Symmetries, Groups, Algebras, Connections, and Gauge Fields.

1 Symmetries and Groups

Many laws of Nature are invariant under certain transformations, like translations, reflexions, rotations, time reversal, charge conjugation, Lorentz transformations, space-time diffeomorphisms, ... In classical mechanics the well-known Noether theorems link symmetries of space and time to conservation laws, for example homogeneity and isotropy are related to momentum and angular momentum conservation.

Obviously, the composition of two symmetry transformations must result in a symmetry transformation again, moreover, it turns out that the axioms of group theory provide a natural framework for symmetry transformations.

Group axioms: The combination $(G, \circ)$ of a set G and a binary operation $\circ$, usually called multiplication, defined on G is called a group, if

1. $g \circ h \in G \; \forall \, g, h \in G$ (closure),
2. $(g \circ h) \circ f = g \circ (h \circ f)$ (associativity),
3. there exists a unit element e so that $g \circ e = e \circ g = g \; \forall \, g \in G$,
4. $\forall \, g \in G$ there exists an inverse element g^{-1}, so that $g \circ g^{-1} = g^{-1} \circ g = e$.

If, in addition to 1. - 4., a commutative law $g \circ h = h \circ g \; \forall \, g, h \in G$ holds, the group is called commutative or Abelian. In the following we shall mostly use the shorthand notation G for $(G, \circ)$, with the operation $\circ$ implicitly included.

Example: The group O(3) of orthogonal transformations in 3-dimensional Euclidean space R^3, subgroup of the Galilei and the Lorentz group. The group elements act as matrices M on the elements $\vec{x} = (x, y, z)$ of R^3, leaving the inner product - and, in consequence, the norm - invariant, i. e. $|M\vec{x}| = |\vec{x}|$. In index notation this reads

$$M^i{}_k \, x^k \, M^i{}_l \, x^l = \delta_{kl} \, x^k \, x^l, \tag{1}$$

from which we may conclude

$$M^i{}_k \, M^i{}_l = \delta_{kl} \qquad \text{or} \qquad M^T \, M = \mathbf{1}, \tag{2}$$

where M^T is the transposed matrix and $\mathbf{1}$ is the unit matrix. In other words, the transposed of an orthogonal matrix is its inverse. There is an immediate consequence of this property for the determinant

$$\det M = \det M^T = \det M^{-1} = \frac{1}{\det M}, \tag{3}$$

which means that $\det M = \pm 1$. The two values ± 1 of the determinant in O(3) have an important geometric meaning: matrices, whose determinant is equal to $+1$ leave the orientation of an arbitrary basis in R^3 unchanged, whereas matrices with determinant -1 invert the orientation. The former ones, often called "even" transformations, form the subgroup of rotations, denoted by SO(3), the group of special orthogonal transformations. A subgroup is a subset of a group, which satisfies the group axioms. The "odd" transformations with determinant -1 do not form a subgroup, because they do not contain the unit element. They can be decomposed into a rotation and a space inversion, the latter one represented by the matrix $P = \mathrm{diag}(-1, -1, -1)$.

The rotation group SO(3) Because of the symmetry of the indices k and l, (2) is a set of 6 equations for the 9 components of a 3×3 matrix, so that there remain 3 independent components of a rotation matrix. Therefore SO(3) is a three-parameter group. A convenient choice of the parameters are the so-called Euler angles, $0 \leq \alpha, \gamma < 2\pi$, $0 \leq \beta < \pi$ with the following meaning: An arbitrary rotation matrix $R \in SO(3)$ can be written as the composition of three subsequent rotations by an angle α around the axis $\vec{e}_3$, followed by a rotation by an angle β around the rotated axis $\vec{k} = R(\alpha)\vec{e}_1$ and finally by an angle γ around the new axis $\vec{e}_3'$,

$$R(\alpha, \beta, \gamma) = R_{\vec{e}_3{}'}(\gamma) \cdot R_{\vec{e}_1{}'}(\beta) \cdot R_{\vec{e}_3}(\alpha). \tag{4}$$

More explicitly,

$$R = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5}$$

Group manifolds Every group element can be represented as a point (α, β, γ) in the parameter space. On the other hand, every $R \in SO(3)$ is equivalent to a single rotation around a certain axis $\vec{e}$. With this interpretation in mind, we can visualize the group as a solid sphere: $\vec{e}$ is characterized by the angles ϑ and φ in spherical coordinates, and the rotation angle $0 \leq \chi \leq \pi$ determines the radial coordinate, so we have a sphere of radius π. As rotations around an axis $\vec{e}$ by an angle π are equivalent to rotations by π around the antiparallel axis $-\vec{e}$, the antipode points of the sphere must be identified. This is the compact group manifold of SO(3). Manifolds are typical for groups, which depend in a continuous way on their parameters, so that the group multiplication is a continuous operation (differentiable). Groups of this kind were studied first 1893 by Sophus Lie.

Exercises:

1. Verify the group axioms for orthogonal matrices!

2. Consider two matrices A and B, describing a rotation by an angle α around the x-axis and by an angle β around the y-axis. Calculate AB and BA and their respective rotation axes!

Subgroups, Cosets As mentioned above, a subgroup $K \subset G$ is a subset of G that satisfies the group axioms. The **left coset** of the element $g \in G$ with respect to the subgroup K is the set

$$gK := \{gk : g \in G,\ k \in K \subset G\}, \tag{6}$$

where k runs over all elements of K. Analogously the right cosets are defined as

$$Kg := \{kg : g \in G,\ k \in K \subset G\}. \tag{7}$$

From a left (right) coset of an element g an arbitrary element gk (kg) can be chosen as a representative. The cosets of a group are disjoint, i. e. $gK \cap hK = \emptyset$, when g and h are representatives of different cosets, and the union of all cosets is the group G, $\bigcup_g gK = G$.

If for a subgroup $N \subset G$ the left and the right cosets are identical, $gN = Ng\ \forall\, g \in G$, the subgroup is called **normal subgroup** or invariant subgroup. In this case a multiplication of cosets can be introduced. Let $\nu_i = g_i N$ and $\nu_j = g_j N$ be two cosets. In the product of two representatives, $g_i n_1\, g_j n_2$, with $n_1,\ n_2 \in N$, $n_1 g_j$ can be written as $g_j n_3$, with n_3 being another element of N, leading to

$$\nu_i\, \nu_j = g_i N\, g_j N = g_i g_j\, NN = g_i g_j\, N = \nu_{ij}$$

in shorthand notation, where ν_{ij} is the coset of $g_i g_j$. It is easy to see that the other group axioms are satisfied with $e = N$ and $\nu_i^{-1} = g_i^{-1} N$. The group of cosets of a group G with respect to a normal subgroup N is called the factor group or quotient group G/N. The **centre** of a group, the set of all elements that commute with everything,

$$Z = \{z \in G : gz = zg \; \forall \, g \in G\}, \tag{8}$$

is the maximal Abelian normal subgroup.

Example: O(3) and SO(3) are non-abelian groups, $SO(3) \subset O(3)$. Convenient representatives of the cosets are the unity matrix $\mathbf{1}$ and the parity transformation $P = -\mathbf{1}$. Both of them commute with all rotation matrices $R \in SO(3)$, so the latter one is a normal subgroup and the factor group $O(3)/SO(3) = \{\mathbf{1}, P\}$ is isomorphic to $Z_2 = \{1, -1\}$.

2 Lie Groups

Lie groups, which O(3) and SO(3) are representative examples of, are described by

1. $(G, \circ)$ is a group.
2. G is an analytic manifold of dimension $d_G = n$, n is the number of parameters, there are local coordinates x^i, $g = g(x^1, x^2, \ldots, x^n)$, with g depending analytically on them.
3. The mapping $(g(x^i), g(y^i)) \to g(x^i)\, g^{-1}(y^i)$ is analytical, i. e. together with $g(x^i)$ also $g^{-1}(x^i)$ is analytical, and the multiplication is analytical.

(3.) means that when a group element $f(p_1, p_2, \ldots, p_n)$ is composed with $g(q_1, q_2, \ldots, q_n)$ to give an element $f \circ g(r_1, r_2, \ldots, r_n)$, then the parameters r_i are analytic functions of p_i and q_i, and the parameters of the inverse element f^{-1} are analytic functions of p_i.

Usually the parametrization is chosen in such a way that the unit element lies in the coordinate origin, $e = g(0, 0, \ldots, 0)$. This is the case for SO(3) and the Euler angles, where $R(0, 0, 0) = \mathbf{1}$ and further $R(\alpha, \beta, \gamma)^{-1} = R(\pi - \gamma, \beta, \pi - \alpha)$ (modulo 2π).

Another example: The **Lorentz group** as a matrix group is the set of 4×4 matrices which are orthogonal in the sense of the Minkowski metric $\eta_{ik} = \pm \mathrm{diag}(-1, 1, 1, 1)$, leaving inner products of the form $x^i y^k \eta_{ik}$ invariant. This is equivalent with the relations

$$L^i{}_m \, L^k{}_n \, \eta_{ik} = \eta_{mn}, \qquad \text{or} \qquad L^T \eta L = \eta \tag{9}$$

for Lorentz matrices L. (9) is a set of 10 equations for the total of 16 matrix elements, so there are 6 independent parameters, namely a velocity vector $\vec{v}$ and 3 rotation angles. A matrix satisfying (9) can be written in the form

$$L = \begin{pmatrix} \gamma & \mathbf{a}^T \\ \mathbf{b} & M \end{pmatrix},$$

where $\mathbf{b}$ is a column vector, $\mathbf{a}^T$ a row vector, and M is a 3×3 matrix. From the invariance property of L and its inverse follows

$$\begin{array}{lll} \gamma^2 - \mathbf{a}^2 = \gamma^2 - \mathbf{b}^2 = 1, & M\mathbf{a} = \gamma\mathbf{b}, & \mathbf{b}^T M = \gamma \mathbf{a}^T, \\ M^T M = \mathbf{a}\mathbf{a}^T + 1, & MM^T = \mathbf{b}\mathbf{b}^T + 1. & \end{array}$$

Pure rotations form the subgroup SO(3), pure velocity transformations do not. The matrix form of the latter ones can be derived from the well-known form of a Lorentz transformation with velocity v in a certain coordinate direction, say x:

$$L^i{}_k = \begin{pmatrix} \gamma & \gamma v & 0 & 0 \\ \gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{10}$$

with the usual $\gamma = (1-v^2)^{-1/2}$. Here time transforms as $t \rightarrow \gamma t + \gamma v x$, in the case of a transformation by an arbitrary velocity $\vec{v}$ it is almost obvious that this transformation generalises to $t \rightarrow \gamma t + \gamma \vec{v}\vec{x}$, this determines the first row and the first column of the general Lorentz matrix. For the generalization of the spatial submatrix we observe that in (10) the coordinate in the direction of motion is multiplied by γ, whereas the coordinates orthogonal to it are unchanged. In the general case we decompose any coordinate vector $\vec{x}$ into a component along the velocity vector and an orthogonal component and multiply the former one by γ,

$$\vec{x} \rightarrow \gamma \frac{\vec{x}\vec{v}}{|\vec{v}|^2}\vec{v} + \vec{x} - \frac{\vec{x}\vec{v}}{|\vec{v}|^2}\vec{v} + \gamma\vec{v}t.$$

In consequence, the matrix of an arbitrary velocity transformation has the form

$$L^i{}_k = \begin{pmatrix} \gamma & \gamma \mathbf{v}^T \\ \gamma \mathbf{v} & 1 + \frac{\gamma^2}{1+\gamma}\mathbf{v}\mathbf{v}^T \end{pmatrix}.$$

Addition of velocities: Consider a particle moving at a speed $\bar{\mathbf{w}}$ in an inertial system $\bar{I}$. What is its speed in an inertial system I, moving at a speed $\mathbf{v}$

w. r. to $\bar{I}$? The coordinates $\mathbf{x}$ and t in I are related to $\bar{\mathbf{x}}$ and $\bar{t}$ in $\bar{I}$ in the following way

$$\mathbf{x} = \bar{\mathbf{x}} + \frac{\gamma^2}{\gamma+1}(\bar{\mathbf{x}}\mathbf{v})\mathbf{v} + \gamma\mathbf{v}\bar{t}, \tag{11}$$

$$t = \gamma\bar{t} + \gamma(\mathbf{v}\bar{\mathbf{x}}). \tag{12}$$

Inserting $\bar{\mathbf{x}} = \bar{\mathbf{w}}\bar{t}$, we obtain for $\mathbf{u} = \mathbf{x}/t =: \bar{\mathbf{w}} \circ \mathbf{v}$

$$\mathbf{u} = \frac{\frac{\bar{\mathbf{w}}}{\gamma} + \frac{\gamma}{\gamma+1}(\mathbf{v}\bar{\mathbf{w}})\mathbf{v} + \mathbf{v}}{1 + \mathbf{v}\bar{\mathbf{w}}}.$$

This is the general form of the addition theorem of 3-velocities, the generalisation of the well-known formula $u = \frac{v+w}{1+vw}$.

Now we can analyse the composition of two velocity transformations: The product of two matrices $L_{\mathbf{v}}$ and $L_{\bar{\mathbf{w}}}$ is not of the symmetric form of a pure velocity transformation, so it obviously contains some rotation. To calculate the latter one, we carry out a transformation by $-\mathbf{u} = -\bar{\mathbf{w}} \circ \mathbf{v}$ on the combined Lorentz transformation and find a rotation matrix:

$$R(\alpha) = L_{\bar{\mathbf{w}}}\, L_{\mathbf{v}}\, L_{-\mathbf{u}},$$

according to a rotation angle α

$$\cos\alpha = \frac{(1+\gamma_{\mathbf{u}}+\gamma_{\mathbf{v}}+\gamma_{\bar{\mathbf{w}}})^2}{(1+\gamma_{\mathbf{u}})(1+\gamma_{\mathbf{v}})(1+\gamma_{\bar{\mathbf{w}}})} - 1 = \frac{\gamma_{\mathbf{v}}+\gamma_{\bar{\mathbf{w}}}}{1+\gamma_{\mathbf{v}}\gamma_{\bar{\mathbf{w}}}}.$$

This is the **Thomas rotation**, arising when two non-parallel pure velocity transformations are composed.
Exercise: Decompose the product $L_w L_v$ of two velocity transformations with $\mathbf{w} = (w, 0, 0)$ and $\mathbf{v} = (0, v, 0)$ into a product $L_R L_u$ of a rotation and a pure velocity transformation. Calculate $\mathbf{u}$ and the rotation angle.

Components of the Lorentz group manifold: The transformations with determinant equal to $+1$ and ${L^0}_0 > +1$ form the subgroup $\mathcal{L}_+^\uparrow$, the group of proper, orthochronous Lorentz transformations, without space or time reflexion. The cosets $\mathcal{L}_+^\downarrow = T\mathcal{L}_+^\uparrow$ (det$= -1$, ${L^0}_0 < -1$, $\mathcal{L}_-^\uparrow = P\mathcal{L}_+^\uparrow$ (det$= -1$, ${L^0}_0 > 1$) and $\mathcal{L}_-^\downarrow = PT\mathcal{L}_+^\uparrow$ (det$= +1$, ${L^0}_0 < -1$), constructed by application of space and time reflexion P and T, are all disjoint, the Lorentz group manifold has four non-connected components. Furthermore, the absolute value of the velocity being strictly less than one (i. e. $|v_i| < 1$, $i = 1, 2, 3$), the Lorentz group is an example of a non-compact Lie group.

3 Representations

Under a representation of a group we understand its action on a linear space. The examples considered so far are matrix representations of SO(3) in 3-dimensional euclidian space and of the Lorentz group SO(3,1) in Minkowski space. But, beside on vectors, Lorentz transformations can be applied to other objects, like tensors and spinors. To make notations more precise, consider the elements of a group G as points on the group manifold, with a "multiplication rule" $g_1 \circ g_2 = g_3$. A (linear) representation is an association of a transformation $L(g)$ on a certain space, the representation space, to each $g \in G$, which reproduces the abstract calculation rules of the group.

$$L(g_1 \circ g_2) = L(g_1)\, L(g_2), \qquad L(g^{-1}) = L(g)^{-1}, \qquad L(e) = \mathrm{id}.$$

Examples: (1) **Contragredient** representation: Consider a matrix L acting on the components of a vector with respect to a basis $\{e_i\}$ in the form

$$\bar{v}^i = L^i{}_k\, v^k.$$

When $\bar{v}^k$ are to be the components of the original vector in a different basis $\{\bar{e}_i\}$ (passive transformation), we may write

$$\mathbf{v} = v^i\, e_i = \bar{v}^i\, \bar{e}_i = L^i{}_k\, v^k\, \bar{e}_i$$

and express the "old" basis in terms of the "new" one,

$$e_k = L^i{}_k\, \bar{e}_i.$$

Assuming a basis transformation by a matrix $\tilde{L}$,

$$\bar{e}_i = \tilde{L}_i^k\, e_k,$$

we find

$$L^i{}_k\, \tilde{L}_i^j = \delta_k^j,$$

in matrix notation

$$\tilde{L} = \left(L^{-1}\right)^T.$$

This is called the contragredient representation to the representation by the matrices L.

Under a change of bases, $\{e_i\} \to \{\bar{e}_i = S_i{}^k e_k\}$ the transformation matrix L goes over into $\tilde{L} = S\, L\, S^{-1}$. Matrix representations related in this form are called **equivalent**.

(2) **Tensor representations, Kronecker product**. Take two representations $g \to T'_g$ and $g \to T''_g$ in two vector spaces $\mathbf{V}'$ and $\mathbf{V}''$. For the Kronecker product acting on $\mathbf{V}' \otimes \mathbf{V}''$ we assume that tensor components T^{ik} transform like products of vector components $v'^i v''^k$, where $v' \in \mathbf{V}'$ and $v'' \in \mathbf{V}''$,

$$v'^{\bar{\imath}} v''^{\bar{\alpha}} = T'^{\bar{\imath}}_{g\ k} T''^{\bar{\alpha}}_{g\ \beta} v'^k v''^\beta = (T'_g \otimes T''_g)^{\bar{\imath}\bar{\alpha}}{}_{k\beta} v'^k v''^\beta .$$

(3) **Direct sum** arrange the components v'^i and v''^α to column vectors which transform according to

$$\begin{pmatrix} v'^{\bar{\imath}} \\ v''^{\bar{\alpha}} \end{pmatrix} = \begin{pmatrix} T'^{\bar{\imath}}_{g\ k} & 0 \\ 0 & T''^{\bar{\alpha}}_{g\ \beta} \end{pmatrix} \begin{pmatrix} v^k \\ v^\beta \end{pmatrix} .$$

The block matrix is written as $T'_g \oplus T''_g$.

3.1 Reducible and irreducible representations

In example (3) the representation acts in both vector spaces $\mathbf{V}'$ and $\mathbf{V}''$ independently, each one is invariant under the action of the transformations. A representation is called **reducible**, if there are invariant subspaces of the representation space, otherwise it is called **irreducible**. Note that not every reducible representation splits into a direct sum of representations, when it does, it is called fully reducible and its matrix has block diagonal form. The matrix of a not fully reducible representation has the typical form

$$\begin{pmatrix} T'_g & A_g \\ 0 & T''_g \end{pmatrix} .$$

Obviously the vectors $(v', 0)$ are transformed among themselves, i.e. $\mathbf{V}'$ is an invariant subspace, but not so $\mathbf{V}''$.

