

Symmetry, groups and quantum field theory

Notes

John O'Neill

Note to the reader (assuming there is one...).

This document is just an overview of the subject, composed essentially of notes I have made to remind myself of the math behind the main principles of modern physics. I should say, the mathematical and physical concepts which *are* the main principles of modern physics. No attempt to teach anything or to indicate any real calculations of cross sections or whatever. It presents, for me, the mathematics of physics and the *concepts* you need to know in order to understand the principles presented in simpler books, such as for laymen.

*Note: The Minkowski **metric signature** used will be $(+1, -1, -1, -1)$. Among authors cited, only Robinson and Susskind use the opposite. See Table 2.*

Licensed under a Creative Commons Attribution 4.0 International License.



Table of Contents

1. Lie groups.....	5
1.1. Really short preview.....	5
1.2. Groups and representations.....	5
2. Generators and Lie algebras.....	8
2.1. Lie algebras.....	8
2.2. Covering groups and subalgebras.....	11
2.3. Summary – The basics of Lie groups.....	12
3. Lie groups for physics.....	12
3.1. Orthogonal groups.....	13
3.2. Unitary groups.....	13
3.3. 2-dimensional rotations – $SO(2)$ and $U(1)$	14
3.4. 3-dimensional rotations – $SO(3)$ and $SU(2)$	14
3.4.1. $SU(2)$	14
3.4.2. $SO(3)$	16
3.5. Irreducible representations of $SU(2)$ – spin.....	16
3.6. Lorentz transformations.....	18
3.6.1. The Lorentz algebra.....	19
3.6.2. The $(0,0)$ representation.....	20
3.6.3. The $(\frac{1}{2},0)$ representation -- spinors.....	21
3.6.4. The $(0,\frac{1}{2})$ representation.....	22
3.6.5. Geometric picture of spin.....	23
3.6.6. Weyl and Dirac spinors and chirality.....	25
3.6.7. The $(\frac{1}{2},\frac{1}{2})$ representation.....	27
3.7. The Poincaré group.....	28
4. Finding Lagrangians from symmetry.....	29
4.1. Scalar particles.....	29
4.2. Spin $\frac{1}{2}$ particles – fermions.....	30
4.3. Spin 1 particles.....	30
5. Field theory of free fields.....	31
5.1. Why QFT? Problems with QM and SR.....	31
5.2. Some reminders.....	31
5.2.1. Action and the Euler-Lagrange equation.....	32
5.2.2. Fourier transforms.....	32
5.2.3. Harmonic oscillators and ladder operators.....	33
5.2.4. Angular momentum.....	35
5.3. A wee bit of philosophy.....	35
5.4. Interlude: a word on Fock spaces.....	36
5.5. The ubiquity of simple harmonic oscillators.....	37
5.6. Action and energy in field theory.....	41
5.7. EM field Lagrangian and Maxwell's equations.....	44
5.8. The Schrödinger and Heisenberg pictures.....	45
5.9. Second quantization.....	46

5.10. Relativistic scalar Lagrangian (Klein-Gordon).....	49
5.10.1. Classical relativistic scalar Lagrangian.....	49
5.10.2. Complexification.....	50
5.10.3. Convert fields to operators by imposing commutation relations.....	51
5.10.4. From parameters to wave functions to fields to	52
5.10.5. Number operators, Hamiltonian and energy.....	52
5.10.6. Creation and annihilation operators.....	53
5.10.7. Normal ordering.....	54
5.10.8. Résumé of two quantizations.....	54
5.10.9. Klein-Gordon particles are bosons.....	56
5.10.10. Continuity, currents and negative energy.....	56
5.10.11. Comparison to a SHO.....	57
5.10.12. The Feynman propagator.....	57
5.11. Spin- $\frac{1}{2}$ particles – the Dirac equation.....	58
5.11.1. Spin, helicity and chirality.....	60
5.11.2. Interlude – negative energies.....	62
5.11.3. Solutions to the Dirac equation.....	63
5.11.4. The Dirac equation in QFT.....	65
5.11.5. Feynman propagator.....	66
5.12. Spin-1 particles – photons.....	66
6. Symmetries and the Standard Model.....	69
6.1. The standard model (SM).....	69
6.2. The Core Theory.....	72
6.3. Noether's theorem and currents.....	72
6.4. Isospin.....	75
6.5. Gauge symmetry.....	76
6.6. Interaction Lagrangians – U(1) symmetry and QED.....	77
6.7. SU(n) gauge invariance.....	81
6.8. SU(2) gauge invariance.....	82
6.9. SU(3) symmetry and QCD.....	83
6.10. Physical results of gauge symmetry.....	86
6.11. Spontaneous symmetric breaking and the Higgs mechanism.....	86
6.12. Electroweak interactions.....	91
6.12.1. La recette.....	91
6.12.2. La cuisine.....	92
6.13. Summary – Forces and particles.....	96
6.14. So what's the big deal?.....	97
6.15. Construction of SU(3) multiplets – the Eightfold Way.....	98
7. Scattering.....	106
7.1. The interaction picture.....	106
7.2. The S-matrix and ordering.....	108
7.3. Contractions and Wick's theorem.....	109
7.4. Putting it all together – the Dyson series.....	112
7.5. The final step – Feynman diagrams.....	113
8. Path integral formulation.....	115
9. Renormalization.....	118

10. Physics and finance.....	121
11. Review and summing up.....	123
11.1. Really short summary.....	123
11.2. Lie groups and algebras – Forces and particles.....	123
11.3. Examples.....	124
11.4. QFT.....	124
11.5. Why second quantization?.....	125
12. Annex: Math symbols.....	125
13. Bibliography.....	126

I hope that these notes will help you (well, me, at least) to understand how the simultaneous requirements of classical and Quantum Mechanics, Special Relativity and symmetry (Lie groups) lead via second quantization (and with help from the mathematics of simple harmonic oscillators) to QFT, the physics of multiple particles and interactions between them. Two starting points, to show where we are headed:

- “The world is made of quantum fields.”
- QFT is part of quantum mechanics, not something separate.

1. Lie groups

Let's start with a super-brief overview, to get an idea of where we are going.

1.1. Really short preview

For physicists, 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. Such groups are abstract things, but we can study specific examples, called **representations**. Since the Lie group and therefore the representation are continuous, the latter can be generated incrementally from the identity operator by using operators called **generators**. The generators are derivatives of Lie group elements (in a representation) and form a vector space called a **Lie algebra**. The generators are especially interesting to physicists because they represent physical observables such as energy or linear or angular momentum. They thus provide links between transformations and observables. This allows us to deduce properties of the world around us from properties of mathematical groups, which are thus a window onto the world.¹

If you don't find the coming overview brief enough, you can skip over to section 2.3 on page 12. Otherwise, let's see more details. Take a deep breath...²

1.2. Groups and representations

A **group**, call it $(G, *)$, is a set, G , of abstract elements, g_i , and some operation, $*$, on them subject to certain conditions:³

- closure: $\forall g_1, g_2 \in G, g_1 * g_2 \in G$
- associativity: $\forall g_1, g_2, g_3 \in G, (g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$
- identity: $\exists g \in G$, denoted e , $\mid \forall g \in G, e * g = g * e = g$
- inverse: $\forall g \in G, \exists h \in G \mid g * h = h * g = e$

The number of elements is the **order**, or dimension, of the group. It may be infinite.

A **Lie group** is a group which is also a differentiable manifold, meaning that its elements are organized continuously and smoothly (as opposed to elements of discrete groups). For closure, the group operator must induce a differentiable map of the manifold onto itself. Every group element A induces a map that takes any element B to another element $C = A * B$, and this map must be differentiable.

¹ I think Pythagoras would have loved this!

² The following presentation of Lie groups is based mostly on Jeevanjee and Robinson.

³ Math symbols explained in annex 12.

Being continuous, a Lie group is **parametrized** by one or more continuous variables, their number being the **order** of the group.

There is some inconsistency in the literature. Many authors do not use the term order, which can lead to confusion between the dimension of the vector space of objects acted on and that of the elements of the group. The order, or dimension, of the group is in general not the same as the dimension of the vector space. The group $SO(2)$ operates on a 2-dimensional vector space but has only one element, the angle of rotation in the plane, so is of group order 1. The group of permutations of three objects, a discrete (non-continuous) group, contains six elements (the permutations) acting on only three objects.

A **representation** can be thought of as an instantiation of an abstract group, representing the elements of the Lie group (or algebra) as operators on a linear vector space. The objects transformed are then the elements of the vector space. More rigorously, a representation is a map between any abstract *element* g of a group G and a linear *transformation* $R(g)$ of some vector space

$$R : g \rightarrow D_R(g)$$

in such a way that the group properties are preserved:

- $D_R(e) = 1$,
- $D_R(g_1)D_R(g_2) = D_R(g_1 * g_2)$.

Note that it is the identification of each element of an abstract group manifold with a linear *transformation* of a vector space – an operation in the vector space. The elements of the representation tell how to transform the vector space, for instance, to rotate it or give it a boost. In physics especially, it is convenient to represent the transformations 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.⁴ Different representations of a specific group may have different dimensions.

In order to study groups of transformations on a vector space V , let's start simply.

$$\mathcal{L}(V)$$

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:⁵

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

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

The most general group will take any point in V to any other point in V . If V has dimension n , then the group elements can be represented by $n \times n$ matrices and these must be non-singular, or **invertible**. Such a group is the *largest* and most general Lie group in n dimensions. This group is defined as the **general linear group** of a vector space V , denoted by $GL(n)$ ⁶.

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

Then, according to whether the elements are real or complex, we have the **real** or **complex general linear group in n dimensions**.⁷

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

⁴ We are leaving out an awful lot here, such as the definition of an invariant subspace.

⁵ Jeevanjee, 25.

⁶ Jeevanjee, 118; Robinson, 76.

⁷ Jeevanjee, 118. The scalar field C has only its transformation properties which are concerned by the group.

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

If the determinant of each element is +1, these groups become the **special linear groups**

$$SL(n, \mathbb{R}) \quad \text{or} \quad SL(n, \mathbb{C}).$$

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.⁸ Examples are the inner or scalar product, and the Minkowski metric. The set of **isometries** $Isom(V)$, consists of operators T which “preserve” $(\cdot|\cdot)$, meaning that

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

If the condition

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

holds (it's **positive-definite**), then $(\cdot|\cdot)$ is called the **inner product** and the vector space is an **inner-product space**.⁹ If the inner product is interpreted as a length, then an isometry can be seen as preserving lengths. It is also a group. Now things are looking interesting.

From here on, let's consider only isometries and apply them to three different vector spaces. These are all subsets of n-dimensional general linear groups.¹⁰ Since the conditions on each reduce the number of degrees of freedom and so the number of independent variables, these groups will not have the same order, or group dimension, as the dimensions of the vector space where they operate.

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

$$T^T T = 1 \quad \text{or} \quad T^T = T^{-1}, \quad (1.3)$$

which is the **orthogonality condition**. (Since the T are matrices, 1 must be the identity matrix.) This is easily seen, as a transformation of an inner product takes place as follows.

$$v'^2 = v T^T T v = v^2.$$

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

$$T^\dagger T = 1 \quad \text{or} \quad T^\dagger = T^{-1} \quad (1.4)$$

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 ,

$$\eta(v, w) = v_0 w_0 - v_1 w_1 - v_2 w_2 - v_3 w_3 = v^T \eta w,$$

which is a non-degenerate Hermitian form, but not an inner product, since it is not positive-definite. It could have any number of negative and positive terms. Of course, we are interested in the 4-dimensional spacetime case (1,4). Then its group of isometries must satisfy

$$\eta(v, w) = \eta(\Lambda v, \Lambda w) = v^T \Lambda^T \eta \Lambda w,$$

so that

⁸ Or a one-form operating on a vector ...?

⁹ Jeevanjee, 34.

¹⁰ I think...

$$\Lambda^T \eta \Lambda = \eta, \quad (1.5)$$

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 metric of Minkowski space*.

It is denoted by **O(1, n-1)** in this convention¹¹, so **O(1,3)** is the 4-dimensional Minkowski spacetime of Special Relativity (SR).

Summary: The matrix representations of isometric (distance-preserving) subgroups of the general linear group $GL(V)$, acting on the n-dimensional vector space V, are the orthogonal and unitary matrices, and the Lorentz transformations – $O(n)$, $U(n)$ and $O(1, 3)$. In general, the parameter n is not the order, or dimension, of the group. We will see that $SO(3)$ has a dimension (order) of 3, but $SO(2)$ of only one.

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

$$\det(T) = +1. \quad (1.6)$$

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

2. Generators and Lie algebras

The generators of a Lie group make up the Lie algebra of the group.

2.1. Lie algebras

The transformation corresponding to a Lie group is specified by a number (order) of parameters. For instance, rotation in 2-dimensional real space, $SO(2)$, is specified by a single parameter, the angle of rotation, and so is of order 1. So the number of parameters is not necessarily the dimension of the space acted upon by the group. Since it is continuous, a *representation* of a Lie group can be generated by infinitesimal operations starting from the identity element. So taking each parameter $\delta\theta_a$ as an infinitesimal increment of parameter θ_a , close to the identity,

$$D(\theta_a) = I + i\delta\theta_a T^a, \quad (2.1)$$

where the factor i is included so the **infinitesimal generator** T^a in the representation will be Hermitian. Note that as soon as we talk about generators, we are referring to a representation of a Lie group, not the abstract group. Applying a generator multiple (infinite) times leads to Lie group elements in the representation:

$$D(g(\theta)) = e^{i\theta_a T^a}. \quad (2.2)$$

Here, θ_a is a parameter, and T^a is an infinitesimal generator of the Lie group and a member of the Lie algebra of the Lie group (definition coming). There is one generator per parameter, so the number of generators (the group order) is constant for a given group, even though the dimension of a representation and so that of *each matrix* T^a in the representation may vary. You may soon tire of hearing this repeated, but it is important to distinguish between the dimension of the group (the order) and the dimension of the representation, which may vary. Examples of representations of different dimensions of the group $SU(2)$ are shown in Table 1 on p.17.

¹¹ Jeevanjee uses the unusual (to me) convention of $O(n-1,1)$, putting time at the end, for some reason beyond my fathoming.

From equation (2.2), we see that the generator, T^a is basically the derivative evaluated at $\theta_a = 0$:

$$T^a = -i \frac{\partial D}{\partial \theta_a} \Big|_{\theta_a=0}. \quad (2.3)$$

From (2.1) and (2.3), one can “see” that a Lie algebra is the **tangent space** to the group at the identity.¹²

As an example, consider the case of SO(2), rotations in 2 dimensions, with only a single parameter, the angle:

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

Then

$$-i \frac{dD}{d\theta} \Big|_{\theta=0} = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} = T_R, \quad (2.4)$$

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

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

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 only slightly less evident for arbitrary angles.

Formally, a **Lie algebra**¹³ is a vector space \mathfrak{g} together with a binary operator, called the **Lie bracket**, $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$. 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 for $X, Y, Z \in \mathfrak{g}$.
- Anticommutativity: $[X, Y] = -[Y, X] \forall X, Y \in \mathfrak{g}$.
- The Jacobi identity: $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0 \forall X, Y, Z \in \mathfrak{g}$.

The Lie bracket is not necessarily associative. The **Lie bracket** $[\cdot, \cdot]$ tells us how to combine these matrices. A Lie algebra is “closed under commutators”¹⁴: If X and Y are elements of a Lie algebra, then so is

$$[X, Y] = XY - YX.$$

Starting with the Jacobi identity, one can show that the commutator of the generators, X^i , obeys

$$[X^i, X^j] = if^{ij}_k X^k, \quad (2.5)$$

where the **structure constants** f^{ij}_k of the Lie algebra are independent of the representation, even though the elements X^i are not.¹⁵ This may seem confusing until you recognize that a generator X^i is really a matrix with elements $[X^i]_{ab}$ where subscripts a and b for the matrix elements run from 1 to n_R , the dimension of the representation, which is arbitrary; whereas the generator superscripts i, j and k run from 1 to the order (dimension) of the group, which is constant and a property of the group. There is therefore a fixed number of structure constants for each group. For instance, we will see that the structure constants for SO(3) and SU(2) are the Levi-Civita symbols.

- The **Lie algebra** is the vector space the infinitesimal generators, which are subject to the Lie

¹² Meaning the space of tangents to elements of the group.

¹³ Schwichtenberg, PS, 45.

¹⁴ Schwichtenberg, PS, 155.

¹⁵ Maggiore, 15.

bracket defined by the structure constants. “The Lie algebra is encoded within the commutation relations of the generators.”¹⁶

The elements of the Lie algebra are operators (matrices), not numbers, and they do not commute. So although closure requires

$$e^{i\alpha_i X_i} e^{i\beta_j X_j} = e^{i\delta_k X_k},$$

we can not do a simple sum of the exponents of the left side of the equation. Instead, we can use power series to derive the **Baker-Campbell-Hausdorff formula**:

$$e^{i\alpha_i X_i} e^{i\beta_j X_j} = e^{i(\alpha_i X_i + \beta_j X_j) - \frac{1}{2}[\alpha_i X_i, \beta_j X_j]}. \quad (2.6)$$

If we think of the *parameter space* of the group as a *vector space*, in which each point designates a specific element of the group, then the generators specify the space near the identity. Like the 3-d unit vectors \hat{i} , \hat{j} and \hat{k} , which are also defined near the identity (zero) but form a set of basis vectors, the generators form a basis for the entire vector space of the parameters. Said the other way around, the parameters specify a point in the vector space in terms of the generators. Such a point in the vector space corresponds to a specific element of the group, such as a rotation through a certain angle.

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.

$$[C, X] = 0. \quad (2.7)$$

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. They therefore can be used to *label* the representation. An example is $J^2 = \sum_i J_i^2$ for rotations.

Within a Lie algebra, the set of mutually commuting generators, which therefore can be diagonalized simultaneously, is called the **Cartan subalgebra**. The number of such generators is the **rank** of the subalgebra.

So the number of parameters of a Lie group, its **order**, is a constant property of the group. It is also the number of Lie **generators** in the Lie algebras of the group. The number of mutually commuting generators (Casimir operators) is the **rank** of the Lie algebra. Although a representation may have an arbitrary dimension, the order and rank are constant across all representations (and Lie algebras) of a Lie group. Since the Cartan generators are mutually commuting, they share a set of eigenvectors in the vector space. The eigenvectors of the Cartan generators span the space and form a basis for the representation. Their eigenvalues, which vary in number according to the dimension of the representation, can be used to label the corresponding states within the representation, just as the Casimir element provides a label for each representation. Examples will follow.

Among those groups important to physics (more later), $SO(n)$ and $SU(n)$ apply to spaces of n real or complex dimensions. This n is not the order of the group, but it does fix (but does not equal) the number of generators. A subset of these are Cartan generators. An arbitrary $SU(n)$ group will always have $n^2 - 1$ **generators** and be of **rank** $n - 1$, whereas an arbitrary $SO(n)$ group will have $\frac{n(n-1)}{2}$ generators. In different representations, the number of generators remains the same although the dimension of the representation may vary.

