}) \cdots \otimes (\underset{\sim}{L}\,\underset{\sim}{e}_{i_r}) \quad (3.19)$$

where

$$T = T^{i_1 i_2 \cdots i_r} \underset{\sim}{e}_{i_1} \otimes \underset{\sim}{e}_{i_2} \otimes \cdots \otimes \underset{\sim}{e}_{i_r}.$$

The definitions of $\underset{\sim}{L}_r$ (3.3) and $\underset{\sim}{L}^r$ (3.19) imply that

$$\overline{\underset{\sim}{T}}(\overline{\underset{\sim}{\alpha}}) = \underset{\sim}{T}(\underset{\sim}{\alpha}) \quad (3.20)$$

where $\overline{\underset{\sim}{T}} = \underset{\sim}{L}^r(\underset{\sim}{T})$ and $\overline{\underset{\sim}{\alpha}} = \underset{\sim}{L}_r(\underset{\sim}{\alpha})$, provided that $\underset{\sim}{L}$ is non-singular so that $\underset{\sim}{L}_r$ is defined. The subspaces of antisymmetric and symmetric r-tensors are denoted by $V^{[n^r]}$ and $V^{(n^r)}$, respectively. For brevity, we call antisymmetric r-tensors, r-vectors.

If $\underset{\sim}{T}^k$ is a tensor of symmetry class $\{k\}$ and $\underset{\sim}{\alpha}_h$ a cotensor of symmetry class $\{h\}$, then

$$\underset{\sim}{T}^k(\underset{\sim}{\alpha}_h) = 0 \qquad \text{if } \{k\} \neq \{h\}. \quad (3.21)$$

It follows that

$$\underset{\sim}{T}(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}(\underset{\sim}{T}) = \sum_{\{k\}} \underset{\sim}{T}^k(\underset{\sim}{\alpha}_k) \quad (3.22)$$

if $\underset{\sim}{T} = \sum_{\{k\}} \underset{\sim}{T}^k$, $\underset{\sim}{\alpha} = \sum_{\{h\}} \underset{\sim}{\alpha}_h$.

4. THE GRASSMAN ALGEBRA

Let V^n be an n-dimensional vector space and consider the direct sum (Grassman space)

$$G = V_{[1]} \oplus V_{[n]} \oplus V_{[n^2]} \oplus \cdots \oplus V_{[n^n]} \tag{4.1}$$

of the spaces of r-covectors, $r = 0, 1, 2, \ldots, n$ having the common carrier space V^n. The space $V_{[1]}$ is the space of scalars, and $V_{[n]} = V_n$ is the space of covectors. Then

$$\dim(G) = 1 + n + \binom{n}{2} + \binom{n}{3} = (1+1)^n = 2^n. \tag{4.2}$$

Let $\underset{\sim}{\alpha}_r$ and $\underset{\sim}{\beta}_s$ be an r-covector and an s-covector, respectively, and let $\underset{\sim}{J}_{[r]}$ be the idempotent linear transformation (antisymmetrizer) that projects V_{n^r} into $V_{[n^r]}$:

$$\underset{\sim}{J}_{[r]}(V_{n^r}) = V_{[n^r]}. \tag{4.3}$$

The exterior (Grassman) product $\underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s$ is the $(r+s)$-covector defined by

$$\underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s = \frac{(r+s)!}{r!\, s!} \underset{\sim}{J}_{[r+s]}(\underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s). \tag{4.4}$$

In words, the exterior product of an r-covector and an s-covector is the antisymmetric part of their tensor product times the numerical factor $(r+s)/r!\,s!$. From the associativity of the tensor product and the property

$$\underset{\sim}{J}_{[r+s]}(\underset{\sim}{\alpha}_r \otimes \underset{\sim}{\beta}_s) = \underset{\sim}{J}_{[r+s]}(\underset{\sim}{J}_{[r]}\underset{\sim}{\alpha} \otimes \underset{\sim}{J}_{[s]}\underset{\sim}{\beta})$$

it follows that

$$\underset{\sim}{\alpha}_r \vee (\underset{\sim}{\beta}_s \vee \underset{\sim}{\gamma}_t) = (\underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s) \vee \underset{\sim}{\gamma}_t \tag{4.5}$$

so that the exterior product defined by (4.4) is associative, but

$$\underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s = (-)^{rs}\, \underset{\sim}{\beta}_s \vee \underset{\sim}{\alpha}_r. \tag{4.6}$$

Also, $\underset{\sim}{\alpha}_r \vee (\underset{\sim}{\beta}_s + \underset{\sim}{\gamma}_s) = \underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s + \underset{\sim}{\alpha}_r \vee \underset{\sim}{\gamma}_s$, and $\underset{\sim}{\alpha}_r \vee (\lambda \underset{\sim}{\beta}_s) = (\lambda \underset{\sim}{\alpha}_r) \vee \underset{\sim}{\beta}_s = \lambda \underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s$. Thus, if $\underset{\sim}{\alpha} = \langle \underset{\sim}{\alpha}_0, \underset{\sim}{\alpha}_1, \ldots, \underset{\sim}{\alpha}_n \rangle$ and $\underset{\sim}{\beta} = \langle \underset{\sim}{\beta}_0, \underset{\sim}{\beta}_1, \ldots, \underset{\sim}{\beta}_n \rangle$ are any two elements of the Grassman space G and we define their product by

$$\underset{\sim}{\alpha} \vee \underset{\sim}{\beta} = \langle \underset{\sim}{\alpha}_0 \underset{\sim}{\beta}_0,\ \sum_{r+s=1} \underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s,\ \sum_{r+s=2} \underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s,\ \ldots,\ \sum_{r+s=n} \underset{\sim}{\alpha}_r \vee \underset{\sim}{\beta}_s \rangle, \tag{4.7}$$

then the bilinear mapping $\underset{\sim}{\Gamma}$: $G \times G \to G$ with $\underset{\sim}{\Gamma}(\underset{\sim}{\alpha}, \underset{\sim}{\beta}) = \underset{\sim}{\alpha} \vee \underset{\sim}{\beta}$ determines the linear associative <u>Grassman algebra</u> $(\underset{\sim}{\Gamma}, G)$. A corresponding algebra is defined in the same way in the space

$$G^c = V^{[1]} + V^{[n]} + V^{[n^r]} + \ldots + V^{[n^n]} \quad (4.8)$$

which we denote by $(\underset{\sim}{\Gamma}, G^c)$. Since G^c is the conjugate of G, if $\underset{\sim}{\alpha} \in G$ and $\underset{\sim}{v} \in G^c$, $\underset{\sim}{v}(\underset{\sim}{\alpha})$ is defined and given by

$$\underset{\sim}{v}(\underset{\sim}{\alpha}) = \underset{\sim}{\alpha}(\underset{\sim}{v}) = \sum_{r=0}^{n} \underset{\sim}{\alpha}_r(\underset{\sim}{v}^r), \quad (4.9)$$

where

$$\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_0 \oplus \underset{\sim}{\alpha}_1 \oplus \ldots \oplus \underset{\sim}{\alpha}_n, \quad \underset{\sim}{v} = \underset{\sim}{v}^1 \oplus \ldots \oplus \underset{\sim}{v}^n.$$

A set of vectors (covectors) $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r$ in V^n is linearly independent if and only if the r-vector (covector) $\underset{\sim}{v}_1 \vee \underset{\sim}{v}_2 \vee \ldots \vee \underset{\sim}{v}_r$ is different from zero.

An r-vector $\underset{\sim}{w}_r \in V^{[n^r]}$ is <u>simple</u> if and only if there exists a set of r-vectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r$ such that $\underset{\sim}{w}_r = \underset{\sim}{v}_1 \vee \underset{\sim}{v}_2 \vee \ldots \vee \underset{\sim}{v}_r$.

If $\underset{\sim}{w}$ is an r-vector and $\underset{\sim}{v}$ a vector, then $\underset{\sim}{v}$ is a <u>divisor</u> of $\underset{\sim}{w}$ if and only if $\underset{\sim}{w} \vee \underset{\sim}{v} = \underset{\sim}{v} \vee \underset{\sim}{w} = 0$. It is known that $\underset{\sim}{v}$ is a divisor of $\underset{\sim}{w}$ if and only if there exists an (r-1)-vector $\underset{\sim}{u}$ such that $\underset{\sim}{w} = \underset{\sim}{u} \vee \underset{\sim}{v}$.

Let $\underset{\sim}{\alpha}$ be any r-cotensor. Then with respect to each argument of $\underset{\sim}{\alpha}$, say the p^{th}, there is a set of $n^{(r-1)}$ covectors $\underset{\sim}{\alpha}^{(p)}_{i_1 i_2 \ldots i_{r-1}}$ defined by

$$\alpha^{(p)}_{i_1 i_2 \dots i_r}(u) = \alpha(e_{i_1}, e_{i_2}, \dots, e_{i_{p-1}}, u, e_{i_{p+1}}, \dots, e_{i_r}), \tag{4.10}$$

e_i a basis. The number of linearly independent covectors in this set is called the p^{th} rank of α. The p^{th} rank of r-tensors and of general mixed tensors is defined in the obvious analogous way. Every rank of an r-covector or of an r-vector has one and the same value, which is called simply its rank. The rank of a 2-covector (2-tensor) is always an even number. If $2s$ is the rank of the r-covector α, then α is expressible as the sum of s simple 2-covectors:

$$\alpha = \beta_1 \vee \gamma_1 + \beta_2 \vee \gamma_2 + \dots + \beta_s \vee \gamma_s, \quad \text{if rank } (\alpha) = 2s, \tag{4.11}$$

where the β_p, γ_p, $p = 1, 2, \dots, s$ are linearly independent. Thus, if the β_p, γ_p are the first $2s$ elements of a basis in V^n, the matrix of components of α with respect to such a basis has the values

$$||\alpha_{ij}|| = \text{diag}\,(Q, \dots Q, 0, 0, \dots 0), \tag{4.12}$$

where Q is the 2×2 matrix $\begin{Vmatrix} 0 & 1 \\ -1 & 0 \end{Vmatrix}$.

If α_r is an r-covector and v^r an r-vector, their scalar product $\alpha_r \cdot v^r = v^r \cdot \alpha_r$ is defined by

$$\alpha_r \cdot v^r = (1/r!)\, \alpha_r(v^r). \tag{4.13}$$

The interior product of an (r+s)-covector and an s-vector is then defined by

$$(\underset{\sim}{\alpha}_{r+s} \wedge \underset{\sim}{v}^{s}) \cdot \underset{\sim}{w}^{r} = \underset{\sim}{\alpha}_{r+s} \cdot (\underset{\sim}{v}^{s} \vee \underset{\sim}{w}^{s}), \quad \text{for all } \underset{\sim}{w}^{r}. \tag{4.14}$$

The interior product of an (r+s)-vector and an s-covector is defined by

$$(\underset{\sim}{v}^{r+s} \wedge \underset{\sim}{\alpha}_{s}) \cdot \underset{\sim}{\beta}_{r} = \underset{\sim}{v}^{r+s} \cdot (\underset{\sim}{\beta}_{r} \vee \underset{\sim}{\alpha}_{s}), \quad \text{for all } \underset{\sim}{\beta}_{r}. \tag{4.15}$$

5. DUALITY

Let $\underset{\sim}{e} \neq 0$ be an arbitrary n-covector with carrier space V^n. Every n-covector with an n-dimensional carrier space is simple; therefore, there exists a linearly independent set of covectors $\underset{\sim}{e}^i$, $i = 1, 2, \ldots, n$ such that

$$\underset{\sim}{e} = \underset{\sim}{e}^1 \vee \underset{\sim}{e}^2 \vee \ldots \vee \underset{\sim}{e}^n. \tag{5.1}$$

Let $\underset{\sim}{E}$ denote the corresponding n-vector defined by the reciprocal set $\underset{\sim}{e}_i$ of vectors:

$$\underset{\sim}{E} = \underset{\sim}{e}_1 \vee \underset{\sim}{e}_2 \vee \ldots \vee \underset{\sim}{e}_n. \tag{5.2}$$

If the $\underset{\sim}{e}_i$ are the basis vectors in V^n, then the corresponding tensor components of $\underset{\sim}{e}$ and $\underset{\sim}{E}$ given by

$$\begin{aligned} E^{i_1 i_2 \ldots i_n} &= \underset{\sim}{E}(\underset{\sim}{e}^{i_1}, \underset{\sim}{e}^{i_2}, \ldots, \underset{\sim}{e}^{i_n}), \\ e_{i_1 i_2 \ldots i_n} &= \underset{\sim}{e}(\underset{\sim}{e}_{i_1}, \underset{\sim}{e}_{i_2}, \ldots, \underset{\sim}{e}_{i_n}) \end{aligned} \tag{5.3}$$

have the values $E^{12\ldots n} = +1$, $e_{12\ldots n} = +1$. The value of every other component of $\underset{\sim}{E}$ and $\underset{\sim}{e}$ is determined by the values of these two components and the antisymmetry of $\underset{\sim}{E}$ and $\underset{\sim}{e}$. The components $E^{i_1 i_2 \ldots i_n}$ and $e_{i_1 i_2 \ldots i_n}$ are called the <u>permutation symbols.</u> Note that every linearly independent set of vectors $\underset{\sim}{e}_i$ (i.e., every basis) determines a corresponding $\underset{\sim}{E}$ and $\underset{\sim}{e}$. These should perhaps be distinguished by writing $\underset{\sim}{E}(i)$ and $\underset{\sim}{e}(i)$ to indicate their dependence on the basis $\underset{\sim}{e}_i$. The components of a given $\underset{\sim}{E}(i)$, $\underset{\sim}{e}(i)$ with respect to another basis $\underset{\sim}{e}_{i'}$, say, do not, in general, have the values given by the permutation symbols (5.3). Rather, they are determined by the general relations

$$\begin{aligned} \underset{\sim}{E}(i') &= (\det \underset{\sim}{S})\, \underset{\sim}{E}(i), \\ \underset{\sim}{e}(i') &= (\det \underset{\sim}{S})^{-1}\, \underset{\sim}{e}(i), \end{aligned} \tag{5.4}$$

where $S^i{}_{j'}$ is the matrix which defines the $\underset{\sim}{e}_{i'}$ as a linear combination of the set $\underset{\sim}{e}_i$:

$$\underset{\sim}{e}_{j'} = S^i{}_{j'}\, \underset{\sim}{e}_i. \tag{5.5}$$

It follows that the mixed tensor $\underset{\sim}{\delta} = \underset{\sim}{E}(i) \otimes \underset{\sim}{e}(i)$ is independent of i. The mixed tensor $\frac{1}{r!}\underset{\sim}{\delta}$ so defined is the identity transformation $\frac{1}{r!}\underset{\sim}{\delta}$: $V^{[n^n]} \to V^{[n^n]}$. More generally, the identity transformation $(1/r!)\,\delta_r$: $V^{n^r} \to V^{n^r}$ has tensor components with respect to an arbitrary basis $\underset{\sim}{e}_i$ in V^n given by

$$\delta^{i_1 i_2 \cdots i_r}_{j_1 j_2 \cdots j_r} = (1/n-r)!)E^{i_1 i_2 \cdots i_r k_1 \cdots k_{n-r}}\, e_{j_1 j_2 \cdots j_r k_1 \cdots k_{n-r}} \tag{5.6}$$

where $\underset{\sim}{E}$ and $\underset{\sim}{e}$ are the n-vector and n-covector defined in terms of an arbitrary linearly independent set of vectors.

Formulas like (5.6) point up the need for a more efficient and condensed notation when dealing with components of r-vectors and r-covectors. For

$$\alpha^{\cdots\, i_1 i_2 \cdots i_r}_{\cdots}$$

where α is antisymmetric in the indices $i_1 \ldots i_r$, let us write $\alpha^{\cdots}_{\cdots (i)}$, and for the contracted product such as occurs in (5.6), let us write

$$(1/r!)\, \alpha^{\cdots}_{\cdots i_1 i_2 \cdots i_r}\, \beta^{\cdots i_1 i_2 \cdots i_r}_{\cdots} = \alpha^{\cdots}_{\cdots (i)}\, \beta^{\cdots (i)}_{\cdots} \qquad (5.7)$$

where $\underset{\sim}{\alpha}$ and $\underset{\sim}{\beta}$ are general mixed tensors. Thus, for example, the scalar product defined in (4.13) is given, in the condensed notation, in terms of components by

$$\underset{r}{\alpha} \cdot \overset{r}{v} = \alpha_{r(i)}\, v^{r(i)},$$

and, more briefly, by $\alpha_{(i)} v^{(i)}$, when the value of r is clear from the context or unimportant for the meaning of the term.

For each choice of $\underset{\sim}{e}$ and $\underset{\sim}{E}$ we can show that the mappings

$$\underset{\sim}{D}: V^{[n^r]} \rightarrow V_{[n^{n-r}]},$$
$$\underset{\sim}{D}': V_{[n^r]} \rightarrow V^{[n^{n-r}]} \tag{5.8}$$

defined by

$$\underset{\sim}{D}(\underset{\sim}{v}^r) = \underset{\sim}{e} \wedge \underset{\sim}{v}^r,$$
$$\underset{\sim}{D}'(\underset{\sim}{\alpha}_r) = \underset{\sim}{E} \wedge \underset{\sim}{\alpha}, \tag{5.9}$$

have the property

$$\underset{\sim}{D}\,\underset{\sim}{D}' = \underset{\sim}{\delta}^r = \underset{\sim}{D}'\,\underset{\sim}{D}. \tag{5.10}$$

Thus, $\underset{\sim}{D}$ and $\underset{\sim}{D}'$ are 1-1 and onto. $\underset{\sim}{D}\underset{\sim}{v}^r$ is called the <u>dual</u> of $\underset{\sim}{v}^r$ and $\underset{\sim}{D}'\underset{\sim}{\alpha}_r$ is called the <u>dual</u> of $\underset{\sim}{\alpha}_r$. It must be kept in mind that the duality isomorphism between the spaces of r-vectors and (n-r)-covectors established by D and D' depends on the choice of basis used to define $\underset{\sim}{E}$ and $\underset{\sim}{e}$. Because of the relations (5.4), two bases $\underset{\sim}{e}_i$ and $\underset{\sim}{e}_{i'}$ determine different isomorphisms unless $\underset{\sim}{e}_i$ and $\underset{\sim}{e}_{i'}$ are related by a transformation with determinant +1; i.e., by a unimodular transformation.

The interior and exterior products of r-covectors and r-vectors introduced above and the duality mappings based on an n-vector and n-covector are related to the classical cross product of Gibbs' vector analysis in the following way:

<u>Definition of cross product:</u>

$$\underset{\sim}{\alpha} \times \underset{\sim}{\beta} = \underset{\sim}{D}'(\underset{\sim}{\alpha} \vee \underset{\sim}{\beta})$$
$$\underset{\sim}{u} \times \underset{\sim}{v} = \underset{\sim}{D}(\underset{\sim}{u} \vee \underset{\sim}{v}).$$

6. QUADRATIC FORMS

A 2-cotensor

$$\underset{\sim}{g}: V^n \times V^n \to R \tag{6.1}$$

is <u>symmetric</u> if

$$\underset{\sim}{g}(\underset{\sim}{v}, \underset{\sim}{u}) = \underset{\sim}{g}(\underset{\sim}{u}, \underset{\sim}{v}). \tag{6.2}$$

A symmetric 2-cotensor is called a <u>quadratic form</u>. Its rank with respect to either argument has a common value called the <u>rank</u> of g. If the rank is n, g is <u>nonsingular</u>; otherwise, g is <u>singular</u>. If $\underset{\sim}{g}(\underset{\sim}{v}, \underset{\sim}{v}) > 0$ (< 0) for all $\underset{\sim}{v} \neq 0 \in V^n$, $\underset{\sim}{g}$ is <u>positive (negative) definite</u>; otherwise, g is <u>indefinite</u>. A basis $\underset{\sim}{e}_i$ can always be found such that $g_{ij} = g(e_i, e_j) =$ diag $(1, 1, \ldots, -1, -1, \ldots 0, 0)$. The diagonal matrix defined by $\underset{\sim}{g}$ in this way is called the <u>signature</u> of g. It is also called the <u>canonical form</u> of the matrix of components g_{ij}.

If $\underset{\sim}{g}$ is nonsingular, then the mapping (cf., the discussion in §2)

$$\underset{\sim}{g}^*: V^n \to V_n \tag{6.3}$$

defined by

$$\underset{\sim}{g}^*(\underset{\sim}{v})(\underset{\sim}{w}) = \underset{\sim}{g}(\underset{\sim}{v}, \underset{\sim}{w}) \tag{6.4}$$

is 1-1 and onto. In physical theories and in Riemannian geometry where a particular nonsingular quadratic form $\underset{\sim}{g}$ plays a central and dominant role, it is customary to identify the elements $\underset{\sim}{v} \in V^n$ and their images $\underset{\sim}{g}(\underset{\sim}{v})$ under $\underset{\sim}{g}$, and to regard them merely as different representations of the "same" vector. Here we shall denote the covector $\underset{\sim}{g}(\underset{\sim}{v})$ by $\underset{\sim}{v}^{\dagger}$ and the vector $\underset{\sim}{g}^{*-1}(\underset{\sim}{\alpha})$ by $\underset{\sim}{\alpha}^{\dagger}$. Also, while it is customary to denote the components of the inverse of $\underset{\sim}{g}^*$ by $g^{\alpha\beta}$, we shall denote $\underset{\sim}{g}^{*-1}$ by $\underset{\sim}{g}^{\dagger}$ and its components by $g^{\dagger\alpha\beta}$.

These are special case of the general rules of tensor notation concerning the "raising and lowering" of indices by transvection with a fundamental symmetric, nonsingular 2-cotensor and its inverse. More generally now, "raising" all the indices of an r-covector defines an isomorphism

$$\underset{\sim}{g}^*: \quad V_{[n^r]} \rightarrow V^{[n^r]} \qquad (6.5)$$

defined by

$$\alpha^{\dagger\, i_1 i_2 \ldots i_r} = g^{\dagger i_1 j_1} g^{\dagger i_2 j_2} \ldots g^{\dagger i_r j_r} \alpha_{j_1 j_2 \ldots j_r} \qquad (6.6)$$

$$\underset{\sim}{\alpha}^{\dagger} = \underset{\sim}{g}^*(\underset{\sim}{\alpha}) ,$$

and "lowering" all the indices of an r-vector defines the inverse transformation

$$g^{*-1}: \; V^{[n^r]} \rightarrow V_{[n^r]}, \tag{6.7}$$

and one has

$$\underset{\sim}{a}^{\dagger\dagger} = \underset{\sim}{g}^{*-1}(\underset{\sim}{a}^{\dagger}) = \underset{\sim}{a}. \tag{6.8}$$

Every non-singular quadratic form in V^n determines such an isomorphism between the spaces $V^{[n^r]}$ and $V_{[n^r]}$. The isomorphisms $\underset{r}{g^*}$ determined in this way by different quadratic forms are distinct, as are the duality transformations $\underset{\sim}{D}(i)$ and $\underset{\sim}{D}(i')$ determined by linearly independent sets $\underset{\sim}{e}_i$ and $\underset{\sim}{e}_{i'}$ not related by a unimodular transformation.

Let

$$g = (1/n!)\, E^{i_1 i_2 \cdots i_n} E^{j_1 j_2 \cdots j_n} g_{i_1 j_1} \cdots g_{i_n j_n} \tag{6.9}$$

denote the determinant of the quadratic form $\underset{\sim}{g}$. Note that the value of the determinant depends on the basis used to define $\underset{\sim}{E}$.

Two ordered linearly independent sets of vectors $\underset{\sim}{v}_i$ and $\underset{\sim}{v}_{i'}$ in V^n are said to have the same orientation if they are related by a transformation with positive determinant; otherwise, they are said to have opposite orientation. A V^n together with an ordered linearly independent set of vectors in it is an oriented n-dimensional vector space $\vec{V}^n$. $(V^n, \underset{\sim}{v}_i)$

and (V^n , $\underset{\sim}{v}_{i'}$) determine the same $\vec{V}^n$ provided $\underset{\sim}{v}_i$ and $\underset{\sim}{v}_{i'}$ have the same orientation; otherwise, they are regarded as different oriented V^n, say $\vec{V}^n$ and $\overleftarrow{V}^n$.

Consider now a $\vec{V}^n$ and set

$$\underset{\approx}{D} = (\epsilon\sqrt{|g|}\,)\ \underset{\sim}{D}(i) \tag{6.10}$$

where $\epsilon = +1$ if the set $\underset{\sim}{e}_i$ used to define the dual transformation has the same orientation as $\vec{V}^n$ and $\epsilon = -1$ otherwise. This definition of $\underset{\approx}{D}$ is independent of the basis used to define $\underset{\sim}{D}(i)$. One then has

$$\underset{\sim}{g}_r^{*-1}\ \underset{\approx}{D} = (-)^{r(n-r)}\ \frac{g}{|g|}\ \underset{\approx}{D}^{-1}\ \underset{\sim}{g}_r^{*}\ . \tag{6.11}$$

7. CHAINS AND COCHAINS

Let A^n denote an n-dimensional affine space with translations V^n. We call elements of V^n vectors and the elements of A^n points. Points are denoted by lightface, lower-case Latin letters p, q, ..., etc. We write

$$\underset{\sim}{v}(p) = p + \underset{\sim}{v}$$

for the image of p under the translation $\underset{\sim}{v}$, and

$$\underset{\sim}{v} = p - q$$

for the unique element of V^n defined by the pair of points (p, q). We say that the vector $\underset{\sim}{v}$ points from q to p.

