ANNALS OF PHYSICS: **23**, 233–283 (1963)

Geometric Definition of Gauge Invariance*

ELIHU LUBKIN

Department of Physics, Brown University, Providence 12, Rhode Island

I. INTRODUCTION

This paper is akin to several which have appeared recently (*1–3*), in which the possible physical applicability of mathematical structures is discussed broadly, in the absence however of particular suggestions relating to the current experimental situation. Easily available Lie group theorems (*4–6*) are stated as facts. Further work on conservation laws and structures based on global properties of Lie groups would, after detailed study of the one-dimensional, rotation, and Lorentz groups, entail exhaustive studies of the simple Lie groups as presented in their classification (*7*, *8*, *1*). The exhaustive study of component structures may not easily yield to known mathematical tools, if the difficulties of Speiser and Tarski (*3*) are real.

II. HISTORICAL INTRODUCTION: GENERAL RELATIVITY, ELECTROMAGNETIC THEORY, AND THE YANG-MILLS *B*-FIELD

The feeling that all the standards of reference or fiduciae used to describe the physical situation at a point in space-time are themselves determined by the surroundings is the foundation of Einstein's theory of general relativity. The set of possible fiduciae at a point being given, or physically determined, we set out to compare fiduciae at different points. If the concept of point is merely a device to sort out well-separated phenomena, then the tie connecting one point to an infinitesimally neighboring point should dominate: any statement made about all points at once should rest upon integration, not directly upon axioms. Thus, Weyl saw the essence of general relativity in the determination of the displacement of the fiduciae of physics from one point to another along a curve, by the integration of an infinitesimal displacement. These fiduciae are restricted to be special vector-space bases, and I also abide by this restriction.

Einstein's fiduciae were the tetrads of vectors orthonormal in the Lorentz metric, physically defined by rulers or clocks, and light signals. I call these "inertial frames" (*9*, *10*), and generalize by designating any family of special vector-space bases a family of *frames*.

* Supported in part by the U. S. Atomic Energy Commission under Contract AT(30-1)-2262.

Weyl tried to introduce new fiduciae. He first suggested that the standard or gauge of length be included among the pointwise fiduciae, and found a vector potential subject to gauge transformations (*11*). When complex fields were introduced into physics, Weyl felt the choice of a null phase to be a more reasonable pointwise fiducial choice, and again found a vector potential subject to gauge transformations (*12*). (In this paper he deals with a two-component spinor field so that the complex numbers should not seem arbitrary.)

Classifications of particles have brought many more groups, describable as freedom of choice among pointwise fiduciae, into theoretical physics, and the scope of general relativity is consequently enlarged. This was recognized for choice of orthonormal triads in isotopic spin space, by Yang and Mills (*13*). The idea was finally presented in formal isolation by Utiyama (*14*), who discusses gravitation, electromagnetic theory, and the Yang-Mills theory as examples.

The present paper covers much the same ground that was covered by Utiyama, but with the intent of emphasizing the similarity of the mathematical object in view with the objects of differential geometry. Properties which follow directly from the geometrical concepts are presented without Lagrangians. The problems of obtaining massed particles and of avoiding massless particles—major stumbling blocks to the applicability of Yang-Mills theories—and problems arising from quantization of Lagrangians, will scarcely be mentioned. Conservation laws emerging from global topology of the groups, may be new.

III. SKETCHY DEFINITION OF CONNECTED BUNDLE. ISOTROPY. STRONG HOMOGENEITY. NORMALIZATION OF THE GROUP REPRESENTATION TO A STANDARD FORM

At each point x of space-time, let there be a vector space, and among the bases of this vector space let there be a class of distinguished bases, the "frames." The matrix transformations $f'^{(\alpha)} = L^{(\alpha)}{}_{(\beta)}f^{(\beta)}$ interchanging these frames will be required to constitute a group of matrices, $\mathbf{L}$. The corresponding abstract Lie group will be "the group" or "the symmetry group." That the points x are points of space-time will not usually be essential, so that I will speak of a manifold, rather than of space-time.

There will be a rule for displacing vectors linearly along curves in the manifold, called variously a connection, a parallel displacement, or a displacement. The connection will be tied to the structure of frames, by requiring the displaced of a frame to be a frame.

If any set of bases were designated "frames," then since the matrices in $f'^{(\alpha)} = L^{(\alpha)}{}_{(\beta)}f^{(\beta)}$ and in $f^{(\alpha)} = L'^{(\alpha)}{}_{(\beta)}f'^{(\beta)}$ are inverse, and since $L^{(\alpha)}{}_{(\beta)} = \delta^{(\alpha)}_{(\beta)}$ if $f = f'$, the set $\mathbf{L}$ of matrices for interchanging frames would possess the inverse matrices and the unit matrix, but might not be closed under matrix multiplication. That $\mathbf{L}$ be a group is therefore equivalent to closure under matrix multiplication. This is a condition of isotropy: Let L and L' belong to $\mathbf{L}$, so that there exist frames f_1, f_2, f_3, f_4 with $f_1 = Lf_2$ and $f_3 = L'f_4$. If f_4 is an L image, $f_4 = Lf_5$, then $f_3 = L'Lf_5$, and $L'L$ is in $\mathbf{L}$. But f_1 is known to be an L image. If f_4 is not, then that difference of f_4 from f_1 is a violation of isotropy, defined as the independence of the set of matrices obtained by taking all frames to one from choice of the one frame. Equivalently, that the set $\mathbf{L}$ of matrices relating pairs of frames be obtained in its entirety if the pairs be restricted by arbitrarily fixing one of the frames.

The following homogeneity condition is required: the groups of linear transformations defined at different points by the interchange of frames must be isomorphic groups of linear transformations. The matrices L for interchange of frames at one point may then differ from the L' at another point only by an over-all similarity transformation, $L' = SLS^{-1}$.

That even this difference be not allowed will be called "strong homogeneity." Suppose that we question strong homogeneity, and allow the matrices $L(A)$ which generate arbitrary frames from one frame at A to differ from those appropriate to point B by a similarity transformation, $L(B) = SL(A)S^{-1}$. Choose frame f_0 at A, and let the displaced of these along AB to B be called $f_0'^{(\alpha)}$. The $f_0'^{(\alpha)}$ constitute a frame at B, by our condition that the displaced of a frame be a frame. Every frame $f^{(\alpha)}$ at A is of form $f^{(\alpha)} = L(A)^{(\alpha)}{}_{(\beta)}f_0{}^{(\beta)}$. The displacement is assumed to preserve linear combinations, whence the displaced $f'^{(\alpha)}$ of $f^{(\alpha)}$ is $f'^{(\alpha)} = L(A)^{(\alpha)}{}_{(\beta)}f_0'^{(\beta)}$, is a frame at B, and therefore of form $L_1(B)^{(\alpha)}{}_{(\beta)}f_0'^{(\beta)}$, where $L(A) = L_1(B)$, and the set of all $L(A)$ coincides with that of all $L(B)$, and we may take $S = 1$: Strong homogeneity follows from the assumptions that the ~~parallel~~ displacement acts linearly and that the displaced of a frame is a frame.

Although now $L(B) = L(A)$, we might have $L(B) = L(A) = SLS^{-1}$ in terms of an abstract equivalent representation L not associated with any point. But if $f_0 = Sg_0$ and $f = Sg$, then $g = Lg_0$, so that a redefinition of the frames normalizes the representation to form L. The condition that f-frames be preserved under parallel displacement is equivalent to the similar condition on the g-frames, and any definition of the f-frames determines the g-frames through $f = Sg$, and vice versa.

The structure of identical vector spaces attached to the points of a manifold, together with a symmetry group, is known as a vector bundle. An algebraic term for a vector space with a restricted symmetry group of linear transformations acting on it is a "representation"; the mathematical object in question is therefore more precisely named, "a connected bundle of representations."

IV. PHYSICAL MOTIVATION FOR THE DEFINITIONS. NEAR AND FAR

The mathematical situation presented here, concerning vector fields with components $v_{(\alpha)}(x)$, might be compatible with quantum mechanics; field operators usually bear indices to signify simple vector-space properties. However, the separation between the vector spaces at different points x would seem to run counter to the spirit of quantum mechanics, where effects at different points are bound by interference within a degree of freedom. The structure of a connected vector bundle may need basic modification to meet the needs of quantum mechanics, and may appear unmodified only in classical limits.

The role of spatial coordinates may even be an approximate feature—an introduction of a subdivision of some structure into more or less separated substructures, with the coordinates an approximate concept expressing degree of separation of such substructures. This primitive meaning of coordinates in locating different "objects," and its violation by the large extension of the objects of quantum mechanics on the spatial coordinates, both make me uneasy about accepting them as basic parameters. Of course, nearly everyone feels compelled to be unhappy with the x^μ.

If one ignores such uneasiness, one may write classical field theories based on the x^μ by writing Lagrangian densities, and one may then use "conventional quantization procedures," as is well known.

But if the x^μ do signify merely gross separation, going from one point to the next is complex, and anything connected with an extended path has an even more derivative status; the idea of relating vectors at different points without a connection is then repugnant. One would not necessarily shun x^μ-based Laggrangian densities, but the implicit approximation should necessitate connections for adequate support.

I mention two other possible definitions of connected bundle, and indicate why I have not used them.

First, consider the proposal that the image under parallel displacement of the frames at a point be a family of bases at the new point such that the interchanges of bases within the new family of bases constitute an equivalent representation to that given originally, but that this new family of bases be not required to coincide with the frames at the new point. This is equivalent to dropping the restriction on the parallel displacement that the displaced of a frame be a frame; the connection no longer serves the function of transporting the physical fiduciae represented by the frames. In other words, starting from any preconception of the frames, we seek the most general displaced of these, and form the union, and use that for the frames, thereby reestablishing the former definition.

Second, consider the proposition that if a certain unique rule is used for infinitesimal displacements from a point, it is not guaranteed that after a loop displacement to the original point, the same rule for further infinitesimal displacements is to be obeyed. Then regard the point at the end of the loop different from that at the beginning. This analogue of Riemann surface is so simple here, because here there is no process of continuation which generates itself from a local element. Rejection of multivaluedness may also be founded on the speculation affirming the derivative status of the space-time point, for then the points are physically bonded by the connection, and the bonds may not be altered when we in thought follow along a series of other bonds which loops back.

V. AXIOMATIC DEFINITION OF CONNECTED BUNDLE

A connection requires a vector bundle for its support; a vector bundle is defined over a manifold. "Manifold" and "bundle" are already implicit in the notions "coordinate" and "component." For formal sharpening of these concepts, see ref. *15*; the following remarks may suffice.

A. MANIFOLD

A manifold is a space which may be expressed as a union of open patches U_i, where the points of the ith patch may be labeled by coordinates $(x^1, \cdots, x^m, i)$. The open sets and as much differential structure as is desired are to be given faithfully by the coordinates.

The transformations $f^{\mu}{}_{ji}(x^1, \cdots, x^m, i) = (x'^1, \cdots, x'^m, j)$ for relating different coordinates must satisfy certain obvious requirements so that the open sets and differential structure defined by different coordinatizations agree, and so that the overlapping patches are seamed together properly: The mapping defined by f_{ji} from U_i into U_j will be said to seam U_j onto U_i. The manifold is completely given in coordinates only when besides the U_i, the seaming functions f_{ji} are also given.

For any i and j, it is required that there exist indices $1, \cdots, r$ such that seaming functions $f_{i1}, f_{12}, \cdots, f_{r-1,r}, f_{rj}$ are given; the manifold is to be connected in the topological sense—I distinguish this once from the use of "connected" in connection with affine connection. For the seams to be symmetric, it is required that for each f_{ji} there be an f_{ij} with $f_{ij}f_{ji} = 1$ on the domain of f_{ji}. The seams are smoothed out by requiring that $f_{kj}f_{ji}f_{ik} = 1$ where defined. This is seen to be the familiar "group property" condition of transitivity $\bar{\bar{x}}(x) = \bar{\bar{x}}(\bar{x}(x))$, or $f_{ji}f_{ik} = f_{jk}$ where defined, if one multiplies on the left by f_{jk}.

An equivalence of two coordinatizations is a complete set of admissible mappings f_{ji} and f_{ij} between the patches U_i and U_j of the two coordinatizations where each $(x^1, \cdots, x^m, i)$ is imaged in the other coordinatization, and vice versa, and the ensemble of both patchworks, their seaming functions supplemented with the new f_{ji} and f_{ij} functions, form a coordinatization.

Each point in a coordinatization is required to possess a neighborhood which lies entirely within one patch, the coordinates having been required to give the topology *faithfully*. If the open sets defined in patches U_i were thrown together by union without this condition, then interesting examples of bifurcating and, let us face it, multifurcating structures would be included; e.g., two one-dimensional patches $(x, 1)$ and $(x, 2)$, $0 < x < 1$, with $f_{12}(x, 1) = (x, 2)$ for $0 < x < \frac{1}{2}$.

The exclusion of the branching structures is a condition of homogeneity, inasmuch as the sites of branching are obviously special points, possible pathological objects which branch symmetrically everywhere excluded!

B. Representation Bundle. Frames

At each point x of a manifold there is given an n-dimensional vector space, $V(x)$. There is an open covering by patches U_i, not necessarily the patches used to define the manifold itself, with the vectors given at point x in the ith patch by symbols $(v_{(1)}, \cdots, v_{(n)}, x, i)$. The set of all these as x ranges over the ith patch and the $v_{(\alpha)}$ range freely is the vectored patch V_i. The open sets and as much differential structure as desired are specified by these numbers. For these concepts to be invariant, the transformation functions will have to be continuous and differentiable a sufficient number of times.

An admissible transformation of free vectors is obtained through an n by n matrix $L_{(\alpha)}{}^{(\beta)}$ belonging to $\mathbf{L}$, a fixed Lie group of n by n matrices; $v'_{(\alpha)} = L_{(\alpha)}{}^{(\beta)} v_{(\beta)}$.

An admissible transformation $L_{(\alpha)}{}^{(\beta)}(x)$ of the set of vectors in all the $V(x)$ as x ranges over some set is continuous from the point x into $\mathbf{L}$, and differentiable as may times as desired for the purpose of defining the concept of derivative invariantly to a sufficient order. The jth vectored patch V_j is said to be seamed onto the ith vectored patch by an admissible transformation $L_{(\alpha)}{}^{(\beta)}{}_{ji}(x)$ defined for x in $U_i \cap U_j$, if $L_{(\alpha)}{}^{(\beta)}{}_{ji}(x)$, abbreviated $L_{ji}(x)$, maps from V_i into V_j thus: $v'_{(\alpha)} = L_{(\alpha)}{}^{(\beta)}{}_{ji}(x) v_{(\beta)}$, with $(v_{(1)}, \cdots, v_{(n)}, x, i)$ taken to $(v'_{(1)}, \cdots, v'_{(n)}, x, j)$. The component bundle is given by the patchwork of U_i's together with a given seaming function $L_{ji}(x)$ on each (nonempty) $U_i \cap U_j$. The seams are made symmetric by requiring $L_{ij}(x)$ to be inverse to $L_{ji}(x)$, for x in $U_i \cap U_j$. The seams are made smooth or transitive by requiring that $L_{kj}(x)L_{ji}(x)L_{ik}(x) = 1$, for x in $U_i \cap U_j \cap U_k$. This may be restated as a group property, and the corresponding statement for longer chains of L_{ji}'s follows.

The above is a component representation bundle.

A strict equivalence of component bundles on the same manifold is given by L_{ji} and L_{ij} seaming the two different patchworks together, so that these and the seaming functions given for the separate bundles together satisfy the conditions for seaming functions.

The structure of distinguished bases or frames is already present. If we write the vector $v(x)$ given in a patch by $(v_{(1)}, \cdots, v_{(n)}, x, i)$ as $v_{(\alpha)}(x, i)f^{(\alpha)}(x, i)$, where $(f^{(\alpha)}(x, i))_{(\beta)} = \delta^{(\alpha)}_{(\beta)}$, the basis $f^{(\alpha)}(x, i)$ at x associated with the ith patch is made explicit. By $f^{(\alpha)}(x, i)$ is meant the invariant vectors which assume components $\delta^{(\alpha)}_{(\beta)}$ in the ith patch; for x in $U_i \cap U_j$, $v(x) = v_{(\alpha)}(x, i)f^{(\alpha)}(x, i) = v_{(\alpha)}(x,j)f^{(\alpha)}(x,j)$, with $v_{(\alpha)}(x,j) = L_{(\alpha)}{}^{(\beta)}{}_{ji}(x)v_{(\beta)}(x, i)$, and $f^{(\alpha)}(x, j) = L^{(\alpha)}{}_{(\beta)ji}(x) \cdot f^{(\beta)}(x, i)$, where $L^{(\alpha)}{}_{(\beta)}L_{(\gamma)}{}^{(\beta)} = \delta^{(\alpha)}_{(\beta)}$. Also $f'^{(\alpha)} = L^{(\alpha)}{}_{(\beta)}f^{(\beta)}$; on the f bases,

$(f'^{(\alpha)})_{(\beta)} = L^{(\alpha)}{}_{(\beta)}$: The general frame is the ordered set of the columns of an $\mathbf{L}^{(\alpha)}{}_{(\beta)}$-matrix.

The frames are therefore a general way of speaking of a group representation by matrices. This is a triviality, as it is obvious that when one starts with components one is speaking also of bases, and that one cannot avoid speaking of components if one wishes to apply a matrix, but it is yet worth remarking, inasmuch as the original language suggested by inertial frames and the derived Lorentz transformations in which one starts with the frames may at first thought seem less general than it is.

C. Admissible Connection

The connection is to respect the differential structure of the bundle, which has been created for its support, and of the Lie group $\mathbf{L}$. Thus, given a smooth curve $x(\lambda)$, the $L_{(\alpha)}{}^{(\beta)}(\lambda)$ matrix for transporting vectors from $x(0)$ to $x(\lambda)$ is to be continuous and differentiable in λ;

$$L_{(\alpha)}{}^{(\beta)}(\lambda) = \delta^{(\alpha)}_{(\beta)} + \frac{dL_{(\alpha)}{}^{(\beta)}}{d\lambda}(0) + o(\lambda)$$

in the limit $\lambda \to 0$, and

$$\frac{dL_{(\alpha)}{}^{(\beta)}}{d\lambda}(0) = c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x(0))\,\frac{dx^{\mu}(\lambda)}{d\lambda},$$

and since we wish bonds between points, not curves, to be basic, the $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$ are to be functions of x independent of the curves $x(\lambda)$ to which they are applied.