For example: Any representation of $SU(2)$ has 3 generators of which one is the Cartan generator, taken to be J_3 . For a j representation, where j corresponds to the Casimir operator J^2 of value $j(j+1)$ (not to the dimension of the representation, which is $2j+1$), the generators are $(2j+1) \times (2j+1)$ matrices. The eigenvalues of the Cartan generator run in integral increments from $-j$ to $+j$. Although the number of Cartan operators is

¹⁶ Blundell and Lancaster, 84.

constant across representations, the number of their eigenvectors depends on the dimension of the representation.

A special representation, the **adjoint representation**, may be defined in term of the structure constants by

$$[T^a]_{bc} = -if_{abc}. \quad (2.8)$$

This definition has two important results:

- The adjoint representation as well as its matrices has the same dimension as the structure constants and, therefore, the dimension (number of parameters) of the group, e.g., $8(n^2 - 1)$ for SU(3) or 3 for SU(2).
 - The adjoint representation is the (only) one in which the number of eigenvectors of the Cartan generators (the simultaneously diagonalizable generators) is the dimension of the group, This is equivalent to saying the number of eigenvectors is the same as the number of generators, which is constant and independent of the representation. The dimension of each eigenvector is that of the representation.
 - So in the adjoint representation, the number of generators = the order of the group = the dimension of the representation = the number of eigenvectors of the Cartan generators = the number of weight vectors (to be defined soon).
- The Jacobi identity then may be used to derive (summed over c)

$$[T^a, T^b] = if_{abc}T^c. \quad (2.9)$$

The **fundamental representation** of SO(n) or SU(n) is the one consisting of $n \times n$ matrices.

Now get this. Classical mechanics can be formulated in terms of **Poisson brackets**

$$\{f, g\} = \sum_i \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i}, \quad f, g \in C(P)$$

where f and g are observables in a $2n$ -dimensional phase space P of n generalized coordinates q_i and the n conjugate momenta p_i . The vector space $C(P)$ is the set of all complex-valued, infinitely differentiable functions on P . The Poisson bracket is obviously antisymmetric and brute force will show that it obeys the Jacobi identity, which means the set of observables constitutes “...one huge Lie algebra.”¹⁷ Taking the Poisson bracket as the Lie bracket of the Lie algebra, it can be the basis for forming a representation of the Lie algebra in a Hilbert space \mathcal{H} “Thus the set of all observables in quantum mechanics forms a Lie algebra, which is one of our main reasons for studying Lie algebras...”¹⁸ You can’t ask for more motivation than that!

2.2. Covering groups and subalgebras

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.

Put the other way around, a covering group is the *unique simply-connected Lie group* corresponding to a given

¹⁷ Jeevanjee, 167.

¹⁸ Jeevanjee, 167.

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.

2.3. Summary – The basics of Lie groups

Let's sum up:

- A **Lie group** corresponds to a set of continuous transformations which are characterized by a certain number of **parameters**, e.g., rotation angles, which is always the same for the group and its representations. This is the **order** of the group. The dimension of an $SU(n)$ group is $(n^2 - 1)$, 8 for $SU(3)$ or 3 for $SU(2)$; of an $SO(n)$ group, $\frac{n(n-1)}{2}$.
- From the equation for a generator (2.3), there is one **generator** per parameter so the number of generators is also constant and is equal to the order of the group. However, ...
- ...different **representations** of the same group may have different dimensions.
- Repeat: Although a representation may have arbitrary dimensions, it always has the same number of generators.
 - Caveat: Do NOT confuse the order, or dimension, of a group (constant) and the dimension of a representation (variable).
- The subset of the generators which commute and are mutually diagonalizable are the **Cartan generators**. Their number, called the **rank** of the group, is also a constant for all representations.
- The **eigenvectors** of the Cartan generators span the space of a representation and therefore are of the same dimension as the representation. The eigenvalues can label the corresponding states. **Casimir operators** commute with every generator and are used to label the representations.
 - $SO(3)$ example: order = number of parameters = $\frac{3(3-1)}{2} = 3$; so three generators = J_1, J_2, J_3 ; Cartan generator = J_3 , so rank = 1; Casimir operator J^2 .
- The **adjoint representation** is special because it has the same dimension as the group. It can be defined in terms of the structure constants, as in (2.9).
- The **fundamental representation** of a group $SO(n)$ or $SU(n)$ is the one consisting of $n \times n$ matrices.
- There is only one simply-connected Lie group corresponding to each Lie algebra. It is called the **covering group** and may cover many representations. We will see that $SU(2)$ is the covering group for $SO(3)$ as well as for the restricted Lorentz group Λ_+^\dagger .

3. Lie groups for physics

Our general method of study will be

- 1) start with an example of a group, e.g., 2×2 matrices and rotation;
- 2) derive the Lie algebra, i.e., the commutation relations;
- 3) use the Casimir element and the Cartan subalgebra of the Lie algebra to look at different representations and search for a simply-connected group, the **covering group**.

In ordinary n -dimensional space, we are interested in transformations which conserve distances (inner products) between two points – isometries. So orthogonal (real) and unitary (complex) transformations are important. The most interesting Lie groups for physics are $SO(n)$ 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)

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

which is a kind of distance between two points. So we add the $O(1,3)$ group to the list of potentially important groups for physics. It will turn out to be somewhat more complex than this.

3.1. 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 in (1.3), by the condition,

$$O^T O = I \quad (3.1)$$

and a **special orthogonal group** by (3.1) and

$$\det(O) = +1. \quad (3.2)$$

The first condition guarantees the conservation of lengths

$$x_1^2 + x_2^2 + x_3^2 + \dots$$

and the second keeps only rotations (not reflections), a rotation being a continuous linear operator which takes orthonormal bases to orthonormal bases.¹⁹

Adding to this equation (2.2) for the group element in terms of generators, we see that

$$e^{i\theta X^T} e^{i\theta X} = 1 \Rightarrow X^T = -X.$$

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

3.2. 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 Hermitian conjugate U^\dagger must satisfy

$$U^\dagger U = 1. \quad (3.3)$$

A **special unitary group**, also satisfies

$$\det(U) = +1, \quad (3.4)$$

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 (3.1) and (3.3) that a unitary operator is an isometry of both real and complex inner product spaces.²⁰

For physics, we define the infinitesimal generator U of the Lie algebra as

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

so that the fundamental equation for isometries

$$(e^{i\theta U})^\dagger e^{i\theta U} = 1$$

yields

¹⁹ Jeevanjee, 125, 151.

²⁰ Jeevanjee, 120-121.

$$U^\dagger = U$$

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

3.3. 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 in the complex plane. Since $e^{i\theta}e^{i\phi} = e^{i(\theta+\phi)}$, it is obviously a group.

Rotations in two dimensions are the unit circle, S^1 , and can be represented by 2-d matrices which are elements of either $SO(2)$ or $U(1)$. The fundamental representation (2x2 matrices) of the $SO(2)$ group consists of rotations in terms of the sine and cosine of the angle of rotation. which satisfy (3.1) and (3.2).

$$R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (3.5)$$

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

$$R_\theta = e^{i\theta} = \cos(\theta) + i\sin(\theta). \quad (3.6)$$

One can map (3.6) to a real matrix by using the 2-d identity matrix

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (3.7)$$

to show that

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

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

$$U(1) \cong SU(2).$$

There is a one-to-one, invertible correspondence between their elements. One can interpret them as the same group with different labels for the elements.

3.4. 3-dimensional rotations – SO(3) and SU(2)

The conditions (3.1) and (3.2) for an orthogonal group are satisfied by 3x3 matrices which are simple extensions of equation (3.5), forming the fundamental representation of $SO(3)$. Again there is an isomorphism – with $SU(2)$.

3.4.1. SU(2)

Let's consider the $j=1/2$ case by writing a 2x2 matrix as follows, where the parameters are complex numbers,

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (3.8)$$

with conditions (3.4) requiring

$$ad - bc = 1. \quad (3.9)$$

Then

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

and unitarity (3.3) requires $U^- = U^\dagger$, so this and (3.9) can be used to solve for c and d to give

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}. \quad (3.10)$$

This now contains two complex numbers and so four real ones. But the requirement that the determinant be +1 takes away one degree of freedom, so we are left with three free parameters for the SU(2) group.²¹ This is consistent with the fact that an arbitrary SU(n) group will always have $n^2 - 1$ **generators**.

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

$$U = e^{i\theta J_i}$$

and applying requirements (3.3), now requires that

$$J_i^\dagger = J_i,$$

meaning J_i is Hermitian (the reason for the factor of i in the exponent). Equation (3.4) requires a zero trace.²² So the generators of SU(2) must be traceless Hermitian matrices. A possible basis in terms of 2x2 matrices is the triplet

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (3.11)$$

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

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \quad (3.12)$$

so we can define the generators as $J_i \equiv \frac{1}{2}\sigma_i$ in order to get

$$[J_i, J_j] = i\epsilon_{ijk}J_k \quad (3.13)$$

which expresses the **Lie bracket of SU(2)** and shows the structure constants to be the elements of the Levi-Civita symbol. Looks a lot like the matrix elements of angular momentum, doesn't it?

Note well that factor of $\frac{1}{2}$ we have just included in the definition of the generators. It was required by the commutators (Lie brackets). It also will pop up in the rotational part of the Lorentz transformation. In both cases (really the same), it is responsible for the unexpected behavior of spinors under rotation.

Alternatively, we could start with the J_3 eigenstate and employ the inverse of the ladder operators

$$J_1 = \frac{1}{\sqrt{2}}(J^- + J^+) \quad \text{and} \quad J_2 = \frac{1}{\sqrt{2}}(J^- - J^+)$$

to *derive* the matrix elements $[J_a]_{m'm} = \langle j, m' | J_a | j, m \rangle$ in terms of $\frac{1}{2}$ the Pauli matrices (3.11).²³

$$J_{1/2}^1 = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \frac{\sigma^1}{2}, \quad J_{1/2}^2 = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \frac{\sigma^2}{2}, \quad J_{1/2}^3 = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\sigma^3}{2} \quad (3.14)$$

²¹ Robinson, 88.

²² A corollary to Jacobi's identity, $\det(e^A) = e^{\text{tr}(A)}$, shows that the trace of a determinant 1 matrix must be zero. Schwichtenberg, 46.

²³ Robinson, 92-93.

That was for the case of $j = \frac{1}{2}$. For $j=1$, a set of basis vectors for the generators acting on \mathbb{C}^3 may be obtained the same way, by starting with the following obvious definition of J_3 and calculating the other matrix elements by expressing them in terms of J^+ and J^- :

$$J_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad J_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (3.15)$$

3.4.2. 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 (a member of the associated Lie algebra) as

$$O = e^{i\theta J}. \quad (3.16)$$

Putting this together with conditions (3.1) and (3.2) gives

$$J^T + J = 0 \quad \text{and} \quad \text{Tr}(J) = 0, \quad (3.17)$$

which can be satisfied by the following generators:²⁴

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (3.18)$$

These are 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 are quite obviously not basis vectors. They can be derived easily from, e.g.,

$$J_3 = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

by using equation (2.3) for the generators in terms of derivatives evaluated at $\theta = 0$.

We can write them compactly using the Levi-Civita symbol:

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

which satisfies the Lie brackets

$$[J_i, J_j] = i\epsilon_{ijk} J_k \quad (3.19)$$

Comparison with (3.13) shows the structure constants to be the same as for SU(2). Since SO(3) and SU(2) have the *same* Lie algebra, they can have a common covering. Or one may be the covering group of the other.

3.5. Irreducible representations of SU(2) – spin

Our goal now is to build representations of this group for different dimensions. We insist again that any such representation has the same order, three, in spite of the number of dimensions. The irreducible representations are the ones of particular interest. Since SU(2) is equivalent (homomorphic) to S^3 , the three-sphere, it is *the* simply-connected group corresponding to this Lie algebra and is therefore its covering group.

One more time, although the representations will have different dimensions, the number of parameters or generators is a constant, the order of the group, three in this case ($4^2 - 1$). We know from (3.13) and the classical

²⁴ Schwichtenberg, PS, 44.

QM study of angular momentum that we can build one Casimir element for this Lie algebra, i.e., which commutes with every generator in the group, 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 in (3.15), 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, from the $J_i \equiv \frac{1}{2}\sigma_i$ and the Pauli matrices (3.11),

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

We see that the representations can be labeled by the Casimir operator values, $j(j+1)$, i.e., 2 or $3/4$.

Again, 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 once again can define two operators

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

and use the commutation relations (3.13) to show that they are the usual ladder or step (raising and lowering) operators used, e.g., for angular momentum. So we finish with the representations for SU(2) shown in Table 1. 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	$j(J^2)$	Cartan eigenvalues, $m(J_3)$	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).

Note that the representations have different dimensions, but all have the same group order, 3 in this case. So SU(2) has representations in different dimensions.

Looking ahead a wee bit, we can say that *particles of different spin belong to different representations of a rotation group*.

3.6. Lorentz transformations

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

$$\Lambda^T \eta \Lambda = \eta. \quad (3.20)$$

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

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

meaning Λ_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 Λ_+^\uparrow , which has determinant $+1$ and $\Lambda_0^0 \geq +1$. 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,

Λ_P and Λ_T , so the entire Lorentz group may be represented by P

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

Consider an infinitesimal vector transformation²⁵

$$\Lambda_\nu^\mu = \delta_\nu^\mu + \omega_\nu^\mu$$

and require conservation of a length in Minkowski spacetime

$$\eta_{\mu\nu} x'^\mu x'^\nu = \eta_{\mu\nu} (\Lambda_\rho^\mu x^\rho) (\Lambda_\sigma^\nu x^\sigma) = \eta_{\mu\nu} x^\mu x^\nu.$$

Ignoring square terms in ω_ν^μ leads to

$$\omega_\nu^\mu = -\omega_\mu^\nu,$$

so the Lorentz group is represented by a 4x4 antisymmetric matrix and therefore has six independent parameters. These may be taken as three rotation angles and the three components of the boost velocity \vec{v} .

Rotations are then simply 3-d rotations (3.18) tucked into the Minkowski spatial part:

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

for example, expressed so as to make clear the matrix form,

$$[R_z(\theta_z)]_\nu^\mu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_z & \sin\theta_z & 0 \\ 0 & -\sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3.21)$$

We use equation (2.3) to calculate the generator

$$[J_z(\theta_z)]_\nu^\mu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (3.22)$$

We can do the same thing for the boosts, starting with the hyperbolic form for the boost along the x-axis:

²⁵ Following treatment based on Schwichtenberg, PS, 66ff.

$$[B_x(\phi_x)]_\nu^\mu = \begin{pmatrix} \cosh\phi_x & -\sinh\phi_x & 0 & 0 \\ -\sinh\phi_x & \cosh\phi_x & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and the generators are

$$[K_1]_\nu^\mu = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [K_2]_\nu^\mu = \begin{pmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, [K_3]_\nu^\mu = \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad (3.23)$$

A general restricted Lorentz transformation is then of the form

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta} + i\vec{K}\cdot\vec{\phi}}. \quad (3.24)$$

Brute-force calculation shows that the commutation relations of the six generators are

$$[J_i, J_j] = i\epsilon_{ijk}J_k, \quad [J_i, K_j] = i\epsilon_{ijk}K_k, \quad [K_i, K_j] = -i\epsilon_{ijk}J_k. \quad (3.25)$$

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

$$N_i^\pm = \frac{1}{2}(J_i \pm iK_i), \quad (3.26)$$

so that the commutation relations become

$$[N_i^+, N_j^+] = i\epsilon_{ijk}N_k^+, \quad [N_i^-, N_j^-] = i\epsilon_{ijk}N_k^-, \quad [N_i^+, N_j^-] = 0. \quad (3.27)$$

Now a revelation: N_i^+ and N_i^- each satisfy the commutation relations (3.19) for the Lie algebra of SU(2). So the complexified²⁶ Lie algebra for the restricted Lorentz group consists of *two copies of 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 representations (0,0), (1/2,0), (0,1/2) and (1/2,1/2), 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 Λ_+^\uparrow . A very important group indeed!

3.6.1. The Lorentz algebra

Starting with Lorentz generators matrices for the 4-d Minkowski space from (3.23), we can use the metric to lower one index, which gives, e.g.,

$$[K_1]_{\mu\nu} = \eta_{\mu\rho}[K_1]_\nu^\rho = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \text{ etc.}$$

The $[K_i]_{\mu\nu}$ are thus a set of antisymmetric imaginary matrices. as are the rotation matrices such as

²⁶ Complexification due to the factor i in equation (3.26).

$$[J_1]_{\mu\nu} = \eta_{\mu\rho} [J_1]_{\nu}^{\rho} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}.$$

This also means they are all Hermitian.

We can adopt a different notation for the same thing by assuming the subscript on J_i and K_i as another dimension of the matrices and write $J^{\mu\nu}$ as the generator for rotations between the μ and ν dimensions (directions), so that J^{10} is a boost in the x direction and J^{23} a rotation mixing y and z, i.e., around the x-axis. In other words,

$$K^i \rightarrow J^{i0} \quad \text{and} \quad J^i \rightarrow J^{jk}.$$

Such a general antisymmetric Hermitian matrix can then be written

$$(J^{\mu\nu})^{ab} = -i(\eta^{\mu a} \eta^{\nu b} - \eta^{\nu a} \eta^{\mu b}),$$

as can be derived by brute force. One can then somewhat more laboriously show that²⁷

$$[J^{\mu\nu}, J^{\rho\lambda}] = i(\eta^{\lambda\mu} J^{\rho\nu} - \eta^{\lambda\nu} J^{\rho\mu} - \eta^{\rho\mu} J^{\lambda\nu} + \eta^{\rho\nu} J^{\lambda\mu}), \quad (3.28)$$

which is therefore the Lie-algebra bracket for the Lorentz group. It defines the **Lorentz algebra**.

An otherwise arbitrary set of matrices which satisfies the *anticommutation* relation

$$\{\gamma^{\mu}, \gamma^{\nu}\} = -2\eta^{\mu\nu} \mathcal{I}, \quad (3.29)$$

is said to constitute a **Clifford algebra**. If we define

$$S^{\mu\nu} = \frac{i}{4} [\gamma^{\mu}, \gamma^{\nu}], \quad (3.30)$$

with the γ^{μ} subject to the Clifford algebra condition, then $S^{\mu\nu}$ is obviously antisymmetric. It also satisfies the Lorentz-algebra bracket (3.28):²⁸

$$[S^{\mu\nu}, S^{\rho\lambda}] = i(\eta^{\lambda\mu} S^{\rho\nu} - \eta^{\lambda\nu} S^{\rho\mu} - \eta^{\rho\mu} S^{\lambda\nu} + \eta^{\rho\nu} S^{\lambda\mu}).$$

So $S^{\mu\nu}$ is a general form for generators of the Lorentz algebra – provided the γ^{μ} matrices satisfy the Clifford-algebra anticommutation relation (3.29).

3.6.2. 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.

3.6.3. The (½,0) representation -- spinors

Use the Pauli matrices (3.11) as basis matrices, so

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

From the definitions of N_i^+ and N_i^- (3.26) with $N_i^- = 0$, one finds

$$J_i = \frac{1}{2} \sigma_i, \quad K_i = \frac{-i}{2} \sigma_i \quad (3.31)$$

²⁷ Robinson, 131.