A common task is the reduction of the Kronecker product of two representations, the result is the **Clebsch-Gordan series**.
Example: Transformation of the electromagnetic field tensor. We may arrange the components as a 6-vector $(\vec{E}, \vec{B})$. Under spatial rotations $\vec{E}$ and $\vec{B}$ transform separately, SO(3) is represented as a direct sum. Boosts, on the other hand, do not leave the spaces spanned by $(\vec{E}, \vec{0})$ and $(\vec{0}, \vec{B})$ invariant.

3.2 Schur's lemma

In many cases, particularly in higher dimensions, an index-free notation is more convenient. In this notation a representation $g \to T_g$ of a group $G =$

$\{g, \ldots\}$ in a vector space $\mathbf{V}$ over the *complex* numbers, is a linear, nonsingular mapping of $\mathbf{V}$ onto itself, characterized by

$$T_g(\alpha v + \beta w) = \alpha T_g v + \beta T_g w, \qquad v, w \in \mathbf{V}, \qquad \alpha, \beta \in \mathbf{C},$$

with the representation property

$$T_{g_1} T_{g_2} = T_{g_1 g_2}, \qquad\qquad T_e = \mathrm{id}_{\mathbf{V}}.$$

Reducibility means the existence of a nontrivial linear subspace $\mathbf{V}_1 \subset \mathbf{V}$, so that $T_g \mathbf{V}_1 \subset \mathbf{V}_1$. Two representations $g \to T_g$ and $g \to T'_g$ in $\mathbf{V}$ and $\mathbf{V}'$ are **equivalent**, $T_g \simeq T'_g$, if there exists a one-to-one linear mapping $S : \mathbf{V} \to \mathbf{V}'$, such that

$$T_g = S^{-1} T'_g S \quad \forall g \in G.$$

The relation $ST_g = T'_g S$ can be visualized in a commutative diagram.

If S is a not necessarily one-to-one mapping $\mathbf{V} \to \mathbf{V}'$ with $T'_g S = ST_g$, called an intertwiner, the image $S\mathbf{V} = \mathbf{V}'_1 \subseteq \mathbf{V}'$ is a linear subspace of $\mathbf{V}'$, which is invariant under T'_g;

$$T'_g \mathbf{V}'_1 = T'_g S \mathbf{V} = S T_g \mathbf{V} = S\mathbf{V} = \mathbf{V}'_1.$$

Also the set $\mathbf{V}_0 \subset \mathbf{V}$ of vectors, which are mapped to the zero vector by S is an invariant subspace of $\mathbf{V}$, because from $\{0'\} = T'_g S \mathbf{V}_0 = S T_g \mathbf{V}_0$ follows $T_g \mathbf{V}_0 \subset \mathbf{V}_0$.

From this we deduce a theorem, the first part of Schur's lemma.
For two irreducible representations $g \to T_g$ and $g \to T'_g$ in the spaces $\mathbf{V}$ and $\mathbf{V}'$ and S being a linear mapping $\mathbf{V} \to \mathbf{V}'$ with $ST_g = T'_g S$, either S vanishes identically or it is one-to-one and the representations are equivalent.
When T_g and T'_g are supposed to be irreducible, $\mathbf{V}'_1$ or $\mathbf{V}_0$ must coincide with $\{0'\}$ or $\mathbf{V}'$, or with $\{0\}$ or $\mathbf{V}$, respectively. $\mathbf{V}_0 = \mathbf{V}$ or $\mathbf{V}' = \{0'\}$ means $S \equiv 0$, $\mathbf{V}_0 = \{0\}$ and $\mathbf{V}'_1 = \mathbf{V}'$ implies invertibility of S and equivalence of $g \to T_g$ and $g \to T'_g$.

The second part of Schur's lemma is
If $g \to T_g$ is a representation in $\mathbf{V}$ and $S : \mathbf{V} \to \mathbf{V}$ is a linear mapping that commutes with all T_g, then S is either a multiple of the identity mapping or T_g is reducible.
Proof: Consider the linear subspace $\mathbf{V}_s \subset \mathbf{V}$ of eigenvectors v of S with eigenvalue s, ($Sv = sv$). Due to $ST_g = T_g S$ we have $ST_g v = T_g S v = sT_g v \in \mathbf{V}_s$, and $\mathbf{V}_s$ is invariant under all T_g. As an invariant subspace $\mathbf{V}_s$ must coincide with $\mathbf{V}$ if T_g is irreducible; $Sv = sv \; \forall v \in \mathbf{V}$ means $S = s \cdot \mathrm{id}_{\mathbf{V}}$. Note that here we need a vector space over the complex numbers, in the reals the eigenvalue s need not exist!

4 Lie algebras

4.1 Infinitesimal rotations

Consider group elements close to the unit element. Finite elements can be constructed by composition of infinitesimal ones. To do this, we make use of the fact that we can compose group elements and that we can do calculus on a Lie group.

We write $R = 1 + \Omega$ for a small rotation, where the elements of the matrix Ω are small of first order, i. e. higher orders are negligible. From orthogonality $R\,R^T = 1$ we obtain $\Omega + \Omega^T = 0$, Ω is antisymmetric. Such a matrix can be written in the form $\Omega = \vec{\alpha}\vec{\Lambda}$ with the vector $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ pointing into the direction of the rotation axis and a length equal to the rotation angle, and $\Lambda = (\Lambda_1, \Lambda_2, \Lambda_3)$ being a formal 3-vector of the matrices

$$\Lambda_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Lambda_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad \Lambda_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In this basis of the linear space of antisymmetric 3×3 matrices we have for the $\mu\nu$ element of the matrix Λ_λ

$$\Lambda_{\lambda\mu\nu} = (\Lambda_\lambda)_{\mu\nu} = -\varepsilon_{\lambda\mu\nu}.$$

In first order the transformation equation $\vec{x}' = R\vec{x}$ goes over to $x'^\mu = x^\mu - \varepsilon^{\lambda\mu\nu}\,\alpha^\lambda\,x^\nu$ or $\vec{x}\,' = \vec{x} + \vec{\alpha} \times \vec{x}$.

To establish the relation between infinitesimal and finite transformations we write a finite rotation $R(\vec{\alpha})$ as

$$R(\vec{\alpha}) = R\left(\frac{\vec{\alpha}}{N}\right) R\left(\frac{\vec{\alpha}}{N}\right) = \ldots = \left[R\left(\frac{\vec{\alpha}}{N}\right)\right]^N,$$

for N large enough we can write $R(\vec{\alpha}/N) \approx 1 + \vec{\alpha}\vec{\Lambda}/N$ and in the limit $N \to \infty$

$$R(\vec{\alpha}) = e^{(\vec{\alpha}\vec{\Lambda})} = \left(e^{\vec{\alpha}\vec{\Lambda}/N}\right)^N.$$

Example: Calculate the matrices $e^{\alpha\Lambda_1}$, $e^{\beta\Lambda_2}$, and $e^{\gamma\Lambda_3}$.

Rotations $R(\tau\vec{\alpha}_0) = \exp(\tau\vec{\alpha}_0\vec{\Lambda})$ for fixed $\vec{\alpha}_0$ and variable τ form a one-parameter subgroup; for $\tau = 0$ we get the unity, for $\tau = 1$ the rotation matrix $\exp(\vec{\alpha}_0\vec{\Lambda})$. Each matrix of the form $\vec{\alpha}_0\vec{\Lambda}$ is the generator of such a one-parameter subgroup, with sums and real multiples of generators being generators as well. Thus the generators form a 3-dimensional vector space with Λ_1, Λ_2, Λ_3 providing a basis. (They are the generators of the rotations around the coordinate axes.)

Matrix multiplication leads out of this vector space, because the product of two antisymmetric matrices is not antisymmetric in general. The commutator, however,

$$[A, B] = AB - BA = -[B, A]$$

is antisymmetric, because

$$[A, B]^T = [B^T, A^T] = -[A^T, B^T] = -[A, B].$$

The commutator of two generators is a generator again and may be written in the form $\vec{\alpha}\vec{\Lambda}$. For $A = \vec{m}\vec{\Lambda}$ and $B = \vec{n}\vec{\Lambda}$ we get

$$[\vec{m}\vec{\Lambda}, \vec{n}\vec{\Lambda}] = (\vec{m} \times \vec{n})\vec{\Lambda},$$

or, for the above basis,

$$[\Lambda_\mu, \Lambda_\nu] = \varepsilon_{\mu\nu\lambda}\Lambda_\lambda.$$

With $A \circ B = [A, B]$ we have a product on the vector space of generators. Due to the antisymmetry and the Jacobi identities for commutators,

$$[[A, B], C] + [[C, A], B] + [[B, C], A] = 0,$$

we have the rules

$$A \circ B = -B \circ A, \quad (A \circ B) \circ C + (C \circ A) \circ B + (B \circ C) \circ A = 0.$$

A vector space with a product satisfying these rules is called a Lie algebra. In general, for a basis $\{X_A\}$ the relation

$$X_A \circ X_B = C^D{}_{AB}\, X_D$$

defines the algebra. The structure constants $C^D{}_{AB}$ (structure tensor) must satisfy the rules

$$C^D{}_{AB} = -C^D{}_{BA}, \qquad C^D{}_{AB}\, C^E{}_{CD} + C^D{}_{CA}\, C^E{}_{BD} + C^D{}_{BC}\, C^E{}_{AD} = 0.$$

The generators of the defining representation of SO(3) form a three-dimensional Lie algebra over the reals with structure tensor $\varepsilon_{\lambda\mu\nu}$.

4.2 Lie algebra and representations of SO(3)

Consider the orthogonal transformations of R^3 as special representation of the Lie group $SO(3)$. It is obviously irreducible, because there are no invariant subspaces.

Assume a representation $g \to T_g$ of $SO(3)$ in a vector space $\mathbf{V}$. To a one-dimensional subgroup $g(\tau)$ with $g(0) = e$ a matrix group $T_{g(\tau)}$ is associated. For small τ we have

$$T_{g(\tau)} \approx id_{\mathbf{V}} + \tau t,$$

where

$$t = \frac{\partial}{\partial \tau} T_{g(\tau)} \Big|_{\tau=0}$$

is the generator of the considered subgroup. We will show that the generators of such one-dimensional subgroups form a vector space and find a basis of three generators t_μ, generating the rotations around the coordinate axes and satisfying the commutation relations

$$[t_\mu, t_\nu] = \varepsilon_{\mu\nu\lambda}\, t_\lambda.$$

The generators form also a Lie algebra, the structure of which is isomorphic to the already known one (with the exception of the trivial representation, where $t_\mu = 0$).

Thus the problem, how to find all irreducible representations of $SO(3)$, is reduced to the determination of the representations of the three generators of the Lie algebra.

To prove this claim we consider a rotation, given by the matrix $R(\vec{\alpha})$ in the basis e_μ. In relation to the basis $\bar{e}_\mu$, which originates from e_μ by the rotation S, $\bar{e}_\mu = S_{\mu\nu}\, e_\nu$, the considered rotation is given by the matrix $S\, R(\vec{\alpha})\, S^{-1}$. As $\vec{\alpha}$ becomes $\vec{\alpha}' = S\vec{\alpha}$ in the new basis,

$$S\, R(\vec{\alpha})\, S^{-1} = R(S\vec{\alpha})$$

must hold. With $R(\vec{\alpha}) =: S_g(\vec{\alpha})$ and $S =: S_h$ considered as representation matrices of the abstract group elements $g(\vec{\alpha})$ and h in the defining representation, we have

$$h\, g(\vec{\alpha})\, h^{-1} = g(S_h\vec{\alpha}).$$

For an arbitrary representation in a vector space $\mathbf{V}$ this means

$$T_h\, T_{g(\vec{\alpha})}\, T_{h^{-1}} = T_{g(S_h\vec{\alpha})}. \tag{13}$$

Replacing $\vec{\alpha}$ by $\tau\vec{\alpha}$ and considering small τ, we obtain

$$T_{g(\tau\vec{\alpha})} \approx id_{\mathbf{V}} + \tau t, \qquad t := \frac{\partial}{\partial \tau} T_{g(\tau\vec{\alpha})} \Big|_{\tau=0}.$$

With

$$t_\mu := \frac{\partial}{\partial \alpha_\mu} T_{g(\vec{\alpha})}\bigg|_{\alpha=0} \qquad \text{and} \qquad \vec{t} := (t_1, t_2, t_3)$$

we have

$$t = \alpha_\mu t_\mu = \vec{\alpha}\,\vec{t}.$$

The generators t_μ of rotations around the axes form indeed a vector space spanned by t_1, t_2, and t_3.

Inserting the expansion of $T_{g(\tau\vec{\alpha})}$ into (**??**) and replacing $\vec{\alpha}$ by $\tau\vec{\alpha}$, $\tau \ll 1$ yields

$$T_h\,\vec{\alpha}\vec{t}\,T_h^{-1} = (S_h\vec{\alpha})\vec{t}.$$

Now we assume also h close to the unit element, i. e. in the form $h(\tau\vec{\beta})$ with $\tau \ll 1$, so that

$$T_h \approx id_{\mathbf{V}} + \tau\vec{\beta}\vec{t}, \qquad T_h^{-1} \approx id_{\mathbf{V}} - \tau\vec{\beta}\vec{t}$$

and, according to the behaviour of a vector under infinitesimal rotations,

$$S_h\,\vec{\alpha} = \vec{\alpha} + \tau\vec{\beta} \times \vec{\alpha}.$$

Inserting this we get

$$[\vec{\beta}\vec{t}, \vec{\alpha}\vec{t}] = (\vec{\beta} \times \vec{\alpha})\vec{t}$$

or

$$[t_\mu, t_\nu] = \varepsilon_{\mu\nu\lambda} t_\lambda.$$

These commutation relations hold for any representation, this proves the above claim.

The key relation can be written in a different form. $\vec{\alpha}$ being arbitrary, we can specialize to

$$T_h\,t_\mu\,T_h^{-1} = (S_h)_{\nu\mu}\,t_\nu,$$

or, with $h \to h^{-1}$ and $\left(S_h^{-1}\right)_{\nu\mu} = (S_h)_{\mu\nu}$, to

$$T_h^{-1}\,\vec{t}\,T_h = S_h\,\vec{t}.$$

Generally a triple $\vec{v}$ of operators on V satisfying the relation

$$T_h^{-1}\,\vec{v}\,T_h = S_h\,\vec{v}$$

is called a vector operator. Another relation is

$$[\vec{v}, \vec{\beta}\vec{t}] = \vec{v} \times \vec{\beta}.$$

The square $v^2 = v_\mu v_\mu$ is invariant under the representation, as

$$v^2 = (S_h \vec{v})^2 = T_h^{-1}\, \vec{v}\, T_h\, T_h^{-1}\, \vec{v}\, T_h = T_h^{-1}\, v^2\, T_h,$$

and from this follows

$$T_h\, v^2 = v^2\, T_h.$$

v^2 commutes with all operators T_h of the representation. If the representation is irreducible in V, then v^2 is a multiple of the unit $id_{\mathbf{V}}$ according to Schur's lemma.

For $\vec{v} = \vec{t}$ in particular, we get the Casimir operator

$$C := t^2,$$

which commutes with all representation operators. For $SO(3)$ we have

$$\Lambda_1^2 + \Lambda_2^2 + \Lambda_3^2 = -2 \cdot \mathbf{1}.$$

4.3 Lie algebras of Lie groups

Consider an arbitrary n-dimensional Lie group $\mathcal{G}$ with n parameters β_A and with the elements $g(\beta_A) = g(\beta_1, \ldots, \beta_n)$ and $g(0) = e$. $g(\tau) := g(\beta(\tau))$ defines a curve in the group manifold. For curves through the unit element, $g(\beta_A(0)) = e$, in the vicinity of e the representation $g \to T_g$ is approximated by

$$T_{g(\tau)} \approx id_{\mathbf{V}} + \tau t,$$

where

$$t := \frac{\partial}{\partial \tau} T_{g(\tau)}\bigg|_{\tau=0} = \frac{\partial \beta_A}{\partial \tau}\bigg|_{\tau=0} \frac{\partial}{\partial \beta_A} T_{g(\beta)}\bigg|_{\beta=0}.$$

t is the generator of a one-parameter subgroup in the representation. The finite transformations of this subgroup are given by $\exp(\tau t)$, with the multiplication in the subgroup defined by $\exp(\tau_1 t)\, \exp(\tau_2 t) = \exp[(\tau_1 + \tau_2) t]$. An arbitrary curve $g(\tau)$ with $g(0) = e$ is not a subgroup in general.

If $g(\tau)$, $g_1(\tau)$ are two curves through e, then also the products $g(\tau) g_1(\tau)$ form a curve through e. In the representation T_g we have for infinitesimal τ

$$g(c\tau) g_1(\tau) \to T_{g(c\tau)}\, T_{g_1(\tau)} \approx id_{\mathbf{V}} + \tau(ct + t_1),$$

t_1 is a generator like t. The generators thus form a vector space $L_{\mathbf{V}}$, spanned by

$$t_A := \frac{\partial}{\partial \beta_A} T_{g(\beta)}\bigg|_{\beta=0}, \qquad A = 1, \ldots, n.$$

The products $g(\tau)g_1(\tau)$ constructed from two one-parameter subgroups do not form a subgroup in general, rather only a curve through e. This is reflected by the representation of the two subgroups, generally

$$\exp(\tau t)\exp(\tau t_1) \neq \exp(\tau t + \tau t_1).$$

To show that the generators build up a Lie algebra with the commutator, we consider, in addition to a curve $g(\tau)$ through e the elements $h\, g(\tau)\, h^{-1}$ with an arbitrary $h \in \mathcal{G}$, forming a curve through e, too. For small τ we have

$$T_h\, T_{g(\tau)}\, T_h^{-1} \approx id_{\mathbf{V}} + \tau T_h t T_h^{-1},$$

so, together with $t \in L_{\mathbf{V}}$ also $T_h t T_h^{-1} \in L_{\mathbf{V}}$ is a generator. If we write $h = g_1(\tau)$ and assume a small τ, so that

$$T_{g_1(\tau)} \approx id_{\mathbf{V}} + \tau t_1, \qquad T_{g_1(\tau)}^{-1} \approx id_{\mathbf{V}} - \tau t_1,$$

then the ensuing relation

$$L_{\mathbf{V}} \ni (T_h\, t\, T_h^{-1} - t)/\tau \approx [t_1, t]$$

shows that $L_{\mathbf{V}}$ is indeed a Lie algebra.

The association $t \to T_h\, t\, T_h^{-1}$ is a one-to-one mapping of $L_{\mathbf{V}}$ onto itself, the adjoint action of h on L_V. If the representation is faithful (The relation $g \circ h \leftrightarrow T_g\, T_h$ is an isomorphism), this operator is denoted by Ad_h and $h \to \mathrm{Ad}_h$ is a representation of $\mathcal{G}$ in the n-dimensional space $L_{\mathbf{V}}$, the adjoint representation. For $SO(3)$ it is accidentally equal to the defining one.

