Symmetry and Lie groups in physics

Notes

John O'Neall

Table of Contents

1. Lie groups....2
2. Lie algebras....3
3. Irreducible groups and representations....5
4. Lie groups for physics....5
   4.1. Orthogonal groups....6
   4.2. Unitary groups....6
   4.3. 2-dimensional rotations – SO(2) and U(1)....7
   4.4. 3-dimensional rotations – SO(3) and SU(2)....7
      4.4.1. SU(2)....7
      4.4.2. SO(3)....8
      4.4.3. SU(2) and SO(3)....9
   4.5. Irreducible representations of SU(2)....9
   4.6. Lorentz transformations....10
      4.6.1. The (0,0) representation....12
      4.6.2. The (½,0) representation....12
      4.6.3. The (0,½) representation....13
      4.6.4. Weyl and Dirac spinors....13
      4.6.5. The (½,½) representation....14
   4.7. The Poincaré group....15
5. Finding Lagrangians....15
   5.1. Scalar particles....15
   5.2. Spin ½ particles – fermions....16
   5.3. Spin 1 particles....16
   5.4. Interaction Lagrangians – U(1) symmetry....17
   5.5. SU(2) and SU(3) gauge invariance....19
   5.6. Noether's theorem and currents....20
6. Bibliography....22

1. Lie groups

Overview: A ***Lie group*** is a continuous transformation group. Its elements may be transformations such as rotations or translations but not reflections, as these are not continuous. Since it is continuous, it can be generated incrementally from the identity operator by using operators called ***generators***. The generators are derivatives of Lie group elements with respect to their parameters and are members of a vector space called a ***Lie algebra***. The vector space of the infinitesimal generators of a Lie group forms a Lie algebra. The generators are especially interesting to physicists because they often represent physical observables such as momentum or angular momentum. They thus provide links between transformations and observables.

Somewhat more rigorously: A group is a set of transformations which is closed and invertible. A ***Lie group*** is a group which is also a differentiable manifold, meaning its elements are organized continuously and smoothly (as opposed to discrete groups). In addition, the group operator must induce a differentiable map of the manifold onto itself. (Every group element A induces a map that takes any element of the group B to another element of the group, C = AxB, and this map must be differentiable.)

In order to study groups of transformation on vector space, let's start simply:

is defined as representing the set of all linear operators on the vector space V. A linear operator is a function T from V to V itself which satisfies the linearity condition:[1]

T(cv + w) = cT(v) + T(w).

Now we will take subsets and subsets of subsets under certain conditions to reach the groups of interest in physics.

As a first subset of $\mathcal{L}(V)$ consider only those linear operators which are ***invertible***, defined as the ***general linear group*** of a vector space V, denoted by $GL(V)$[2].

$$GL(V) \subset \mathcal{L}(V).$$

Consider matrix representations of this group by assuming V has a scalar field C and is of dimension n. Then, according to whether C is real or complex, we get the ***real*** or ***complex general linear group in n dimensions***.[3]

$GL(V) \ \rightarrow \ GL(n, \mathbb{R})$ (real, n-dimensional field)

$GL(V) \ \rightarrow \ GL(n, \mathbb{C})$ (complex, n-dimensional field)

Going further, consider subgroups of these operating on vector spaces possessing a ***non-degenerate Hermitian form*** $(\cdot|\cdot)$, a function which assigns a scalar value to an ordered pair of vectors. Examples are the inner or scalar product, or the Minkowski metric. The set of ***isometries*** $Isom(V)$, consists of operators T which "preserve" $(\cdot|\cdot)$, meaning that

$$(Tv|Tw) = (v|w) \ \forall \ u, w \in V. \qquad (1.1)$$

If the condition

$(v|v) > 0$ for all $v \in V$, with $v \neq 0$ (1.2)

holds (it's positive-definite), then $(\cdot|\cdot)$ is called the ***inner product*** and the vector space is an ***inner-product space***.[4] If the inner product is interpreted as a length, an isometry preserves lengths. It is also a group. Now things are looking interesting.

So henceforth let's consider only isometries and apply them to three different vector spaces. These are all

1 Jeevanjee, 25.
2 Ibid, 118.
3 Ibid, 118. But what does a scalar field C mean here?
4 Ibid, 34.

subsets of n-dimensional general linear groups.[5]

1) If V is a ***real inner-product space***, then in an orthonormal basis, an isometric transformation operator T must obey

$$[T]^T[T] = 1 \quad \text{or} \qquad [T]^T = [T]^{-1}, \tag{1.3}$$

which is the ***orthogonality condition***. (The brackets indicate matrices. We could write the equation without them.)

2) If instead V is a ***complex inner-product space***, we find

$$[T]^\dagger = [T]^{-1} \tag{1.4}$$

where $[T]^\dagger$, the ***adjoint*** (or ***Hermitian conjugate***), is the transpose of the complex conjugate of the matrix. Such operators are termed ***unitary*** and in an orthonormal basis are represented by unitary matrices. Obviously, a unitary operator is an isometry of both a real and a complex inner-product space.

3) The vector space V may be a ***real vector space with a Minkowski metric*** η,

$$\eta(v_1, v_2) = t_1 t_2 - x_1 x_2 - y_1 y_2 - z_1 z_2,$$

which is a non-degenerate Hermitian form, but not an inner product, since it is not positive-definite. Then its group of isometries must satisfy

$$\Lambda^T \eta \Lambda = \eta, \tag{1.5}$$

where Λ is the 4-dimensional Lorentz transformation. In fact, this can be taken as a definition of the Lorentz transformations. The ***Lorentz group*** is then the set of all transformations which preserve the inner product of Minkowski space. It is denoted by ***O(1, n-1)*** in this convention[6], so ***O(1,3)***, in the 4-dimensional Minkowski space-time of Special Relativity (SR).

Summary: The matrix representations of isometric subgroups of the general linear group $GL(V)$, depending on V, are the orthogonal or unitary matrices, or the Lorentz transformations – $O(n)$, $U(n)$ and $O(1,3)$.

Both $O(n)$ and $U(n)$ have subgroups characterized as special, meaning simply that they contain only those matrices whose determinant is +1.

$$det(T) = +1. \tag{1.6}$$

They are called the ***special orthogonal group***, $SO(n)$, and the ***special unitary group***, $SU(n)$.

2. Lie algebras

Since it is continuous, a Lie group can be generated by infinitesimal operations starting from the identity elements

$$g(\epsilon) = I + \epsilon X, \tag{2.1}$$

where X is the ***generator***, because applying this multiple (infinite) times leads to Lie group elements

$$g(\theta) = e^{\theta X}. \tag{2.2}$$

The object X is a generator of the Lie group and therefore a member of the Lie algebra of the Lie group. So one can say in a more mathematically precise way that a Lie algebra is related to a Lie group as follows:

> For a given Lie group G (given by n x n matrices), the Lie algebra $\mathfrak{g}$ of G is given by those n x n matrices X such that $e^{\theta X}$ is a member of G for all real θ. For symbol lovers, the Lie algebra $\mathfrak{g}$ of a given matrix Lie group $G \subset GL(n, \mathbb{C})$ is

5 I think...

6 Jeevanjee uses the unusual (to me) convention of O(n-1,1), putting time at the end, for some reason beyond my fathoming.

$$\mathfrak{g} = \{X \in M_n(\mathbb{C}) \mid e^{\theta X} \in G \,\forall\, \theta \in \mathbb{R}\},$$

where $M_n(\mathbb{C})$ is the group of $n \times n$ matrices with complex entries.

From equation (2.2), it is clear that the generator, X is just the derivative evaluated at $\theta = 0$:

$$\frac{dR}{d\theta}\bigg|_{\theta=0} = Xe^{\theta X}\bigg|_{\theta=0} = X. \qquad (2.3)$$

Alternatively, from (2.1) and (2.4), one can "see" (and derive) that a Lie algebra is the ***tangent space*** to the group at the identity.