Curvature will entail derivatives of the $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$; currents, derivatives of the curvature; and we shall take derivatives of the currents, so we shall need fourfold differentiability of the infinitesimal components of connection, $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$. For this to have invariant meaning, the functions $f^{\mu}{}_{ji}(x^1, \cdots, x^m, i)$ and $L_{(\alpha)}{}^{(\beta)}{}_{ji}(x, i)$ in the definitions of manifold and bundle must be at least fourfold differentiable.

The $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$, which give the infinitesimal displacements, are sufficient to define the finite displacements along smooth curves contained in a single vectored patch, because if such a path is divided into n segments each with coordinate differences less than some mesh $\delta \sim 1/n$, then the displacements are given individually with error $o(\delta)$, and hence the matrix product of the $\delta^{(\beta)}_{(\alpha)} + c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)\Delta x^{\mu}$ taken in the proper order will differ from the finite displacement by an error $\sim no(\delta)$, or $no(1/n) \to 0$ as $n \to \infty$.

A vector is transported across a patch seam by stopping at any point in the seam, and switching over to the same invariant vector at that point in the new patch.

Since $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)\,dx^{\mu} = c_{(\alpha)}{}^{(\beta)}(x)$ is to be an invariant matrix form,

$$c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x'^1, \cdots, x'^m, j) = \frac{\partial x^\nu}{\partial x'^\mu} c_{(\alpha)}{}^{(\beta)}{}_{\nu}(x^1, \cdots, x^m, i). \tag{1}$$

If displacement along a curve C from x to x', lying entirely within each of several patches U_i, induces the transformation $v_{(\alpha)}(x', i) = L_{(\alpha)}{}^{(\beta)}{}_i(x', x)v_{(\beta)}(x, i)$, when expressed in the ith patch, then

$$L_{(\alpha)}{}^{(\beta)}{}_i(x', x) = L_{(\alpha)}{}^{(\gamma)}{}_{ij}(x')L_{(\gamma)}{}^{(\delta)}{}_j(x', x)L_{(\delta)}{}^{(\beta)}{}_{ji}(x). \tag{2}$$

$$\begin{aligned} c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x, i) &= L_{(\alpha)}{}^{(\gamma)}{}_{ij}(x)c_{(\gamma)}{}^{(\delta)}{}_{\mu}(x, j)L_{(\delta)}{}^{(\beta)}{}_{ji}(x) \\ &\quad + \frac{\partial L_{(\alpha)}{}^{(\delta)}{}_{ij}(x)}{\partial x^\mu} L_{(\delta)}{}^{(\beta)}{}_{ji}(x), \end{aligned} \tag{3}$$

or

$$c_{(\alpha)}{}^{(\beta)}(x, i) = L_{(\alpha)}{}^{(\gamma)}{}_{ij}(x)c_{(\gamma)}{}^{(\delta)}(x, j)L_{(\delta)}{}^{(\beta)}{}_{ji}(x) + dL_{(\alpha)}{}^{(\delta)}{}_{ij}(x)L_{(\delta)}{}^{(\beta)}{}_{ji}(x),$$

or

$$c'_{(\alpha)}{}^{(\beta)}{}_{\mu}(x) = L_{(\alpha)}{}^{(\gamma)}(x)L^{(\beta)}{}_{(\delta)}(x)c_{(\gamma)}{}^{(\delta)}{}_{\mu}(x) + L^{(\beta)}{}_{(\delta)}(x)\frac{\partial L_{(\alpha)}{}^{(\delta)}(x)}{\partial x^\mu}.$$

The first term alone would make the connection a tensor; the second is the usual term proportional to the gradient of the transformation, and not involving the original components of connection. Note that the term $L^{(\beta)}{}_{(\delta)}\, \partial L_{(\alpha)}{}^{(\delta)}/\partial x^\mu$ may replaced by $-\partial L^{(\beta)}{}_{(\delta)}/\partial x^\mu L_{(\alpha)}{}^{(\delta)}$ or by half the sum of these.

Conversely, if the transformation law relating quantities in different vectored patches labeled by i is given for the $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x, i)\, dx^\mu$ then the matrices $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x, i)\Delta x^\mu$ for corresponding intervals along a curve C from x to x' must agree to first order in the linear transformations defined on invariant vectors, whence the properly ordered matrix product must converge to $L_i(x', x)$'s related by the finite transformation law (2) for a curve C in $U_i \cap U_j$.

Since

$$L_{(\alpha)}{}^{(\beta)}(\lambda) = \delta^{(\beta)}_{(\alpha)} + \lambda c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x(0))\frac{dx^\mu}{d\lambda}(0) + o(\lambda)$$

for a curve $x(\lambda)$, each $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$ for fixed x and μ is a matrix in the infinitesimal algebra of **L**. If a basis of matrices $E_{I(\alpha)}{}^{(\beta)}$ is chosen in this Lie algebra, we may write

$$c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x) = c^I{}_{\mu}(x)E_{I(\alpha)}{}^{(\beta)}, \tag{4}$$

which separates the free parameters $c^I{}_\mu(x)$ from the restriction on $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x)$ that it belong to the given Lie algebra of matrices. The $c^I{}_\mu(x)$ for fixed I and x are of course components of a covariant tangent vector; in invariant notation, $c_{(\alpha)}{}^{(\beta)}(x) = c^I(x)E_{I(\alpha)}{}^{(\beta)}$.

The extraction of an E_I basis from the law of transformation between vectored patches is complicated, but if $L_{(\alpha)}{}^{(\beta)}{}_{ij}(x) = \delta^{(\beta)}_{(\alpha)}$, then $\partial L_{(\alpha)}{}^{(\beta)}{}_{ij}(x)/\partial x^{\mu} = \lambda^{I}{}_{ij\mu}(x)E_{I(\alpha)}{}^{(\beta)}$, and $c^{I}{}_{\mu}(x, i) = c^{I}{}_{\mu}(x, j) + \lambda^{I}{}_{ij\mu}(x)$.

D. Displacement Along a Curve

Suppose the $c_{(\alpha)}{}^{(\beta)}{}_{\mu}(x, i)$ are given, where i designates a fine patch—the intersection of a manifold patch, and a bundle patch—subject to the laws (1), (3) in fine-patch overlaps, for enough fine patches to cover the manifold. Is the connection itself—the law for displacing vectors along smooth curves—thereby determined? It is:

The curve C runs from x to x', the vector v at x is displaced along C into the vector v' at x', if we already have the invariant connection. What we actually have is that by computing ordered-product line integrals of infinitesimal components of connection within fine patches, and changing notation for the point or vector at a finite set of seam points, the object $(x^1, \cdots, x^m, v_{(1)}, \cdots, v_{(n)}, i)$ is taken to the object $(x'^1, \ldots, x'^m, v'_{(1)}, \cdots, v'_{(n)}, j)$, where the ith fine patch lies over the initial invariant point x, the jth over x'. There is a parallel statement for a second fine patchwork that is appropriately related to the first by seaming functions, in such a way that we may unite both patchworks to form a single patchwork.

Both laws for displacement are therefore alternative ways to obtain the displacement for the case of the redundant patchwork. Thus, the theorem rests directly on the question of whether the mixed procedure of computing integrals within fine patches and changing notation at selected seam points in a single patchwork is independent of choice of patches and seam points—except for the obvious change of notation for the initial and final vectors due to changing frames at the start and finish.

We wish to show that the matrix for displacement in the first manner is equal to that for rephrasing, displacement in the second manner, and then reverse rephrasing. The content of Eq. (2) is that this is true if both manners do not involve internal seam points, the path there being completely contained in a single fine patch in both manners.

More generally, there are some seam points in both manners. Let those points which are seam points in one manner but not in the other constitute trivial seam points in the other, so that there is only one set of seam points $x_1, \cdots, x_s$. The curve C from x to x' is a succession of segment curves, C_0 from x to x_1, C_1 from x_1 to $x_2, \cdots, C_s$ from x_s to x'. For the rth segment, it is known from Eq. (2) that the matrix $D_1(r)$ yielding displacement in the first manner equals $R_{12,r}(x_{r+1})D_2(r)R_{21,r}(x_r)$, where $R_{21,r}(x_r)$ is the matrix for rephrasing from the first to the second manner at x_r in the patch known to wholly contain C_r. The entire displacement in the first manner is $D_1 = D_1(s)R_1(x_s) \cdots D_1(1)R_1(x_1)D_1(0)$, where $R_1(x_r)$ is the patch-to-patch rephrasing of the first manner at the rth seaming point. Therefore, $D_1 = R_{12,s}(x')D_2(s)R_{21,s}(x_s) \cdots R_{12,1}(x_2)D_2(1) \cdot R_{21,1}(x_1)R_1(x_1)R_1(x_1)R_{12,0}(x_1)D_2(0)R_{21,0}(x)$. If $R_{21,r}(x_r)R_1(x_r)R_{12,r-1}(x_r) = R_2(x_r)$, then $D_1 = R_{12,s}(x')D_2(s) \cdots D_2(1)R_2(x_1)D_2(0)R_{21,0}(x)$, whence $D_1 = R_{12,s}(x') \cdot D_2R_{21,0}(x)$; q.e.d.

The essential point is the equation, $R_{21,r}(x_r)R_1(x_r)R_{12,r-1}(x_r) = R_2(x_r)$. If x_r is a trivial seaming point for the second manner, $R_2(x_r) = 1$, and the equation is the three-patch seaming axiom. If $R_1(x_r) = 1$, the same axiom applies if one uses, e.g., $R_{21,r}(x_r) = (R_{12,r}(x_r))^{-1}$, the two-patch seaming axiom. If x_r is a true seaming point for both manners, the equation is the true four-patch analogue of the two- and three-patch seaming axioms. Finally, if the seam is a manifold, not a bundle, seam, the matrices are all 1; the manifold patches are only to make up for the expression of the matrix-product integral displacement in terms of literal coordinate differences.

VI. ALGEBRA

Constructions used in matrix representation theory extend quite simply to the theory of representation bundles. I mean such things as field-automorphic image of the matrices, homomorphism, direct sum, tensor product, and direct product.

On the level of matrix representations, we have matrix group(s) $\mathbf{L}_i$ (there may be only one), and a rule for building a matrix L out of a list, L_i in $\mathbf{L}_i$, so that the matrices L constitute a group $\mathbf{L}$, and such that this mapping from the list of $\mathbf{L}_i$ onto $\mathbf{L}$ preserves Lie group properties. On the level of representation bundles, we consider a single manifold, with representation space(s) $V_i(x)$ at each point x, with $\mathbf{L}_i$ acting on $V_i(x)$, and each with a connection $c_i(x)$, and similarly for $\mathbf{L}$, $V(x)$, $c(x)$. How may the $L_i \to L$ construction be related to the seaming of patches and to a possible $c_i(x) \to c(x)$ construction?

There are really two questions. First, given $\mathbf{L}_i$ and $c_i(x)$ so that we already have a connected bundle for each i over a common manifold, how do we extend an $\mathbf{L}_i \to \mathbf{L}$ construction to build a connected $\mathbf{L}$-bundle over the manifold? Second, given a connected $\mathbf{L}$-bundle such that $\mathbf{L}$ may be regarded as the image of an $\mathbf{L}_i \to \mathbf{L}$ construction, can one produce $\mathbf{L}_i$-bundles over the manifold of the $\mathbf{L}$-bundle and $c_i(x)$'s so that the L-bundle and connection reappear as the answer to the first question?

A generally affirmative answer to the first question is easily obtained, as follows. The matrix $L_i{}^{(\alpha)}{}_{(\beta)}$ is equivalent to an i-frame, the ordered list of its columns. The construction, $(L_i) \to L$, is therefore already a construction which assigns to a family of frames, one for each i, a resultant $\mathbf{L}$-frame. At the same time, the L_i matrices used in seaming patches map to an L matrix. The axioms which declare that chains of L_i seaming matrices at x beginning and ending in the same coordinate patch are the identity at x are inherited by the L matrices owing to the group-homomorphic character postulated for the $(L_i) \to L$ construction. Finally, the connection, given as a law for assigning a matrix L_i for each i to a smooth path, is obviously extended by assigning that matrix L to the path which is the algebraic construct of the L_i ; the displacement is defined so that parallel displacement and the algebraic construction are commutative.

The answer to the second question is generally negative, except for the general class of constructions $(L_i) \to L$ which are uniquely invertible. In the latter case, each $L \to L_i$ will give an $\mathbf{L}_i$-bundle in the manner of the first question. Even if $(L_i) \to L$ is a local isomorphism, however, $\mathbf{L}$-bundles may be found which cannot be built algebraically from $\mathbf{L}_i$-bundles.

Examples

1. Automorphism of Numbers. Here there is only one "$\mathbf{L}_i$", which we therefore call "$\mathbf{L}$." The original matrices $L_{(\alpha)}{}^{(\beta)}$ and $L^{(\alpha)}{}_{(\beta)}$ have their matrix elements replaced by nontrivially automorphic images, $L_{(\dot{\alpha})}{}^{(\dot{\beta})}$ and $L^{(\dot{\alpha})}{}_{(\dot{\beta})}$. For the reals, this is impossible; for the complexes, there is complex conjugation, $L_{(\dot{\alpha})}{}^{(\dot{\beta})} = (L_{(\alpha)}{}^{(\beta)})^*$. Matrix elements do not necessarily belong to an order-complete field, so that it is conceivable that this option of automorphism be more interesting than complex conjugation, without going to characteristic p. An infinitesimal element $c^I E_{I(\alpha)}{}^{(\beta)}$ is taken to $c^{\dot{I}} E_{\dot{I}(\dot{\alpha})}{}^{(\dot{\beta})}$, where $c^{\dot{I}}$ is the automorphic image of c^I. The one-one character of this operation makes the answer to the second question affirmative. The mapping of matrices may be regarded as a consequence of automorphic mapping $v_{(\alpha)} \to v_{(\dot{\alpha})}$ of vector components, and the requirement that the mappings of matrices and vectors be commutative.

2. Homomorphism. Again, only one "$\mathbf{L}_i$." There is a homomorphism $L_{(\alpha)}{}^{(\beta)} \to M_A{}^B$; $L^{(\alpha)}{}_{(\beta)} \to M^A{}_B$ of the matrices, which as in the general remarks may be considered a mapping of $\mathbf{L}$-frames onto $\mathbf{M}$-frames, and the seaming functions and connection matrices are directly imaged. But a single $\mathbf{L}$-vector at x might have no specific $\mathbf{M}$-vector image. This cannot be smoothed over by construction: consider the two-one homomorphism of the spin ½ representation of rotations onto the spin 1 representation.

3. Admissible Vector Homomorphism onto. This is the special case of $\mathbf{L} \to \mathbf{M}$ homomorphism which is borne by vectors. Let us use "u" for a typical vector associated with the $\mathbf{L}$-bundle, "v" for the M-bundle. Then we are given a linear mapping $u \to v$ such that if $L \to M$, then $Lu \to Mv$. If a u-basis is chosen which extends a basis of the kernel, and if the images of the vectors used to extend a kernel basis are used as a basis in the image space, the matrices $\mathbf{L}$ leave the kernel invariant and are therefore reduced (though not necessarily completely reduced), and the matrices $\mathbf{M}$ coincide with that diagonal bloc of the matrices $\mathbf{L}$ which acts on the (arbitrary) complement of the kernel. Such choice of basis has been legitimated at the end of Section III. Explicitly: $u_{(\alpha)} \to v_A$ by a linear map; $v_A = h_A{}^{(\alpha)} u_{(\alpha)}$.

The action of L: $(Lu)_{(\alpha)} = L_{(\alpha)}{}^{(\beta)} u_{(\beta)}$.
of M: $(Mv)_A = M_A{}^B v_B$.

The condition of admissibility:

$$L_{(\alpha)}{}^{(\beta)} u_{(\beta)} \to M_A{}^B h_B{}^{(\beta)} u_{(\beta)} = h_A{}^{(\alpha)} L_{(\alpha)}{}^{(\beta)} u_{(\beta)},$$

whence

$$M_A{}^B h_B{}^{(\beta)} = h_A{}^{(\alpha)} L_{(\alpha)}{}^{(\beta)}.$$

Let primes designate components on the convenient bases:

$$u'_{(\alpha)} = S_{(\alpha)}{}^{(\beta)} u_{(\beta)}; \qquad v'_A = T_A{}^B v_B.$$

$v' = h'u'$; $Tv = h'Su$; $v = T^{-1}h'Su = hu$, so that $h = T^{-1}h'S$. If we employ the notation $T^C{}_B T_A{}^B = \delta^C_A$, we may write this $h_A{}^{(\alpha)} = T_A{}^B h'_B{}^{(\beta)} S_{(\beta)}{}^{(\alpha)}$. The map $h'_B{}^{(\beta)}$ simply annihilates $u'_{N+1}, \cdots, u'_{(n)}$ if $N - n$ is the dimension of the kernel and if we list the kernel basis last for convenience, and the map preserves the other u' components, so that $h'_B{}^{(\beta)} = \delta_B^{(\beta)}$, except for illegally high values of B; $h_A{}^{(\alpha)} = T^B{}_A \delta_B^{(\beta)} S_{(\beta)}{}^{(\alpha)} = T^B{}_A S_B{}^{(\alpha)}$. Also, $M'_A{}^B \delta_B^{(\beta)} = \delta_{(A)}^{(\alpha)} L'_{(\alpha)}{}^{(\beta)}$, or $M'_A{}^B = L'_A{}^B$, i.e., the $1, \cdots, N$ diagonal block of L'. One can define $h_{(\alpha)}{}^A$, mapping the V space onto the chosen N-dimensional subspace of the U space, so as to be inverse to $h_A{}^{(\alpha)}$ there, so that $h_B{}^{(\alpha)} h_{(\alpha)}{}^A = \delta^A_B$, but $h_{(\beta)}{}^A h_A^{(\alpha)} = \varpi_{(\beta)}{}^{(\alpha)}$, a nonidentical projection when the kernel is not $\{0\}$. $h_{(\alpha)}{}^A = S^B{}_{(\alpha)} T_B{}^A$; $\varpi_{(\beta)}{}^{(\alpha)} = S^B{}_{(\beta)} S_B{}^{(\alpha)}$.

If we regard choice of basis in the image space as trivial, then we may put $T_A{}^B = \delta^B_A$. Then $h_A{}^{(\alpha)} = S_A{}^{(\alpha)}$ and these are the coefficients used for projecting irreducible quantities out of more complex quantities; neglect of T is indeed in order when we consider defining the irreducible quantities.