²⁸ Robinson, 138-9.

so a general transformation is given by

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta} + i\vec{K}\cdot\vec{\phi}} = e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2} + i\vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \quad (3.32)$$

The factors of $\frac{1}{2}$ 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**:

$$\mathcal{X} = \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix}. \quad (3.33)$$

Also, note that equation (3.32) 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

$$\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \sigma_3^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \sigma_1^2 = \sigma_2^2,$$

the expansion for a boost along the z-axis gives

$$\begin{aligned} B_3(\phi) &= e^{\phi \frac{\sigma_3}{2}} = 1 + \frac{1}{2}\phi\sigma_3 + \frac{1}{2}\left(\frac{1}{2}\phi\sigma_3\right)^2 + \dots, \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\phi}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \left(\frac{\phi}{2}\right)^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \dots = \begin{pmatrix} e^{\frac{\phi}{2}} & 0 \\ 0 & e^{-\frac{\phi}{2}} \end{pmatrix} \end{aligned} \quad (3.34)$$

which makes clear the operator is a 2×2 matrix.²⁹ Similar treatment of a rotation leads to almost the usual matrix in terms of sine and cosine of the rotation angle:

$$R_1(\theta) = e^{i\theta \frac{\sigma_1}{2}} = 1 + \frac{1}{2}i\theta\sigma_1 + \frac{1}{2}\left(\frac{i}{2}\theta\sigma_1\right)^2 + \dots = \begin{pmatrix} \cos(\frac{\theta}{2}) & i\sin(\frac{\theta}{2}) \\ i\sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix}. \quad (3.35)$$

For a vector, this is not the angle we expected. But this is not $SO(3)$ and vectors, it is $SU(2)$ and spinors. The fundamental representations of $SU(2)$ and $SO(3)$ apply to different mathematical spaces, a 3-d vector space for $SO(3)$ and a 2-d spinor space for $SU(2)$. The factor of $\frac{1}{2}$ shows that spinors are not normal vectors in spacetime because, after a rotation by 2π , a vector returns to its previous state, whereas a spinor changes by a factor of -1. A set of spinors rotated from zero to 2π corresponds to a set of vectors rotated from zero to 4π . A given rotation angle of the vector maps to two different rotation angles of the spinor. For example, a rotation of the spinor by either $\pi/2$ or $3\pi/2$ has the same effect as a rotation of the vector by π . Mathematically more precisely, the map

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

designates a **homomorphism**³⁰: “... for every $R \in SO(3)$ there correspond exactly two matrices in $SU(2)$ which map to R under ρ .”³¹ It is not an isomorphism because the relation between the two is not invertible. So we see that $SU(2)$ is the **double cover** of $SO(3)$.

Note that this behavior of the spinor is due the factor of $\frac{1}{2}$ which comes from the half-integral value of the spin in equation (3.31). This is one case of a general rule which says:

²⁹ Robinson, 123; Schwichtenberg, PS, 73.

³⁰ Jeevanjee, 176.

³¹ Jeevanjee, 140.

A particle of spin s is invariant under a rotation by $2\pi/s$ radians.³²

Section 3.6.5 will present a geometric interpretation of the relation between the two spaces.

A rotation in space must also take into account the half rotation in spinor space. We know this is the $j=1/2$ representation of $SU(2)$, which behaves like an angular momentum. It is now obvious that this can be considered the **spin** of a spin- $1/2$ particle like an electron, whose spin part is in spinor space. It is total angular momentum – the sum of spatial and spin – that is conserved. But to show this requires working out the effects of the total angular momentum operator and we will not go there.

As already noted at the end of section 3.4.1, $SU(2)$ is the **three-sphere**, S^3 , a 3-dimensional “spherical” space embedded in four dimensions, and is the unique **covering group** for $SO(3)$. $SU(2)$ is said to be the **double-cover** of $SO(3)$, which is seen as half of $SU(2)$.

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.

Let’s just note that all these unit-length conserving groups are in fact the same as³³ S^n , the n -spheres:

- S^1 corresponds to $U(1)$ and $SO(2)$,
- S^3 corresponds to $SU(2)$ and so to each 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.

3.6.4. The $(0, 1/2)$ representation

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

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

so a general transformation is given by

$$\Lambda = e^{i\vec{J}\cdot\vec{\theta} + i\vec{K}\cdot\vec{\phi}} = e^{i\vec{\theta}\cdot\frac{\vec{\sigma}}{2} - \vec{\phi}\cdot\frac{\vec{\sigma}}{2}}. \quad (3.36)$$

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

$$\xi = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}. \quad (3.37)$$

Under rotations, right-chiral spinors transform just like left-chiral spinors by (3.35). But under boosts, they transform by

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

Left and right-chiral spinors transform differently by a negative sign of the exponent 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 direction transforms under a boost along the same direction (and by a

³² Carroll (2024), 242.

³³ “Same as” is too vague. Are they in fact homomorphisms or isomorphisms?

rotation) like a left-chiral one, and vice versa.³⁴ We will see that this requires both right and left-chiral spinors in a valid description of spin- $\frac{1}{2}$ particles.

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

3.6.5. Geometric picture of spin

A kind of stereographic projection, or map, is a heuristic means of relating a space rotation to a spinor.³⁵

Imagine a sphere of radius 1 centered at the origin and describe any point on the sphere by the usual polar-coordinate angles, θ and ϕ , the angles with respect to the z axis and to the x axis in the x-y plane. Map a point on the sphere to the x-y plane through the origin (center of the sphere) by drawing a ray from the top of the sphere, at $x=y=0$ and $z=1$, through the point to be mapped, P, and onto point P' in the x-y plane. Then consider the plane to be the complex plane with x as its real part and y as the imaginary one, so that the projected point is given by one complex number.

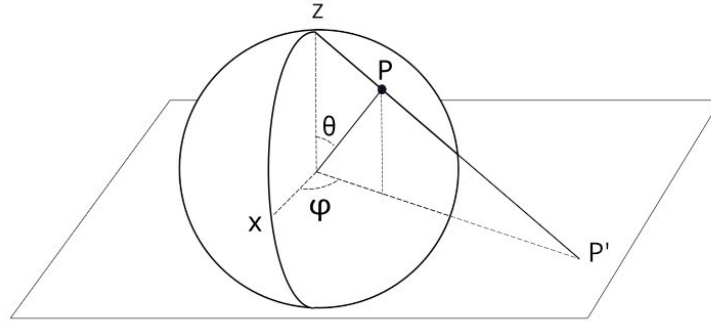


Figure 1. Bloch sphere construction of spinor space.

Since the triangle OZP is isosceles, the “radius” from the origin to the projected point P' is

$$r = r(\theta) = \tan\left(\frac{1}{2}(\pi - \theta)\right) = \cot\frac{\theta}{2},$$

This maps the point (θ, ϕ) on the unit sphere to

$$z = r(\theta)e^{i\phi} = e^{i\phi} \cot\frac{\theta}{2}.$$

This may be written as

$$z = \frac{\alpha}{\beta},$$

where α can be taken as real and all the complex part is relegated to β . Then we may pick

$$\alpha = \cos\frac{\theta}{2} \quad \text{and} \quad \beta = e^{-i\phi} \sin\frac{\theta}{2}.$$

Now let's interpret these numbers as spinor components, which makes the **map of the unit sphere to spinor space** complete:

³⁴ Schwichtenberg, NNQFT, 88.

³⁵ Robinson, 152ff.

$$(\theta, \phi) \rightarrow \psi_R^* = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{-i\phi} \sin \frac{\theta}{2} \end{pmatrix},$$

so that

$$\psi_R = \begin{pmatrix} \cos \frac{\theta}{2} \\ e^{i\phi} \sin \frac{\theta}{2} \end{pmatrix}. \quad (3.38)$$

This indeed can be shown to Lorentz transform like a right-chiral spinor.

Spin “up” then means $(\theta, \phi) = (0, 0)$ and this gives the spinor

$$\psi_R = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |up\rangle = |\uparrow\rangle,$$

whereas “down” means $(\theta, \phi) = (\pi, 0)$, which gives

$$\psi_R = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |down\rangle = |\downarrow\rangle.$$

The x and y directions can be calculated also, with the appropriate results. The factor of $\frac{1}{2}$ the angle is made clear from the projection.

The up and down spin directions are perpendicular (orthogonal, i.e., by an angle of $\frac{\pi}{2}$) in spinor space but correspond to +z and -z in spacetime, which are π radians apart.

So we find again that a complete rotation in spacetime corresponds to a rotation of only π radians in spinor space. A complete rotation in spinor space corresponds to two in spacetime. This 2-to-1 relation between $SO(3)$ and $SU(2)$ is interpreted to mean that $SU(2)$ is the **non-trivial double cover** of $SU(3)$. The projection also shows that spin is a rotation in *spinor space* (the complex plane), not ordinary spacetime.

The stereographic projection provides a geometric picture of spin and the way a direction in space (the sphere), described by $SO(3)$, corresponds to a direction in spinor space (the complex plane), described by $SU(2)$.

Applying a **parity transformation** to the point on the complex plane,

$$\theta \rightarrow \pi - \theta, \phi \rightarrow \phi + \pi,$$

leads to the equivalent left-handed spinor

$$\psi_L = \begin{pmatrix} e^{-i\phi} \sin \frac{\theta}{2} \\ -\cos \frac{\theta}{2} \end{pmatrix},$$

which, as expected, transforms like a left-handed spinor.³⁶

3.6.6. Weyl and Dirac spinors and chirality

Remember that in physics, objects like scalars and vectors are defined by their transformation properties. This is also the case for spinors. Spinors are new and rather strange beasts. They seem to live in two spaces at once, with one foot (so to speak) in 4-d Minkowski space (since they change under Minkowski rotations and boosts) and the other in **two-dimensional spinor space** (since they have their own transformation rules). More

³⁶ Robinson, 154-61, milks this Bloch sphere projection for more information, including taking a complex conjugate to go from right to left-handed spinors and back. We will see this another way in the next section.

precisely, they live in a spinor space, one of which is attached to each point of spacetime.³⁷

Moreover, spinors exist in both left and right-chiral versions. Equations (3.34) and (3.35) show clearly that under transformations of Minkowski space, neither left-chiral nor right-chiral spinors transform like 4-vectors. They transform the same way under rotations, but differently under boosts.

We need to understand more about the relation between the two.

Robinson says of the rest of this section, that it is "... a topic students typically find very frustrating and tedious."³⁸ In my opinion, that's because it is very frustrating and tedious.

Let's change our notation a bit. A left-chiral spinor will be ψ_L and a right-chiral, ψ_R . Since spinors live in their own space, **spinor space**, they have their own **spinor metric**,

$$\epsilon^{ab} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = i\sigma^2, \quad (3.39)$$

which is real. Note that the term on the right is the second Pauli matrix, not a squared term. This is not the Minkowski metric of spacetime but is used similarly, e.g., to raise or lower indices³⁹

$$\psi^a = \epsilon^{ab} \psi_b.$$

This is always possible because ϵ is invariant under Lorentz transformations in the spinor representation.⁴⁰ Thus, a superscripted spinor involves the spinor metric just as a superscripted 4-vector involves the Minkowski metric,

$$x^\mu = \eta^{\mu\nu} x_\nu.$$

Now define two new spinor quantities:

$$\bar{\psi}_L := \epsilon \psi_L^* = i\sigma^2 \psi_L^* \quad \text{and} \quad \bar{\psi}_R := -\epsilon \psi_R^* = -i\sigma^2 \psi_R^*. \quad (3.40)$$

Then use of (3.32) and (3.36) shows that $\bar{\psi}_L$ Lorentz transforms like a right-chiral spinor and $\bar{\psi}_R$ like a left-chiral spinor.⁴¹ Indeed, we may say they are those spinors, for if⁴²

$$\psi_L = \bar{\psi}_R = -i\sigma^2 \psi_R^*,$$

then from this and (3.40),

$$\bar{\psi}_L = i\sigma^2 \psi_L^* = i\sigma^2 (-i\sigma^2 \psi_R^*)^* = i\sigma^2 (-i\sigma^2) \psi_R = \psi_R,$$

since $i\sigma^2$ is real and

$$(-i\sigma^2)(i\sigma^2) = \mathcal{I}.$$

The conversion from ψ_L to $\bar{\psi}_L$ is taken to indicate the operation commonly called **charge conjugation** (reversal),

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \rightarrow \bar{\Psi} = \begin{pmatrix} \bar{\psi}_L \\ \bar{\psi}_R \end{pmatrix} = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix},$$

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

37 A spinor space may be viewed as a parametrized family of spaces, S_x , each isomorphic to $SU(2)$ and one for each point x in Minkowski space B . So, in the lingo of topology, S_x is a fiber "bundled" to Minkowski space B .

38 Robinson, 133.

39 It is customary to use Roman indices for spinors, preferably lower-case (a, b, c...), in order to avoid confusion with l, j, k of basis vectors of Euclidean space.

40 Robinson, 135.

41 Robinson, 125.

42 I may have the signs backward, but they cancel out in the end.

Here's where it gets messy. Let's define super- and sub-scripted spinors as follows:⁴³

$$\psi_L := \psi_a \quad \text{and} \quad \psi_R := \psi^{\dot{a}}. \quad (3.41)$$

Note well the dot over the subscript or subscript which is intended to indicate *right-chiral; no dot, left-chiral*.

Also,

$$\psi_R = \psi^{\dot{a}} = \bar{\psi}_L = \epsilon(\psi_a)^*$$

shows that since the metric raises or lowers the index, complex conjugation converts inverts the chirality; it exchanges dotted and undotted. This is consistent with the notation of (3.41).⁴⁴ So, to take home:

- Complex conjugation interchanges dotted and undotted indices (inverts the chirality).
- The metric ϵ raises and lowers indices.
- Transforming a left or right-chiral spinor into a right or left-chiral requires both actions.

The third point is the reason behind the introduction of the two notations, upper/lower and dotted/undotted.

The spinor metric can be used to construct the inner product

$$\psi_a \psi^a = \psi_a \epsilon^{ab} \psi_b,$$

which is Lorentz invariant (by (3.34) and (3.35)). In addition, terms like

$$\psi^\dagger \psi = \psi_b \epsilon^{ba} \psi_a \quad (3.42)$$

are then also invariant, but only because one of the spinors is its own complex conjugate.

Let's look at the spinor Lorentz transformations of (3.32) and (3.36) 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

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

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

$$\Lambda_{(\frac{1}{2}, 0)} = (e^{i\vec{\theta} \cdot \frac{\vec{\sigma}}{2} + \vec{\phi} \cdot \frac{\vec{\sigma}}{2}}) := \Lambda_a^b \quad (3.43)$$

$$\Lambda_{(0, \frac{1}{2})} = (e^{i\vec{\theta} \cdot \frac{\vec{\sigma}}{2} - \vec{\phi} \cdot \frac{\vec{\sigma}}{2}}) := \Lambda^{\dot{a}}_{\dot{b}} \quad (3.44).$$

Remember that, because of the σ terms, these are matrix equations and applying the operators generally requires expressing the exponential function as a power series. And don't miss 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 (3.26),

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

One more important point: 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. This is rather like in SR, where we are required to use 4-vectors which describe space and time simultaneously. In the spinor case, the solution adopted is to use a **Dirac spinor**

⁴³ Why bother? That tiny dot is almost invisible!

⁴⁴ Dixit Schwichtenber, NNQFT, 99-100.

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_a \\ \psi^{\dot{a}} \end{pmatrix},$$

where a parity transformation gives

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} \rightarrow \Psi' = \begin{pmatrix} \psi_R \\ \psi_L \end{pmatrix}.$$

In spite of appearances, a **Dirac spinor** is a four-component object, composed of two two-component Weyl spinors.

$$\Psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix} = \begin{pmatrix} \psi_{L1} \\ \psi_{L2} \\ \psi_{R1} \\ \psi_{R2} \end{pmatrix}.$$

Its four components exist in spinor space, not 4-d spacetime, and it transforms like two spinors, *not* like a 4-vector, In math speak, it is said to be in the

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

product 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.

3.6.7. The $(\frac{1}{2}, \frac{1}{2})$ representation

An object in this representations has two indices, each one transforming under its own 2-dimensional copy of the Lie algebra $\mathfrak{su}(2)$, so the object is 4-dimensional. Using the Pauli matrices plus the identity as basis 2x2 matrices, such an object can be written as:

$$v_{ab} = v_\nu \sigma_{ab}^\nu = v_0 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + v_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + v_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + v_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (3.45)$$

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 $(\frac{1}{2}, \frac{1}{2})$ representation represents vectors and we can use the simpler vector matrix algebra. In fact, a 4-vector is a **rank 2 spinor** and so has two Cartan generators and two eigenvalues. Whereas a rank-2 tensor has two vector indices, we see now that a rank 2 vector has two spinor indices. In this sense, **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).⁴⁵ 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. As Robinson says: "... the (1/2, 1/2) representation ... is the fundamental defining vector representation of the Lorentz group in 1+3 spacetime dimensions."⁴⁶

Also, note that the determinant of (3.45) is

$$\det(v_{ab}) = (v_0^2) - (v_1^2) - (v_2^2) - (v_3^2),$$

the invariant spacetime "distance" of SR.

3.7. The Poincaré group

The full spacetime symmetry group of nature is the **Poincaré group**, also called the **Inhomogeneous Lorentz group**.

$$\begin{aligned} \text{Poincaré group} &= \text{Lorentz group} + \text{translations} \\ &= \text{rotations} + \text{boosts} + \text{translations}. \end{aligned}$$

Infinitesimal **translations** are represented by the generators

$$P_i = -i\partial_i.$$

The complete Poincaré algebra is then as follows:

$$[J^i, J^j] = i\epsilon^{ijk} J^k, \quad [J^i, K^j] = i\epsilon^{ijk} K^k, \quad [J^i, P^j] = i\epsilon^{ijk} P^k, \quad (3.46)$$

$$[K^i, K^j] = -i\epsilon^{ijk} J^k, \quad [P^i, P^j] = 0, \quad [K^i, P^j] = iH\delta^{ij}, \quad (3.47)$$

$$[J^i, H] = 0, \quad [P^i, H] = 0, \quad [K^i, H] = iP^i. \quad (3.48)$$

The first three equations show that J^i generates rotations of vectors and that K^i and P^i are such vectors. The third line indicates that J^i and P^i , rotations and translations, commute with the generator of time translations (the Hamiltonian) and so are conserved. However boosts, K^i , do not, so are not conserved quantities and are therefore not useful for labeling physical states.⁴⁷

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

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

the square of the mass, supposed to be real and positive. The other is

$$W_\mu W^\mu$$

where W^μ is the Pauli-Lubanski four-vector

$$W^\mu = -\frac{1}{2}\epsilon^{\mu\nu\rho\sigma} J_{\nu\rho} P_\sigma.$$

Mathematically, massive and massless particles must be treated differently. To resume, for massive particles,

$$P_\mu P^\mu = m^2 \quad \text{and} \quad W_\mu W^\mu = -m^2 j(j+1),$$

so their representations are labeled by objects m and j which at this point are just numbers, but will later be identified as mass and spin, $j = 0, \frac{1}{2}, 1, \dots$. Individual states within a representation are labeled by $j_z = -j, -j+1, \dots, j$ and so have $2j+1$ degrees of freedom.

⁴⁵ Schwichtenberg, 83.

⁴⁶ Robinson, 130.

⁴⁷ Maggiore, 35.

This method cannot work for massless particles. It can be shown that massless particles only have one degree of freedom, parametrized by their helicity, the projection of the angular momentum along their direction of motion.