An oriented r-simplex $s_r \subset A^n$ is determined by giving an ordered set of r+1 points $[p_0 p_1 p_2 \ldots p_r]$ called the vertices of s_r. The simplex s_r consists in the set of points p given by

$$p = p_0 + \sum_{i=1}^{r} a^i \underset{\sim}{v}_i, \quad 0 \leq a^i \leq 1, \quad \underset{\sim}{v}_i = p_i - p_0, \tag{7.1}$$

and it is assumed that the vectors $\{\underset{\sim}{v}_i\}$ are linearly independent. The orientation of s_r is given by the orientation of the set of r-vectors $\{\underset{\sim}{v}_i\}$.

The r-vector $\underset{\sim}{s}_r$ of s_r is defined by

$$s_r = \frac{1}{r!} v_1 \vee v_2 \vee \ldots \vee v_r . \tag{7.2}$$

A Euclidean space E^n is an affine space with a positive definite quadratic form $q(u, v)$ in the translation space V^n of E^n.

The <u>r-direction</u> d_r of an r-simplex $s_r \subset E^n$ is defined by

$$d_r = \frac{s_r}{|s_r|_q}, \qquad |s_r|_q = q_r(s_r, s_r), \tag{7.3}$$

where q_r is the quadratic form in $V^{[n^r]}$ whose definition has been given in §5. We call $|s_r|_q$, the <u>r-volume</u> of s_r.

The <u>center</u> p_c of an r-simplex (6.1) is the point defined by

$$p_c = p_0 + \left(\frac{1}{r+1}\right) \sum_{i=1}^{r} v_i . \tag{7.4}$$

An <u>r-cube</u> $t_r \subset A^n$ is the set of points given by

$$p = p_0 + \sum_{i=1}^{r} a^i v_i, \qquad 0 \leq a^i \leq 1, \tag{7.5}$$

where the r vectors v_i are linearly independent, and the orientation of t_r is the orientation determined by the set of r vectors $\{v_i\}$. The point $p_e = p_0 + \frac{1}{2} \Sigma v_i$ in (7.5) is the <u>center</u> of t_r. The r-vector of t_r is $t_r = v_1 \vee v_2 \vee \ldots \vee v_r$, its r-direction is $\underset{\sim}{d}_r = \underset{\sim}{t}_r / |\underset{\sim}{t}_r|_q$, and $|\underset{\sim}{t}_r|_q$ is its r-volume.

An <u>r-chain</u> $c_r \subset A^n$ is a linear combination of a finite number of nonoverlapping r-simplexes $s_{r\alpha}$, $\alpha = 1, 2, \ldots$ with real coefficients. Two r-simplexes c_r and c'_r are non-overlapping if $c_r \cap c'_r$ is the empty set of points in A^n, or the points of an s-simplex, $s < r$. By $0\, s_r$ is meant the empty set of points; by $-1\, s_r$ is meant the simplex comprising the same set of points as the simplex s_r, but with opposite orientation. If $c_r = \sum_\alpha c^\alpha s_{r\alpha}$, and $c'_r = \sum_\alpha c'^\alpha s_{r\alpha}$ are two r-chains in A^n, we define $\lambda c_r + \eta c'_r$ as the r-chain $\sum_\alpha (\lambda c^\alpha + \eta c'^\alpha) s_{r\alpha}$ so that the set C_r of all r-chains in A^n is a linear space.

An r-cochain F is a real-valued linear function, $F\colon C_r \to R$, of r-chains.

8. COCHAINS DEFINED BY INTEGRATION OF r-COVECTOR FIELDS

Let A^n be an affine space with translations V^n. A tensor field in A^n is a mapping $T: A^n \to V$ which assigns to each point $p \in A^n$ a tensor $\underset{\sim}{T}(p)$ in a tensor space V with carrier space V^n.

Let $\underset{\sim}{q}$ be an arbitrary positive definite quadratic form in V^n so that $|\underset{\sim}{T}|_q$ and $|p-q|_q = \underset{\sim}{q}(\underset{\sim}{v}, \underset{\sim}{v})$, $\underset{\sim}{v} = p-q$ define certain norms in V and A^n, respectively, in terms of which we may define the continuity of a tensor field T. If T is continuous with respect to the norm $|\ |_q$, then it is continuous with respect to any other norm $|\ |_{q'}$, defined in this way. We say this to emphasize that the considerations of this section are independent of the choice of the positive definite quadratic form $\underset{\sim}{q}$.

Set

$$\nabla \underset{\sim}{T}(p, v) = \lim_{t \to 0^+} \frac{\underset{\sim}{T}(p + tv) - \underset{\sim}{T}(p)}{t} ;$$

then, if the limit exists, ∇T is linear in $\underset{\sim}{v}$ and hence is a tensor with carrier space V^n; we denote it by $\nabla \underset{\sim}{T}(p)$ and call it the gradient of T at the point p. If $\nabla \underset{\sim}{T}(p)$ is defined for each $p \in A^n$, then ∇T is a tensor field in A^n. Let $R \subset A^n$ be a region of A^n. If T is continuous at each point $p \in R$, we say that the field T is continuous in R. If ∇T exists and is continuous at each point $p \in R$, we say that T is 1-smooth or smooth in R. If $\nabla\nabla T$ exists and is continuous in R, we say that ∇T is 2-smooth in R. Proceeding in this way, we define r-smooth tensor fields in regions R of A^n.

An r-covector field in A^n is a special case

$$\varphi: A^n \to V_{[n^r]} \tag{8.1}$$

of a tensor field defined more generally above. For brevity, we shall call an r-covector field in A^n, an r-form in A^n.

The integral of an r-form over an r-chain,

$$F(c_r, \varphi) = \int_{c_r} \varphi \tag{8.2}$$

is defined as follows. First of all, we set

$$F(c_r, \varphi) = F\left(\sum_\alpha a^\alpha s_{r\alpha}, \varphi\right) = \sum_\alpha a^\alpha \int_{s_{r\alpha}} \varphi, \tag{8.3}$$

so that it suffices to define the integral of φ over an r-simplex s_r. Now for every value of $\epsilon > 0$ an r-simplex s_r can be (subdivided) expressed as an r-chain of the form

$$s_r = \sum_{\beta=1}^{N(\epsilon)} s_{r\beta},$$

where each $s_{r\beta}$ has the same r-direction as s_r, and $\operatorname{diam}_q(s_{r\beta}) < \epsilon \operatorname{diam}_q(s_r)$, (q arbitrary). Let $S_1, S_2, \ldots$ be any sequence of such subdivisions of s_r such that $\epsilon_k \to 0$. For the subdivision S_k, let $p_{k\beta}$ be the center of the simplex $s_{r\beta}$. Set

$$\int_{s_r} \varphi = \lim_{k\to\infty} \sum_{\beta=1}^{N_k} \varphi(p_{k\beta}) \cdot s_{r\beta}. \tag{8.4}$$

An elegant proof that the limit (8.4) exists and is independent of the sequence $S_1, S_2, \ldots$ if φ is continuous in p is given by Whitney. It follows from the definitions (8.2) and (8.4) that every continuous r-form in A^n determines a unique r-cochain F and that $F(a\varphi + b\varphi') = aF(\varphi) + bF(\varphi')$. A cochain defined in this way is called a <u>continuous r-cochain.</u> If the r-form φ is s-smooth, then F is <u>s-smooth.</u>

A continuous r-form φ in $R \subset A^n$ is called <u>regular</u> in R if there exists a continuous (r+1)-form rot φ such that, for every simplex $s_{r+1} \subset R$

$$\int_{\partial s_{r+1}} \varphi = \int_{s_{r+1}} \operatorname{rot} \varphi, \tag{8.5}$$

where ∂s_{r+1} denotes the <u>boundary</u> of the simplex s_{r+1} <u>oriented</u>

as follows: If s_{r+1} has vertices $[p_0 p_1 \cdots p_{r+1}]$, then ∂s_{r+1} is the r-chain given by the sum of r-simplexes

$$\partial s_{r+1} = \sum_{t=0}^{r+1} (-)^t s_r(p_0 p_1 \cdots \hat{p}_{r+1-t}, \cdots, p_{r+1}), \quad (8.6)$$

where $s_r(p_0 p_1 \cdots \hat{p}_{r+1-t}, \cdots, p_{r+1})$ denotes the simplex with ordered vertices $p_0 p_1 \cdots p_{r+1}$ with p_{r+1-t} omitted.

With these definitions, the famous theorems of Gauss, Stokes, Kelvin, Poincaré, and others may be viewed as a special case of the

Divergence theorem: Every smooth r-form φ in $R \subset A^n$ is a regular r-form in R; moreover, for smooth φ,

$$\text{rot}\ \varphi = \partial \vee \varphi = (r+1)\ J_{r+1}(\nabla \varphi). \quad (8.7)$$

A proof of the divergence theorem is not difficult for simplexes and chains. Later, we shall consider r-forms in a smooth manifold, and the definition of $\int \varphi$ will be extended to smooth manifolds and smooth manifolds embedded in a smooth manifold. In this way we get a quick proof of the divergence theorem for a much wider class of regions in A^n.

A regular r-form φ is closed (irrotational) in R if rot $\varphi = 0$. Thus, in other words, an r-form φ is irrotational in R if and only if the continuous r-chain $F(\varphi)$ has the property

$$F(\partial s_{r+1}, \varphi) = \int_{\partial s_{r+1}} \varphi = 0 \quad \text{every } s_{r+1} \subset R.$$

An r-form φ is <u>circulation free</u> in R if there exists a a regular (r-1)-form π such that, throughout R, $\varphi = \text{rot}\,\pi$. The r-form π is called a <u>potential</u> of the circulation free r-form φ. Every circulation free r-form in an arbitrary region $R \subset A^n$ is irrotational,

$$\int_{\partial c_r} \varphi = \int_{\partial c_r} (\text{rot}\,\pi) = \int_{\partial\partial c_r} \pi = 0, \tag{8.8}$$

where we have used the property $\partial\partial c_r = 0$ of the boundary operator ∂. But it is not true that every irrotational r-form φ in an arbitrary region R is circulation free. (Let R be the annulus $p = p_0 + \underset{\sim}{v}$, $a \leq |\underset{\sim}{v}| \leq b$, $a > 0$, $b > 0$, in E^2 and let φ be the 1-form with components (0, 1) in every polar coordinate system for which p_0 is the origin.) Whitney has shown, however, that every irrotational r-form φ in a <u>star-shaped</u> region $R \subset A^n$ is circulation free in A^n, and he has given an explicit construction of a potential π for φ in R.

<u>The potential</u> π <u>of a circulation free r-form</u> φ <u>is not unique.</u> Clearly, if the r-form $\varphi = \text{rot}\ \pi$ in R, then $\varphi = \text{rot}\ \pi'$ in R also where $\pi' = \pi + \text{rot}\ \gamma$ where γ is any regular (r-2)-form.

9. CONTINUOUS r-COVECTOR FIELDS DEFINED BY CERTAIN r-COCHAINS

In the previous section, the continuous r-cochain $F(\varphi)$ was defined for every continuous r-form φ in A^n. In this section, it will be shown how every r-cochain F of a certain class determines a unique r-form $\varphi(F)$. The characterization of this class of cochains and the proof of the existence of $\varphi(F)$ are due to Whitney. I shall sketch here in some detail Whitney's work, for I feel that it has wide applications in continuum mechanics. A very special case of Whitney's theorem to be discussed below will be recognized by experts in continuum mechanics as a new and novel approach to the concept of stress and the existence of a stress tensor. In classical field theories, r-forms represent the most basic and primitive physical quantities. The electromagnetic field, the gravitational field, the charge and current fields, and the stress tensor (a vector-valued 2-form) are familiar examples. But the concept of a field (r-form) is sophisticated indeed (except, perhaps, a 0-form) for it carries with it the rather complicated notion of its r-direction at each point. I believe that the concept of a real valued linear function of r-simplexes or r-chains lies closer to physical intuition than the concept of an r-form.

Therefore, in the part of these lectures which concerns physical theory, the definitions of the basic physical quantities to occur will be given in terms of the values of r-cochains. To make contact with the more traditional view which introduces the gravitational field of force or the electromagnetic field as primitives we shall need the following results.

An r-cochain $F: C_r \to R$ is <u>semi-sharp</u> if

(a) For each bounded region $R \subset A^n$, there exists an N_R such that

$$|F(s_r)| \leq N_R \; |\underset{\sim}{s}_r|_q .$$

(b) For each point $p \in A^n$ and $\epsilon > 0$ there exists a ζ such that for any (r+1)-simplex s_{r+1} contained in the r-cube $U_\zeta(p)$ of diameter ζ and center p,

$$|F(\partial s_{r+1})| \leq \epsilon \, |\underset{\sim}{s}_{r+1}|_q .$$

(c) One may choose ζ in (b) such that for any r-simplex s_r and vector $\underset{\sim}{v} \in V^n$ (the translation space of A^n)

$$|F(T_{\underset{\sim}{v}} s_r - s_r)| \leq \epsilon \, |\underset{\sim}{s}_r| \text{ if } s_r \subset U_\zeta(p), \; |\underset{\sim}{v}| \leq \zeta .$$

In (c) $T_{\underset{\sim}{v}} s_r$ is the simplex consisting in the set of points of s_r translated by $\underset{\sim}{v}$ and having the same orientation as s_r.

Theorem (Whitney): If F is a semi-sharp r-cochain, there exists a continuous r-covector field φ such that

$$F(c_r) = \int_{c_r} \varphi.$$

In other words, every semi-sharp r-cochain is a continuous r-cochain.

The proof of the theorem rests in part on the following

Lemma (Whitney): Let φ be a real valued function of simple r-vectors such that

(1) φ is homogeneous of degree one:

$$\varphi(a\underset{\sim}{v}_r) = a\varphi(\underset{\sim}{v}_r), \quad \underset{\sim}{v}_r = \underset{\sim}{v}_1 \vee \underset{\sim}{v}_2 \vee \dots \vee \underset{\sim}{v}_r$$

(2) $\sum_{\beta=0}^{r+1} \varphi(\underset{\sim}{s}_{r\beta}) = 0,$ for every (r+1)-simplex s_{r+1}, where the boundary of $\underset{\sim}{s}_{r+1}$ is given by $\sum_{\beta=0}^{r+1} \underset{\sim}{s}_{r\beta}$.

(In words, the last condition reads, the sum of the values of φ on the r+1-oriented r-vectors of the faces of every (r+1)-simplex is zero.) Then there exists a unique r-covector $\underset{\sim}{\varphi}$ such that $\underset{\sim}{\varphi} \cdot \underset{\sim}{v}_r = \varphi(\underset{\sim}{v}_r)$.

In other words, every φ with properties (1) and (2) is linear in $\underset{\sim}{v}_r$ and defines a unique r-covector $\underset{\sim}{\varphi}$. The uniqueness of $\underset{\sim}{\varphi}$ is immediate. I present a somewhat simpler proof of the linearity of φ than given by Whitney.

Set

$$F(\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_r) = \varphi(\underset{\sim}{v}_1 \vee \underset{\sim}{v}_2 \vee \ldots \vee \underset{\sim}{v}_r).$$

Then, by (1), F is homogeneous of degree 1 in each argument. We show that it must be linear. Suppose $r = 1$, and consider a 2-simplex with faces having 1-directions, $\underset{\sim}{u}$, $-\underset{\sim}{v}$, and $\underset{\sim}{v}-\underset{\sim}{u}$. Then, by (2),

$$F(\underset{\sim}{u}) + F(\underset{\sim}{v}-\underset{\sim}{u}) + F(-\underset{\sim}{v}) = 0,$$

and using (1),

$$F(\underset{\sim}{v}-\underset{\sim}{u}) = F(\underset{\sim}{v}) + F(-\underset{\sim}{u}),$$

or, setting $\underset{\sim}{w} = -\underset{\sim}{u}$,

$$F(\underset{\sim}{v}+\underset{\sim}{w}) = F(\underset{\sim}{v}) + F(\underset{\sim}{w}),$$

which proves that F is linear if $r = 1$. The general case $r > 1$ is illustrated sufficiently by the case $r = 2$. Consider the 3-simplex s with $\underset{\sim}{v}_i = p_i - p_0$, $i = 1, 2, 3$, $p_0 p_1 p_2 p_3$ the vertices of s. It follows from (2) that

$$F(\underset{\sim}{v}_2-\underset{\sim}{v}_3, \underset{\sim}{v}_1-\underset{\sim}{v}_3) + F(\underset{\sim}{v}_1, \underset{\sim}{v}_2) + F(\underset{\sim}{v}_2, \underset{\sim}{v}_3) + F(\underset{\sim}{v}_3, \underset{\sim}{v}_1) = 0 \tag{9.1}$$

for all linearly independent v_1, v_2, v_3. In (9.1) replace the arguments v_1, v_2, v_3 by v_1, av_2, and av_3, respectively. Then, for every $a > 0$,

$$F(a(v_2-v_3), v_1-av_3) + F(v_1, av_2) + F(av_2, av_3) = 0. \quad (9.2)$$

Now use (1) to obtain

$$aF(v_2-v_3, v_1-av_3) + aF(v_1, v_2) + a^2F(v_2, v_3) + aF(v_3, v_1) = 0. \quad (9.3)$$

Divide (9.3) by a. The limit of the resulting expression for $a \to 0$ is

$$F(v_2-v_3, v_1) + F(v_1, v_2) + F(v_3, v_1) = 0. \quad (9.4)$$

Now using (1), the antisymmetry of F, and setting $-v_3 = w$, one gets

$$F(v_2+w, v_1) = F(v_2, v_1) + F(w, v_1),$$

which proves the linearity of F in its first argument. By antisymmetry, it is linear also in the second argument. Hence, F is an antisymmetric, multilinear function of two vectors; i.e., φ is a 2-covector.

Proof of the main theorem: Let F be a semi-sharp cochain so that (a), (b), and (c) hold. Choose any point p and an r-direction $\underset{\sim}{d}$. Let $s_1, s_2, \ldots$ be a sequence of r-simplexes containing p and having the common r-direction $\underset{\sim}{d}$; $\underset{\sim}{d} = \underset{\sim}{s}_i / |\underset{\sim}{s}_i|_q$, $i = 1, 2, \ldots$. Suppose that diam $(s_i) \to 0$. Set

$$\varphi(p, \underset{\sim}{d}) = \lim_{i \to \infty} F(s_i) / |\underset{\sim}{s}_i|_q . \tag{9.5}$$

Existence and uniqueness of the limit is proved as follows: Let s be any r-simplex with r-direction $\underset{\sim}{d}$ such that $s \subset U_\xi(p)$. Then one can choose an r-cube τ containing p such that $T_{\underset{\sim}{v}_k} \tau \subset s$, $k = 1, 2, \ldots, n$ and

$$\left| s - \sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau \right| \leq \epsilon_1, \tag{9.6}$$

for every $\epsilon_1 > 0$. Then, by property (a) of F,

$$\left| F\left(s - \sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau \right) \right| \leq N \left| s - T_{\underset{\sim}{v}_k} \tau \right|_q \leq N \epsilon_1 . \tag{9.7}$$

By property (c) of F, it follows that, for each value of k,

$$\left| F\left(T_{\underset{\sim}{v}_k} \tau - \tau \right) \right| < \epsilon |\tau|_q, \tag{9.8}$$

so that by summing (9.8) over the s values of k,

$$\left| F\left(\sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau - s\tau \right) \right| \leq \sum_{k=1}^{n} \left| F\left(T_{\underset{\sim}{v}_k} \tau - \tau \right) \right| \leq n\epsilon |\tau|_q \leq \epsilon |s|_q . \tag{9.9}$$

Therefore,

$$\left| F(s)|\tau| - F(\tau)|s| \right| = \left| F(s) - F\left(\sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau\right) |\tau| + F\left(\sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau\right) |\tau| - F(\tau)|s| \right|$$

$$\leq N\epsilon_1 |\tau| + \left| F\left(\sum_{k=1}^{n} T_{\underset{\sim}{v}_k} \tau\right) |\tau| - F(\tau)|s| \right|$$

$$\leq N\epsilon_1 |\tau| + |\tau| \left| F\left(\Sigma\, T_{\underset{\sim}{v}_k} \tau\right) - nF(\tau) \right| + \left| nF(\tau)|\tau| - F(\tau)|s| \right|$$

$$\leq N\epsilon_1 |\tau| + \epsilon |\tau|\ |s| + |\tau| N(n|\tau| - |s|)$$

$$\leq N\epsilon_1 |\tau| + |\tau|\ |s|\epsilon + N\epsilon_1 |\tau| .$$

Dividing by $|\tau|\ |s|$, one gets

$$\left| \frac{F(s)}{|s|} - \frac{F(\tau)}{|\tau|} \right| \leq 2\, \frac{\epsilon_1 N}{|s|} + \epsilon .$$

For each s, choose $\epsilon_1 = |s|\epsilon/(2N)$. Then,

$$\left| \frac{F(s)}{|s|} - \frac{F(\tau)}{|\tau|} \right| \leq 2\epsilon . \qquad (9.10)$$

The inequality (9.10) holds for every $s \subset U_\zeta(p)$; hence, for any pair s and $s' \subset U_\zeta(p)$

$$\left| \frac{F(s)}{|s|} - \frac{F(s')}{|s'|} \right| \leq 4\epsilon . \tag{9.11}$$

This proves the existence and uniqueness of the limit in (9.5). Using (9.11) it is also easy to see that $\varphi(p, \underset{\sim}{d})$ is continuous in the argument p:

$$|\varphi(p, \underset{\sim}{d}) - \varphi(p', \underset{\sim}{d})| = \left| \lim_{i \to \infty} \frac{F(s_i)}{|s_i|} - \lim_{j \to \infty} \frac{F(s_j')}{|s_j'|} \right| , \tag{9.12}$$

where s_i, $i = 1, 2, \ldots$ is a sequence of r-simplexes with r-direction $\underset{\sim}{d}$ each of which contains the point p and s_j', $j = 1, 2, \ldots$ is a sequence with r-direction $\underset{\sim}{d}$, each of which contains the point p'. We may choose $s_i = T_{\underset{\sim}{v}} s_i'$, where $\underset{\sim}{v} = p - p'$. Then (9.12) becomes

$$|\varphi(p, \underset{\sim}{d}) - \varphi(p', \underset{\sim}{d})| = \lim_{i = \infty} \left| \frac{F(T_{\underset{\sim}{v}} s_i') - F(s_i')}{|s_i'|} \right| . \tag{9.13}$$

Using property (c) of F we now get

$$|\varphi(p, \underset{\sim}{d}) - \varphi(p', \underset{\sim}{d})| < \epsilon, \quad \text{for all } |\underset{\sim}{v}| = |p - p'| \leq \xi, \quad s_i' \subset U_{\xi}(p). \tag{9.14}$$

From the definition of the Riemann integral $\int_s f(p)dp$ of a continuous function $f: A^n \to R$, it now follows that,

$$F(s) = \int_s \varphi(p, \underset{\sim}{d})dp \qquad (9.14)$$

for every simplex s with r-direction $\underset{\sim}{d}$. Cut s into simplexes $s_1, s_2, \ldots, s_n$ of diam $< \xi$. For each s_i let $p_i \in s_i$. Then

$$\left| F(s_i) - \int_{s_i} \varphi(p, \underset{\sim}{d})dp \right| \leq \left| F(s_i) - \varphi(p_i, \underset{\sim}{d}) \; |s_i| \right|$$
$$+ \left| \int_{s_i} [\varphi(p_i, \underset{\sim}{d}) - \varphi(p, \underset{\sim}{d})]dp \right|$$
$$\leq 5 \epsilon \; |s_i|;$$

hence,

$$\left| F(s) - \int_s \varphi(p, \underset{\sim}{d})dp \right| \leq 5 \epsilon \; |s|;$$

which proves (8.14).

To this point, only the properties (a) and (c) of F have been used. Property (c) is now invoked to prove that φ has the properties (1) and (2) of the Lemma. Now $\varphi(p, \underset{\sim}{d})$ has been defined only for r-directions $\underset{\sim}{d}$; i.e., only for simple r-vectors such that $|\underset{\sim}{d}| = 1$. Define $\varphi(p, \underset{\sim}{v})$ for all simple r-vectors $\underset{\sim}{v}$ by setting

$$\varphi(p, \underset{\sim}{v}) = |\underset{\sim}{v}| \; \varphi(p, v/|\underset{\sim}{v}|).$$

Since, by definition, $\varphi(p, \underset{\sim}{d}) = - \varphi(p, -\underset{\sim}{d})$,

$$\varphi(p, a\underset{\sim}{v}) = a\varphi(p, \underset{\sim}{v}) \tag{9.15}$$

for all real a. Hence, $\varphi(p, \underset{\sim}{v})$ satisfies the first condition of the lemma.