As an example, consider the case of SO(2), rotations in 2 dimensions:

$$SO(2) = \left\{ \begin{pmatrix} \cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \right\} \bigg| \theta \in [0, 2\pi).$$

Then

$$\frac{dR}{d\theta}\bigg|_{\theta=0} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = X, \qquad (2.4)$$

is the generator, a 2x2 matrix. For the starting vector $r_0 = (1, 0)$,

$$Xr_0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

which is a vector pointing in the direction of the change in r_0 under a rotation. It is evident that this works also for r at (0,1), (-1,0) or (0,-1) and slightly less evident for arbitrary angles.

A more abstract, mathematical definition of a ***Lie algebra***[7] is a vector space $\mathfrak{g}$ together with a binary operator, called the ***Lie bracket***, $[,] : \mathfrak{g} \; x \; \mathfrak{g} \to \mathfrak{g}$. The binary operator satisfies the following conditions:

- Bilinearity: [aX + bY, Z] = a[X,Z] + b[Y,Z] and [Z, aX + bY] = a[Z,X] + b[Z,Y] , for arbitrary numbers a,b and such that X,Y,Z ϵ $\mathfrak{g}$.
- Anticommutativity: [X,Y] = -[Y,X] $\forall$ X,Y ϵ $\mathfrak{g}$
- The Jacobi identity: [X,[Y,Z]] + [Z,[X,Y]] + [Y,[Z,X]] = 0 $\forall$ X,Y ϵ $\mathfrak{g}$

In words, again, a Lie algebra is composed of all n-dimensional complex matrices such that the exponential of such a matrix times a real number is a member of the corresponding Lie group. The matrices of the Lie algebra are then the generators of the Lie group through this exponential. In addition, the Lie algebra is associated with an operation, called the ***Lie bracket [,]***, which tells us how to combine these matrices. In other words, a Lie algebra is "closed under commutators"[8]: If X and Y are elements of a Lie algebra, then so is

[X,Y] = XY – YX.

One can derive a general description of a Lie algebra of isometries from (1.1) by using $T = e^{\theta X}$ to write it as

$$(Tv|Tw) = (e^{\theta X}v|e^{\theta X}w) = (v|w).$$

Differentiating with respect to T and then setting $\theta = 0$ and re-arranging a bit yields

$$(Xv|w) = -(v|Xw) \;\forall\; v.w \in V. \qquad (2.5)$$

This is the ***fundamental equation of a Lie algebra of isometries***. But look out, physicists don't define it quite this way, as is shown in section 4.2.

7 Schwichtenberg, PS, 45.
8 Ibid., 155.

3. Irreducible groups and representations

A topological space is ***simply connected*** if it is path-connected (not disjoint) and every path between any two points can be continuously transformed into any other path between those points. This means that the space can contain no holes. There is an important fact, which we just have to accept:

There is only one simply-connected Lie group corresponding to each Lie algebra.

This "mother" group to potentially many other Lie groups sharing the same Lie algebra is called the ***covering*** group and is said to ***cover*** the other groups.[9]

Put the other way around, a covering group is the unique simply-connected Lie group corresponding to a given Lie algebra. Any other Lie group which might correspond to this Lie algebra is not simply connected; it is said to be covered by the covering group – group to group.

A ***representation*** can be thought of loosely as an instantiation of an abstract group in a vector space. More rigorously, a representation is a map between any group element g of a group G and a linear transformation R(g) of some vector space

$$g \to R(g)$$

in such a way that the group properties are preserved. Note that it is the identification of each point of an abstract group manifold with a linear *transformation* of a vector space – an operation, not a point in the vector space. In physics especially, it is convenient to represent the generators by a set of matrices. A representation is ***irreducible*** when it is a representation of a group G on a vector space V which has no invariant subspace besides the zero space {0} and V itself. Different representations of a group may have different dimensions.

A ***Casimir element*** C is built from elements of the Lie algebra in such a way that it commutes with every generator X of the group.

$$[C, X] = 0. \qquad (3.1)$$

Schur's lemma then says that it must be a multiple of the identity, so Casimir elements provide linear operators with a constant value for each representation and so can be used to *label* the representation. An example is J^2 for rotations.

Within a Lie algebra, the set of diagonal generators is called the ***Cartan subalgebra***. Just as the Casimir element provides a label for each representation, the Cartan elements provide labels within a representation. Examples will follow.

The general method of study will be

1) start with an example of a group, e.g., 2x2 matrices and rotation;
2) derive the Lie algebra;
3) use the Lie algebra to look at different representations and search for a simply-connected one, the ***covering group***. [HUH??]

4. Lie groups for physics

In ordinary n-dimensional space, we are interested in transformations which conserve distances between two points – isometries. So orthogonal (real) and unitary (complex) transformations are important. The most interesting Lie groups for physics are $\mathbb{Z}_1$, $\mathbb{Z}_2$, $SO(n)$ and $SU(n)$. In particular, $U(1)$, $SU(2)$ and $SU(3)$ apply respectively to the EM, weak and the strong forces.

In addition, SR requires the invariance of the interval (the Minkowski metric)

$$s^2 = t^2 - x^2 - y^2 - z^2,$$

9 One is tempted to call the covering group the mother group and the Lie algebra the father group, all the covered groups being their offspring. Jeez.

which is a kind of distance between two points. So we add the O(1,3) group to the list of important groups for physics.

4.1. Orthogonal groups

Since an ***orthogonal*** transformation is a linear transformation on a real vector space V that preserves inner products, it is therefore the group of isometries on V and is defined, as we have seen, by the condition,

$$O^T O = I \tag{4.1}$$

and a ***special orthogonal group*** by (4.1) and

$$det(O) = +1. \tag{4.2}$$

The first condition guarantees the conservation of lengths

$$x_1^2 + x_2^2 + x_3^2 + ...$$

and the second keeps only rotations (not reflections), a rotation being a continuous linear operator which takes orthonormal bases to orthonormal bases.[10]

Of course (4.1) for the group element is consistent with (2.5) for the generator as the former means

$$e^{\theta X^T} e^{\theta X} = 1 \Rightarrow J^T + J = 0.$$

So we see that the Lie algebra of O(n) is the set of $n \times n$ antisymmetric matrices.

4.2. Unitary groups

A ***unitary*** transformation is similar to an orthogonal transformation in that it preserves inner products, but in a complex vector space. So, as an isometry, a unitary group's adjoint $U^\dagger$ must satisfy

$$U^\dagger U = 1. \tag{4.3}$$

A ***special unitary group***, also satisfies

$$det(U) = +1, \tag{4.4}$$

Every isometry of a complex inner product space is unitary, and vice versa. In an orthonormal basis, a unitary operator is represented by a unitary matrix. It is clear by (4.1) and (4.3) that a unitary operator is an isometry of both real and complex inner product spaces.[11]

Again, from (2.5), this is equivalent to ***generators***

$$X^\dagger = -X.$$

<u>But</u> physicists do not define the group in terms of the generator as $g = e^{\theta X}$. but include a *factor I* in the exponent:

$$g = e^{i\theta X},$$

so that the fundamental equation for isometries yields

$$X^\dagger = X$$

and the matrices are ***Hermitian*** and so can represent physical observables.

10 Jeevanjee, 125.
11 Jeevanjee, 120-121.

4.3. 2-dimensional rotations – SO(2) and U(1)

The unitary group $U(1)$ is just multiplication by a phase $e^{i\theta}$ and represents rotation by an angle θ.