Note that all this has been shown from consideration of Poincaré invariance; we have not yet written a single Lagrangian. Indeed, some authors introduce QFT by considering how Lie groups, Lorentz transformations and representations of the Poincaré group predict the existence of massive particles of spin j with $2j + 1$ degrees of freedom, and massless particles of definite helicity and only one degree of freedom.⁴⁸

4. Finding Lagrangians from symmetry

Standard notation, at least for us, uses Φ for scalar fields, Ψ for spinor fields and A and \mathcal{W} for vector fields.

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:⁴⁹

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.⁵⁰

The overriding central constraining factor is, of course, that the action must be Lorentz invariant, which will be satisfied if the Lagrangian is Lorentz invariant (although this is not a necessary condition), for instance, if it is a scalar.

Note that Lagrangians for fields are functions of the components of the field, not of momenta and coordinates of particles. More on this in section 5.6.

We must consider several cases. For the moment, we consider only real, classical particles.

4.1. 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 because they are not Lorentz-invariant, so the D term goes out. The F term is odd in terms of both the field and the derivative, so out with it.
- 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

$$\mathcal{L} = C\Phi^2 + E\partial_\mu\Phi\partial^\mu\Phi. \tag{4.1}$$

⁴⁸ Maggiore, 36-40; Weinberg, 62-74.

⁴⁹ Schwichtenberg, PS, 97-98.

⁵⁰ Except for the Higgs field (6.55).

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

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

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 (4.1), which gives us an identification for the variables C and D. Therefore, we can say the Lagrangian for scalars is

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \Phi \partial^\mu \Phi - m^2 \Phi^2). \quad (4.2)$$

From this Lagrangian, the Euler-Lagrange equations (5.34) give

$$(\partial_\mu \partial^\mu + m^2) \Phi = 0, \quad (4.3)$$

the **Klein-Gordon equation**.

4.2. Spin 1/2 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

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

The 4x4 γ^μ matrices are defined in terms of the identity and the Pauli matrices σ_i

$$\gamma^0 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \quad \gamma_i = \begin{bmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{bmatrix}.$$

This is the **Dirac representation**. With $A = -m$ and $B = 1$, this gives

$$\mathcal{L} = -m \bar{\Psi} \Psi + i \bar{\Psi} \gamma^\mu \partial_\mu \Psi = \bar{\Psi} (i \gamma^\mu \partial_\mu - m) \Psi \quad (4.4),$$

the **Dirac Lagrangian**.⁵¹ A more direct and intuitive “derivation” of the Dirac equation will be shown in section 5.11.

4.3. Spin 1 particles

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

$$\mathcal{L}_{Proca} = C_1 \partial^\mu A^\nu \partial_\mu A_\nu + C_2 \partial^\mu A^\nu \partial_\nu A_\mu + C_3 A^\mu A_\mu + C_4 \partial^\mu A_\mu,$$

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

$$2C_3 A^\rho = 2C_1 \partial_\sigma \partial^\sigma A^\rho + 2C_2 \partial^\rho (\partial_\sigma A^\sigma).$$

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

$$m^2 A^\rho = \partial_\sigma (\partial^\sigma A^\rho - \partial^\rho A^\sigma). \quad (4.5)$$

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**

⁵¹ See Robinson, 167ff, for a much more complete (and laborious) derivation of this Lagrangian.

$$F^{\sigma\rho} := \partial^\sigma A^\rho - \partial^\rho A^\sigma \quad (4.6)$$

which gives

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

The Lagrangian for massless spin 1 is then

$$\mathcal{L}_{Maxwell} = -\frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) = \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \quad (4.7)$$

For a massive spin 1 field,

$$\mathcal{L}_{Proca} = -\frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) + m^2 A_\mu A^\mu = \frac{1}{4}F^{\mu\nu}F_{\mu\nu} + m^2 A_\mu A^\mu. \quad (4.8)$$

5. Field theory of free fields

Why fields? It's all because of QM and SR.⁵²

5.1. Why QFT? Problems with QM and SR

Think about single-particle QM. Squeeze a single particle into a space smaller than its Compton wavelength, \hbar/m . By the Uncertainty Principle, its momentum is then $\gg m$, which means that pair production can take place ... and we no longer have a single particle. But classical wave mechanics cannot handle such a case of particle production or creation. The Schrödinger equation for an electron has no place for the creation of an electron-positron pair.

If that were not serious enough, the Schrödinger equation is obviously not Lorentz-invariant since it includes spatial derivatives quadratically but time derivatives linearly.

Using the unitary time-evolution operator e^{-iHt} , it is possible to derive an amplitude for evolution from a state at $t=0$ over a space-like interval, i.e., $\delta x > \delta t$, and the result is not (quite) the zero it should be!⁵³ This result is clearly not acceptable to SR. More simply, you can't measure an event here and now and then measure an event light-years away but one second later -- if "one second later" even means anything in this case (which it doesn't).

In EM theory, we measure fields at a point and consider the propagation of waves through space. The notion of fields provides **locality** to the equations, a means of connecting two points. Needing just this in QM, we suspect that we need to use fields similar to those of EM. But since measurable quantities are represented by operators in QM, it seems we need **operator-valued fields**. That is the core assumption of quantum field theory. But first, a reminder of a bit of math.

5.2. Some reminders

Here are some reminders of previously studied subjects which we will need.

5.2.1. Action and the Euler-Lagrange equation

The basic variables of classical mechanics are the position and momentum of a particle, plus its non-spatial properties such as spin. A free particle is represented by a wave function $\psi(x)$ which is evaluated at a position which is a single point in space, the position presumably occupied by the particle. Its Lagrangian is then

⁵² The ideas of the next three paragraphs are borrowed from B&L, chapter 8, 75-76; and Zee, 3-4.

⁵³ Lancaster and Blundell, 75-76.

$$L = \frac{1}{2}m\dot{x}^2 + V(x).$$

Its momentum is defined by

$$p(x) = \frac{\partial L}{\partial \dot{x}}$$

at the position x . The Action is defined in terms of the Lagrangian

$$A = \int L(q, \dot{q}) dt, \quad (5.1)$$

so the principle of least (or at least extremized) Action says that

$$\delta A = \int \delta L(q, \dot{q}) dt = 0.$$

We then use differentiation by parts to find

$$\delta A = \int \left(\frac{\partial L}{\partial q} \delta q + \frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) dt.$$

Choose

$$u = \frac{\partial L}{\partial \dot{q}} \text{ and } dv = \delta \dot{q} = \frac{d(\delta q)}{dt},$$

then since $d(uv) = (du)v + u dv$, we get

$$\delta A = \int \left[\left(\frac{\partial L}{\partial q} \delta q + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta q \right) - \delta q \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \delta \dot{q} \right) \right] dt.$$

We suppose the middle term reduces to zero upon integration, since the end points are considered stationary.

So for $\delta A = 0$, we must have the **Euler -Lagrange equation**⁵⁴

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0. \quad (5.2)$$

This is used to write the equation of motion of the particle, information about its dynamics.

5.2.2. Fourier transforms

A function $f(x)$ can be transformed into a form which describes the distribution of frequencies of the original function. It is a complex-valued (forward) **Fourier transform**:

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx. \quad (5.3)$$

The **inverse transform** is:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx} dk. \quad (5.4)$$

The forward transform is an integral of e^{-ikx} , the inverse, of e^{ikx} (using the Minkowski metric signature +---). It is easy to deduce the three-dimensional version of this:

$$F(\vec{k}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(\vec{x}) e^{-i\vec{k} \cdot \vec{x}} d^3x.$$

⁵⁴ This somewhat over-hasty derivation can be improved on, but it gets across the basic steps.

In QM, we use these equations to transform back and forth between the the position (\vec{x}) and momentum (\vec{p}) representations. To go to relativistic four dimensions, we write the four-momentum in terms of the temporal and spatial frequencies.

$$p^\mu = (E/c, \vec{p}) = (\hbar\omega, \hbar\vec{k}) = \hbar k^\mu$$

and use $\hbar=c=1$, which gets us the kx in the exponent of the two equations:

$$kx = \omega t - \vec{k} \cdot \vec{x}. \quad (5.5)$$

This is the Lorentz-invariant product of the four-vectors

$$x^\mu = (t, \vec{x}) \quad \text{and} \quad k^\mu = (\omega, \vec{k}). \quad (5.6)$$

So

$$kx = k^\mu x_\mu = g_{\mu\nu} k^\mu x^\nu, \quad (5.7)$$

means

$$k^2 = k^\mu k_\mu = \omega^2 - \vec{k}^2 = m^2, \quad (5.8)$$

which is the SR dispersion relation

$$E^2 = p^2 + m^2,$$

The SR 4d version then looks a lot like the 1d version except that now it's over four dimensions. With the (+---) metric signature, one-dimensional $-ikx$ goes to four-dimensional $kx = +\omega t - \vec{k} \cdot \vec{x}$, so

$$F(k) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} f(x) e^{ikx} d^4x \quad (5.9)$$

and

$$f(x) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} F(k) e^{-ikx} d^4k. \quad (5.10)$$

Different authors may distribute the factors of 2π differently.

5.2.3. Harmonic oscillators and ladder operators

A classical harmonic operator, a mass attached to a spring, obeys Hooke's law:

$$F = -kx = m \frac{d^2x}{dt^2}, \quad (5.11)$$

of which the solution is

$$x(t) = A \sin(\omega t) + B \cos(\omega t)$$

where

$$\omega = \sqrt{\frac{k}{m}} \quad (5.12)$$

is the angular frequency of oscillation. Since the potential energy can be expressed by

$$V(x) = \frac{1}{2} kx^2 = \frac{1}{2} m\omega^2 x^2,$$

the quantum oscillator has for its Schrödinger equation

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2\psi = E\psi. \quad (5.13)$$

The Hamiltonian can then be written in terms of the momentum operator $\hat{p} = -i\hbar d/dx$.

$$\hat{H} = \frac{1}{2m}[\hat{p}^2 + (m\omega x)^2]. \quad (5.14)$$

Since the momentum and position operators do not commute, we can not factor this simply. But define

$$\hat{a} = \frac{1}{2\hbar m\omega}(m\omega\hat{x} + i\hat{p}_x) \quad \text{and} \quad \hat{a}^\dagger = \frac{1}{2\hbar m\omega}(m\omega\hat{x} - i\hat{p}_x), \quad (5.15)$$

whose product is

$$\hat{a}\hat{a}^\dagger = \frac{1}{2\hbar m\omega}[\hat{p}^2 + (m\omega x)^2] - \frac{i}{2\hbar}[\hat{x}, \hat{p}].$$

The standard canonical commutation relation

$$[\hat{x}, \hat{p}] = i\hbar, \quad (5.16)$$

which is the **canonical commutation relation** leads to

$$\hat{a}\hat{a}^\dagger = \frac{1}{\hbar\omega}\hat{H} + \frac{1}{2} \quad \text{and} \quad \hat{a}^\dagger\hat{a} = \frac{1}{\hbar\omega}\hat{H} - \frac{1}{2}, \quad (5.17)$$

so

$$[\hat{a}, \hat{a}^\dagger] = 1. \quad (5.18)$$

The Hamiltonian can be written in terms of these new operators \hat{a}_\pm as

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}). \quad (5.19)$$

If

$$\hat{H}\psi = E\psi$$

we find that

$$\hat{H}(\hat{a}_+\psi) = (E + \hbar\omega)(\hat{a}_+\psi) \quad \text{and} \quad \hat{H}(\hat{a}_-\psi) = (E - \hbar\omega)(\hat{a}_-\psi), \quad (5.20)$$

so \hat{a}^\dagger can be seen as a raising operator, boosting a state of energy E_n to E_{n+1} . Similarly, \hat{a} is a lowering operator, reducing a state of energy E_n to E_{n-1} . So an energy eigenstate ψ_n of energy E_n can be expressed in terms of the zero-state energy state ψ_0 by

$$\psi_n(x) = A_n(\hat{a}^\dagger)^n\psi_0(x),$$

where A_n is a normalization constant. In QFT, the operators \hat{a}^\dagger and \hat{a} will be considered **creation** and **annihilation operators**. due simply to the commutation relations (5.18). All this comes from the commutation relation for the operators and the definition of the Hamiltonian in terms of them. In addition, equation (5.19) tells us that $\hat{a}^\dagger\hat{a}$ can be interpreted as the energy of the oscillator in increments of $\hbar\omega$ -- a **number operator**.

Now we are ready for QFT --- well, almost.

5.2.4. Angular momentum

The classical definition of angular momentum \mathbf{L} is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}$$

and inserting the QM momentum operator $\hat{p} = -i\hbar d/dx$ yields

$$[L_x, L_y] = i\hbar L_z$$

and so on, or more succinctly,

$$[L_i, L_j] = i\hbar \epsilon_{ijk} L_k.$$

Recall that we saw in (3.13) that this is the Lie bracket for SU(2). From these commutation relations, we can show similarly to the method of harmonic oscillators, that the ladder operators defined by

$$L_{\pm} = L_x \pm iL_y$$

are ladder operators for eigenvalues of the z-component of angular momentum, L_z . If f is an eigenstate of L_z , and the square of the **total angular momentum** (the Casimir operator)

$$L^2 = L_x^2 + L_y^2 + L_z^2,$$

then for

$$L^2 f = \lambda f \quad \text{and} \quad L_z f = \mu f$$

we find that

$$L_z(L_{\pm} f) = (\mu \pm \hbar)(L_{\pm} f).$$

Now we are ready for spin. Well, almost...

5.3. A wee bit of philosophy

First, two basic assumptions:

- Without going into ontology, we will assume that there is a reality “out there” to observe, that what we perceive with our senses (and instruments) really exists in some meaningful and useful way.
- Understanding it will only be possible if it is governed by universal natural laws⁵⁵ and so we also shall assume their existence. We speak of a **law** when we have determined that a given previous event (called the cause) in certain specific and well-defined conditions always gives rise to the same subsequent event (the result) across countless observations. The word “always” here means “every time we have observed it”.⁵⁶ So we consider that this will always be true in the future too. The Sun will indeed appear to rise tomorrow morning.

Now we want to use QFT to study systems of multiple particles. But what is a particle?

The first point should be amended with the provision that we see – obviously – what we see and not what we do not. And we only see, or as physicists prefer to say, *observe* properties of particles, not objects devoid of properties. Science philosopher Paul Teller calls this notion **primitive thisness**, the idea that properties have to be properties of something, of some substratum of the particle which is never explained, but which allows us to mentally identify separate particles.⁵⁷ The idea smacks of the notion of essence of ancient philosophers. An electron is not a just a point particle, a neutral thingy, it is an assembly of properties -- mass, charge and spin, in addition to its position and momentum. Without those properties, there is nothing left we can identify as an electron. So rather than assume that point particles exist, we can admit that we only discern particles by their properties.

Standard treatment of a two-particle system labels each one, say particle 1 and particle 2, and then identifies them by their properties. In fact, Teller points out, they really have two roles, that of bearing properties and that of

⁵⁵ This idea is as old as Epicurus or Lucretius.

⁵⁶ This second assumption comes in fact from Hume, among others, so we're in good company.

⁵⁷ Teller, 28.

supplying identity. Notions of symmetry under swapping of particles also assume we can identify individual particles. Teller uses a vocabulary which employs **quanta** instead of particles. His quanta can not be labeled, merely aggregated. Teller counters this labeling by preferring use of the concept of Fock space, which only counts the number of quanta in each state, as we shall see later in section 5.5. And quanta do not behave like fields, says Teller.

In addition to primitive thisness, Teller is bothered (Who is not?) by the notion of **superposition**, a well-known bogey of QM. Teller claims to "... understand superimposed properties in terms of **propensities**. If the property Q is the superposition of properties P₁ and P₂, then Q is a property in its own right, but it also includes a propensity to yield as a measurement result the superimposed properties P₁ or P₂, when the right 'measurement' activating conditions are in place, with the probabilities given by the probability amplitudes".⁵⁸ I'm not sure this interpretation helps me a lot.

Teller's goal is to show that we can avoid the "field" in QFT, that we can describe multi-particle systems without any mention of fields. In order to do this, he starts with a multi-particle Fock space. He then assumes the existence of a number operator in terms of a creation operator and its Hermitian adjoint with the commutation relations of a harmonic oscillator (though he does not mention that). These elements lead to a mathematics of multi-particle Fock spaces built by multiple applications of creation operators without mentioning fields. The procedure adopted is the inverse of that of second quantization, starting with only Fock space vectors and the definition of the number and lowering operators; all the rest follows.

The entire Fock space can be built up from the vacuum state by use of the creation operator and results in:

$$|\dots n_i \dots\rangle = \Pi_i \frac{1}{\sqrt{n_i!}} (\dots a_i^\dagger \dots) |0\rangle. \quad (5.21)$$

Interestingly, Srednicki, as well as Lancaster and Blundell, introduces multi-particle QM in the same way.⁵⁹

5.4. Interlude: a word on Fock spaces

Informally, a Fock space is the sum of a set of Hilbert spaces representing zero particle states, one particle states, two particle states, and so on, where all the particles in each state are identical.⁶⁰ If the identical particles are bosons, the n-particle states are vectors in a symmetrized tensor product of n single-particle Hilbert spaces H. If the identical particles are fermions, the n-particle states are vectors in an anti-symmetrized tensor product of n single-particle Hilbert spaces H. A general state in Fock space is a linear combination of n-particle states, one for each n. In this way, it can represent multi-particle states, a goal of QFT.

Formally, the **Fock space** is the (Hilbert) direct sum of tensor products of copies of a single-particle Hilbert space H. Meaning that it is a sum of states of n particles, each of which is a properly symmetrized tensor product of n single-particle Hilbert spaces. The zero-particle state is simply a complex scalar. Then a general state s given by

$$|\Psi\rangle_\nu = a|0\rangle \oplus \sum_i a_i |\psi_i\rangle + \sum_{ij} a_{ij} |\psi_i, \psi_j\rangle_\nu \oplus \dots$$

where

- $a_i, a_{ij} \dots$ are complex coefficients,
- $|\psi_i\rangle \in \mathcal{H}$ is a state in the single-particle Hilbert space, and
- $|\psi_i, \psi_j\rangle_\nu = a_{ij} |\psi_i\rangle \otimes |\psi_j\rangle + a_{ji} |\psi_j\rangle \otimes |\psi_i\rangle$, are appropriately symmetrized (ν) two-particle Hilbert

⁵⁸ Teller, 8.

⁵⁹ Srednicki, 11; Lancaster & Blundell, 31.

⁶⁰ Most this section is inspired by Wikipedia, Fock space. https://en.wikipedia.org/wiki/Fock_space

states, and so on.

This can also be expressed as saying that a Fock space is

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} \mathcal{F}_N = \mathcal{F}_0 + \mathcal{F}_1 + \mathcal{F}_2 + \dots,$$

where \mathcal{F}_0 is the vacuum state, \mathcal{F}_1 is the single-particle Hilbert space, and the others are, as before appropriately symmetrized multiple Hilbert spaces.

A product state of the Fock space is a state of the form

$$|\Psi\rangle_\nu = |\phi_1, \phi_2, \dots, \phi_n\rangle_\nu = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \dots \otimes |\phi_n\rangle$$

which describes a collection of n particles, one in each state ϕ_i . The general state in a Fock space is a linear combination of product states. The fact that identical particles in QM are indistinguishable requires that all the particles in a given Fock space be identical.