Let $p \in s$ be an interior point of an (r+1)-simplex and let s_λ be s contracted towards p by the factor λ and set $\partial s_\lambda = \sum_i s_{\lambda i}$, where the $s_{\lambda i}$ are the oriented faces of s_λ. Then

$$|\underset{\sim}{s}_\lambda| = \lambda^{r+1}|\underset{\sim}{s}|, \quad |\underset{\sim}{s}_{\lambda i}| = \lambda^r\,|\underset{\sim}{s}_i|.$$

Let $s_\lambda \subset U_\zeta(p)$ for $\lambda \leq \lambda_0$ and let $\underset{\sim}{d}_i$ be the r-direction of s_i. Then

$$\left|\sum_i \varphi(p, \underset{\sim}{s}_{\lambda i}) - \int_{s_\lambda} \varphi(p, \underset{\sim}{d})dp\right| = \left|\sum_i \int_{s_{\lambda i}} [\varphi(p, \underset{\sim}{d}_i) - \varphi(q, \underset{\sim}{d}_i)]dq\right|$$

$$\leq \sum_i \epsilon\, |\underset{\sim}{s}_{\lambda i}| = \epsilon \sum_i \lambda^r\, |\underset{\sim}{s}_i|$$

$$\leq \epsilon\, \lambda^r\, |\partial \underset{\sim}{s}|.$$

Also, by property (b),

$$\left|\int_{\partial s_\lambda} \varphi(p, \underset{\sim}{d})dp\right| = |F(\partial s_\lambda)| \leq \epsilon\, |\partial \underset{\sim}{s}_\lambda| \leq \epsilon\lambda^r |\partial \underset{\sim}{s}|.$$

Hence,

$$\left|\sum_i \varphi(p, \underset{\sim}{s}_{\lambda i})\right| \leq 2\epsilon\lambda^r |\partial \underset{\sim}{s}|$$

and dividing by λ^r we get

$$\left|\sum_i \varphi(p, \underset{\sim}{s}_i)\right| \leq 2\epsilon\, |\partial \underset{\sim}{s}|.$$

Since ϵ is arbitrary, $\varphi(p, \underset{\sim}{v})$ has property (2) of the lemma. It follows that φ is an r-covector for each value of p. Define $\underset{\sim}{\varphi}(p)$ by

$$\varphi(p, \underset{\sim}{v}_r) = \underset{\sim}{\varphi}(p) \cdot \underset{\sim}{v}_r, \tag{9.16}$$

and denote the corresponding r-covector field by φ (i. e., φ is a continuous r-form). Then

$$F(s) = \int_s \varphi(p, \underset{\sim}{d}) dp = \int_s \varphi \tag{9.17}$$

which is the assertion of the theorem.

<u>Remark on notation.</u> Let $d\underset{\sim}{s}(p)$ be an r-vector with r-direction $\underset{\sim}{d}_r$ such that the r-volume of any $s' \subset s$ with r-direction $\underset{\sim}{d}_r$ is given by $\int_{s'} |d\underset{\sim}{s}(p)|$. Then, the integral of the r-form φ is also denoted by

$$F(s) = \int_s \varphi = \int_s \underset{\sim}{\varphi}(p) \cdot d\underset{\sim}{s}(p) = \int_s \underset{\sim}{\varphi} \cdot d\underset{\sim}{s}. \tag{9.18}$$

Other expressions for F(s) are as follows: Let $\underset{\sim}{D}(i)$ and $\underset{\approx}{D}$ be the duality transformations defined with respect to an arbitrary n-covector and quadratic form $\underset{\sim}{q}$. The first of these is independent of q and the second depends only on the orientation of the n-covector used to define $\underset{\sim}{D}(i)$. Set

$$\underset{\approx}{D}\underset{\sim}{\varphi} = \underset{\sim}{\tilde{\varphi}}, \qquad \underset{\approx}{D}'\, d\underset{\sim}{s} = d\underset{\sim}{\tilde{s}},$$

$$\underset{\sim}{D}(i)\underset{\sim}{\varphi} = \underset{\sim}{\hat{\varphi}}, \quad \underset{\sim}{D}'(i)d\underset{\sim}{s} = d\underset{\sim}{\hat{s}}.$$

Then,

$$F(s) = \int_s \varphi = \int_s \underset{\sim}{\varphi}\cdot d\underset{\sim}{s} = (-)^{n(n-r)} \int_s \underset{\sim}{\hat{\varphi}}\cdot d\underset{\sim}{\hat{s}}$$

$$= (-)^{n(n-r)} \int_s \underset{\sim}{\tilde{\varphi}}\cdot d\underset{\sim}{\tilde{s}}. \tag{9.19}$$

The quantities $\underset{\sim}{\hat{\varphi}}$, $\underset{\sim}{\tilde{\varphi}}$, $d\underset{\sim}{\hat{s}}$, and $d\underset{\sim}{\tilde{s}}$ depend on a quadratic form and an n-covector. These could even be chosen as continuous functions of p. But the integral $\int_s \varphi$ is independent of any quadratic form or n-form $D(i, p)$; its value depends only on the r-form $\varphi: A^n \to V_{[n^r]}$ and on the oriented r-simplex s. To use the representations (9.19) masks this independence and the simplicity of the definition of $\int_s \varphi$. Nevertheless, many of the standard and traditional formulas of classical vector analysis rest upon the possibility of these alternative representations, and to exhibit the relation between some of the results in the physical theory involving r-forms and known classical results, it is necessary to introduce expressions like (9.19). In particular, in classical vector analysis, the (n-r-1)-forms defined by

$$\begin{aligned} \operatorname{div} \hat{\varphi} &= D \operatorname{rot} \varphi, \\ \operatorname{div} \tilde{\varphi} &= D \operatorname{rot} \varphi, \end{aligned} \tag{9.20}$$

are called the natural and absolute divergence of φ, respectively. The divergence theorem or the definition of a regular r-form Eq. (7.5) then appears in the following guises:

$$\int_{s} \underset{\sim}{\varphi} \cdot \underset{\sim}{ds} = \int_{s} \text{rot}\, \underset{\sim}{\varphi} \cdot \underset{\sim}{ds}, \tag{9.21}$$

$$= (-)^{n(n-r)} \int_{s} (\text{div}\, \underset{\sim}{\hat{\varphi}}) \cdot \underset{\sim}{d\hat{s}}, \tag{9.22}$$

$$\int_{s} \underset{\sim}{\hat{\varphi}} \cdot \underset{\sim}{d\hat{s}} = (-)^{n} \int_{s} (\text{div}\, \underset{\sim}{\hat{\varphi}}) \cdot \underset{\sim}{d\hat{s}}, \tag{9.23}$$

and in similar guises with $\underset{\sim}{D}$ replaced by $\underset{\approx}{D}$.

10. SMOOTH MANIFOLDS

In the physical theory considered in the lectures, it is not assumed that space or space-time is an affine space, and we require a definition of r-covector fields and a theory of integration of r-forms in a smooth manifold. We sketch here the theory as presented by Whitney.

An n-dimensional smooth manifold M^n is a mathematical system of the following sort. M^n is a connected topological space with open sets $U, U', \ldots$ together with a collection of <u>coordinate systems</u> χ_i, i in some index set. Each coordinate system is a homeomorphism

$$\chi_i: O_i \to M^n, \quad O_i \subset A^n \tag{10.1}$$

of an open set in A^n into M^n, A^n a fixed n-dimensional affine space. A finite or denumerable set of the <u>coordinate patches</u> $U_i = \chi_i(O_i)$ cover M^n. If $U_i \cap U_j \neq 0$, then $\chi_i^{-1} \circ \chi_j = \chi_{ij}$ is defined in some $O_{ij} \subset O_j$ and $\chi_{ij}: O_{ij} \to A^n$ is a mapping from an open set in A^n into some other open set in A^n. We require that the <u>gradient</u> of χ_{ij} defined by

$$\underset{\sim}{\nabla} \chi_{ij}(p, \underset{\sim}{v}) = \lim_{t \to 0+} \frac{\chi_{ij}(p + t\underset{\sim}{v}) - \chi_{ij}(p)}{t} \tag{10.2}$$

exist and be continuous in p throughout its domain. It can be shown that $\chi_{ij}(p, \underset{\sim}{v})$ is linear in $\underset{\sim}{v}$; hence, $\underset{\sim}{\nabla} \chi_{ij}(p, \underset{\sim}{v}) = \underset{\sim}{\nabla} \chi_{ij}(p) \cdot \underset{\sim}{v}$ and $\underset{\sim}{\nabla} \chi_{ij}(p)$ is a linear transformation in V^n, the translation space of A^n. The coordinate systems of a smooth manifold have the property that $\nabla \chi_{ij}$ has rank n at each p in the domain of χ_{ij}. The manifold M^n is s-smooth if each χ_{ij} is smooth (i.e., if $\nabla^s \chi_{ij}$ for all i and j, where defined, exists and is continuous).

If a subset of the U_i exists such that the U_i cover M^n and such that the Jacobians det $|\nabla \chi_{ij}|$ are all positive, then M^n is orientable and the set of coordinate systems χ_i orient M^n. Any coordinate system χ'_j related to one of the χ_i by a transformation with positive Jacobian is a preferred coordinate system of the oriented M^n.

If an s-smooth manifold M^n is defined in terms of coordinate systems χ_i, call any mapping $\chi = \chi_i \circ \varphi\colon O \rightarrow M^n$ obtained by composing any χ_i with an r-smooth, $r \geq s$, homeomorphism $\varphi\colon O \rightarrow O_i$, $O \subset A^n$, an admissible coordinate system for M^n. Henceforth, by a coordinate system of M^n, is meant any admissible coordinate system for M^n.

A mapping f: $M \rightarrow M'$ of one k-smooth manifold M into another M' is s-smooth, $s \leq k$, if $\chi_i^{-1} \circ f \circ \chi_j$, where defined, is s-smooth. When M' = R (real line), f is a real-valued s-smooth function in M; when M = R, or a connected open set in $O \subset R$, f is an <u>s-smooth parametrized curve</u> in M'.

Consider all the smooth curves in M^n defined by f, f', ... which contain a given point $x \in M^n$. We may assume, without loss in generality, that the domain of each f, f', ... contains the point $0 \in R$ and that f(0) = x. Let χ_i be a coordinate system of M^n such that $x \in U_i$. (We say that χ_i is a coordinate system about x.) Then $f_i = \chi_i^{-1} \circ f$ is defined near 0. Let

$$\dot{\underset{\sim}{f}}_i = \frac{df_i}{dt}(0). \qquad (10.3)$$

Then, by definition, $\dot{\underset{\sim}{f}}_i$ is a vector in V^n, the translation space of A^n. Call two smooth curves through x <u>equivalent</u> (in particular, <u>tangent</u>) if $\dot{\underset{\sim}{f}}_i = \dot{\underset{\sim}{f}}'_i$. Since $\dot{\underset{\sim}{f}}_i = \underset{\sim}{\nabla} \chi_{ij}(\dot{f}'_j)$, this definition of equivalence is independent of the coordinate system. By a vector $\underset{\sim}{v}(x)$ at x in M^n we mean an equivalence class of smooth parametrized curves in M^n which pass through x. We call the vector $\dot{\underset{\sim}{f}}_i \in V^n$ a <u>representation</u> of the corresponding set of equivalent curves $\underset{\sim}{v}(x)$. The sum $\underset{\sim}{v}(x) + \underset{\sim}{u}(x)$

and the multiple $\lambda \underset{\sim}{v}(x)$ of vectors at x in M^n is defined by addition and multiplication of their representations in V^n. From the linearity of the law of transformation of the representations corresponding to two coordinate systems, we see that the definition of $\underset{\sim}{v}(x) + \underset{\sim}{u}(x)$ and of $\lambda \underset{\sim}{v}(x)$ is independent of coordinate system. The set of all vectors $\underset{\sim}{v}(x)$ at x in M^n forms an n-dimensional vector space $V^n(x)$, the <u>tangent space of M^n at x.</u> Each coordinate system χ_i of M^n about x defines an isomorphism

$$\underset{\sim}{\nabla} \chi_i(p): \; V^n \to V^n(x), \quad p = \chi_i^{-1}(x), \tag{10.4}$$

where $\underset{\sim}{\nabla} \chi_i(p, \underset{\sim}{\dot{f}}_i) = \underset{\sim}{v}(x)$, $\underset{\sim}{v}(x)$ the equivalence class of smooth curves having the representation $\underset{\sim}{\dot{f}}_i$ in the coordinate system χ_i. The n-dimensional vector spaces $V^n(x)$ and $V^n(x')$, $x \neq x'$, though isomorphic, are distinct. In general, $\underset{\sim}{v}(x) + \underset{\sim}{u}(x')$ is not defined. One could define $\underset{\sim}{v}(x) + \underset{\sim}{u}(x')$ by adding their representations in some coordinate system χ_i, provided x and x' were both in U_i; but such a definition is not independent of χ_i.

A mapping

$$f: \; M^m \to M^n, \qquad m \leq n \tag{10.5}$$

of one smooth manifold into another is regular if the following holds. Let $x = f(X)$, and let χ_i be a coordinate system of M^m about X, χ_α a coordinate system of M^n about x. Then $f_{\alpha i} = \chi^{-1}_\alpha \circ \chi_i$ maps some $R \subset A^m$ into A^n. Let

$$\nabla f_{\alpha i}(p, \underset{\sim}{v}) = \nabla f_{\alpha i}(p) \cdot \underset{\sim}{v} = \lim_{t \to 0^+} \frac{f_{\alpha i}(p+\underset{\sim}{v}t) - f_{\alpha i}(p)}{t}, \qquad (10.6)$$

$p \in A^m$, $\underset{\sim}{v} \in V^m$. Then, by definition, $\underset{\sim}{\nabla} f_{\alpha i}(p)$ is a linear transformation of V^m into V^n. The mapping f is regular at X if $\nabla f_{\alpha i}$ is continuous at $p = \chi_i^{-1}(X)$ and has rank m. This definition of regularity at X is independent of the coordinate systems χ_i and χ_α. f is regular in an open set $Q \subset M^m$ if it is regular at each $X \subset Q$; it is regular if it is regular for all $X \in M^m$. Now in M^m and M^n, the coordinate systems χ_i and χ_α establish the isomorphisms $\nabla\chi_i$ and $\nabla\chi_\alpha$ between V^n and $V^n(x)$, and between V^m and $V^m(X)$, respectively. The gradient of the mapping ∇f is then defined by

$$\nabla f = \nabla\chi_\alpha \circ \nabla f_{\alpha i} \circ (\nabla\chi_i)^{-1}. \qquad (10.7)$$

The gradient of f is defined for every $X \in M^n$, and its definition is independent of coordinate systems. Its value at a point X of M^n defines a linear transformation $\nabla f(X)$: $V^m(X) \to V^n(x)$, $x = f(X)$, of the tangent space of M^m at X into the tangent space of M^n at the point $f(X)$.

11. INTEGRATION IN MANIFOLDS

A tensor whose carrier space is the tangent space $V^n(x)$ at a point x of a smooth manifold M^n is called a tensor at x. A tensor field in a manifold is a function

$$T:\ M^n \to \left\{ \overset{k}{\underset{n}{V^n}}{}_{\ell}(x);\quad x \in M^n \right\}$$

such that $T(x) \in \overset{k}{\underset{n}{V^n}}{}_{\ell}(x)$. (The values of k and ℓ are the same for all x; and $T(x)$ is a tensor at x.) The field T is continuous or s-smooth if its representation $\overset{*}{\chi}_i(T)$ defined by

$$\overset{*}{\chi}_i(T)(p, \underset{\sim}{v}_1, \underset{\sim}{v}_2, \dots, \underset{\sim}{\alpha}^1, \underset{\sim}{\alpha}^2, \dots) =$$

$$T(x, \underset{\sim}{\nabla}\chi_i \circ \underset{\sim}{v}_1, \underset{\sim}{\nabla}\chi_i \circ \underset{\sim}{v}_2, \dots, \underset{\sim}{\nabla}\chi_i^{-1} \cdot \underset{\sim}{\alpha}{}^{1}_{2}, \dots), \tag{11.1}$$

$p = \chi_i^{-1}(x)$, $\quad v_k \in V^n$, $\quad \alpha^k \in V_n$, in any coordinate system χ_i of M^n is continuous or s-smooth. This definition of smoothness is independent of the coordinate system in an r-smooth manifold provided $s \leq r$.

The gradient ∇f of a regular mapping $f:\ M^m \to M^n$ considered in the previous section is an example of the more general concept of a multi-point tensor. We may view ∇f as a bilinear function

$$\nabla f:\ V^m(X) \times V_n(x) \to R.$$

It is a mixed tensor field over M^m. In general, it is not a field in M^n because only in special subcases of $m = n$ is ∇f defined at every point of M^n.

Let φ be an r-cotensor field in M^n and let $f: M^m \to M^n$ be regular. Then we define a corresponding r-cotensor field $f^*(\varphi)$ in M^m by

$$(f^*\varphi)(X, \underset{\sim}{U}_1, \underset{\sim}{U}_2, \ldots, \underset{\sim}{U}_r) = \varphi(x, \underset{\sim}{\nabla} f \underset{\sim}{U}_1, \underset{\sim}{\nabla} f \underset{\sim}{U}_2, \ldots, \underset{\sim}{\nabla} f \underset{\sim}{U}_r), \quad (11.2)$$

where each $U_k \in V^m(X)$, $x = f(X)$, and $\underset{\sim}{\nabla} f \underset{\sim}{U}_k = \nabla f(X, \underset{\sim}{U}_k) \in V^n$.

Now $\int_s \varphi$ has been defined only for simplexes and linear combinations of simplexes in an affine space A^n. We wish to extend this definition to curvilinear simplexes and certain open sets in a manifold. This can be done following Whitney. First of all, consider an n-form φ in A^n. Say that φ is summable over the open set $R \subset A^n$ if there exists a polyhedron P with the property that for every $\epsilon > 0$

$$\left| \int_Q \varphi \right| < \epsilon \quad \text{for all polyhedra } Q \subset R - P. \qquad (11.3)$$

Then if φ is summable over R, there exists a number $\int_R \varphi$ such that

$$\left| \int_Q \varphi - \int_R \varphi \right| < \epsilon \quad \text{if } P \subset Q \subset R. \qquad (11.4)$$

This defines $\int_R \varphi$ for certain φ and open sets $R \subset A^n$. If, with respect to an arbitrary positive definite (metric) in the V^n or A^n, $R = \sup\{Q\colon\ Q \subset R,\ Q \text{ a polyhedron}\} < \infty$ and $|\varphi| < \infty$, then φ is summable over R and, therefore, $\int_R \varphi$ is defined. We need also the fact that if φ is summable over R, then so also is $\pi\varphi$ where π is any bounded continuous real-valued function in R.

The next result we need is the <u>transformation formula.</u> <u>Let</u> f <u>be any one-one regular mapping of the open set</u> $R \subset A^n$ <u>onto the open set</u> $R' \subset A'^n$ <u>with</u> $\overline{J}_f(p) > 0$ <u>in</u> R. <u>Let</u> φ <u>be a continuous n-form in</u> R' <u>summable over</u> R'. <u>Then</u> $f^*\varphi$ <u>is summable over</u> R, <u>and</u>

$$\int_R f^*\varphi = \int_{R'} \varphi . \qquad (11.5)$$

Next we define $\int_{M^n} \varphi$ for any compact smooth oriented manifold M^n. Let $\operatorname{spt}(\varphi)$ denote the closure of the set of points $x \in M^n$ where $\varphi(x) \neq 0$. If $\operatorname{spt}(\varphi) \subset U_i$ where U_i is a coordinate patch of M_n, χ_i a preferred coordinate system of M^n, then we set

$$\int_{M^n} \varphi = \int_{O_i} \chi_i \varphi, \quad \operatorname{spt}(\varphi) \subset \chi_i(O_i). \qquad (11.6)$$

That this definition of $\int_{M^n} \varphi$ is independent of the coordinate system follows from the transformation formula (11.5)

Suppose next that $\mathrm{spt}(\varphi)$ is not contained in any U_i. In this case, let $\sum_i \pi_i$ be a partition of unity with the following property. Let the finite set of coordinate patches U_i, $i = 1, 2, \ldots, N$ cover M^n, where the χ_i orient M^n. Then one can construct a set of real-valued smooth functions π_i in M^n such that $\pi_i(x) = 0$, $x \notin U_i$, and $\sum_i \pi_i(x) = 1$, $x \in M^n$. Express $\varphi(x)$ as $\sum_i \varphi_i(x)$, $\varphi_i(x) = \pi_i(x)\varphi(x)$. Then $\int_{U_i} \varphi_i$ is defined as above, and we set

$$\int_{M^n} \varphi = \sum_i \int_{U_i} \varphi_i. \tag{11.7}$$

The definition (11.7) can be shown to be independent of the partition of unity $\Sigma\, \pi_i$ and the coordinate systems χ_i.

Now let R be any open subset of an oriented manifold M^n such that $\overline{R}$ is compact. Let φ be any continuous n-form defined in a neighborhood of $\overline{R}$. Then there exists a finite number of coordinate patches U_i which cover $\overline{R}$, χ_i preferred coordinate systems of M^n. Define the π_i as above. Then, for some neighborhood U of $\overline{R}$, $\sum_i \pi_i(x) = 1$, $x \in U$. Set $\varphi_i = \pi_i \varphi$ and define

$$\int_R \varphi_i = \int_{\chi_i^{-1}(R)} \chi_i \varphi_i = \int_R \varphi = \sum_i \int_R \varphi_i . \tag{11.8}$$

Again, it can be shown that this definition of $\int_R \varphi$ is independent of the coordinate systems χ_i and of the partition of unity $\Sigma\, \pi_i$.

We record here Whitney's version of the divergence theorem in manifolds. For the complete definition of a standard n-manifold, I refer the reader to Whitney's book. It can be remarked that every smooth image of a polyhedron in A^n is a standard n-manifold, and every regular region in the sense of Kellogg is a standard n-manifold. To state the theorem we need at least the following partial description of a standard n-manifold. There is a connected compact topological space, $\overline{M}$, a closed subset $\partial\overline{M}$ of $\overline{M}$ (call it the boundary of M) and a closed subset $\partial_0\overline{M}$ of $\partial\overline{M}$ (call it the edges and vertices of $\partial\overline{M}$). $\overline{M} - \partial\overline{M}$ is a smooth oriented n-dimensional manifold, and $\partial M - \partial_0 M$ is a finite collection of oriented (n-1)-dimensional smooth manifolds. At each point of $x \in (\partial M - \partial_0 M)$ there is defined an outward normal vector $\underset{\sim}{v}(x) \neq 0$, which, by continuity, is defined at points of $\overline{M} - \partial\overline{M}$ near x. The orientation of the smooth manifold $\overline{M} - \partial\overline{M}$ is fixed by the ordered set $\{v(x), v_1(x), \dots, v_{n-1}(x)\}$ of vectors, where the last $n-1$ elements determine the orientation of the part of $\partial\overline{M} - \partial_0\overline{M}$ containing the point x. Let φ be an n-form such that

(a) φ is defined, continuous, and bounded in $\partial\overline{M} - \partial_0\overline{M}$, and is regular in $\overline{M} - \partial\overline{M}$,

(b) φ is summable over $\partial\overline{M} - \partial_0\overline{M}$,

(c) rot φ is summable over $\overline{M} - \partial\overline{M}$.

Then

$$\int_{\partial\overline{M}-\partial_0\overline{M}} \varphi = \int_{\overline{M}-\partial\overline{M}} \text{rot } \varphi. \tag{11.9}$$

Henceforth, when the meaning is clear from the context, we shall abbreviate (11.9) to read

$$\int_{\partial\overline{M}} \varphi = \int_{\overline{M}} \text{rot } \varphi. \tag{11.10}$$

When φ is smooth and M is 2-smooth, rot φ is given by rot $\varphi(x) = \chi_i^* \partial V \varphi$, where χ_i is any coordinate system about x and $\partial V \varphi$ is defined in (7.7).

Let $f\colon M^r \to M^n$ be any regular mapping of an oriented r-dimensional smooth manifold into an n-dimensional smooth manifold and let φ be an r-form in M^n. Then $f^*\varphi$ is an r-form in M^r and we set

$$\int_{R'} \varphi = \int_R f^*\varphi, \qquad R' = f(R). \tag{11.11}$$

This defines the integral of an r-form φ in M^n, $r \leq n$ over every r-dimensional smooth manifold or piece R' of a smooth manifold in M^n. If $\overline{R}' = f(\overline{R})$, and $\overline{R} \subset M^r$ is a standard manifold, then

$$\int_{\partial \overline{R}'} \varphi = \int_{\overline{R}'} \operatorname{rot} \varphi \tag{11.12}$$

provided the (r-1)-form φ is continuous in some neighborhood of $\overline{R}'$, $f^*\varphi$ is regular in R, $f^*\varphi$ is summable over $\partial\overline{R}-\partial_0\overline{R}$, and $f^* \operatorname{rot} \varphi$ is summable over R.