Note that rotations in two dimensions are just the unit circle, S^1, and can be represented by 2-d matrices which are elements of either $SO(2)$or $U(1)$. From (4.1) and (4.2), the ***SO(2)*** group can be represented by rotations in terms of the sine and cosine of the angle of rotation.

$$\mathrm{R}_\theta = \begin{pmatrix} \cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}. \quad (4.5)$$

However, constant distances in 2-d can also be represented by unit complex numbers in 1-d. This is the ***U(1)*** group. Such a number is represented by

$$R_\theta = e^{i\theta} = cos(\theta) + isin(\theta). \quad (4.6)$$

One can map (4.6) to a real matrix by using the 2-d identity matrix and

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \text{ and } \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (4.7)$$

to show that

$$R_\theta = cos(\theta)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + sin(\theta)\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix},$$

which is identical to the equation for 2-d SO(2), showing that there is an ***isomorphism*** between the two groups U(1) and SO(2).

4.4. 3-dimensional rotations – SO(3) and SU(2)

The conditions (4.1) and (4.2) for an orthogonal group can be satisfied by 3x3 matrices which are simple extensions of equation (4.5), forming a representation of SO(3).

4.4.1. SU(2)

Just as SO(2) rotations could be represented in terms of complex numbers by U(1), we would like to describe SO(3) in terms of a unitary group. In this case, we must use 4-d complex numbers called ***quaternions***, as there are no 3-d complex numbers.[12] The unit-length constraint reduces the 4 degrees of freedom to three, as needed for 3-d rotations. We can extend the idea of the complex number from one to four dimensions by defining

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$$

and

$$\mathbf{ijk} = -1.$$

Quaternions can be written as 2x2 matrices in terms of the basis vectors

$$\mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \mathbf{i} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \mathbf{j} = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} \text{ and } \mathbf{k} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}. \quad (4.8)$$

Then a *unit* quaternion may be written as

$$q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$$

with

$$q^\dagger q = 1$$

12 Schwichtenberg, PS, 2018, 33. Why not?

and

$$det(q) = a^2 + b^2 + c^2 + d^2 = 1.$$

In this way, the set of unit quaternions can be written as 2x2 matrices which satisfy

$$U^\dagger U = 1 \qquad \text{and} \qquad det(U) = 1 \tag{4.9}$$

which fulfills the requirements (4.3) and (4.4) for a unitary group, in this case, $SU(2)$, because of (4.9).

In order to associate this with the rotation of a 3-d vector in space, we set a=0 and identify only the imaginary part with the 3-d vector

$$v = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}. \tag{4.10}$$

Keeping the imaginary and real parts separate under arbitrary rotations requires transformations of the type

$$v' = qvq^{-1}.$$

With this requirement satisfied, unit quaternions can indeed describe 3-d rotations by equation (4.10). And this gives us a 2-d unitary representation of rotations in three dimensions.[13]

Let's look at its Lie algebra. Starting with its expression in terms of generators iJ_i

$$U = e^{iJ_i}$$

and applying requirements (4.9), now requires that

$$J_i^\dagger = J_i,$$

meaning J_i is Hermitian (the reason for the factor of I in the exponent). The second of equations (4.9) requires a zero trace. So the generators of SU(2) must be traceless Hermitian matrices. A possible basis in terms of 2x2 matrices is the trio

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{4.11}$$

which are the ***Pauli matrices***. These satisfy the commutation relation

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \tag{4.12}$$

so if we define the generators as $J_i \equiv \frac{1}{2}\sigma_i$, we have

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{4.13}$$

which expresses the ***Lie algebra*** $\mathfrak{su}(2)$ ***of SU(2)***.

4.4.2. SO(3)

Looking for the Lie algebra of SO(3), we know that every element O of the group can be written in terms of a generator (member of the associated Lie algebra) as

$$O = e^{i\theta J}, \tag{4.14}$$

where the factor i ($= \sqrt{-1}$) is included to make things nicer later on.

Putting this together with equations (4.1) and (4.2) gives

$$J^T + J = 0 \quad \text{and} \quad tr(J) = 0, \tag{4.15}$$

which, after including the "physicist's i", can be satisfied by the following basis vectors:[14]

13 But I can't see that we have discussed the Lie algebra for SU(2) yet.

14 Schwichtenberg, PS, 44.

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (4.15)$$

These are basis vectors for the generators of the group SO(3) and so elements of $\mathfrak{so}(3)$, its Lie algebra, the vector space of 3x3 antisymmetric matrices. They can be written compactly using the Levi-Civita symbol:

$$(J_i)_{jk} = -i\epsilon_{ijk},$$

which satisfies the Lie brackets

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \quad (4.16)$$

This equation *defines* a set of transformations on a vector space and thereby the *Lie algebra* for SO(3), $\mathfrak{so}(3)$, by defining the value of the Lie bracket $[J_i, J_j]$. It is the same as equation (4.13), the Lie algebra $\mathfrak{su}(2)$ of SU(2). So SO(3) and SU(2) have the same Lie algebra.

Now we would like to know which, if either, is the covering group for these two Lie groups.

4.4.3. SU(2) and SO(3)

One can show that a unit vector in SO(3) corresponds to two different unit vectors in SU(2). It turns out that a rotation of a unit quaternion by an angle θ corresponds to a rotation of the corresponding vector by an angle 2θ, so for a single complete rotation of the unit quaternion, the vector goes around twice. Stated the other way around, to a given rotation angle of the vector there correspond two different rotation angles of the unit quaternion. For example, a rotation of the unit quaternion by either $\pi/2$ or $3\pi/2$ corresponds to a rotation of the vector by π. More precisely: For the map

$$\rho : SU(2) \to SO(3),$$

which is a ***homomorphism***[15], "… for every $R \epsilon SO(3)$ there correspond exactly two matrices in SU(2) which map to R under ρ."[16]

Now SU(2) is the three-sphere, S^3, a 3-dimensional "spherical" space embedded in four dimensions. So it and indeed all Lie groups can be considered multi-dimensional spaces. Since the three-sphere is easily seen to be simply connected, SU(2) is the ***covering group*** for SO(3). SU(2) is said to be the ***double-cover*** of SO(3), which is seen as half of SU(2). In a sense, SU(2) is more complete than SO(3).

We have now used methodology steps 1) through 3), deriving the Lie algebra from an example (twice, in fact) of a group and then using that to identify the covering group. Now we can use the Lie algebra to consider other representations.

All these unit-length conserving groups are in fact the same as[17] S^n, the n-spheres:

- S^1 corresponds to U(1) and SO(2),
- S^3 corresponds to SU(2) and so to half of SO(3), SU(2) being the fundamental or covering group.

It would be handy to refer to them as such, S^n, but history has decided otherwise.

4.5. Irreducible representations of SU(2).

The irreducible representations are the ones of particular interest. We have seen that since SU(2) is equivalent (isomorphic) to S^3, the three-sphere, it is *the* simply-connected group corresponding to this Lie algebra and so is its covering group.

We can build one Casimir element for this Lie algebra, i.e., which commutes with every generator in the group,

15 Jeevanjee, 176.
16 Ibid, 140.
17 "Same as" is too vague. Are they in fact homomorphisms or isomorphisms?

the J_i, and that is

$$J^2 = J_1^2 + J_2^2 + J_3^2$$

in three dimensions. From the definitions of the J_i it is easy to show that

$$J_{3-d}^2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 2I,$$

twice the identity matrix. In two dimensions

$$J_{2-d}^2 = \begin{pmatrix} 3/4 & 0 \\ 0 & 3/4 \end{pmatrix} = \frac{3}{4}I.$$

We see that the representations can be labeled by the Casimir operator values, 2 or ¾.

Similarly, there is one Cartan element, a diagonal operator, which we usually take as J_3, labeling the element by its eigenvalue m.

Following standard QM methods, we can define two other operators

$$J_+ = J_1 + iJ_2 \qquad \text{and} \qquad J_- = J_1 - iJ_2$$

and show that they are the usual ladder operators used, e.g., for angular-momentum. So we finish with the representations for SU(2) shown in the following table. The representations are labeled by j, the number associated with the Casimir operator, and the different elements by m, the eigenvalues of the Cartan operator, J_3.