In both types of homomorphism, the $L \to M$ map defines an infinitesimal map $E \to E'$, and $c_I E^I \to c_I E'^I$. In the case of a homomorphism which is locally isomorphic, all the c_I are effective, and not otherwise.

4. Isomorphism. In the case of isomorphism $\mathbf{L} \to \mathbf{M}$ of the matrix groups, the first and second questions are obviously answered in the affirmative, as we have homomorphisms both ways, although it is not necessary that there be one-one correspondence of vectors: e.g., when the vector spaces are of unequal dimension.

5. (Global) Vector Isomorphism onto. The second question has been answered affirmatively in Section III, where the phrasing "normalization of the representation to standard form" was used. The vectors here correspond, by means of a similarity transformation.

6. Direct Sum. By a similarity transformation $\mathbf{L} \to S\mathbf{L}S^{-1}$, the matrices can be standardized to diagonal bloc form. That this can be done for the matrices is the algebraic notion of direct sum; since the set $\mathbf{L}$ is independent of the point

x by strong homogeneity, one x-independent similarity transformation S does this at once for all points of the manifold. Then all seaming matrices and displacement matrices may be considered interchangeably, either as direct sum matrices, or else as ordered lists of matrices, one acting in each direct summand space; the first and second questions are answered affirmatively, as a special case of vector-borne isomorphism. In particular, even when there are several isomorphic direct summands, the restriction on the parallel displacement that its matrix always belong to the group **L** prevents mixture of direct summands. Isolation of a single direct factor is an example of vector-borne homomorphism.

7. *Direct Product.* Direct sum of matrices is the obvious vehicle for direct product of the abstract groups of the matrices, in the sense that ordered sets of group elements are the objects, and these are multiplied coordinatewise. The situation is confused by the use of "direct product" for the tensor product of matrices; by direct product of matrices is meant the direct sum where the groups belonging to the separate summands are independent. It is only for this special case of direct sum that the association of the summands is trivial, for although the transformations in patching or displacement of the summand-space vectors occur within the separate spaces without mixing, the several vectors are transformed at once by a common group element, which may be taken to be the direct sum matrix L itself. If the group is a direct product, however, this correlation of the transformations is empty.

8. *Tensor Product.* As for all the operations discussed here except for the direct homomorphism of matrices, the operation of tensor product may be based on the vector spaces. The multilinear functions $t(v_1, \cdots, v_r)$ on a list of vector spaces V_i, expressed on a basis through coefficients $t^{(\alpha_1, \cdots, \alpha_r)}$, thus: $t(v_1, \cdots, v_r) = t^{(\alpha_1, \cdots, \alpha_r)} v_{1(\alpha_1)} \cdots v_{r(\alpha_r)}$, are the tensors. Invariance of the function value $t(v_1, \cdots, v_r)$ to choice of basis implies that the tensor coefficients suffer the transformation $\prod_i L_i{}^{(\alpha_i)}{}_{(\beta_i)}$ when the vector components transform via $L_{i(\alpha_i)}{}^{(\beta_i)}$. If $r = 1$, we have the case of the dual of a vector space. The map $(L_{1(\alpha_1)}{}^{(\beta_1)}, \cdots, L_{r(\alpha_r)}{}^{(\beta_r)}) \rightarrow \prod_i L_i{}^{(\alpha_i)}{}_{(\beta_i)}$ already implies an assignation of a frame to a choice of r frames; in fact, the basis belonging to the $t^{(\alpha_1, \cdots, \alpha_r)}$ components is, say, a lexicographic ordering of those multilinear functions $f_{(\alpha_1, \cdots, \alpha_r)}$ which are defined by $f_{(\alpha_1, \cdots, \alpha_r)}(f_1{}^{(\beta_1)}, \cdots, f_r{}^{(\beta_r)}) = \delta^{(\beta_1)}_{(\alpha_1)}, \cdots \delta^{(\beta_r)}_{(\alpha_r)}$ in terms of the vector-space bases. A dualization is usually applied to define "tensor product" of vectors, so as to have tensor product of matrices reduce simply to product of the components, rather than that of those of the contragredient matrices.

For the basic infinitesimal operators, we have

$$E_I{}^{(\alpha_1, \ldots, \alpha_r)}{}_{(\beta_1, \cdots, \beta_r)} = \sum_j \prod_{i \neq j} \delta^{(\alpha_i)}_{(\beta_i)} E_{I_j}{}^{(\alpha_j)}{}_{(\beta_j)},$$

with $E_{I_i}{}^{(\alpha_i)}{}_{(\beta_i)} = -E_{I_i(\alpha_i)}{}^{(\beta_i)}$. The c^I in $c^I E_I$ are preserved, as the tensor product is defined to be linear. The application of c^{I}'s in parallel displacement and in covariant derivative will consequently follow the familiar rule of treating each index separately and then summing the separate terms thus obtained.

It is conventional when speaking of the algebra of tensors to introduce possible powers of various determinants in the transformation law for the tensor components, and thus speak of various kinds of weighted tensors and pseudotensors. Since the determinant to some power is a one-dimensional representation, the discussion of weights is comprised in the general tensor-product discussion. Determinants are introduced, more explicitly, by taking the totally antisymmetric part (homomorphism) of the tensor product of an n-dimensional vector space with itself, n times. Pseudocharacteristics associated with disconnection of the group into components are also examples of tensor products, wherein one matrix factor consists of a representation which is the identity on the identity component.

If the given representation is equivalent by a similarity transformation to matrices $L^{(\alpha_1,\cdots,\alpha_r)}{}_{(\beta_1,\cdots,\beta_r)} = \prod_i L_i{}^{(\alpha_i)}{}_{(\beta_i)}$, then the vectors with components may be regarded as transformed in consequence of their definition as multilinear functions on r vector spaces, and since we are speaking of vector-based isomorphism, this representation of the t-vector space as a space of tensors does not conflict with the seaming functions or the displacement; both the first and second questions have affirmative answers with respect to representing the given t's as tensors. Again, either one group may act in the separate factor spaces by various representations, or several groups may act independently, the second case being a special case of the first wherein the over-all group is a direct product. If the group acting on the vectors is imaged homomorphically, not isomorphically, to obtain the group acting on the tensors, then the second question may have no answer in the form of inverse image vector bundles, for a suitably chosen example of a tensor bundle.

Therefore the algebraic operations involved in classifying semisimple groups in terms of simple groups, and those of building representations by reduction of tensor products of others, are applicable to a representation bundle whenever applicable to the algebraic matrix representation itself, except that certain bundles of algebraic constructs may exist without inverse image bundles of their algebraic ingredients, if the construction is many-one.

VII. SPINORS AND VECTORS

Some interesting examples of the algebraic relationships between representation bundles are obtained by considering the relation between spinors u^a over 4-space, and ordinary tangent vectors $v_{(\mu)}$, given with respect to frames. Everyone is familiar with the construction $v^{(\mu)} = \sigma^{(\mu)}{}_{\dot{a}b}\, u^{\dot{a}b}$, wherein the Pauli matrices 1, σ_x , σ_y , σ_z map spin tensors to vectors. If the $v^{(\mu)}$ are allowed to be complex, the map is generally one-one; real $v^{(\mu)}$ correspond to Hermitean $u^{\dot{a}b}$. The construction involves automorphism (complex conjugation), tensor product, and an isomorphism given by the Pauli matrix coefficients. It may be varied by prefixing the spinor side of the equation by powers of the determinant of the $u^{\dot{a}b}$ matrix or of its complex conjugate, whereby one defines variously "weighted" vectors (*16*).

The full linear group acting on 2-spinors has 8 real dimensions. The subgroup of determinant 1 has 6 real dimensions, and is the familiar spin-Lorentz group. The two other dimensions are a real scaling, and a phase factor. Therefore, whatever vectors form the image of an algebraic structure built on spinors, the matrix group which acts on the vectors must be a subgroup of the direct product of the Lorentz group and two real one-dimensional groups, in a neighborhood of the identity. The complex conjugation in $u^{\dot{a}b}$'s vitiates the phase-factor degree of freedom, so that the $\sigma^{(\mu)}{}_{\dot{a}b}\, u^{\dot{a}b}$ construction leads to a group of only seven real dimensions. By use of determinantal prefixes, this can be lowered to six or raised to eight, or the seventh scale-factor dimension may be replaced by the phase-factor dimension. The extra dimensions beyond six correspond to ordinary gauge degrees of freedom, and when the representing matrices function as a connection, a separate electromagnetic potential is introduced for each such extra degree of freedom. The phase-factor degree of freedom has a different global character from the real-stretch degree of freedom, but this difference is effective only if we use it to build magnetic monopoles.

In ref. *10*, the relation between coefficients of connection $k^{a}{}_{b(\mu)}$ and $c^{(\nu)}{}_{(\varpi\mu)}$ for transporting spinors and vectors was discussed; the rest of this paragraph should dispel the mystery in such discussions. Vector frames were the inertial frames interrelated by the usual six-parameter Lorentz group, but spinors were handled freely, so that the full linear eight-real-parameter group was tacitly assumed for them. Since the two extra degrees of freedom available for the spinor connection are associated with an over-all numeric factor, they appeared as coefficients of the unit Pauli matrix. The construction of vectors destroyed the phase factor, so only the real-scaling degree of freedom appeared, beyond the six from the Lorentz group, as an extra degree of freedom for the vector connection derived from the spinor connection. A true analogue of the restriction of the vector transformations to Lorentz transformations by choice of inertial-frame tetrads is the choice of inertial-frame spinor pairs (u_1, u_2) such that the matrices involved in a frame transformation $(u_1, u_2) \to (u'_1, u'_2)$ have determinant 1. Then the infinitesimal components of connection for spinors do not have the extra degrees of freedom.

VIII. TANGENT SPACES

Notation for tangent-space vectors is set out here.

Covariant tangent vectors at $x^{\mu} = x_0^{\mu}$ or at $x'^{\mu} = x'^{\mu}_0$, $\mu = 1, \cdots, m$, whose components transform according to $v'_{\mu}(x'_0) = (\partial x^{\nu}/\partial x'^{\mu})_{(x=x_0)} v_{\nu}(x_0)$, *are con-*

structible by defining the vectors to be the sets of functions f differentiable at the point x_0 with the functions in any one set possessing a common value $v_\mu(x_0) = (\partial f/\partial x^\mu)(x_0)$ for their respective derivatives, when these are computed in terms of one common set of coordinates, x^μ. These vectors form an m-dimensional vector space, the tangent space of covariant vectors at x_0. The bases implied by the component notation are the ordered set of $e^\mu(x_0)$, where $e^\mu(x_0)$ is that set of functions differentiable at x_0 which coincide with the μth coordinate function x^μ to first order in the x^ν, at x_0. If the dual or reciprocal basis to the e^μ be written (e_μ), then a contravariant tangent vector field $u(x)$ may be expanded, $u(x) = u^\mu(x)e_\mu(x)$, and more familiarly, $u = dx$, $u^\mu = dx^\mu$, $dx = dx^\mu e_\mu$. The application of a variable contravariant vector to a covariant vector yields a differential form $u^\mu v_\mu$, or $dx^\mu(x)(\partial f/\partial x^\mu)(x)$. If v_μ is actually a gradient, $v_\mu = \partial f/\partial x^\mu$ in a neighborhood, then one function "f" may be used in a whole neighborhood. A coordinate transformation introduces new bases (e'^μ) for the covariant tangent spaces, and therefore new dual bases (e'_μ) for the contravariant tangent spaces. The transformation at any one point is given by an arbitrary nonsingular m by m matrix, barring arbitrary restrictions on the choice of curvilinear coordinates, but the over-all transformation—a nonsingular m by m matrix function of the coordinates—must satisfy integrability conditions. If the group for frame transformations is to be other than the full linear group in m-dimensional space, and even in that case, if we wish freedom from the integrability conditions, then the (e^μ, e_μ) bases are not to be regarded as the frames.

In general, then, the frames $f^{(\mu)}$ will be other vector space bases, although $\mu = 1, \cdots, m$. Thus, $v = v_\mu e^\mu = v_{(\mu)}f^{(\mu)}$; $u = u^\mu e_\mu = u^{(\mu)}f_{(\mu)}$, where $(f_{(\mu)})$ is the frame of contravariant vectors dual to $(f^{(\mu)})$. In order to completely divorce the bookkeeping of curvilinear coordinate transformations from the language which describes the geometry, it is convenient to define each $f^{(\mu)}$ to be the same covariant vector when a transformation of coordinates induces a transformation $e^\mu \to e'^\mu$ of the coordinate-bound bases. Thus, if

$$f^{(\mu)}(x^1, \cdots, x^m) = f^{(\mu)}{}_\nu(x^1, \cdots, x^m)e^\nu(x^1, \cdots, x^m)$$

and if $(x^1, \cdots, x^m) \to (x'^1, \cdots, x'^m)$ are corresponding m-tuples in a transformation of coordinates, so that also

$$f^{(\mu)}(x'^1, \cdots, x'^m) = f'^{(\mu)}{}_\nu(x'^1, \cdots, x'^m)e'^\nu(x'^1, \cdots, x'^m),$$

with

$$e'^\nu = \frac{\partial x'^\nu}{\partial x^\varpi}\, e^\varpi,$$

then

$$f'^{(\mu)}{}_\nu \frac{\partial x'^\nu}{\partial x^\varpi} = f^{(\mu)}{}_\varpi, \quad \text{or} \quad f'^{(\mu)}{}_\nu = f^{(\mu)}{}_\varpi \frac{\partial x^\varpi}{\partial x'^\nu}.$$

If the tangent spaces are introduced by the computation of derivatives, the coordinate-bound bases will enter first, and this will be signalled by the use of unparenthesized indices. If a representation bundle built on the tangent spaces is also to be discussed, then the coefficients ${f^{(\mu)}}_\nu$ for mediating between the differentiations and the algebra will assume importance equal to that of the connection itself.

IX. CURVATURE

A. Definition

The structure defined by vectored patches is presented in a highly implicit form: The connection mapping the space at x onto the space at $x' \neq x$ (the path a smooth curve which does not cross or touch itself) is given by the identity matrix if basic frames are introduced along the curve by parallel displacement from the start; also all the infinitesimal components of connection may be annulled at any one point x, if basic frames about x are introduced by displacing a basic frame at x parallelly along a complete family of curves radiating from x.

If, however, a smooth curve loops back to its start, the parallel displacement of a frame around it may yield a new frame; the loop displacement of a general vector referred to a fixed frame defines an admissible linear transformation $v'_{(\alpha)} = {R_{(\alpha)}}^{(\beta)} v_{(\beta)}$ at the base point (start and finish). If a different frame is chosen as basis at the base point, $\bar{v}_{(\alpha)} = {L_{(\alpha)}}^{(\beta)} v_{(\beta)}$, then ${\bar{R}_{(\alpha)}}{}^{(\beta)} = {L^{(\gamma)}}_{(\alpha)} {R_{(\gamma)}}^{(\delta)} {L_{(\delta)}}^{(\beta)}$: the ${R_{(\alpha)}}^{(\beta)}$ transform under frame transformations as mixed-tensor components, and the matrix $[{R_{(\alpha)}}^{(\beta)}]$ is given only up to conjugation with an arbitrary **L**-matrix, unless a base frame is specified together with the base point; this is the case $x = x'$ of the theorem of Section V, D. A change of base point introduces no further change in the R-matrix than conjugation by an **L**-ma-

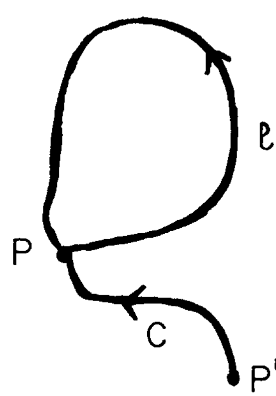

FIG. 1. Change of base point

trix. The rule assigning $R_{(\alpha)}{}^{(\beta)}$-transformations to loops (with base point and base frame) will be called curvature, and the matrix $[R_{(\alpha)}{}^{(\beta)}]$, the curvature around the loop.

The ambiguity of **L**-conjugations of the curvature matrices is not serious: Consider the class of all loops beginning and ending at one common base point bearing one base frame, (P, f), then change both, (P', f'). The curvatures sustain merely one over-all **L**-conjugation, if the loop based at P' corresponding to loop l at P is constructed by going from P' to P along a fixed curve C, then describing l, then describing C backwards: $C^{-1}lC$. (This is the usual argument for showing that the use of a base point in one-dimensional homotopy is not a difficulty.) There is no ambiguity at all, of course, when **L** is a commutative group, e.g., in the study of two-dimensional metric spaces, and in ordinary gauge invariance as met in electromagnetic theory.

B. Characterization of Connected Bundles by Curvature

If two equivalent **L**-bundles over the same manifold equipped with connections assign the same curvature matrices to all loops from some base point x_0 with base frame f_0, may the equivalence be established so that the connections agree?

A proof that this may be done will be sketched. Draw a family of curves radiating out from the base point x_0, so that to each point x there is one curve C drawn from x_0 to x. The two displacements along C from x_0 to x will in general introduce different frames, f_1 and f_2, at x; by the frames f_1 and f_2, is meant the columns of the displacement matrices. By applying an appropriate transformation $L(C, x)$ to the basic frame at x in the patchwork for the second bundle, the columns of the second displacement matrix may be made to agree with those of the first, however, so that now $f_1 \to f_2$ is given by the identity matrix. The fact that if another curve C' is drawn from x_0 to x, the displacement around the loop $C'^{-1}C$, C followed by the reverse of C', is given by the same curvature matrix for both connections then has the consequence that the change of basic frame $L(C, x)$ at x is in fact independent of the curve C. If one imagines the manifold cut so that, except for cuts, it is covered by one huge patch, the $L(C, x)$ construction is seen to do the job of establishing the desired equivalence with identity of connections, for the C' paths may even be allowed to cross cuts; the $C'^{-1}C$ argument covers this because the assumption of equal curvature matrices was made for all loops, even such as are not homotopic to points. One could also make $L(C, x)$ constructions in patches, and then use the curvature assumption to prove that the seaming functions are all right. Note that if the $L(C, x)$ transformation of the basis at x in the second structure is not made first, before C' is compared to C, then the proof is snarled by the circumstance that x-based curvatures may differ from the one structure to the equivalent one by an **L**-conjugation.

This proof of equivalence based on equal curvature does not apply to tangent bundles, unless the role of the vectors as tangent vectors is ignored; there is the question of comparison of the transformations $f^{(\alpha)}{}_{\mu}(x)$ expressing the frames relative to the coordinate-bound bases $e^{\mu}(x)$. On the other hand, there is some freedom to modify these by coordinate transformations.