A useful and convenient basis for a Fock space is the **occupation number basis**.⁶¹ Given a basis $\{|\psi_i\rangle\}_{i=0,1,2,\dots}$ of H , we can denote the state with n_i particles in state $|\psi_i\rangle$, going up to n_k particles in state $|\psi_k\rangle$ and none beyond by defining

$$|n_0, n_1, \dots, n_k\rangle = |\psi_0\rangle^{n_0} |\psi_1\rangle^{n_1} \dots |\psi_k\rangle^{n_k}, \quad (5.22)$$

with n_i being the number of (identical) particles in quantum state i . This state can be constructed from creation and annihilation operators as in (5.21). For bosons, n_i can take on any zero or positive value; for fermions, only 0 or 1. Note that we only see agglomerations of identical particles with no way of identifying any one of them, which satisfies Mr Teller's dislike of primitive thisness, mentioned in section 5.3.

5.5. The ubiquity of simple harmonic oscillators

Now for a look at where we are going. It all revolves around the way in which simple harmonic oscillators pop up in QFT, bringing along their extraordinarily useful math.⁶²

Let's start by comparing quantum fields with classical ones. A particle in classical physics is an object at a specific location in space-time. In QM, a particle wave function is a function in space of one particle, its coordinates and/or momenta. A classical field is an entity which exists everywhere in space-time, even where it has zero value. The wave function of a quantum field is like none of these. A quantum field has a wave function which is not a function of one point \vec{x} but of the field configuration *in all of space*, a function of a function, or **functional**, which we write as

$$\Psi(\{\phi(x)\}).$$

The variable of the wave function here is the whole quantum field configuration, $\{\phi(x)\}$, so the wave function is the probability amplitude for a complete configuration of a field in all of space. Then $|\Psi[\phi(x)]|^2$ is the probability of finding the field in configuration $\{\phi(\vec{x})\}$, where \vec{x} ranges over all space for *this* configuration. Conceptually, Ψ measures the probability of simultaneously measuring the value of ϕ throughout space, even though we cannot do this in reality.

A classical scalar field is an object which maps real space (\mathbf{R}^3) to a real number (\mathbf{R}); the quantum wave function of the whole field maps space to a complex number (\mathbf{C}). That the basic variables of the QFT wave function be the fields is evident in the definition of the conjugate momentum (density), where the derivative is not

⁶¹ Also called the occupancy number basis.

⁶² This and the next section are based on Sean Carroll, "The biggest ideas in the universe: fields", <http://www.preposterousuniverse.com/blog/2020/05/19/the-biggest-ideas-in-the-universe-9-fields/>. Also Carroll (2024), chap. 4.

with respect to the time differential of a position variable but to that of a field.

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}.$$

Okay, but how to get particles out of all these fields? Let's consider the simplest case, non-interacting fields, or "free" fields. Just as in one dimension we can use Fourier transforms, as in equation (5.3), to write any function as a sum of sine waves, so in 3-d, we can write the spatial field configuration as a sum of plane waves. We call these **modes**, each of which will be specified by a value of \vec{k} , the wave number, and the "height", or amplitude, of the wave. Remember that the transform integrates (or sums) over all values of position, so the modes no longer depend on that. In this way, we have passed from a position (\vec{x}) representation to a momentum (\vec{k}) one.

Now there are three sorts of energy for each mode:

- kinetic energy = $\frac{1}{2}$ (rate of change of field over time)²,
- **gradient energy** = $\frac{1}{2}$ (rate of change of field over space)²,
- potential energy = $\frac{1}{2} m$ (field value)².

These are really energy densities in QFT.

$$\rho(x, t) = \frac{1}{2}(\partial_t \phi)^2 + \frac{1}{2}(\partial_x \phi)^2 + V(\phi)$$

The new, non-classical term, the gradient energy, as in (5.56), is due to the fact that the argument of the wave function, the field, varies over space as well as time. It can be considered a shear term. Of course, kinetic and gradient energy can be taken together as a 4-d energy term representing rate of change over spacetime. So the 4-d Lorentz-invariant Lagrangian density for one mode is WHERE[5.4/115 a free-scalar [?WHAT?]] Changing the height, h , of the plane waves will change all three energies as the square of h , which is the same behavior as that of a simple harmonic oscillator (SHO). The system of a single mode can therefore be thought of as a SHO and that is a Good Thing, because we know the solutions to the Schrödinger equation for a SHO. But don't forget, we are considering equations for free particles, which SHOs are not, as they are due to a potential energy term.

The wave function of a complex field can be broken down via a Fourier transform into modes corresponding each to a single value of momentum (\vec{k}), where each mode depends on a single variable, h , and no longer on position \vec{x} . We can consider each mode as being fixed by the value of \vec{k} and being a function of the height, $f_{\vec{k}}(h)$. We have thus replaced a complicated problem by a sum of simple ones. These will include the solutions to the equations of motion for free particles of spin 0, $\frac{1}{2}$ and 1.

Now we can use all the tricks we learned about raising, lowering and number operators for SHOs in QM (section 5.2.3), and these lead us to an interpretation. Energy eigenvalues are equally spaced, being given by

$$E_n = \left(n + \frac{1}{2}\right) \hbar \omega, \quad (5.23)$$

so by analogy, we may consider each level n as corresponding to a wave function Ψ_n of n identical particles, each one of energy $\hbar \omega$. Then Ψ_0 is the wave function for no particles at all, the vacuum, represented by the ket $|0\rangle$ in Hilbert space.

This is all for free particles, without any interactions. We can consider a slightly more complicated model, including some interactions. Spoiler: It's SHOs all the way down.

Consider a linear chain of N objects, each of mass m , and connected by springs. The normal position of particle j is x_j and it can be displaced by a small distance x_j . The Hamiltonian for the system is

$$\hat{H} = \sum_j \frac{\hat{p}_j^2}{2m} + \frac{1}{2} K (\hat{x}_{j+1} - \hat{x}_j)^2. \quad (5.24)$$

Since the structure is a sum of oscillators, its overall behavior is likely represented by a sum of waves of varying periodicity. So it is reasonable to look at what is called its **reciprocal space**, i.e., its decomposition in terms of frequencies of vibration.⁶³ The quantities which measure the wave distribution are simply the Fourier transforms from \hat{x} and \hat{p} representations to the (ω, \vec{k}) representation. By doing the Fourier transforms and enforcing periodic boundary conditions, one can define raising and lowering operators in the reciprocal space and derive a Hamiltonian of the same form as equation (5.111) summed over the different modes of the particles.⁶⁴

$$\hat{H} = \sum_{k=1}^N \hbar \omega_k \left(\hat{a}_k^\dagger \hat{a}_k + \frac{1}{2} \right). \quad (5.25)$$

The real-space coupled, oscillating masses behave in reciprocal space, the (ω, \vec{k}) representation, as if they were independent harmonic oscillators, – as if uncoupled! These **modes** of the particles are also called **phonons**. Each such phonon mode can possess energy in an integral number of quanta – just like particles can. Now what we have is a number of particles in particular energy quantum states, with n_i particles in the i^{th} energy state. So we could write the state in the so-called **occupation-number representation** as

$$|n_1, n_2, \dots, n_N\rangle.$$

In order to satisfy the goal of representing multi-particle states, QFT uses Fock spaces, a **Fock space** being an extension of Hilbert space to include many-particle states. For the relation between Fock spaces and the occupation-number representation, see section 5.4.

The phonon modes are now comparable to a (product) system of N independent oscillators, each one possessing n_{p_m} quanta of energy $\hbar\omega$. So we see an analogy between completely different systems, one of harmonic oscillators and one of identical particles.⁶⁵ We will use an identification like this in quantum field theory when we consider **second quantization**. In the case of the EM field, the phonons will be photons, showing how a *field can exhibit particle characteristics*.

Identical particles	SHO
Particles in momentum states	Quanta in oscillators
m^{th} momentum mode p_m	k^{th} oscillator
$E = \sum_{m=1}^N n_{p_m} E_{p_m}$	$E = \sum_{k=1}^N n_k \hbar \omega_k$

Using creation and annihilation operators like those of equations (Error: Reference source not found) and (Error: Reference source not found) but with subscripts to denote the different particles, one can show that the reciprocal coordinate is given by⁶⁶

⁶³ Reciprocal space, http://goodwin.chem.ox.ac.uk/goodwin/TEACHING_files/l1_handout.pdf.

⁶⁴ Lancaster and Blundell do it, 25-27.

⁶⁵ Lancaster and Blundell, 30.

⁶⁶ Lancaster and Blundell, 27, problem (2.3).

$$\hat{x}_j = \frac{1}{\sqrt{N}} \left(\frac{\hbar}{N} \right)^{1/2} \sum_k \frac{1}{(2\omega_k)^{1/2}} [\hat{a}_k e^{ikja} + \hat{a}_k^\dagger e^{-ikja}]. \quad (5.26)$$

This represents a time-independent **position operator** and is called the **mode expansion** of the scalar field.⁶⁷

Thus,

QFT can be interpreted as a theory of particles, because the discrete set of solutions to the Schrödinger equation⁶⁸ with correspondingly discrete equally-spaced energies can be thought of as states of different numbers of particles, each being an excitation of the field.

When we do this with the EM field, the energy states are called photons.⁶⁹ Now we can view the SHO ladder operators as creation and annihilation operators. We will be seeing this statement again. We will also see how the $\frac{1}{2}\hbar\omega$ terms can be considered the non-zero (!) value of the energy of the vacuum.

We have been considering non-interacting fields, without any interactions between particles. But it is impossible to observe particles without using interactions in order to do so. Also, interactions between particles are local in space. At $x = \pm\infty$, i.e., when the particles are infinitely far apart, any interaction term goes to zero.

Nutshell: “One quantum field can be thought of as a superposition of different numbers of particles.” So we can “invent” QFT either by starting with particles or with fields.

One more time:

Free-particle fields can be expressed as sums of SHOs, each of momentum \vec{k} . We can therefore use all the paraphernalia of diagonal and ladder operators to find eigenvalues of energy, which we can then interpret not just as quanta of energy or excitations of the field, but as particles.

Or:

“We describe the Universe by combining fields to form a Lagrangian density $\mathcal{L}[\phi(x)]$. Our canonical quantization process often allows us to quantize these fields leading to a Universe pictured as a vacuum acted on by field operators \hat{a}^\dagger . The excitations of the vacuum that the field operators produce are particles and antiparticles.”⁷⁰

Or again:

“Every particle and every wave in the Universe is simply an excitation of a quantum field that is defined over all space and time.”⁷¹

So much for an introduction to the idea; now on to the nitty-gritty.

5.6. Action and energy in field theory

Quantum Lagrangians may be constructed from symmetry considerations, as we have done in preceding paragraphs.. They may also be constructed by analogy with classical ones through the process of quantization, such as by replacing momentum, p_x , by the operator $-i\hbar\partial_x$.

Leonard Susskind cites four observed principles of physical laws.⁷²

- The **action principle**, which depends on the Lagrangian, without which there is no conservation of

⁶⁷ You do NOT have to remember this.

⁶⁸ Or the equations of motion derived from the Lagrangian by using the Euler-Lagrange equations.

⁶⁹ Lancaster and Blundell, 27.

⁷⁰ Lancaster & Blundell, 154.

⁷¹ Ibid, 1.

⁷² Susskind and Friedman (2017), 256-259.

energy or momentum or the relationship between them and symmetry.⁷³

- **Locality**, which means that an event at one place can immediately affect only nearby places. This requires continuity of the Lagrangian and fields, which in turn says that the Lagrangian depends on the field and its first derivatives with respect to position.
- **Lorentz invariance**, so that the equations of motion remain the same in every Lorentz-related reference frame, as required by SR. The involved 4-vectors may change, but the relations between them remain the same; they are then said to be **Lorentz covariant**. Lorentz invariance of the equations of motion is guaranteed if the Lagrangian is a scalar.
- **Gauge invariance**, such as for the EM field, where there are many equivalent descriptions of a system differing by arbitrary functions of position.

Quantization of classical Lagrangians must take these principles into account.

Let's start by comparing quantum fields with classical ones or non-relativistic ones.

- A particle in classical physics is an object with coordinates of one point in space. The basic variable is $x(t)$, which is a function of time. (How could it be a function of itself?) The Euler-Lagrange equation for a particle is

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i} \quad (2.3)$$

In these equations, the coordinates q_i are the spatial location of the particle and so represent degrees of freedom of the particle. In general, it is moving, so the q_i are functions of time and \dot{q}_i is non-zero.

- In QM, a particle wave function is a function in space of one particle, its coordinates and/or momenta. We can construct its Lagrangian in terms of these variables and use the Euler-Lagrange equation (2.3) to solve for its state of motion. The probability that it occupy a specific position in space is given by the square of the wave function at that position.
- A classical field is an entity with its own equations of motion (such as Maxwell's equations) but which exists in all space-time, even where it has zero value.⁷⁴

The wave function of a quantum field is like none of these. A quantum field has a wave function which is not a function of one point \vec{x} like a classical field, but of the field configuration *in all of space*, which we write as

$$\Psi(\{\phi(x)\}).$$

In short, $x^\mu \rightarrow \phi(x^\mu)$. The variable of the wave function here is no longer the position x . Instead, that it is replaced by the whole quantum field $\{\phi(x)\}$, so the wave function is what is called a **functional**. For this reason, the canonical momentum is no longer the derivative with respect to $\frac{\partial x}{\partial t}$, but the **canonical momentum density**, of which the derivative is with respect to $\frac{\partial \phi}{\partial t} = \dot{\phi}$:

$$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}. \quad (5.27)$$

Here we have introduced the notion of momentum density. Locality means a field depends on its value and its derivatives in the neighborhood of a point, so we can consider its value at every point and then integrate to get the total field. Also, Lorentz invariance demands an equal treatment of space and time variables, which suggests a **density** in time and space for the **Lagrangian** as well as the **Hamiltonian**, and so the **energy**, given by

⁷³ Susskind and Friedman (2017), 331.

⁷⁴ Uh...

italicized letters:

$$L = \int \mathcal{L}(x, t) d^3x dt, \quad H = \int \mathcal{H}(x, t) d^3x dt, \quad E = \int \rho(x, t) d^3x dt. \quad (5.28)$$

The action is

$$Action = \int dt L(\phi, \dot{\phi}, \partial_x \phi) = \int dt \int d^3x \mathcal{L}(\phi, \dot{\phi}, \partial_x \phi) \quad (5.29)$$

and

$$\mathcal{H} = \pi \dot{\phi} - \mathcal{L}. \quad (5.30)$$

In QFT, the fields are considered to be variables, so the canonical commutation relation is

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\delta^{(4)}(x - y), \quad (5.31)$$

again in terms of the field. The **Euler-Lagrange equations for a field**, taking into account the spatial variations of the field can be calculated from the minimum of the Action in the same way, using differentiation by parts.⁷⁵

The Euler-Lagrange equations in Minkowski space then are given by

$$\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} = \frac{\partial \mathcal{L}}{\partial \phi}. \quad (5.32)$$

Then $|\Psi[\phi(x)]|^2$ is the probability of finding the field in configuration $\{\phi(\vec{x})\}$.

A quick comparison of these matters is given in the following table.

	Particle	Field
Basic variable	\vec{x}	$\phi(x)\}$
Canonical momentum	$p(x) = \frac{\partial L}{\partial \dot{x}}$	$\pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}}$
Wave function	$\Psi(x)$	$\Psi(\{\phi(x)\})$, also written $\Psi(\phi)$
Euler-Lagrange equation	$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0$	$\frac{\partial}{\partial x^\mu} \frac{\partial \mathcal{L}}{\partial \phi_{,\mu}} = \frac{\partial \mathcal{L}}{\partial \phi}$
Canonical commutation relation	$[\hat{x}, \hat{p}_x] = i\hbar$	$[\hat{\phi}(x), \hat{\pi}(y)] = i\delta^{(4)}(x - y)$

Consider the complex version of the Klein-Gordon Lagrangian:

$$\mathcal{L} = \frac{1}{2} [(\partial_\alpha \phi^\dagger \partial^\alpha \phi - m^2 \phi^\dagger \phi)] = \frac{1}{2} [(\dot{\phi}^\dagger \dot{\phi} - \nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi)].$$

The conjugate (canonical) momentum density is then

$$\pi^\mu(x) = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi(x))} = \partial^\mu \phi(x).$$

So the Hamiltonian density is

$$\mathcal{H} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \dot{\phi}^r - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} + \frac{\partial \mathcal{L}}{\partial \dot{\phi}^\dagger} \dot{\phi}^\dagger - \mathcal{L} =$$

⁷⁵ Klauber, 18-19.

$$\frac{1}{2}(\dot{\phi}^\dagger \dot{\phi} + \nabla \phi^\dagger \cdot \nabla \phi + \mu^2 \phi^\dagger \phi) = \frac{1}{2}(\partial_\mu \phi^\dagger \partial^\mu \phi + \mu^2 \phi^\dagger \phi), \quad (5.33)$$

where r indicates fields considered different, in this case ϕ and ϕ^\dagger . The Euler-Lagrange equations lead to two equations of motion, one for each field:

$$\begin{aligned} (\partial_\mu \partial^\mu + \mu^2)\phi &= 0, \\ (\partial_\mu \partial^\mu + \mu^2)\phi^\dagger &= 0. \end{aligned}$$

We could also have derived this equation by using the classic dispersion relations

$$p^\mu p_\mu - m^2 = 0$$

and substituting the standard momentum operator $p_\mu \rightarrow i\hbar \partial_\mu$, letting the result act on a field as in (4.2). Solutions are Fourier expansions of the position field in terms of momentum fields ($-ikx = -i(Et - \vec{p} \cdot \vec{x})$), the sums or integrals of plane-wave equations:

$$\phi(x) = \int \frac{d^3 \vec{k}}{\sqrt{2(2\pi)^3 \omega_k}} a(\vec{k}) e^{-ikx} + \int \frac{d^3 \vec{k}}{\sqrt{2(2\pi)^3 \omega_k}} b^\dagger(\vec{k}) e^{ikx} := \phi^+ + \phi^-, \quad (5.34a)$$

$$\phi^\dagger(x) = \int \frac{d^3 \vec{k}}{\sqrt{2(2\pi)^3 \omega_k}} b(\vec{k}) e^{-ikx} + \int \frac{d^3 \vec{k}}{\sqrt{2(2\pi)^3 \omega_k}} a^\dagger(\vec{k}) e^{ikx} := \phi^{\dagger+} + \phi^{\dagger-}. \quad (5.34b)$$

(Don't forget, these fields are operators and so should have hats on.)