With $h = g_1(\tau)$, $\tau \ll 1$ we can consider the derivative of the adjoint map,

$$T_h\, t\, T_h{}^{-1} \approx t + \tau[t_1, t].$$

The derivative

$$\lim_{\tau \to 0} \frac{T_h\, t\, T_h{}^{-1} - t}{\tau} = [t_1, t]$$

is for each $t_1 \in L_V$ a mapping $L_{\mathbf{V}} \to L_{\mathbf{V}}$, $t \to [t_1, t]$, called the adjoint action of t_1 and denoted by

$$\mathrm{ad}_{t_1} t = [t_1, t].$$

It satisfies the Leibniz rule as a consequence of the following commutator relation,

$$\mathrm{ad}_t[t_1, t_2] = [\mathrm{ad}_t t_1, t_2] + [t_1, \mathrm{ad}_t t_2].$$

With the aid of the adjoint representation of a Lie algebra we can construct the symmetric, second-order Killing-Cartan tensor,

$$g_{AB} := \mathrm{Tr}(\mathrm{ad}_{X_A} \mathrm{ad}_{X_B}) = C^C{}_{DA}\, C^D{}_{CB},$$

which may be expressed in terms of the structure constants. Like the latter ones, it is invariant under the adjoint representation. For semisimple groups (each Abelian normal subgroup is discrete - a simple group ia a group, whose only normal subgroups are the trivial group and the group itself) the matrix g_{AB} is invertible and in each representation the Casimir operator

$$C := g^{AB}\, t_A t_B$$

is invariant, i. e. $T_g^{-1} C T_g = C$. With $t^A := g^{AB} t_B$ the further operators

$$\mathrm{Tr}(\mathrm{ad}_{X_A} \mathrm{ad}_{X_B} \ldots \mathrm{ad}_{X_C}) t^A t^B \ldots t^C = C^E{}_{DA}\, C^F{}_{EB} \ldots C^D{}_{GC}\, t^A t^B \ldots t^C$$

are invariant. In irreducible representations they are all multiples of the unit operator.

4.4 Unitary irreducible representations of SO(3)

are formulated in Hilbert spaces, which play a fundamental role in quantum mechanics and quantum field theory. (Conservation of scalar products, norms, probabilities,. . .)

A Hilbert space $\mathbf{H}$ is a (finite or infinite dimensional) vector space over the complex numbers with a scalar product

$$\begin{aligned}
&\langle x, \alpha y_1 + \beta y_2 \rangle = \alpha \langle x, y_1 \rangle + \beta \langle x, y_2 \rangle, \qquad \alpha, \beta \in \mathbf{C} \\
&\langle x, y \rangle = \langle y, x \rangle^* \\
&||x||^2 := \langle x, x \rangle > 0 \quad \forall x \neq 0.
\end{aligned}$$

A representation $g \to T_g$ of a Lie group in a Hilbert space is called unitary, if $\forall g \in \mathcal{G}$, $\forall x, y \in \mathbf{H}$

$$\langle T_g x, T_g y \rangle = \langle x, y \rangle,$$

i. e. when the operators T_g leave the scalar product invariant. Infinitesimal unitary operators are given by

$$T_{g(\tau)} \approx \mathrm{id}_{\mathbf{H}} + \tau t,$$

where from unitarity follows the relation

$$\langle tx, y \rangle + \langle x, ty \rangle = 0.$$

This means the generators are antihermitian operators, whereas $\pm it$ are hermitian,

$$\langle \pm itx, y \rangle = \langle x, \pm ity \rangle.$$

They are called the hermitian generators of the associated one-parameter unitary subgroup of the representation.

The adjoint (hermitian conjugate) operator $A^\dagger$ to an operator A is given by

$$\langle A^\dagger x, y\rangle = \langle x, Ay\rangle.$$

Hermitian operators satisfy $A^\dagger = A$, antihermitian ones $A^\dagger = -A$, and unitary ones $A^\dagger = A^{-1}$. The eigenvalues of hermitian, antihermitian, and unitary operators are real, imaginary, or lie on the unit circle, respectively.

Assume a finite group and a scalar product $\langle\ ,\ \rangle_0$ in the representation space. Then

$$\langle x, y\rangle := \sum_{g\in\mathcal{G}} \langle T_g x, T_g y\rangle_0$$

is an invariant scalar product:

$$\langle T_{g'}x, T_{g'}y\rangle = \sum_g \langle T_g T_{g'} x, T_g T_{g'} y\rangle_0 = \sum_g \langle T_{gg'}x, T_{gg'}y\rangle_0 = \sum_{g''} \langle T_{g''}x, T_{g''}y\rangle_0 = \langle x, y\rangle.$$

For compact Lie groups the sum can be replaced by an integral. From this follows the (specialization of the) theorem (to finite groups) that
Each irreducible representation of a compact Lie group is equivalent to a unitary representation.

According to a further theorem *all the irreducible representations of a compact Lie group in a Hilbert space are finite dimensional, so we can restrict the search for all representations of $SO(3)$ to finite dimensional ones.*

The generators t_μ are antihermitian operators in a finite dimensional Hilbert space, the hermitian generators are

$$J_\mu := it_\mu = J_\mu^\dagger$$

with the commutation relations

$$[J_\mu, J_\nu] = i\varepsilon_{\mu\nu\lambda} J_\lambda$$

and $\forall x \in \mathbf{H}$ and $\mu = 1, 2, 3$

$$\langle x, J_\mu^2 x\rangle = \langle J_\mu x, J_\mu x\rangle \geq 0.$$

According to Schur's lemma in irreducible representations ($\mathbf{H}$ is a complex vector space) the Casimir $\vec{J}^2$ is a multiple of the unit operator,

$$\vec{J}^2 = \lambda\, \mathrm{id}_{\mathbf{H}}, \qquad \lambda \geq 0.$$

The irreducible representations can be found with the aid of the eigenvalue spectrum of one of the hermitian generators, say J_3. It is convenient to pass over to the combinations

$$J_\pm := J_1 \pm i J_2 \quad \text{and} \quad J_3$$

with

$$[J_+, J_-] = 2J_3, \quad [J_3, J_\pm] = \pm J_\pm, \quad \vec{J}^2 = J_\pm J_\mp \mp J_3 + J_3^2.$$

Now consider the eigenvectors of J_3, forming a complete orthogonal system in $\mathbf{H}$. Assume x_m to be a normalized eigenvector to the eigenvalue m,

$$J_3\, x_m = m\, x_m, \qquad ||x_m|| = 1.$$

Then

$$\begin{aligned} &J_3\, J_\pm\, x_m = (m \pm 1)\, J_\pm\, x_m, \\ &\langle J_\pm\, x_m, J_\pm\, x_m \rangle = \langle x_m, J_\mp\, J_\pm\, x_m \rangle = \lambda \mp m - m^2. \end{aligned}$$

From this follows that either $J_\pm\, x_m$ is the zero vector or $m \pm 1$ is an eigenvalue of J_3, too. The representation being finite dimensional, there are only finitely many eigenvalues, the largest of which may be denoted by j. For a normalized eigenvector x_j

$$J_+ x_j = 0 \qquad \text{and} \qquad \lambda = j + j^2$$

must hold. In the series of eigenvectors $J_- x_j$, $J_-^2 x_j$, ... to the eigenvalues j, $j-1$, ... after $N-1$ applications of J_- a least eigenvalue j' must be reached,

$$J_3 (J_-)^{N-1} x_j = j' (J_-)^{N-1} x_j,$$

so that

$$(J_-)^N x_j = 0.$$

From the above relations follows

$$\lambda = j^2 + j = j'^2 - j', \qquad j - j' + 1 = N$$

or $(j + j')(j - j' + 1) = 0$, i. e. $j' = -j$ and thus $2j + 1 = N$, a positive integer. For j and λ the possible values

$$j = 0, \tfrac{1}{2}, 1, \tfrac{3}{2}, 2, \ldots \qquad \lambda = j(j+1) = 0, \tfrac{3}{4}, 2, \tfrac{15}{4}, 6, \ldots$$

arise. The eigenvalues and eigenvectors of J_3 are

$$j, j-1, \ldots, -j+1, -j, \qquad x_j, J_- x_j, \ldots, (J_-)^{2j} x_j.$$

These vectors are orthogonal, i.e. linearly independent, so they span a $2j+1$ dimensional subspace of $\mathbf{H}$, which is invariant under $J_\pm$, J_3, $\vec{J}^2$. In irreducible representations this subspace must be equal to $\mathbf{H}$.

Up to equivalence, an irreducible representation of $SO(3)$ is determined uniquely by the maximal eigenvalue j of the operator J_3 or the eigenvalue $j(j+1)$ of the Casimir operator $\vec{J}^2$, j can assume only the values 0, 1/2, 1, ...

Normalized eigenvectors x_m with $||x_m|| = 1$ are determined up to a phase factor. For them

$$J_3 x_m = m x_m, \qquad\qquad J_\pm x_m = \rho_\pm(m) x_{m\pm 1}$$

with

$$|\rho_\pm(m)|^2 = j(j+1) \mp m - m^2.$$

From $J_\pm^\dagger = J_\mp$ and the orthogonality of x_m follows

$$\rho_\pm(m) = \langle x_{m\pm1}, J_\pm x_m\rangle = \langle J_\mp x_{m\pm1}, x_m\rangle = \langle x_m, J_\mp x_{m\pm1}\rangle^* = \rho_\mp^*(m\pm1).$$

This is compatible with the above relation, so we can choose the phase

$$\rho_\pm(m) = {}_{+}\sqrt{j(j+1) \mp m - m^2},$$

the associated basis $\{x_m\}$ is called a canonical basis. In this basis the operators are represented by the following matrices

$$J_3 = \mathrm{diag}(j, j-1, \ldots, -j+1, -j), \qquad \vec{J}^2 = j(j+1)\,\mathbf{1},$$

$$J_+ = \begin{pmatrix} 0 & \rho_+(j-1) & & & 0 \\ & 0 & \rho_+(j-2) & & \\ & & & \ddots & \\ & & 0 & & \rho_+(-j) \\ & & & \ddots & \\ & 0 & & & 0 \end{pmatrix}$$

$$J_- = \begin{pmatrix} 0 & & & & \\ \rho_-(j) & 0 & & & 0 \\ & \rho_-(j-1) & 0 & & \\ & & \ddots & \ddots & \\ & 0 & \rho_-(-j+1) & 0 & \end{pmatrix}$$

The simplest cases:
$j = 0$: trivial representation
$j = 1$: 3-dimensional representation,

$$J_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{i\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix},$$

$$\vec{J}^2 = 2 \cdot \mathbf{1}.$$

This representation is equivalent to the defining representation, J_3 is the diagonalized form of $i\Lambda_3$.
$j = 2$: 5-dimensional representation, equivalent to the representation in the space of traceless symmetric tensors $T^{\mu\nu}$,
$j = 1/2$: 2-dimensional representation with the generators $\vec{J} = 1/2\,\vec{\sigma}$ (spinor representation), where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli spin matrices with the commutation relations

$$[\sigma_\mu, \sigma_\nu] = 2i\varepsilon_{\mu\nu\lambda}\sigma_\lambda$$

and the anticommutator (Clifford algebra) relation

$$\{\sigma_\mu, \sigma_\nu\} = \sigma_\mu\sigma_\nu + \sigma_\nu\sigma_\mu = 2\delta_{\mu\nu}\mathrm{id}.$$

Together this gives the product

$$\sigma_\mu\sigma_\nu = \delta_{\mu\nu}\mathrm{id} + i\varepsilon_{\mu\nu\lambda}\sigma_\lambda$$

(formally the multiplication rules of Hamilton's quaternions). This representation is also called the fundamental one.

Knowing the representations of the Lie algebra $so(3)$, what remains to do is to construct the representation matrices of finite rotations. To represent a group element $g(\vec{\alpha})$ we consider the one-parameter group $g(\tau\vec{\alpha})$. With $t = \vec{\alpha}\vec{t}$ being the generator of this subgroup the desired representation is

$$g(\vec{\alpha}) \to T_{g(\vec{\alpha})} = \exp(\vec{\alpha}\vec{t}) = \exp(-i\vec{\alpha}\vec{J}).$$

$j = 0$: trivial, $j = 1$: defining representation.
$j = 1/2$:

$$\vec{\alpha}\vec{J} = \frac{1}{2}\vec{\alpha}\vec{\sigma} = \frac{1}{2}\alpha_\mu\sigma_\mu,$$

$$(\vec{\alpha}\vec{\sigma})^2 = \alpha_\mu\sigma_\mu\alpha_\nu\sigma_\nu = \frac{1}{2}\alpha_\mu\alpha_\nu(\sigma_\mu\sigma_\nu + \sigma_\nu\sigma_\mu) = \vec{\alpha}^{\,2}\cdot\mathbf{1}.$$

Writing $\vec{\alpha} = \alpha\vec{n}$ with $\vec{n}^{\,2} = 1$ we get

$$(-i\vec{\alpha}\vec{J})^2 = -\left(\frac{\alpha}{2}\right)^2\cdot\mathbf{1}, \qquad (-i\vec{\alpha}\vec{J})^3 = \left(\frac{\alpha}{2}\right)^3\vec{n}\vec{\sigma},\ \ldots$$

Therefore from the expansion

$$\exp(-i\vec{\alpha}\vec{J}) = \mathbf{1}\cos\frac{\alpha}{2} - i\vec{n}\vec{\sigma}\sin\frac{\alpha}{2} =: U(\vec{\alpha})$$

we find

$$U(\vec{\alpha}) = \begin{pmatrix} \cos\frac{\alpha}{2} - in_3\sin\frac{\alpha}{2} & -i(n_1 - in_2)\sin\frac{\alpha}{2} \\ -i(n_1 + in_2)\sin\frac{\alpha}{2} & \cos\frac{\alpha}{2} + in_3\sin\frac{\alpha}{2} \end{pmatrix}.$$

This matrix is unitary with determinant equal to one
($\det U = \exp[\mathrm{Tr}(-i\vec{\alpha}\vec{\sigma}/2)] = 1$).

The "representation" by $U(\vec{\alpha})$ has the property that, in contrast to the tensor representations (integer j), the composition of two rotations by an angle π does not lead to unity, but to $-\mathrm{id}$. So the matrices $U(\vec{\alpha})$ form a group only when the underlying range of α is extended to $0 \le |\vec{\alpha}| \le 2\pi$, thus covering the set of rotations twice – to each rotation $g(\vec{\alpha})$ two matrices $U(\vec{\alpha})$ and $-U(\vec{\alpha}) = U(\vec{\alpha} + 2\pi\frac{\vec{\alpha}}{\alpha})$ are associated and this is not a representation in the strict sense.

4.5 The group $SU(2)$

For $\vec{\alpha}$ varying over $0 \le |\vec{\alpha}| \le 2\pi$ the matrices $U(\vec{\alpha})$ form the group $SU(2)$ of all unitary, unimodular (det=1) matrices. Namely, for each complex 2×2 matrix

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

unitarity demands $c = -\lambda b^*$, $d = \lambda a^*$, $|a|^2 + |b|^2 = 1$, $|\lambda| = 1$, unimodularity restricts to $\lambda = 1$, so that

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1. \tag{14}$$

From this follows $|a| \le 1$, so that there is exactly one α, $0 \le \alpha \le 2\pi$, for which $\mathrm{Re}\,a = \cos\frac{\alpha}{2}$. A unique $\vec{n}$ arises from $\mathrm{Im}\,a = -n_3\sin\frac{\alpha}{2}$, $\mathrm{Re}\,b = -n_2\sin\frac{\alpha}{2}$, $\mathrm{Im}\,b = -n_1\sin\frac{\alpha}{2}$.

If we interpret the real and imaginary parts of a and b as cartesian coordinates in $\mathbf{R}^4$, we see that the group manifold of $SU(2)$ is the unit sphere S^3

in $\mathbf{R}^4$. $U \in SU(2)$ and $-U$ belonging to one and the same rotation, we can consider $SO(3)$ as S^3 with antipode points identified. After the restriction $0 \leq |\vec{\alpha}| \leq \pi$ the identification can be dropped, except at the boundary sphere $|\vec{\alpha}| = \pi$.

After the identification leading to $SO(3)$ a curve $g(\tau)$ going from $\alpha = 0$ to $\alpha = 2\pi$ becomes closed and not contractible to e, in contrast to curves restricted to the lower half sphere. When this occurs, the manifold is called multiply connected. Here we have two classes of curves, one being continuously contractible to a point, the other not. $SO(3)$ is therefore twofold connected. The transition to $SU(2)$ - which undoes the identification - makes every curve continuously contractible, $SU(2)$ is simply connected, it is called the universal covering of $SO(3)$.

A Lie group and its universal covering have the same Lie algebra and in a sufficiently small neighborhood of the unit they are isomorphic, at a large scale there is a homomorphism $SU(2) \to SO(3)$, where the discrete invariant subgroup $Z_2 = \{1, -1\}$ is mapped to $e \in SO(3)$, $SO(3) \simeq SU(2)/Z_2$.

4.6 Tensor and spinor formulations

To construct a local isomorphism between $SO(3)$ and $SU(2)$, we associate to each 3-vector $\vec{x}$ a traceless hermitian 2×2 matrix $X = \vec{x}\vec{\sigma}$. $\vec{x}$ being real and σ_μ traceless, we have

$$X = X^\dagger \qquad \text{and} \qquad \operatorname{tr} X = 0.$$

Reversely each traceless hermitian matrix can be written in the form $X = \vec{x}\vec{\sigma}$ with real $\vec{x}$. $\vec{x}$ can be retrieved from X by

$$\vec{x} = \frac{1}{2} \operatorname{tr} X\vec{\sigma},$$

because $\operatorname{tr} \sigma_\mu \sigma_\nu = 2\delta_{\mu\nu}$. Further

$$X^2 = \vec{x}^{\,2} \cdot \mathbf{1}, \qquad\qquad \det X = -\vec{x}^{\,2}.$$

Consider the adjoint action of $U \in SU(2)$ on the element X of the Lie algebra $su(2)$ (linear combination of the generators σ_i),

$$X' = UXU^{-1} = UXU^\dagger.$$

The matrix X' is hermitian and traceless again,

$$(X')^\dagger = (UXU^\dagger)^\dagger = UX^\dagger U^\dagger = UXU^\dagger = X',$$

$$\operatorname{tr} X' = \operatorname{tr} UXU^{-1} = \operatorname{tr} U^{-1}UX = \operatorname{tr} X = 0$$

and defines a linear transformation $\vec{x} \to \vec{x}\,' = R\vec{x}$, which is orthogonal due to

$$(\vec{x}\,')^2 \cdot \mathbf{1} = X'^2 = UXU^{-1}UXU^{-1} = UX^2U^{-1} = \vec{x}\,^2 \cdot \mathbf{1}$$

or

$$-(\vec{x}\,')^2 = \det X' = \det U \det X \det U^\dagger = \det X = -\vec{x}\,^2.$$

The corresponding orthogonal matrix of the transformation

$$x'_\mu = R_{\mu\nu}\, x_\nu$$

can be found by comparison of

$$X' = x'_\mu \sigma_\mu = R_{\mu\nu} x_\nu \sigma_\mu$$

with $U x_\nu \sigma_\nu U^\dagger$:

$$R_{\mu\nu}\sigma_\mu = U\sigma_\nu U^\dagger = U\sigma_\nu U^{-1}.$$

Multiplication by σ_ρ from the left and taking the trace gives explicitly

$$R_{\rho\nu} = \frac{1}{2} \operatorname{tr} \sigma_\rho U \sigma_\nu U^\dagger = \frac{1}{2} \operatorname{tr} \sigma_\rho U \sigma_\nu U^{-1}.$$

Also U can be expressed explicitly in terms of R, namely, for each 2×2 matrix the identity

$$\sigma_\mu M \sigma_\mu = 2 \operatorname{tr} M \cdot \mathbf{1} - M$$

holds. By multiplication of $R_{\mu\nu}\sigma_\mu$ by σ_ν we find

$$R_{\mu\nu}\sigma_\mu\sigma_\nu = U\sigma_\nu U^\dagger \sigma_\nu = U(2 \operatorname{tr} U^\dagger \cdot \mathbf{1} - U^\dagger) = (2 \operatorname{tr} U)U - \mathbf{1}.$$

Taking the trace gives

$$2(\operatorname{tr} U)^2 = 2(1 + \operatorname{tr} R),$$

hence

$$U = \pm \frac{\mathbf{1} + R_{\mu\nu}\sigma_\mu\sigma_\nu}{2\sqrt{1 + \operatorname{tr} R}}.$$

These formulae show the local equivalence of the adjoint representation of $SU(2)$ with the defining representation of $SO(3)$.