C. Infinitesimal Curvature

In the case of an infinitesimal loop, the curvature matrix must differ from the unit matrix infinitesimally, and the infinitesimal difference will belong to the matrix Lie algebra of the matrix group **L**. Since any displacement, not necessarily infinitesimal, followed by its reverse, leads to a unit curvature matrix, it follows that the lowest order in which infinitesimal curvature may appear (and does) is the second order of coordinate-difference infinitesimals. This is just as in the case of Riemannian geometry (*17*, *18*). It may be verified by direct computation that displacement of a vector around a parallelogram of displacements δ, ϵ, $-\delta$, $-\epsilon$ is $\delta^{(\beta)}_{(\alpha)} + R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}t^{\mu\nu}$, to this order, where

$$t^{\mu\nu} = \tfrac{1}{2}(\delta^{\mu}\epsilon^{\nu} - \delta^{\nu}\epsilon^{\mu}),$$

and

$$R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu} = -\frac{\partial c_{(\alpha)}{}^{(\beta)}{}_{\nu}}{\partial x^{\mu}} + \frac{\partial c_{(\alpha)}{}^{(\beta)}{}_{\mu}}{\partial x^{\nu}} + c_{(\alpha)}{}^{(\gamma)}{}_{\mu}\, c_{(\gamma)}{}^{(\beta)}{}_{\nu} - c_{(\alpha)}{}^{(\gamma)}{}_{\nu}\, c_{(\gamma)}{}^{(\beta)}{}_{\mu}. \qquad (5)$$

Since the curvature matrix for the loop is a coordinate invariant, and is even frame invariant so long as one keeps the frame at the base point fixed, $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}$ necessarily transforms in the manner expected by the writing of its indices. Note that its definition does not involve the concept of absolute tangent-vector differential, i.e., of the tangent-space connection, Γ, unless of course the vectors $v_{(\alpha)}$ and connection $c_{(\alpha)}{}^{(\beta)}{}_{\mu}$ themselves refer to tangent vectors, as is the case in Riemannian geometry.

D. Algebraic Identities

We come now to identities satisfied by the $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}$. There is the obvious antisymmetry,

$$R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu} + R_{(\alpha)}{}^{(\beta)}{}_{\nu\mu} = 0,$$

and the condition that the matrix belong to the Lie algebra,

$$R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x) = R^{I}{}_{\mu\nu}(x)E_{I(\alpha)}{}^{(\beta)}, \qquad (6)$$

for appropriate μ, ν-antisymmetric coefficients $R^{I}{}_{\mu\nu}(x)$. For particular Lie algebras, this could be given instead by conditions of symmetry on the indices α, β. If we put $c_{(\alpha)}{}^{(\beta)}{}_{\mu} = c^{I}{}_{\mu}E_{I(\alpha)}{}^{(\beta)}$ into the definition of $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}$, then the Lie-algebra components of curvature and connection are seen to be related by

$$R^I{}_{\mu\nu} = \frac{\partial c^I{}_\nu}{\partial x^\mu} - \frac{\partial c^I{}_\nu}{\partial x^\nu} + c^I{}_{JK}\, c^J{}_\mu\, c^K{}_\nu\,, \tag{7}$$

where the $c^I{}_{JK}$ are the constants of structure appropriate to the $E_{I(\alpha)}{}^{(\beta)}$, namely,

$$E_{J(\alpha)}{}^{(\gamma)} E_{K(\gamma)}{}^{(\beta)} - E_{K(\alpha)}{}^{(\gamma)} E_{J(\gamma)}{}^{(\beta)} = c^I{}_{JK} E_{I(\alpha)}{}^{(\beta)}. \tag{8}$$

We have seen that one set $c^I{}_\mu(x)$ of Lie-albegra components of connection may be used to construct many concrete matrix realizations of a connected bundle of representations, and the $R^I{}_{\mu\nu}(x)$ form for the curvature would be useful in at once defining the infinitesimal curvature for such a family of bundles.

If the given infinitesimal curvature vanishes over a cap, then the constructed infinitesimal curvature (see F) also vanishes, whence displacement around the loop bounding the cap is given by the identity matrix. For spaces with vanishing infinitesimal curvature, the displacement operator depends only on the homotopy class of the curve. Now choose a base point in each simply connected patch and introduce basic frames throughout the patch by parallel displacement along a complete system of radiating curves. This is independent of choice of the curves, and displacement from the base point x_0 in a patch is given by the unit matrix. But then displacement in a patch from x to x' is given by the unit matrix for displacement from x_0 to x' through x, and the displacement in general reduces to the product of seaming matrices. Therefore, the theory of connected **L**-bundles with null infinitesimal curvature reduces to that of **L**-bundles without a connection.

E. Bianchi Identities

Bianchi identities hold, and may be expressed in various forms. The simplest form is obtained by introducing frames about x so that the $c^I{}_\mu(x) = 0$, whence the $c^I{}_{JK} c^J{}_\mu c^K{}_\nu$ term and its first derivatives all vanish at x, so that the verification of

$$\frac{\partial R^I{}_{\mu\nu}(x)}{\partial x^\varpi} + \frac{\partial R^I{}_{\nu\varpi}(x)}{\partial x^\mu} + \frac{\partial R^I{}_{\varpi\mu}(x)}{\partial x^\nu} = 0 \tag{9}$$

reduces to the cancellation of only six terms. By using explicit differentiation for covariant differentiation for tangent vectors, the invariant form $R^I{}_{\mu\nu,\varpi}$ + cycl.$(\mu, \nu, \varpi) = 0$ is obtained;

$$\frac{\partial R^I{}_{\mu\nu}}{\partial x^\varpi} + R^J{}_{\mu\nu}\, c^K{}_\varpi\, c^I{}_{JK} + \text{cycl. } (\mu, \nu, \varpi) = 0. \tag{10}$$

It is not only not necessary to have a connection for the tangent spaces to state Bianchi identities, but such a connection refuses to participate in them, even if it is given.

F. A Peculiar Integral. The Gauss-Bonnet Theorem

We have seen that the curvature mapping which assigns matrices to loops is a characterization of an **L**-bundle, but the class of sufficiently smooth loops is such an unwieldy object that this characterization is not likely to be useful, unless the curvature can be specified more simply. One would hope that the finite could be reduced to the infinitesimal, and that it should suffice to give the infinitesimal curvature $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x)$, supplemented of course by matrices $R_{(\alpha)}{}^{(\beta)}(\ell)$ for a set of loops ℓ sufficient to generate the homotopy classes. I have tried to define an integral whereby the curvature around a loop homotopic to a point may be expressed in terms of the infinitesimal curvature over a retraction cap, bounded by the loop, with indifferent success.

Figure 2 shows the way in which the loop is subdivided, so that infinitesimal loops covering the retraction cap come into consideration. Unfortunately, if frames are arbitrarily specified beforehand over the cap, and the infinitesimal curvature $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}$ is given relative to them, the process of parallel displacement along the curve of Fig. 2d or along any other scheme of subdivision, starting

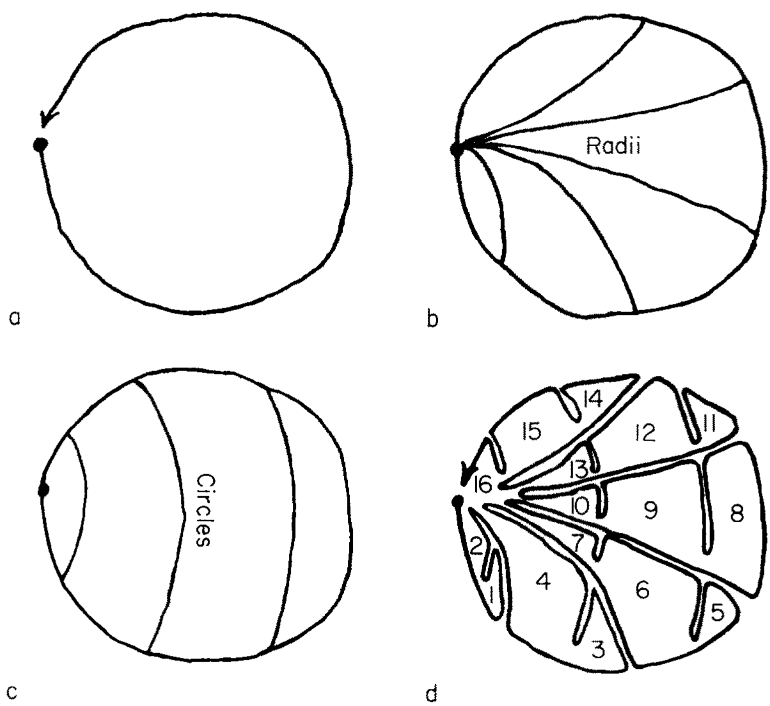

Fig. 2. Subdivision of a loop which bounds a cap

from the frame given as basis at the base point, will lead to a frame at an infinitesimal loop other than the given one, whence the displacement about the infinitesimal loop will not be obtained through the given $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x)$ function, but will be obtained through some **L**-matrix conjugate of it. It is futile to try to remedy this by altering the subdivision construction, for any integral which leads to the over-all curvature matrix $R_{(\alpha)}{}^{(\beta)}(\ell)$ of the entire loop ℓ must show its **L**-conjugation dependence on the choice of basic frame at the base point, so that some propagation of this frame over the entire retraction cap must be a necessary part of the definition of the integral.

Imagine that we are given the *connection* over the cap, minimally, that we are given the radial components, $c_{(\alpha)}{}^{(\beta)}{}_{\mu}\delta^{\mu}$ for radial δ^{μ}, of the connection over the cap. Displace the base frame parallel to the family of radii, to obtain the new frames over the cap. Explicitly, the connection from the base point to a point x along a radius will be

$$C_{(\alpha)}{}^{(\beta)}(x)$$
$$= \lim_{\delta' s\to 0} c_{(\alpha)}{}^{(\alpha_n)}(x)\delta^{\mu}\cdot c_{(\alpha_n)}{}^{(\alpha_{n-1})}{}_{\mu_n}(x_n)\delta_n{}^{\mu_n}\cdots c_{(\alpha_1)}{}^{(\alpha_0)}{}_{\mu_1}(x_1)\delta_1{}^{\mu_1}c_{(\alpha_0)}{}^{(\beta)}{}_{\mu_0}(x_0)\delta_0{}^{\mu_0},$$

so that the constructed frame $f(x)$ at x is related to the original, given frame $f'(x)$ at x by $f^{(\alpha)}(x) = C_{(\beta)}{}^{(\alpha)}(x)f'^{(\beta)}(x)$, or using our notation for the transposed inverse matrix, $f'^{(\alpha)}(x) = C^{(\alpha)}{}_{(\beta)}(x)f^{(\beta)}(x)$. Since the new frames $f^{(\beta)}(x)$ are to be used as bases, $(f^{(\beta)}(x))_{(\gamma)} = \delta^{(\beta)}_{(\gamma)}$, and $(f'^{(\alpha)}(x))_{(\gamma)} = C^{(\alpha)}{}_{(\gamma)}(x)$; the radially integrated connection defines the given frames in terms of the constructed frames. Let the given curvature be $R'_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x)$. Then on the new frames, the infinitesimal curvature R is given by $R_{(\delta)}{}^{(\gamma)}{}_{\mu\nu}(x) = C_{(\delta)}{}^{(\alpha)}(x)R'_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x)C^{(\gamma)}{}_{(\beta)}(x)$, and it is this constructed infinitesimal curvature which is to be used in the integral. Note that the given functions which have actually been needed so far are the $c_{(\alpha)}{}^{(\beta)}{}_{\mu}\delta^{\mu}$ along a family of radial tangent vectors δ^{μ}, and the given $R'_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x)$ over the cap; the notions of given and constructed frames are conceptual devices which entail no further given functions.

The integral itself is the limit defined by the product of the curvature matrices around the infinitesimal loops in the ordering explained by the numbers in Fig. 2d. Thus, $R_{(\alpha)}{}^{(\beta)}(\ell)$ is the fine-mesh limit of products of the form $R_{16(\alpha)}{}^{(\alpha_{15})}\cdot R_{15(\alpha_{15})}{}^{(\alpha_{14})}\cdots R_{1(\alpha_1)}{}^{(\beta)}$, where $R_{i(\alpha)}{}^{(\beta)} = \delta^{(\beta)}_{(\alpha)} + R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}(x_i)t_i^{\mu\nu}$, x_i is any mean coordinate for the ith element of area, and $t_i^{\mu\nu}$ is $\frac{1}{2}(\delta^{\mu}\epsilon^{\nu} - \delta^{\nu}\epsilon^{\mu})$, or simply $\delta^{\mu}\epsilon^{\nu}$, for rectangular elements of area spanned by two coordinate-difference infinitesimals.

It is obvious that in parallel displacement around the curve of Fig. 2d, all displacements along the incursions cancel, and that therefore $R_{(\alpha)}{}^{(\beta)}(\ell)$ is obtained. The construction is needed so that the fact that the boundary of an infinitesimal element is described not at once, but in two separate segments, should not matter, and so that one knows which of the many **L**-conjugate curvatures

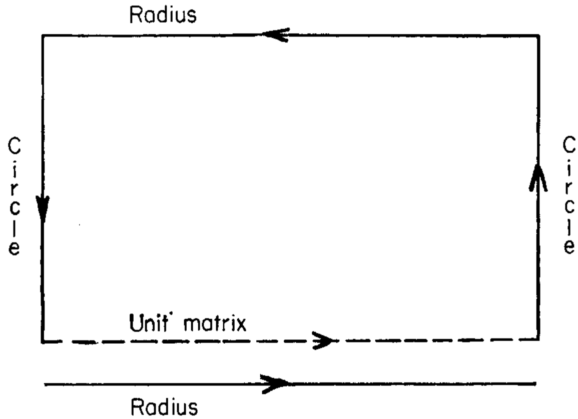

FIG. 3. Building a proper element of area

about the element should be used. In fact, the construction renders all matrices for radial displacements unit matrices, so that only the displacements along circles count, and these are included in one continuous description of each infinitesimal element if one adds extra ineffective radial displacements. The curvature matrix to be used when an infinitesimal loop is thus described in following the path of Fig. 2d is that reckoned relative to the frames introduced by parallel displacement along the radii; i.e., the R-matrices computed from the given R'-matrices and the integral connection matrices, as already explained.

The unsatisfactory aspect of this integral is the fact that a great deal must be known about the components of connection themselves for its computation, so that the given infinitesimal curvature may be suitably corrected prior to the actual ordered-product integration of the infinitesimal curvature; and, as has already been remarked, it is difficult to see how else the **L**-conjugation, which the entire result, ${R_{(\alpha)}}^{(\beta)}(\ell)$, must suffer when the frame at the base point is changed, may enter.

If we are interested in the curvature around all loops homotopic to a point with a common base point and base frame, which are contained in one surface, or all loops which bound caps which are ruled by a common family of curves radiating from a common base point, one construction of new frames by displacement from the base point along the radial curves will suffice for the very large class of loops described, but the class of all loops is so much larger that the advantage gained is small. If the underlying manifold is two-dimensional, of course, the class of such loops is exhaustive.

The difficulties disappear, of course, if the group **L** is commutative, for then **L**-conjugation is ineffective, $R' = R$, the order of factors does not matter, and the integral is the limit of

$$\prod_i (1 + R_{i\mu\nu}t_i^{\mu\nu})$$

or of

$$\prod_i \exp (R_{i\mu\nu}t_i^{\mu\nu}) = \exp \sum_i R_{i\mu\nu}t_i^{\mu\nu};$$

thus

$$[R_{(\alpha)}{}^{(\beta)}(\ell)] = \exp \iint [R_{(\gamma)}{}^{(\delta)}]_{\mu\nu}(x) \frac{\partial(x^\mu, x^\nu)}{\partial(u, v)} du\, dv.$$

This is the Gauss-Bonnet theorem. For example, if the group is the phase-factor group of electromagnetic theory, an **L**-matrix is a numeric factor $e^{i\varphi}$, an infinitesimal element an imaginary number, $R_{\mu\nu}(x)$ is of form $i\frac{1}{2}f_{\mu\nu}(x)$, where $f_{\mu\nu}(x)$ is real, and

$$R(\ell) = \exp i \frac{1}{2} \iint f_{\mu\nu}(x) \frac{\partial(x^\mu, x^\nu)}{\partial(u, v)} du\, dv, \quad \text{or} \quad \exp i \iint f_{uv}(x)\, du\, dv,$$

the integral itself giving the phase angle. The fact that the integral is the same over different retraction caps for the same loop is Gauss' theorem in electromagnetic theory applied to magnetic rather than to electric charges. The historical statement of the Gauss-Bonnet theorem is for two-dimensional metric spaces, where the frames are orthonormal pairs of tangent vectors, the group **L** is the commutative group of matrices $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$, the infinitesimal matrices are of the form $\theta \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, and the infinitesimal curvature is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ times $\theta_{\mu\nu}(x)$. We can simplify

$$\theta_{\mu\nu}(x) \frac{\partial(x^\mu, x^\nu)}{\partial(u, v)} du\, dv$$

to $2\theta_{uv}(u, v)\, du\, dv$ by going to u, v coordinates. The finite curvature is

$$\exp 2 \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \iint_{\text{cap}} \theta_{uv}(u, v)\, du\, dv,$$

so that the angle of rotation suffered by a vector on parallel displacement around a loop is given by $2 \iint_{\text{cap}} \theta_{uv}(x)\, du\, dv$; $2\theta_{uv}$ is the u, v-coordinate based Gaussian curvature density. (The notation for electromagnetic theory may be brought into this real form by emphasizing real and imaginary parts.)