Dimension	j (J^2)	m (J_3)	n-sphere	Casimir eigenvalues
1	0	0		j(j+1) = 0
2	1/2	-1/2 , 1.2	S^1	j(j+1) = 3/4
3	1	-1,0,1		j(j+1) = 2
4	3/2	-3/2, -1/2,1/2, 3/2	S^3	j(j+1) = 15/4

Table 1. Representations of SU(2).

4.6. Lorentz transformations

As already noted (1.5), a Lorentz transformation Λ must conserve the Minkowski metric η and so must satisfy

$$\Lambda^T \eta \Lambda = \eta. \qquad (4.17)$$

From this alone one can deduce that the determinant of Λ must be ± 1 and also that

$$\Lambda_0^{\ 0} = m\sqrt{1 + \sum_i (\Lambda_i^{\ i})^2},$$

meaning $\Lambda_0^{\ 0}$ is either $\geq +1$ or ≤ -1. These two constraints together give four combinations of which only one can be generated from the identity element by infinitesimal transformations, as required by a Lie algebra. This is the so-called ***proper orthochronous Lorentz group***, also referred to as the ***restricted Lorentz group***, represented by the symbol $\Lambda_+^\uparrow$, The word "proper" here refers to the +1 value of the determinant and orthochronous means that the direction of time is not changed. The other three sub-categories can be reached by parity and time-reversal transformations of the restricted Lorentz group, so the entire Lorentz group may be represented by

$$O(1,3) = \{\Lambda_+^\uparrow, \Lambda_P\Lambda_+^\uparrow, \Lambda_T\Lambda_+^\uparrow, \Lambda_P\Lambda_T\Lambda_+^\uparrow\}.$$

Rotations are then simply 3-d rotations tucked into the Minkowski spatial part:

$$\Lambda_{rot} = \begin{pmatrix} 1 & 0 \\ 0 & R_{3x3} \end{pmatrix}$$

with generator

$$J_i = \begin{pmatrix} 1 & 0 \\ 0 & J_i^{3d} \end{pmatrix}. \qquad (4.18)$$

For a boost along the x-axis, the *generator*[18] defined by

$$\Lambda^\mu{}_\rho = \delta^\mu{}_\rho + \epsilon K^\mu{}_\rho$$

must satisfy (4.17), which leads to

$$K^T\eta = -\eta K. \qquad (4.19)$$

Following the method of (4.18), we write the general generator

$$X = \begin{pmatrix} X_{00} & a \\ b & X^{3d} \end{pmatrix}$$

and plug it into (4.17) to show the general form for a generator of O(1,3) to be

$$X = \begin{pmatrix} 0 & a \\ a & X^{3d} \end{pmatrix}, \text{ with } X^{3d} \in \mathfrak{o}(3), a \in \mathcal{R}^{n-1}.$$

Now we can use our previous knowledge of the Lorentz transformations to deduce the form of a generator of a boost along the x-axis, which does not affect the y or z coordinates:

$$K_1 = \begin{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} & \\ & \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{pmatrix}.$$

Plug this into (4.18) to find one way to write the first of the following set of generators, the others being found similarly.

$$K_1 = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}. \qquad (4.20)$$

A general restricted Lorentz transformation is then of the form

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta}+i\vec{K}\cdot\vec{\phi}}, \qquad (4.21)$$

with the commutation properties

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \qquad [J_i, K_j] = i\epsilon_{ijk}K_k, \qquad [K_i, K_j] = -i\epsilon_{ijk}J_k.$$

Note that the third of these commutators says two boosts are equivalent to a rotation. Combine these generators linearly by defining two *complexified* generators

18 Schwichtenberg, PS, 68.

$$N^{\pm} = \frac{1}{2}(J_i \pm iK_i), \tag{4.22}$$

so the commutation relations become

$$[N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+, \quad [N_i^-, N_j^-] = i\epsilon_{ijk}N_k^-, \quad [N_i^+, N_j^-] = 0. \tag{4.23}$$

Now, N_i^+ and N_i^- each satisfy the commutation relations (4.16) for the Lie algebra of SU(2). So the complexified[19] Lie algebra for the restricted Lorentz group consists of *two copies* of $\mathfrak{su}(2)$, the Lie algebra for SU(2), which is therefore the covering group for the restricted Lorentz group. We can label the irreducible representations of each of the two SU(2) groups by its Casimir variable j, as in Table 1. Denoting the representations by (j^+, j^-), we have the (0,0) representation, the (½,0) representation, the (0,½) representation and the (½,½) representation, to name only those of particular interest in physics.

We have now found that ***SU(2)*** is the ***covering group both*** for the special orthogonal group ***SO(3)*** and for the restricted Lorentz group $\mathbf{\Lambda}_+^{\uparrow}$. A very important group indeed!

4.6.1. The (0,0) representation

From Table 1, j=0 for a 1-dimensional representation. This means every matrix is just 1 so nothing changes and this is the scalar representation.

4.6.2. The (½,0) representation

Use the Pauli matrices (4.11) as basis matrices, so

$$N_i^+ = \frac{1}{2}\sigma_i.$$

From the definitions of N_i^+ and N_i^-, one finds

$$J_i = \frac{1}{2}\sigma_i, \qquad K_i = \frac{-i}{2}\sigma_i$$

so a general transformation is given by

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta} + i\vec{K}\cdot\vec{\phi}} = e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2} + \vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \tag{4.24}$$

The factors of ½ show that this is the representation of the double cover of the Lorentz group. Note that these are in terms of complex 2×2 matrices. These act on two-component objects called – heads up, here! – ***left-chiral spinors***:

$$\mathcal{X} = \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix}. \tag{4.25}$$

Also, note that equation (4.24) is a matrix equation and the operator is defined by its Taylor series

$$e^M = \sum_{n=0}^{\infty} \frac{M^n}{n!}.$$

Using the Pauli matrices properties, including

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{and} \qquad \sigma_3^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

the expansion for a boost along the z-axis gives

19 Complexification due to the factor I in equation (4.22).

$$B_3(\phi) = e^{\phi\frac{\sigma_3}{2}} = 1 + \frac{1}{2}\phi\sigma_3 + \frac{1}{2}\Big(\frac{1}{2}\phi\sigma_3\Big)^2 + ... = \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix}, \qquad (4.26)$$

which makes clear the operator is a 2x2 matrix. Similar treatment of a rotation leads to almost the usual matrix in terms of sine and cosine of the rotation angle.

$$R_3(\theta) = e^{i\theta\frac{\sigma_3}{2}} = 1 + \frac{1}{2}i\theta\sigma_3 + \frac{1}{2}\Big(\frac{i}{2}\theta\sigma_3\Big)^2 + ... = \begin{pmatrix} cos(\frac{\theta}{2}) & isin(\frac{\theta}{2}) \\ isin(\frac{\theta}{2}) & cos(\frac{\theta}{2}) \end{pmatrix}, \qquad (4.27)$$

after using

$$i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This shows that spinors are not normal vectors in spacetime because, after a rotation by 2π, a spinor changes by a factor of -1.

4.6.3. The (0,½) representation

By a calculation similar to that of the last paragraph, we can find for this representation that J_i has the same value, but not K_i.

$$J_i = \frac{1}{2}\sigma_i, \qquad K_i = \frac{i}{2}\sigma_i,$$

so a general transformation is given by

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta}+i\vec{K}\cdot\vec{\phi}} = e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}-\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \qquad (4.28)$$

Again, this is a 2×2 matrix representation which acts on similar but different two-dimensional objects called ***right-chiral spinors***:

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}. \qquad (4.29)$$

Right-chiral spinors transform just like left-chiral spinors by (4.27). But under boosts, they transform by

$$B_3(\phi) = e^{-\phi\frac{\sigma_3}{2}} = 1 - \frac{1}{2}\phi\sigma_3 + \frac{1}{2}\Big(\frac{1}{2}\phi\sigma_3\Big)^2 + ... = \begin{pmatrix} e^{-\frac{\phi}{2}} & 0 \\ 0 & e^{\frac{\phi}{2}} \end{pmatrix}.$$

Left and right-chiral spinors transform differently by a negative sign under boosts along a given direction. A parity transformation, or mirroring, in the same direction, reverses the sign in the same way. So a right-chiral spinor mirrored along a given direct transforms under a boost along the same direction (and by a rotation) like a left-chiral one, and vice versa.[20]

The generic name for both types of spinors is ***Weyl spinors***.