12. CURVILINEAR r-COCHAINS

Let M^n be an n-dimensional smooth manifold with coordinate systems $\chi_i : O_i \to U_i \subset M^n$. If $s \subset O_i$ is an r-simplex in A^n, we call $\chi_i(s)$ a <u>curvilinear r-simplex</u> in M^n. A <u>curvilinear r-chain</u> in M^n is an expression of the form $c = \Sigma a^i s_i$, where each s_i is a curvilinear simplex in M^n, $s_i \cap s_j$ = sum of chains of lower dimension, a^i are positive or negative integers or zero. We drop the adjective "curvilinear" when the meaning is clear without it. The sum of r-chains in M^n is defined in the same way as the sum of r-chains in A^n. An <u>r-cochain</u> in M^n is a linear function of r-chains in M^n. If $c = \Sigma a^i s_i$, then $\int_{s_i} \varphi$ is defined as in the last section and we set

$$\int_c \varphi = \Sigma a^i \int_{s_i} \varphi. \tag{12.1}$$

The value of $\int_c \varphi$ is independent of its representation as a sum of r-simplexes. Thus, every r-form φ in M^n, summable over every curvilinear r-simplex in M^n, determines a unique r-cochain in M^n by (12.1). If the (r-1)-form φ is regular in a neighborhood of c, then

$$\int_{\partial c} \varphi = \int_c \operatorname{rot} \varphi, \tag{12.2}$$

and if φ is smooth, then rot $\varphi = \partial \vee \varphi = \chi_i(\partial \vee \chi_i^{-1}\varphi)$. (Cf., §8 for the definition of $\partial \vee \gamma$, γ an r-form in A^n.) By the triangulation theorem for smooth manifolds, every compact oriented manifold can be expressed as a curvilinear r-chain of the form $M^r = \Sigma s_i$. Thus, for any compact r-dimensional manifold M^r in M^n and any regular (r-1)-form φ in M^n, we have a relation like (12.2) with c replaced by M^r. We say that an r-cochain defined by integration of an r-form φ in M^n is continuous or s-smooth accordingly as the r-form φ is continuous or s-smooth.

The three properties (a), (b), and (c) of §9 which characterize a semi-sharp r-cochain in an affine space A^n do not have an immediate invariant significance for r-cochains in a manifold. But since they are purely local conditions, using the coordinate systems of a manifold, the properties can be stated in terms of the inverse images $\chi_i^{-1}(s) = s'$ of curvilinear r-simplexes contained in some $U_i \subset M^n$ and of the corresponding values of the r-cochain. Thus, if F is an r-cochain in M^n and we set $F(s) = F'(s')$, $s' = \chi_i^{-1}s$, then F is called semi-sharp if and only if F' is semi-sharp. It follows that semi-sharp r-cochains and continuous r-cochains in a manifold are in 1-1 correspondence. Every semi-sharp r-cochain in M^n determines a unique continuous r-form φ in M^n such that $F(c) = \int_c \varphi$ and, conversely, every continuous r-form φ in M^n defines a unique semi-sharp r-cochain by this same rule of association.

13. MOTIONS AND INVARIANT COCHAINS

Let M^m and M^n denote smooth manifolds of the dimensions indicated. Let $O \subset R$ denote a connected open set of real numbers. Then a regular mapping

$$f: M^n \times O \to M^m, \quad m \geq n+1 \tag{13.1}$$

is called a parametrized motion of M^n in M^m. The trajectory of a point $X \in M^n$ is the smooth parametrized curve $f_X: O \to M^m$, $f_X(t) = f(X, t)$. The orbit of X is the set of points $\{x;\ x = f(X, t),\ t \in O\}$.

When $m = n+1$, define a motion of M^n as a smooth mapping

$$f': M^{n+1} \to M^n \tag{13.2}$$

such that $\nabla f'$ has rank n in M^{n+1}. The orbit of a point $X \in M^n$ is the set of points $\{x;\ f'(x) = X\}$. Using the implicit function theorem, it can be shown that every connected piece of an orbit is a smooth parametrized curve in M. Thus, when $m = n+1$, there is no distinction locally between motions and parametrized motions of M^n in M^{n+1}. But the orbits of a motion may not be connected so that motions and parametrized motions in the large are not in 1-1 correspondence. Moreover, the existence of motions depends on the character of the manifolds M^n and M^m. In the following we shall be concerned only with local properties of motions and parametrized motions only.

Assume now that $m = n + 1$. Then (12.1) has a unique smooth inverse f^{-1}: $M^{n+1} \to M^n \times O$ which consists in two smooth mappings f': $M^{n+1} \to M^n$ and T: $M^{n+1} \to O$, where $\nabla f'$ has rank n throughout its domain and ∇T has rank 1. Thus, every parametrized motion $(m = n + 1)$ determines a unique motion f', but not conversely. Two parametrized motions with the same orbits determine one and the same motion by this construction. Through each point $x \in U = f(M^n \times O)$ in the range of f there passes one and only one trajectory. Let $\underset{\sim}{v}(x) \in V^n(x)$ denote the tangent vector of the trajectory through x. From the regularity assumption, the vector field v with these values in U is continuous. We call this vector field in U, the <u>velocity field of the parametrized motion</u> f. If f': $M^{n+1} \to M^n$ is a motion, then at each point x in the domain of f', $\nabla f'$ has one and only one linearly independent proper vector $\underset{\sim}{w}(x) \neq 0$ such that

$$\underset{\sim}{\nabla} f'(x) \cdot \underset{\sim}{w}(x) = 0. \qquad (13.3)$$

The velocity field of a motion is the equivalence class of smooth vector fields in the domain of f' such that $\underset{\sim}{v}(x) \neq 0$ is a proper vector of $\nabla f'(x)$ with proper value zero at each point.

Let f_t: $M^n \to M^{n+1}$ be defined by $f_t(X) = f(X, t)$. Then

$$T_\tau = f_{t+\tau} \circ f_t^{-1}, \quad t,\ t+\tau \in O \tag{13.4}$$

is a 1-1 regular mapping

$$T_\tau: U_t \to U_{t+\tau} \tag{13.5}$$

of an open set $U_t \subset M^{n+1}$ onto an open set $U_{t+\tau} \subset M^{n+1}$. We have

$$T_0 = 1 \text{ (the identity)}, \quad T_\tau \circ T_{\tau'} = T_{\tau+\tau'}. \tag{13.6}$$

Let s_t be any curvilinear r-simplex in U_t. Then $s_{t+\tau} = T_\tau(s_t)$ is a curvilinear r-simplex in $U_{t+\tau}$. Let φ be an r-form in M^{n+1}. Then

$$F(s_t, \tau) = \int_{s_{t+\tau}} \varphi \tag{13.7}$$

is defined for all $s_t \subset U_t$. If $F(s_t, \tau)$ is independent of τ for all such s_t, we say that φ is an <u>invariant r-form under the motion</u> f, or that the corresponding r-cochain $F(s_t)$, defined by integration of φ is an <u>invariant r-cochain under the motion</u> f.

<u>Theorem</u>: Let $F(s_t, \tau)$ be defined by (13.7). Let v be the velocity field of the parametrized motion f. Then, if the r-form φ and the (r-1)-form $v \wedge \varphi$ are regular in the range of f, $F(s_t, \tau)$ is differentiable in τ for each s_t and

$$\frac{dF(s_t, \tau)}{d\tau} = \int_{s_{t+\tau}} [v \wedge \operatorname{rot} \varphi + \operatorname{rot}(v \wedge \varphi)]. \qquad (13.8)$$

To prove (13.8) consider the value of

$$\Delta F = F(s_t, \tau + \tau') - F(s_t, \tau) = \int_{s_{t+\tau+\tau'}} \varphi - \int_{s_{t+\tau}} \varphi .$$

The quantity ΔF differs from the integral of φ over the boundary $\partial(s_{t+\tau} \times I)$, $I = \{t;\ \tau \leq t \leq \tau + \tau'\}$, by an integral of φ over $\partial s_{t+\tau} \times I$:

$$\Delta F = \int_{\partial(s_{t+\tau} \times I)} \varphi + \int_{(\partial s_{t+\tau}) \times I} \varphi$$

$$= \int_{s_{t+\tau} \times I} \operatorname{rot} \varphi + \int_{(\partial s_{t+\tau}) \times I} \varphi ,$$

where we have used the regularity of φ. But the integrals on the right can be expressed as iterated integrals,

$$\Delta F = \int_I d\tau'' \int_{S_{t+\tau}} v \wedge (\text{rot } \varphi) + \int_I d\tau'' \int_{\partial S_{t+\tau}} v \wedge \varphi$$

$$= \int_I d\tau'' \int_{S_{t+\tau}} v \wedge \text{rot } \varphi + \int_I d\tau'' \int_{S_{t+\tau}} \text{rot } (v \wedge \varphi), \qquad (13.9)$$

where we have used the regularity of $v \wedge \varphi$. Thus, dividing (13.9) by τ'' and taking the lim $\tau'' \to 0$ yields the formula (13.8).

It follows that if φ and $v \wedge \varphi$ are regular, then φ and the r-cochain $F = \int \varphi$ are invariant under the parametrized motion f with velocity field v if and only if

$$\mathcal{L}_v \varphi = v \wedge \text{rot } \varphi + \text{rot } (v \wedge \varphi) = 0, \qquad (13.10)$$

at each point in the range of f. The continuous r-form $\mathcal{L}_v \varphi$ is called the Lie derivative of φ with respect to the vector field v. Let $\underset{\sim}{D}(x)$ be the duality transformation defined by any linearly independent set of vectors at x. Set $\hat{\varphi} = \underset{\sim}{D}(\varphi)$ and $\text{div } \hat{\varphi} = \underset{\sim}{D}(\text{rot } \varphi)$. The dual of $\mathcal{L}_v \varphi$ is then given by

$$\underset{\sim}{D}(\mathcal{L}_v \varphi) = (\text{div } \hat{\varphi}) \times v + \text{div } (\hat{\varphi} \times v). \qquad (13.11)$$

If the vectors used to define $\underset{\sim}{D}$ are the tangent vectors to the coordinate curves of a coordinate system, then rot φ and div φ are given by the differentiation formulas $\partial \wedge \varphi$ and $\partial \wedge \hat{\varphi}$

provided that φ is smooth. But the formulas (13.8) and (13.11) hold under the weaker hypotheses made in deriving them.

14. PHYSICAL UNITS, PHYSICAL DIMENSIONS, AND PHYSICAL QUANTITIES

Weyl has remarked:

"However, not only points are required to be represented by reproducible symbols, but also every other kind of geometric entity, and when passing to physics all sorts of physical quantities like velocities, forces, field strengths, wave functions, and what not, expect a similar symbolic treatment. One often acts as though once the points have been submitted to it by fixing a frame of reference for them, all these other things will follow suit without necessitating further provisions. This is certainly not true; at least further units of measurement have to be fixed at random so as to make the scheme of reference complete."

For the purposes of these lectures, the following formal definitions will be adopted. For a discussion of the historical use and development of the concepts and mathematics of physical units and dimensions, see Ericksen's Appendix on Tensor Fields in The Classical Field Theories, Vol. III/1, Handbuch der Physik.

A physical unit $\underset{\approx}{U}$ is a vector different from zero in a one-dimensional vector space U called a <u>unit space</u>. Thus, any physical unit $\underset{\approx}{U} \in U$ is a basis for U. A function $f: U \rightarrow L$

with values in some linear space is said to have <u>physical dimension</u> $[\underline{\underline{U}}^n]$ if

$$f(a\underset{\approx}{U}) = a^{-n} f(\underset{\approx}{U}); \qquad (14.1)$$

i.e., if f is a homogeneous function of degree -n. More generally now, if $f: U_1 \times U_2 \times \dots \times U_n \to L$ is homogeneous of degree $-n_k$ in the argument U_k, we write

$$\text{phys. dim. } f = \left[\underline{\underline{U}}_1^{n_1} \, \underline{\underline{U}}_2^{n_2} \cdots \underline{\underline{U}}_k^{n_k}\right]. \qquad (14.2)$$

In these lectures we restrict the use of the term <u>physical quantity</u> and use it only to mean the following. A physical quantity is an r-cochain in some manifold M^n with values which are homogeneous functions of a set $\underset{\approx}{U}_a$, $a = 1, 2, \dots, N$ of independent physical units. Thus, if F is a physical quantity, it has a definite physical dimension, and if F is semi-sharp or continuous, there exists an r-form in M^n such that

$$F(c, \underset{\approx}{U}_1, \dots, \underset{\approx}{N}_N) = \int_c \varphi(\underset{\approx}{U}_1, \dots, \underset{\approx}{U}_N), \qquad (14.3)$$

and we write

$$\text{phys. dim. } F = \text{phys. dim. } \varphi = \left[\underline{\underline{U}}_1^{n_1} \cdots \underline{\underline{U}}_N^{n_N}\right], \qquad (14.4)$$

with suitable values for the exponents $n_1, \dots, n_N$.

In particular, if phys. dim. $F = [\underline{\underline{U}}]$, we write,

$$F(c, \underset{\approx}{U}) = \int_c \varphi(\underset{\approx}{U}) . \tag{14.5}$$

The real number $F(c, \underset{\approx}{U})$ is called the amount of the physical quantity F in c measured in units $\underset{\approx}{U}$; $\varphi(\underset{\approx}{U})$ is called the density of the physical quantity F measured in units $\underset{\approx}{U}$.

Our use of the term "physical quantity" in these lectures lies closest to the concept and definition of an "extensive" quantity as that term is used in the literature of thermostatics.

LECTURE I

ELECTRIC CHARGE AND MAGNETIC FLUX

"In the application of mathematics to the calculation of electrical quantities, I shall endeavour in the first place to deduce the most general conclusions from the data at our disposal, and in the next place to apply the results to the simplest cases that can be chosen."

J. C. Maxwell

I.1. INTRODUCTORY REMARKS

The electrodynamics of elastic media finds important engineering applications in the phenomena of photoelasticity and piezoelectricity. The classical theories of these effects and of many others are embraced as special cases by the general theory of the electromagnetic field in material media to be presented in these lectures. The lectures will stress an orderly introduction of the physical concepts and laws of nature which are common to a broad class of special theories

We begin with a mathematical theory of electricity and magnetism and only later introduce the mechanical and kinematical concepts of length, time, force, stress, energy, and momentum. It is common knowledge that modern concepts of length, time, and simultaniety had their orgins in the theory of electricity and magnetism and that these concepts are not entirely consistent with the Newtonian viewpoint. One objective of the treatment of the electromagnetic field given here is to trace again these origins.

I. 2. ELECTRIC CHARGE AND MAGNETIC FLUX

It is possible to understand the classical equations of Maxwell and Lorentz governing the electric, magnetic, charge, and current fields in terms of only three primitive concepts:

1) the distribution of magnetic flux
2) the distribution of electric charge
3) the universal relation between these two distributions

For this purpose we begin by taking the concept of an <u>event</u> as primitive and undefined, just as a point is primitive and undefined in Euclidean geometry. We call the set of all events, <u>event-space</u>, and denote it by $\mathcal{E}$. We lay down the first of some eighteen assumptions or principles to be introduced in the course of the lectures.

A1: The set of all events is an orientable 4-dimensional smooth manifold.

Events will be denoted by lower case Greek letters ξ , ξ', ξ'', etc. The coordinates of an event are denoted by ξ^{α}, where the index α ranges over the values 1, 2, 3, and 4. Greek lower case indices will always have this same range.

To each oriented 2-dimensional submanifold $\mathcal{E}^2 \subset \mathcal{E}$ we assign a real number $F(\mathcal{E}^2, \underset{\approx}{\Phi})$ called the <u>magnetic flux</u> through $\mathcal{E}^2$, where $\underset{\approx}{\Phi}$ is a <u>unit of magnetic flux.</u>

A2. Magnetic flux is a continuous 2-cochain in $\mathcal{E}$ with phys. dim. $[\underline{\underline{\Phi}}]$.

It follows from A2 that there exists an <u>electromagnetic field</u> φ , a continuous 2-form in $\mathcal{E}$, such that

$$F(\mathcal{E}^2, \underset{\approx}{\Phi}) = \int_{\mathcal{E}^2} \varphi(\underset{\approx}{\Phi}) \quad ,$$

$$\text{phys. dim. } \varphi = [\underset{=}{\Phi}] \quad . \tag{2.1}$$

A3: (Faraday's Law of Magnetic Induction) The magnetic flux of every cycle $\mathcal{C}^2 \subset \mathcal{E}$ vanishes:

$$\oint_{\mathcal{C}^2} \varphi = 0 \tag{2.2}$$

From Faraday's law of magnetic induction follows the existence of electromagnetic potentials α , regular 1-forms in $\mathcal{E}$, such that the magnetic flux through any $\mathcal{E}^2$ is given by

$$F(\mathcal{E}^2, \underset{\approx}{\Phi}) = \oint_{\partial\mathcal{E}^2} \alpha(\underset{\approx}{\Phi}) \quad , \tag{2.3}$$

where,

$$\text{phys. dim. } \alpha = [\underset{=}{\Phi}] \quad , \tag{2.4}$$

$$\text{rot } \alpha = \varphi \quad . \tag{2.5}$$

The distribution of magnetic flux does not determine a unique electromagnetic potential for if α satisfies (2.3) for every $\mathcal{E}^2$, then so also does α' given by

$$\alpha' = \alpha + \text{rot } \beta \quad , \tag{2.6}$$

where β is an arbitrary regular 0-form in $\mathcal{E}$ (scalar field).

Next we introduce the definition and certain properties of electric charge in precise analogy with magnetic flux. To each oriented 3-dimensional submanifold of events $\mathcal{E}^3 \subset \mathcal{E}$ we assign a real number $C(\mathcal{E}^3, \underset{\approx}{Q})$ called the electric charge of $\mathcal{E}^3$ measured in units of electric charge $\underset{\approx}{Q}$.

A4: Electric charge is a continuous 3-cochain in $\mathcal{E}$ and phys. dim. $C = [\underset{=}{Q}]$.

According to A4 there exists a 3-form χ in $\mathcal{E}$, the charge-current field , such that

$$C(\mathcal{E}^3, \underset{\approx}{Q}) = \int_{\mathcal{E}^3} \chi(\underset{\approx}{Q}) \ , \qquad (2.7)$$

$$\text{phys. dim. } \chi = [\underset{=}{Q}].$$

A5: (Law of Conservation of Charge) The electric charge of every cycle C^3 is zero:

$$\oint_{C^3} \chi = 0 \ . \qquad (2.8)$$

It follows from the law of conservation of charge that there exist charge-current potentials η , 2-forms in $\mathcal{E}$, such that

$$C(\mathcal{E}^3, \underset{\approx}{Q}) = \oint_{\partial \mathcal{E}^3} \eta(\underset{\approx}{Q}) \quad , \tag{2.9}$$

$$\text{phys. dim. } \eta = [\underline{\underline{Q}}] \, , \tag{2.10}$$

$$\text{rot } \eta = \chi \; . \tag{2.11}$$

The charge distribution does not determine a unique charge-current potential. If η satisfies (2.9) for every $\mathcal{E}^3$, then so also does η' given by

$$\eta' = \eta + \text{rot } \beta \, , \tag{2.12}$$

where β is an arbitrary 1-form in $\mathcal{E}$.

When the electromagnetic field and the charge-current field are not only continuous but regular, then Faraday's law of induction and the law of conservation of electric charge are equivalent to the local conditions

$$\begin{aligned} \text{rot } \varphi &= 0 \, , \\ \text{rot } \chi &= 0 \, . \end{aligned} \tag{2.13}$$

Moreover, if φ and χ are smooth, these local conditions are expressible in terms of the differential equations,

$$\begin{aligned} \partial \vee \varphi &= 0 \, , \\ \partial \vee \chi &= 0 \, . \end{aligned} \tag{2.14}$$

But the smoothness of the distributions of magnetic flux and electric charge assumed in A2 and A4 is already so severe as to rule out electromagnetic fields and charge-current fields commonly considered in applications. Weaker assumptions sufficiently general to include all the applications are the following alternatives to A2 and A4. Call an r-form in $\mathcal{E}$ piecewise regular if $\mathcal{E}$ can be subdivided into a finite number of standard manifolds, $\mathcal{E} = \bigcup_i \mathcal{E}_i$, $\mathcal{E}_i \cap \mathcal{E}_j = 0$, such that the r-form is regular in each. Then replace A2 and A4 by

A2': There exists a piecewise regular electromagnetic potential α such that the magnetic flux measured in units of $\underset{\approx}{\Phi}$ through any $\mathcal{E}^2$ is given by

$$F(\mathcal{E}^2, \underset{\approx}{\Phi}) = \oint_{\partial\mathcal{E}^2} \alpha(\underset{\approx}{\Phi}) \quad . \tag{2.1'}$$

A4': There exists a piecewise regular charge-current potential η such that the electric charge in $\mathcal{E}^3$ measured in units $\underset{\approx}{Q}$ is given by

$$C(\mathcal{E}^3, \underset{\approx}{Q}) = \oint_{\partial\mathcal{E}^3} \eta(\underset{\approx}{Q}) \quad . \tag{2.4'}$$

Faraday's law of magnetic induction and the law of conservation of electric charge now follow as consequences of A2' and A4'. In each $\mathcal{E}_i$ the electromagnetic field and the charge-current field are given by

$$\varphi = \text{rot}\,\alpha \quad ,$$
$$\chi = \text{rot}\,\eta \quad . \tag{2.15}$$

II. 3. DISCUSSION

No definitions of "magnetic flux per unit area" or of "electric charge per unit volume" have been given. No concept of length, area, volume, or of time has been used nor can be inferred from the definitions and properties assigned thus far to magnetic flux and electric charge. The definitions and properties of magnetic flux and electric charge given here in this generality must be attributed to Bateman and Kottler.

If a coordinate system is introduced in about an event ξ, and a choice is made for the unit of electric charge, then the charge-current field χ has a value at ξ represented by components $\chi_{\alpha\beta\mu}(\xi, \underset{\approx}{Q})$. If the coordinates and the unit of charge are transformed, these components of the charge-current field undergo the transformation

$$\chi_{\alpha'\beta'\mu'}(\xi, \underset{\approx}{Q'}) = Q^{-1}\; s^{\alpha}{}_{\alpha'}\; s^{\beta}{}_{\beta'}\; s^{\mu}{}_{\mu'}\; \chi_{\alpha\beta\mu}(\xi, \underset{\approx}{Q}) , \tag{3.1}$$

where $\underset{\approx}{Q}' = \underset{=}{Q}\ \underset{\approx}{Q}$, and $s^{\alpha}_{\alpha''}$ are the components of the gradient of the coordinate transformation. Similarly, the components of the electromagnetic field φ under the transformation

$$\varphi_{\alpha'\beta'}(\xi, \underset{\approx}{\Phi}') = \underset{=}{\Phi}^{-1}\ s^{\alpha}_{\alpha'}\ s^{\beta}_{\beta'}\ \varphi_{\alpha\beta}(\xi, \underset{\approx}{\Phi}) \quad , \qquad (3.2)$$

where $\underset{\approx}{\Phi}' = \Phi\ \underset{\approx}{\Phi}$. These laws of transformation for the components of the electromagnetic and charge-current fields were discovered, after some trial and error, by Lorentz . From the present viewpoint, these transformation formulas are consequences of the definitions of magnetic flux and electric charge. We see that these transformation formulas merely reflect the fact that electric charge and magnetic flux, relative to given units, are numbers assigned to certain sets of events, and that these numbers do not depend in any way upon coordinate systems, observers, measuring apparatus, clocks, rigid rods, or anything else of that nature.

that $K(\varphi^2)$ and $K(\varphi^1)$ are linearly independent and have a common factor. By P3 this common factor cannot be ϵ^2, else we would have $\hat{K}(\varphi^2, \varphi^2) = 0$; hence the common factor must be ϵ^3 (i. e., a covector proportional to ϵ^3) . Thus $K(\varphi^2) = \pi \vee \epsilon^3$ for some value of the covector π. But φ^1 and $K(\varphi^2)$ have a common factor also since $E(\varphi^1, K(\varphi^2)) = \hat{K}(\varphi^1, \varphi^2) = E(\varphi^2, K(\varphi^1)) = 0$, where we have used the symmetry property P2. This common factor cannot be ϵ^4; hence it must be proportional to ϵ^1. Thus $K(\varphi^2) = a\, \epsilon^3 \vee \epsilon^1$ for some value of a. Now set $\varphi^3 = \epsilon^3 \vee \epsilon^4$. Then $K(\varphi^3)$ has a factor in common with $K(\varphi^2)$. This factor cannot be ϵ^3 (P3 again) ; hence we can set $K(\varphi^3) = \pi' \vee \epsilon^1$ for some value of π'. But $\pi' \vee \epsilon^1$ has a factor in common with $\epsilon^2 \vee \epsilon^3$ and this cannot be ϵ^3; hence $K(\varphi^3) = b\, \epsilon^1 \vee \epsilon^2$ for some value of b. Choosing now as a basis in $V^4(\xi)$ the four linearly independent vectors $e^a(\xi)$ reciprocal to the four covectors $\epsilon_a(\xi)$ we find that the components of K have the canonical values set forth in (3. 2) except that in place of some of the 1's and -1's there appear the unknown scalar values of a and b. But from the condition $K^2 = -I$ we may deduce that the values of a and of b must both be +1, and the theorem is proved.