A spinor u transforms under a rotation according to

$$u \to u' = U(\vec{\alpha})\, u.$$

The scalar product in the sense of the unitary geometry, invariant under this transformation, is

$$\langle u, v \rangle = u_1^* v_1 + u_2^* v_2,$$

where $u = (u_1, u_2)$ in the canonical basis. In analogy to the ε tensor there is an ε spinor, defining an invariant bilinear form

$$\varepsilon^{AB} u_A v_B = u_1 v_2 - u_2 v_1.$$

In contrast, the scalar product is sesquilinear.

Like tensor representations in arbitrary vector spaces, we may investigate higher-order spinors and their transformations. To this end we consider Kronecker products $U(\vec{\alpha}) \otimes U(\vec{\alpha}) \otimes \ldots$ and their reduction.

A simple example is the reduction of the representation $g(\vec{\alpha}) \to U(\vec{\alpha}) \otimes U(\vec{\alpha})$. With $u = (u_1, u_2)$ and $v = (v_1, v_2)$ the components of $u \otimes v$ are

$$(u_1 v_1, u_1 v_2, u_2 v_1, u_2 v_2).$$

If further $u' = Uu$ and $v' = Uv$, then, with U given by (**??**).

$$\begin{pmatrix} u'_1 v'_1 \\ u'_1 v'_2 \\ u'_2 v'_1 \\ u'_2 v'_2 \end{pmatrix} = \begin{pmatrix} a^2 & ab & ab & b^2 \\ -ab^* & |a|^2 & -|b|^2 & a^* b \\ -ab^* & -|b|^2 & |a|^2 & a^* b \\ b^{*2} & -a^* b^* & -a^* b^* & a^{*2} \end{pmatrix} \begin{pmatrix} u_1 v_1 \\ u_1 v_2 \\ u_2 v_1 \\ u_2 v_2 \end{pmatrix}. \tag{15}$$

From this we may read off that for the antisymmetric part

$$u'_1 v'_2 - u'_2 v'_1 = \left(|a|^2 + |b|^2\right)(u_1 v_2 - u_2 v_1) = u_1 v_2 - u_2 v_1,$$

it transforms according to the trivial representation. In the subspace of the symmetric second-order spinors $u_{(A} v_{B)}$ we choose the basis $(u_1 v_1, (u_1 v_2 + u_2 v_1)/\sqrt{2}, u_2 v_2)$, then (**??**) becomes

$$\begin{pmatrix} (u'_1 v'_2 - u'_2 v'_1)/\sqrt{2} \\ u'_1 v'_1 \\ (u'_1 v'_2 + u'_2 v'_1)/\sqrt{2} \\ u'_2 v'_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & a^2 & \sqrt{2}\, ab & b^2 \\ 0 & -\sqrt{2}\, ab^* & |a|^2 - |b|^2 & \sqrt{2}\, a^* b \\ 0 & b^{*2} & -\sqrt{2}\, a^* b^* & a^{*2} \end{pmatrix} \begin{pmatrix} (u_1 v_2 - u_2 v_1)/\sqrt{2} \\ u_1 v_1 \\ (u_1 v_2 + u_2 v_1)/\sqrt{2} \\ u_2 v_2 \end{pmatrix}.$$

This is already the complete reduction, for it is easy to see that for infinitesimal rotations around the 3-axis ($b = 0$, $a \approx 1 - i\alpha/2$) the generator J_3 of the arising 3-dimensional representation acquires the form $\mathrm{diag}(1, 0, -1)$, which characterizes the irreducible representation of weight 1. Moreover, the form of the generators $J_\pm$ shows that the chosen basis is a canonical one.

In the same way one can form the symmetric part $u_{(A} u_B \ldots u_{C)}$ of a higher-order spinor. To see that the space of symmetric spinors of a certain order p is irreducible, we first find out its dimension by counting the independent components of a spinor. We choose a basis in which the first p_1

indices are equal to 1 – the remaining $p_2 = p - p_1$ are then equal to 2. As p_1 can be $0, 1, \dots p$, there are $p+1$ independent components, the space is $p+1$ dimensional. Now we investigate the eigenvalues of J_3. In the space of all spinors of order p an infinitesimal rotation around the 3-axis has the form

$$
\begin{aligned}
&U(\tau\vec{e}_3) \otimes \dots \otimes U(\tau\vec{e}_3) \approx \\
&\mathbf{1} \otimes \dots \otimes \mathbf{1} - \frac{i\tau}{2}(\sigma_3 \otimes \mathbf{1} \otimes \dots \otimes \mathbf{1} + \ \dots \ + \mathbf{1} \otimes \dots \otimes \mathbf{1} \otimes \sigma_3), \\
&J_3 = \frac{1}{2}(\sigma_3 \otimes \dots \otimes \mathbf{1} + \ \dots \ + \mathbf{1} \otimes \dots \otimes \sigma_3).
\end{aligned} \tag{16}
$$

If $u^\pm$ is an eigenspinor of J_3 to the eigenvalue $\pm 1/2$, then $u^\pm \otimes \dots \otimes u^\pm$ belongs to the subspace of totally symmetric spinors of order p and is an eigenspinor to the eigenvalue $\pm p/2$. From the known spectrum of J_3 follows that also the eigenvalues $p/2-1, \dots, -p/2+1$ and the associated eigenspinors must occur. (It is easy to see that $u^+_{(A} u^+_B \dots u^-_{C)}$ with p_1 factors u^+ and p_2 factors u^- is an eigenspinor to the eigenvalue $(p_1 - p_2)/2$.) It follows that the dimension of the representation in the space of totally symmetric spinors must be at least equal to $2(p/2) + 1 = p + 1$. Indeed it is an irreducible representation with the weight $j = p/2$.

Now we can construct the explicit form of the representation matrix of finite rotations for each weight j. Symmetric spinors of order p ($p = 2j$) transform like $u_A u_B \dots u_C$, the independent components are

$$
(u_1)^p, \ (u_1)^{p-1} u_2, \ \dots, \ u_1 (u_2)^{p-1}, \ (u_2)^p. \tag{17}
$$

A rotation transforms them to the expressions $(u'_1)^p, \ \dots, \ (u'_2)^p$, where

$$
u'_1 = a u_1 + b u_2, \qquad u'_2 = -b^* u_1 + a^* u_2 \tag{18}
$$

and simple multiplication yields the elements of the representation matrix. To obtain it in unitary form, we must introduce an invariant scalar product and a norm and normalize the above expressions in such a way that the square of the norm of an element of $u_A u_B \dots u_C$ appears as a sum of absolute squares of the monomials in (**??**) (with normalizing denominators N_i). Unitarity means conservation of the norm

$$
\left| \frac{(u_1)^p}{N_1} \right|^2 + \left| \frac{(u_1)^{p-1} u_2}{N_2} \right|^2 + \dots = \left| \frac{(u'_1)^p}{N_1} \right|^2 + \left| \frac{(u'_1)^{p-1} u'_2}{N_2} \right|^2 + \dots \tag{19}
$$

The scalar product $\langle u, u \rangle = u^*_A u_A$ is invariant, and so is

$$
\frac{1}{p!} u^*_A u^*_B \dots u^*_C u_A u_B \dots u_C = \frac{1}{p!} (u^*_A u_A)^p = \frac{1}{p!} \langle u, u \rangle^p \tag{20}
$$

and gives a suitable square of a norm in the space of symmetric spinors. To express it in the above form as sum of absolute squares of monomials we make use of the binomial theorem ($p_2 = p - p_1$).

$$\frac{1}{p!}(u_1^* u_1 + u_2^* u_2)^p = \frac{1}{p!}\sum_{p_1=0}^{p} \left(\begin{array}{c} p \\ p_1 \end{array} \right) (u_1^* u_1)^{p_1} (u_2^* u_2)^{p_2} =$$
$$\sum_{p_1} \frac{(u_1^*)^{p_1}(u_2^*)^{p_2}}{\sqrt{p_1! p_2!}} \frac{(u_1)^{p_1}(u_2)^{p_2}}{\sqrt{p_1! p_2!}}. \tag{21}$$

From this we can directly read off the normalization of the monomials, up to phase factors. The matrix elements of the representations of rotations are got by expanding

$$\frac{(u_1')^{p_1}(u_2')^{p_2}}{\sqrt{p_1! p_2!}} = \frac{(a u_1 + b u_2)^{p_1}(-b^* u_1 + a^* u_2)^{p_2}}{\sqrt{p_1! p_2!}} =$$
$$\frac{1}{\sqrt{p_1! p_2!}} \sum_q \binom{p_1}{q} (a u_1)^q (b u_2)^{p_1 - q} \sum_\ell \binom{p_2}{\ell} (-b^* u_1)^\ell (a^* u_2)^{p_2 - \ell} = \tag{22}$$
$$\sqrt{p_1! p_2!} \sum_q \frac{a^q u_1^q \, b^{p_1 - q} u_2^{p_1 - q}}{q! \, (p_1 - q)!} \sum_\ell \frac{(-b^*)^\ell u_1^\ell \, (a^*)^{p_2 - \ell} u_2^{p_2 - \ell}}{\ell! \, (p_2 - \ell)!}$$

and reading off the coefficients of $(u_1)^{q_1}(u_2)^{q_2}/\sqrt{q_1! q_2!}$.

After reordering according to the canonical basis, $p_1 = j + m$, $p_2 = j - m$, $m = -j, \ldots, j$, we have

$$\sqrt{(j+m)!(j-m)!} \sum_{q,\ell} \frac{a^q \, b^{j+m-q} (-b^*)^\ell (a^*)^{j-m-\ell}}{q!(j+m-q)!\ell!(j-m-\ell)!} \, u_1^{q+\ell} \, u_2^{2j-q-\ell}. \tag{23}$$

In terms of j and m, on the left-hand side of (**??**) we have the element $(u_1')^{j+m}(u_2')^{j-m}$. An analogous enumeration $u_1^{j+n} u_2^{j-n}$ on the right-hand side is achieved by the substitution $q = j + n - \ell$, $n = -j, \ldots, j$, so that

$$\frac{(u_1')^{j+m}(u_2')^{j-m}}{\sqrt{(j+m)!(j-m)!}} = \sqrt{(j+m)!(j-m)!} \times \tag{24}$$
$$\sum_{n,\ell} (-1)^\ell \frac{a^{j+n-\ell} \, (a^*)^{j-m-\ell} \, b^{m-n+\ell} \, (b^*)^\ell}{(j+n-\ell)!(m-n+\ell)!\ell!(j-m-\ell)!} \, u_1^{j+n} \, u_2^{j-n}.$$

From this comparison we can read off the matrix elements

$$D^{(j)}_{mn}(\vec{\alpha}) = \tag{25}$$
$$\sum_\ell (-1)^\ell \frac{\sqrt{(j+m)!(j-m)!(j+n)!(j-n)!}}{(j-m-\ell)!(j+n-\ell)!(m-n+\ell)!\ell!} a^{j+n-\ell} (a^*)^{j-m-\ell} b^{m-n+\ell} (b^*)^\ell.$$

In the sum over the integer $\ell \in N$, $0 \leq \ell \leq j-m$, all the values, which would lead to factorials of negative numbers, have to be omitted. In the canonical basis the transformation is

$$T_g|jm\alpha\rangle = \sum_n D^{(j)}_{nm}(g)|jn\alpha\rangle, \tag{26}$$

where α enumerates multiply occurring m's in reducible representations. By reduction of the Kronecker "powers" of the two dimensional spinor representation we may get all the irreducible representations of $SO(3)$, therefore the denotation "fundamental representation".

4.7 Representations of the Lorentz group

4.7.1 The adjoint action of the Lorentz group

In the framework of the Lorentz group the generators of the subgroup of rotations appear as the three 4 by 4 matrices

$$M_i = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda_i \end{pmatrix}, \qquad i = 1, 2, 3. \tag{27}$$

with Λ_i being the well-known generators of $SO(3)$. In the Lie-Algebra there are three further elements, generating boosts in the direction of the axes,

$$N_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; N_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; N_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \tag{28}$$

For the commutators we find

$$[M_i, M_j] = \epsilon_{ij}{}^k M_k, \quad [N_i, N_j] = -\epsilon_{ij}{}^k M_k, \quad [N_i, M_j] = \epsilon_{ij}{}^k N_k. \tag{29}$$

In the six-dimensional basis $\{N_i, M_i\}$ the adjoint action of the generators (= commutators with the other generators) can be written in form of the following 6 by 6 matrices

$$\mathrm{ad}_{M_i} = \begin{pmatrix} \Lambda_i & 0 \\ 0 & \Lambda_i \end{pmatrix}, \qquad \mathrm{ad}_{N_i} = \begin{pmatrix} 0 & \Lambda_i \\ -\Lambda_i & 0 \end{pmatrix}. \tag{30}$$

For the Cartan-Killing metric $g_{ij} = \mathrm{Tr}(\mathrm{ad}_{t_i}\,\mathrm{ad}_{t_j})$ with $t_i \equiv N_i$ for $i = 1, 2, 3$ and $t_i = M_i$ for $i = 4, 5, 6$ we find

$$g_{ij} = 4\begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \tag{31}$$

where $\mathbf{1}$ is the 3 by 3 unit matrix. Accordingly there is a Casimir operator $C_1 = \vec{N}^{\,2} - \vec{M}^{\,2}$. In the four-dimensional representation C_1 is a multiple of the unit matrix, $C_1 = 3\cdot\mathbf{1}$, in the adjoint representation it is $4\cdot\mathbf{1}$. Beside this, there is a further matrix commuting with all the matrices in (**??**), namely

$$C_2 = \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}. \tag{32}$$

Consider a finite element $\exp(\chi N_1)$ of the adjoint representation of the Lorentz group, generated by the boost generator in the x direction N_1

$$\exp(\chi N_1) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \mathrm{ch}\chi & 0 & 0 & 0 & -\mathrm{sh}\chi \\ 0 & 0 & \mathrm{ch}\chi & 0 & \mathrm{sh}\chi & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & \mathrm{sh}\chi & 0 & \mathrm{ch}\chi & 0 \\ 0 & -\mathrm{sh}\chi & 0 & 0 & 0 & \mathrm{ch}\chi \end{pmatrix}. \tag{33}$$

If we write as usual $\mathrm{ch}\chi = \gamma$ and $\mathrm{sh}\chi = v\gamma$ and apply this matrix to a 6-vector $(\vec{E}, \vec{B})$, formed by a combination of the electric and magnetic field, we find the usual Lorentz transformation formulae for the fields:

$$\begin{pmatrix} \vec{E} \\ \vec{B} \end{pmatrix} \rightarrow \begin{pmatrix} E_x \\ \gamma E_y - v\gamma B_z \\ \gamma E_z + v\gamma B_y \\ B_x \\ \gamma B_y + v\gamma E_z \\ \gamma B_z - v\gamma E_y \end{pmatrix}. \tag{34}$$

In the representation space of electromagnetic 6-vectors the second Casimir operator generates a "duality transformation" $(\vec{E}, \vec{B}) \leftrightarrow (-\vec{B}, \vec{E})$. Under this transformations the Maxwell equations of the free electromagnetic field are invariant, the existence of electric charges and non-existence of magnetic ones, however, breaks this duality invariance.

The considered real representation is irreducible, but the spaces of real linear combinations of the complex vectors $(\vec{E} \pm i\vec{B})$ transform separately, and thus span invariant subspaces.

Consider the complex basis

$$\vec{M}^{\pm} = \frac{1}{2}\left(\vec{M} \pm i\vec{N}\right) \tag{35}$$

of the Lorentz Lie algebra with the commutation relations

$$[M^{\pm}_{\mu}, M^{\pm}_{\nu}] = \epsilon_{\mu\nu\lambda} M^{\pm}_{\lambda}, \qquad [M^{+}_{\mu}, M^{-}_{\nu}] = 0. \tag{36}$$

The Lie algebra decomposes into the direct sum of two three-dimensional Lie algebras $L^{\pm}$, $L = L^{+} \oplus L^{-}$. Both L^{+} and L^{-} have the structure of the algebra $so(3)$ of the rotation group.

The Lie algebra of the real Lorentz group is made up of the real superpositions of M and N. The split is possible only by making use of i, i. e. of coefficients that would appear in the Lie algebra of the complex Lorentz group, which is isomorphic to $SO(4, \mathbf{C})$, as the sign does not play a role in the complex group. Locally the latter one is isomorphic to the product of two complex rotation groups $SO(3, \mathbf{C})$.

The real linear combinations of $M^{\pm}$ build up the Lie algebra of $SO(4, \mathbf{R})$, which in turn is locally isomorphic to $SO(3, \mathbf{R}) \times SO(3, \mathbf{R})$. Irreducible representations of the Lorentz group can thus be characterized by first complexifying it, making use of the isomorphism to $SO(4, \mathbf{C})$ and the local product decomposition of the latter one, and finally restricting to $SO(4, \mathbf{R})$ and its decomposition:

$$\begin{aligned} L \to L^{c} &\cong SO(4, \mathbf{C}) \cong SO(3, \mathbf{C}) \times SO(3, \mathbf{C}) \\ &\to SO(4, \mathbf{R}) \cong SO(3, \mathbf{R}) \times SO(3, \mathbf{R}). \end{aligned}$$

Accordingly, the irreducible representations of the Lorentz group by matrices $D^{(j,j')}$ can be classified by two indices j and j' of two copies of $SO(3)$ representation matrices. $D^{(j,j')}$ are $(2j+1)(2j'+1)$ dimensional.

From the Casimir operators

$$(\vec{M}^{\pm})^2 = \vec{M}^2 - \vec{N}^2 \pm 2i\vec{M}\vec{N} \tag{37}$$

of the two rotation groups we can read off the Casimir operators

$$C_1 = \vec{M}^2 - \vec{N}^2 \quad \text{and} \quad C_2 = \vec{M}\vec{N} \tag{38}$$

of the Lorentz group in the representations $D^{(j,j')}$.

In this formalism infinitesimal Lorentz transformations are given by

$$L(\vec{v}, \vec{\alpha}) \approx \mathbf{1} + \vec{\alpha}(\vec{M}^{+} + \vec{M}^{-}) - i\vec{v}(\vec{M}^{+} - \vec{M}^{-}) = \mathbf{1} + (\vec{\alpha} - i\vec{v})\vec{M}^{+} + (\vec{\alpha} + i\vec{v})\vec{M}^{-}, \tag{39}$$

finite transformations by

$$L(\vec{v}, \vec{\alpha}) = D^{(j,j')}(\vec{v}, \vec{\alpha}) = D^{(j)}(\vec{\alpha} - i\vec{v}) \otimes D^{(j')}(\vec{\alpha} + i\vec{v}). \tag{40}$$

Note that

$$L(\vec{v}, \vec{\alpha}) \neq \exp[(\vec{\alpha} - i\vec{v})\vec{M}^{+} + (\vec{\alpha} + i\vec{v})\vec{M}^{-}]\ !$$

The one-parameter subgroup connecting $L(\vec{v}, \vec{\alpha})$ with the unit element e is not given by the curve $(\vec{v}(\tau), \vec{\alpha}(\tau)) = (\tau\vec{v}, \tau\vec{\alpha})$, because rotations and boosts do not commute and v^i are not additive parameters.