4.6.4. Weyl and Dirac spinors

Spinors are new and rather strange beasts. They do not live in the 4-d Minkowski space of SR, but in ***two-dimensional spinor space***. Equations (4.26) and (4.27) show clearly that left-chiral spinors do not transform like 4-vectors and neither do right-chiral spinors. Since they live in their own space, ***spinor space***, they have their own ***spinor metric***,

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad (4.30)$$

20 Schwichtenberg, QFT, 88.

which is not the Minkowski metric of space-time but is used similarly, e.g., to raise or lower indices

$$\mathcal{X}^a = \epsilon^{ab}\mathcal{X}_a,$$

which may be taken as the definition of a superscripted spinor. The spinor metric can be used to construct the inner product

$$\mathcal{X}_a\mathcal{X}^a = \mathcal{X}_a\epsilon^{ab}\mathcal{X}_b,$$

which is Lorentz invariant (by (4.26) and (4.27)). In addition, terms like

$$\xi^\dagger\mathcal{X} = \xi_b\epsilon^{ba}\mathcal{X}_a$$

are then also invariant.

Let's change our notation a bit, paying special attention to the dots[21] on the super- and subscripts, so that undotted spinors are left-chiral and dotted ones right-chiral:[22]

$$\mathcal{X} := \mathcal{X}_a \qquad \text{and} \qquad \xi := \xi^{\dot{a}}$$

Then define the complex conjugates as inverting the chirality:

$$(\mathcal{X}_a)^\dagger = \mathcal{X}_{\dot{a}} \quad \text{and} \qquad (\xi^{\dot{a}})^\dagger = \xi^a,$$

which makes the notation consistent.[23]

We can also define two new spinor quantities in terms of left and right spinors

$$\mathcal{X}_L^C = \epsilon\mathcal{X}_L^* \quad \text{and} \qquad \xi_R^C = -\epsilon\xi_R^*,$$

which transform “backwards” under a boost, i.e., $\mathcal{X}_L^C$ transforms like a right-chiral spinor and ξ_R^C transforms like a left-chiral spinor. This is taken to indicate the operation commonly called ***charge conjugation*** (reversal):

$$\Psi = \begin{pmatrix}\mathcal{X}_L\\ \xi_R\end{pmatrix} \to \Psi^C = \begin{pmatrix}\xi_R^C\\ \mathcal{X}_L^C\end{pmatrix} = \begin{pmatrix}\xi_L\\ \mathcal{X}_R\end{pmatrix},$$

although in fact what it does is transform a left-chiral into a right-chiral.

Let's spell out the results in all their grisly detail:

$$\mathcal{X}' = \mathcal{X}_a' = \Lambda_a^{\ b}\mathcal{X}_b = (e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}+\vec{\phi}\cdot\frac{\vec{\sigma}}{2}})X_b,$$

for left-chiral spinors and

$$\xi' = \xi'^{\dot{a}} = \Lambda^{\dot{a}}_{\ b}\xi^b = (e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}-\vec{\phi}\cdot\frac{\vec{\sigma}}{2}})\xi^b,$$

for right-chiral spinors. So finally, the spinor Lorentz transformation operators are as follows:

$$\Lambda_{(\frac{1}{2}.0)} = (e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}+\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}) := \Lambda_a^{\ b} \tag{4.31}$$

$$\Lambda_{(0,\frac{1}{2})} = (e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2}-\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}) := \Lambda^{\dot{a}}_{\ b} \qquad (4.32).$$

Once again, remember that, because of the σ terms, these are matrix equations. Again, note the dot on the superscript a on the last term of the second equation.

Under a parity transformation, the generator J_i is unchanged (angular momentum being a pseudovector), but K_i changes sign. This means that, from equation (4.22),

$$N_i^+ \leftrightarrow N_i^-.$$

So in order to maintain the validity of equations under a parity transformation, we need to have both a left-chiral and a right-chiral spinor. The solution adopted is to use a ***Dirac spinor***

21 Sometimes indistinguishable from fly spots. It's time now for us older scientists to get out the ol' magnifying glass.
22 Called the Van der Waerden notation. Why do we bother with all this?
23 Dixit Schwichtenber, NNQFT, 99-100.

$$\Psi = \begin{pmatrix} \mathcal{X} \\ \xi \end{pmatrix},$$

where a parity transformation gives

$$\Psi = \begin{pmatrix} \mathcal{X} \\ \xi \end{pmatrix} \to \Psi' = \begin{pmatrix} \xi \\ \mathcal{X} \end{pmatrix}.$$

In spite of appearances a Dirac spinor is a 4-dimensional object, being composed of two 2-d spinors. Its four components exist in spinor space, not 4-d spacetime, and it transforms like two spinors, *not* like a 4-vector, It is said to be in the

$$(\tfrac{1}{2}, 0) \oplus (0, \tfrac{1}{2})$$

representation. In particular, a boost along the z-axis has the formalism

$$B_3^D(\phi) = \begin{pmatrix} e^{\frac{\phi}{2}} & 0 & 0 & 0 \\ 0 & e^{-\frac{\phi}{2}} & 0 & 0 \\ 0 & 0 & e^{-\frac{\phi}{2}} & 0 \\ 0 & 0 & 0 & e^{\frac{\phi}{2}} \end{pmatrix} = \begin{pmatrix} \Lambda_{(\frac{1}{2},0)} & 0 \\ 0 & \Lambda_{(0,\frac{1}{2})} \end{pmatrix},$$

which does not look at all like the Lorentz transformation of a 4-vector in terms of hyperbolic functions.

4.6.5. The (½,½) representation

An object in this representations has two indices, each one transforming under its own 2-dimensional copy of the Lie algebra $\mathfrak{su}(2)$. Using the Pauli matrices plus the identity as basis 2x2 matrices, such an object can be written as:

$$v_{ab} = v_\nu \sigma^\nu_{ab} = v_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + v_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + v_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Going through the separate transformations shows that the resulting transforms could equally well be expressed by a 4-vector formalism, such as this boost along the z-axis:

$$\begin{pmatrix} v'_0 \\ v'_1 \\ v'_2 \\ v'_3 \end{pmatrix} = \begin{pmatrix} cosh(\phi) & 0 & 0 & sinh(\phi) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ sinh(\phi) & 0 & 0 & cosh(\phi) \end{pmatrix} \begin{pmatrix} v_0 \\ v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

So the (½, ½) representation represents vectors and we can use the simpler vector matrix algebra. In fact, a 4-vector is a ***rank 2 spinor.*** Whereas a rank-2 tensor has two vector indices, we see now that a vector has two spinor indices. So *spinors are more fundamental than 4-vectors*, since 4-vectors are not appropriate for describing all physical systems on a fundamental level (electrons, for instance). Hence the (approximate) expression that a vector is the square root of a rank-2 tensor and a spinor is the square root of a vector.

4.7. The Poincaré group

The full space-time symmetry group of nature is the ***Poincaré group***, the Lorentz group plus translations, or rotations plus boosts plus translations. Infinitesimal translations are represented by the generators

$$P_i = -i\partial_i.$$

The Poincaré group has two Casimir operators, of which the more useful is

$$P_\mu P^\mu := m^2$$

(Yes, that m.). Also, the number $j_1 + j_2$ of the representation (j_1, j_2) can be used as a label and is called the spin, so the $(0,0)$ representation is the spin 0 representation; the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$, spin 1\2; the $(\frac{1}{2}, \frac{1}{2})$, spin 1.