Now, second quantization requires the canonical commutation relation of the field to be those of (5.31). Imposing this on the plane-wave equations shows that similar commutation relations must hold for the coefficients $\hat{a}(\vec{k})$ and $\hat{b}(\vec{k})$, now attired with the appropriate head-wear:

$$\begin{aligned} [a(\vec{k}), a^\dagger(\vec{k}')] &= [b(\vec{k}), b^\dagger(\vec{k}')] = \delta_{\vec{k}\vec{k}'} && \text{(discrete);} \\ &= \delta(\vec{k} - \vec{k}'). && \text{(continuous).} \end{aligned}$$

These are the same as the commutation relations for the ladder operators of SHOs, which should not be surprising given the form of the plane-wave solutions. Plugging the solutions (5.34) into the Hamiltonian density (5.33) shows that the free-particle Hamiltonian $H = \int \mathcal{H} d^3x$ becomes

$$H = \sum_k \omega_k \left(N_a(\vec{k}) + \frac{1}{2} + N_b(\vec{k}) + \frac{1}{2} \right),$$

where the **number operators** are

$$N_a(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}); \quad N_b(\vec{k}) = b^\dagger(\vec{k})b(\vec{k}).$$

A number operator is interpreted as returning the number of particles in the state, i.e., of that energy. And the a 's and b 's are **creation** and **destruction operators**.

$$\begin{aligned} a^\dagger(\vec{k})|n_{\vec{k}}\rangle &= \sqrt{n_{\vec{k}} + 1}|n_{\vec{k}} + 1\rangle \\ a(\vec{k})|n_{\vec{k}}\rangle &= \sqrt{n_{\vec{k}}}|n_{\vec{k}} - 1\rangle \\ b^\dagger(\vec{k})|\bar{n}_{\vec{k}}\rangle &= \sqrt{\bar{n}_{\vec{k}} + 1}|\bar{n}_{\vec{k}} + 1\rangle \quad b^\dagger(\vec{k})|\bar{n}_{\vec{k}}\rangle = \sqrt{\bar{n}_{\vec{k}} + 1}|\bar{n}_{\vec{k}} + 1\rangle \\ b(\vec{k})|\bar{n}_{\vec{k}}\rangle &= \sqrt{\bar{n}_{\vec{k}}}| \bar{n}_{\vec{k}} - 1\rangle. \end{aligned}$$

So in order to pass from particle mechanics to field mechanics, we make the following correspondences:

- $t \rightarrow x^0$, $x^i(t) \rightarrow \phi(x^i)$, so $x^\mu \rightarrow \phi(x^\mu)$
- $L, H \rightarrow \mathcal{L}, \mathcal{H}$ = Lagrangian and Hamiltonian densities
- The conjugate momentum of the field is given by (Error: Reference source not found).
- The Euler-Lagrange equations, the definition of the Hamiltonian and the action are those of (5.32), (5.30) and (5.29).

5.7. EM field Lagrangian and Maxwell's equations

For the EM field, making the Lagrangian Lorentz invariant is accomplished by constructing it from Lorentz-invariant quantities, scalars built from the components of the Lorentz-covariant EM field tensor $F_{\mu\nu}$, defined in terms of the 4-d vector potential in equation (4.6). The Lagrangian without charges or currents then is usually written

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}. \quad (5.35)$$

Doing the sums and remembering minus signs from raising or lowering indices, \mathcal{L} reduces to⁷⁶

$$\mathcal{L} = \frac{1}{2}(E^2 - B^2). \quad (5.36)$$

The factors of $\frac{1}{2}$ or $\frac{1}{4}$ are conventions which are without physical importance. Substituting the vector potential components from (4.6) leads to the Lagrangian for the EM field

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu). \quad (5.37)$$

This is the equation for no charges or current. If such are present, we must add a term so that the **Lagrangian for the EM field with currents** is

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) + J^\mu(x)A_\mu(x), \quad (5.38)$$

where

$$J^\mu = (\rho, J^i), \quad (5.39)$$

the charge density-current four-vector. From this Lagrangian, the Euler-Lagrange equations, which in this context are

$$\frac{\partial}{\partial x^\nu} \frac{\partial \mathcal{L}}{\partial A_{\mu,\nu}} = \frac{\partial \mathcal{L}}{\partial A_\mu},$$

may be used to derive Maxwell's equations starting with the vector potential in the EM field tensor and the requirement of gauge invariance, which in turn requires that the current 4-vector obey the **continuity equation**⁷⁷

$$\partial_\mu J^\mu = 0. \quad (5.40)$$

⁷⁶ Susskind and Friedman (2017) 339-341.

⁷⁷ Susskind (2017), 341-346. The factor of $\frac{1}{4}$ is a convention, according to Susskind.

5.8. The Schrödinger and Heisenberg pictures

The classic view of QM is the so-called Schrödinger picture. The time-dependent Schrödinger equation in the case where the Hamiltonian is independent of time leads to the unitary time-evolution operator $\mathbf{U}(t)$ such that

$$|\Psi(t)\rangle = \mathbf{U}(t)|\Psi(0)\rangle$$

with

$$\mathbf{U}(t) = e^{-\frac{i}{\hbar}Et}.$$

In the Heisenberg picture, however, the state is independent of time, so taking the Schrödinger state at zero,

$$|\Psi_H(t)\rangle = \hat{U}_S^\dagger(t)|\Psi_S(t)\rangle = e^{i\hat{H}t/\hbar}|\Psi_S(t)\rangle = |\Psi_S(0)\rangle. \quad (5.41)$$

The expectation value of a Hermitian operator L in terms of the time-evolution operator then can be written as

$$\langle L \rangle = \langle \Psi(t) | L | \Psi(t) \rangle = \langle \Psi(0) | \mathbf{U}^\dagger(t) L \mathbf{U}(t) | \Psi(0) \rangle,$$

which we can group in two ways⁷⁸

$$\langle L \rangle = \langle \Psi(0) | \mathbf{U}^\dagger(t) [L] \mathbf{U}(t) | \Psi(0) \rangle = \langle \Psi^S(t) | L^S | \Psi^S(t) \rangle \quad (5.42)$$

or

$$\langle L \rangle = \langle \Psi(0) | [\mathbf{U}^\dagger(t) L \mathbf{U}(t)] | \Psi(0) \rangle = \langle \Psi^H | L^H(t) | \Psi^H \rangle, \quad (5.43)$$

both being equivalent in result. But the grouping of (5.42) represents a constant operator in acting on wave functions changing in time, whereas (5.43) represents a time-dependent operator acting on time-constant wave functions. The former, to which we are most used, is called the **Schrödinger picture** of QM (denoted by a superscript S); the latter, the **Heisenberg picture** (superscript H). The Heisenberg picture is the one most used in field theory. This only applies to (passive) time evolution, not (active) translations in space or time.⁷⁹

To go from the Schrödinger to the Heisenberg picture,

- replace $\Psi(t)$ by $\Psi(0)$, and
- define the Heisenberg operator by

$$L^H(t) = \mathbf{U}^\dagger(t) L^S \mathbf{U}(t). \quad (5.44)$$

The first of these two steps means

$$\Psi^S(t) = U(t)\Psi(0) = U(t)\Psi^H,$$

so the second step requires

$$\Psi^H = U^\dagger(t)\Psi^S. \quad (5.45)$$

Differentiate by parts and use $\frac{dL^S}{dt} = 0$ to get

$$\frac{dL^H(t)}{dt} = -\frac{i}{\hbar}[L^H(t), H], \quad (5.46)$$

where H in the commutator means the Hamiltonian. This is the **Heisenberg equation of motion**. In classical mechanics in terms of Poisson brackets, $\{\cdot, \cdot\}$,

⁷⁸ Lancaster and Blundell, 74.

⁷⁹ Lancaster and Blundell, 80.

$$\frac{du}{dt} = \frac{\partial u}{\partial q_i} \dot{q}_i + \frac{\partial u}{\partial p_i} \dot{p}_i + \frac{\partial u}{\partial t} = \{u, H\} + \frac{\partial u}{\partial t}, \quad (5.47)$$

where the last term is usually zero. Comparison of these two equations suggests that part of quantization will be the replacement

$$\{u, H\} \rightarrow \frac{-i}{\hbar} [u, H]. \quad (5.48)$$

Incidentally, the expectation value for any quantum field – scalar, spinor or vector – is zero, so fields themselves are unmeasurable. Fortunately, this is not the case with operators such as energy, momentum or charge.⁸⁰ But measurable quantities have the same expectation value in both pictures.

5.9. Second quantization

There are two methods of introducing QFT, each being useful according to the circumstances.

- The more historical formulation is **canonical quantization**, which seeks to quantify a classical theory while maintaining its structure, particularly its symmetries. It is based on the correspondence (5.48) between classical Poisson brackets and QM commutations. Or on the replacement of the variable momentum by an operator. This was discussed in section 5.1.
- The more modern method is the **path integral** formulation, which is a generalization of the action principle of classical mechanics, based on Feynman's notion of a sum over all possible paths.

One method may be more useful than the other depending on the situation studied. We will consider the path integral formulation in section 8.

In order to pass from classical to quantum mechanics, **first quantization** quantizes momenta and coordinates by promoting them to be operators. It results in non-commutation of the coordinates and their respective conjugate momenta.

$$[q^i, p_j] = i\hbar\delta_j^i \quad \text{with} \quad p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}^i} \quad (5.49)$$

The result is the equation for energy (Schrödinger's equation), in terms of position and momentum as operators, and introducing a *new* object, the **wave function**.

$$H\Psi = i\hbar\partial_t\Psi.$$

The main characteristics of the wave function which interest us are:

- It's a **field**, having a value at each point in space.
- It's a probability wave, an amplitude, the modulus-square of which gives the probability that the particle be in the corresponding state of motion (Born rule).
- Permitted values of variables are given by eigenfunctions of the wave function operated on by the appropriate operator.
- Although this is not always clearly stated, the wave function represents one **particle** moving in space, whose coordinates \vec{x} are functions of time, t .

Formally, the wave function is related to the state vector in Hilbert space, which is complex and conserves inner products. Coordinates and momenta operate on the particle represented by the state vector. We now have equations of NRQM.

⁸⁰ Klauber, 190. He says there are exceptions "... more advanced areas of QFT."

Making NRQM relativistic is done by writing (via informed guesswork) an appropriate Lagrangian which is Lorentz-invariant and which reduces to the NRQM value under non-relativistic conditions. Lorentz invariance imposes certain behaviors on the Lagrangian, simplifying its confection. This allows us to write Lorentz-invariant Lagrangians of particles for various values of spin -- Klein-Gordon, Dirac and Proca, as we saw in section 4. Each result is an equation where conjugate positions and momenta are non-commuting operators. From each Lagrangian we use the Euler-Lagrange equations to derive the equation of motion for a wave function describing a single free particle moving in space, so the coordinates \vec{x} are functions of time. Now we have NRQM \rightarrow RQM.

Next, we do almost the same thing we did before, but with an extraordinarily important difference. As we have noted, the Schrödinger equation handles the coordinates quadratically but time linearly. Also, since first quantization, the coordinates are considered operators, but time is not. The (easier) solution to these difficulties is to backtrack and no longer consider the coordinates as operators, but to take the *particle* as the operator.⁸¹

Second quantization (or field quantization) quantizes the field (wave function), interpreted as being a particle. So the non-commutation is taken to be that of the field itself and its own canonical momentum, π .

$$[\phi(x, t), \pi(x', t')] = i\hbar\delta(t - t')\delta(x - x') \quad \text{with} \quad \pi = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \quad (5.50)$$

In words, a quantum field and its associated canonical momentum do not commute, but every other pair of fields and momentum densities does. Note that the operator fields are parametrized by the spacetime coordinates, but those coordinates variables do not describe the state of a stationary or moving particle, but simply the place in space where the field is evaluated. Whereas position coordinates of particles evolve in time, those of fields do not.

But how do these operators operate to produce particles?

The answer is, we can write these fields in terms of other operators, notably, creation, annihilation and counting (number) operators. A creation operator converts a state of n particles into one of $n+1$ particles. In particular, it creates a single particle on the vacuum state.

$$a^\dagger(\vec{k})|0\rangle = |1\rangle$$

Multiple creations then lead to any numbers of particles in an energy state or different energy states. We describe this using a **Fock space**. (More later ... maybe.)

If this procedure seems somewhat arbitrary, it is worth remembering that neither Newton nor Maxwell nor Einstein derived their respective equations from other theories, but rather by letting experimental results and general principles inform inspired guesswork (and lots of math).

Summarizing, we recall that the *canonical quantization* from classical physics to quantum field theory goes through two conceptual steps of quantization of the Lagrangian:

1. The first step promotes conjugate pairs of variables, in particular \vec{x} and \vec{p} , to operators, which is equivalent to imposing commutation relations on them. This procedure takes us from classical to quantum mechanics by the **canonical commutation relation**:

$$[\hat{x}, \hat{p}] = i\hbar. \quad (5.51)$$

It's called canonical commutation because it concerns the commutation of canonical (or conjugate) variables. Since the operators must operate on something, we must introduce a new notion, the wave function (or state vector) with the properties we all know and love, including that it is a field.

2. Then the jump to QFT requires considering the fields themselves not as states but as operators, which must themselves obey commutation relations (5.50). Since they are essentially operators for the

81 Robinson, 143.

creation, annihilation and counting of particles, they operate on the vacuum ket.

The requirements of SR also must be taken into account and this can be done before or after quantization. In other words, “relativize” and “quantize” – in either order.

We will employ these concepts by following the schema of Lancaster and Blundell. wherein **canonical quantization** is a series of steps for going from a classical field theory to a quantum one. They see the method as comprised five steps.⁸²

- I. Start by preparing a Lagrangian for a classical relativistic field, considered as complex. This step generally involves consideration of symmetries plus some creative guesswork. Write this as a Lagrangian density. It is then useful to employ the Euler-Lagrange equation to derive the equations of motion and find their solutions, which will be interpreted as fields.
- II. Evaluate the momentum and Hamiltonian densities in terms of fields using (Error: Reference source not found) and (Error: Reference source not found). Find the energy in terms of number operators.
- III. Quantize the fields by imposing commutation relations (5.50) on them, thus treating the fields as *operators*. Enforcement of the commutation relations on the fields will lead to commutation relations for the amplitudes.
- IV. Work out the fields in terms of creation and annihilation operators based on the amplitudes. These would allow creation, annihilation and, so, mutation of particles, the basic steps of particle interactions.
- V. Use normal ordering (section 5.10.7) to avoid infinite energies.

As an example, consider a massive scalar field (meaning one with mass, not necessarily a very huge one). Note that we are using the **metric signature**

$(+, -, -, -)$.

Would that there were a standard (preferably, this one). There almost is, at least among field theorists. Table 2 tells it all. The last column is explained later in section 5.11.

⁸² Adapted from Lancaster and Blundell, 98ff.

Author(s)	metric signature	most-used γ matrix basis
Barr et al.	+---	both
Blundell & Lancaster	+---	chiral
Carroll	-+++	
Griffiths	+---	Dirac
Jeevanjee	+++− (!)	NA
Klauber	+---	Dirac
Maggiore	+---	chiral
Peskin & Schroeder	+---	chiral
Robinson	-+++	chiral
Schwartz	+---	chiral
Schwichtenberg	+---	chiral
Srednicki	-+++	
Susskind ⁸³	-+++	NA
Zee	?	both

Table 2. Comparison of metric signatures and γ bases by author.

Among QFT specialists, (+1,-1,-1,-1) does seem to be the standard (Sorry, Messrs Robinson and Srednicki) – to the extent that there is one.

5.10. Relativistic scalar Lagrangian (Klein-Gordon)

We start with the classical version.

5.10.1. Classical relativistic scalar Lagrangian

The standard classical relativistic Lagrangian density for a massive scalar field starts from equation (5.30) with a potential energy term as follows:

$$\mathcal{L} = \frac{1}{2}[\partial_\mu \phi(x)]^2 - \frac{1}{2}m^2[\phi(x)^2]. \quad (5.52)$$

This form was initially proposed because it is a scalar composed of Lorentz-invariant four-vectors and it represents the SR dispersion relation linking energy, momentum and mass. Its ultimate adoption is due to the fact that it works.

Comparing this to the classical Lagrangian for a particle taken as a SHO,

⁸³ Susskind claims (-1,1,1,1) is generally used in GR and he seems to be right. It is also used by Carroll, Schutz and Hartle, but not by Lambourne or Collier.

$$\mathcal{L}_{SHO} = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}m\omega^2 x^2,$$

shows clearly that instead of considering the spatial position of our object (particle), we are talking about the configuration of a field $\phi(x)$. SO SHOULD WE NOT be talking about $\{\phi(x)\}$, the complete field over all of space (x)?

Since this is the classical relativistic Lagrangian, the field is real. The Euler-Lagrange equations really are easy to solve for the equation of motion of this Lagrangian and the result is the **Klein-Gordon equation**.⁸⁴

$$(\partial_\mu \partial^\mu + m^2)\phi = 0, \quad (5.53)$$

with

$$\mu^2 = \frac{m^2 c^2}{\hbar^2} = m^2$$

in natural units ($c = \hbar = 1$).⁸⁵

5.10.2. Complexification

In the above equation, ϕ is a real, classical field, not a wave function. For QFT, we must consider the field as being **complex**, not real. As a result, we must distinguish between the field ϕ and its complex conjugate ϕ^\dagger , so the **Lagrangian density** takes on the slightly more complex form

$$\mathcal{L} = \frac{1}{2}[(\partial_\alpha \phi^\dagger \partial^\alpha \phi - m^2 \phi^\dagger \phi)] = \frac{1}{2}[(\dot{\phi}^\dagger \dot{\phi} - \nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi)], \quad (5.54)$$

where the upper zero indicates spin (which we will see shortly) and the lower, a free particle.

We can calculate the **conjugate (canonical) momentum density**

$$\pi^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi(x))} = \partial^\mu \phi(x), \quad (5.55)$$

whose timelike component is just $\partial^0 \phi(x)$. Then the Hamiltonian density is

$$\begin{aligned} \mathcal{H} &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \dot{\phi}^r - \mathcal{L} = \frac{\partial \mathcal{L}}{\partial \dot{\phi}} \dot{\phi} + \frac{\partial \mathcal{L}}{\partial \dot{\phi}^\dagger} \dot{\phi}^\dagger - \mathcal{L} = \\ &= \frac{1}{2}(\dot{\phi}^\dagger \dot{\phi} + \nabla \phi^\dagger \cdot \nabla \phi + \mu^2 \phi^\dagger \phi) = \frac{1}{2}(\partial_\mu \phi^\dagger \partial^\mu \phi + \mu^2 \phi^\dagger \phi), \end{aligned} \quad (5.56)$$

where r indicates a sum over fields considered different, in this case ϕ and ϕ^\dagger . Now we see that the energy has three components – kinetic energy of motion in time, what we have called gradient energy in section 5.5 and potential energy (the mass term). Because the two fields ϕ and ϕ^\dagger are considered separate, the Euler-Lagrange equation leads to a Klein-Gordon equation for each field.⁸⁶ From here on, ϕ represents a wave.

$$(\partial_\mu \partial^\mu + \mu^2)\phi = 0 \quad (5.57a)$$

$$(\partial_\mu \partial^\mu + \mu^2)\phi^\dagger = 0 \quad (5.57b).$$

Care must be taken not to confuse the μ for mass and those in the subscripts and superscripts. Two different equations means two solutions, each of which can be expressed as a Fourier series of plane waves for the

⁸⁴ Klauber, 42.