The representation matrix of $L(\vec{v}, \vec{\alpha})$ is

$$D^{(j,j')}(\vec{v}, \vec{\alpha}) = D^{(j)}(\vec{\alpha})\, D^{(j)}(-i\vec{u}) \otimes D^{(j')}(\vec{\alpha})\, D^{(j')}(i\vec{u}), \tag{41}$$

where $\vec{u}$ is

$$\vec{u} = \operatorname{arth} |\vec{v}| \cdot \frac{\vec{v}}{|\vec{v}|}, \tag{42}$$

$D^{(j)}(-i\vec{u})$ and $D^{(j')}(i\vec{u})$ are rotations by an imaginary angle.

Example: The subgroup $SO(3)$ ($\vec{v} = 0$)

$$D^{(j,j')}(0, \vec{\alpha}) = D^{(j)}(\vec{\alpha}) \otimes D^{(j')}(\vec{\alpha}) = D^{(j+j')}(\vec{\alpha}) \oplus \ldots \oplus D^{(j-j')}(\vec{\alpha}). \tag{43}$$

Here the representation of the Lorentz group becomes reducible with the exception of $j = 0$ or $j' = 0$. $j = 1$, $j' = 0$ and $j = 0$, $j' = 1$ denote the defining representation of $SO(3)$.

4.7.2 Spinorial representations and local isomorphism to $SL(2, \mathbf{C})$

The simplest nontrivial representations are $j = \frac{1}{2}$, $j' = 0$ and $j = 0$, $j' = \frac{1}{2}$, they are a system of fundamental representations. Choose $j = \frac{1}{2}$, $j' = 0$:

$$D^{(\frac{1}{2},0)}(\vec{v}, \vec{\alpha}) = e^{-\frac{i}{2}\vec{\alpha}\vec{\sigma}}\, e^{\frac{1}{2}\vec{u}\vec{\sigma}}. \quad (\neq e^{-\frac{1}{2}(\vec{\alpha}-i\vec{v})\vec{\sigma}}\ !) \tag{44}$$

With

$$\vec{u}\vec{\sigma} = \begin{pmatrix} u_3 & u_1 - iu_2 \\ u_1 + iu_2 & -u_3 \end{pmatrix}, \quad \text{and} \quad (\vec{u}\vec{\sigma})^2 = u^2 \cdot \mathbf{1},$$

where $u = (u_1^2 + u_2^2 + u_3^2)^{\frac{1}{2}}$, we find

$$e^{-\frac{1}{2}\vec{u}\vec{\sigma}} = \begin{pmatrix} \operatorname{ch}\frac{u}{2} - \frac{u_3}{u}\operatorname{sh}\frac{u}{2} & \left(-\frac{u_1}{u} + i\frac{u_2}{u}\right)\operatorname{sh}\frac{u}{2} \\ \left(-\frac{u_1}{u} - i\frac{u_2}{u}\right)\operatorname{sh}\frac{u}{2} & \operatorname{ch}\frac{u}{2} + \frac{u_3}{u}\operatorname{sh}\frac{u}{2} \end{pmatrix} = D^{(\frac{1}{2})}(-\vec{u}\vec{\sigma}).$$

Introducing the unit vector $n^i = u^i/u$ we can write

$$D^{(\frac{1}{2})}(u, \vec{n}) = \begin{pmatrix} \operatorname{ch}\frac{u}{2} - n_3\operatorname{sh}\frac{u}{2} & (-n_1 + in_2)\operatorname{sh}\frac{u}{2} \\ -(n_1 + in_2)\operatorname{sh}\frac{u}{2} & \operatorname{ch}\frac{u}{2} + n_3\operatorname{sh}\frac{u}{2} \end{pmatrix} \tag{45}$$

The determinant of this matrix is equal to one, so $D^{(\frac{1}{2})}(u, \vec{n}) \in SL(2, \mathbf{C})$, the group of complex 2 by 2 matrices with unit determinant.

A four-vector x^i can be uniquely transformed into a 2×2 matrix X in the following way

$$X := x^0 \cdot \mathbf{1} + \vec{x}\vec{\sigma} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x_3 \end{pmatrix}. \tag{46}$$

X is hermitian for real x^i, $\operatorname{tr} X = 2x^0$. We can introduce the formal four-vectors

$$\{\sigma_i\} = \{\mathbf{1}, \vec{\sigma}\} \quad \text{and} \quad \tilde{\sigma}^i = \sigma_i \qquad (\neq \eta^{ik}\sigma_k), \tag{47}$$

then

$$x^i = \frac{1}{2}\operatorname{tr}(X\tilde{\sigma}^i). \tag{48}$$

The determinant is

$$\det X = (x^0)^2 - \vec{x}^2 = x^i\, x_i. \tag{49}$$

When we multiply a hermitian X by an arbitrary complex unimodular matrix $A \in SL(2, \mathbf{C})$ in the following way,

$$X' = AXA^\dagger, \tag{50}$$

then X' is also hermitian and we can associate to it the real four-vector

$$x'^i = \frac{1}{2}\operatorname{tr}(X'\tilde{\sigma}^i), \tag{51}$$

and

$$x'^i\, x'_i = \det X' = \det X = x^i\, x_i, \tag{52}$$

because $\det A = 1$. The norm of x^i is thus conserved.

Now the Lorentz matrix can be expressed in terms of A:

$$L^i{}_k = \frac{1}{2}\operatorname{tr} A\, \sigma_k\, A^\dagger\, \tilde{\sigma}^i, \tag{53}$$

and vice versa

$$A = \pm \frac{L^i{}_k\, \sigma_i\, \tilde{\sigma}^k}{\sqrt{\det L^i{}_k\, \sigma_i\, \tilde{\sigma}^k}}. \tag{54}$$

One Lorentz matrix L corresponds to the matrices $\pm A \in SL(2, \mathbf{C})$, so $SL(2, \mathbf{C})$ is a double covering of the proper orthochronous Lorentz group.

5 Group Manifolds

5.1 Manifolds

A manifold M is a topological Hausdorff space with local diffeomorphic mappings to R^n – coordinates $\{x^i\}$, $i = 1, \ldots, n$.
Curve $x^i(s)$, tangent vector $v^i = \mathrm{d}x^i(s)/\mathrm{d}s$, $v = v^i(s)\partial_i$ in the coordinate basis. Tangent space T_pM, cotangent space T_p^*M, diffeomorphism $f : M \to N$ or $M \to M$, $f : x^i \to y^i$. A diffeomorphism induces mappings in the tangent and cotangent spaces.
Push-forward $f_* : T_p \to T_{f(p)}$. f maps a curve $x^i(s)$ to $y^i(x^j(s))$, f_* maps the corresponding tangent vectors:

$$\frac{\mathrm{d}x^i}{\mathrm{d}s}\frac{\partial}{\partial x^i} \to \frac{\partial y^i}{\partial x^j}\frac{\mathrm{d}x^j}{\mathrm{d}s}\frac{\partial}{\partial y^i} = \bar{v}^i\frac{\partial}{\partial y^i},$$

in components

$$v^i\,\frac{\partial}{\partial x^i} \to v^i\,\frac{\partial y^j}{\partial x^i}\,\frac{\partial}{\partial y^j}. \tag{55}$$

In T^*M, the space of covectors or one-forms, the duals of vectors, f induces a mapping $T^*_{f(p)}M \to T^*_pM$, the pull-back, defined by

$$\langle f^*\omega, v\rangle = \langle \omega, f_*v\rangle. \tag{56}$$

For the components this means

$$\left\langle f^*(\omega_i \mathrm{d}x^i), v^j\frac{\partial}{\partial x^j}\right\rangle = \left\langle \omega_i\,\mathrm{d}y^i, v^k\frac{\partial y^j}{\partial x^k}\frac{\partial}{\partial y^j}\right\rangle = \omega_i\,v^j\,\frac{\partial y^i}{\partial x^j},$$

so that

$$(f^*\omega)_i = \omega_j\,\frac{\partial y^j}{\partial x^i}. \tag{57}$$

5.2 Invariant vector fields on group manifolds

Define a left translation $L_a : G \to G$ and a right translation $R_a : G \to G$ for a certain element a of a Lie group G by

$$L_a g = ag, \qquad R_a g = ga \qquad \forall g \in G. \tag{58}$$

The diffeomorphic maps L_a and R_a induce maps in the tangent spaces, $L_{a*} : T_gG \to T_{ag}G$, R_{a*} analogous.

Left-invariant vector fields X on a group manifold G are defined by

$$L_{a*}X|_g = X|_{ag}\,, \tag{59}$$

in coordinates

$$L_{a*}X\big|_g = X^\mu(g)\,\frac{\partial x^\nu(ag)}{\partial x^\mu(g)}\,\frac{\partial}{\partial x^\nu}\bigg|_{ag} = X^\nu(ag)\,\frac{\partial}{\partial x^\nu}\bigg|_{ag}. \tag{60}$$

A vector $v \in T_eG$ defines a unique left-invariant vector field X_v throughout G by

$$X_v\big|_g = L_{g*}v, \qquad g \in G$$

and a left-invariant vector field X defines a unique vector $v = X\big|_e \in T_eG$.

In consequence there is a vector space isomorphism between the set of left-invariant vector fields $\mathbf{g}$ and T_eG. To establish an algebra isomorphism we have to find an equivalent to the commutator in the language of vector fields.

Tangent vectors on a manifold act as differential operators on functions, calculating their derivative in a certain direction. In a coordinate basis the action of a vector (field) $V(x) = V^i(x)\partial_i$ on a function is given by

$$V[f] = V^i(x)\partial_i f(x) \in \mathbf{R}. \tag{61}$$

This is again a scalar field which can be acted upon once more by a vector field, so we can define in a natural way a commutator

$$[U,V][f] = UV[f] - VU[f] = (V^i{}_{,k}\,U^k - U^i{}_{,k}\,V^k)f_{,i}, \tag{62}$$

so the commutator is again a vector field, called the Lie bracket of U and V. In terms of coordinates this is

$$[U,V]^i = V^i{}_{,k}\,U^k - U^i{}_{,k}\,V^k. \tag{63}$$

Left-invariant vector fields are closed under the Lie bracket,

$$L_{a*}[X,Y]_g = [L_{a*}X\big|_g\,, L_{a*}Y\big|_g] = [X,Y]_{ag}, \tag{64}$$

so $\mathbf{g}$ can be identified with the Lie algebra of G and the latter one can be understood as an algebra of infinitesimal group elements in the neighborhood of any finite group element. The construction of left-invariant vector fields from a basis in T_eG manages to attach the Lie algebra to every element of G.

Example: Take the curve through the unity of $SO(3)$ given by $P(s) := \exp(s\Lambda_i)$, in fact a one-parameter subgroup, and act on it with a fixed group element g, given by the rotation matrix $R(\vec{\alpha})$ from the left, then

$$Q(s) := L_gP(s) = R(\vec{\alpha})P(s) \tag{65}$$

is a curve passing through the element g. To find the tangent vector at g, $t_i = L_{g*}\Lambda_i$, we write $Q(s)$ as $g \equiv R(\vec{\alpha})$, followed by a rotation $S(s)$ and consider the limit $s \to 0$. From $Q(s) = R(\vec{\alpha})P(s) = S(s)R(\vec{\alpha})$ we find

$$S(s) = R(\vec{\alpha})P(s)R(\vec{\alpha})^{-1}. \tag{66}$$

For $s \ll 1$, i.e. close to unity on the curve $P(s)$ and close to $R(\vec{\alpha})$ on $Q(s)$, $\exp(s\Lambda_i) \approx \mathbf{1} + s\Lambda_i$ and $S(s) \approx \mathbf{1} + st_i$, where the generator t_i represents the tangent vector to $Q(s)$ at $R(\vec{\alpha})$. From (**??**) we find that $t_i = \mathrm{Ad}_g\Lambda_i$ and further that $[t_i, t_j] = \varepsilon_{ij}{}^k t_k$ everywhere on the group manifold.

5.3 Frames and structure equation

Take a basis $\{V_1, V_2, \ldots, V_n\}$ of T_eG of an n-dimensional group manifold. By push-forward we can construct n linearly independent left-invariant vector fields $\{X_1, X_2, \ldots, X_n\}$ at each point of G, $X_\mu\big|_g = L_{g*}V_\mu$. $\{X_\mu\}$ form a basis of every tangent space T_gG. The Lie bracket = commutator is again an element of T_gG, so it can be expanded in terms of $\{X_\mu\}$,

$$[X_\mu, X_\nu] = c_{\mu\nu}{}^{\lambda}\, X_\lambda. \tag{67}$$

Due to the invariance of the Lie bracket = commutator under the left action L_g, the structure constants are independent of g.

Define a basis of left-invariant one-forms $\{\theta^\mu\}$ dual to $\{X_\mu\}$, $\langle\theta^\mu, X_\nu\rangle = \delta^\mu_\nu$. The exterior derivative is a two-form with the following action on two basis vectors

$$\begin{aligned}
\mathrm{d}\theta^\mu(X_\nu, X_\lambda) \quad &= X_\nu[\theta^\mu(X_\lambda)] - X_\lambda[\theta^\mu(X_\nu)] - \theta^\mu([X_\nu, X_\lambda]) \\
&= X_\nu[\delta^\mu_\lambda] - X_\lambda[\delta^\mu_\nu] - \theta^\mu(c_{\nu\lambda}{}^\kappa X_\kappa) = -c_{\nu\lambda}{}^\mu.
\end{aligned}$$

From this follows the **Maurer-Cartan structure equation**

$$\mathrm{d}\theta^\mu = -\frac{1}{2}\, c_{\nu\lambda}{}^\mu \theta^\nu \wedge \theta^\lambda. \tag{68}$$

The **Maurer-Cartan form or canonical form**
is defined as a Lie algebra-valued one-form $\theta:\ T_gG \to T_eG$ by

$$\theta:\ X \to (L_{g^{-1}})_* X = (L_g)_*^{-1}\, X, \qquad X \in T_gG. \tag{69}$$

Take a basis $\{V_\mu\}$ of T_eG, a basis $\{X_\mu\}$ of T_gG, generated by $X_\mu\big|_g = L_{g*}V_\mu$, a basis $\{\theta^\mu\}$ of T^*_gG, and a vector $Y = Y^\mu X_\mu \in T_gG$. Then

$$\theta(Y) = Y^\mu\theta(X_\mu) = Y^\mu\, L_{g*}^{-1}[L_{g*}V_\mu] = Y^\mu V_\mu, \tag{70}$$

θ simply replaces the basis vectors X_μ by V_μ, so we may write it explicitly as

$$\theta = V_\mu \otimes \theta^\mu. \tag{71}$$

From the structure equation follows

$$\mathrm{d}\theta + \frac{1}{2}[\theta \wedge \theta] = -\frac{1}{2} V_\mu \otimes c_{\nu\lambda}{}^\mu \theta^\nu \wedge \theta^\lambda + \frac{1}{2} c_{\nu\lambda}{}^\mu V_\mu \otimes \theta^\nu \wedge \theta^\lambda = 0. \tag{72}$$

A straightforward way to introduce the Maurer-Cartan form for matrix groups is to define it as

$$\theta = g^{-1}\mathrm{d}g \tag{73}$$

at every point g of the group manifold. Take, as an example, the one-parameter subgroup of elements

$$g(\phi) = \exp(\phi\Lambda_3) = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

of rotations around the 3-axis of SO(3). We see that

$$\begin{aligned} g^{-1}(\phi)\,\mathrm{d}g(\phi) &= \begin{pmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -\sin\phi & -\cos\phi & 0 \\ \cos\phi & -\sin\phi & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathrm{d}\phi \\ &= \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathrm{d}\phi = \Lambda_3 \otimes \mathrm{d}\phi \end{aligned}$$

is independent of ϕ. When it acts on a tangent vector to the curve $g(\phi)$ at ϕ, it gives Λ_3, the tangent vector at e.

From the Maurer-Cartan form $g^{-1}\mathrm{d}g$ we may read off left-invariant one-forms for every $g \in G$ and having found a basis $\{\theta^\mu\}$ at every point we can construct a (left-invariant) metric

$$\mathrm{d}s^2 = \theta^\mu \otimes \theta^\nu \delta_{\mu\nu} \tag{74}$$

on all of the group. It extends the Cartan-Killing metric from the Lie-algebra T_eG to the whole manifold. With the aid of it one constructs further a left invariant measure, the left Haar measure $\mathrm{d}\mu_H$ for compact groups. It satisfies

$$\int_G \mathrm{d}\mu_H(g)\, f(g) = \int_G \mathrm{d}\mu_H(g)\, f(ag) \quad \forall a. \tag{75}$$

Analogously one can construct a right Haar measure. For $SU(2)$, for example, the Haar measure is the usual measure on the sphere S^3 (round metric).

Example: An $SO(3)$ matrix e^Λ, where

$$\Lambda := \vec{\alpha}\vec{\Lambda} = \alpha \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} + \beta \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Explicitly

$$\Lambda = \begin{pmatrix} 0 & -\gamma & \beta \\ \gamma & 0 & -\alpha \\ -\beta & \alpha & 0 \end{pmatrix} \quad \text{and} \quad \Lambda^2 = \begin{pmatrix} -\beta^2-\gamma^2 & \alpha\beta & \alpha\gamma \\ \alpha\beta & -\alpha^2-\gamma^2 & \beta\gamma \\ \alpha\gamma & \beta\gamma & -\alpha^2-\beta^2 \end{pmatrix}, \tag{76}$$

and with the definition

$$n^2 := \alpha^2 + \beta^2 + \gamma^2 \tag{77}$$

we find for the higher powers of Λ

$$\Lambda^3 = -n^2\Lambda, \quad \Lambda^4 = -n^2\Lambda^2, \quad \Lambda^5 = n^4\Lambda, \quad \Lambda^6 = n^4\Lambda^2, \dots$$

so that

$$g = e^\Lambda = \mathbf{1} + \frac{\sin n}{n}\,\Lambda + \frac{1-\cos n}{n^2}\,\Lambda^2. \tag{78}$$

Further

$$g^{-1} = \mathbf{1} - \frac{\sin n}{n}\,\Lambda + \frac{1-\cos n}{n^2}\,\Lambda^2 \tag{79}$$

and

$$\begin{aligned} \mathrm{d}g = {} & \frac{n\sin n + 2\cos n - 2}{n^3}\,\Lambda^2\,\mathrm{d}n + \frac{1-\cos n}{n^2}\,\mathrm{d}\Lambda^2 + \\ & \frac{n\cos n - \sin n}{n^2}\,\Lambda\,\mathrm{d}n + \frac{\sin n}{n}\,\mathrm{d}\Lambda. \end{aligned} \tag{80}$$

After some calculations we find the Maurer-Cartan form in the coordinates (α, β, γ):

$$g^{-1}\mathrm{d}g = \left[\frac{n-\sin n}{n^2}\,\mathrm{d}n\,\alpha_i + \frac{\sin n}{n}\,\mathrm{d}\alpha_i - \frac{1-\cos n}{n^2}\,(\vec{\alpha}\times\mathrm{d}\vec{\alpha})_i\right]\otimes\Lambda_i. \tag{81}$$

In an analogous way we can define a “square” of left-invariant vector fields, called the Laplacian,

$$\Delta = \delta^{\mu\nu} X_\mu X_\nu. \tag{82}$$

With the vectors interpreted as differential operators this is indeed the usual Laplace operator associated with the left-invariant metric.