Standard notation, at least for us, uses Φ for scalar fields, Ψ for spinor fields and A for vector fields.

5. Finding Lagrangians

One of the goals of these studies of groups is to derive proper Lagrangians for the systems of physics. Another is to identify behavior of intrinsic properties of particles like spin, isospin or color, and to infer the existence of mediator particles, or gauge fields. Although no reason is known for it, aside from the fact that the results give the correct equations of motion, two general principles of Lagrangian construction must be observed:[24]

1. The Lagrangian may only contain the lowest non-trivial derivatives, meaning first or second order. Sometimes the second order is necessary in order to maintain Lorentz invariance.
2. For free fields or particles, we must stop at second order in the field.

The central requirement is, of course, that the Action be Lorentz invariant, which will be satisfied if the Lagrangian is Lorentz invariant (although this is not a necessary condition), i.e., if it is a scalar.

We must consider several cases.

5.1. Scalar particles

Scalar fields transform according to the (0,0) representation of the Lorentz group. A Lagrangian (density) obeying the above rules would be of this form:

$$\mathcal{L} = A\Phi^0 + B\Phi + C\Phi^2 + D\partial_\mu\Phi + E\partial_\mu\Phi\partial^\mu\Phi + F\Phi\partial_\mu\Phi + G\Phi\partial_\mu\partial^\mu\Phi.$$

Consider the various terms:

- The A term is just a constant and so has no effect on the equations of motion.
- Odd powers of ∂_μ are forbidden, so the D term goes out. Apparently, the F term also.
- The B term can be ignored as it becomes a constant after use of the Euler-Lagrange equations and so changes nothing physical.
- After integration by parts, assuming the fields go to zero at infinity, a term like $G\Phi\partial_\mu\partial^\mu\Phi$ would be just like the F term and so is redundant.

This leaves us with only

$$\mathcal{L} = C\Phi^2 + E\partial_\mu\Phi\partial^\mu\Phi. \tag{5.1}$$

Another approach to Lorentz invariance could start with the SR dispersion relation

$$E^2 = p^2 + m^2 \to -\hbar^2\nabla^2 + m^2.$$

This suggests that we do something like take the square of the Schrödinger equation. If we do that and put in the QM operator for the momentum, the result looks like (5.1), which gives us an identification for the variables C and D. Therefore, we can say the Lagrangian for scalars is

$$\mathcal{L} = \frac{1}{2}\big(\partial_\mu\partial^\mu\Phi - m^2\Phi\big). \tag{5.2}$$

From this Lagrangian, the Euler-Lagrange equations give

$$\big(\partial_\mu\partial^\mu + m^2\big)\Phi = 0, \tag{5.3}$$

the ***Klein-Gordon equation***.

24 Schwichtenberg, PS, 97-98.

5.2. Spin ½ particles – fermions

By similar but somewhat more laborious reasoning based on Lorentz invariance of Dirac spinors, we find the Lagrangian for the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation to be of the form

$$\mathcal{L} = A\Psi^\dagger\gamma_0\Psi + B\Psi^\dagger\gamma_0\gamma^\mu\partial_\mu\Psi = A\bar{\Psi}\Psi + B\bar{\Psi}\gamma^\mu\partial_\mu\Psi.$$

With A = -m and B = I, this gives

$$\mathcal{L} = -m\bar{\Psi}\Psi + i\bar{\Psi}\gamma^\mu\partial_\mu\Psi = \bar{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi \qquad (5.4),$$

the ***Dirac Lagrangian***.

5.3. Spin 1 particles

From the $(\frac{1}{2}, \frac{1}{2})$ representation, similar considerations lead to a general invariant form

$$\mathcal{L}_{Proca} = C_1\partial^\mu A^\nu\partial_\mu A_\nu + C_2\partial^\mu A^\nu\partial_\nu A_\mu + C_3A^\mu A_\mu + C_4\partial^\mu A_\mu,$$

where A is an arbitrary, as yet unidentified vector field. We can ignore the 4th term, as it does not affect the equations of motion. Passing all this through the Euler-Lagrange equations gives

$$2C_3A^\rho = 2C_1\partial_\sigma\partial^\sigma A^\rho + 2C_2\partial^\rho(\partial_\sigma A^\sigma).$$

Adjusting the constants ($C_1 = -C_2 = 1/2, C_3 = m^2$) gives the ***Proca equation***:

$$m^2A^\rho = \partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma). \qquad (5.5)$$

For a photon, with spin 1 and mass = 0, this gives

$$\partial_\sigma(\partial^\sigma A^\rho - \partial^\rho A^\sigma) = 0,$$

which is the inhomogeneous Maxwell equation without electric currents, justifying the choice of constants. This is commonly written in terms of the electromagnetic field tensor

$$F^{\sigma\rho} := \partial^\sigma A^\rho - \partial^\rho A^\sigma \qquad (5.6)$$

which gives

$$\partial_\sigma F^{\sigma\rho} = 0.$$

The Lagrangian for massless spin 1 is then

$$\mathcal{L}_{Maxwell} = \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu) = \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \qquad (5.7)$$

For a massive spin 1 field,

$$\mathcal{L}_{Proca} = \frac{1}{2}(\partial^\mu A^\nu\partial_\mu A_\nu - \partial^\mu A^\nu\partial_\nu A_\mu) + m^2A_\mu A^\mu = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + m^2A_\mu A^\mu. \; (5.8)$$

5.4. Interaction Lagrangians – U(1) symmetry

Symmetry considerations are not only valid for finding the form of free-particle Lagrangians, they can also be used to find interaction terms. It is obvious that a unitary transformation

$$\Psi \rightarrow \Psi' = e^{i\alpha}\Psi$$

does not change the Dirac Lagrangian. Since this transformation does not affect spacetime, it represents an ***internal symmetry*** of the system, say an electron. It clearly affects the field everywhere in spacetime and so is a ***global*** transformation. But SR and its speed limit forbid such a global transformation, so let's try a local transformation by letting α depend on the coordinates. Then

$$\Psi \to \Psi' = e^{i\alpha(x)}\Psi \quad \text{and} \qquad \bar{\Psi} \to \bar{\Psi}' = \bar{\Psi}e^{-i\alpha(x)}. \tag{5.9}$$

Now the derivative in the Lagrangian causes an extra term under the transformation and

$$\mathcal{L}'_{Dirac} = \mathcal{L}_{Dirac} - (\partial^{\mu}a(x))\bar{\Psi}\gamma_{\mu}\Psi, \tag{5.10}$$

so the Dirac spin-½ Lagrangian by itself is not invariant under local U(1) transformations.[25] The extra term should remind us of the Dirac 4-current from QFT, $\bar{\psi}\gamma^{\mu}\psi$, multiplied by a partial derivative.