⁸⁵ Sometimes referred to as “God units”. Robinson, 143,

⁸⁶ Klauber, 49.

discrete case

$$\phi(x) = \sum_k \frac{1}{\sqrt{2V\omega_k}} a(\vec{k}) e^{-ikx} + \sum_k \frac{1}{\sqrt{2V\omega_k}} b^\dagger(\vec{k}) e^{ikx} := \phi^+ + \phi^- \quad (5.58a)$$

$$\phi^\dagger(x) = \sum_k \frac{1}{\sqrt{2V\omega_k}} b(\vec{k}) e^{-ikx} + \sum_k \frac{1}{\sqrt{2V\omega_k}} a^\dagger(\vec{k}) e^{ikx} := \phi^{\dagger+} + \phi^{\dagger-} \quad (5.58b)$$

or as an integral for the continuous case

$$\phi(x) = \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_k}} a(\vec{k}) e^{-ikx} + \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_k}} b^\dagger(\vec{k}) e^{ikx} := \phi^+ + \phi^- \quad (5.59a)$$

$$\phi^\dagger(x) = \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_k}} b(\vec{k}) e^{-ikx} + \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_k}} a^\dagger(\vec{k}) e^{ikx} := \phi^{\dagger+} + \phi^{\dagger-}. \quad (5.59b)$$

Note that the second term in each solution is only found in solutions of the equations resulting from the Lagrangian (5.54) of relativistic quantum mechanics and not those of the Schrödinger equation. Because the energy operator

$$\hat{E} = i\partial_t,$$

(with $\hbar = 1$) solutions of these equations with $-ikx = -i(Et - \vec{p} \cdot \vec{x})$ in the exponent result in positive energies, but those with $+ikx = i(Et - \vec{p} \cdot \vec{x})$, in **negative** energies. So the distinction here is that the + terms, ϕ^+ and $\phi^{\dagger+}$, correspond to RQM states of *positive* energy (terms proportional to $e^{-ik \cdot x}$) whereas the - terms, ϕ^- and $\phi^{\dagger-}$ correspond to states of *negative* energy. So $\phi^{\dagger+}$ can not be the complex conjugate of ϕ^+ , but a different field altogether.⁸⁷ In fact, since \dagger changes the sign of the ikx exponent, $(\phi^-)^\dagger = \phi^{\dagger+}$. The + or - superscripts have nothing to do with whether they operate on particles or antiparticles. This is all exceedingly messy, but the message to take home about creation and annihilation operators will be derived soon as (5.65).

5.10.3. Convert fields to operators by imposing commutation relations

This is step III of second quantization.⁸⁸ Do this by imposing the commutation relations of equations (Error: Reference source not found) on the fields and by writing them as operators (with hats on them). Assuming the same value for time, this gives field commutators

$$[\hat{\phi}(\vec{x}, t), \hat{\pi}(\vec{y}, t)] = i\delta^{(3)}(\vec{x} - \vec{y}). \quad (5.60)$$

Or, in 4-d,

$$[\hat{\phi}(x), \hat{\pi}(y)] = i\delta^{(4)}(x - y). \quad (5.61)$$

Starting with the commutation relation for fields (5.60), plug in equations (5.58) with the x system corresponding to (ω, \vec{k}) and the y system to (ω', \vec{k}') and expand. The result shows that the coefficients, the a's and b's, are themselves operators and must satisfy the commutation relations⁸⁹

$$\begin{aligned} [a(\vec{k}), a^\dagger(\vec{k}')] &= [b(\vec{k}), b^\dagger(\vec{k}')] = \delta_{\vec{k}\vec{k}'} \quad (\text{discrete}); \\ &= \delta(\vec{k} - \vec{k}') \quad (\text{continuous}). \end{aligned} \quad (5.62)$$

⁸⁷ Klauber, 50.

⁸⁸ Or second postulate. Klauber, 157.

⁸⁹ Klauber, 52-53.

From commutation relations for the *fields*, we have deduced commutation relations for the *amplitudes*, or *coefficients*. This depends on the form of the solutions for the fields as sums of modes (Fourier series, or mode expansions).

5.10.4. From parameters to wave functions to fields to ...

Let's pause a moment to consider just what is going on.

In classical mechanics, we start with either Newton's laws or a Lagrangian in terms of coordinates and their derivatives (or momenta) and we can solve them for the behavior of the system in terms of these spacetime-dependant variables. Schrödinger quantized classical mechanics by making operators out of these quantities and this makes some of them non-commutating. As operators, they then needed something on which to operate, so the wave function (or state vector) was born. But what was that? With the Born rule, the absolute square of the wave function was seen to represent the probability of the system's being in a certain state, in which the state variables may take on eigenvalues given by the momentum and position operators operating on the state vector.

So now we have a wave function. But with the invention of QFT, second quantization forced the quantization of the wave functions of the field by the same trick of turning them into operators and imposing commutation rules on them. So what do they operate on? Enter a new state vector. It's a bit confusing here, because the Dirac bracket convention had already been used in ordinary QM, but it is used again to denote the "thingy" operated on by the field operators. These new QFT state vectors will be used to calculate expectation values of observables.

Since the fields are operators and evolve in time – by equations (5.58) or (5.59) – then we are in the **Heisenberg picture**. Discussions of the Dirac and Proca equations will also be in the HP. In fact, it can be shown that the Klein-Gordon field equation (5.53) is equivalent to the Heisenberg equation of motion (5.46).⁹⁰ We will see later, in section 7.1, that the time-evolution operator in the IP (interaction picture, a special case of the HP) is governed by the free part only of the Hamiltonian (or Lagrangian).

5.10.5. Number operators, Hamiltonian and energy

Once the amplitudes $a(\vec{k})$, $a^\dagger(\vec{k}')$, $b(\vec{k})$, $b^\dagger(\vec{k}')$ are known to satisfy the above commutation relations, then we can plug the solutions (5.58) into the Hamiltonian density (5.56) and calculate the Hamiltonian

$$H = \int \mathcal{H} d^3x.$$

The result is the free-particle Hamiltonian

$$H = \sum_k \omega_k \left(N_a(\vec{k}) + \frac{1}{2} + N_b(\vec{k}) + \frac{1}{2} \right), \quad (5.63)$$

expressed in terms of number operators.⁹¹

$$N_a(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}); \quad N_b(\vec{k}) = b^\dagger(\vec{k})b(\vec{k}). \quad (5.64)$$

As Klauber says, these two equations "lie at the heart of QFT..."⁹² Again, (5.63) refers to a free particle of spin 0. We shall see very soon that the a terms correspond to particles and the b terms to antiparticles.

In words, in an occupation number representation, $N_a(\vec{k})$ is the number operator whose eigenvalue $n_a(\vec{k})$ is the

⁹⁰ Klauber, 79-80.

⁹¹ The following notation is from Klauber, 54-55.

⁹² Klauber, 54.

number of a particles with 3-momentum \vec{k} in the state. The two extra factors of $\frac{1}{2}$ in (5.63) are considered to represent the **zero-point energy** (ZPE) or vacuum energy (VE). It has its origin here in the non-commutation of the fields and their conjugate momenta from second quantization. Although this is often interpreted as virtual particles popping in and out of existence, there is nothing in the equations for the number operators (which contain the time dependence in QFT) for any variation in the number, which however should vary constantly if virtual particles are ceaselessly created and destroyed. In addition, we have calculated (5.63) for a free field with no interactions with anything else, meaning no virtual particles are appearing or disappearing.⁹³ At least, not as far as we know at this point.

The form of equation (5.63) shows that all energies in QFT are positive, even for the b particles, which are antiparticles, as we shall soon see.

5.10.6. Creation and annihilation operators

Using the commutation relations (5.62) and the number operators (5.64), one can show, as for a SHO that the a's and b's are creation and destruction operators as follows:

$$\begin{aligned} a^\dagger(\vec{k})|n_{\vec{k}}\rangle &= \sqrt{n_{\vec{k}} + 1}|n_{\vec{k}} + 1\rangle \\ a(\vec{k})|n_{\vec{k}}\rangle &= \sqrt{n_{\vec{k}}}|n_{\vec{k}} - 1\rangle \\ b^\dagger(\vec{k})|\bar{n}_{\vec{k}}\rangle &= \sqrt{\bar{n}_{\vec{k}} + 1}|\bar{n}_{\vec{k}} + 1\rangle \\ b(\vec{k})|\bar{n}_{\vec{k}}\rangle &= \sqrt{\bar{n}_{\vec{k}}}\bar{n}_{\vec{k}} - 1\rangle. \end{aligned} \quad (5.65)$$

This result also is due to second quantization, which has changed non-QFT constant coefficients into operators with commutation operators.⁹⁴ This is clear from

$$\hat{n}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger\hat{a}\hat{a}^\dagger|n\rangle = (\hat{a}\hat{a}^\dagger + 1)\hat{a}^\dagger|n\rangle = (n + 1)\hat{a}^\dagger|n\rangle,$$

where the third step is due to the commutation relation. And since the coefficients are creation and annihilation operators, equations (5.58) and (5.59) show that the fields ϕ and ϕ^\dagger are also operators. This is important, yea, astounding:

Quantum fields are **operator fields**, operating on states such as the vacuum state.

Note that use of these operators on a state gives a new state. So the operators do not have eigenvalues and do not represent observables. Furthermore, since the a's and b's are operators, then $\phi(x)$ and $\phi^\dagger(x)$, as defined in (5.58) and (5.59) must also be operators. Since these equations are mode expansions of SHO solutions, it should come as no surprise that we come upon number operators and ladder operators.

Consider again the solutions of equations (5.58) and (5.59). We will confirm in (5.70) that the a's are particle operators, whereas the b's are antiparticle operators. From equations (5.65), we see which are creation and which annihilation⁹⁵ operators. So there are four cases.

⁹³ Klauber, 55-56.

⁹⁴ Even if we have not always put the hat on the a and b operators.

⁹⁵ Or should we say "terminator" to be more up to date? ;)

Field operator from equation (5.58) and (5.59)	Term ⁹⁶	contains	action
ϕ	ϕ^+	ae^{-ikx}	destroys particles
	ϕ^-	$b^\dagger e^{ikx}$	creates antiparticles
ϕ^\dagger	$\phi^{\dagger+}$	be^{-ikx}	destroys antiparticles ⁹⁷
	$\phi^{\dagger-}$	$a^\dagger e^{ikx}$	creates particles

Table 3. Particle/antiparticle creation/annihilation operators

So we have that

- ϕ is the **total lowering operator**, since it destroys particles and creates antiparticles; and
- ϕ^\dagger is the **total raising operator**, since it destroys antiparticles and creates particles.

Putting in a minus sign for antiparticles, the **total particle-number operator** is then

$$N(\phi) = \sum_{\vec{k}} (N_a(\vec{k}) - N_b(\vec{k})). \quad (5.66)$$

Now we can see why we need QFT. It can handle particle creation and annihilation and, so, particles' changing into one another. But it can also – importantly – handle multi-particle states.

5.10.7. Normal ordering

If we use the Hamiltonian of equation (5.63), the terms $\frac{1}{2}$ when integrated over space will lead to an infinite energy, which most physicists find quite embarrassing. A way to avoid this problem is to use **normal ordering**, which means rearranging a string of operators so that creation operators are on the left and annihilation operators on the right, otherwise in the original order. This procedure supposedly corrects an ambiguity in the order of terms in the classical Lagrangian and effectively removes the infinities, which have their origin in the non-commutation of certain a 's and b 's.⁹⁸ But since it pretends momentarily that the creation and annihilation operators temporarily commute – which they do not – the procedure is complex. In fact, the non-commutation of the operators is fundamental elsewhere in QFT.

While normal ordering is questionable as a means of eliminating infinite vacuum energies, it can be legitimate as well as useful if one takes account accurately of the commutation relations for the operators exchanged. This is the subject of the use of Wick's theorem of contractions (section 7.3).

5.10.8. Résumé of two quantizations

After the nitty-gritty, let's look at an overview of what we have done and found out about quantization.

- Start with a relativistic Lagrangian density for particles

96 These are Klauber's notation. Everybody else uses ϕ^+ to mean creation and ϕ^- annihilation, as will we in future paragraphs. This will agree with a^\dagger and b^\dagger as creation and a and b as annihilation operators.

97 I admit to not understanding this, the second part. If negative exponent means positive energy, these are particles, not antiparticles. However, b seems to prime, saying they are antiparticles...

98 Klauber, 60-61; Lancaster and Blundell, 104-105.

$$\mathcal{L} = \frac{1}{2}[\partial_\mu \phi(x)]^2 - \frac{1}{2}m^2[\phi(x)^2] \quad (5.52)$$

and make the fields, ϕ , complex:

$$\mathcal{L} = \frac{1}{2}[(\partial_\alpha \phi^\dagger \partial^\alpha \phi - m^2 \phi^\dagger \phi)] = \frac{1}{2}[(\dot{\phi}^\dagger \dot{\phi} - \nabla \phi^\dagger \cdot \nabla \phi - m^2 \phi^\dagger \phi)]. \quad (5.54)$$

Derive the canonical momentum of the field

$$\pi^\mu(x) = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi(x))} = \partial^\mu \phi(x) \quad (5.55)$$

and use it to write the Hamiltonian density

$$\begin{aligned} \mathcal{H} &= \frac{\partial \mathcal{L}}{\partial \dot{\phi}^r} \dot{\phi}^r - \mathcal{L} = \frac{\partial L}{\partial \dot{\phi}} \dot{\phi} + \frac{\partial L}{\partial \dot{\phi}^\dagger} \dot{\phi}^\dagger - \mathcal{L} = \\ &= \frac{1}{2}(\dot{\phi}^\dagger \dot{\phi} + \nabla \phi^\dagger \cdot \nabla \phi + \mu^2 \phi^\dagger \phi) = \frac{1}{2}(\partial_\mu \phi^\dagger \partial^\mu \phi + \mu^2 \phi^\dagger \phi). \end{aligned} \quad (5.56)$$

Use this or the Euler-Lagrange equations to write the Klein-Gordon equation (wave equation, equivalent to the Schrödinger equation)

$$(\partial_\mu \partial^\mu + \mu^2)\phi = 0 \quad (5.57a)$$

$$(\partial_\mu \partial^\mu + \mu^2)\phi^\dagger = 0 \quad (5.57b).$$

Then solve these for the plane-wave solutions

$$\phi(x) = \sum_k \frac{1}{\sqrt{2V\omega_k}} a(\vec{k}) e^{-ikx} + \sum_k \frac{1}{\sqrt{2V\omega_k}} b^\dagger(\vec{k}) e^{ikx} := \phi^+ + \phi^- \quad (5.58a)$$

$$\phi^\dagger(x) = \sum_k \frac{1}{\sqrt{2V\omega_k}} b(\vec{k}) e^{-ikx} + \sum_k \frac{1}{\sqrt{2V\omega_k}} a^\dagger(\vec{k}) e^{ikx} := \phi^{\dagger+} + \phi^{\dagger-} \quad (5.58b)$$

or the continuous solutions of equations (5.59). We have now done **first quantization**, converting a particle equation to a wave equation and so passing from classical to quantum mechanics. We now deal with wave fields.

II. Quantize the *fields* by imposing the commutation relations

$$\{\phi^i(x, t), \pi_j(y, t)\} = \delta_j^i \delta(x - y) \rightarrow$$

$$\{[\phi^r(x, t), \pi_s(y, t)] = i\hbar \delta_s^r \delta(x - y); [\phi^r, \phi^s] = [\pi_r, \pi_s] = 0 \quad (\text{Error: Reference source not found})$$

on the fields ϕ of the plane-wave solutions (5.58). This requirement leads to commutation relations on the *amplitudes*:

$$\begin{aligned} [a(\vec{k}), a^\dagger(\vec{k}')] &= [b(\vec{k}), b^\dagger(\vec{k}')] = \delta_{\vec{k}\vec{k}'} \quad (\text{discrete}); \\ &= \delta(\vec{k} - \vec{k}') \quad (\text{continuous}). \end{aligned} \quad (5.62)$$

The commutation relations constitute **second quantization**, showing how waves can give rise to particles.

III. Use these results to evaluate the Hamiltonian from the Hamiltonian density and find the energy states:

$$H = \sum_k \omega_k \left(N_a(\vec{k}) + \frac{1}{2} + N_b(\vec{k}) + \frac{1}{2} \right), \quad (5.63)$$

which look like those of a SHO in terms of a number operator

$$N_a(\vec{k}) = a^\dagger(\vec{k})a(\vec{k}); \quad N_b(\vec{k}) = b^\dagger(\vec{k})b(\vec{k}). \quad (5.64)$$

IV. Show that the amplitude operators can be used as creation and annihilation operators of particles.

V. Use normal ordering. More on this in section 7.3.

The number operators represent observables and the a 's will be used in propagators and interaction theory.

5.10.9. Klein-Gordon particles are bosons

Looking at equations (5.65) for how creation operators raise the number of particles in a state, we see that, just as in the case of raising operators for the NRQM SHO, we can write any state (ignoring b for the moment) as

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle.$$

This equation says that multiple particles occupy the same state $|n\rangle$ and so, according to the Fermi exclusion principle, Klein-Gordon particles must be **bosons**.

5.10.10. Continuity, currents and negative energy

Using the standard QM method of multiplying the Schrödinger equation by the complex conjugate of the wave function, subtracting the complex conjugate of that and doing some differentiating by parts, we can derive the probability current for the K-G equation in QFT.⁹⁹

$$\rho = j^0 = i \left(\frac{\partial \phi}{\partial t} \phi^\dagger - \frac{\partial \phi^\dagger}{\partial t} \phi \right) \quad (5.67)$$

and

$$j^\mu = i (\phi^{\cdot\mu} \phi^\dagger - \phi^{\dagger,\mu} \phi) \quad (5.68)$$

so that the continuity equation is then

$$j^\mu_{,\mu} = \partial_\mu j^\mu = 0. \quad (5.69)$$

Now things get interesting. Insert the K-G solutions (5.58) into the probability density (5.67), then derive the **effective probability density**

$$\bar{\rho} = \langle \phi_1 \phi_2 \dots | \rho | \phi_1 \phi_2 \dots \rangle,$$

supposing that the bras and kets are eigenstates of \vec{k} , i.e., energy-momentum. The somewhat surprising and interesting result is¹⁰⁰

$$\bar{\rho} = \sum_{\vec{k}} \frac{1}{V} \{ a^\dagger(\vec{k})a(\vec{k}) - b^\dagger(\vec{k})b(\vec{k}) \} = \sum_{\vec{k}} \frac{1}{V} \{ N_a(\vec{k}) - N_b(\vec{k}) \}. \quad (5.70)$$

The minus sign on the second term comes partly from that on the second term of equation (5.67) and partly from derivatives of the negative-energy b states. The term looks like negative probability as well as negative energy, but if we multiply this by the appropriate charge q , then the second term makes perfect sense as charge density

⁹⁹ Klauber, 45-46.

¹⁰⁰ Klauber, 69.

if the b particles are interpreted as **antiparticles**. This is therefore what is done.

We will say more about negative energies in section 5.11.2.

5.10.11. Comparison to a SHO

Similarities of K-G fields with a SHO are obvious and interesting, but no proof of QFT. Why? Well...¹⁰¹

- Both SHOs and QFT have raising and lowering operators with similar behaviors on states.

However,

- SHOs are examples of bound states with a non-zero potential energy and hence subject to a force, whereas the K-G equation applies to free states with no potential or force.
- A SHO is a single particle with different energy states, whereas QFT can handle states of many particles in each energy state.
- A SHO represents a real particle oscillating in place with a non-sinusoidal wave function (Hermite polynomial), whereas QFT considers complex, sinusoidal waves which move through space.

So although SHOs may help interpreting or justifying QFT, they do not offer a rigorous derivation of it.