5.4 Bundles

A bundle E is locally a Kronecker product of a basis manifold M and a fibre F. If F is a vector space, the bundle is called vector bundle. In local coordinates an element u of E has coordinates (p, f), where p^i are coordinates on M and f^k are components of a vector in F.

The canonical projection π is a mapping $E \to M$, $\pi u = p$, $\pi^{-1} p = F$. A section σ is a mapping $M \to E$, $p \to (p, f(p))$, such that $\pi\sigma = \mathrm{id}$. Sections determine vector fields $f(p)$.

The structure group G of a bundle is a Lie group that acts from the left on the fibres and has the following property. Take a set $\{U_i\}$ of open coverings of M with diffeomorphisms $\phi_i : U_i \times F \to \pi^{-1}(U_i)$ with $\pi\phi_i(p, f) = p$, i. e. a local trivialization of the bundle, mapping $\pi^{-1}(U_i)$ onto $U_i \times F$. If we write $\phi_i(p, f) = \phi_{i,p}(f)$, then $\phi_{i,p} : F \to F_p$ is a diffeomorphism. On $U_i \cap U_j \neq \emptyset$ $t_{ij}(p) := \phi_{i,p}^{-1} \circ \phi_{j,p} : F \to F$ is an element of G. ϕ_i and ϕ_j are related by a smooth map

$$t_{ij} : U_i \cap U_j \to G, \quad \phi_j(p, f) = \phi_i(p, t_{ij}(p) f). \tag{83}$$

t_{ij} is called the transition function, G the structure group. Bundles are denoted by (E, π, M, F, G). In addition to the left action there is also a right action of G.

A simple example: A cylinder can be described as a product bundle of S^1 and the interval $I = [-1, 1]$ with the structure group consisting only of the unit element e. On the overlap of arbitrary intervals the transition function is the identity. A Moebius strip is locally a product of an interval $U \subset S^1$ and the interval I, but it has no global product structure. The transition functions are the identity or a reflection P of I. In this case of a nontrivial bundle the structure group is $\{e, P\} \simeq Z_2$.

A principal bundle is a fibre bundle with the fibre F identical to the structure group. Principal bundles with structure group G are also called G-bundles and denoted by $P(M, G)$.

6 Connections on principal bundles

Definition: Assume $u \in P(M, G)$, $p = \pi(u)$. The **vertical subspace** $V_u P$ of the tangent space $T_u P$ is the subspace of vectors tangent to G.

Construction: Take an element of the Lie algebra, $A \in \mathbf{g}$. The right action

$$R_{\exp(tA)} u = u \exp(tA)$$

defines a curve in P through u. As

$$\pi(u) = \pi\left(u\,\exp(tA)\right) = p,$$

the curve lies in G. Define a vector $A^{\#} \in T_uP$ by

$$A^{\#} f(u) = \frac{\mathrm{d}}{\mathrm{d}t} f\left(u \exp(tA)\right)\bigg|_{t=0}, \qquad f : P \to \mathbf{R}. \tag{84}$$

$A^{\#}$ is tangent to the fibre G_p at u, $A^{\#} \in V_uP$.

In this way we construct a vector field $A^{\#}$, the so-called **fundamental vector field**, generated by A. There is a vector space isomorphism $\# : \mathbf{g} \to V_uP$, $A \to A^{\#}$. We have
(i) $\pi_* X = 0 \;\; \forall X \in V_uP$,
(ii) $[A^{\#}, B^{\#}] = [A, B]^{\#}$.

The horizontal subspace H_uP of T_uP is a complement. A connection on P is a unique separation of the tangent space T_uP into the vertical subspace V_uP and the horizontal subspace H_uP, such that
(i) $T_uP = H_uP \oplus T_uP$,
(ii) a smooth vector field X on P is separated into smooth vector fields $X^H \in H_uP$ and $X^V \in V_uP$ as $X = X^H + X^V$,
(iii) $H_{ug}P = R_{g*}H_uP$ for arbitrary $u \in P$, $g \in G$.

(iii) states that the horizontal subspaces H_uP and $H_{ug}P$ on the same fibre are related by R_{g*}. Thus the subspace H_uP at u generates all the horizontal subspaces on the same fibre.

The **connection one-form** is a Lie-algebra-valued one-form $\omega \in \mathbf{g} \otimes T^*P$ that defines as a projection of T_uP onto the vertical subspace $V_uP \simeq \mathbf{g}$.

(i) $\omega(A^{\#}) = A$, $A \in \mathbf{g}$,

(ii) ${R_g}^*\omega = \mathrm{Ad}_{g^{-1}}\omega$,

that is ${R_g}^*\omega_{ug}(X) = \omega_{ug}(R_{g*}X) = g^{-1}\omega_u(X)g$. The pullback acts on the form index, the adjoint action on the Lie-algebra indices.

To show consistency with the projection property of a connection, define the horizontal subspace H_uP as the kernel of ω,

$$H_uP = \{X \in T_uP | \omega(X) = 0\}. \tag{85}$$

Show that $R_{g_*}H_uP = H_uP$ (consistence with the connection property): Take $X \in H_uP$, construct $R_{g_*}X \in T_{ug}P$.

$$\omega(R_{g_*}X) = R_g{}^*\omega(X) = g^{-1}\omega(X)g = 0,$$

as $\omega(X) = 0$. Accordingly $R_{g_*}X \in H_{ug}P$. This connection is called **Ehresmann connection.** The connection one-form ω thus separates T_uP into $H_uP \oplus V_uP$ in accordance with the connection axioms.

Given an open covering $\{U_i\}$ of M and a local section σ_i in each chart U_i, a **local connection form** is a Lie-algebra-valued one-form $\mathcal{A}_i$ on U_i,

$$\mathcal{A}_i = \sigma_i{}^*\omega \in \mathbf{g} \otimes \Omega^1(U_i). \tag{86}$$

Conversely, from a given local one-form $\mathcal{A}_i$ we can construct a connection one-form ω, such that $\mathcal{A}_i = \sigma_i{}^*\omega$. Define a $\mathbf{g}$-valued one-form ω_i in P,

$$\omega_i = g_i^{-1}\pi^*\mathcal{A}_i\, g_i + g_i^{-1}\mathrm{d}_P g_i, \tag{87}$$

d_P is the exterior derivative on P and g_i is the local trivialization defined by $\phi_i^{-1}(u) = (p, g_i)$ for $u = \sigma_i(p)g_i$.

First show that $\sigma_i^*\omega_i = \mathcal{A}_i$. Take $X \in T_PM$.

$$\begin{aligned}\sigma_i^*\omega_i(X) &= \omega_i(\sigma_{i*}X) = \pi^*\mathcal{A}_i(\sigma_{i*}X) + \mathrm{d}_P g_i(\sigma_{i*}X)\\ &= \mathcal{A}_i(\pi_*\sigma_{i*}X) + \mathrm{d}_P g_i(\sigma_{i*}X).\end{aligned}$$

We have made use of $\sigma_{i*}X \in T_{\sigma_i}P$ and $g_i = e$ at σ_{i*}. We further note that $\pi_*\sigma_{i*} = \mathrm{id}_{T_p(M)}$, $\mathrm{d}_P g_i(\sigma_{i*}X) = 0$, since $g \equiv e$ along $\sigma_{i*}X$. Thus $\sigma_{i*}\omega(X) = \mathcal{A}_i(X)$.

Next show that ω_i satisfies the axioms of a connection one-form.
(i) $X = A^\# \in V_uP$, $A \in \mathbf{g} \Rightarrow \pi_*X = 0$.

$$\begin{aligned}\omega_i(A^\#) &= g_i^{-1}\mathrm{d}_P g(A^\#) = g(u)^{-1}\left.\frac{\mathrm{d}g(u\exp(tA))}{\mathrm{d}t}\right|_{t=0}\\ &= g(u)^{-1}g(u)\left.\frac{\mathrm{d}\exp(tA)}{\mathrm{d}t}\right|_{t=0} = A.\end{aligned}$$

(ii) $X \in T_uP$, $h \in G$.

$$R_h^*\omega_i(X) = \omega(R_{h*}X) = g_{iuh}^{-1}\,\mathcal{A}_i(\pi_*R_{h*}X)g_{iuh} + g_{iuh}^{-1}\mathrm{d}_P g_{iuh}(R_{h*}X).$$

Since $g_{iuh} = g_{iu}h$ and $\pi_*R_{h*}X = \pi_*X$ ($\pi R_h = \pi$), we have

$$R_h^*\omega_i(X) = h^{-1}g_{iu}^{-1}\,\mathcal{A}_i(\pi_*X)g_{iu}h + h^{-1}g_{iu}^{-1}\mathrm{d}_P g_{iu}(X)h = h^{-1}\omega_i(X)h,$$

where we have noted that

$$g_{iuh}^{-1}\mathrm{d}_P g_{iuh}(R_{h*}X) = g_{iuh}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}g_{i\gamma(t)h}\Big|_{t=0} =$$

$$h^{-1}g_{iu}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}g_{i\gamma(t)}\Big|_{t=0} h = h^{-1}g_{iu}^{-1}\mathrm{d}_P g_{iu}(X)h.$$

Here $\gamma(t)$ is a curve through $u = \gamma(0)$, whose tangent vector at u is X.

Hence the $\mathbf{g}$-valued one-form ω_i indeed satisfies $\mathcal{A}_i = \sigma_i^*\omega_i$ and the axioms of a connection one-form.

For uniqueness of ω throughout P the relation $\omega_i = \omega_j$ on the overlaps $U_i \cap U_j$ must hold. From this follow the transformation properties of the local forms $\mathcal{A}_i$.

Lemma: Consider the local sections σ_i and σ_j on the neighborhoods U_i and U_j of the principal bundle $P(M,G)$. For $X \in T_P M$, $p \in U_i \cap U_j$, $\sigma_{i*}X$ and $\sigma_{j_*}X$ satisfy

$$\sigma_{j_*}X = R_{t_{ij}\,*}(\sigma_{i*}X) + \left(t_{ij}^{-1}\mathrm{d}t_{ij}(X)\right)^{\#}. \tag{88}$$

Proof: From $\sigma_j(p) = \sigma_i(p)t_{ij}(p)$ we derive (γ is a curve with $\gamma(0) = p$, $\dot\gamma(0) = X$.)

$$\begin{aligned}\sigma_{j_*}(X) &= \frac{\mathrm{d}}{\mathrm{d}t}\sigma_j(\gamma(t))\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left\{\sigma_i(t)t_{ij}(t)\right\}\Big|_{t=0}\\ &= \frac{\mathrm{d}}{\mathrm{d}t}\sigma_i(t)\cdot t_{ij}(p) + \sigma_i(p)\frac{\mathrm{d}}{\mathrm{d}t}t_{ij}(t)\Big|_{t=0}\\ &= R_{t_{ij}\,*}(\sigma_{i*}X) + \sigma_j(p)t_{ij}^{-1}(p)\frac{\mathrm{d}}{\mathrm{d}t}t_{ij}(t)\Big|_{t=0}.\end{aligned}$$

(Assuming G is a matrix group, we have $R_{g_*}X = Xg$.) Further

$$t_{ij}^{-1}(p)\mathrm{d}t_{ij}(X) = t_{ij}^{-1}(p)\frac{\mathrm{d}}{\mathrm{d}t}t_{ij}(t)\Big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t}\left[t_{ij}^{-1}(p)t_{ij}(t)\right]\Big|_{t=0} \in T_e(G) \simeq \mathbf{g}.$$

(Note that $t_{ij}^{-1}(p)t_{ij}(\gamma(t)) = e$ at $t = 0$.) The second term of $\sigma_{j*}X$ represents the vector field $\left(t_{ij}^{-1}\mathrm{d}t_{ij}(X)\right)^{\#}$ at $\sigma_j(p)$.

The compatibility condition is obtained by applying the connection form ω to (**??**):

$$\sigma_j^*\omega(X) = R_{t_{ij}}^*\omega(\sigma_{i*}X) + t_{ij}^{-1}\mathrm{d}t_{ij}(X) = t_{ij}^{-1}\omega(\sigma_{i*}X)t_{ij} + t_{ij}^{-1}\mathrm{d}t_{ij}(X),$$

which yields

$$\mathcal{A}_j = t_{ij}^{-1}\mathcal{A}_i\, t_{ij} + t_{ij}^{-1}\mathrm{d}t_{ij}. \tag{89}$$

Example: A $U(1)$ bundle over M. Assume $\mathcal{A}_i$ and $\mathcal{A}_j$ to be local connection forms on the overlapping charts U_i and U_j and a transition function

$$t_{ij}(p) = \exp[i\chi(p)], \qquad \chi(p) \in R$$

in $U_i \cap U_j$. Then

$$\mathcal{A}_j(p) = t_{ij}(p)^{-1}\mathcal{A}_i\, t_{ij}(p) + t_{ij}(p)^{-1}\mathrm{d}t_{ij}(p) = \mathcal{A}_i(p) + \mathrm{i}\,\mathrm{d}\chi(p).$$

This is the usual gauge transformation law for the electromagnetic vector potential $A_\mu = -\mathrm{i}\mathcal{A}_\mu$,

$$(A_j)_\mu = (A_i)_\mu + \partial_\mu\chi.$$

7 Curvature

7.1 The covariant derivative

of a vector-valued r-form $\phi \in \Omega^r(P) \otimes V$, $\phi : TP \otimes \ldots \otimes TP \to V$ is a mapping $\Omega^r(P) \to \Omega^{r+1}$, defined as

$$D\phi(X_1, \ldots, X_{r+1}) \equiv \mathrm{d}_P\phi(X_1^H, \ldots, X_{r+1}^H) \tag{90}$$

for $X_1, \ldots, X_{r+1} \in T_uP$.

7.2 The curvature two-form

Ω is the covariant derivative of the connection one-form ω,

$$\Omega \equiv D\omega \in \Omega^2 \otimes \mathbf{g}. \tag{91}$$

Behavior under right-translation:

$$R_a^*\Omega = a^{-1}\Omega\, a, \qquad a \in G. \tag{92}$$

Proof: Note that $(R_{a*}X)^H = R_{a*}(X^H)$. R_a preserves the horizontal subspaces, by virtue of the definition of the latter as kernel of the connection and $\mathrm{d}_P\, R_a^* = R_a^*\, \mathrm{d}_p$ for X^H, Y^H.

$$\begin{aligned} R_a^*\,\Omega(X,Y) &= \Omega(R_{a*}X, R_{a*}Y) = \mathrm{d}_P\omega\left((R_{a*}X)^H, (R_{a*}Y)^H\right) \\ &= \mathrm{d}_P\omega(R_{a*}X^H, R_{a*}Y^H) = R_a^*\mathrm{d}_P\omega(X^H, Y^H) \\ &= \mathrm{d}_P R_a^*\omega(X^H, Y^H) = \mathrm{d}_P(a^{-1}\omega\, a)(X^H, Y^H) \\ &= a^{-1}\mathrm{d}_P\omega(X^H, Y^H)\, a = a^{-1}\Omega(X,Y)\, a. \end{aligned}$$

(a is a constant element and hence $\mathrm{d}_P a = 0$).

Cartan's structure equation for ω and Ω.

$$\begin{aligned} \Omega(X,Y) &= \mathrm{d}_P\omega(X,Y) + [\omega(X),\omega(Y)], & (93)\\ \Omega &= \mathrm{d}_P\omega + \omega\wedge\omega. & (94)\end{aligned}$$

Proof: There are three cases:
1) $X, Y \in H_uP$.
$\omega(X) = \omega(Y) = 0$. $\Omega(X,Y) = \mathrm{d}_P\omega(X^H, Y^H) = \mathrm{d}_P\omega(X,Y)$.

2) $X \in H_uP$, $Y \in V_uP$: $\Omega(X,Y) = 0$.
$\omega(X) = 0$, so we need to prove $\mathrm{d}_P\omega(X,Y) = 0$.
$\mathrm{d}_P\omega(X,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]) = X\omega(Y) - \omega([X,Y])$. $Y \in V_uP$, so there is an element $V \in \mathbf{g}$, s. t. $Y = V^{\#}$. Then $\omega(Y) = V$ is constant, hence $X\omega(Y) = X \cdot V = 0$.
The vector field Y is generated by $g(t) = \exp(tV)$, so

$$[X,Y] = \lim_{t\to 0}\frac{1}{t}(R_{g(t)_*}X - X).$$

Since $R_{g_*}H_uP = H_{ug}P$, $R_{g_*}X$ is horizontal and so is $[X,Y] \Rightarrow \omega([X,Y]) = 0$.

3) $X, Y \in V_uP$: $\Omega(X,Y) = 0$.
In this case $\mathrm{d}_P\omega(x,Y) = X\omega(Y) - Y\omega(X) - \omega([X,Y]) = -\omega([X,Y]) = -[\omega(X),\omega(Y)]$.

7.3 The local form of the curvature

is the pull-back

$$\mathcal{F} = \sigma^*\Omega \tag{95}$$

for a local section σ on a chart U. Expressed in terms of the gauge potential $\mathcal{A}$, $\mathcal{F}$ becomes

$$\mathcal{F} = \mathrm{d}\mathcal{A} + \mathcal{A}\wedge\mathcal{A}, \tag{96}$$

where d is the exterior derivative on M.

The action on vectors of TM:

$$\mathcal{F}(X,Y) = \mathrm{d}\mathcal{A}(X,Y) + [\mathcal{A}(X),\mathcal{A}(Y)]. \tag{97}$$

Proof: With $\mathcal{A} = \sigma^*\omega$, $\sigma^*\mathrm{d}_P\omega = \mathrm{d}\sigma^*\omega$, and $\sigma^*(\zeta\wedge\eta) = \sigma^*\zeta\wedge\sigma^*\eta$ we find from Cartan's structure equation

$$\mathcal{F} = \sigma^*(\mathrm{d}_P\omega + \omega\wedge\omega) = \mathrm{d}\sigma^*\omega + \sigma^*\omega\wedge\sigma^*\omega = \mathrm{d}\mathcal{A} + \mathcal{A}\wedge\mathcal{A}, \tag{98}$$

in components

$$\mathcal{F}_{\mu\nu} = \partial_\mu \mathcal{A}_\nu - \partial_\nu \mathcal{A}_\mu + [\mathcal{A}_\mu, \mathcal{A}_\nu]. \tag{99}$$

$\mathcal{F}$ is called the (Yang-Mills) field strength. In the overlap of two coordinate neighborhoods $U_i \cap U_j$

$$\mathcal{F}_j = \mathrm{Ad}_{t_{ij}} \mathcal{F}_i = t_{ij}^{-1} \mathcal{F}_i \, t_{ij}. \tag{100}$$

7.4 Bianchi identities

When ω and Ω are given in terms of the basis $\{T_i\}$ of $\mathbf{g}$, $\omega = \omega^i T_i$ and $\Omega = \Omega^i T_i$, then

$$\Omega^i = \mathrm{d}_P \omega^i + f^i_{jk} \, \omega^j \wedge \omega^k. \tag{101}$$

Exterior differentiation yields

$$\mathrm{d}_P \Omega^i = f^i_{jk} \, \mathrm{d}_p \omega^j \wedge \omega^k - f^i_{jk} \, \omega^j \wedge \mathrm{d}_p \omega^k. \tag{102}$$

From $\omega(X) = 0$ for a horizontal vector X we have

$$D\Omega(X, Y, Z) = \mathrm{d}_P \, \Omega(X^H, Y^H, Z^H) = 0,$$

where $X, Y, Z \in T_u P$. This is the Bianchi identity

$$D\Omega = 0, \tag{103}$$

in local form

$$\mathrm{d}\mathcal{F} + [\mathcal{A}, \mathcal{F}] = 0. \tag{104}$$

This is the form of the homogenous Maxwell and Yang-Mills equations.