X. COVARIANT DERIVATIVE

If a vector field is defined and differentiable at x, then its absolute differential is defined as the difference between the first-order increment of its components and the first-order increment of its components obtained by parallel displacement from x,

$$d_{\mathrm{abs}}v_{(\alpha)} = dv_{(\alpha)} - d_{\|}v_{(\alpha)} ,$$

or

$$d_{\mathrm{abs}}\, v_{(\alpha)} = \left(\frac{\partial v_{(\alpha)}}{\partial x^{\mu}} - c_{(\alpha)}{}^{(\beta)}{}_{\mu}\, v_{(\beta)}\right) dx^{\mu},$$

and the coefficients are the components of the "covariant derivative";

$$v_{(\alpha),\mu} = \frac{\partial v_{(\alpha)}}{\partial x^{\mu}} - c_{(\alpha)}{}^{(\beta)}{}_{\mu}\, v_{(\beta)} . \tag{11}$$

Since

$$d_{\mathrm{abs}}v_{(\alpha)} = v_{(\alpha),\mu}\, dx^{\mu} \doteq (v(x + dx) - v_{\|}(x, dx))_{(\alpha)}$$

is the difference of two vectors at $x + dx$, to first order, it transforms as a vector at $x + dx$, and since it has only first-order magnitude, it transforms as a vector at x, to linear order in the dx^{μ}. $v_{(\alpha),\mu}$ is therefore a second-rank tensor, an element of the vector space constructed by tensor product of the given vector space and the tangent space, and with the index μ referring to a coordinate-dependent (e^{μ}) basis in the tangent space, if the dx^{μ} designate literal increments of the coordinates.

In order to displace $v_{(\alpha),\mu}$ or, equivalently, to extract a second covariant derivative, it is necessary that a connection be given for the bundle of tangent spaces. Such a connection will be assumed when needed below, and will be distinguished by the notations $\Gamma_{[\nu]}{}^{[\varpi]}{}_{\mu}$, $\Gamma_{\nu}{}^{\varpi}{}_{\mu}$, etc., from the more general notion of infinitesimal components $c_{(\alpha)}{}^{(\beta)}{}_{\mu}$ of an arbitrary connection. (Thus, an index free of parentheses is to be used with a coordinate-bound tangent-space basis; an index in round parentheses is to be used with a frame, and may or may not refer to a tangent space; an index in square parentheses is to be used with a frame basis for a tangent space.)

XI. TORSION AND CONNECTED TANGENT SPACES

Infinitesimal curvature is often introduced in Riemannian geometry by permuting two successive covariant derivations, rather than by the parallel displacement of the vector space at a point around infinitesimal loops. By applying (11) to the $v_{(\alpha),\mu}$ tensor, we find that

$$v_{(\alpha),\mu,\nu} = \frac{\partial v_{(\alpha),\mu}}{\partial x^{\nu}} - c_{(\alpha)\mu}{}^{(\beta)\varpi}{}_{\nu}\, v_{(\beta),\varpi} ,$$

where the connection $c_{(\alpha)\mu}{}^{(\beta)\varpi}{}_{\nu}$ for tensors is given by the one-index-at-a-time rule, as discussed in the section on algebra,

$$c_{(\alpha)\mu}{}^{(\beta)\varpi}{}_{\nu} = c_{(\alpha)}{}^{(\beta)}{}_{\nu}\delta_{\mu}^{\varpi} + \delta_{(\alpha)}^{(\beta)}\Gamma_{\mu}{}^{\varpi}{}_{\nu} .$$

By substituting (11) itself and collecting the coefficients of $v_{(\beta)}$ and of $v_{(\beta),\varpi}$, one obtains

$$v_{(\alpha),\mu,\nu} - v_{(\alpha),\nu,\mu} = -R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}v_{(\beta)} - T_{\mu}{}^{\varpi}{}_{\nu}v_{(\alpha),\varpi} , \tag{12}$$

where $R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}$ is as before, and in particular depends only on c, not on Γ—unless c depends on Γ—and where

$$T_{\mu}{}^{\varpi}{}_{\nu} = -\Gamma_{\mu}{}^{\varpi}{}_{\nu} + \Gamma_{\nu}{}^{\varpi}{}_{\mu} , \tag{13}$$

is the "torsion." There are many ways to see that all the quantities written in (12) are the components of tensors of the type expected in concordance with the notation of positioned indices.

As is well known, and as is obvious through the circumstance that the $T_{\mu}{}^{\varpi}{}_{\nu}$ are coordinate-based components of a tangent-space tensor, the $\Gamma_{\mu}{}^{\varpi}{}_{\nu}$ cannot in general be set zero at even a point, by a choice of coordinates, although we argued generally that, given a group of matrices $\mathbf{L}$ acting between distinguished bases or frames $f^{[\mu]}$ in the tangent spaces, the $\Gamma_{[\mu]}{}^{[\varpi]}{}_{\nu}$ or the $\Gamma_{[\mu]}{}^{[\varpi]}{}_{[\nu]}$ could be put to zero at any one point, the construction establishing this situation being choice of frames in a neighborhood of the point by parallel displacement of the basic frame at the point along a complete system of radiating curves.

It may be instructive to consider how these circumstances may coexist. The key point is that transformation of coordinate-bound bases $e'^{\mu} = L^{\mu}{}_{\nu}e^{\nu}$ induced by a coordinate transformation is $L^{\mu}{}_{\nu} = \partial x'^{\mu}/\partial x^{\nu}$, and is therefore restricted by integrability conditions; $\partial L^{\mu}{}_{\nu}/\partial x^{\varpi}$ is $\partial^2 x'^{\mu}/\partial x^{\varpi}\partial x^{\nu}$, and is therefore necessarily (ν, ϖ)-symmetric. The general law for transformation of components of connection given for $\mathbf{L}$-transformations by (3), and its derivation, are in fact valid for any change of basis. If the basis at the base point x is unchanged, then $c'_{(\alpha)}{}^{(\beta)}{}_{\mu} = c_{(\alpha)}{}^{(\beta)}{}_{\mu} + \partial L_{(\alpha)}{}^{(\beta)}/\partial x^{\mu} = c_{(\alpha)}{}^{(\beta)}{}_{\mu} - \partial L^{(\beta)}{}_{(\alpha)}/\partial x^{\mu}$, and $\Gamma'_{\nu}{}^{\varpi}{}_{\mu} = \Gamma_{\nu}{}^{\varpi}{}_{\mu} - \partial^2 x'^{\varpi}/\partial x^{\mu}\partial x^{\nu}$. Since we may put $x'^{\varpi} = x^{\varpi} + \frac{1}{2}s_{\mu}{}^{\varpi}{}_{\nu}x^{\mu}x^{\nu}$ for arbitrary (μ, ν)-symmetric $s_{\mu}{}^{\varpi}{}_{\nu}$, we have the familiar fact that the symmetric part of $\Gamma_{\nu}{}^{\varpi}{}_{\mu}$ may be annulled at a point, and then $\Gamma_{\nu}{}^{\varpi}{}_{\mu} = -\frac{1}{2}T_{\nu}{}^{\varpi}{}_{\mu}$. We therefore have a second geometric role for the torsion: its vanishing is the necessary and sufficient condition which allows the transformation of bases needed to annul the coefficients of connection at a point to be effected via coordinate-bound bases. More graphically, if the torsion doesn't vanish at a point, then parallel displacement of a frame along a complete system of radiating curves will result in a family of frames which cannot mesh smoothly into a coordinate system in the neighborhood of the point—this applying for frames in the tangent spaces of contravariant vectors, otherwise it makes no sense.

Since the transformation $f^{[\mu]}(x) = f^{[\mu]}{}_{\nu}(x)e^{\nu}(x)$ or $f_{[\mu]}(x) = f_{[\mu]}{}^{\nu}(x)e_{\nu}(x)$ on the one hand, with $f_{[\mu]}{}^{\nu}f^{[\mu]}{}_{\varpi} = \delta^{\nu}_{\varpi}$, and $f_{[\mu]}{}^{\nu}f^{[\varpi]}{}_{\nu} = \delta^{[\varpi]}_{[\mu]}$, and the torsion $T_{\nu}{}^{\mu}{}_{\varpi}$, on the other, are both features peculiar to the tangent spaces, and since $f_{[\mu]}{}^{\nu}(x)$ clearly contains the entire relationship of the **L**-group geometry to the coordinates, the torsion must depend on the $f_{[\mu]}{}^{\nu}(x)$. Formulas for this are as follows.

Since $f_{[\mu]}{}^{\nu}(x)$ is a transformation of bases, (3) applies:

$$\Gamma_{[\alpha]}{}^{[\beta]}{}_{\mu} = f_{[\alpha]}{}^{\nu} f^{[\beta]}{}_{\varpi} \Gamma_{\nu}{}^{\varpi}{}_{\mu} + f^{[\beta]}{}_{\nu} \partial f_{[\alpha]}{}^{\nu}/\partial x^{\mu}.$$

By a suitable coordinate transformation, it is possible to render $f^{[\mu]}{}_{\nu} = \delta^{\mu}_{\nu}$, and thereby shorten the writing:

$$\Gamma_{[\alpha]}{}^{[\beta]}{}_{\mu} = \Gamma_{\alpha}{}^{\beta}{}_{\mu} + \partial f_{[\alpha]}{}^{\beta}/\partial x^{\mu} = \Gamma_{\alpha}{}^{\beta}{}_{\mu} - \partial f^{[\beta]}{}_{\alpha}/\partial x^{\mu} = \Gamma_{\alpha}{}^{\beta}{}_{\mu} + \frac{1}{2}\left(\frac{\partial f_{[\alpha]}{}^{\beta}}{\partial x^{\mu}} - \frac{\partial f^{[\beta]}{}_{\alpha}}{\partial x^{\mu}}\right).$$

If frames are constructed to annul the $\Gamma_{[\alpha]}{}^{[\beta]}{}_{\mu}$ and coordinates are chosen to annul the (α, μ)-symmetric part of $\Gamma_{\alpha}{}^{\beta}{}_{\mu}$ so that $\Gamma_{\alpha}{}^{\beta}{}_{\mu} = -\tfrac{1}{2}T_{\alpha}{}^{\beta}{}_{\mu}$, then

$$\begin{aligned} T_{\alpha}{}^{\beta}{}_{\mu} &= 2\partial f_{[\alpha]}{}^{\beta}/\partial x^{\mu} = -2\partial f^{[\beta]}{}_{\alpha}/\partial x^{\mu} = \partial(f_{[\alpha]}{}^{\beta} - f^{[\beta]}{}_{\alpha})/\partial x^{\mu} \\ &= \tfrac{1}{2}(T_{\alpha}{}^{\beta}{}_{\mu} - T_{\mu}{}^{\beta}{}_{\alpha}) = \frac{\partial f_{[\alpha]}{}^{\beta}}{\partial x^{\mu}} - \frac{\partial f_{[\mu]}{}^{\beta}}{\partial x^{\alpha}} = \frac{\partial f^{[\beta]}{}_{\mu}}{\partial x^{\alpha}} - \frac{\partial f^{[\beta]}{}_{\alpha}}{\partial x^{\mu}}, \end{aligned}$$

the last form being a curl. If all the normalization conditions are dropped, one has also

$$\begin{aligned} T_{[\mu]}{}^{[\nu]}{}_{[\varpi]} &= \Gamma_{[\varpi]}{}^{[\nu]}{}_{[\mu]} - \Gamma_{[\mu]}{}^{[\nu]}{}_{[\varpi]} + f_{[\varpi]}{}^{\rho} f_{[\mu]}{}^{\sigma}\left(\frac{\partial f^{[\nu]}{}_{\rho}}{\partial x^{\sigma}} - \frac{\partial f^{[\nu]}{}_{\sigma}}{\partial x^{\rho}}\right), \\ T_{\mu}{}^{\nu}{}_{\varpi} &= f^{[\rho]}{}_{\varpi} f_{[\sigma]}{}^{\nu} \Gamma_{[\rho]}{}^{[\sigma]}{}_{\mu} - f^{[\rho]}{}_{\mu} f_{[\sigma]}{}^{\nu} \Gamma_{[\rho]}{}^{[\sigma]}{}_{\varpi} + f_{[\rho]}{}^{\nu}\left(\frac{\partial f^{[\rho]}{}_{\varpi}}{\partial x^{\mu}} - \frac{\partial f^{[\rho]}{}_{\mu}}{\partial x^{\varpi}}\right). \end{aligned}$$

The role of the $f^{[\mu]}{}_{\nu}$ in the case of the tangent bundles suggests that they be tied into the axiomatic structure. For example, the condition that $\Gamma_{[\mu]}{}^{[\nu]}{}_{\varpi}$ be a universal function of the $f^{[\mu]}{}_{\nu}$ and the $\partial f^{[\mu]}{}_{\nu}/\partial x^{\varpi}$ homogeneous of first degree in the latter, is very strong: when **L** is the group of orthogonal transformations relative to some nonsingular numeric symmetric matrix (metric) $\eta_{[\mu\nu]}$, this condition establishes the $\Gamma_{\mu}{}^{\nu}{}_{\varpi}$ as the usual Christoffel symbols corresponding to the gradient-basis components $g_{\mu\nu} = f^{[\varpi]}{}_{\mu} f^{[\rho]}{}_{\nu} \eta_{[\varpi\rho]}$ of the metric (*10*).

XII. CONSERVED CURRENT DENSITIES

Since

$$\frac{\partial}{\partial x^{\mu}}\left(\frac{\partial R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}}{\partial x^{\nu}}\right)$$

vanishes, we may try to interpret the single literal divergence of the infinitesimal curvature as a conserved current density, if the manifold is space-time. However, these quantities are not invariant.

The following theorem suggests how to take care of the coordinate dependence. Namely, if the $t^{\mu\nu\varpi\cdots}$ are coordinate-bound components of a completely contravariant tensor density completely antisymmetric in the superscripts, then $\partial t^{\mu\nu\varpi\cdots}/\partial x^\mu = (\text{div } t)^{\nu\varpi\cdots}$ are also the coordinate-bound components of a tensor density completely antisymmetric in the superscripts. Since $\partial^2 t^{\mu\nu\varpi\cdots}/\partial x^\nu \partial x^\mu = 0$, we have div div $t = 0$.

A scalar density s transforms so that $s\, dx^1 \cdots dx^m = s'\, dx'^1 \cdots dx'^m$; i.e., $s' = |\, det\, [\partial x/\partial x'] \,|\, s$, and the tensors formed by the tensor product of the scalar densities with the various mixed tensors are called tensor densities. The covariant derivative of tensor densities, when a connection is given, is usually defined so that $t^{\mu\nu\varpi\cdots}{}_{,\mu}$ and $\partial t^{\mu\nu\varpi\cdots}/\partial x^\mu$, under the above conditions, differ only by a tensor linear in the torsion; and the two agree if the connection is taken torsionless. More fundamentally, the whole notion of a parallel displacement for tangent vectors is irrelevant to this section, because the literal divergences may be directly verified to be tensors; only the $c_{(\alpha)}{}^{(\beta)}{}_\mu$ are involved in the covariant derivatives in this section.

We go to a metric so as to be able to raise indices, $R_{(\alpha)}{}^{(\beta)\mu\nu} = g^{\mu\varpi} g^{\nu\rho} R_{(\alpha)}{}^{(\beta)}{}_{\varpi\rho}$. Then $|\det g_{..}|^{1/2}$, formed from the determinant of the components of the covariant metric tensor, is a convenient scalar density, and

$$r_{(\alpha)}{}^{(\beta)\mu\nu} = |\det g_{..}|^{1/2} R_{(\alpha)}{}^{(\beta)\mu\nu}$$

are the components of a tensor density. If

$$i_{(\alpha)}{}^{(\beta)\mu} = \partial r_{(\alpha)}{}^{(\beta)\mu\nu}/\partial x^\nu, \tag{14}$$

then $i_{(\alpha)}{}^{(\beta)\mu}$ is also a tensor density, satisfying the literal equation of continuity,

$$\partial i_{(\alpha)}{}^{(\beta)\mu}/\partial x^\mu = 0, \tag{15}$$

if the basic frames are kept fixed, and only the coordinate-bound tangent-space indices are studied. Since $i_{(\alpha)}{}^{(\beta)\mu}$ is not frame-invariant, I call it a current pseudodensity, in analogy to the pseudotensor of energy and momentum of Einstein, and in conformity with the language of Yang and Mills.

$$i_{(\alpha)}{}^{(\beta)\mu} = i^{I\mu} E_{I(\alpha)}{}^{(\beta)}$$

separates the algebraically independent components, and

$$\partial i^{I\mu}/\partial x^\mu = 0.$$

If we define the current density by

$$j_{(\alpha)}{}^{(\beta)\mu} = j^{I\mu} E_{I(\alpha)}{}^{(\beta)} = r_{(\alpha)}{}^{(\beta)\mu\nu}{}_{,\nu}\,, \tag{16}$$

then $\partial j_{(\alpha)}{}^{(\beta)\mu}/\partial x^\mu \neq 0$ in general; even $j_{(\alpha)}{}^{(\beta)\mu}{}_{,\mu} \neq 0$, in general, although if the latter equation held, it would not be a satisfactory conservation law.

By writing out the covariant derivatives, one sees how much the current differs from the pseudocurrent:

$$\dot{j}_{(\alpha)}{}^{(\beta)\mu} = \dot{i}_{(\alpha)}{}^{(\beta)\mu} - c_{(\alpha)}{}^{(\gamma)}{}_{\nu} r_{(\gamma)}{}^{(\beta)\mu\nu} + c_{(\gamma)}{}^{(\beta)}{}_{\nu} r_{(\alpha)}{}^{(\gamma)\mu\nu},$$

or

$$\dot{j}^{I\mu} = \dot{i}^{I\mu} + c^{I}{}_{JK} r^{J\mu\nu} c^{K}{}_{\nu},$$

where the $c^{I}{}_{JK}$ are given by (8).

We therefore have

$$\partial \dot{j}^{I\mu}/\partial x^{\mu} = \partial(c^{I}{}_{JK} r^{J\mu\nu} c^{K}{}_{\nu})/\partial x^{\mu}. \tag{17}$$

The integral over a 4-volume of $\partial \dot{j}^{I\mu}/\partial x^{\mu}$ may therefore be transformed, not only to a 3-boundary integral of $\dot{j}^{I\mu}$, but also to a 3-boundary integral of

$$c^{I}{}_{JK} r^{J\mu\nu} c^{K}{}_{\nu} \equiv k^{I\mu}.$$

If there are two times when the 3-spatial volume integrals of $k^{I\mu}$ vanish or coincide and a spatial 2-surface "at infinity" over which the time-integrated flux of $k^{I\mu}$ vanishes, then although $\partial \dot{j}^{I\mu}/\partial x^{\mu}$ doesn't necessarily vanish within the bounded 4-volume, its integral does, and we obtain direct conservation of the $\dot{j}^{I\mu}$ current density, for these circumstances. If situations where $k^{I\mu}$ vanishes sufficiently at spatial infinity are viewed as more general than situations where also matter is sufficiently smeared out or dissected so that the 3-volume $k^{I\mu}$ integrals at fixed times vanish or coincide, the following argument may be employed. Vanishing of the time integral of the 2-surface integral of $k^{I\mu}$ would render the similar integral of $\dot{i}^{I\mu}$ equal to that of $\dot{j}^{I\mu}$. When we have this vanishing at spatial infinity, the invariant $\dot{j}^{I\mu}$ surface integrals therefore measure geometrically meaningful fluxes which are conserved together with the 3-volume integrated charge pseudodensity given by $\dot{i}^{I\mu}$. This gives the 3-volume integrated pseudocurrent density the meaning of a conserved geometrically meaningful quantity, although the density itself is highly frame-dependent. The vanishing of $c^{I}{}_{JK} r^{J\mu\nu} c^{K}{}_{\nu}$ at spatial infinity may occur independently of choice of frames if the infinitesimal curvature vanishes there, or through choice of sufficiently parallel frames over the surface to make the $c^{K}{}_{\nu}$ zero or small.