II. 4. THE LIGHT CONE

Two quadratic forms q and q' in any vector space V^m are conformal if $q = a q'$; i. e., if one be proportional to the other. A quadratic form with signature $\pm(1, 1, \ldots, 1, -1)$ determines a cone. The cone is the set of all vectors such that $q(v, v) = 0$. Every quadratic form in the same conformally equivalent class determines the same cone. Conversely, we can show that the cone of a quadratic form of signature $(1, 1, \ldots, 1, -1)$ determines the quadratic form uniquely up to a factor; i. e., determines the conformal class of the form. To see this recall that any symmetric bilinear form $q(u, v)$ is determined by its values $q(w, w)$ because $q(u, v) = \frac{1}{2}(q(u+v, u+v) - q(u, u) - q(v, v))$. Now suppose that $q'(u, u) = 0$ whenever $q(u, u) = 0$. It is required to show that $q' = a q$. Let $e_i, i=1, \ldots n$ be a basis such that $q(e_i, e_j) = \mathrm{diag}(1, 1, \ldots, 1, -1)$. Then $q'(e_A \pm e_n, e_A \pm e_n) = 0$ for $A = 1, 2, \ldots n-1$ because $e_A \pm e_n$ belongs to the cone of q. From these conditions on q' we deduce that $q'(e_A, e_n) = 0$, and that $q'(e_A, e_A) + q'(e_n, e_n) = 0$. But $u = \frac{1}{2}(e_A + e_B) + e_n$ is also an element of the cone of q for $A \neq B$; hence, $q'(u, u) = 0$. If this condition be expanded, then using the properties of q' already established we get $q'(e_A, e_B) = 0$, for $A \neq B$. Therefore, $q'(e_i, e_j) = \text{diag}(a, a, \ldots, a, -a)$, which proves the assertion.

The adjoint $q\dagger$ of a quadratic form q with signature $(1, 1, 1, -1)$ determines a cone of covectors in $V_4(\xi)$ satisfying $q^\dagger(\alpha, \alpha) = 0$.

Theorem II. 4. 1: There exists a unique cone in $V^4(\xi)$ such that, for all covectors $\alpha_1, \alpha_2, \beta_1$, and β_2,

$$\hat{K}(\alpha_1 \vee \beta_1 \,,\, \alpha_2 \vee \beta_2 \,,\, \underset{\approx}{\Phi}, \underset{\approx}{Q}) = \pm [q^\dagger(\alpha_1, \alpha_2)\, q^\dagger(\beta_1, \beta_2) - q^\dagger(\alpha_1, \beta_2)\, q^\dagger(\alpha_2, \beta_1)] \,, \tag{4. 1}$$

for one of the quadratic forms $q(v, v)$ in the conformal class which determines the cone.

Proof: Let e_a be a Lorentz frame at ξ and let ϵ^a be the reciprocal basis. The definition of the quadratic form K in $V_{[4^2]}(\xi)$ depends linearly upon the choice of the 4-vector used to define the duality transformation. Choose for this 4-vector, $E = e_1 \vee e_2 \vee e_3 \vee e_4$. Let natural units be chosen for charge and magnetic flux so that $K^2 = -I$. For this choice of basis, of E, and of the units of charge and flux we have the following values for the components of $\hat{K}$:

$$\hat{K}^{(\alpha\beta)(\mu\tau)} = \hat{K}(\epsilon^{\alpha} \vee \epsilon^{\beta}, \epsilon^{\mu} \vee \epsilon^{\tau}, \underset{\approx}{\Phi}, \underset{\approx}{Q}) = \mathrm{diag}(-1, 1, -1, 1, -1, 1).$$

The rows and columns of the six by six matrix of independent components of $\hat{K}$ are labeled according to the same convention (14), (23), (24), (31), (34), (12) used in (II. 3. 2). Now let Latin indices range over the values 1, 2, and 3. One easily verifies that

$$\hat{K}^{ijkl} = q^{\dagger ik} q^{\dagger jl} - q^{\dagger il} q^{\dagger jk} = \delta^{ik} \delta^{jl} - \delta^{il} \delta^{jk} \quad (4.3)$$

From this equation we may deduce that

$$q_{ij} = \delta_{ij} \det \| q_{rs} \| \, . \quad (4.4)$$

On taking the determinant of this equality we find that $\det \| q_{rs} \| = \pm 1$; hence,

$$q_{rs} = \pm \delta_{rs} \, , \quad q^{\dagger rs} = \pm \delta^{rs} \, . \quad (4.5)$$

Now use

$$K^{ijk4} = q^{\dagger ik} q^{\dagger j4} - q^{\dagger i4} q^{\dagger jk} = 0 \quad (4.6)$$

and (4.5) to deduce that

$$q^{\dagger j4} = 0 \, . \quad (4.7)$$

Finally, from (4.5), (4.7), and

$$K^{i4j4} = q^{\dagger ij} q^{\dagger 44} - q^{\dagger i4} q^{\dagger j4} = - \delta^{ij} \quad (4.8)$$

one gets

$$\begin{aligned} q^{\dagger 44} &= -1 \text{ when } q^{\dagger ij} = +\delta^{ij} \, , \\ q^{\dagger 44} &= +1 \text{ when } q^{\dagger ij} = -\delta^{ij} \, . \end{aligned} \quad (4.9)$$

This proves the theorem and also shows that if $\epsilon_a(\xi)$ is a Lorentz frame at ξ, then the cone determined by K (the light cone) is given by

$$q = a\left[\sum_{i=1}^{3} \epsilon_i \otimes \epsilon_i - \epsilon_4 \otimes \epsilon_4 \right], \qquad (4.10)$$

where a is an arbitrary factor of proportionality.

II. 5. THE ELECTROMAGNETIC SYMMETRY GROUP

The aether tensor $K(\xi)$ determines the electromagnetic symmetry group as follows. Every non-singular transformation

$$S: V^4(\xi) + \underset{=}{Q} + \underset{\equiv}{\Phi} \rightarrow V^4(\xi) + \underset{=}{Q} + \underset{\equiv}{\Phi} \qquad (5.1)$$

induces a non-singular linear transformation

$$T_S: W \rightarrow W \qquad (5.2)$$

in the tensor space $W = V_{[4^2]}(\xi) \otimes V^{[4^2]}(\xi)\ [\underset{=}{Q}\ \underset{=}{\Phi}]$ of which K is an element. The transformation T_S is defined by the condition

$$\bar{K}(\bar{\alpha} \vee \bar{\beta}, \bar{\underset{\approx}{Q}}, \bar{\underset{\approx}{\Phi}}) \cdot \bar{u} \vee \bar{v} = K(\alpha \vee \beta, \underset{\approx}{Q}, \underset{\approx}{\Phi}) \cdot u \vee v, \qquad (5.3)$$

where $(\bar{\underset{\approx}{Q}}, \bar{\underset{\approx}{\Phi}}, \bar{u}, \bar{v}) = (S(\underset{\approx}{Q}), S(\underset{\approx}{\Phi}), S(u), S(v))$, $\bar{K} = T_S(K)$, and $\bar{\alpha}$ and $\bar{\beta}$ are defined by the corresponding conditions $\bar{\alpha} \cdot \bar{v} = \alpha \cdot v$, $\bar{\beta} \cdot \bar{v} = \beta \cdot v$. The transformation S is an element of the electromagnetic symmetry group if and only if

$$T_S(K) = K \quad , \tag{5.4}$$

i.e., if and only if the aether tensor is an invariant tensor under the induced transformation T_S. But if K is invariant under T_S, then so also is the light cone determined by K. Hence the electromagnetic symmetry group is a subgroup of the transformations (5.1) for which

$$T'_S(q) = a\,q \quad , \tag{5.5}$$

where T'_S is the linear transformation induced by S in the space $V_{(42)}(\xi)$ of symmetric dimensionless 2-cotensors at ξ. The group of transformations S defined by the condition (5.5) is the <u>conformal Lorentz group</u>. The electromagnetic symmetry group is a subgroup of the conformal Lorentz group.

It has been established that the aether tensor has the representation

$$K = k\,[\,(e_2 \vee e_3) \otimes (\epsilon^1 \vee \epsilon^4) - (e_1 \vee e_4) \otimes (\epsilon^2 \vee \epsilon^3)$$

$$+ \text{ cyclic permutations of } 1, 2, 3.\ ] \quad , \tag{5.6}$$

where k is a scalar with phys. dim. $[\underset{\approx}{Q}\underset{\approx}{\Phi}^{-1}]$. The condition $k=1$ defines natural units for $\underset{\approx}{Q}$ and $\underset{\approx}{\Phi}$. The e_α are the elements of a Lorentz frame at ξ, and $\epsilon^{\alpha}(e_{\beta}) = \delta^{\alpha}_{\beta}$. The transformations S in (5.1) have the form

$$S = \left\| \begin{matrix} s & & \\ & \underline{\underline{\Phi}}^{-1} & \\ & & \underline{\underline{Q}}^{-1} \end{matrix} \right\| \qquad (5.7)$$

where $s: V^4(\xi) \to V^4(\xi)$ is a non-singular linear transformation of the tangent space at ξ, and $\underline{\underline{\Phi}}$ and $\underline{\underline{Q}}$ are the ratios of the new and old units of flux and charge. From the representation (5.7) of K we deduce that a necessary condition for K to be invariant under T_S is that

$$\underline{\underline{Q}}\,\underline{\underline{\Phi}}^{-1}\,(\text{sign det } s) = +1 . \qquad (5.8)$$

In other words, the transformation S must be a proper transformation, and the absolute value of the product $\underline{\underline{Q}}\,\underline{\underline{\Phi}}^{-1}$ must be 1. One now convinces himself that the two conditions (5.5) and (5.8) are both necessary and sufficient that K be invariant under T_S.

II. 6. DISCUSSION

Taking the relation between the distributions of electric charge and magnetic flux as a starting point, we have shown how this relation determines a unique cone of directions at each event. We have defined an electromagnetic symmetry group as the invariance group of the aether tensor which describes completely how the distributions of charge and flux are related in the classical theory. This symmetry group turns out to be a certain subgroup of the transformations in a 6-dimensional vector space, the direct sum of the tangent space $V^4(\xi)$ of the manifold of events, and the two unit spaces of electric charge and magnetic flux. If the units of electric charge and magnetic flux are held fixed, then the subgroup of the electromagnetic symmetry group defined by that condition turns out to be the proper conformal Lorentz group of transformations of $V^4(\xi)$. On the other hand, an improper conformal Lorentz transformation of $V^4(\xi)$ when accompanied by an improper transformation of the unit space $\underline{\Phi} \otimes \underline{\Omega}^c$ of determinant -1 is an electromagnetic symmetry element. Thus, oddly enough, the statement commonly made that Maxwell's equations are invariant under the Lorentz group of transformations (transformations which leave a quadratic form of signature (1, 1, 1, -1) invariant) is not quite correct for two opposing accounts. First of all, this statement is usually made under the agreement that the units of charge and flux have been fixed upon and not subject to transformation. In this case, the Lorentz group does not contain the corresponding subgroup of electromagnetic symmetry transformations. Moreover, if the units are held fixed, since only proper transformations of $V^4(\xi)$ are then contained in the corresponding subgroup of electromagnetic symmetry transformations, this subgroup does not contain the full Lorentz group. On the other hand, if transformations of the units of electric charge and magnetic flux are properly taken into account (they are certainly of equal importance to the transformations of the tangent space) we have seen that Maxwell's equations are invariant under certain improper transformations of $V^4(\xi)$ provided they are accompanied by an appropriate improper transformation of the unit space $\underline{\underline{\Phi}} \otimes \underline{\underline{\Omega}}$.

LECTURE III

ACTION AND GRAVITY

"These physical hypotheses, however, are entirely alien from the way of looking at things which I adopt, and one object which I have in view is that some of those who wish to study electricity may, by reading this treatise, come to see that there is another way of treating the subject, which is no less fitted to explain the phenomena, and which, though in some parts it may appear less definite, corresponds, as I think, more faithfully with our actual knowledge, both in what it affirms and in what it leaves undecided."

J. C. Maxwell

III. 1. INTRODUCTORY REMARKS

An essential qualitative difference between electric charge and gravitational or inertial mass is that charge appears to occur in Nature with either sign, but mass occurs with but one sign. Thus charge and mass are essentially different qualities of matter requiring essentially different mathematical representations. How can we introduce in a natural way this essential difference within the present framework of concepts and mathematics? We have seen in the first two lectures how Maxwell's equations for the electromagnetic field and the charge-current field can be viewed as conditions upon the distributions of two physical quantities, electric charge and magnetic flux. Each of these physical quantities has been represented mathematically as a linear function of oriented, submanifolds of events; the first by a linear function of 3-dimensional submanifolds, the second by a linear function of 2-dimensional submanifolds. In neither of these cases would it be natural or even possible to introduce a condition that the distributions be positive everywhere (or negative). The situation is different, however, in the case of a linear function of 4-dimensional submanifolds in an orientable embedding space of 4-dimensions. The 4-dimensional oriented curvilinear simplexes in event-space can be divided into two equivalence classes of similarly oriented simplexes. This cannot be done with simplexes of lower dimension. Thus it is natural to seek a theory and representation of mass connected in some way with the simplexes of dimension four in $\mathcal{E}$.

If a physical quantity is represented by a 4-cochain in event-space, it is possible to introduce the condition that the value of the cochain, for any fixed choice of the physical unit, has the same sign on every similarly oriented 4-simplex in $\mathcal{E}$. It turns out that mass itself is not the natural quantity which is of uniform sign on such a class, but rather a quantity we shall call action which we may think of as related to the integral of the mass of a 3-simplex over a 1-simplex in a time-like direction. These remarks are intended only to guide the intuition, since formal definitions of mass and action are given later.

III. 2. ACTION AND GRAVITY

We introduce two more unit spaces $\underline{\underline{A}}$ and $\underline{\underline{G}}$ alongside $\underline{\underline{Q}}$ and $\underline{\underline{\Phi}}$, and two corresponding physical quantities called <u>action</u> and <u>gravity</u>.

A8: Action is a continuous 4-cochain in event-space with phys. dim. $[\underline{\underline{A}}]$.

It follows that there exists a continuous 4-form μ in such that the action $A(\mathcal{E}^4, \underset{\approx}{A})$ of a given oriented

4-dimensional submanifold $\mathcal{E}^4$ measured in units of action $\underset{\approx}{A}$ is given by

$$A(\mathcal{E}^4, \underset{\approx}{A}) = \int_{\mathcal{E}^4} \mu(\underset{\approx}{A}) \quad , \tag{2.1}$$

and

$$\text{phys. dim. } \mu = [\underset{=}{A}] \, . \tag{2.2}$$

Now by A1 , $\mathcal{E}$ is orientable. Thus it is possible to divide the oriented simplexes of $\mathcal{E}$ into two equivalence classes of similarly oriented 4-simplexes. Denote these two classes by $\mathcal{E}^+$ and $\mathcal{E}^-$.

A8: If S and S' are any two 4-simplexes in the same orientation class ($\mathcal{E}^+$ or $\mathcal{E}^-$) , then action has the property

$$A(S, \underset{\approx}{A})\, A(S', \underset{\approx}{A}) \geqslant 0 \quad . \tag{2.3}$$

Now every continuous 4-cochain in $\mathcal{E}$ has a potential. Thus there exists a regular 3-form π in $\mathcal{E}$ such that

$$A(\mathcal{E}^4, \underset{\approx}{A}) = \oint_{\partial\mathcal{E}^4} \pi(\underset{\approx}{A}) \quad . \tag{2.4}$$

As in the case of the electromagnetic and charge-current potentials, π is not uniquely determined by the distribution of action. If π satisfies (2. 4) then so also does

$$\pi' = \pi + \operatorname{rot} \beta \ , \tag{2. 5}$$

where β is any regular 2-form in $\mathcal{E}$.

<u>Gravity</u> is a physical quantity with phys. dim. $[\underset{=}{G}]$.

A9: Gravity is a continuous 1-cochain in event-space.

It follows from A9 that there exists a unique <u>gravitational field</u> γ, a 1-form in $\mathcal{E}$, such that the gravity $G(\mathcal{E}^1, \underset{\approx}{G})$ of any curve in $\mathcal{E}$ measured in units of gravity $\underset{\approx}{G}$ is given by

$$G(\mathcal{E}^1, \underset{\approx}{G}) = \int_{\mathcal{E}^1} \gamma(\underset{\approx}{G}) \tag{2. 6}$$

$$\text{phys. dim. } \gamma = [\underset{=}{G}] \ .$$

A10: The gravitational field is circulation free.

Accordingly, if $\mathcal{C}^1$ is any closed smooth curve in then

$$G(\mathcal{C}^1, \underset{\approx}{G}) = \oint_{\mathcal{C}^1} \gamma(\underset{\approx}{G}) = 0 \ . \tag{2. 7}$$

It follows from A10 that there exist gravitational potentials ψ , regular 0-forms (scalar fields) in $\mathcal{E}$, such that

$$G(\mathcal{E}^1, \underset{\approx}{G}) = \oint_{\partial\mathcal{E}^1} \psi(\underset{\approx}{G}) = \psi(\xi^+) - \psi(\xi^-) , \qquad (2.8)$$

and

$$\text{phys. dim. } \psi = [\underset{=}{G}] . \qquad (2.9)$$

In (2.8), ξ^+ and ξ^- are the end points of the oriented curve $\mathcal{E}^1$. The gravitational potential is determined by the distribution of gravity only to within an additive constant. What we have called the gravity of a curve is equal to the difference between the values of the gravitational potential at its two end points.

III. 3. THE GRAVITATIONAL AETHER RELATION

Guided now by the way that the distributions of electric charge and magnetic flux are related in Maxwell's theory, we shall introduce a connection between the distributions of action and gravity which, up to this point, have been treated as independent.

A.11: There exists a linear transformation

$$L: V_4(\xi)\ [\underset{=}{G}] \rightarrow V_{[43]}(\xi)\ [\underset{=}{A}] \tag{3.1}$$

of the space of 1-forms in $\mathcal{E}$ with phys. dim. $[\underset{=}{G}]$ into the space of 3-forms in $\mathcal{E}$ with phys. dim. $[\underset{=}{A}]$ such that $L(\gamma)$ is a potential of action.

We shall call L the <u>gravitational aether tensor</u> so as to distinguish it from the (electromagnetic) aether tensor K which plays the analogous role in relating the distributions of charge and flux.

Let $E \neq 0$ be an arbitrary 4-vector in $V^{[4^4]}(\xi)$ and set $D(\alpha) = \text{dual}\ \alpha = E \wedge \alpha$, where α is any r-covector at ξ. Then $\hat{L}$ defined by

$$\hat{L}(\alpha, \beta, \underset{\approx}{G}, \underset{\approx}{A}) = D(\beta) \cdot L(\alpha, \underset{\approx}{G}, \underset{\approx}{A}) \tag{3.2}$$

is a bilinear form in $V_4(\xi)$ with phys. dim. $[\underset{=}{A}\underset{=}{G}^{-1}]$.

A12: The bilinear form $\hat{L}$ defined by the gravitational aether tensor L is symmetric and has signature $\pm(1, 1, 1, -1)$.

The definition of $\hat{L}$ in terms of L depends on the choice of a 4-vector E. Since all such 4-vectors are proportional, L determines $\hat{L}$ up to a factor. Thus the gravitational aether tensor determines , not a quadratic form, but a cone of directions in $V_4(\xi)$ defined by $\hat{L}(\alpha, \alpha) = 0$. We call this cone the <u>gravitational</u> <u>cone</u>. The connection between gravity and electromagnetism is established in part by

A13: The light cone and the gravitational cone coincide at every event.

III. 4. GRAVITATIONAL INVARIANCE GROUP AND THE METRICAL STRUCTURE OF EVENT-SPACE

The gravitational invariance group is defined in the same way as the electromagnetic invariance group with K replaced by L. Thus we consider the set of all transformations

$$S: V^4(\xi) \oplus \underline{\underline{A}} \oplus \underline{\underline{G}} \to V^4(\xi) \oplus \underline{\underline{A}} \oplus \underline{\underline{G}} \qquad (4.1)$$

of the six-dimensional vector space consisting in the direct sum of the tangent space at ξ and the two unit spaces $\underline{\underline{A}}$ and $\underline{\underline{G}}$. Each such transformation induces a linear transformation

$$T_S : W \rightarrow W \tag{4.2}$$

in the tensor space

$$W = V_{[43]} \otimes V^4(\xi) \, [\underline{\underline{A}}\,\underline{\underline{G}}^{-1}] , \tag{4.3}$$

of which L is an element. The gravitational invariance group is the set of all transformations S such that

$$T_S(L) = L \quad . \tag{4.4}$$

The transformations S are of the form

$$S = \begin{Vmatrix} s & & \\ & \underline{\underline{A}}^{-1} & \\ & & \underline{\underline{G}}^{-1} \end{Vmatrix} \quad . \tag{4.5}$$

Since L defines a unique cone in $V_4(\xi)$, it follows that the transformations s in (4.5) must be a subgroup of the conformal Lorentz transformations. It follows from A12 and A13 that there exists a Lorentz frame e_a at ξ such that $L(\xi)$ is given by

$$L(\xi) = \ell[\epsilon^1 \vee \epsilon^2 \vee \epsilon^3 \otimes e_4 + \epsilon^1 \vee \epsilon^2 \vee \epsilon^4 \otimes e_3 + \epsilon^2 \vee \epsilon^3 \vee \epsilon^4 \otimes e_1 + \epsilon^3 \vee \epsilon^1 \vee \epsilon^4 \otimes e_2] , \tag{4.6}$$

where ℓ is a constant and

$$\text{phys. dim. } \ell = [\underline{\underline{A}}\,\underline{\underline{G}}^{-1}] \quad . \tag{4.7}$$

One sees from this representation of L that necessary conditions for S to be a symmetry element are

$$(\underline{\underline{A}}\,\underline{\underline{G}}^{-1})\ \mathrm{sign}(\det s) > 0\ , \tag{4.8}$$

and

$$|\det s| = \underline{\underline{A}}^2\,\underline{\underline{G}}^{-2}\ . \tag{4.9}$$

One then verifies that the three conditions, s a conformal Lorentz transformation, (4.8), and (4.9), are both necessary and sufficient that S be a symmetry element.

Consider now the symmetric tensor $\hat{L}$ defined in (3.2) in terms of L and an arbitrary 4-vector $E \neq 0$. For every choice of E, det L < 0. Let $E = e_1 \vee e_2 \vee e_3 \vee e_4$ and $\mathcal{E} = \epsilon^1 \vee \epsilon^2 \vee \epsilon^3 \vee \epsilon^4$, where e_a and ϵ^a are reciprocal sets and e_a is a Lorentz frame at ξ. Now we have

$$L:\ V_4(\xi)\,[\underline{\underline{G}}] \rightarrow V^4(\xi)\,[\underline{\underline{A}}]\ ,$$

$$\det L = \mathcal{E} \cdot [\,\hat{L}(\epsilon^1) \vee \hat{L}(\epsilon^2) \vee \hat{L}(\epsilon^3) \vee \hat{L}(\epsilon^4)\,] \tag{4.10}$$

and,

$$\text{phys. dim. } \det \hat{L} = [\underline{\underline{A}}^4\,\underline{\underline{G}}^{-4}] \tag{4.11}$$

If $E' = a\,E$, then $\hat{L}' = a\,\hat{L}$ and we see that

$$\det \hat{L}' = a^4\,a^{-2} \det \hat{L} = a^2 \det \hat{L}\ . \tag{4.12}$$

It follows that the symmetric tensor $g^\dagger$ defined by

$$g^{\dagger} = \frac{\hat{L}}{\sqrt{-\det \hat{L}}} \qquad (4.13)$$

depends only on the orientation of E and

$$\text{phys. dim. } g^{\dagger} = [\underline{\underline{A}}^{-1} \underline{\underline{G}}], \quad \det g < 0 \ . \qquad (4.14)$$

Thus we see that the gravitational aether tensor determines a unique (up to sign) symmetric tensor field $g^{\dagger}$ and cotensor field g ($g^{\dagger\alpha\beta} g_{\beta\tau} = \delta^{\beta}_{\tau}$) with signature $\pm(1, 1, 1, -1)$. The gravitational symmetry group can be characterized alternatively as the subgroup of proper transformations (4.5) such that

$$T_S(g) = g \quad , \qquad (4.15)$$

where T_S is the linear transformation induced by S in the space of symmetric 2-cotensors at ξ having phys. dim. $[\underline{\underline{A}}\underline{\underline{G}}^{-1}]$. If g is smooth, then it is seen that the gravitational aether relation endows event-space with a smooth (pseudo) Riemannian structure defined by a fundamental <u>metric tensor</u> $\pm g$ with phys. dim. $[\underline{\underline{A}}\underline{\underline{G}}^{-1}]$, $\det g < 0$.

LECTURE IV

MOTION OF CONTINUOUS MEDIA

"The idea of motion implies the existence of some means of recognizing again and again the entity that moves. By extending the idea of a mathematical point we have the concept of a moving point which we shall call an electrical point and we may start with the fundamental hypothesis that three independent quantities (α, β, γ) are sufficient to specify an electrical point and distinguish it from others."