To show this relation, operate with σ^* on (**??**). Left-hand side

$$\sigma^* \, \mathrm{d}_P \Omega = \mathrm{d} \cdot \sigma^* \Omega = \mathrm{d}\mathcal{F}.$$

Right-hand side

$$\sigma^*(\mathrm{d}_P \omega \wedge \omega - \omega \wedge \mathrm{d}_\mathrm{P} \omega) = \mathrm{d}\sigma^* \omega \wedge \sigma^* \omega - \sigma^* \omega \wedge \mathrm{d}\sigma^* \omega = \mathrm{d}\mathcal{A} \wedge \mathcal{A} - \mathcal{A} \wedge \mathrm{d}\mathcal{A}.$$

From this follows eq. (**??**)

$$\mathcal{D}\mathcal{F} := \mathrm{d}\mathcal{F} + \mathcal{A} \wedge \mathcal{F} - \mathcal{F} \wedge \mathcal{A} = \mathrm{d}\mathcal{F} + [\mathcal{A}, \mathcal{F}] = 0,$$

where the action of $\mathcal{D}$ on a $\mathbf{g}$-valued p-form η on M is defined by

$$\mathcal{D}\eta = \mathrm{d}\eta + [\mathcal{A}, \eta]. \tag{105}$$

Note that $\mathcal{D}\mathcal{F} = \mathrm{d}\mathcal{F}$ for $G = U(1)$ (Maxwell).

8 Covariant differentiation and curvature on vector bundles

8.1 Covariant differentiation

Take a vector bundle with the fibre F being a vector space $\mathbf{V}$ of arbitrary dimension, independent of the dimension of the basis manifold. The covariant derivative $D\psi$ of a section ψ of the bundle (= vector field) is a bundle-valued one-form. Acting on a tangent vector $u \in T(M)$ it gives another section,

$$(D\psi)(u) = D_u \psi, \tag{106}$$

the covariant derivative of ψ along u. At a point p of the bundle

$$D\psi_p \in \mathbf{V} \otimes T_p^*. \tag{107}$$

A concrete differentiation or connection is specified by its action on a section basis $\{b_I\}$ (in the following capital indices denote vector indices in $\mathbf{V}$, lower-case indices are tangent space indices)

$$Db_I = \omega^J{}_I \otimes b_J, \qquad D_u b_I = \omega^J{}_I(u)\, b_J. \tag{108}$$

In a basis $\{e_i\}$ of tangent vectors and a co-basis $\{e^i\}$ we can write

$$\omega^J{}_I = A^J{}_{Ii}\, e^i, \qquad \omega^J{}_I(e_i) = A^J{}_{Ii}, \qquad D_{e_i}\, b_I = A^J{}_{Ii}\, b_J. \tag{109}$$

For an arbitrary section ψ with $\psi = \psi^I b_I$ the coefficients ψ^I are ordinary functions,

$$D\psi = (D\psi)^I \otimes b_I, \tag{110}$$

and

$$(D\psi)^I := \mathrm{d}\psi^I + \omega^I{}_J\, \psi^J \tag{111}$$

are ordinary one-forms with

$$D_{e_i}\psi = \psi^I{}_{;i}\, b_I, \qquad \psi^I{}_{;i} = e_i(\psi^I) + A^I{}_{Ji}\, \psi^J. \tag{112}$$

From the action on $\{b_I\}$ follows the action on the dual basis $\{b^I\}$:

$$\begin{aligned}
&D\langle b^I, b_J\rangle = \langle Db^I, b_J\rangle + \langle b^I, Db_J\rangle = D\,\delta^I{}_J = 0\\
&\Rightarrow \langle Db^I, b_J\rangle = -\langle b^I, Db_J\rangle = -\langle b^I, \omega^K{}_J \otimes b_K\rangle = -\omega^K{}_J\,\langle b^I, b_K\rangle\\
&= -\omega^I{}_J = -\omega^I{}_K\,\delta^K{}_J = -\omega^I{}_K\langle b^K, b_J\rangle = \langle -\omega^I{}_K \otimes b^K, b_J\rangle
\end{aligned}$$

From this we get

$$Db^I = -\omega^I{}_K \otimes b^K. \tag{113}$$

In the special case $F = \mathbf{V} \otimes \mathbf{V}^*$, when the fibre is the space of tensors of type $(1,1)$, an element of this fibre defines a mapping $\mathbf{V}_p \to \mathbf{V}_p$ by contraction. In the case of the tangent bundle the elements of $T^1_1 = T(M) \otimes T^*(M)$ define a mapping $T_p \to T_p$ and can be also interpreted as vector-valued one-forms and we can also define an exterior covariant derivative D (only for the tangent bundle!). For a section ψ of T^1_1

- $D\psi$ is a T^1_1 - valued one-form,
- $D \wedge \psi$ is a T^1 - valued two-form.

For a *bundle-valued n-form* $\psi \otimes \alpha$ with ψ being a section of the vector bundle and α an n-form, the exterior covariant derivative is the bundle-valued $n+1$ - form

$$D \wedge (\psi \otimes \alpha) = (D\psi) \wedge \alpha + \psi \otimes \mathrm{d}\alpha. \tag{114}$$

The section

$$\delta = \delta_I{}^J\, b^I \otimes b_J \tag{115}$$

gives rise to the identical map $\mathbf{V}_p \to \mathbf{V}_p$ and its covariant derivative is equal to zero. In the case of tangent bundles the section $\delta = e^i \otimes e_i$ can be seen as a section of T^1_1 as well as a vectorial one-form with $D\delta = 0$ but

$$\begin{aligned} D \wedge \delta = (De_i) \wedge e^i + e_i \otimes \mathrm{d}e^i = \omega^j{}_i \otimes e_j \wedge e^i + e_i \otimes \mathrm{d}e^i = \\ e_i \otimes \omega^i{}_j \wedge e^j + e_i \otimes \mathrm{d}e^i = e_i \otimes (\mathrm{d}e^i + \omega^i{}_j \wedge e^j) =: e_i \otimes \theta^i = \theta. \end{aligned} \tag{116}$$

$$\theta := D \wedge \delta \tag{117}$$

is a vector-valued two-form, called the torsion form and (**??**) is called the first Cartan structure equation. The action of the torsion form on a pair of tangent vectors is

$$\theta(u, v) = D_u v - D_v u - [u, v]. \tag{118}$$

Generally a section ψ of $\mathbf{V} \otimes \mathbf{V}^*$ has components $\psi^I{}_J$, $\psi = \psi^I{}_J\, e_I \otimes e^J$. In the case of the tangent bundle the indices can be written as i, j, so that $\psi^I{}_{J;k} \equiv \psi^i{}_{j;k}$ and j and k can be antisymmetrized. When ϕ is a vectorial one-form,

$$D \wedge \phi = \mathrm{Alt}\, D\phi + \mathrm{C}(\phi \otimes \theta), \tag{119}$$

where "Alt" means antisymmetrization and "C" means contraction. The explicit action of a bundle-valued one form ϕ on a pair of vectors is

$$(D \wedge \phi)(u, v) = D_u(\phi(v)) - D_v(\phi(u)) - \phi([u, v]). \tag{120}$$

8.2 Parallel transport and curvature

A connection allows to transport a vector in $\mathbf{V}_p$ to an infinitely close fibre $\mathbf{V}_q$. A section ψ with values ψ_p, ψ_q, which are parallel in the sense of the connection, satisfies $D_u\psi = 0$, where u is an infinitesimal vector from p to q (value in p = value in q, transported to p). $D\psi = 0$ globally means that ψ is covariantly constant. This equation does not have nontrivial solutions in general, because it gives rise to the nontrivial integrability conditions

$$D \wedge D\psi = 0. \tag{121}$$

Interestingly the $\mathbf{V}$-valued two-form $D \wedge D\psi$ depends at every point homogeneously linearly only on ψ, not on its derivatives. Take a function f

$$\begin{aligned}
&D \wedge D(f\psi) = D \wedge (\psi \otimes \mathrm{d}f + f\, D\psi) = \\
&D\psi \wedge \mathrm{d}f + \psi \otimes \mathrm{d} \wedge \mathrm{d}f + \mathrm{d}f \wedge D\psi + f\, D \wedge D\psi = f\, D \wedge D\psi,
\end{aligned}$$

because the first and the third term cancel and the second is zero. thus $D \wedge D$ defines a two-form Ω with values in $\mathbf{V} \otimes \mathbf{V}^*$. In each point the values $\Omega(u, v)$ belong to $\mathbf{V}_p \otimes \mathbf{V}_p^*$ in such a way that the contraction of Ω with ψ is equal to $\Omega(\psi) = D \wedge D\psi$. Ω is called the curvature form of the connection. It is antisymmetric in its tangent vector arguments $\Omega(u, v) = -\Omega(v, u)$ and from (**??**) follows

$$\Omega(u, v)(\psi) = D_u(D_v\psi) - D_v(D_u\psi) - D_{[u,v]}\psi. \tag{122}$$

This relation is called the Ricci identity.

With respect to the section basis $\{b^I \otimes b_J\}$ Ω has the decomposition

$$\Omega = \Omega^J{}_I \otimes b^I \otimes b_J,$$

where $\Omega^J{}_I$ are ordinary two-forms, forming the *curvature matrix*. For $\Omega(b_I) = \Omega^J{}_I \otimes b_J$ holds

$$\Omega(b_I) = D \wedge D\, b_I = D \wedge (\omega^J{}_I \otimes b_J) = (\mathrm{d} \wedge \omega^J{}_I) \otimes b_J - \omega^J{}_I \wedge \omega^K{}_J \otimes b_K.$$

From this follows Cartan's second structure equation

$$\Omega^J{}_I = \mathrm{d} \wedge \omega^J{}_I + \omega^J{}_K \wedge \omega^K{}_I \quad \text{or} \quad \Omega = \mathrm{d} \wedge \omega + \omega \wedge \omega. \tag{123}$$

For Ω holds the second Bianchi identity

$$D \wedge \Omega \equiv 0 \quad \text{or} \quad \mathrm{d} \wedge \Omega + \omega \wedge \Omega - \Omega \wedge \omega \equiv 0. \tag{124}$$

For a tangent bundle from Cartan's first structure equation

$$\theta^i = \mathrm{d} \wedge e^i + \omega^i{}_j \wedge e^j \tag{125}$$

follows

$$D \wedge \theta = C(\Omega \wedge \delta), \quad \text{i. e.} \quad (D \wedge \theta)^i = \Omega^i{}_j \wedge e^j, \tag{126}$$

the first Bianchi identity.

Zero torsion leads to a symmetry property of the Riemann tensor, namely

$$R_{ijkl} + R_{iklj} + R_{iljk} = 0.$$

8.3 Fibre metric, G-structures

In Riemsnnian geometry there is a metric in the tangent bundle of a manifold. This can be generalized to arbitrary vector bundles in form of a fibre metric. A metric is a section of the bundle $(M, \mathbf{V}^* \otimes \mathbf{V}^*)$, which allows to introduce an inner product, a norm, a notion of orthogonality,...

Alternatively, a metric can be constructed by introduction of orthogonal bases $\{b_A\}$, so that $\gamma(b_A, b_B) = \gamma_{AB} = \pm\delta_{AB}$. Two orthogonal bases $\{b_A\}$ and $\{b_B\}$ at one point of M are related by a transformation $b_A = S^{\bar{B}}{}_A\, b_{\bar{B}}$. S lies in a (pseudo-)orthogonal subgroup G of $GL(n)$, the group of nonsingular $n \times n$ matrices, defined by the invariance of the scalar product. S can vary from point to point in M. Thus a fibre metric can be defined by declaring at every point a basis as orthonormal, then all the other orthonormal bases are obtained by application of local transformations $S \in G$.

As a generalization, G need not be an orthogonal group, but an arbitrary Lie subgroup of $GL(n)$. If we choose at every point p a basis $\{b_A\}$ of $\mathbf{V}_p$ and apply to it all transformations $S \in G$, we obtain a class of G-bases in each fibre $\mathbf{V}_p$. The totality of these bases defines a *G-structure*. When we start with a different basis system, not belonging to the above class, we obtain a different G-structure.

In Riemannian geometry the connection does not violate orthogonality. The generalization to arbitrary vector bundles with fibre metric γ is the postulate

$$D\gamma := \mathrm{d}\gamma - \omega^T\gamma - \gamma\,\omega = 0. \tag{127}$$

This is equivalent to the Leibniz rule

$$u(\gamma(\phi, \psi)) = \gamma(D_u\phi, \psi) + \gamma(\phi, D_u\psi). \tag{128}$$

Proof: The right-hand side is

$$\gamma_{AB}(u^i(\partial_i\phi^A + \omega_i{}^A{}_C\phi^C), \psi^B) + \gamma_{AB}(\phi^A, u^i(\partial_i\psi^B + \omega_i{}^B{}_C\psi^C)) =$$

$$u^i\gamma_{AB}[(\partial_i\phi^A,\psi^B)+\omega_i{}^A{}_C(\phi^C,\psi^B)+(\phi^A,\partial_i\psi^B)+\omega_i{}^B{}_C(\phi^A,\psi^C)]=$$
$$u^i[\partial_i(\gamma_{AB}(\phi^A,\psi^B))-(\partial_i\gamma_{AB})(\phi^A,\psi^B)+\omega_i{}^C{}_A\gamma_{CB}(\phi^A,\psi^B)+\omega_i{}^C{}_B\gamma_{AC}(\phi^A,\psi^B)]$$
$$=u^i[\partial_i(\gamma(\phi,\psi))-(\partial_i\gamma)(\phi,\psi)+((\omega_i^T\gamma)_{AB}+(\gamma\omega_i)_{AB})(\phi^A,\psi^B)]=$$
$$u(\gamma(\phi,\psi))-u^i[\partial_i\gamma-\omega_i^T\gamma-\gamma\omega_i](\phi,\psi)=u(\gamma(\phi,\psi))-D_u\gamma(\phi,\psi).$$

The integrability condition for covariant constance of the metric is

$$D\wedge D\gamma=0,\qquad \Omega^T\gamma+\gamma\,\Omega=0. \tag{129}$$

With the definition

$$\Omega_{AB}=\gamma_{AC}\,\Omega^C{}_B$$

this gives rise to another symmetry property of the Riemann tensor.

For arbitrary G-structures the connection should be compatible with it. Given a G-basis field $\{b_A\}$, translation of $\{b_A|_p\}$ at the point p to the point q should lead to a G-basis $\{b_A|_q\}$ in q. $\{b_A|_p+Db_A\}$ should differ from $\{b_A|_q\}$ only infinitesimally, namely by a transformation $S=\mathbf{1}+\omega\in G$. From this we derive a postulate for the covariant derivative D: When $D_ub_A=\omega^B{}_A(u)\,b_B$, then the matrix $\omega(u)$ belongs for each u and for each G-basis to the Lie algebra of G. A G-connection ω is a Lie algebra valued one-form w. r. to G-bases.

- If $G=GL(n)$, there is no restriction, the Lie algebra contains all nonsingular $n\times n$ matrices.
- When G is orthogonal, the Lie algebra consists of antisymmetric matrices.
- In the case $G=\{1\}$ the Lie algebra vanishes, $\omega\equiv 0$. There is only one G basis in a $G-$structure and $\Omega=0$.

Remark: A $\{1\}$-connection defines a parallelism at a distance, each b_A is covariantly constant. Vice versa, however, $\Omega\equiv 0$ does not uniquely determine a $\{1\}$-structure: $\{b_A\}$ and $\{\bar b_A=S^B{}_A\,b_B\}$ with a constant matrix $S^B{}_A$ are both $\{1\}$-structures. A connection with $\Omega\equiv 0$ can have nonvanishing torsion.

8.4 Curvature and torsion of G-connections

G connections have values in the Lie algebra $\{G_a\}$ with $[G_a,G_b]=C^c_{ab}\,G_c$,

$$\Omega=\Omega^aG_a,\qquad \Omega^a=\mathrm{d}\omega^a+\frac{1}{2}\,C^a_{bc}\,\omega^b\wedge\omega^c. \tag{130}$$

On the tangent bundle also torsion can be defined. One may ask, whether for a given G there is a torsion-free G-connection.

- For (pseudo-)orthogonal groups vanishing torsion determines the G-connection uniquely.
- For some groups vanishing torsion is not possible - "essential torsion".
- For other groups there are more than one torsion-free G-connections, for example for the symplectic group, which is characterized by conserving the symplectic form.

Among the groups with $S \in G$, $\det S = 1$ for $n > 2$ (pseudo-)orthogonal groups are the only ones with unique torsion-free G-connection. (For $n = 2$ $G = \{e^\Lambda, \Lambda \in \mathbf{R}\}$ is a counterexample.)

9 Gauge theories

Recall that in a certain basis $\{b_I\}$ the covariant derivative D on a vector bundle is determined by the connection matrix ω. Consider a change of bases

$$\bar{b}_I = S^J{}_I \, b_J. \tag{131}$$

Then

$$Db_I = D(S^J{}_I \, \bar{b}_J) = \left((\mathrm{d}S)^J{}_I + S^K{}_I \, \bar{\omega}^J{}_K\right) \otimes \bar{b}_J. \tag{132}$$

On the other hand

$$Db_I = \omega^K{}_I \, b_K = \omega^K{}_I \, S^J{}_K \otimes \bar{b}_J. \tag{133}$$

Comparison of (**??**) and (**??**) leads to

$$(\mathrm{d}S)^J{}_I + \bar{\omega}^J{}_K S^K{}_I = S^J{}_K \, \omega^K{}_I,$$

multiplication with $(S^{-1})^I{}_L$ gives the typical behavior of a connection under a change of basis,

$$\bar{\omega}^J{}_L = S^J{}_K \, \omega^K{}_I \, (S^{-1})^I{}_L - (\mathrm{d}S)^J{}_I (S^{-1})^I{}_L,$$

in matrix notation

$$\bar{\omega} = S \, \omega \, S^{-1} - (\mathrm{d}S) \, S^{-1}. \tag{134}$$

For the curvature the corresponding relation is

$$\bar{\Omega} = S \, \Omega \, S^{-1} = (\mathrm{Ad}S)(\Omega).. \tag{135}$$

Ω transforms under the adjoint representation of G, it is gauge covariant, but ω is not, it can be made locally equal to zero at every point by certain transformations S. For $\Omega = 0$ there is of course always a gauge with $\omega \equiv 0$ everywhere.

Thus for the covariant derivative of a section ψ holds

$$D\,\psi = \mathrm{d}\psi + \omega\psi, \tag{136}$$

it commutes with changes of bases,

$$\bar{D}\,S\psi = S\,D\psi,$$

whereas

$$\mathrm{d}\,S\psi = S\,\mathrm{d}\psi + (\mathrm{d}S)\psi \neq S\,\mathrm{d}\psi, \tag{137}$$

which means that the ordinary exterior derivative does not commute with transformations $S \in G$.