It may be objected that the physical fluxes through a bounding 2-surface in 3-space are borne by the very same mysterious particles which lend doubt to the possibility of the vanishing of $c^{I}{}_{JK} r^{J\mu\nu} c^{K}{}_{\nu}$ throughout a 3-volume, even at special times, and therefore that the conservation-law interpretation is probably empty. There are two answers to this objection. First, the objection does not apply to time intervals in which no particle crosses the surface, and this may be a long time for a surface sufficiently distant from matter. Second, for a scattering experiment in which we introduce fluxes of separated particles at spatial infinity, each particle may be surrounded by a large sphere, and the spatial 2-surface may be replaced by a spacelike 2-surface composed of the disconnected surfaces of these moving spheres, together with another surface surrounding the interaction region, the particulate spheres being shed as the interaction region is entered, and donned as the interaction region is left. This argument should however be broken by any good theory of broken symmetries (*19*), because the necessity for the symmetries to appear broken comes not only from the fact of rest masses and their inequality, but also from violation of current conservation laws. The argument is also broken if the interior zones of particles are connected to points which may lie outside the apparent macroscopic bounding surfaces; for further remarks on the possibility of conservation laws under these circumstances see Section XVII.

In the case of commutative groups, of course, the $c^I{}_{JK}$ vanish, $i^{I\mu} = j^{I\mu}$, and there is no complication.

$j^{I\mu}{}_{,\mu}$ can be calculated as follows:

$$\frac{\partial_j{}^{I\mu}}{\partial x^\mu} = c^I{}_{JK}\left(\frac{\partial r^{J\mu\nu}}{\partial x^\mu}c^K{}_\nu + r^{J\mu\nu}\frac{\partial c^K{}_\nu}{\partial x^\mu}\right)$$

$$= c^I{}_{JK}(-j^{J\nu}c^K{}_\nu + c^J{}_{LM}r^{L\nu\varpi}c^M{}_\varpi c^K{}_\nu + \tfrac{1}{2}r^{J\mu\nu}c^K{}_{LM}c^L{}_\mu c^M{}_\nu + \tfrac{1}{2}r^{J\mu\nu}R^K{}_{\mu\nu}).$$

The two terms in the middle may be seen to cancel if the (ϖ, ν)-antisymmetry of the first is employed to M, K antisymmetrize, and if the Jacobi identity

$$c^I{}_{JK}c^J{}_{LM} + c^I{}_{JL}c^J{}_{MK} + c^I{}_{JM}c^J{}_{KL} = 0$$

is then invoked. Then $j^{I\mu}{}_{,\mu} = \partial j^{I\mu}/\partial x^\mu + c^I{}_{JK}j^{J\mu}c^K{}_\mu$ immediately gives

$$j^{I\mu}{}_{,\mu} = \tfrac{1}{2}c^I{}_{JK}r^{J\mu\nu}R^K{}_{\mu\nu}\,. \tag{18}$$

The possibility of other currents, conserved in the above sense, arises whenever an antisymmetric tangent-space tensor is available. One such is the dual of the curvature tensor,

$$^d r_{(\alpha)}{}^{(\beta)\varpi\rho\cdots} = \tfrac{1}{2}\epsilon^{\mu\nu\varpi\rho\cdots}R_{(\alpha)}{}^{(\beta)}{}_{\mu\nu}\,, \tag{19}$$

where the dots signify extra indices, in the event that the manifold has more than 4 dimensions. No metric is needed. $^d i_{(\alpha)}{}^{(\beta)\varpi\cdots} = \partial\, ^d r_{(\alpha)}{}^{(\beta)\varpi\rho\cdots}/\partial x^\rho$ is "literally conserved" in the sense that $\partial\, ^d i_{(\alpha)}{}^{(\beta)\varpi\cdots}/\partial x^\varpi = 0$, the interpretation in terms of conservation applying in a normal way only to the case of four-dimensional space-time. $^d i$ is related to $^d j_{(\alpha)}{}^{(\beta)\varpi\cdots} = {}^d r_{(\alpha)}{}^{(\beta)\varpi\rho\cdots}{}_{,\rho}$ in the same way that i is related to j, but $^d j = 0$ is a form of the Bianchi identities, Section IX, E. Therefore, when $^d i = {}^d j$ on a 3-boundary, the conserved quantity is identically zero; where $^d i = {}^d j$ at spatial infinity, the net change of the conserved 3-space integral of $^d i$ is zero. In the same sense, then, that one obtains conserved charges, the dual charges vanish, and their conservation is empty. "Dual charge" will however be employed nontrivially in a modified sense, in the sequel.

In the case of commutative groups, the vanishing of the dual current density is unambiguous; for the typical example of electromagnetic theory, it is the vanishing of the magnetic pole current density.

The triviality of the dual currents seems to be coextensive with the conservation of the currents, in the shallow sense that, in a situation where it is hard to argue for one, it is likewise hard to argue for the other. This may indicate that when the conservation of some currents is broken, but survives in an approximate form, the corresponding dual currents, even in the present sense, may vanish only to the same approximation.

It the torsion does not vanish, and if indices may be raised and lowered and plain tensors converted to tensor densities, there may be more currents to consider. Thus, there is a vector $T_\mu{}^\mu{}_\nu = T_\nu$, $T^\mu = g^{\mu\nu}T_\nu$, $t^\mu{}_\varpi{}^\nu = g^{\mu\sigma}g^{\nu\rho}g_{\varpi\tau}T_\sigma{}^\tau{}_\rho \, |\det g_{..}|^{1/2}$, and $t^{\mu\nu} = t^\mu{}_\varpi{}^\nu T^\varpi$. The frame-dependent $t^\mu{}_{[\varpi]}{}^\nu$ may conceivably be impressed in this way. One could consider an ad hoc torsion, superimposed upon Riemannian geometry, or torsion arising from a nonsymmetric $g_{\mu\nu}$ (*23*, *17*).

XIII. THE THREE INGREDIENTS IN THE STRUCTURE OF A LIE GROUP. ARE THERE PHYSICAL PHENOMENA CORRELATED WITH THE NONINFINITESIMAL STRUCTURE?

A connected simply connected Lie group is determined by its infinitesimal algebra. More generally, the component of identity of an arbitrary Lie group may be obtained from that connected simply connected Lie group which is generated by its infinitesimal algebra, by mutually identifying the elements of a discrete closed normal subgroup. The connected simply connected group and this identification or factorization are the universal covering group of the component of the identity, and the covering map. Each element of a discrete normal subgroup commutes with the whole identity component, because it commutes with infinitesimal elements; the discrete normal subgroup which is the kernel of this construction is therefore commutative. It is also isomorphic to the "fundamental group" of homotopy equivalence classes of loops in the final, factor group; the fundamental group is therefore commutative. The isomorphism relating the discrete normal subgroup and the fundamental group after factorization is easily visualized if we imagine the simply connected group with the points of the discrete subgroup marked, and widen the class of loops through the identity to encompass curves drawn from one element of the discrete subgroup to another.

The entire Lie group consists in general of a number of components, each one a manifold isomorphic to the component of the identity, which is a normal subgroup. If we mutually identify the elements of the component of the identity, then we obtain another discrete group, which it seems reasonable to call the component group. Since the direct product of the component of the identity with the component group is a Lie group with the same component of identity and component group as the given group, and since it is easy to give examples of Lie groups which are not isomorphic to the direct product so generated, the relation of the other components to that of the identity is not adequately represented by the component group, but as I know no general characterization of the interrelationship between the components, I will emphasize the component group in the sequel.

By forming the direct product of a connected Lie group with a group, each point of which constitutes an open set, any such group may be impressed as component group in conjunction with an arbitrary component of identity, but this construction is too trivial to be of interest. In fact, each matrix in **L** is equivalent to a pair of matrices, one belonging to the component of identity $\mathbf{L}_1$, and the second belonging to a discrete group, $\mathbf{L}_2$. The connected **L**-bundle is therefore equivalent to a connected $\mathbf{L}_1$-bundle and a connected $\mathbf{L}_2$-bundle, over a common manifold. Furthermore, since $\mathbf{L}_2$ is discrete, there are no infinitesimal elements, and therefore no infinitesimal components of connection, so so that the $\mathbf{L}_2$-connection along curves degenerates into the formation of products of seaming matrices, and the connected $\mathbf{L}_2$-bundle is really only a plain $\mathbf{L}_2$-bundle. Therefore it is only if it is impossible to find an element in each component which commutes with the whole identity component—I will then say that **L** is "locked" —that the component structure becomes of real interest.

In summary, although a Lie group is given adequately in the immediate neighborhood of the identity by its infinitesimal algebra, its global structure is not thereby determined: a discrete normal subgroup of the determined, connected, simply connected Lie group with the given infinitesimal structure is needed to specify the structure of the component of the identity, another discrete group is needed to specify the structure of components, and yet more information is needed to measure locking.

The infinitesimal curvature and the current densities derived from infinitesimal quantities are all elements of Lie algebras. If all quantities of physical interest are already included in the curvature and current densities, then physical interest in the global structure of the symmetry group would be divorced from questions concerning the parallel displacement. Thus an interesting question is raised: Is the entire Lie group only artificially involved in the discussion of the physical properties related to admissible connections, and should not the proper object cited in a physical discussion be only the group's infinitesimal algebra, or are there really physical features related to admissible connections involving the two discrete groups defined by a Lie group?

In the absence of a Lagrangian and in the present context, I do not know how to obtain laws which sound as if they might be physical, unless I aim at conservation laws. I therefore seek examples of conservation laws stated in terms of the two discrete groups only. It is not likely that the examples I give are exhaustive.

XIV. GLOBAL CONSERVATION LAWS

On the physical side of the picture, I will consider an approach characteristic of phenomenological particle physics: a large region in ordinary space-time with the usual simple topological properties assigned to empty space-time in physics, surrounding a small, mysterious region containing matter, and therefore of dubious topological character. The mysterious zone, perhaps together with some ordinary, vacuous space, is separated from the large, ordinary zone by the surface of a sphere, and this for a period of time. It is therefore suggested that, at any one time, or on any cut from the outside, ordinary region with a spacelike hypersurface, the two-dimensional spherical boundary or "bag" which sequesters the mysterious region be the object of study.

Any discrete mathematical object associated with a bag will be constant as the bag is displaced continuously with time, in such a way as to keep the mysterious matter inside, and therefore the discrete mathematical object will define a conserved property. By continuously deforming the bag in space in such a way as to keep the same mysterious matter inside it, the discrete object is not altered, so that the conserved property is really a characteristic of the matter, not of the bag.

When two zones of matter intermingle, bags which sequester each separately become impossible. However, if before two zones of matter intermingle, both are surrounded by one large bag, the quantity appropriate to the two zones of matter together may be obtained, and seen to be conserved through the intermingling. In order to obtain the law for "adding" the conserved quantity, the way to obtain the quantity appropriate to the large bag from the quantities borne by the two small bags prior to the interminingling must be ascertained.

This general scheme will be applied to obtain conservation laws related to the nonsimple-connectedness of the component of identity as represented by the fundamental group, but I have found no way to apply it in the discussion of conservation laws related to the structure of components, which will therefore be run on a more ad hoc basis.

On the mathematical side of the picture, one must consider how the lack of simple-connectedness and the lack of connectedness of the group of matrices **L** can affect the structure of connected **L**-bundles. The connection directly associated group elements to paths, but, as we have seen, in a highly frame-dependent way: I recall again that the connection may be rendered by the unit matrix for any nonclosed path which does not cross itself, and note further that the matrix effecting the connection from a fixed starting point to a continuously moving endpoint which proceeds along a path depends discontinuously on that moving endpoint, owing to the jumps introduced by seaming matrices when the moving point goes from one patch into another, and this even if the path is a loop. The matrix effecting the connection around a whole closed loop, however, depends only on the base point and the choice of basic frame at that point, and obviously varies continuously as the loop is deformed continuously, and is therefore a fit object for our attention. We have already seen that these "curvature" matrices for a common base point and frame, arbitrarily chosen, determine the connected bundle up to isomorphism, so that an exhaustive study of these matrices would characterize the connected **L**-bundle completely. Although no claim is made here to such exhaustive study, this remark is adduced to further motivate the subsequent confinement of attention to the curvature matrices assigned to loops.

For each loop then, there is a group element, i.e., a matrix in **L**. If one loop and base frame is deformed continuously to another, its group element proceeds continuously to that of the other; the elements in the component group corresponding to homotopic loops coincide. Single loops, as representative of homotopy classes in the manifold, are therefore already related to the component structure of the group, because even though they correspond to single elements of the group, a single element determines the component in which it lies.

In order to explore the nonsimple-connectedness of the component of the identity of the group, however, we need loops in the component of identity of the group, and since a group element comes from a loop in the manifold, we need a loop of loops in the manifold. This immediately suggests a sectioned torus, but the physical picture motivates us to obtain a loop of loops rather from the surface of a sphere.

XV. COMPONENT OF THE IDENTITY: DUAL CHARGE

To obtain a loop of loops from the surface of a sphere, or a "bag," where all the loops share a common base point, the following method is suggested: Use the circles marked on the family of planes passing through a tangent through the base point. At first sight, it seems that two orientations must also be given: a sense of description for any one of these circles, and an ordering of the loops thus obtained by, e.g., specifying for any two loops which is to be considered later in the loop of loops. However, owing to the simple physical interpretation of the bag, we may consider the outer zone of extended, ordinary space to be obviously known and distinguishable from the inner zone, so that we may regard the expression "outer normal" as having unambiguous meaning. We may also adopt a single right-hand rule for the whole of a large zone of ordinary space, exterior to several small bags, and therefore we may also use a right-hand rule. Then by requiring the beginning infinitesimal loops to be right-handed with respect to the outer normal, so that the concluding infinitesimal loops are left-handed with respect to the outer normal, the ordering of the loops is determined by the sense of description of any one of them. Now, suppose such a figure of sensed circles is rigidly rotated on a sphere by rotating the tangent line, leaving the base point fixed. For a rotation through a straight angle, the figure obtained coincides with that obtained by the alternative method of reversing the sense of all the loops or, equivalently, of any one of them. A quantity so associated to the figure of loops that it must be constant under a continuous deformation of the figure is

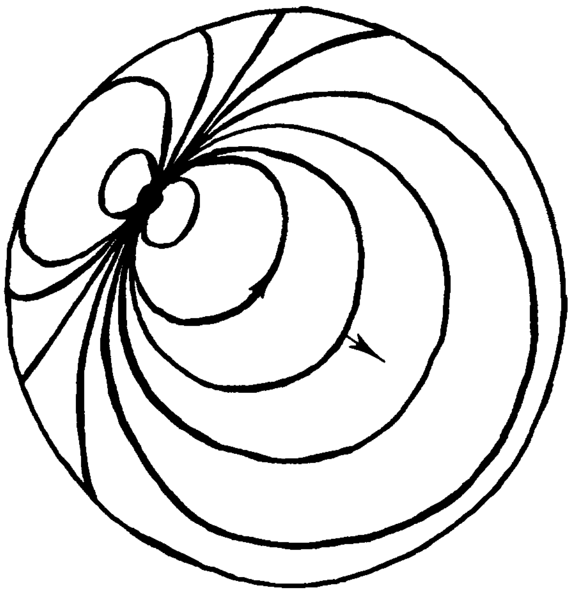

FIG. 4. Figure of loops on a sphere

therefore indifferent to the sense of all or any of the loops, and it is also evidently insensitive to the choice of a tangent through the base point, since that may be continuously rotated.

The quantity of interest is, of course, the homotopy class of the loop in the group consisting of the displacement or curvature matrices generated by the loop of loops: The loop in the group will in general be continuously deformed and nontrivially, because it has infinitesimal degrees of freedom, as the loop of loops is continuously deformed, and will therefore depend on the details of the construction of the loop of loops, but the homotopy class of this loop in the group is therefore invariant. The homotopy class in the group associated to the surface of a sphere by the construction of a figure of loops therefore depends only on the meaning of the right-hand outer normal rule, and on the choice of frame at the base point. It may depend on position of the base point and on choice of frame only insofar as the possible new frame may not be obtainable from the original one by a continuous process. This source of ambiguity is therefore impossible if the group has only one component. But even in the general case, since it is impossible to obtain a curvature matrix outside of the component of identity through a displacement about a loop which is homotopic to a point, and since each loop of the figure is in fact homotopic to a point, the specification of a connected component within the family of all frames at a base point—which I call a generalized orientation (see next section)—is transferable uniquely to all points on the sphere, and it is only this aspect of the choice of base point and base frame which may affect the final homotopy class assigned to the sphere.

What more can be said generally about how the homotopy-class quantity may depend on choice of frame at the base point? We have already noted that the homotopy-class quantity may depend only on the generalized orientation of the frame. A change of generalized orientation of frame will conjugate all curvature matrices, and will take a loop of such matrices in the group **L** into a group-theoretically conjugate loop. The homotopy relationships between conjugate loops must be the same as those between the original loops, since conjugation of all matrices in **L** is a Lie-group isomorphism, but the conjugate loops may belong to homotopy classes distinct from the original ones: a change of orientation may induce a nontrivial automorphism in the fundamental group. In the example where **L** is the circle, the fundamental group is that of the integers, with only one possible nontrivial automorphism: an integer is replaced by its negative. Since these integers will signify a dual charge, this is to say that it is possible for the sign of dual charge to depend on a generalized orientation convention to be imposed from the large region of ordinary space, on all the little bags surrounding clumps of matter, simultaneously. In the example where **L** is the three-dimensional group of rotations in integral spin representation, the fundamental group has only two elements, and no nontrivial automorphism.