H. Bateman

IV.1. INTRODUCTORY REMARKS

It is worth emphasizing that nothing in the preceding lectures depends in any way upon the concept of motion or the concept of a material medium. Thus it stands independently of whatever is now said about continuous media and motion.

One of the earliest questions raised by the new mechanics of special relativity theory concerned the concept and definition of rigid motion. In classical mechanics, a rigid motion of any set of material points is defined by the condition that the distance between every pair of points in the set at each instant of time remain invariant in time. In the Minkowski manifold of special relativity theory, an instant of time is not defined. We have given, rather, a cone of directions at each event. Born was the first to consider the problem of extending or transferring the classical notion of a rigid motion to the new kinematical setting of Minkowski space. The concept and definition of a rigid motion of a continuous medium is an essential preliminary to a relativistic theory of elastic response. Essential also to an extension of classical continuum mechanics to the more general space-time manifolds of general relativity is the concept and definition of velocity and acceleration. The objective in this lecture will be to show how the counterparts of these classical kinematical ideas can be defined in a natural way in a manifold of events in which there is given at each event a cone of directions corresponding to a metric field g with $\det g < 0$, phys. dim. $g = [\underline{\underline{A}}\,\underline{\underline{G}}^{-1}]$.

IV. 2. MOTION OF MATERIAL MEDIA

We shall consider only 3-dimensional material media.

A 3-dimensional material medium is an orientable smooth manifold of dimension 3, which we denote by $\mathcal{M}$, together with any additional structure which may be assigned to it. In what follows, the only properties of $\mathcal{M}$ which will be used are those which follow from the definition of a smooth manifold. The points of $\mathcal{M}$ are called <u>material points</u> and will be denoted by X, X', X'', etc. A motion of $\mathcal{M}$ is a smooth mapping

$$f: \mathcal{E}' \to \mathcal{M} \tag{2.1}$$

of a set of events $\mathcal{E}'$ onto $\mathcal{M}$ such that f has rank three at every point in the domain of f and such that the proper vector of with proper value 0 is time-like. The orbit of a material point X is the set of events $\{\xi; f(\xi) = X\}$ experienced by the point X and is called the <u>world-line</u> of X. As remarked in the preliminaries §12, from what has been assumed and the implicit function theorem it follows that locally the world-line of every material point is a smooth parametrizable curve in $\mathcal{E}$. In other words, a motion may be represented locally as a mapping

$$f^{\dagger}: \mathcal{M} \times \mathcal{O} \to \mathcal{E}, \tag{2.2}$$

where $\mathcal{O}$ is an open set of real numbers. The gradient of $f^\dagger$ with respect to the parameter, $\nabla_\tau f^\dagger$ is a proper vector of ∇f with proper value zero; it is tangent to the world-line of the point X at the event $f^\dagger(X, \tau)$. The world-velocity field of the motion is the normalized field of time-like tangent vectors

$$v = \frac{\nabla_\tau f^\dagger}{\sqrt{|g(\nabla_\tau f^\dagger, \nabla_\tau f^\dagger)|}}, \qquad (2.3)$$

It should perhaps be pointed out that the direction of v, either forward or backward, depends on the parametrization of the world-lines. As yet we have introduced no assumptions which distinguish the future from the past. Thus one should keep in mind the dependence of v upon the parametrization. Let $v^\dagger$ be the covector set in correspondence with v by g; i. e., defined by $v^\dagger(u) = g(v, u)$. Then the world-acceleration of the motion is defined by

$$a^\dagger = (\text{rot } v^\dagger) \wedge v \ . \qquad (2.4)$$

Note that $a^\dagger(v) = 0$. Note also that phys. dim. $v = [\underline{\underline{T}}^{-1}]$, phys. dim. $v^\dagger = [\underline{\underline{A}}\underline{\underline{G}}^{-1}\ \underline{\underline{T}}^{-1}]$, phys. dim. $a^\dagger = [\underline{\underline{A}}\underline{\underline{G}}^{-1}\ \underline{\underline{T}}^{-2}]$, where we introduce the notation $\underline{\underline{T}} = +\sqrt{\underline{\underline{A}}\ \underline{\underline{G}}^{-1}}$ for the time dilation which accompanies a transformation $\underset{\approx}{A}' = \underline{\underline{A}}^{-1}\ \underset{\approx}{A}$,

$\underset{\approx}{G}' = \underset{=}{G}^{-1} \underset{\approx}{G}$ of the units of action and gravity. Only the sign of the acceleration is affected by the choice of these units and it is independent of the parametrization of the motion.

In what follows it proves convenient to introduce the symbol s which has the value plus one or minus one such that $sg(u,u) < 0$ if u is a time-like vector. The value of s depends on the units of action and gravity. We may indicate this dependence by writing phys. dim. $s = [\underset{=}{A}\underset{=}{G}^{-1}\underset{=}{T}^{-2}]$.

IV. 3. RIGID MOTIONS

The gradient ∇f of a motion (2.1) determines a linear mapping

$$\nabla f \colon W_3(X) \longrightarrow V_4(\xi) \quad , \quad X = f(\xi) \tag{3.1}$$

of the tangent space in the material manifold $\mathcal{M}$ into the space of covectors at ξ. In terms of ∇f and g we define a symmetric bilinear function of material covectors C^{-1} by

$$C^{-1}(\Omega, \Lambda) = s\, g^{\dagger}(\nabla f(\Omega), \nabla f(\Lambda)) \tag{3.2}$$

and show that C^{-1} is positive definite. Set $\omega = \nabla f(\Omega)$. By hypothesis, $\nabla f \cdot v = 0$, where v is the velocity vector and is time-like. But then $\omega(v) = g^{\dagger}(\omega, v^{\dagger}) = 0$, where $v^{\dagger} \cdot u = g(v,u)$, and $v^{\dagger}$ is time-like. Hence the set of image vectors

$\nabla f(W_3)$ consists in covectors each of which is normal to the time-like covector $v^{\dagger}$. It follows that $sg^{\dagger}$ restricted to this subspace of $V_3(\xi)$ is positive definite. Therefore, C^{-1} defined in (3. 2) is positive definite, and phys. dim. $C^{-1} = [\underline{\underline{T}}^{-2}]$. It is seen that a motion of $\mathcal{M}$ in $\mathcal{E}$ determines a set of positive definite quadratic forms $\{C(\xi);\ f(\xi) = X\}$ in each tangent space $W^3(X)$ of $\mathcal{M}$. A motion of $\mathcal{M}$ is <u>locally rigid</u> at X if and only if this set of forms consists in a single element; i. e., if and only if $f(\xi) = f(\xi') \rightarrow C(\xi) = C(\xi')$. The above definition of a locally rigid motion is independent of any parametrization of the motion. If a parametrization is given, more familiar geometric results emerge, but it is important to distinguish those results which depend on the parametrization from those which do not.

Let $\tau: \mathcal{E}' \rightarrow R$ be a regular mapping of the events in a neighborhood of $\xi \in \mathcal{E}'$ into the real numbers such that $\nabla\tau$ has rank one and such that each member of the family of level surfaces $\tau(\xi) =$ constant is space-like. We may choose τ such that $s\nabla\tau \cdot v = -1$. Then $|\tau(\xi_2) - \tau(\xi_1)|$, where ξ_1 and ξ_2 are events experienced by a given material point X is the interval of proper time measured along the world-line of X between the events ξ_1 and ξ_2. We may view (2. 2) as a one-parameter family of embeddings of $\mathcal{M}$ in $\mathcal{E}$. The image $f^\dagger(\mathcal{M}, \tau)$, for each value of τ, is a smooth 3-dimensional surface in $\mathcal{E}$, and the surfaces $f^\dagger(\mathcal{M}, \tau)$ and $f^\dagger(\mathcal{M}, \tau')$ have no points in common if $\tau \neq \tau'$.

In §12 of the preliminaries it was shown how a motion of $\mathcal{M}$ determines a one-parameter family of motions of defined by

$$T_\tau(\xi) = f^\dagger_{t+\tau} \circ f^{\dagger\,-1}_t(\xi) . \tag{3. 3}$$

The gradient, ∇T_τ, of this point transformation is a linear transformation of the tangent space $V^4(\xi)$ onto the tangent space $V^4(\xi_\tau)$, $\xi_\tau = T_\tau(\xi)$. Let the quadratic form g_τ be defined by

$$g_\tau(u, v, \xi) = s g(\nabla T_\tau(u), \nabla T_\tau(v), T_\tau(\xi)) \tag{3. 4}$$

and suppose that the parameter t is proper time. Then the Lie derivative of the metric field g with respect to the velocity field v of the motion is given by

$$\underset{v}{\mathcal{L}} g = s \left. \frac{dg_\tau}{d\tau} \right|_{\tau=0} . \tag{3. 5}$$

Let a superposed dot denote differentiation with respect to t. On verifies by direct calculation that

$$\overset{\circ}{C}(U, W, X, \tau) = s\int_V g(P_v(u) , P_v(w), f^{\dagger}(X, \tau)), \quad (3.6)$$

where $P_v = I + s\, v \otimes v^{\dagger}$ is the projection of $V^4(\xi)$ onto the subspace $V^3(\xi, v)$ of vectors at ξ normal to the velocity, and u is the unique solution of $P_v \nabla f^{\dagger}(U) = P_v \cdot u$. It follows from (3.6) that a necessary and sufficient condition that a motion be locally rigid at X is that the restriction of the Lie derivative of the metric field with respect to the velocity field of the motion to the subspaces $V^3(\xi, v)$, $f(\xi) = X$ vanish.

The differential equations expressing this condition have the form

$$B_{\alpha\beta} = v^{\dagger}_{(\alpha;\beta)} + v^{\dagger}_{(\alpha} a_{\beta)} = 0 , \quad f(\xi)=X. \quad (3.6)$$

where a semi-colon denotes covariant differentiation, $v^{\dagger}_{\alpha}$ is the velocity field (covariant components of v) , and the a_{α} are the components of the acceleration defined in (2.4). The tensor B is called Born's rate of strain tensor.

V.4. MEASURES OF RELATIVE STRAIN AND ROTATION

The transformation

$$F_\tau(\xi_t) = P_v \bullet \nabla_{\mathcal{M}} f^\dagger \cdot \nabla f(\xi_t) \tag{4.1}$$

is a non-singular linear transformation

$$F_\tau(\xi_t) : V^3(\xi_t, v) \to V^3(\xi_{t+\tau}, v) \tag{4.2}$$

of the subspace of space-like vectors normal to the velocity at ξ_t onto the onto the corresponding space at the event $\xi_{t+\tau}$. In each of these subspaces we have the induced metric

$$g_v = -s(g + s\ v^\dagger \otimes v^\dagger) . \tag{4.3}$$

For brevity, let us denote the inner product $g_v(u, w, \xi_t)$ by $(u, w)_t$, where u and w are elements of $V^3(\xi, v)$. We define the adjoint transformation $F_\tau^\dagger$ by the condition

$$(u, F_\tau \circ w)_{t+\tau} = (F_\tau^\dagger \circ u , w) . \tag{4.4}$$

Then

$$C_\tau(\xi_t) = F_\tau^\dagger \circ F_\tau(\xi_t) \tag{4.5}$$

is positive definite and self-adjoint in the sense that

$$(C_\tau u, u)_t = (u, C_\tau^\dagger u)_t = (F_\tau^\dagger F_\tau u, u)_t =$$

$$= (F_\tau u, F_\tau u)_t > 0 \tag{4.6}$$

if $u \in V^3(\xi, v)$, $u \neq 0$, and

$$C_0(\xi_t) = P_v(\xi_t) . \tag{4.7}$$

We call $C_\tau(\xi_t)$ the <u>relative strain measure.</u> Let $D_\tau(\xi_t)$

be the self-adjoint positive definite square root of $C_\tau(\xi_t)$ so that $D_\tau^2 = C_\tau$ and $D_\tau = D_\tau^\dagger$. Set

$$F_\tau = R_\tau \circ D_\tau \,, \qquad (4.8)$$

so that

$$R_\tau = F_\tau \circ D_\tau^{-1} \,. \qquad (4.9)$$

Then we may assert that R_τ is isometric in the sense that

$$(R_\tau u, \; R_\tau w)_{t+\tau} = (u, w)_t \,. \qquad (4.10)$$

This follows from the observation that

$$R_\tau R_\tau^\dagger = F_\tau D_\tau^{-1} (D_\tau^{-1})^\dagger F_\tau^\dagger = P_v \,. \qquad (4.11)$$

Thus the transformation R_τ is a measure of the relative rotation of the material about X in the configuration at ξ_t and the configuration at $\xi_{t+\tau} = T_\tau(\xi_t)$.

In terms of components one has the following relation between $C_\tau(\xi_t)$ and the tensor C defined in (3.2).

$$C_\tau{}^\alpha{}_\mu(\xi_t) = g_v^{\alpha\epsilon}(\xi_t) \, C_{AB}(\xi_{t+\tau}) \, f^A{}_{,\epsilon}(\xi_t) f^B{}_{,\mu}(\xi_t) \,,$$

where $g_v^{\alpha\beta} = s(g^{\alpha\beta} + s\, v^\alpha v^\beta)$. (4.12)

By definition, a motion is rigid if and only if $C_{AB}(\xi_{t+\tau})$ is independent of τ. But this is the only term on the right-hand side of (4.12) which depends on τ; hence, a necessary and sufficient condition that a motion be locally rigid at X is that the relative strain measure $C_\tau(\xi_t)$ be independent of τ.

LECTURE V

MASS, STRESS, ENERGY, ENTROPY, AND MOMENTUM

"'It will be remembered that Faraday, when studying the curvature of lines of force in electrostatic fields, had noticed an apparent tendency of adjacent lines to repel each other, as if each tube of force were inherently disposed to distend laterally; and that in addition to this repellent or diverging force in the transverse direction, he supposed an attractive or contractile force to be exerted at right angles to it, that is to say, in the direction of the lines of force.'"

"'Of the existence of these pressures and tensions Maxwell was fully persuaded; and he determined analytical expressions suitable to represent them. The tension along the lines of force must be supposed to maintain the pondermotive force which acts on the conductor on which the lines of force terminate; and it may therefore be measured (in the system of units we are now using) by the force which is exerted on unit area of the conductor, i.e., $\epsilon E^2/8$ or $DE/8$. The pressure at right angles to the lines of force must then be determined so as to satisfy the condition that the aether is to be in equilibrium."

E. T. Whittaker

V.1. INTRODUCTORY REMARKS

Nothing considered thus far depends in any way upon definitions or properties of stress, energy, entropy, or momentum. The electromagnetic field and the gravitational field have been introduced as densities of physical quantities called magnetic flux and gravity. They have not been defined in the traditional way in terms of the force exerted on charge, current, and mass; nor have any other "mechanical" attributes been assigned them. The principles of balance of energy, momentum, and angular momentum, and the elementary concepts of force and inertia which suffice as a suitable framework for classical continuum mechanics do not extend in a completely natural way to the more general space-time manifolds considered here. We seek a system of mechanical principles sufficiently general to include Einstein's relation between the curvature of space-time and the distribution of energy and momentum, yet specific enough to embrace a theory of motion and entropy production in a continuum interacting with an electromagnetic field.

V. 2. INERTIAL MASS

Recall that action is a 4-cochain in $\mathcal{E}$ and that

$$A(\mathcal{E}^4, \underset{\approx}{A}) = \int_{\mathcal{E}^4} \mu(\underset{\approx}{A}) \tag{2.1}$$

where μ is the density of action.

Let a motion of a material medium $\mathcal{M}$ be parametrizable so that we are given a congruence of world-lines by a mapping

$$f^\dagger: \mathcal{M} \times \mathcal{O} \to \mathcal{E} \tag{2.2}$$

such that the tangent field $\overset{\circ}{\xi} = \nabla_t f^\dagger$ (where t is the parameter) is everywhere time-like, $s\,\overset{\circ}{g}(\dot{\xi}, \dot{\xi}) < 0$. Define J by

$$J = \frac{\det \nabla f^\dagger}{\sqrt{-sg(\dot{\xi}, \dot{\xi})}} , \tag{2.3}$$

$$\text{phys. dim. } J = [\underset{=}{T}^{-1}] . \tag{2.4}$$

It follows from the Jacobi identies that

$$\operatorname{div}(J^{-1} v) = (J^{-1} v^a)_{,a} = 0 . \tag{2.5}$$

Let $\hat{\mu}$ denote the dual of the action density μ. Then we define the density of inertial mass ρ by

$$\mu = J^{-1} \rho \tag{2.6}$$

Let v be the velocity field of a motion of $\mathcal{M}$. A tensor field in the world-tube of $\mathcal{M}$ is said to be invariant under the motion if its Lie derivative with respect to v vanishes. It is said to be absolutely invariant under the motion if its Lie derivative with respect to any field λ v proportional to the velocity vanishes. If φ is an r-form,

$$\underset{w}{\mathcal{L}} \varphi = \text{rot}(w \wedge \varphi) + w \wedge \text{rot}\, \varphi \tag{2.7}$$

from which we see that an invariant r-form is absolutely invariant if and only if

$$w \wedge \varphi = 0 . \tag{2.8}$$

In particular, if a scalar field is invariant under a motion, it is absolutely invariant under that motion.

Recall that in the previous lecture we showed that a motion of $\mathcal{M}$ was locally rigid about a point X if and only if the induced metric g_v were invariant under the motion. With these remarks in mind we come to the important concept of the constitutive relation for the inertial mass of a continuous medium. We consider a class of material media for which the value of the inertial mass $\rho(\xi)$ at a given event is given by

$$\rho(\xi) = U(g(\xi), \nabla f^{\dagger}(\xi), \omega_a(\xi), X(\xi)) \qquad (2.9)$$

where ω_a, $a = 1, 2, \ldots N$ is a set of fields in the world-tube of $\mathcal{M}$ called state variables. Different functions U define different materials of the class. We say that the state of the material about X in invariant if the motion is rigid and if $\mathcal{L}\omega_a = 0$, and absolutely invariant if the motion is rigid and $\underset{\lambda v}{\mathcal{L}}{}^{v}\omega_a = 0$ for arbitrary λ.

A 14. (Principle of Material Indifference) The inertial mass is invariant under the motion whenever the state is invariant under the motion.

Replacement Theorem: If the constitutive function U for the inertial mass satisfies the principle of material indifference and invariance of each of the state variables ω_a implies its absolute invariance, then

$$U(g, \nabla f^{\dagger}, \omega_a, X) = U(g_v, \nabla f^{\dagger}, \omega_a, X). \qquad (2.10)$$

Proof: Set $g = g_v + r\, v^\dagger \otimes v^\dagger$ and insert this value for g in U. The replacement theorem is equivalent to the assertion that the resulting value of U is independent of the value of r. Assuming, as we do, that U is differentiable in g, this is equivalent to the condition

$$\frac{dU}{dr} = \frac{\partial U}{\partial g_{\alpha\beta}} v^\dagger_\alpha v^\dagger_\beta \mathrel{\overset{\circ}{=}} 0 . \qquad (2.11)$$

But if U satisfies the principle of material indifference, then

$$\frac{\partial U}{\partial g_{\alpha\beta}} \underset{\lambda v}{\mathcal{L}} g_{\alpha\beta} = 0 \qquad (2.12)$$

for arbitrary λ if v is the velocity field of a locally rigid motion. But for a rigid motion we have

$$\underset{\lambda v}{\mathcal{L}} g_{\alpha\beta} = 2\lambda a_{(\alpha} v^\dagger_{\beta)} + 2\lambda_{,(\alpha} v^\dagger_{\beta)} , \qquad (2.13)$$

where a_α are the components of the acceleration. Choosing $\lambda = 0$ at a point and $\lambda_{,\alpha}$ proportional to $v^\dagger_\alpha$ at that point, one sees that (2.12) is sufficient for the replacement of g by g_v.

But g_v is uniquely determined by the gradient of $f^\dagger$ and the Born strain measures $C_{AB} = s\,(g_v)_{\alpha\beta}\, f^{\dagger\alpha}{}_{,A}\, f^{\dagger\beta}{}_{,B}$ which proves the corollary: Every constitutive relation of the type (2.9) is equivalent to one of the form

$$\rho(\xi) = U^*(C(\xi),\ \nabla f^{\dagger}(\xi),\ \omega_a(\xi)) \qquad (2.18)$$

provided that invariance of the state variables ω_a implies their absolute invariance.

V. 3. ENTROPY AND THE CLAUSIUS-DUHEM INEQUALITY

Alongside the four basic physical quantities, magnetic flux, electric charge, gravity, and action we now set a fifth and last called entropy.

A 15: Entropy is a continuous 4-cochain in event-space.

$$S(\mathcal{E}^4, \underset{\approx}{S}) = \int_{\mathcal{E}^4} \sigma(\underset{\approx}{S}) \quad , \qquad (3.1)$$

where σ is the entropy field and $\underset{\approx}{S}$ is the unit of entropy.

$$\text{phys. dim. } \sigma = [\underset{=}{S}] \ ,$$

and

$$S(\mathcal{E}^4, \underset{\approx}{S})\ S(\mathcal{E}'^4, \underset{\approx}{S}) \geqslant 0 \qquad (3.2)$$

if $\mathcal{E}^4$ and $\mathcal{E}'^4$ are similarly oriented.

Let a parametrized motion of a continuous medium be given and consider the entropy $S(\mathcal{E}^4(\tau), \underset{\approx}{S})$ of $\mathcal{E}^4(\tau) = T_\tau(\mathcal{E}^4)$ where $\mathcal{E}^4$ is a fixed oriented submanifold in the world-tube of the medium and suppose that the parameter is proper time.

Then

$$\left.\frac{dS(\mathcal{E}^4(\tau), \underset{\approx}{S})}{d\tau}\right|_{\tau=0} = \int_{\mathcal{E}^4} \mathcal{L}_v \sigma = \oint_{\partial\mathcal{E}^4} v \wedge \sigma \,. \quad (3.3)$$

Define the entropy flux relative to the material, q/θ and the rate of production of entropy Φ/θ by setting

$$\oint_{\partial\mathcal{E}^4} v \wedge \sigma = -\oint_{\partial\mathcal{E}^4} h/\theta + \int_{\mathcal{E}^4} \Phi/\theta \quad , \quad (3.4)$$

where $\theta \neq 0$ is a factor later to be identified with the absolute temperature. It is assumed that $\hat{h} \circ v = 0$, and

$$\text{phys. dim.}\left(\frac{h}{\theta}\right) = \text{phys. dim.}\left(\frac{\Phi}{\theta}\right) = [\underset{=}{S}\underset{=}{T}^{-1}]. \quad (3.5)$$

Thus the sign of the entropy produced in $\mathcal{E}^4$ defined by

$$\Delta(\mathcal{E}^4, \underset{\approx}{S}\underset{\approx}{T}^{-1}) = \int_{\mathcal{E}^4} \Phi(\underset{\approx}{S}\underset{\approx}{T}^{-1})/\theta \quad (3.6)$$

depends on the sign of the unit of entropy, the orientation of $\mathcal{E}^4$, and the direction of time, since the definition of Φ/θ compares the entropy of a set of events $\mathcal{E}^4$ and a set $\mathcal{E}^4(\tau)$

obtained from $\mathcal{E}^4$ by translation along the world-lines of a material medium. The sign of the product

$$\Delta^*(\mathcal{E}^4, \underset{\approx}{S}^2 \underset{\approx}{T}^{-1}) = S(\mathcal{E}^4, \underset{\approx}{S})\Delta(\mathcal{E}^4, \underset{\approx}{S}\underset{\approx}{T}^{-1}) \qquad (3.7)$$

however, is independent of the orientation of $\mathcal{E}^4$ and the unit of entropy.

A 16: (The Clausius-Duhem Inequality). For every $\mathcal{E}^4$, either

$$\Delta^*(\mathcal{E}^4, \underset{\approx}{S}^2 \underset{\approx}{T}^{-1}) \geqslant 0,$$

or,

$$\Delta^*(\mathcal{E}^4, \underset{\approx}{S}^2 \underset{\approx}{T}^{-1}) \leqslant 0 \qquad (3.8)$$

depending only on the direction of the time step $d\tau$ in the definition of the rate of production of entropy.

None of our assumptions before (3.8) allow one to distinguish between past and future events along a given world-line. But the Clausius-Duhem inequality provides this distinction if the inequality sign in (3.8) holds for at least one $\mathcal{E}^4$. The Clausius-Duhem inequality asserts that the entropy of a set of events $\mathcal{E}^4(\tau)$ obtained by <u>advancing</u> the set $\mathcal{E}^4$ into the future along the world-lines of a material medium is never less that the entropy of $\mathcal{E}^4$ plus whatever entropy has flowed across the lateral boundary of the corresponding segment of the world-tube.

V. 4. SOME DIFFERENTIAL INDENTITIES

Before considering equations of motion we digress to record some results which will ease the formal manipulations.