Similarly, the Proca Lagrangian for a massless spin-1 particle has a global symmetry when

$$A_{\mu} \to A'_{\mu} = A_{\mu} + a_{\mu}.$$

If we make the transformation local, by making a_{μ} depend on the coordinates, it is not a symmetry of the massless Proca Lagrangian. But if a_{μ} is replaced by the derivative (gradient) of an arbitrary function $\alpha(x)$, it is indeed a symmetry.

$$A_{\mu} \to A'_{\mu} = A_{\mu} + \partial_{\mu}\alpha(x). \tag{5.11}$$

(This is of course the usual gauge symmetry of the electromagnetic field.) If we define the transformation

$$T_{\alpha}(A^{\mu}) = A^{\mu}(x) - \partial^{\mu}\alpha(x), \tag{5.12}$$

with the map Φ from U(1) to T_{α}

$$e^{i\alpha(x)} \to T_{\alpha}. \tag{5.13}$$

Then the identity $\alpha(x) = 0$ means

$$T_0(A^{\mu}) = A^{\mu} - \partial^{\mu}0 = A^{\mu}$$

and

$$\begin{aligned} T_{\alpha+\beta}(A^{\mu}) &= (A^{\mu} - \partial^{\mu}(\alpha(x) + \beta(x)) = (A^{\mu} - \partial^{\mu}\alpha(x)) - \partial^{\mu}\beta(x) = T_{\beta}(T_{\alpha}(A^{\mu})) \\ &= (T_{\beta} \otimes T_{\alpha})(A^{\mu}) \end{aligned} \tag{5.14}$$

just as

$$\Phi(e^{i\alpha(x)}e^{i\beta(x)}) = \Phi(e^{i\alpha(x)+i\beta(x)}) = T_{\alpha+\beta} = T_{\alpha} \otimes T_{\beta} = \Phi(e^{i\alpha(x)}) \otimes \Phi(e^{i\beta(x)})$$

Equation (5.14) shows that (5.11) is a homomorphism to U(1)and so a representation of the U(1) group for the Proca equation.[26]

Hoping that the partial derivatives in (5.10) and (5.11) are related, let's consider an object with one foot in each representation and see how it transforms under (5.11).

$$A_{\mu}\bar{\Psi}\gamma^{\mu}\Psi \to (A_{\mu} + \partial_{\mu}a(x))\bar{\Psi}\gamma^{\mu}\Psi = A_{\mu}\bar{\Psi}\gamma^{\mu}\Psi + \partial_{\mu}a(x)A_{\mu}\bar{\Psi}\gamma^{\mu}\Psi,$$

which contains a second term which looks a lot like the extra term in (5.10) only multiplied by an additional factor A_{μ}. Since this new term looks contains elements of both the Dirac and Proca (vector) Lagrangians, It requires also the free part of the vector field if the equation is to be complete. Let's anticipate the result some and make an interaction Lagrangian from the Dirac and the massless Proca equations plus the additional interaction term, with a coupling constant g thrown in because it will be handy later.

$$\mathcal{L}_{Dirac+Proca+int} = \bar{\psi}(i\gamma_{\mu}\partial^{\mu} - m)\psi + qA_{\mu}\bar{\psi}\gamma^{\mu}\psi - \frac{1}{2}(\partial^{\mu}A^{\nu}\partial_{\mu}A_{\nu} - \partial^{\mu}A^{\nu}\partial_{\nu}A_{\mu}). \tag{5.15}$$

Now the Lagrangian is invariant under local SU(1) transformation, which requires both the Dirac and the Proca transforms, (5.9) and (5.11), applied simultaneously. The first term in (5.15) is the Dirac Lagrangian; the second, the Dirac 4-current multiplied by the gauge (vector) potential; and the third, the Maxwell Lagrangian for a massless particle.

25 Be careful. Some texts use $\Psi \to \Psi' = e^{-i\alpha(x)}\Psi$, which changes the sign of the interaction term.

26 Thanks to Gaussian97 at Physics Forums, https://www.physicsforums.com/threads/symmetry-of-qed-interaction-lagrangian.995118/. Homomorphisms are explained by Jeevanjee, 138 and 188.

We use the derivative of a function to pass from its value at one point in spacetime to a point at an infinitesimal distance from there. If however, the coordinates themselves are changing, we need to know more, specifically, how things, such as basis vectors, change as a function of the change in the coordinates. Such objects are called ***connections*** (not to be confused with the Christoffel connections of GR, even those have a logically similar function). Equation (5.11) looks like such a change of coordinates. A connection may be defined in terms of a so-called covariant derivative which is just the ordinary derivative plus the connection term. Here, we can define a ***covariant derivative***

$$D^\mu = \partial^\mu - igA^\mu. \tag{5.16}$$

On the one hand, this is just a notational device to make the equation simpler: It avoids our writing the interaction term specifically. On the other, it tells us we need to know more than just the derivative to know how a locally-invariant field changes from one place to another. For some reason, it is called the ***minimal coupling rule***.

Now the Lagrangian is the correct ***Lagrangian for the quantum field theory of electrodynamics***, QED:

$$\begin{aligned}\mathcal{L}_{Dirac+Proca+int} &= \bar{\Psi}(i\gamma_\mu D^\mu - m)\Psi - \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu)\\ &= \bar{\Psi}(i\gamma_\mu D^\mu - m)\Psi - \frac{1}{4}(F^{\mu\nu}F_{\mu\nu}). \qquad (5.17)\end{aligned}$$

All QED (as well as classical EM, of course) comes from Lorentz and U(1) invariance + a bit of imagination. We could write the equation using the covariant derivative for the components of the vector potential also, but the differences cancel out. Note that equation (5.15) and both variants of (5.17) represent the same physical system.

In quite a similar way, one can construct the locally gauge-invariant Lagrangians for a massive charged scalar field with a massless vector field and for a massive vector field with a massless one.[27] So locally gauge-invariant Lagrangians are available for the interaction of a scalar, Dirac or vector field with a massless vector field.

5.5. SU(2) and SU(3) gauge invariance

Recall from section 5.4 that in order to ensure the local gauge invariance under U(1) symmetry for a spin-½ particle, two other things were required:[28]

- the introduction of a massless vector (spin-1) field, which we take to represent a photon, including its free Lagrangian, and
- addition of an interaction term depending on both types of fields (or, equivalently, the use of the the covariant derivative), as well as the transformation properties of the vector field (the same ones used for the vector and scalar potentials of classical electromagnetism).

This was expressed in equations (5.11) through (5.17).

We can go further, although the calculations are more laborious.

In order to study local gauge invariance for two equal-mass Dirac fields, we express them as one two-component column vector, so under U(2) symmetry. This then requires not two, but three massless vector fields. It also requires adding a term (a cross product) to the equation for $F^{\mu\nu}$. Here, the equal-mass requirement makes the identification of the fields as particles difficult without the notion of symmetry breaking.[29]

And lo, behold, requiring local gauge invariance for three equal-mass Dirac fields, expressed as a three-component column vector and so under U(3) symmetry, requires the addition of eight massless vector fields. Here the three Dirac particles are accepted to be three ***quarks*** of the same ***flavor*** (and so mass) but different ***colors*** (red, blue, green) and the vector fields to be the ***gluons*** of the strong interaction force. Each gluon forms the source for a color current, in the sense of a Noether current.[30] Voilà ***quantum chromodynamics***.

27 Schwichtenberg, PS, 142-3.
28 Griffiths, 360.
29 Ibid, 365.
30 Ibid, 368/

In this case, the state vector[31] representing three equal-mass particles is

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \end{pmatrix}, \quad \bar{\psi} = \begin{pmatrix} \bar{\psi}_1 & \bar{\psi}_2 & \quad \bar{\psi}_3 \end{pmatrix}$$

so the Lagrangian looks just like the one-particle case but is in fact a three-component column vector each element of which is a four-component Dirac spinor. (Got that?) The symmetry group of this beast is U(3) with

$$U^\dagger U = 1,$$

so U may be written in terms of a Hermitian matrix H

$$U = e^{iH}$$

which in turn may be expressed in terms of nine real numbers a_i and θ and the eight 3x3 Gell-Mann matrices λ as

$$H = \theta \mathbf{1} + \lambda \cdot \mathbf{a},$$

so

$$U = e^{i\theta} e^{i\lambda \cdot a}.$$

The first part is U(1) and the equation expresses $U(3) = U(1) \otimes SU(3)$. So finally we want to transform the Lagrangian in such a way that it is invariant under local SU(3) gauge transformations.

$$\psi \to S\psi, \qquad \text{with} \qquad S \equiv e^{-iq\lambda \cdot \phi(x)}.$$