When the homotopy classes of two bags are to be added, they are to be based on generalized orientations introduced at once on both bags from the simply connected surrounding zone of ordinary space. From the general discussion of the preceding section, it follows that we have obtained a conserved quantity associated with the homotopy classes in the group's identity component (or any one component). These classes have a group law of multiplication: the product of two classes is the class of the loop in the group formed by successively describing loops in the two classes, so that the conserved quantity is itself an element of a group. There remains the question of ascertaining the law of combination or addition which gives the conserved quantity to be assigned to two bags together from those which belong to the separate bags. It is hard to see how this law could be anything other than the group law of combination in the fundamental group of homotopy classes, and this is what in fact it is.

Figure 5 shows how we combine the quantities contained by two bags: a satisfactory figure on an enveloping bag is evolved continuously from the concatenation of the two loops of loops belonging to the originally separate small bags. Since this induces the group law of combination in the fundamental group of homotopy classes, the law of combination is indeed the group law of the fundamental group. Since this law of combination is commutative, the word "addition" is justified in the usual sense that the term is taken to imply commutativity.

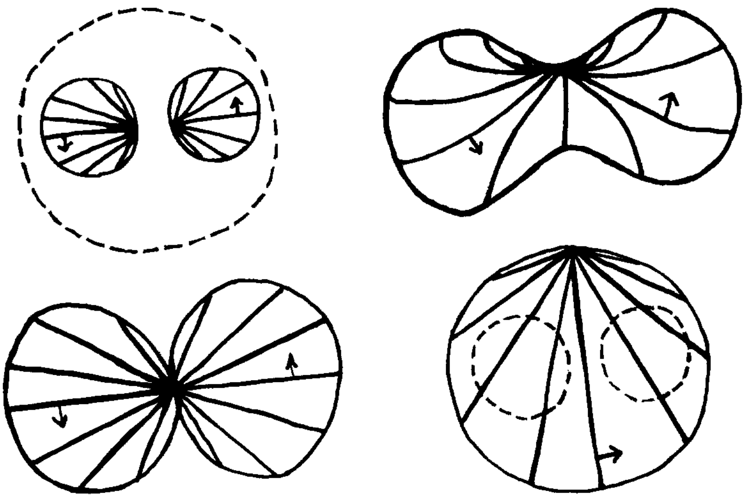

FIG. 5. Fusion of two figured bags

In the construction shown, the small bag on the left comes first. A similar construction made with the small bag on the right coming first leads to an equivalent figure on the large bag. Therefore, we see again that the law of combination must be commutative.

I will call this conserved quantity "dual charge," for the following reason. Since the final point loop is regarded as the boundary of the whole bag punctured at the hinge point of the construction, the displacement about that final point loop—albeit the identity—may be regarded as a surface integral of the curvature over the entire bag, in the sense of the integral defined in Section IX, F. If such a spatial bag were a boundary of a spatial volume, then such a surface integral could be written as the volume integral of the time component of the divergence of the dual infinitesimal curvature; i.e., the dual charge. If such a bag is actually a boundary, the homotopy class of the construction is the identity; the dual charge is zero. If, however, we imagine a small bag in physical space, but admit the possibility of "wormholes" (*20*, *21*), then in a crude, macroscopic sense the bag is a boundary, but microscopically it is not. Therefore, the homotopy-class quantity obtained by examining only the surface of the bag is confused, on the macroscopic level, with what appears to be the total dual charge contained inside the bag. This generalizes the quantization of magnetic charge found by Dirac (*22*).

The quantization, i.e., discreteness, of the dual charge is given directly by that of the fundamental group of group-loop homotopy classes. In the usual way of describing electromagnetic theory, however, the scale factor between the vector potential and displacement becomes involved in this quantization law, so that the object quantized by this argument is the product of the elementary electric and magnetic charges. From this, one then deduces the separate quantization of both types of charge. It is interesting to speculate that some dynamical theories for other groups may involve similar extensions and interpretations of what seems in this abstract presentation to be only quantization of dual charge.

If one ignores complex fields and the usual interpretation of charge, and uses the real line rather than the circle of phase factors as the one-dimensional symmetry group which gives rise to electromagnetism, e.g., by going back to the original gauge invariance of Weyl (*11*), then the remarks about magnetic charges do not actually apply to electromagnetism, the fundamental group then being trivial. This is one unpalatable way to account for the lack of known magnetic monopoles. Another is to ignore possibilities of topological complication of ordinary space-time. There is very likely no real problem, however, as it is likely that understanding of their theoretical properties might explain why magnetic monopoles are unlikely to appear in particle detectors.

If one confines oneself to simple Lie groups, after the line and its nonsimply-connected image, the circle, with its fundamental group isomorphic to the additive group of the integers, there comes the three-dimensional group of rotations. The simply-connected representative is given by any half-integral spin representation. Well-known nonsimply-connected representations are afforded by the integral spin representations. For these, however, there are only two homotopy classes: all loops not homotopic to a point are homotopic to each other, so that the group of dual charges is here the two-element or "yes, no" group. Thus, the isotopic dual charge associated with a Yang-Mills field associated with integral isotopic spin has the option of behaving exactly like a generalized parity —by which I mean any multiplicative quantum number with eigenvalues ± 1.

If we use the general method of producing the nonsimply-connected rotation groups from the spin $\frac{1}{2}$ faithful representation of the simply-connected group by factorization of that matrix group by a discrete central subgroup, it becomes clear that there is no other possible fundamental group. In fact, the matrices are 2 by 2 unitary matrices of determinant 1. An arbitrary complex matrix commuting with all these—an irreducible representation—is a multiple by a complex number λ of the identity, by Schur's lemma. To belong to the group, it must have determinant 1, which already gives $\lambda^2 = 1$; $\lambda = \pm 1$: The entire center of the group has only two elements, whence the fundamental group of any rotation group can have no more.

It is likely that there are three-dimensional rotation groups whose dual charges may be of greater interest than that associated with isotopic spin; the rotations in ordinary space, for example. The fact that a dichotomic, yes-no quantity develops in a mathematically natural way may be of interest in relation to the problem of understanding those dichotomic quantities which are usually introduced a priori. The wildest speculation in this direction would be to hope for an understanding of the Pauli exclusion principle on other grounds than anticommutation relations; I mean no more than to emphasize that the occupation number of a mode is given by "yes, no" rather than by an integer. Since "yes" or "no" is assigned to a bag which may be very small, one may imagine a coarse-grained simply-connected space cut up in meshes of various sizes, with mutually consistent distributions of "yes" and "no" assigned to each mesh, so that the production of a field with different wavelength modes but satisfying an exclusion principle is not hard to imagine. If the rotation group is that of ordinary space, it should be possible to see why a phenomenological exclusion-principle field introduced in this way should have half-integral spin. The fact that it must be built on topological flaws in the space out of the rotation-relevant part of the Christoffel-symbol field applicable to an integral-spin representation lends a pleasant air of weirdness to this project.

Although the source of a dichotomic quantity is different, a similar suggestion has been made by J. A. Wheeler (*24*).

Anticommutation relations appear more mysterious than commutation relations because if we had only the latter, we could look upon the algebraic laws of quantum field theory as that major part of a definition of a group of (unitary) transformations afforded by the commutators of its infinitesimal elements, the important information about the latter being presumably these commutation relations, in analogy to the known theory of Lie groups and Lie algebras. Therefore, we would have an over-all geometric view of the algebraic laws of quantum field theory, the hope of which is greatly damaged by the fact of anticommutation relations.

The different possible laws of addition for dual charge are severely limited by a theorem on commutative groups: A commutative group generated by a finite set of its elements is isomorphic to a direct sum of a finite number of groups each isomorphic to the group of integers, and of a finite commutative group. The finite commutative group may be represented as a direct sum of a finite set of the groups V_j, integers modulo τ_j, and this may be so arranged that each τ_{j+1} is divisible by τ_j (*25*). The theorem applies because the fundamental group of a Lie group is finitely generated. The following argument for this was communicated to be by Dr. L. Greenberg: A connected Lie group is homeomorphic to the direct product of a compact manifold and a Euclidean space (*26*), whence the fundamental group is that of the compact manifold, and that that is finitely generated follows from the introduction of a finite simplicial network (*27*).

There are discrete conserved quantities besides dual charge, but if one limits discussion to the assignment of elements in the fundamental group to bags by devices in which a bag is made to produce a loop of loops, one will not find them: I call the abstract loop of loops a sectioned torus, and regard the specific introduction of a loop of loops on the sphere as a continuous mapping of the sectioned torus into the sphere (surface). Deform a loop on the sphere to a point, and deform the neighboring loops in the loop of loops so as to keep the mapping of the torus continuous. Since one loop on the torus now maps to a point, that loop may itself be squashed to a point, converting the abstract sectioned torus to a sectioned sphere. Two figures introduced by mapping the abstract sphere, with sections marked on it, into the spatial sphere, will yield equal elements in the fundamental group if one can be continuously deformed into the other. Therefore, the element in the fundamental group depends only on the homotopy class of the mapping of one sphere into another. These classes form the second homotopy group of the sphere, which is isomorphic to the group of integers, its generator being represented by the particular construction given here. The powers of this generator correspond to the possible reversal of the right-hand rule, counting the dual charge more than once, or perhaps not at all, and are therefore not interesting.

I now sketch a wormhole construction of a bundle and connection which realizes an arbitrary dual charge—were the bag homotopic to a point, the loop in the group induced by a figure of loops on the bag would also be homotopic to a point, and the dual charge, zero, whence some kind of wormhole is essential.

Excise a sphere in each of two 3-spaces, and paste the spaces together at these spherical cuts. Split the spheres concentric with the seam sphere into hemispheres, by one plane, and make these hemispheres overlap slightly. Mark a radius in the splitting plane, which penetrates each sphere in one point. Let the seaming matrix at that point be the identity, and execute the desired loop in the matrix group as the point runs around the circle marked by the splitting plane. The infinitesimal components of connection may be quite arbitrary; e.g., they may vanish: Dual charge is a phenomenon to be reckoned with even in **L**-bundles without connections.

The construction is easily modified to yield a pair of opposite dual charges by excising both spheres from the same 3-space.

It should be noted that the conservation of ordinary charge, which depended on the transformation of a volume to a surface integral on a remote surface, has a clear meaning as a conservation law only if all the wormholes which begin in the gross "inside" of the surface also end in it, so that in fine detail, the surface really does bound a volume. The formal transformation to a surface integral otherwise involves very many surfaces close to or grossly "inside" matter, if one admits topological complications. If one wishes to handle electric charge as ordinary, not dual, charge, one must either regard that as coming from charge density, or from electric flux through wormholes (*20*). If the latter, an isolated electron must join a distant positive charge by a wormhole, and the discussion of the conservation of ordinary charge based on integrals of densities would not apply in electromagnetism.

The discussion of charge from a generalized geometrodynamic viewpoint, in which current densities are put to zero, could perhaps isolate the charge due to the trapping of flux by topological complications. If densities are not put to zero, one could not approach the ordinary charge as a topological invariant of a bag, inasmuch as continuous distortion of the bag would continuously alter the content of charge, and make the contribution due to trapped flux inherently undefinable.

XVI. COMPONENT GROUP: GENERALIZED ORIENTATION

It has already been noted in Section XIV that the homomorphism of loops to the matrices giving the effect of displacement around them, called curvature, induces a homomorphism of the group of homotopy classes of loops in the manifold into the component group, "discrete curvature." Confinement of the loops to a simply connected patch leads only to the identity, so that the discrete curvature mapping explores the **L**-bundle, not the details of the connection. If **L** is locked, the connection may however be materially conditioned by the discrete curvature. Physical interpretation of infinitesimals associated with the connection may therefore invest discrete-curvature properties with physical meaning, and especially conserved quantities related to the discrete curvature should interact with other physical quantities in characteristic ways.

If the matrix group **L** has several components, then that is already reflected in the family of frames at one point of the manifold of an **L**-bundle: The family of frames may be obtained from any one frame by applying **L**, and the topology of **L** obviously induces a topology in this family of frames, in which each component is obtained from a component of **L**.

These components in the family of frames at a point may be called generalized orientations, as they generalize the concept of orientation in Riemannian geometry. The vector spaces there are the tangent spaces of the basic m-dimensional manifold, **L** is the group of m by m orthogonal real matrices, and the frames are those bases orthonormal in the metric. The matrices **L** fall into two components, characterized by determinants 1, -1, and so the frames at any point fall into two components. If two frames belong to the same component, they are said to share the same orientation; otherwise, to have opposite orientations. If the discrete curvature is trivial, as it is when one limits ones loops to lie within a simply connected region, the orientations may be propagated consistently from any point, and one speaks of orientability of a portion of the space or of all of it. If the discrete curvature is nontrivial, the space being in particular not simply connected, the space is said to be nonorientable.

In the case of a generalized orientation, there may be levels of orientability intermediate between complete orientability and nonorientability. To clarify the range of possibilities, it will be best to make the definitions a little more explicit, by emphasizing the requirement that the loops in the manifold share a common base point with a common base orientation. (If one defines the curvature by transporting the frame around the loop, one winds up with a final frame which will not generally coincide with the initial frame of the next loop and therefore the definition of a product of loops by successive description is threatened. This difficulty has been circumvented by referring variously transported vectors to one base frame at the base point.) The discrete-curvature image of the loops in the component group will constitute a subgroup, invariant to continuous change of the universal base point if the initial orientation is propagated continuously, and which will suffer conjugation by an arbitrary component-group element if the initial orientation is arbitrarily changed. This class of conjugate subgroups will be called the **L**-bundle's deorientizer. The different kinds of orientability are given by the distinct conjugate classes of subgroups of the component group, the number of orientations available in any one of these kinds of orientability is given by the index in the component group of any one of the conjugate subgroups in the deorientizer, and so, in particular, divides the order of the component group. The language is obviously simplified for commutative component groups.

The limitation of possible deorientizers to classes of conjugate subgroups of the component group is supported from below by a complete set of examples of such generalized deorientizers built upon structures in a 3-space manifold. Suppose the subgroup which we wish to set up as representative of a deorientizer has g generators. Excise g nonintersecting tori which share a common equatorial plane from Euclidean 3-space, and introduce g new Euclidean 3-spaces, each with one excised torus. Cut each space in half along the equatorial plane, so that the $g + 1$ spaces produce $2g + 2$ pieces, each of which is homeomorphic to Euclidean 3-space, and therefore a possible simple patch. Seam the g toroidal cuts in the one 3-space to the toroidal cuts in the g other spaces, producing what I will call g toroidal wormholes. Seam the planar cuts in each of the $g + 1$ spaces back to what was joined before the cuts were introduced. For **L**-matrices in the seamings, use the unit matrix, except for the circular planar cuts across the doughnut holes of the g toroidal wormholes, where in each pair of circular cuts associated with a toroidal wormhole an **L**-matrix corresponding to a component group generator is employed. Imagine all the pieces extended slightly into the pieces to which they are seamed, so that we may speak of the seam overlap regions. Then there are never more than three overlapping patches on a point, and that only for points near the edge of the planar circular cuts. The three-patch seaming axiom requires that the **L**-matrices used in a pair of circular cuts associated with the same toroidal wormhole, in the same sense, coincide; the two-patch axiom requires only that the inverse matrix be employed in threading a toroidal wormhole in the reversed sense.

If, instead of toroidal wormholes, we built simple crosscaps by excising g spheres in one 3-space, identifying diametrically opposite points, and then employing g matrices $L_1, \cdots, L_g$ on the g crosscaps, the construction would be limited by the equations $L_i^2 = 1$ coming from the 2-patch seaming axiom, although this could also be traced to the 3-patch seaming axiom with some manipulation. This "fault" is, of course, intrinsic to crosscaps, in that the curve homotopy group associated with any one crosscap is a 2-element group, and not a free cyclic group.

It has already been remarked in Section XV that the specification of the dual charge assigned to a bag surrounded by simply-connected ordinary space on the outside depends not only on a right-hand rule, but possibly also on a choice of generalized orientation in this simply connected ordinary region. It should also be noted that this possibility exists for the case of the curvature and the current densities. In fact, if a quantity q is given by an element $q^I E_{I(\alpha)}{}^{(\beta)}$ in the infinitesimal algebra, then it will be transformed to the conjugate $q^I F E_I F^{-1} = q'^I E_I$, where F is the matrix which takes the components of a vector on the old frame to those on the newly oriented frame, and this represents some component distinct from that of the identity. The quantities q^I will be unchanged only if $FE_I F^{-1} = E_I$, for all I. If for each possible change of orientation such an F can be found, then the group generated by such F's commutes with the component of identity, and its direct product with the component of identity is isomorphic to the given group **L**. Therefore the locked involvement of a component group necessitates a material dependence of infinitesimal-algebra quantities q^I, including infinitesimal curvature and charge density, on choice of orientation.

In order to develop the possibility of related conservation laws, I again consider a picture of ordinary 3-space with one or several matter-containing pockets cut out by bags. The vacuous region is simply connected, and so its deorientizer, defined, of course, on the basis of only those loops which lie entirely within the vacuous region, is trivial. If the interdiction on the crossing of any one bag is dropped, however, the deorientizer may no longer be trivial, and in that case it will specify a nontrivial property of the matter in the bag. More precisely, the bags must be considered to move in time so as to surround world-zones of separated lumps of matter, and the vacuous zone of space-time must be cut off from interminglings of such lumps of matter by early and late spacelike boundaries, with the conservation law applicable only to the limited time interval between, so that the interdictions on the crossing of bags may not be subverted by motion in time. Such moving bags may be called timelike cylinders; annihilation would correspond to the possibility of a bent cylinder, i.e., of timelike

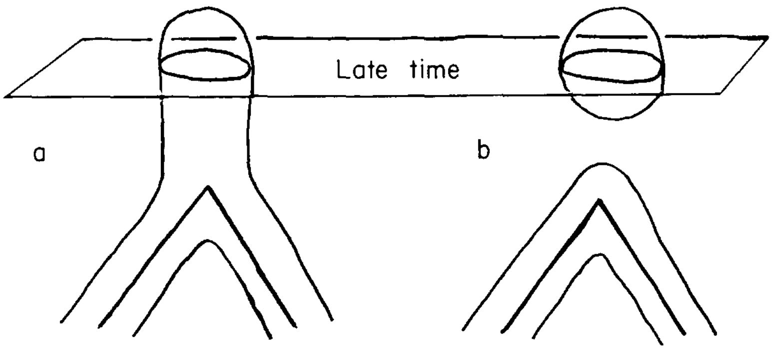

FIG. 6. Dependence of deorientizers on a boundary which extends in time

cylinders which meet in a latest bag out of which is ejected matter without deorientizing effect.