Let $t^{(a)}$, $a = 1, 2, \ldots N$ denote the ordered set of components in some coordinate system of the manifold of a differentiable tensor field, or the components of a set of such fields in the manifold. Then the components of the Lie derivative of the set of fields with respect to the vector field w with components w^{α} are given in terms of the $t^{(a)}$ and the partial derivatives of the $t^{(a)}$ by

$$\underset{w}{\pounds}\, t^{(a)} = w^{\alpha}\, t^{(a)}{}_{,\alpha} + F^{(a)\beta}_{(b)\alpha}\, t^{(b)}\, w^{\alpha}{}_{,\beta} \qquad (4.1)$$

where a comma followed by an index denotes partial differentiation with respect to the corresponding coordinate. The coefficients $F^{(a)\alpha}_{(b)\beta}$ are constants independent of the coordinates. All partial derivatives in the formula (4.1) may be replaced by the corresponding covariant derivatives and the formula remains true. Thus we also have

$$\underset{w}{\mathcal{L}}\, t^{(a)} = w^{\alpha}\, t^{(a)}{}_{;\alpha} + F^{(a)\beta}_{(b)\alpha}\, t^{(b)}\, w^{\alpha}{}_{;\beta}\,, \tag{4.2}$$

where a semi-colon denotes covariant differentiation. The equivalence of (4.1) and (4.2) follows from

$$t^{(a)}{}_{;\alpha} = t^{(a)}{}_{,\alpha} - F^{(a)\beta}_{(b)\mu}\, t^{(b)} \left\{ {\mu \atop \alpha\beta} \right\},$$

$$w^{\alpha}{}_{;\beta} = w^{\alpha}{}_{,\beta} + \left\{ {\alpha \atop \mu\beta} \right\} w^{\mu}\ . \tag{4.3}$$

In particular,

$$\underset{w}{\mathcal{L}}\, g_{\alpha\beta} = w^{\mu}\, g_{\alpha\beta;\mu} + 2\, w_{(\alpha;\beta)}$$

$$= 2\, w_{(\alpha;\beta)}\ , \tag{4.4}$$

since $g_{\alpha\beta;\mu} = 0$. The Lie derivative and the partial derivative commute:

$$\underset{w}{\mathcal{L}}\, t^{(a)}{}_{,\alpha} = (\underset{w}{\mathcal{L}}\, t^{(a)})_{,\alpha} \tag{4.5}$$

Suppose $f^{\dagger}: \mathcal{M} \times \mathcal{O} \to \mathcal{E}$ is a parametrized motion of a continuous medium. Then $\nabla f^{\dagger}$ has components which we now denote by $\xi^{\alpha}{}_{,A}$ and $\mathring{\xi}^{\alpha}$. Now suppose that

$$f^{\dagger}_{\lambda}: \mathcal{M} \times \mathcal{O} \to \mathcal{E} \tag{4.6}$$

is a one-parameter family of such motions of $\mathcal{M}$ such that $f^\dagger_\lambda$ is smooth in all its arguments. Set

$$w^a = \frac{\partial f^\dagger}{\partial \lambda} \quad . \tag{4.7}$$

Then we may assert that

$$\overset{*}{(\nabla f^\dagger_\lambda)} = \frac{\partial(\nabla f^\dagger)}{\partial \lambda} + \underset{w}{\mathcal{L}}(\nabla f^\dagger) = 0 \ , \tag{4.8}$$

or, in terms of components,

$$\overset{*}{\xi^a{}_{,A}} = \frac{\partial(\xi^a{}_{,A})}{\partial \lambda} + \underset{w}{\mathcal{L}}\, \xi^a{}_{,A} = 0 \ ,$$

$$\overset{*}{\mathring{\xi}} = \frac{\partial \mathring{\xi}^a}{\partial \lambda} + \underset{w}{\mathcal{L}}\, \dot{\xi}^a = 0 \quad . \tag{4.9}$$

Consider next a 4-form ν in event-space and its integral over an oriented $\mathcal{E}^4$ in $\mathcal{E}$.

$$N(\mathcal{E}^4) = \int_{\mathcal{E}^4} \nu \tag{4.10}$$

The submanifold may be taken as small as one pleases so that, in particular, let us assume that it lies entirely within a single coordinate patch. Let $\hat{\nu}$ denote the dual of ν. Suppose that the value $\hat{\nu}(\xi)$ in any coordinate system is given

$$\mathcal{V}(\xi) = \epsilon K(g_{\alpha\beta}(\xi),\ g_{\alpha\beta,\gamma}(\xi),\ g_{\alpha\beta,\gamma\mu}(\xi),\ \ldots) \tag{4.11}$$

where K is a smooth function of its arguments independent of the coordinate system and $\epsilon = \pm 1$ depending on the orientation of the basis vectors. The functions $K = \sqrt{-g}\; R$, $K' = \sqrt{-g}\, R^{\alpha\beta} R_{\alpha\beta}$, where $R_{\alpha\beta} = R_{\mu\alpha\beta}{}^{\mu}$; $R = g^{\dagger\alpha\beta} R_{\alpha\beta}$, and $R_{\mu\alpha\beta}{}^{\tau}$ is the curvature tensor of $g_{\alpha\beta}$ are examples of functions K satisfying our hypotheses. We show that for every such function one has the identity

$$\left(\frac{\delta \hat{\mathcal{V}}}{\delta g_{\alpha\beta}}\right)_{;\beta} = 0 \tag{4.12}$$

where

$$\frac{\delta \hat{\mathcal{V}}}{\delta g_{\alpha\beta}} = \frac{\partial \hat{\mathcal{V}}}{\partial g_{\alpha\beta}} - \left(\frac{\partial \hat{\mathcal{V}}}{\partial g_{\alpha\beta,\gamma}}\right)_{,\gamma} + \left(\frac{\partial \hat{\mathcal{V}}}{\partial g_{\alpha\beta,\gamma\delta}}\right)_{,\gamma\delta} - \ldots \tag{4.13}$$

is the Hamiltonian or Lagrangian derivative of $\mathcal{V}$ with respect to g . This well-known and oft quoted theorem may be proved as follows. Let $T_\tau : \mathcal{E} \to \mathcal{E}$ be any one-parameter family of point transformations such that $T_o = I$ and let $w^a = \partial \bar{\xi}^a(\xi, \tau)/\partial\tau \,\big|\, \tau=0$. Let $T_\tau(\mathcal{E}^4) = \mathcal{E}^4_\tau$. Then

$$\dot{N} = \frac{d}{d\tau} \int_{\mathcal{E}^4_\tau} \mathcal{V} \,\Bigg|_{\tau=0} = \int_{\mathcal{E}^4_0} \pounds_w \mathcal{V}$$

$$= \oint_{\partial \mathcal{E}_0^4} w \wedge \nu \quad .$$

Thus $\overset{\circ}{N}$ vanishes whenever $w = 0$ on the boundary $\partial \mathcal{E}^4$ of the region $\mathcal{E}^4$. But this implies that

$$\int_{\mathcal{E}^4} \frac{\delta \nu}{\delta g_{\alpha\beta}} \underset{w}{\pounds} g_{\alpha\beta} = 0$$

if w vanishes on the boundary of $\mathcal{E}^4$ together with its derivatives up to order (m-1), where m is the order of the highest derivative of g appearing in the function K.

Using (4.4) we then see that

$$\int_{\mathcal{E}^4} \frac{\delta \nu}{\delta g_{\alpha\beta}} w_{(\beta;\alpha)} = \int_{\mathcal{E}^4} \left[\left(\frac{\delta \nu}{\delta g_{\alpha\beta}} w_\alpha \right)_{;\beta} - \left(\frac{\delta \nu}{\delta g_{\alpha\beta}} \right)_{;\beta} w^\alpha \right] \qquad (4.14)$$

$$= - \int_{\mathcal{E}^4} \left(\frac{\delta \nu}{\delta g_{\alpha\beta}} \right)_{;\beta} w^\alpha = 0$$

and this must hold for every vector field w in $\mathcal{E}^4$ which vanishes with its derivatives up to order (m-1) on the boundary. We then conclude that we must have (4.12) at each interior point of $\mathcal{E}^4$, which proves the assertion.

V. 5. EQUATIONS OF MOTION OF A CONTINUOUS MEDIUM

Recall that in classical mechanics the equations of motion of systems with a finite number of degrees of freedom with generalized coordinates $q^{(a)}$, $a = 1, 2, \ldots N$ may be stated in the form

$$\delta \int_{t_1}^{t_2} T \; dt = - Q_{(a)} \, \delta q^{(a)} \tag{5.1}$$

for variations $\delta q^{(a)}$ consistent with the constraints; *i.e.*, for all variations $\delta q^{(a)}$ of a specified class. The coefficients $Q_{(a)}$ are the generalized forces. The specification of a particular dynamical system consists in laying down constitutive relations for the kinetic energy T and the generalized forces in terms of the coordinates $q^{(a)}$ and their time derivatives. Here we shall propose Lagrange equations for a continuous medium in a general space-time manifold for which, at each event, we are given a cone of directions $g(\xi)$ that varies smoothly from point to point.

Consider the action of a segment of a world-tube of the material medium $\mathcal{M}$,

$$A(\mathcal{E}^4, \underset{\approx}{A}) = \int_{\mathcal{E}^4} \mu(\underset{\approx}{A}) \quad , \tag{5.1}$$

and suppose that the motion of $\mathcal{M}$ is given by $f_0^\dagger : \mathcal{M} \times \mathcal{O} \to \mathcal{E}$ where $f_\lambda^\dagger : \mathcal{M} \times \mathcal{O} \to \mathcal{E}$ is a one-parameter family of neighboring comparison motions of $\mathcal{M}$. Let

$$w = \text{grad}_\lambda \, f^\dagger \quad . \tag{5.2}$$

Suppose that

$$\mu = J^{-1} \rho \quad , \tag{5.3}$$

where the inertial mass ρ is given by a constitutive relation of the form

$$\rho = U(g_{\alpha\beta}, \xi^a{}_{,A}, \overset{\circ}{\xi}{}^a, \omega_{(a)}, X) . \tag{5.4}$$

The generalization to materials for which U might depend on higher derivatives of $f^\dagger$ and the metric tensor is not too different from the special case (5.4). More definite and special theories follow on specific choices for the state variables $\omega_{(a)}$, but for the moment we are interested in exposing the common structure of all such dynamical systems. Let $g_\lambda(\xi) = g(\xi, \lambda)$ and $\omega_{(a)\lambda} = \omega_{(a)}(\xi, \lambda)$ be fields such that $g(\xi, 0) = g(\xi)$ and $\omega_{(a)}(\xi, 0) = \omega_{(a)}(\xi)$. Let

$$\delta g_{\alpha\beta} = \left. \frac{\partial g_{\alpha\beta}}{\partial \lambda} \right|_{\lambda=0} , \delta \omega_{(a)} = \left. \frac{\partial \omega_{(a)}}{\partial \lambda} \right|_{\lambda=0} \tag{5.5}$$

and call these quantities the variations of g and $\omega_{(a)}$.

For each comparison state and motion of $\mathcal{M}$, the function

$$A_\lambda = A(\mathcal{E}_\lambda, A) = \int_{\mathcal{E}_\lambda} \mu(g_\lambda, \nabla f_\lambda^\dagger, \omega_{(a)\lambda})$$

is differentiable with respect to the parameter λ. Its derivative is given by

$$\overset{*}{A}_0 = \int_{\mathcal{E}_0} \overset{*}{\mu} = \int_{\mathcal{E}_0} \left[\frac{\partial \mu}{\partial g_{\alpha\beta}} \overset{*}{g}_{\alpha\beta} + \frac{\partial \mu}{\partial \omega_{(a)}} \overset{*}{\omega}_{(a)} \right] \qquad (5.7)$$

where we have used the identities (4.8), and the superposed star denotes the linear operator $\partial / \partial \lambda + \underset{w}{\mathcal{L}}$.

In view of (5.7) we define the generalized forces $Q^{\alpha\beta}$ and $Q^{(a)}$ by setting

$$\delta A = \overset{*}{A}_0 = - \int_{\mathcal{E}_0} \left[\frac{1}{2} Q^{\alpha\beta} \overset{*}{g}_{\alpha\beta} + Q^{(a)} \overset{*}{\omega}_{(a)} \right]. \qquad (5.8)$$

This is the Lagrange equation for the dynamical system with constitutive equation (5.4).

The Lagrange equation (5.8) is a variational equation to be satisfied for some specified class of variations $(w, \delta g, \delta\omega_{(a)})$ of the motion, the metric, and the state variables. As in the case of point mechanics, we adopt the view that constitutive equations for the generalized forces $Q^{\alpha\beta}$ and $Q^{(a)}$ expressing these quantities as functions or functionals of the motion and state variables in addition to the constitutive relation for the inertial mass are required to specify the dynamical system. What we have before us is merely a framework into which one can fit most every more definite special case of a theory of motion and deformation of a continuous medium. This remains to be demonstrated by examples. The nature of the system and of the special theory depends on the physical and geometrical properties assigned the state variables $\omega_{(a)}$ and upon the detailed nature of the constitutive relations for the generalized forces. But we shall assume that in every special case, these constitutive relations are consistent with

A 17: (The Principle of Local Determinism) The value of each generalized force $(Q^{\alpha\beta}, Q^{(a)})$ at an event ξ is uniquely determined by the values of the fields g, $\nabla f^{\dagger}$, and $\omega_{(a)}$ at events ξ' not later than ξ belonging to the same world line.

V. 6. THE ELECTROMAGNETIC AND GRAVITATIONAL STRESS-ENERGY-MOMENTUM TENSORS

As in the case of the mechanics of systems with finite degrees of freedom, part or all of the generalized forces may possess a potential. What we do now is essentially to classify the generalized forces into types; a given generalized force is considered the sum of forces of various types. Of these special types of forces, we consider first electromagnetic forces.

Consider the scalar function

$$\hat{\mu}_{(\varphi)} = -\frac{k}{2a}\sqrt{-g}\,(\varphi, \varphi) \tag{6.1}$$

$$\text{phys. dim. } \hat{\mu}_{(\varphi)} = [\underline{\underline{A}}],$$

where (φ, φ) denotes the quadratic form in the space of 2-covectors defined by the metric field g (and $g^{\dagger}$) as in §5 of the preliminaries. The constant k with phys. dim. $[\underline{\underline{Q}}\bar{\underline{\underline{\Phi}}}^{-1}]$ is the constant appearing in the aether tensor K of (II. 2) so that $-k^2 I = K^2$, and a is the fine structure constant.

$$\text{phys. dim. } a = [\underline{\underline{Q}}\ \underline{\underline{\Phi}}\ \underline{\underline{A}}^{-1}] \quad . \tag{6.2}$$

The function $\hat{\mu}_{(\varphi)}$ of the fields g and φ is the dual of a 4-form $\mu_{(\varphi)}$ whose integral over an oriented $\mathcal{E}^4$ is the electromagnetic action of that set. The symmetric tensor field defined by

$$T^{\alpha\beta}_{(\varphi)} = 2 \frac{\delta\mu}{\delta g_{\alpha\beta}} \tag{6.3}$$

is the electromagnetic stress-energy-momentum tensor.

The covariant divergence of the electromagnetic stress-energy-momentum tensor yields the electromagnetic Poynting identity

$$f^{\alpha}_{(\varphi)} = T^{\alpha\beta}{}_{;\beta} = \frac{1}{a} g^{\alpha\beta} \hat{\chi}^{\gamma} \varphi_{\beta\gamma} \quad . \tag{6.4}$$

Its value is the generalized Lorentz force $f^{\alpha}_{(\varphi)}$ which is a measure of the rate at which electromagnetic energy and momentum is converted to other forms of energy and momentum.

In analogy with the above, we define the density of gravitational field action $\hat{\mu}_{(\gamma)}$ by

$$\hat{\mu}_{(\gamma)} = \frac{1}{2} a' \sqrt{-g}\, (\gamma, \gamma) \quad . \tag{6.5}$$

The gravitational constant a' has

$$\text{phys. dim. } a' = [\underline{\underline{G}}^{-1}] . \tag{6.6}$$

The gravitational stress-energy-momentum tensor is defined by

$$T^{\alpha\beta}_{(\gamma)} = 2 \frac{\delta \hat{\mu}_{(\gamma)}}{\delta g_{\alpha\beta}} , \tag{6.7}$$

and the gravitational Poynting identity is

$$f^{\alpha}_{(\gamma)} = T^{\alpha\beta}_{(\gamma);\beta} = a' \hat{\mu}_{()} g^{\dagger\alpha\beta} \gamma_{\beta} . \tag{6.8}$$

The generalized gravitational force $f^{\alpha}_{(\gamma)}$ is a measure of the rate at which gravitational field energy and momentum is converted into other forms of energy and momentum.

Finally, let

$$\hat{\mu}_{(g)} = a'' \sqrt{-g} \; K(g_{\alpha\beta}, g_{\alpha\beta,\gamma}, g_{\alpha\beta,\gamma\delta}, \ldots) \tag{6.9}$$

where K is a scalar function of the metric tensor and its derivatives having phys. dim. $[\underline{\underline{A}}^{-1}\underline{\underline{G}}]$ and independent of any other field quantities. We call $\hat{\mu}_{(g)}$ the action of curvature. We assume that $\hat{\mu}_{(g)}$ has phys. dim. $[\underline{\underline{A}}]$ so that the relativity constant a'' has

$$\text{phys. dim. } a'' = [\underline{\underline{G}}]. \tag{6.10}$$

Later to exhibit Einstein's theory of motion and gravitation as a special case one chooses $\mathbb{K} = R$, where R is the scalar curvature. But whatever choice be made for the function $\mathbb{K}$ one has the Bianchi identity (or a special case of that identity)

$$f^{\alpha}_{(g)} = T^{\alpha\beta}_{(g);\,\beta} = 0\,, \qquad\qquad (6.11)$$

$$T^{\alpha\beta}_{(g)} = 2\,\frac{\delta\,\hat{\mu}_{(g)}}{\delta\,g_{\alpha\beta}}\,,$$

satisfied by the Einstein stress-energy-momentum tensor $T^{\alpha\beta}_{(g)}$.

V. 7. THE INTRINSIC STRESS-ENERGY-MOMENTUM TENSOR

It has been assumed that the action density $\hat{\mu}$ is of the form

$$\hat{\mu} = J^{-1} \rho \tag{7.1}$$

where the density of inertial mass ρ satisfies the principle of material indifference and J is defined in terms of the metric field and the motion of a material medium by

$$J = \frac{\det \nabla f^{\dagger}}{\sqrt{-sg(\nabla_t f^{\dagger}, \nabla_t f^{\dagger})}} . \tag{7.2}$$

The intrinsic stress-energy-momentum tensor is defined by

$$T^{\alpha\beta}_{(\mu)} = 2 \frac{\delta \hat{\mu}}{\delta g_{\alpha\beta}} . \tag{7.3}$$

Substituting from (7.1) for $\hat{\mu}$ one gets

$$T^{\alpha\beta}_{(\mu)} = -s\,\mu\,v^{\alpha}\,v^{\beta} + 2\,J^{-1} \frac{\partial \rho}{\partial g_{\alpha\beta}} \tag{7.4}$$

$$= -s\,\mu\,v^{\alpha}\,v^{\beta} + T^{\alpha\beta}_{(e)}$$

where $T^{\alpha\beta}_{(e)}$ is the <u>elastic stress tensor.</u> If the state variables $\omega_{(a)}$ are such that the replacement theorem applies, then the velocity vector v is a characteristic direction of the elastic stress tensor and

$$T^{\alpha\beta}_{(e)}\,v^{\dagger}_{\alpha} = 0 . \tag{7.5}$$

Hence, the velocity vector is a characteristic vector of the intrinsic stress-energy-momentum tensor,

$$T^{\alpha\beta}_{(\mu)} \; v^{\dagger}_{\beta} = \hat{\mu} \; v^{\alpha} , \tag{7.6}$$

and the proper value is precisely the action density itself.

V. 8. SOME FURTHER CONSTITUTIVE RELATIONS AND THE ENTROPY EQUATION

If $T^{\alpha\beta}_{(b)}$, $b = (\mu , \varphi , \gamma, g)$ is a stress-energy-momentum tensor, we call

$$\hat{\epsilon}_{(b)} = s \; T^{\alpha\beta}_{(b)} \; v^{\dagger}_{\alpha} \; v^{\dagger}_{\beta} \tag{8.1}$$

the density of relative energy of the corresponding type. Thus $\hat{\epsilon}_{(\mu)} = \hat{\mu}$ is the intrinsic relative energy, $\hat{\epsilon}_{(\varphi)}$ is the relative electromagnetic energy density, etc. The adjective "relative" is used because the definition of the energy density depends on the velocity of matter at that event, and such energy (except in the case of the intrinsic energy) is a relative quantity in the sense that it depends upon the motion.

Now set

$$Q^{\alpha\beta} = T^{\alpha\beta}_{(\varphi)} + T^{\alpha\beta}_{(\gamma)} + T^{\alpha\beta}_{(g)} + Q^{\alpha\beta}_{(d)} , \tag{8.2}$$

which serves to define the dissipative stress-momentum tensor, $Q^{\alpha\beta}_{(d)}$.

A 18: The relative dissipative energy vanishes.

$$Q^{\alpha\beta}_{(d)} v^{\dagger}_{\alpha} v^{\dagger}_{\beta} = 0 . \qquad (8.3)$$

Assumption A 18 implies that if we define a complete stress-energy-momentum tensor T by

$$\begin{aligned} T &= T_{(\mu)} + Q \\ &= T_{(\mu)} + T_{(\varphi)} + T_{(\gamma)} + T_{(g)} + Q_{(d)} \end{aligned} \qquad (8.4)$$

then the complete relative energy is given by

$$\hat{\epsilon} = s\, T^{\alpha\beta} v^{\dagger}_{\alpha} v^{\dagger}_{\beta} = \hat{\mu} + \hat{\epsilon}_{(\varphi)} + \hat{\epsilon}_{(\gamma)} + \hat{\epsilon}_{(g)} . \qquad (8.5)$$

Thus according to A 18 the only forms of relative energy are intrinsic, electromagnetic, gravitational, and metrical.

Consider now the case where one of the state variables $\omega_{(a)}$ in the constitutive function for the inertial mass is the entropy field σ. We now write the constitutive relation (2.9) in the form

$$\mu = J^{-1} U(g , \nabla f^{\dagger} , \sigma , \omega_{(a'')}) . \qquad (8.6)$$

Alternatively, and preferably, let $\tilde{\sigma} = J\hat{\sigma}$ be the absolute scalar measure of entropy and set

$$\hat{\mu} = J^{-1} \; U^{*}(g, \nabla f^{\dagger}, \tilde{\sigma}, \omega_{(a')}) . \qquad (8.6^{*})$$

The Lagrange equation (5.8) takes the more explicit form

$$\delta A = \delta \int_{\mathcal{E}^4} \mu = - \int_{\mathcal{E}^4} [\tfrac{1}{2} Q^{\alpha\beta} \overset{*}{g}_{\alpha\beta} - J^{-1} \theta \overset{*}{\tilde{\sigma}} + Q^{(a')} \overset{*}{\omega}_{(a')}] , \qquad (8.7)$$

where θ is the absolute temperature and

$$\text{phys. dim. } \theta = [\underline{\underline{A}} \; \underline{\underline{S}}^{-1}] . \qquad (8.7)$$

According to A 18, the dissipative stress-momentum tensor has a unique decomposition of the form

$$Q^{\alpha\beta}_{(d)} = S^{\alpha\beta}_{(d)} + s(h^{\alpha} v^{\beta} + h^{\beta} v^{\alpha}) \qquad (8.8)$$

where $S^{\alpha\beta}_{(d)} v^{\dagger}_{\beta} = 0$ is the dissipative stress, and $h^{\alpha} = Q^{\alpha\beta}_{(d)} v^{\dagger}_{\beta}$, $h^{\alpha} v^{\dagger}_{\alpha} = 0$, is the heat flux.

Suppose now that the Lagrange equation must hold for arbitrary variations of the class

$$\delta g_{\alpha\beta} = 0 , \quad \overset{*}{\omega}_{(a')} = 0 = \delta\omega_{(a')} + \underset{w}{\mathcal{L}} \omega_{(a')} . \qquad (8.10)$$

Necessary conditions are

$$T^{\alpha\beta}{}_{;\beta} = 0 \ , \qquad \theta = \frac{\partial \rho}{\partial \tilde{\sigma}} \ . \qquad (8.11)$$

Substituting for T and $Q_{(d)}$ from (8.3) and (8.9) in $(8.11)_1$ one gets

$$- s \mu \overset{\circ}{v}{}^{a} + s J^{-1} \dot{\rho} \, v^{a} + T^{a\beta}_{(e);\beta} + f^{a}_{(\varphi)} + f^{\alpha}_{(\gamma)}$$

$$+ S^{a\beta}_{(d);\beta} + s (h^{a} v^{\beta} + h^{\beta} v^{a})_{;\beta} = 0 \ , \qquad (8.12)$$

where a superposed dot denotes $(\ )_{,a} v^{a}$ (i.e., differentiation with respect to proper-time along a world-line of the motion).

The equation

$$v_{a} T^{a\beta}{}_{;\beta} = 0 \ , \qquad (8.13)$$

a consequence of $(8.11)_1$, will be called the <u>entropy equation</u>.