ω is a "compensation field", it compensates $(\mathrm{d}S)\psi$ in (**??**). In theories the physical content of which does not depend on x-dependent transformations, $S(x)$ are called *local gauge transformations*, G is called the *gauge group*. Theories of this kind are called *Gauge theories*. The coefficients ${\omega^I}_k$ and ${\Omega^I}_{J\,kl}$ of $\omega^I = {\omega^I}_k\,\mathrm{d}x^k$ and ${\Omega^I}_J = {\Omega^I}_{J\,kl}\,\mathrm{d}x^k \wedge \mathrm{d}x^l$ are called *gauge potentials* and *gauge field strengths*, respectively.

The difference quotient

$$\frac{1}{t}[S(x(t)) - S(x(0))]\,S^{-1}(x(0))$$

corresponding to $(\mathrm{d}S)S^{-1}$ describes a curve passing through the unit element of G, so the matrix $(\mathrm{d}S)S^{-1}$ represents the tangent vector at unity, i. e. an element of the Lie algebra.

Example 1: $M =$ space-time, $E =$ tangent bundle, $G = \mathbf{R}^+ \times O(n) = \{e^\Lambda O : \Lambda \in \mathbf{R}, O \in O(n)\}$, the Lie algebra

$$\mathcal{L}G = \{a \cdot \mathbf{1} + M : a \in \mathbf{R}, M = -M^T\} \tag{138}$$

consists of the antisymmetric matrices and multiples of the unit matrix. Denote a G basis by $\{e_a\}$ and a co-basis by $\{\omega^a\}$. Scaling transformations act in the following way:

$$e_a = e^{\Lambda(x)}\,\bar{e}_a, \qquad \omega^a = e^{-\Lambda(x)}\,\bar{\omega}^a. \tag{139}$$

For the metric tensor

$$g = \delta_{ab}\,\omega^a \otimes \omega^b, \quad \text{or} \quad \bar{g} = \delta_{ab}\,\bar{\omega}^a \otimes \bar{\omega}^b \tag{140}$$

follows the conformal transformation

$$\bar{g} = e^{2\Lambda(x)}\, g. \tag{141}$$

The G-structure distinguishes a metric, with respect to which the G-bases are orthogonal, up to a conformal factor (conformal structure on M). A choice of the conformal factor means a gauge of measuring rods and clocks.

A G-connection is defined with respect to $\{e_a\}$ by the connection matrix

$$\omega = \alpha \cdot \mathbf{1} + \omega' \tag{142}$$

with curvature matrix

$$\Omega = \varphi \cdot \mathbf{1} + \Omega' \tag{143}$$

(unique split into an antisymmetric matrix and a multiple of the unit matrix)

$$\varphi = \mathrm{d} \wedge \alpha, \qquad \Omega' = \mathrm{d}\omega' + \omega' \wedge \omega'. \tag{144}$$

- Action of pure rotations $\in O$:

$$\omega \to \bar{\omega} = O\, \omega\, O^{-1} - \mathrm{d}O \cdot O^{-1}, \quad \Omega \to \bar{\Omega} = O\, \Omega\, O^{-1} \tag{145}$$

  (left action of the gauge group). From this follows

$$\bar{\alpha} = \alpha, \quad \bar{\omega}' = O\, \omega'\, O^{-1} - \mathrm{d}O \cdot O^{-1}, \quad \bar{\varphi} = \varphi, \quad \bar{\Omega}' = O\, \Omega'\, O^{-1}. \tag{146}$$

- Action of scale transformations $e^{\Lambda(x)}\mathbf{1}$:

$$\omega \to \bar{\omega} = \omega - \mathrm{d}\Lambda \cdot \mathbf{1}, \qquad \Omega \to \bar{\Omega} = \Omega \tag{147}$$

  This leads to

$$\bar{\alpha} = \alpha - \mathrm{d}\Lambda, \quad \bar{\omega}' = \omega', \quad \bar{\varphi} = \varphi, \quad \bar{\Omega}' = \Omega'. \tag{148}$$

α and $\varphi = \mathrm{d} \wedge \alpha$ transform formally like the four-potential and the electromagnetic field tensor. 1918 H. Weyl attempted to “geometrize” electromagnetism in this way, but for $\varphi \neq 0$ parallel transport of vectors does not conserve length. This would lead to a “second clock effect”: after an accelerated travel also the rate of a clock could be different than before.

Example 2: M = space-time, $E = (M, \mathbf{C})$, $G = U(1) = \{e^{i\Lambda} : \Lambda \in \mathbf{R}\}$. Sections of this bundle are complex functions. The Lie algebra consists of purely imaginary numbers,

$$\omega = i\alpha, \qquad \Omega = i\varphi = i\mathrm{d} \wedge \alpha. \tag{149}$$

Gauge transformations of a section Φ are $\Phi \to e^{i\Lambda}\Phi$, the corresponding transformations of connection and curvature are

$$\omega \to \bar{\omega} = \omega - i\mathrm{d}\Lambda, \quad \Omega \to \bar{\Omega} = \Omega, \quad \bar{\alpha} = \alpha - \mathrm{d}\Lambda, \quad \bar{\varphi} = \varphi. \tag{150}$$

This reflects the transformation of the electromagnetic four-potential and intensities; the non-integrability of the parallel transport of the phase of a wave function Φ describes the Aharonov-Bohm effect.

Now consider the dynamics of the system: Without electromagnetic interaction the complex scalar field Φ satisfies the free Klein-Gordon equation, which is invariant under gauge transformations $\Phi \to e^{i\Lambda}\Phi$ with constant Λ, like the Lagrangian $\partial_k\Phi^*\partial^k\Phi - m^2\Phi^*\Phi$. To admit *local* gauge transformations, we must introduce a covariant derivative

$$D = \mathrm{d} + i\alpha, \qquad D_k = \partial_k + ieA_k \tag{151}$$

("minimal substitution"). From the coupled Lagrangian

$$L = (D_k\Phi)^*\, D^k\Phi - m^2\Phi^*\Phi - \frac{1}{4}\, F_{ik}F^{ik}, \tag{152}$$

respectively the form

$$L\,\mathrm{d}^4x = (D\Phi)^* \wedge \star D\Phi - m^2\Phi^* \star \Phi - \frac{1}{2e^2}\,\varphi \wedge \star\varphi \tag{153}$$

with

$$\varphi = \frac{e}{2}\, F_{ik}\,\mathrm{d}x^i \wedge \mathrm{d}x^k = \mathrm{d}\alpha \tag{154}$$

we may derive the equations for the scalar field

$$\eta^{ik}D_iD_k\Phi + m^2\Phi = 0 \quad \text{or} \quad \star D \wedge \star D\Phi + m^2\Phi = 0 \tag{155}$$

and the electromagnetic field

$$F^{kl}{}_{;l} = -j^k := -ie[(D^k\Phi^*)\Phi - \Phi^* D^k\Phi], \tag{156}$$

or

$$\mathrm{d} \wedge \star\varphi = -ie^2 \star [(D\Phi^*)\Phi - \Phi^* D\Phi] \tag{157}$$

in terms of forms. In curved space η^{ik} has to be replaced by g^{ik}, but $F_{ik} = A_{k,i} - A_{i,k}$, corresponding to $\varphi = \mathrm{d}\alpha$. $A_{i;k} - A_{k;i}$ is gauge variant, when the space-time connection has torsion.

In the case of a *nontrivial* $U(1)$ bundle there is no global α, so φ is closed, but not exact. An example is the interaction of Φ with a magnetic monopole in flat space. In the rest-frame of a point-like monopole the fields are

$$\vec{E} = \vec{0}, \quad \vec{B} = g\,\frac{\vec{x}}{r^3}, \quad \text{e.g. } \varphi = eg\sin\vartheta\,\mathrm{d}\vartheta \wedge \mathrm{d}\phi. \tag{158}$$

For the four-potential in the neighbourhood of a monopole we make the axisymmetric ansatz

$$\alpha = a(r, \vartheta)\,\mathrm{d}\phi, \tag{159}$$

so that $\mathrm{d}\alpha = \frac{\partial a}{\partial r}\,\mathrm{d}r \wedge \mathrm{d}\phi + \frac{\partial a}{\partial \vartheta}\,\mathrm{d}\vartheta \wedge \mathrm{d}\phi$. From this follows $\frac{\partial a}{\partial \vartheta} = eg \sin\vartheta$ with the solution

$$a = eg\,(K - \cos\vartheta), \tag{160}$$

corresponding to the vector potential

$$\vec{A} = g\,\frac{K - \cos\vartheta}{r\,\sin\vartheta}\,\vec{e}_\phi. \tag{161}$$

$\vec{A}$ has singularities, the locations of which depend on the choice of K. A singularity *must* occur, otherwise the magnetic charge would be equal to zero:

$$4\pi eg = \int_{S^2} \varphi = \int_{\partial S^2} \alpha = 0.$$

Therefore for $g \neq 0$ the two-form φ is closed, but not exact.

The location of the singularity is without physical significance. For $K = \pm 1$, for example, there are two potentials $\alpha_\pm$ with singularities at the south/north pole of each sphere around the monopole. In the overlap they differ by a gauge transformation $\alpha_+ - \alpha_- = 2eg\,\mathrm{d}\phi = \mathrm{d}\Lambda$.

If we consider $i\alpha_\pm$ indeed as connection forms of a vector bundle, they must be related to two section bases $\{b_\pm\}$ which, when the bundle is restricted to S^2, are defined everywhere except the south/north pole. In the rest of S^2 the gauging $b_+ = e^{i\Lambda(x)}\,b_- = e^{2ieg\phi}\,b_-$ must be uniquely possible. Due to the non-uniqueness of the azimuth ϕ this is the case only for integer values of $2eg$. Including $\hbar$ and c, this leads to the Dirac quantization condition

$$\frac{2eg}{\hbar c} \in \mathbf{Z}. \tag{162}$$

The existence of one single magnetic monopole in the universe would give a reason for the discrete quantization of electric charge.

Example 3: M = space-time, $G = SU(2)$, $E = (M, \mathbf{V})$, $\mathbf{V} = \mathbf{C}^2$, the representation space of the fundamental representation of $SU(2)$. When introduced 1954 by Yang and Mills for the description of isospin, this was the first gauge theory on a vector bundle which is not the tangent bundle and has a non-abelian gauge group. The motivation was the following: The electromagnetic interaction left aside, the proton and the neutron are indistinguishable, they can be seen as two states of the "nucleon", related by "rotations" in an internal space, isomorphic to the space of states of a spin

1/2 particle. The strong interaction is invariant under "isospin rotations", like the electromagnetic interaction it should be mediated by a compensation field related to an $SU(2)$-covariant derivative.

The connection and curvature on the isospin bundle are the $su(2)$ elements

$$\omega = \vec{\omega}\,\vec{\tau}, \qquad \Omega = \vec{\Omega}\,\vec{\tau}, \tag{163}$$

where $\tau_i = -\frac{i}{2}\,\sigma_i$ and $\vec{\omega}$ and $\vec{\Omega}$ are triples of coefficients. Each curvature coefficient is expressed in the usual way by its corresponding connection coefficient,

$$\Omega^c = \mathrm{d} \wedge \omega^c + \frac{1}{2}\,\epsilon_{abc}\,\omega^a \wedge \omega^b, \tag{164}$$

sometimes written as

$$\vec{\Omega} = \mathrm{d}\vec{\omega} + \frac{1}{2}\,\vec{\omega} \wedge \vec{\omega}. \tag{165}$$

When the forms ω^a are expressed in terms of their space-time components ${A^a}_\mu\, e^\mu$ and analogously Ω^a in terms of ${F^a}_{\mu\nu}\, e^\mu \wedge e^\nu$, we get the relation between the Yang-Mills field strengths and their potentials (see (**??**))

$$\vec{F}_{\mu\nu} = \partial_\mu \vec{A}_\nu - \partial_\nu \vec{A}_\mu + [\vec{A}_\mu, \vec{A}_\nu], \tag{166}$$

with the free field equations (= Bianchi identities)

$$D \wedge \Omega = 0, \tag{167}$$

in components (see (**??**))

$$\mathrm{d}\vec{F} + [\vec{A}, \vec{F}] = 0. \tag{168}$$

The dynamics can be derived from a Lagrangian, whose expression in terms of forms is

$$L_{YM}\,\mathrm{d}^4x = \frac{1}{2}\,\mathrm{Tr}(\Omega \wedge \star\Omega) = \frac{1}{4}\,\delta_{ab}\,\Omega^a \wedge \star\Omega^b, \tag{169}$$

or, more generally,

$$L_{YM}\,\mathrm{d}^4x = f_{ab}\,\Omega^a \wedge \star\Omega^b. \tag{170}$$

The ensuing equations of motion with the isospin current one-form J are

$$D \wedge \star\Omega^a = (f^{-1})^{ab} \star J_b. \tag{171}$$

A Yang-Mills type theory with gauge group $G = SU(3) \times SU(2) \times U(1)$ describes the contemporary standard model of particle physics.

10 Is the theory of gravity a gauge theory?

Beginning in 1917, Levi-Civita, Weyl, Schouten separated the notion of a connection from the metric. 1922 É. Cartan laid the ground to the calculus of differential forms. So the question, whether or not in the theory of gravity the connection should be Levi-Civita, came up in a natural way. As the form calculus is coordinate-free, the issue of general covariance shifted from coordinate transformations to invariance under vierbein transformations. The metric can be conceived as an $O(3,1)$ structure on the tangent bundle. Is thus general relativity a gauge theory of the Lorentz group? If so, the independent field variables would be the six connection forms $\omega_{ik} = -\omega_{ki}$. This is not true, because ω_{ik} are derived from the four orthonormal basis forms e^i by $\mathrm{d} \wedge e^i = -\omega^i{}_k \wedge e^k$. So ω_{ik} cannot be varied independently. The action integral in terms of forms

$$W = \int \Omega_{ik} \wedge \star(e^i \wedge e^k) \tag{172}$$

contains beside Ω, which is derived from ω, explicitly e^i.

According to the Palatini variation principle, one can ignore the dependence of ω on e and vary independently with respect to ω_{ik} and e^i.

$$\delta W = \int \delta\Omega_{ik} \wedge \star(e^i \wedge e^k) + \int \Omega_{ik} \wedge \delta \star (e^i \wedge e^k) =$$

$$\int (D \wedge \delta\omega_{ik}) \wedge \star(e^i \wedge e^k) + \frac{1}{2} \int \Omega_{ik}\, \epsilon^{ik}{}_{lm}\, \delta(e^l \wedge e^m) =$$

$$\int \mathrm{d} \wedge (\delta\omega_{ik} \wedge \star(e^i \wedge e^k)) + \int \delta\omega_{ik} \wedge D \wedge \star(e^i \wedge e^k) + \int \Omega_{ik} \wedge \epsilon^{ik}{}_{lm} e^l \wedge \delta e^m$$

The resulting field equations are

$$\epsilon^{ik}{}_{lm}\, \Omega_{ik} \wedge e^l = 0 \qquad \text{(or equal to a source term)} \tag{173}$$

and

$$D \wedge \star(e^i \wedge e^k) = 0. \tag{174}$$

(**??**) are Einstein's equations in terms of forms, (**??**) are equivalent to

$$D \wedge (e^i \wedge e^k) = 0 \tag{175}$$

($\epsilon^{ik}{}_{lm}$ is covariantly constant) and $D \wedge e^i = \theta^i$ is the torsion. So (**??**) means

$$\theta^i \wedge e^k - \theta^k \wedge e^i = 0. \tag{176}$$

For $n \geq 4$ this is equivalent to

$$\theta^i = \mathrm{d} \wedge e^i + \omega^i{}_k \wedge e^k = 0. \tag{177}$$

This is the equation relating ω to e. Other Lagrangians and matter Lagrangians containing ω_{ik} (for example, the Dirac field Lagrangian) lead to nonvanishing torsion.

The Palatini principle does not a priori assume zero torsion, torsion is determined dynamically. If matter has zero spin density, there is no torsion and the Palatini principle is equivalent to the usual Einstein-Hilbert principle, spin however, if present, couples to torsion. The theory of gravity based on this principle is called the Einstein-Cartan theory.

Another approach is the assumption of a larger gauge group, containing e^i as connection components - the Poincaré group. In the five-dimensional matrix representation

$$(L, \vec{a}) \leftrightarrow \begin{pmatrix} L & \vec{a} \\ \vec{0} & 1 \end{pmatrix} \tag{178}$$

the Poincaré group appears isomorphic to a subgroup of $GL(5)$. The corresponding connection matrix is

$$\omega = \begin{pmatrix} \omega^i{}_k & \omega^l \\ \vec{0} & 0 \end{pmatrix}, \tag{179}$$

where the basis form e^l is denoted as ω^l. Gauge transformations act in the usual way on the connection, $\bar{\omega} = S\omega S^{-1} - (\mathrm{d}S)S^{-1}$, see equation (**??**), where S is of the form (**??**). We consider two cases

- $S \simeq (L, \vec{0})$. This is a Lorentz transformation of vierbeine for which the connection is an $O(3,1)$ connection,
- $S \simeq (\mathbf{1}, \vec{a})$ (a translation):

$$\bar{\omega}^i{}_k = \omega^i{}_k, \quad \bar{\omega}^i = \omega^i - \mathrm{d}a^i - \omega^i{}_k\, a^k. \tag{180}$$

In the second case $\omega^i{}_k$ do not react to gauge translations, but ω^i do react nontrivially. Can ω^i be identified with the basis forms e^i? For the latter ones no translational gauge transformation was defined, so the e^i can be identified with the connection forms belonging to the translation subgroup of the Poincaré group only in the case of linear independence and in a certain translation gauge. In other words, gauge covariance under the Poincaré gauge group is broken at the kinematic level and is valid only under the subgroup $O(3,1)$.

According to the two types of connection forms, there are the curvature forms ${\Omega^i}_k$ and Ω^i, forming the curvature matrix

$$\Omega = \begin{pmatrix} {\Omega^i}_k & \Omega^l \\ \vec{0} & 0 \end{pmatrix}. \tag{181}$$

Ω^i satisfies the structure equation for torsion,

$$\Omega^i = \mathrm{d} \wedge \omega^i + {\omega^i}_k \wedge \omega^k, \tag{182}$$

but $\Omega^i = \theta^i$ only in the gauge $\omega^i = e^i$. The Bianchi identities for the Poincaré curvature summarizes the Bianchi identities for the $O(3,1)$ curvature ${\Omega^i}_k$ and for torsion.

The behavior under translations is

$$\bar{\Omega}^i_k = {\Omega^i}_k, \qquad \bar{\Omega}^i = \Omega^i - {\Omega^i}_k \, a^k, \tag{183}$$

this means that translation invariance is broken also at the dynamical level, as the action integral contains e^i; it is broken even for Yang-Mills type integrals $\int \Omega \wedge \star \Omega$, because gauge translations act on the $\star$ operation. Gauge invariance under translations would correspond to a conserved Noether current of energy and momentum as generators of space and time translations, but as it is well-known, in general relativity there is no local energy-momentum conservation.

The probably most important application of gauge theory ideas to gravity is the formulation of Ashtekar variables. In this approach space-time is split into space and time to make a canonical formulation in terms of spatial field variables and conjugate momenta possible, the latter ones containing time derivatives. The canonical variables on a space manifold are orthonormal bases (dreibeine, triads) E and (partially independent) connection components A. In these variables the theory works as a gauge theory, the missing dependence of the connection on the metric is imposed by a set of additional conditions, the constraints. The local gauge group is $SU(2)$, the universal covering of the group of local dreibein rotations. The purpose for this reformulation of general relativity is that the canonical pairs (A, E) are suitable for quantization.