The situation here is, however, inherently different from the study by bags of the fundamental group. In that case, the quantity of interest depended directly on the two-dimensional bag, which could be regarded as contained in one instant of time, as interpreted from the outer, "naïve" region of space-time. The difference between deorientizers with or without a cut, however, depends on a boundaryless three-dimensional cut in 4-space, and this is not defined by the two-dimensional bag or bags which it marks on any one spacelike hypersurface drawn in the vacuous zone: the same bag at one instant could be completed into a three-dimensional cut in two different ways, such that the deorientizers introduced by removing the different 3-cuts attached to one bag are different. Thus, in Fig. 6a, the nontriviality of the 3-cut isolating an annihilation-bent line is felt if this cut is opened at the "late time" after the annihilation event, whereas the opening of the hollow 3-cut in Fig. 6b does not release the nontriviality of the annihilation line. The relation of a 2-bag in 3-space to its naïve apparent instantaneous contents is broken, and the deorientizers associated to 3-cuts had perhaps better be presented as labels characteristic of possibly bent world tubes, than as conservation laws.

The addition of the deorientizers of two disjoint 3-cuts, well-separated from the viewpoint of the ordinary, outside region, requires more information than the specification of the separate deorientizers. If the removal of both 3-cuts fails to bring in any essentially new loops because of the possible connection of the mysterious "insides" of the apparently unrelated 3-cuts, then for each subgroup in the first deorientizer there is one in the second such that the two subgroups together generate the combined deorientizer, but if essentially new loops are opened, the combined deorientizer will be larger. In the fusion of 3-cuts, there is never any cancellation of separate deorientizers; there is no notion of the deorientizer of a particle and that of an antiparticle being separate and cancelling, but at best one for a 3-cut isolating a possibly folded world-path.

XVII. THE INTRODUCTION OF GENERALIZED ORIENTATION BY CHARGE CONJUGATION INTO THE ELECTROMAGNETIC CASE

An example of a locked multicomponent group is generated by "charge conjugation" and the usual gauge group of electromagnetic theory. The gauge or phase transformations are, to be specific, $\phi \to \phi e^{i\theta}$, where ϕ is a complex spinless c-number field and θ is a real number modulo 2π, and charge conjugation is complex conjugation of the field ϕ.

In order to represent complex conjugation linearly, the complex numbers will be represented by their real and imaginary parts. Then the phase transformation with parameter θ, or rotation by θ, is effected by the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = R\ (\theta)\,,$$

complex conjugation is effected by the matrix

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

and these generate reflections $S(\theta)$, where the line is which the reflection is made is oriented at angle $\theta/2$ modulo π from the real axis; the $S(\theta)$ like the $R(\theta)$ are indexed by real numbers modulo 2π. The component of identity, consisting of the rotations, is commutative; the reflections form a second component, so that the component group is the two-element group, commutative. The whole group is not, however, commutative; the multiplication table is

$$\begin{aligned} R(\varphi)R(\theta) &= R(\theta + \varphi), \\ R(\theta)S(\varphi) &= S(\theta + \varphi), \\ S(\varphi)R(\theta) &= S(\varphi - \theta), \\ S(\varphi)S(\theta) &= R(\varphi - \theta); \end{aligned}$$

the commutator of a rotation with a reflection is

$$S(\varphi)R(-\theta)S(\varphi)R(\theta) = R(2\theta);$$

the commutator subgroup is the component of the identity. (The group S generated by the commutators with the component of identity or with its infinitesimal elements of one or more representives from each component is a normal subgroup of the component of identity distinct from the identity alone, and possessed of an infinitesimal element provided that **L** is locked, and therefore C is the whole component of identity whenever the infinitesimal algebra of **L** is simple.)

The discussion below will depend only on the abstract group of $R(\theta)$'s and $S(\theta)$'s, and will therefore not be limited to the given realization of this group on the model of a c-number spinless field ϕ.

The discussion of generalized orientation touched on two issues: first, the ambiguity produced in the definition of dual charge and of infinitesimal-algebra quantities, e.g., infinitesimal curvature and current densities, and second, the possibility of new structures, poorly defined in terms of deorientizers and hyperbags, and both these issues are briefly elaborated for the present example. However, since here there is no intermediate between generalized orientability and nonorientability, the mechanism of deorientizers is so poor that it will be dropped.

Our first topic, then, is the matter of possible change of magnetic and electric charge when an S-reflection of the frames is made. It is easily seen that both types of charge do, in fact, reverse: Since $S(\varphi)R(\theta)S(\varphi) = R(-\theta)$, the infinitesimal rotations are simply reversed, so that all infinitesimal algebra quantities, inluding the curvature—i.e., fields—, and the current density, reverse. For the case of magnetic charge, reversal is already implied in the reversal of the magnetic flux, but a direct argument is also easy: The original magnetic charge is given relative to a geometric figure of loops on a bag which is unaltered by the S-reflection of frames, but the $R(\theta)$'s assigned to loops are replaced by $R(-\theta)$'s whence the loop induced in the rotation component of the group, the circle, is the loop of inverses of the original loop taken in the original, not the reversed, order, which in this case belongs to the inverse element of the fundamental group, that is, to opposite dual charge. In this sense, then, the S-reflections, and in particular, $S(0)$, are charge conjugations.

I will now discuss three deorientizing structures: a crosscap and a toroidal wormhole, as defined in the previous section, and a toroidal crosscap.

The general objection to use of a crosscap, in that the 3-space fundamental group is then only a two-element group and will not support a complicated deorientizer, does not apply here, because each reflection $S(\theta)$ actually squares to the identity $R(0)$. We therefore imagine a crosscap bearing, say, an $S(0)$ reflection. The $S(0)$ reflection alone induces a reversal of electric and magnetic charge; more precisely, the electric and magnetic charge defined relative to spatial frames propagated continuously through the crosscap seam considered therefore as a seam of vectored patches only, indeed reverse. Since those continuously propagated frames bear identical ordinary orientation to frames continuous with the original frames within the simply connected region exterior to a large bag around the whole crosscap, the electric and magnetic charges, as defined with respect to one consistently oriented family of frames in the reduced space out of which the crosscap has been excised, are in fact reversed; if one charge (electric or magnetic) remains fixed, while a neighboring charge is threaded through the crosscap and then brought back near to the fixed charge, its charge relative to that of the fixed charge will be found to have reversed. Nevertheless, conservation of charge of both kinds is preserved, the missing charge being made up by a charge assigned to the crosscap itself.

This statement is supported easily for the magnetic charge: The total magnetic charge deduced from a bag surrounding both the small magnetic charge and the crosscap must remain the same when the sign of the magnetic charge on the small magnetic charge has reversed, and the only nonsimply-connected region which may be bagged for the compensating magnetic charge is the zone of the crosscap. It is also possible to see in detail how this happens by drawing Faraday lines of force. Instead of attempting a freehand motion picture here, I note the fact that without an $S(0)$ operation on the lines of force, a system of lines of force which enter a crosscap punches through to another system which leaves, so that no net flux threads a bag surrounding the crosscap, but that if the crosscap is made to bear a charge conjugation operation, then such an entering system punches through to a second entering system, owing to reversal of the sense of the fields, thereby yielding net flux through a bag surrounding the crosscap.

The same image of Faraday lines of force is obviously applicable to the electric charges, so that the answer when the dynamics of the infinitesimal curvature field is governed by Maxwell's equations is clear. But there is no hope that the argument be supported by appealing to a conservation law framed in terms of the volume integral of a density, since the charge which is missing as a density is balanced by an everywhere divergenceless flux. The effective use of charge conjugation thus presents a mechanism for the exchange of charge between a mysterious density and a definite crosscap model.

The density conservation law may, however, be employed indirectly to derive charge conservation, as measured by total flux. If a large bag is drawn to surround the entire event, and this bag is considered as a boundary of the void region exterior to it, then one concludes that the electric flux through it remains constant as the charge (small zone of charge density) threads through the crosscap. This constant flux is originally and also finally the sum of that emerging from a bag containing the charge density and that containing the crosscap region.

If electric and magnetic charges run through a crosscap without a charge conjugation, they do not reverse sign, and the conservation arguments show that the crosscap region does not alter either of its charges. Boundary conditions appropriate to no charge conjugation also make the surface integral of curvature on a bag barely surrounding a crosscap zero, by the cancellation of contributions from elements of area immediately over crosscap-identified points.

The situation for a toroidal wormhole is similar. When there is no charge conjugation across the hole of the doughnut, a magnetic or an electric charge which threads through it does not change sign, of course, and the field which remains stuck to the wormhole is a dipole field, but when there is a charge conjugation, the charge's sign reverses, a change which is compensated by a monopole field on the wormhole. The field emanates both ways from a disc drawn across the hole of the doughnut, which is, of course, of the same artificial significance as a branch cut in analysis, the field being amenable to smooth continuation through such a disc. The picture of Faraday flux lines is most simple if the excised tori have very small volume, but very large doughnut holes, so that the flux in the space in which the charge moves finds it difficult to punch through into the second attached space.

A toroidal crosscap is interesting because it renders the underlying 3-manifold nonorientable in the usual sense of orthogonal 3-space geometry, and thereby behaves differently with electric and magnetic charges. By a toroidal crosscap, I mean the structure formed as follows: Excise a torus from an ordinary 3-space, and then identify diametrically opposite points on, say, the circles of longitude; unlike the case of the toroidal wormhole, no second 3-space or other excised torus in the first 3-space is needed to heal the cut. It is easy to see that, if a little orthogonal triad passes through the cut, it emerges with reversed ordinary orientation. Then, if no S-reflection is used in defining the parallel displacement of phase, there is no change of sign of electric charge produced by threading the cut, but there is change of sign of magnetic charge, because although the definition of outer normal and the $R(\theta)$'s assigned to the loops in a figure of loops remain unchanged when a magnetic charge-containing bag is taken through the cut, the right-hand rule for picking a first loop in the figure of loops is changed. The last possibility—that of changing the sign of electric charge but not changing that of magnetic charge—is achieved by a toroidal crosscap with an S-reflection imposed on the frames for phase.

If $R(\theta)$ is replaced by $R'(\theta)$, shift by θ in the positive sense of the real line, and $S(\theta)$ by $S'(\theta)$, reversal of the real line leaving $\theta/2$ invariant, then the above holds, except that θ is no longer modulo 2π, and the statements about magnetic charge are void, because the fundamental group has only one element, so that the magnetic charge vanishes.

XVIII. LAGRANGIANS

One may explore the classical mechanics and, from the canonical group of that, the quantum mechanics of frame-invariant field functions of the invariant points on connected **L**-bundles, used as Lagrangians. The Lagrangian $-\frac{1}{4}f_{\mu\nu}f^{\mu\nu}$ for electromagnetic theory suggests $g_{IJ}R^{I}{}_{\mu\nu}r^{J\mu\nu}$, where a metric tensor has been assumed to raise the indices, and where the g_{IJ} are numeric coefficients which transform like the E_IE_J under change of the vector-space basis in the Lie algebra. The $c^{L}{}_{IK}c^{K}{}_{JL}$, perhaps multiplied by a constant, may be used for g_{IJ}, always yielding symmetric g_{IJ} and yielding nonsingular g_{IJ} for semisimple Lie algebras.

This is the method by which Yang and Mills obtain a Lagrangian. $g_{IJ}\epsilon^{\mu\nu\varpi\rho}R^I{}_{\mu\nu}R^J{}_{\varpi\rho}$, which needs no metric tensor, is the generalization of the quadratic pseudoscalar invariant of the electromagnetic field. $g_{IJ}j^{I\mu}c^J{}_\mu$ is a generalization of the $j^\mu A_\mu$ term of electromagnetic theory. There are therefore many possible sets of dynamical equations to explore.

XIX. CONCLUDING REMARKS

It may be useful to briefly augment the beginnings which have been suggested. The role of representation bundle as the object underlying an **L**-admissible connection suggests the applicability of the mathematical machinery of bundle theory. The discussion of tangent spaces by trying to build on the easier discussion of the bundle of frames they define, then considering the transformation relating the frames to the coordinate-bound bases, has only been hinted at here. It may not be difficult to mimic the proofs in Riemannian geometry, and thereby give discussions of complete systems of identities for the infinitesimal curvature. It is not clear what consequences can be drawn from the "peculiar integral" in the noncommutative cases. Locking, essential to nontriviality of the multicomponent structure, should be characterized. The discussions of topological peculiarities should be related to a detailed classification of all the possibilities.

Before a quantum theory including provision for variegated topological peculiarities of the underlying manifold (*21*) is at hand, there is yet some possible practical—i.e., quantum-mechanical—use in such considerations, in that the conservation laws they suggest may provide for new reasons for poorly understood quantum numbers. Just as in geometrodynamics, structures built on topological peculiarities are expected to have mass, so that the position of mass in totally gauge-invariant theories may be simpler than the possible masses obtained in sophisticated considerations over simply connected ordinary space-time. Discussions which average over large quantities of topological peculiarities should yield new classical theories with extra densities, analogous to the current and monopole current densities of the totally inhomogeneous Maxwell equations.

The topological peculiarities presented here as associated with small zones of space may, of course, also be big (*28*): new forms of connectivity for ordinary space-time is traditional cosmology.

ACKNOWLEDGMENTS

This work is an offshoot of an unsuccessful attempt to better understand spin ½ by imitating the work of Wheeler and Misner (*20*), with a free neutrino field replacing the free electromagnetic field. Thanks are due to Dr. Sanford Wolfe for acquainting the author with Prof. J. J. Sakurai's stimulating paper on generalized gauge invariance (*29*). The line of work became the explicit demonstration that the idea of general relativity itself was involved, an idea behind all the work in this field, and the declared motivation of the basic paper of Yang and Mills (*13*). Thanks are due to Dr. A. Pais for informing the author that R. Utiyama had already presented the formalism of introducing the general Lie algebra (*14*).

The work continued when the question, in what way do Lie-group properties beyond those contained in the infinitesimal algebra manifest themselves, was raised by the author in contradiction to his own flat statement that "gauge invariance" is a strictly infinitesimal subject, addressed to the High Energy Physics Group at Brown University. Thanks are due to Prof. D. Feldman for his broad view of "high energy physics."

RECEIVED: December 7, 1962

REFERENCES

1. S. L. GLASHOW AND M. GELL-MANN, *Ann. Phys.* (*N. Y.*) **15**, 437 (1961).
2. R. E. BEHRENDS, J. DREITLEIN, C. FRONSDAL, AND B. W. LEE, *Rev. Mod. Phys.* **34**, 1 (1962).
3. D. R. SPEISER AND J. TARSKI, Possible schemes for global symmetry. Institute for Advanced Study, Princeton, preprint (1962).
4. C. CHEVALLEY, "Theory of Lie Groups." Princeton Univ. Press, Princeton, New Jersey, 1946.
5. L. PONTRJAGIN, "Topological Groups." Princeton Univ. Press, Princeton, New Jersey, 1939.
6. N. BOURBAKI, "Elements de Mathematique; Groupes et Algebres de Lie." Hermann, Paris, 1960.
7. G. RACAH, Group theory and spectroscopy. Institute for Advanced Study lecture notes, 1951; available from The Hebrew University, Jerusalem.
8. E. B. DYNKIN, *Am. Math. Soc. Trans.* Ser. 2, **6**, 245 (1957).
9. H. WEYL, *Z. Physik.* **56**, 330 (1929); T. LEVI-CIVITA, *Sitzber. Preuss. Akad. Wiss. Phys.-Math. Kl.* p. 346 (1932); J. A. SCHOUTEN AND D. VAN DANTZIG, *Ann. Math.* **34**, 271 (1933); F. J. BELINFANTE, *Physica* **7**, 305 (1940); W. L. BADE AND H. JEHLE, *Rev. Mod. Phys.* **25**, 714 (1953); D. BRILL AND J. A. WHEELER, *Rev. Mod. Phys.* **29**, 465 (1957); T. W. B. KIBBLE, *J. Math. Phys.* **2**, 212 (1961).
10. E. LUBKIN, Frames and Lorentz invariance in general relativity. UCRL 9668, Univ. of California (1961).
11. H. WEYL, *Ann. Physik* **59.2**, 101 (1919).
12. H. WEYL, *Z. Physik.* **56**, 330 (1929).
13. C. N. YANG AND R. L. MILLS, *Phys. Rev.* **96**, 191 (1954).
14. R. UTIYAMA, *Phys. Rev.* **101**, 1597 (1956).
15. O. VEBLEN AND J. H. C. WHITEHEAD, The foundations of differential geometry. Cambridge Tracts in Math. and Math. Phys., No. 29, 1932, repr. 1953; N. E. STEENROD, "Topology of Fibre Bundles." Princeton Univ. Press, Princeton, New Jersey, 1951.
16. E. M. CORSON, "Introduction to Tensors, Spinors, and Relativistic Wave Equations." Blackie, London, 1953; K. M. CASE, *Phys. Rev.* **107**, 307 (1957).
17. A. EINSTEIN, "The Meaning of Relativity," 5th ed. Princeton Univ. Press, Princeton, New Jersey, 1955.
18. O. VEBLEN, Invariants of quadratic differential forms. Cambridge Tracts in Math. and Math. Phys., No. 24, repr. 1952.
19. J. GOLDSTONE, A. SALAM, AND S. WEINBERG, *Phys. Rev.* **127**, 965 (1962).
20. C. W. MISNER AND J. A. WHEELER, *Ann. Phys.* (*N. Y.*) **2**, 525 (1957).
21. J. A. WHEELER, "Geometrodynamics." Academic Press, New York, 1962.
22. P. A. M. DIRAC, *Proc. Roy. Soc. London* **133**, 60 (1931); *Phys. Rev.* **74**, 817 (1948).
23. R. FINKLESTEIN, *Ann. Phys.* (*N. Y.*) **12**, 200 (1961).
24. J. A. WHEELER, "Geometrodynamics," pp. 93–94. Academic Press, New York, 1962.

25. L. Pontrjagin, "Topological Groups," p. 24. Princeton Univ. Press, Princeton, New Jersey, 1939.
26. K. Iwasawa, *Ann. Math.* **50**, 530 (1949).
27. H. Seifert and W. Threlfall, "Lehrbuch der Topologie." B. G. Teubner, Leipzig, Berlin, 1934.
28. F. Leiber, *Sci. Am.* **206.4**, 128 (1962).
29. J. J. Sakurai, *Ann. Phys.* (*N. Y.*) **11**, 1 (1960).