Using $v^{\dagger}_{a} \overset{\circ}{v}{}^{a} = 0$, $S^{a\beta}_{(d)} v^{\dagger}_{a} = T^{a\beta}_{(e)} v^{\dagger}_{a} = 0$, we find that the entropy equation can be put in the form

$$(\hat{\sigma} v^{a} + \frac{h^{a}}{\theta})_{;\alpha} = \frac{1}{\theta} [S^{a\beta}_{(d)} \gamma^{\dagger}_{(a;\beta)} - h^{a} (\ln|\theta|)_{,a}$$

$$+ s h^{a} \overset{\circ}{v}{}^{\dagger}_{a} - f^{a}_{(\varphi)} v^{\dagger}_{\alpha} - f^{a}_{(\gamma)} v^{\dagger}_{a} - \frac{\partial \hat{\mu}}{\partial \omega_{(a')}} \underset{v}{\mathcal{L}} \omega_{(a')}] \qquad (8.14)$$

Comparing this result with equation (3.4) we see that the dissipation function Φ is given by

$$\Phi = S^{\alpha\beta}_{(d)} \; v^{\dagger}_{(\alpha,\beta)} - h^{\alpha} (\ln|\theta|)_{,\alpha} + s h^{\alpha} v^{\dagger}_{\alpha}$$

$$- f^{\alpha}_{(\varphi)} v^{\dagger}_{\alpha} - f^{\alpha}_{(\gamma)} v^{\dagger}_{\alpha} - \frac{\partial \hat{\mu}}{\partial \omega_{(a')}} \underset{v}{\pounds} \omega_{(a')} \quad (8.15)$$

The Clausius-Duhem inequality and this expression for the dissipation function serve as guides in the construction of special theories.

The dissipation function contains the term $- f^{\alpha}_{(\varphi)} v^{\dagger}_{\alpha}$, which represents the rate at which electromagnetic field energy is converted to other forms of energy. Depending on the nature of the constitutive relations for the charge-current field, some of this energy appears as heat energy so that this term is related to the so-called Joule heat. The dissipation function (8.15) contains also the analogous but less familiar term

$$- f^{\alpha}_{(\gamma)} v^{\dagger}_{\alpha} = - a' \hat{\mu} \dot{\psi} , \quad (8.16)$$

where ψ is a gravitational potential. Thus we may conclude that in the rigid motion of a non-heat conducting, electrically neutral body ($f_{(\varphi)} = h = 0$) entropy is produced at the rate

$$\Phi = - a' \hat{\mu} \dot{\psi} - \frac{\partial \hat{\mu}}{\partial \omega_{(a')}} \underset{v}{\pounds} \omega_{(a')} . \quad (8.17)$$

Thus if the gravitational constant $a' \neq 0$, entropy is produced or absorbed by a particle in every motion with invariant state ($\underset{v}{\pounds}\ \omega_{(a')} = 0$) in which that particle does not move on a surface of constant gravitational potential. In Einstein's theory of gravitation $a' = 0$, so that these remarks do not apply. An alternative when it is not assumed that $a' = 0$, is to assume that one of the state variables $\omega_{(d')}$ is the gravitational potential so that the inertial mass of a particle depends on the gravitational potential which it experiences. Moreover, one can adjust this dependence in such a way that no entropy is produced or absorbed in a rigid motion of the medium in which all the other state variables are invariant. This requires that we set

$$-a'\hat{\mu} - \frac{\partial \hat{\mu}}{\partial \psi} = 0 , \tag{8.18}$$

the general solution of which is

$$\hat{\mu} = \hat{\mu}_o e^{-a'\psi} , \tag{8.19}$$

where $\hat{\mu}_o$ is independent of the gravitational potential. If (8.19) be assumed, the equations of motion (8.12) with all forces set equal to zero except inertial and gravitational forces assume the special form

$$s\,\overset{\circ}{v}{}^{a} = a'(\,g^{\dagger a\beta}\,\psi_{,\beta} + s\,\dot{\psi}\,v^{a}\,)\,. \qquad (8.20)$$

In this case, every material point moves on an orbit independent of its mass. If it be further assumed that $\mathcal{E}$ is an affine space free of curvature, the orbits determined by (8.20) when $a'\psi \ll 1$ lie very close to the classical Newtonian orbits. The perihelia, rather than advancing as in Einstein's theory, recess slowly at a somewhat lesser amount per revolution. Of course, nothing which has been assumed in the general theory requires either that $a' \neq 0$, or that the curvature of event-space vanish.

V. 9. EINSTEIN'S FIELD EQUATIONS

A necessary and sufficient condition that the Lagrange equation (8.7) hold for unrestricted variations of the metric field is

$$T = 0 \qquad (9.1)$$

or, equivalently,

$$-T^{a\beta}_{(g)} = T^{a\beta}_{(\mu)} + T^{a\beta}_{(\varphi)} + T^{a\beta}_{(\gamma)} + Q^{a\beta}_{(d)} \qquad (9.2)$$

If one chooses for the function $\hat{\mu}_{(g)}$ the function $a''\sqrt{-g}\,(R + C)$, where C is the cosmical constant, phys. dim. $C = [\underline{\underline{G}}\underline{\underline{A}}^{-1}]$,

then

$$-T^{\alpha\beta}_{(g)} = a'' \sqrt{-g}\,[\,R^{\alpha\beta} - \frac{1}{2}\, g\dagger^{\alpha\beta}(R + C)\,] \qquad (9.3)$$

and (9. 2) becomes Einstein's relation between the curvature of space-time and the distribution of stress, energy, and momentum. It is of course assumed in Einstein's theory that $a'' \neq 0$, but it is also assumed that what we have called the gravitational constant a' to which $T_{(\gamma)}$ is proportional is equal zero. Under this assumption, the numerical value of the relativity constant can be so chosen that the geodesics of a singular solution of $T_{(g)} = 0$ lie close to the classical orbits of electrically neutral point particles. In Einstein's theory of gravitation, what we have called the gravitational field γ exerts no force on matter and does not influence its motion in any way.

In the special theory of relativity, it is assumed a priori that event-space is affine and that the components of the metrical field (the light cone) are constant in an affine coordinate system of $\mathcal{E}$. Thus, in special relativity, as in classical mechanics, the metrical and temporal structure of the manifold of events is laid down as a postulate at the outset and not affected in any way by the distribution of stress, energy, and momentum. The Lagrange equation holds only for variations consistent with the constraint $\delta g = 0$, which is natural from the viewpoint of the special theory. Thus equation (9.1) does not apply in the special theory. Both the general and the special theory of relativity are embraced as special cases of our general assumptions thus far.

LECTURE VI

ELECTRODYNAMICS OF DIELECTRIC MEDIA

"It was the great merit of H. A. Lorentz that he brought about a change here in a convincing fashion In principle, a field exists, according to him, only in empty space. Matter-considered as atoms-is the only seat of electric charges; between the material particles there is empty space, the seat of the electro-magnetic field, which is created by the position and velocity of the point charges which are located on the material particles. Dielectricity, conductivity, etc., are determined exclusively by the type of mechanical tie connecting the particles, of which the bodies consist. The particle-charges create the field, which, on the other hand, exerts forces upon the charge of the particles, thus determining the motion of the latter according to Newton's law of motion. If one compares this with Newton's system, the change consists in this: action at a distance is replaced by the field, which thus also describes the radiation. Gravitation is usually not taken into account because of its relative smallness; its consideration, however, was always possible by means of the enrichment of the structure of the field; i. e., expansion of Maxwell's law of the field. The physicist of the present generation regards the point of view achieved by Lorentz as the only possible one; at that time, however, it was a surprizing and audacious step, without which the later development would not have been possible."

A. Einstein

V. I. INTRODUCTORY REMARKS

It is significant to the historical development of the principle of relativity that Einstein's famous paper of 1905 was entitled, "Elektrodynamik bewegter Korper". The lectures to this point represent an attempt to summarize by a small number of explicit assumptions, a set of general physical principles common to a large class of more special and specific mathematical theories of the electromagnetic field in material media and of the consequent motion and deformation of such media. These embrace and are consistent with a vast variety of different theories of matter, motion, gravitation, and electromagnetism. They represent but a framework into which one can fit more specific theories characterized by different descriptions of the state of a medium and by different constitutive relations relating that state to the inertial mass and the generalized forces. One of the earliest class of problems in the new relativity mechanics to be attacked by Einstein, Minkowski, Abraham, Bateman, and many others, was the construction of a theory of the electromagnetic field in a moving and deforming medium characterized in part by linear constitutive relations

$$D = \epsilon_\circ E , \quad B = \mu_\circ H , \tag{1.1}$$

for the corresponding medium at rest in some inertial Lorentz frame. Maxwell, Lorentz, Voigt, and others had demonstrated how linear constitutive relations like (1.1) and linear generalizations of them which include the effects of small deformation could account with elegance and simplicity for many of the known optical and electro-mechanical properties of solids and fluids. This early work on the construction of "relativistic" counterparts of known classical constitutive relations like (1.1) for stationary media to the case of moving and deformable media, and the controversy surrounding the apparently contradictory results of different investigators is described in the excellent article by Pauli. It is difficult to perceive in this early work general physical principles not conditioned by the linearity of the underlying classical constitutive relations under consideration and which could be relied upon to guide the construction of the corresponding theory of motion of dielectric media in which the polarization is not a linear function of the electromagnetic field. Thus we have abandoned these earlier methods of reasoning and in this concluding lecture attempt to show how the general principles established thus far can guide the construction and physical interpretation of a more definite special theory of deformable dielectric media which does not rest upon the concept of absolute time and Euclidean space.

V. 2. NON-MAGNETIC PERFECT DIELECTRICS

In ordinary terms, by a perfect dielectric one means a perfect electrical insulator. More formally now, we shall define a perfect dielectric as follows. It is a material medium in the sense used in the previous lectures. Let v be the velocity of the medium, and let η be a charge potential in the world-tube of the medium. Thus the charge of any oriented set $\mathcal{E}^3$ of events experienced by points of the dielectric medium is given by

$$C(\mathcal{E}^3, \underset{\approx}{Q}) = \oint_{\partial\mathcal{E}^3} \eta(\underset{\approx}{Q}) \tag{2.1}$$

In a perfect dielectric medium, there exists a charge-potential η such that

$$\eta \wedge v = 0 . \tag{2.2}$$

If $\hat{\eta}$ denotes the dual of η, then (2.2) is equivalent to

$$\hat{\eta} \vee v = 0 . \tag{2.3}$$

This equation asserts that the velocity vector at each event is a divisor of a charge-potential. Thus there exists a field $\hat{P}$ such that

$$\hat{\eta} = \hat{P} \vee v . \tag{2.4}$$

Now $\hat{P}$ is not uniquely determined by η and v and the relation (2.4) since one may add to any solution of (2.4) for $\hat{P}$

a term proportional to v and obtain another solution. There is only one solution, however, such that

$$\hat{P} \circ v^{\dagger} = 0 \quad , \tag{2.5}$$

and we shall assume as part of the definition of the <u>polarization</u> field $\hat{P}$ that it satisfies (2.5).

It should be remarked that according to A6 , $K(\phi)$, where K is the aether tensor and ϕ is the electromagnetic field, is also a charge potential, whether matter experiences the events in question or not. This is the Lorentz viewpoint. Thus we also have

$$C(\mathcal{E}^3, \underset{\approx}{Q}) = \oint_{\partial \mathcal{E}^3} K(\phi, \underset{\approx}{Q}) \quad . \tag{2.6}$$

But the potential of charge, it must be remembered, is not uniquely determined by the distribution of charge. If this were true, then one would infer that $K(\phi) = \text{dual}(\hat{P} \vee v)$, and this is not implied at all by what has been assumed above. Rather, one can only infer that

$$\oint_{\partial \mathcal{E}^3} K(\phi) = \oint_{\partial \mathcal{E}^3} \text{dual}(\hat{P} \vee v) \quad , \tag{2.7}$$

for every submanifold of events $\mathcal{E}^3$ experienced by the material points of a perfect dielectric. Assuming that $K(\phi)$ and $\hat{P} \vee v$ are regular forms, (2.7) is equivalent to

$$\mathrm{rot}[\, K(\phi)\,] = \mathrm{rot}[\,\mathrm{dual}(\hat{P} \vee v)] , \qquad (2.8)$$

or, taking the dual of this equation,

$$\mathrm{div}[K(\phi)] = \mathrm{div}(\, P \vee v\,) \quad , \qquad (2.8')$$

which we can also write in the form

$$\mathrm{div}\, D = 0 \quad , \quad D = K(\phi) + v \vee \hat{P} . \qquad (2.9)$$

Equation (2.9) is one pair of Maxwell's equations relating the charge, current, electric, and magnetic fields in a perfect dielectric medium.

One sees also from the above that the dual, $\hat{\chi}$, of the charge-current field in a dielectric has the special form

$$\hat{\chi} = -(\mathrm{div}\, \hat{P})\, v + \overset{\circ}{\hat{P}} \qquad (2.10)$$

where $\overset{\circ}{\hat{P}} = \underset{v}{\mathcal{L}} \hat{P}$. We call, $-\mathrm{div}\, \hat{P}$, the density of polarization charge or bound charge, and $\overset{\circ}{\hat{P}}$, the current of polarization. Note that (2.10) does not always correspond to a decomposition of $\hat{\chi}$ into its components along v and normal to v since

$$\overset{\circ}{\hat{P}} \cdot v^{\dagger} = \overset{\circ}{\overline{\hat{P} \cdot v^{\dagger}}} - \hat{P} \cdot \overset{\circ}{v}{}^{\dagger}$$

$$= -\hat{P} \cdot a \, , \qquad (2.12)$$

where $a = \overset{\circ}{v}^{\dagger}$ is the acceleration. Thus the current of polarization in an accelerated medium is not always perpendicular to the velocity as is the polarization.

We next assume for this special theory of dielectrics that the inertial mass $\rho = \mu J$ of a simple dielectric medium is a function

$$\rho = U(g, \nabla f^{\dagger}, \tilde{\sigma}, P, X) \quad . \tag{2.13}$$

In words, the inertial mass is a function of the metric field, the deformation gradient, the entropy, and the polarization. By this assumption one excludes any consideration of many other aspects of dielectrics which might be considered in a more general theory such as, for example, strain gradient effects, or diffusion.

The state variables σ and P do not have the property required in the replacement theorem of Lecture V; viz., that their invariance under the motion implies their absolute invariance. For this reason, we replace the state variables σ and P, which appear most natural in the first instance, by the scalar measures of entropy and polarization defined by

$$\tilde{\sigma} = J \hat{\sigma} \tag{2.14}$$

$$\tilde{\Pi} = J \nabla f^{\cdot} \cdot \hat{P} \quad . \tag{2.15}$$

In terms of components relative to an arbitrary system of coordinates about X in $\mathcal{M}$ and ξ in $\mathcal{E}$, this last equation reads

$$\Pi^A = J\, X^A_{\ ,\alpha}\, \hat{P}^\alpha \quad . \tag{2.16}$$

Since $\hat{P}^\alpha v^\dagger_\alpha = 0$, this equation has a unique solution

$$\hat{P}^\alpha = J^{-1}\, P_v{}^\alpha{}_\beta\, \xi^\beta_{\ ,A}\, \Pi^A \tag{2.17}$$

where $P_v = I - s\, v \otimes v^\dagger$ is the projection of $V^4(\xi)$ onto $V^3(\xi, v)$. Thus $\hat{P}$ is a function of the polarization measures Π, the metric, and the deformation gradient. Also, $\hat{\sigma}$ is a function of the scalar entropy measure $\tilde{\sigma}$, the metric, and the deformation gradient. Therefore, the constitutive relation (2.13) is equivalent to one of the form

$$\wp = U^*(g,\ \nabla f^\dagger,\ \tilde{\sigma},\ \Pi,\ X) \tag{2.18}$$

The polarization measure Π is the dual of a 2-form in the material manifold $\mathcal{M}$, but with respect to $\mathcal{E}$ it may be viewed as a set of 0-forms, or scalar fields in the world-tube of $\mathcal{M}$. Therefore, each of the state variables in the constitutive relation (2.18) for the inertial mass is absolutely invariant under the motion of $\mathcal{M}$ if it is invariant,and the replacement theorem applies. Thus if the principle of material indifference is assumed, every constitutive relation for the inertial mass is equivalent to one of the form

$$\rho = U(C , \nabla f^{\dagger} , \Pi , \tilde{\sigma} , X) \tag{2.19}$$

It can be shown in several ways now that this function U must be independent of $\nabla f^{\dagger}$. One way is as follows. Since σ, C, Π, and ρ are scalar fields in one has

$$\mathop{\mathcal{L}}_{w} q = w^{a}\, \partial_{a} q \ , \quad q = \{\sigma, C, \Pi, \rho\} \tag{2.20}$$

Consider then the vector fields $w^{a}_{(\Gamma)} = \dfrac{\partial \xi^{a}}{\partial X^{\Gamma}}$, where $X^{\Gamma} = (X^{A} , \tau)$, τ the parameter of the motion. One then has

$$\mathop{\mathcal{L}}_{w_{(\Gamma)}} q = w^{a}_{(\Gamma)}\, \partial_{a} q = \frac{\partial q}{\partial X^{\Gamma}} \ . \tag{2.21}$$

But

$$\mathop{\mathcal{L}}_{w} \nabla f^{\dagger} = 0 \ , \tag{2.22}$$

so that computing $\dfrac{\partial \rho}{\partial X^{\Gamma}}$ in each of two alternative ways one gets

$$\mathop{\mathcal{L}}_{w_{(\Gamma)}} \rho = \frac{\partial U}{\partial C_{AB}} \mathop{\mathcal{L}}_{w_{(\Gamma)}} C_{AB} + \frac{\partial U}{\partial \Pi^{A}} \mathop{\mathcal{L}}_{w_{(\Gamma)}} \Pi^{A}$$

$$+ \frac{\partial U}{\partial \tilde{\sigma}} \mathop{\mathcal{L}}_{w_{(\Gamma)}} \tilde{\sigma} + \frac{\partial U}{\partial X^{A}} \mathop{\mathcal{L}}_{w_{(\Gamma)}} X^{A} \tag{2.23}$$

$$= \frac{\partial U}{\partial C_{AB}} \frac{\partial C_{AB}}{\partial X^{\Gamma}} + \frac{\partial U}{\partial \Pi^{A}} \frac{\partial \Pi^{A}}{\partial X^{\Gamma}}$$

$$+ \frac{\partial U}{\partial \xi^{a}_{,\Omega}} \frac{\partial^{2} \xi^{a}}{\partial X^{\Omega} \partial X^{\Gamma}} + \frac{\partial U}{\partial \tilde{\sigma}} \frac{\partial \tilde{\sigma}}{\partial X^{\Gamma}}$$

$$+ \frac{\partial U}{\partial X^{A}} \frac{\partial X^{A}}{\partial X^{\Gamma}} .$$

All terms but one in the last equality cancel each other and we are left with the condition

$$\frac{\partial U}{\partial \xi^{a}_{,\Omega}} \frac{\partial^{2} \xi}{\partial X^{\Omega} \partial X^{\Gamma}} , \tag{2.24}$$

which must hold for all motions and coordinate systems. But this implies that the function U is independent of $\nabla f^{\dagger}$. Thus

$$\rho = U(C, \Pi, \tilde{\sigma}, X) . \tag{2.25}$$

Consider next the Lagrange equation (8.7) for a perfect dielectric and denote the generalized force conjugate to the polarization measure Π by $\mathcal{E}/a$, where a is the fine structure constant. Suppose that the Lagrange equation holds for arbitrary variations of the class

$$\overset{*}{g}_{\alpha\beta} = 0 , \quad \overset{*}{\tilde{\sigma}} = 0 . \tag{2.26}$$

This will imply that

$$\frac{\partial \mu}{\partial \Pi^A} = -\left(\frac{1}{a}\right) \mathcal{E}_A \tag{2.27}$$

The electromagnetic energy-momentum vector in a perfect dielectric has the special form

$$f_\alpha = (1/a)\ \phi_{\alpha\beta} \left[-(\mathrm{div}\ \hat{P})\ v^\beta + \overset{\circ}{\hat{P}}{}^\beta \right] \tag{2.28}$$

On taking the Lie derivative of (2.15) with respect to v one gets the following relation between the current of polarization and $\overset{\circ}{\Pi}$.

$$\overset{\circ}{\hat{P}}{}^\alpha + s\, v^\alpha\ a_\beta\ \hat{P}^\beta = J^{-1}\ P_v{}^\alpha{}_\beta\ \xi^\beta{}_{,A}\ \overset{\circ}{\Pi}{}^A\ . \tag{2.29}$$

Consider next the dissipation function (8.15) for the case in hand. Using the results (2.27), (2.28), and (2.29) just established, we find that the dissipation function can be expressed in the form

$$\Phi = S^{\alpha\beta}\ v^\dagger_{(\alpha;\beta)} - h^\alpha (\ln|\theta|)_{,\alpha} + s\, h^\alpha\ a_\alpha$$

$$+ (1/a)\ \mathcal{E}_{(d)A}\ \overset{\circ}{\Pi}{}^A\ , \tag{2.30}$$

where the dissipative-rotary component $\mathcal{E}_{(d)}$ of the generalized force conjugate to the polarization is given by

$$\mathcal{E}_{(d)A} = \mathcal{E}_A - J^{-1}\ \xi^\alpha{}_{,A}\ \phi_{\alpha\beta}\ v^\beta\ . \tag{2.31}$$

The electromotive intensity at a point in the world-tube of the medium is defined by

$$\underset{v}{E}_{\alpha} = \phi_{\alpha\beta} v^{\beta} . \qquad (2.32)$$

Since

$$\overset{\circ}{\Pi}{}^{A} = J X^{A}{}_{,\alpha} \overset{\circ}{\hat{P}}{}^{\alpha} \qquad (2.33)$$

we find that the polarization term in the dissipation function can be expressed in the alternative form

$$\mathcal{E}_{(d)A} \overset{\circ}{\Pi}{}^{A} = \mathcal{E}_{(d)\alpha} \overset{\circ}{P}{}^{\alpha} , \qquad (2.34)$$

where,

$$\mathcal{E}_{(d)\alpha} = \mathcal{E}_{\alpha} - \underset{v}{E}_{\alpha} \qquad (2.35)$$

$$\mathcal{E}_{\alpha} = J^{-1} X^{A}{}_{,\alpha} \mathcal{E}_{A} .$$

Now let

$$\Delta = \{ S_{(d)} , h , \mathcal{E}_{(d)} \} \qquad (2.36)$$

denote the set of fields whose values determine the rate of entropy production in the medium. The constitutive relations for dielectrics consist in the function U whose value determine the inertial mass, and the relations giving the values $\Delta(\xi)$

at each event in the world-tube of $\mathcal{M}$ in terms of the motion and state of the medium. According to the principle of local determinism, $\Delta\ [\xi(X)]$ is a functional of the values of the field variables (g, $\nabla f^{\dagger}$, σ, P) at events $\xi'(X)$ not later than $\xi(X)$. Whatever may be the explicit form of these constitutive relations, they must be consistent with the condition $\Phi \geqslant 0$.

The simplest constitutive relation for Δ consistent with this Clausius-Duhem inequality is

$$\Delta = 0. \tag{2.36}$$

These are the constitutive relations for a perfectly elastic, thermal insulator, which is transparent, optically passive, and a perfect electrical insulator. The next simplest class of dielectrics shares all these properties except that they may be optically active. The constitutive relations for this class are of the form

$$S_{(d)} = 0 , \quad h = 0 ,$$

$$\mathcal{E}_{(d)} = \Gamma \wedge \overset{\circ}{\hat{P}} , \tag{2.38}$$

where Γ is a 2-form called the <u>gyration coefficient</u> whose value at each event is a function of the values of the deformation gradient and electromagnetic field at that event. This class of media is non-dissipative in the sense that $\Phi = 0$.

If a Lorentz frame exists such that $g_{\alpha\beta} = \text{diag}(1, 1, 1, -1)$ in appropriate units, and if the velocity of the medium <u>relative</u> to this frame as measured by the first three components v^i, $i = 1, 2, 3$ of the velocity vector is every small compared with unity; i.e., $V^2 = \sum_{i=1}^{3} v^i v^i \ll 1$, in the units chosen, and if the rate of change of the inertial mass (internal energy) is everywhere small in the sense that

$$|J^{-1} \dot{\rho}| \; V^2 << |\hat{\mu} \quad \dot{v}|$$

the theory of motion of dielectrics and of the electromagnetic field in them based on (2.38) is indistinguishable for all practical purposes from the classical theory of such media described in ""A dynamical theory of elastic dielectrics"". Thus it is unnecessary to repeat here the way in which simple solutions or approximate solutions of the system of equations proposed here can be constructed and interpreted physically in terms of known qualitative electromagnetic and electromechanical properties of elastic dielectrics. The present treatment is superior to the one given earlier which relies on the concept and properties of absolute time. Whenever it is necessary to treat both inertia and the electromagnetic field side by side as in the case of the dynamics of dielectrics, a classical treatment of the former leads inevitably to the existence of a preferred frame of reference relative to which the aether is at rest. Such a frame is both Galilean (inertial) in the classical sense, and a Lorentz frame from the point of view of the electromagnetic equations. All such frames are at rest relative to one another. While the effects of motion relative to this class of frames is small in some sense (an assumption which must be made for strict logical interpretation of the classical theory), certainly the relativistic treatment given here which does not rest on such an hypothesis, can claim greater simplicity.