Using a covariant derivatives

$$\mathcal{D}_\mu \equiv \partial_\mu + iq\lambda \cdot A_\mu$$

where there are now eight gauge fields A_μ and we want the transformation to function as

$$\mathcal{D}_\mu \psi \to S(\mathcal{D}_\mu \psi).$$

Then in the infinitesimal case, this yields a formula equivalent to ()

$$\vec{A}'_\mu \cong \vec{A}_\mu + \partial_\mu \phi + 2q(\vec{\phi} \times \vec{A}_\mu)$$

with the cross product being

$$(\vec{B} \times \vec{C}) = \sum_{j,k=1}^{8} f_{ijk} B_j C_k$$

where f_{ijk} are the structure constants of SU(3). Finally, the complete Lagrangian for ***chromodynamics*** is

$$\mathcal{L} = i\bar{\psi}(\gamma^\mu \partial_\mu - m)\psi - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} - (q\bar{\psi}\gamma^\mu \lambda \psi) \cdot A_\mu.$$

This all comes from insisting that the Lagrangian be locally invariant under SU(3). The Dirac fields provide eight color currents

$$J^\mu \equiv q(\bar{\psi}\gamma^\mu \lambda \psi)$$

which constitute sources for the color fields A_μ. Equation () is the correct one for the strong interaction, one for each of the six quark flavors.[32]

So in the cases of U(1), U(2) or U(3), one can either write the interaction term or use a covariant derivative with an addition connection term.

---

31 Or whatever it is called...

32 All this derivation based on – almost copied from – Griffiths, EP, 366-369.

Klauber sums up an amazing general rule for QFT:[33]

> If we start with the free Lagrangian and require it to be locally symmetric, then it can only be so if we add to it the particular interaction term(s) that actually describe(s) interactions in the real world.

He adds that local symmetry is essential for renormalization: No gauge invariance, no QFT.

In addition, gauge invariance of the QED Lagrangian requires the photon to have zero mass.

5.6. Noether's theorem and currents

Consider a Lagrangian density $\mathcal{L}(\phi_i, \partial_\mu \phi_i)$ whose field variables ϕ_i undergo infinitesimal global transformations $\delta\phi_i$ which are functions of a parameter α. Use of the Euler-Lagrange equations and integration by parts shows that

$$\delta\mathcal{L} = \partial_\mu \Big(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \Big) \delta\phi_i.$$

If the Lagrangian is invariant under change of a parameter α, then

$$\frac{\partial \mathcal{L}}{\partial \alpha} = 0 \qquad \text{means that} \qquad \partial_\mu J^\mu = 0 \tag{5.18}$$

the latter being the equation of continuity for the ***Noether current*** defined by

$$J^\mu(\phi_i, \partial_\mu \phi_i) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} \frac{\delta \phi_i}{\delta \alpha}. \tag{5.19}$$

Equation (5.18) shows that the current has zero four-divergence $\partial_\mu J^\mu = 0$, so its zeroth component, J^0, integrated over all of space is conserved.

Note that in case of several fields ϕ_i, the RHS of equation (5.19) is to be summed over i. The derivation of this current does not depend on whether α is a constant, so Noether's theorem is equally *valid for global and local transformations*.

Applying this to the Dirac electron shows the current concerned is

$$J^\mu = -g\bar{\Psi}\gamma^\mu\Psi.$$

This is therefore the electric four-current, of which the 0th component is the electric charge. So U(1) symmetry leads to the ***conservation of electric charge***.

If, instead of fields, we had taken the space-time Lagrangian $\mathcal{L}(q_i, \dot{q}_i)$ for a particle and followed a similar procedure, we would have found the following current:

$$J_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} + G,$$

where G is a function whose total time derivative changes the Lagrangian but not the Action. It is usually zero. (In spacetime, this would be the four-divergence in place of the time derivative.)[34] Applying this with various transformations leads to beloved conservation laws, as summarized in the table.

33 Klauber, 296.
34 Blundell & Lancaster, 93.

Transformation	Current	Conserved quantity
Spatial translation $q_i \to q_i + \delta q_I$	$\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q = p \delta q$	Momentum p
Spatial rotation $q_i \to q_i + \epsilon_{ijk} q_j a_k$	$(\vec{p} \times \vec{q}) \cdot \vec{a} = \vec{L} \cdot \vec{a}$	Angular momentum $\vec{L}$
Time translation $t \to t' = t + \epsilon$	$p\dot{q} - \mathcal{L} = \mathcal{H}$	Energy $\mathcal{H}$ (Hamiltonian)
Boost $q_i \to q_i + vt$, so $\dot{q}_i \to \dot{q}_i + v$	$pt - mq$	Uh...

Table 2. Conserved quantities of spacetime transformations

The boost current depends on t, which can be picked so that the current is zero, which is therefore conserved. Uh... ok.

Spacetime transformations of fields are messier because changes come from the fields themselves as well as from the transformed coordinates. The results, though, are consistent with the preceding table. In particular, spacetime translations lead to a tensor current

$$T_\mu^{\ \nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\nu \Phi)} \frac{\partial \Phi}{\partial x_\mu} - \delta_\mu^{\ \nu} \mathcal{L}, \tag{5.20}$$

which looks a lot like the definition of the Hamiltonian in terms of the conjugate momentum and the Lagrangian and so is taken as the ***energy-momentum tensor***.[35] This results in the conservation of energy and each of the three components of momentum. For boosts, an extra term is added to the angular momentum and this is taken to be the spin, but that must wait for more specific considerations (fermions).

The following table summarizes similar results for fields if equation (5.19) is used to calculate the Noether current due to a U(1) transformation on scalar and Dirac fields.

Type of field	Current	Conserved quantity
Free scalar (Klein-Gordon)	$j^\mu = i((\partial_\mu \phi)\phi^\dagger - (\partial_\mu \phi^\dagger)\phi)$	Charge
Free spinor (Dirac)	$j^\mu = \bar{\Psi}\gamma^\mu \Psi$	Charge
Free photon (Proca, m=0)	$j^\mu = 0$	Charge = 0
Full QED Lagrangian (5.15)	$j^\mu = \bar{\Psi}\gamma^\mu \Psi$	Charge

Table 3. Conserved quantities of Noether currents

The results found here using Noether's theorem are the same as those found by the standard QM method starting with the Schrödinger equation. The result for the QED interaction equation is as expected, since photons are chargeless. The proof for the full QED Lagrangian does not depend on the spatial dependence of α in the exponent of the transformation and so is equally true for *local and global* transformations.

Note that all these fields are ***gauge fields***, since they are a function of the exponent in the U(1) transformation, even if that be zero.

Equation (5.19) applied to changes only in the fields, not the coordinates, represents internal symmetries of the system. In particular, for a translation of the field,

35 Schwichtenber, PS, 110.

$$\Phi_i \to \Phi_i + \delta\Phi_i,$$

not of spacetime, it shows conservation of a new quantity, the ***conjugate momentum density***,

$$\pi_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\nu \Phi)}. \tag{5.21}$$

This must be distinguished from the physical momentum density of the field, which is due to invariance under spatial translations.

6. Bibliography

Blundell, Stephen J. and Lancaster, Tom. *Quantum field theory for the gifted amateur*. Oxford: Oxford University Press, 2014.

Griffiths, David. *Introduction to elementary particles*. Second, revised edition. Weinheim: Wiley-VCH, 2008. Print.

Jeevanjee, Nadir. An introduction to tensors and group theory for physicists. 2nd edition. Heidelberg: Springer, 2015. Print.

Klauber, Robert D. *Student friendly quantum field theory.* Fairfield: Sandtrove Press, 2013. Print.

Schwichtenberg, Jakob.[NNQFT] *No-nonsense quantum field theory*. No-nonsense books, 2020. Print.

Schwichtenberg, Jakob. [PF] *Physics from finance.* No-nonsense books, 2020. Print.

Schwichtenberg, Jakob. [PS] *Physics from symmetry*, 2nd edition. Springer, 2018. Print.

Susskind, Leonard and Friedman, Art. *Special relativity and classical field theory: The theoretical minimum*. New York: Basic Books, 2017. Print.