5.10.12. The Feynman propagator

Now we can see how the coefficient operators can be used to construct measurable quantities involving the creation and destruction of virtual particles. The **Feynman propagator** (or just propagator) is a mathematical representation of a virtual particle or antiparticle (such as a force-carrying boson) created at one point in space and time in the vacuum and destroyed at another point. We will use them to propagate a particle from vertex of a Feynman diagram to another. So far, we have only studied the Klein-Gordon equation for bosons, so for the moment we will stick to considering the propagator for a virtual boson.

Suppose a virtual particle is created at y and destroyed at a later time at x . This is physically indistinguishable from a virtual antiparticle of opposite charge going from x to y . In terms of the K-G solutions (5.59), the creation is represented by $\phi^\dagger(y)$ and is followed by an annihilation given by $\phi(x)$, the two forming a **time-ordered operator**

$$T\{\phi(x)\phi^\dagger(y)\} = \phi(x)\phi^\dagger(y), \quad \text{for } t_y < t_x. \quad (5.71)$$

The time-ordering operator, T , not really an operator, just says to put the earliest term on the right. Now we could instead suppose that $\phi(x)$ creates an antiparticle which is later annihilated by $\phi^\dagger(y)$. This **antiparticle** creation-annihilation would be represented by

$$T\{\phi(x)\phi^\dagger(y)\} = \phi^\dagger(y)\phi(x), \quad \text{for } t_x < t_y.$$

In the vacuum state, the transition amplitude for the particle process is given by its vacuum expectation value (VEV):

$$\langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle = \langle 0|\phi(x)\phi^\dagger(y)|0\rangle, \quad (5.72)$$

This reduces to a product of scalars representing the creation and destruction probability amplitudes, respectively, of the virtual particle. We define the **Feynman propagator** by¹⁰²

$$i\Delta_F(x - y) = \langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle. \quad (5.73)$$

¹⁰¹ These criteria are after Klauber, 69-70.

¹⁰² Klauber, 73. Not everybody puts in the factor i on the left-hand side or uses the Δ symbol.

Using the solution expansions (5.59), this can be expressed in terms of commutators of the separate amplitudes of the K-G solutions

$$i\Delta_F(x-y) = \langle 0 | [\phi^+(x), \phi^{\dagger-}(y)] | 0 \rangle$$

for a virtual particle, and

$$i\Delta_F(x-y) = \langle 0 | [\phi^{\dagger+}(y), \phi^-(x)] | 0 \rangle$$

for an antiparticle. Now plug in the K-G solutions (5.59) and use the results of the commutators to get an integral, which can then be converted to a contour integral and then back to the final integral

$$i\Delta_F(x-y) = \frac{i}{(2\pi)^4} \int_{-\infty}^{\infty} \frac{e^{-ik(x-y)}}{k^2 - \mu^2 + i\epsilon} d^4k \quad (5.74)$$

where $\epsilon \rightarrow 0$ after doing the integral.¹⁰³ This is the **Feynman propagator** for spin 0 bosons. It is particularly “simple” in the momentum representation:

$$\Delta_F(k) = \frac{1}{k^2 - \mu^2 + i\epsilon}. \quad (5.75)$$

So from the K-G equation and second quantization, we can calculate a probability amplitude for the creation of virtual particles, a step towards writing the equation for an interaction. Remember, we assume free particles and so no interactions.

The good news is that, at least according to Klauber, all this makes up “... most of the basic principles in QFT.”

5.11. Spin-1/2 particles – the Dirac equation

We saw how to find the Dirac Lagrangian from symmetry considerations, complete with γ matrices, in (4.2).

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi, \quad (5.76)$$

where the adjoint $\bar{\psi} = \psi^\dagger \gamma^0$. We can then use it to derive the Dirac equation.

Alternatively, and perhaps more satisfyingly, the Dirac equation can be derived by starting with the QM dispersion relation

$$E^2 = m^2 + p^2 \rightarrow \partial^2 - m^2.$$

We would like to split this up into two parts using something like $\sqrt{\partial^2 \pm im}$, but the square root of ∂^2 does not mean anything. The historic discovery of the appropriate equation lends some physical understanding. Dirac wanted such a linear equation also to be a solution to the Klein-Gordon equation, and so he considered

$$v^\mu \psi_\mu + m\psi_\mu,$$

but then v^μ would define a vector with a specific direction in space which would not be Lorentz invariant. So the coefficients could not constitute a simple spacetime constant, but must be operators (matrices) of some sort. So he tried

$$(i\gamma^\mu \partial_\mu - m)\psi = 0.$$

In order to compare this with a dispersion relation, one can multiply it by the same quantity with a + sign, the result being

$$-(\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + m^2)\psi = 0,$$

¹⁰³ Details of this laborious calculation are in Klauber, 70-77. Yes, over 8 pages.

which can be rewritten as

$$\frac{1}{2}((\gamma^\mu \gamma^\nu \partial_\mu \partial_\nu + \gamma^\mu \gamma^\nu \partial_\nu \partial_\mu) + m^2)\psi = 0.$$

Exchanging μ and ν in the second term leads to the anticommutator, so

$$(\frac{1}{2}\{\gamma^\mu, \gamma^\nu\} \partial_\mu \partial_\nu + m^2)\psi = 0.$$

If the γ satisfy the condition for a **Clifford algebra** (3.30)¹⁰⁴,

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathcal{I}, \quad (5.77)$$

with η being the Minkowski metric, the result is the desired Klein-Gordon equation,

$$(\delta^2 + m^2)\psi = 0.$$

The Clifford algebra condition (5.77) tells us that the γ^μ must be 4x4 matrices like the metric. The Clifford algebra requirement is satisfied if the γ matrices obey

$$(\gamma^0)^2 = 1, \quad (\gamma^1)^2 = (\gamma^2)^2 = (\gamma^3)^2 = -1$$

with the anticommutation relations

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu := \{\gamma^\mu, \gamma^\nu\} = 0, \quad \mu \neq \nu..$$

All this leads to the **Dirac equation**, the **RQM wave equation for spin-1/2 particles**:

$$(i\gamma^\mu \delta_\mu - m)|\psi\rangle = 0. \quad (5.78)$$

One choice for the **Dirac matrices**, or **gamma matrices**, is the **chiral representation or basis** (or the **Weyl basis**):

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \quad (5.79)$$

where the σ^i are our old friends the **Pauli spin matrices**

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Since the Pauli matrices are 2x2 in dimension, the gammas are 4x4. In all their glory,

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, & \gamma^1 &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \\ \gamma^2 &= \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, & \gamma^3 &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \end{aligned} \quad (5.80)$$

The notation becomes much simpler if we define

$$\sigma^\mu = (I, \vec{\sigma}) \quad \text{and} \quad \bar{\sigma}^\mu = (I, -\vec{\sigma}). \quad (5.81)$$

¹⁰⁴ N.B. The sign of the right-hand side depends on the metric signature. We use (+---). Robinson, for instance, uses the opposite. See Table 2.

(Note the almost-invisible bar over the second quantity, $\bar{\sigma}^\mu$.) Then we can write the chiral basis as

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}. \quad (5.82)$$

Alternatively, one can use the **mass** or **standard** or **Dirac representation or basis**¹⁰⁵:

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix},$$

which is the same as the chiral basis except for γ^0 .

The important point is:

*Physical results are independent of the choice of basis for the gamma vectors.*¹⁰⁶

5.11.1. Spin, helicity and chirality

It is informative to see how the system works in the chiral basis. We represent the operators as 2x2 matrices, knowing this to be a shorthand for 4x4 matrices, so they must operate on four-component vectors. In fact, the vectors are Dirac spinors, four-component spinors each composed of a pair of two-component Weyl spinors.

Let's spell this out completely once:

$$\psi(x) = \begin{pmatrix} \psi_L(x) \\ \psi_R(x) \end{pmatrix} = \begin{pmatrix} \psi_{L1}(x) \\ \psi_{L2}(x) \\ \psi_{R1}(x) \\ \psi_{R2}(x) \end{pmatrix} = \begin{pmatrix} u_L \\ u_R \end{pmatrix} e^{-ip \cdot x} = \begin{pmatrix} u_{L1} \\ u_{L2} \\ u_{R1} \\ u_{R2} \end{pmatrix} e^{-ip \cdot x},$$

the last two terms being for spinors. The subscripts will be explained below.

Using (5.82) and with $p \cdot x = \hat{p}^0 t - \vec{p} \cdot \vec{x}$, the Dirac equation (5.78) gives

$$\left[\begin{pmatrix} 0 & \hat{p}^0 \\ \hat{p}^0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & \sigma \cdot \hat{p} \\ -\sigma \cdot \hat{p} & 0 \end{pmatrix} - \begin{pmatrix} m & 0 \\ 0 & m \end{pmatrix} \right] u(p) e^{-ip \cdot x} = 0,$$

or

$$\begin{pmatrix} -m & \hat{p}^0 - \sigma \cdot \hat{p} \\ \hat{p}^0 + \bar{\sigma} \cdot \hat{p} & -m \end{pmatrix} u(p) = 0, \quad (5.83)$$

where we consider only the spinor part of the solution. Remember that σ is a 2x2 matrix and m is supposed multiplied by a unit matrix. Let's make the usual supposition that the momentum is along the z-axis. Then what this equation represents in all is:

$$\begin{pmatrix} -m & 0 & p^0 - p^3 & 0 \\ 0 & -m & 0 & p^0 + p^3 \\ p^0 + p^3 & 0 & -m & 0 \\ 0 & p^0 - p^3 & 0 & -m \end{pmatrix} u(p). \quad (5.84)$$

For the special case of a *massless* particle, (5.83) reduces to

¹⁰⁵ Griffiths calls it the 'Bjorken and Drell' convention. This is the one also used by Klauber. Schwichtenberg uses both this and the chiral representation, which is also used by Lancaster and Blundell. Actually these two are handy on different occasions.

¹⁰⁶ Robinson, 183.

$$(\hat{p}^0 - \sigma \cdot \hat{p})\psi_R = 0 \quad \text{and} \quad (\hat{p}^0 + \sigma \cdot \hat{p})\psi_L = 0. \quad (5.85)$$

so the equations of the two spinors, ψ_L and ψ_R , are separate and each evolves independently of the other. In particular, one will not change into the other.

Although, we have not yet stated what R and L refer to, each represents a two-component spinor; ψ_L occupies the top two components of the four-component spinor field, ψ_R , the two lower. Starting with the Lorentz algebra of section 3.6.1 and (3.30) and using the chiral representation and the Clifford algebra, one can derive the equations for Lorentz rotations and boosts of the vector $\begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}$.¹⁰⁷ The form of the boost for ψ_L indicates clearly that it is a left-chiral spinor field; similarly, ψ_R transforms like a right-chiral spinor. So in the chiral representation for massless particles, the two chiralities are well separated.

Now if we consider massive particles, (5.83) becomes

$$(\hat{p}^0 - \sigma \cdot \hat{p})\psi_R = m\psi_L, \quad \text{and} \quad (\hat{p}^0 + \sigma \cdot \hat{p})\psi_L = m\psi_R. \quad (5.86)$$

Now the two states are coupled (by the mass) and may oscillate one into the other. If ψ_L and ψ_R are eigenstates of momentum and energy, we may eliminate one of the spinors, say ψ_L to show

$$(\hat{p}^0 + \sigma \cdot \hat{p})(\hat{p}^0 - \sigma \cdot \hat{p})\psi_R = m^2\psi_R, \quad (5.87)$$

which gives the desired dispersion relation

$$(p^0)^2 - (\vec{p})^2 = m^2.$$

Since we may take either of the square roots of this, we may get positive or *negative* energy

$$E = \pm \sqrt{\vec{p}^2 + m^2}.$$

In the extreme relativistic limit of very high energies, we can ignore the mass. So either in this case or that of a massless particle, equation (5.83) reduces via (5.82) to the **Weyl equations**

$$(p \cdot \sigma)u_R = 0 \quad \text{and} \quad (p \cdot \bar{\sigma})u_L = 0.$$

These two equations then lead to

$$\hat{\vec{p}} \cdot \vec{\sigma} u_R = u_R \quad \text{and} \quad \hat{\vec{p}} \cdot \vec{\sigma} u_L = -u_L,$$

where $\hat{\vec{p}} = \frac{\vec{p}}{E}$ is a unit vector in the direction of the particle's motion, so $\hat{\vec{p}} \cdot \vec{\sigma}$ is a way of writing the **helicity operator**.¹⁰⁸ These say that in the extreme relativistic limit the particle's spin is aligned along its direction of motion. By definition, if the eigenvalue of the helicity operator $\hat{\vec{p}} \cdot \vec{\sigma}$ is +1, the spinor is considered to be right-handed (right-chiral); for eigenvalue -1, left-handed.

By convention, the antiparticle solutions of the Dirac equation have negative values of energy. In this case, the equations (5.85) for a *massless* antiparticle become

$$(-|\hat{p}^0| - \sigma \cdot \hat{p})\psi_R = 0 \quad \text{and} \quad (-|\hat{p}^0| + \sigma \cdot \hat{p})\psi_L = 0,$$

so that

$$\frac{\sigma \cdot \hat{p}}{|p|}\psi_R = -\psi_R \quad \text{and} \quad \frac{\sigma \cdot \hat{p}}{|p|}\psi_L = \psi_L, \quad \text{for antiparticles.}$$

The signs of the helicity are inverted from those for positive-energy states,

¹⁰⁷ Robinson, 175-177.

¹⁰⁸ Robinson, 185-6.

Similarly, the **chirality operator** is defined by

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \quad (5.88)$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{only in the chiral basis.} \quad (5.89)$$

The matrix γ^5 can be shown to be invariant under Lorentz transformations. Its eigenstates are

$$\begin{pmatrix} \psi_L \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 \\ \psi_R \end{pmatrix}$$

with eigenvalues -1 and +1, respectively. The **projection operators**

$$P_{\pm} = \frac{1}{2}(1 \pm \gamma^5)$$

project out the right and left-chiral spinors:

$$P_+\psi = \psi_R \quad \text{and} \quad P_-\psi = \psi_L.$$

5.11.2. Interlude – negative energies

Let's be as clear as we can about how we handle seemingly-negative energies.

Suppose the Klein-Gordon equation has solutions like

$$\phi(x) = Ne^{-ip \cdot x} = Ne^{-i(Et - \vec{p} \cdot \vec{x})}.$$

then our convention is to take this to represent an incoming wave. Applying the momentum operator $\hat{p} = -i\partial_x$ returns the momentum $p_x > 0$ (and $E > 0$, by $i\partial_t$). So our convention supposes the wave comes in from the left, moving in the direction of increasing x . Note that it depends on the metric signature chosen (+--- in our case). A positive exponential term would give $p_x < 0$, representing a particle moving in the other direction, i.e., either a particle outgoing particle moving to the right or an incoming antiparticle, as we will now see.

A charged particle in an EM field obeys the classical equation¹⁰⁹

$$m \frac{d^2 x}{d\tau^2} = q F_{\nu}^{\mu} \frac{dx^{\mu}}{d\tau}.$$

If we invert $\tau \rightarrow -\tau$ the equation remains unchanged if we also invert the sign q of the particle. So a charged particle traveling backwards in time is equivalent to the oppositely charged particle moving forwards in time.

The probability-current density for a K-G particle is

$$J^{\mu} = (+q)2|N|^2 p^{\mu} = (+q)2|N|^2 (E, \vec{p})$$

for a particle with positive energy, where N is a normalization constant. If its energy is negative

$$J^{\mu} = (+q)2|N|^2 (-|E|, \vec{p}) = (-q)2|N|^2 (|E|, -\vec{p}).$$

So we can handle a negatively-charged particle *by making it an antiparticle (i.e., reversing its charge), considering its energy to be positive and reversing the sign of its three-momentum.*

$$\begin{pmatrix} -|E| \\ \vec{p} \\ q \end{pmatrix} \rightarrow \begin{pmatrix} |E| \\ -\vec{p} \\ -q \end{pmatrix}. \quad (5.90)$$

¹⁰⁹ This paragraph based on Blundell and Lancaster, 62-63. Theirs is the only book I know which explains this so clearly.

Negative-energy particles are interpreted as positive-energy antiparticles moving in the opposite direction to that of the corresponding particle. The K-G equation with a specific energy then has two solutions¹¹⁰:

$$\phi(x) = \begin{bmatrix} \text{incoming positive-} \\ \text{energy particle} \\ \propto e^{-i(Et-\vec{p}\cdot\vec{x})} \end{bmatrix} + \begin{bmatrix} \text{outgoing negative-} \\ \text{energy particle} \\ \propto e^{+i(Et-\vec{p}\cdot\vec{x})} \end{bmatrix}.$$

Note that this discussion concerns fields considered to have energy -- positive or negative. In QFT, these “fields” will become operators, which don’t have energy. A state is the result of acting on the vacuum with a creation or annihilation operator and this will create either a state with positive energy or annihilate the state to nothing.

5.11.3. Solutions to the Dirac equation

Let’s go ahead and look at the solutions. For now, we will use the chiral, or Weyl, basis. Knowing that solutions of the Dirac equation must also be solutions of the Klein-Gordon equation, we start with the general form of these from (5.58) in terms of plane waves:

$$\psi(x) \simeq u(p)e^{-ipx} + v(p)e^{ipx}, \quad (5.91)$$

where the first term represents solutions for positive-frequency (energy) particles and the second, for negative-frequency, since (reminder: metric signature +---)

$$\hat{E}\psi(x) = i\partial_t\psi(x) = Eu(p)e^{-ipx} + (-E)v(p)e^{ipx}.$$

Then (5.83) becomes two equations, one for positive frequency and one for negative:¹¹¹

$$\begin{pmatrix} -m & p \cdot \sigma \\ p \cdot \bar{\sigma} & -m \end{pmatrix} u(p) = 0 \quad \text{and} \quad \begin{pmatrix} -m & -p \cdot \sigma \\ -p \cdot \bar{\sigma} & -m \end{pmatrix} v(p) = 0. \quad (5.92)$$

In the rest frame, where $p=(m,0,0,0)$, the solutions are

$$u = \begin{pmatrix} \xi \\ \xi \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} \eta \\ -\eta \end{pmatrix}$$

where ξ and η are any two-component spinors, e.g.

$$u_{\uparrow} = \sqrt{m} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u_{\downarrow} = \sqrt{m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_{\uparrow} = \sqrt{m} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad v_{\downarrow} = \sqrt{m} \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}. \quad (5.93)$$

Putting the mass back, let’s make the usual simplification where we take the momentum to be along the z-axis and consider only the spinor part of the solution. Remember that σ is a 2x2 matrix and m is supposed multiplied by a unit matrix. Then the positive-frequency equation (5.92) leads to

$$\begin{pmatrix} -m & 0 & p^0 - p^3 & 0 \\ 0 & -m & 0 & p^0 + p^3 \\ p^0 + p^3 & 0 & -m & 0 \\ 0 & p^0 - p^3 & 0 & -m \end{pmatrix} u(p) = 0. \quad (5.94)$$

The solutions in the p^3 frame can be found to be¹¹²

110 Copied from Blundell and Lancaster, 63.

111 Method of Schwartz, 188-189.

112 Schwartz, 188-189.

$$u(p) = \begin{pmatrix} \begin{pmatrix} \sqrt{E-p^3} & 0 \\ 0 & \sqrt{E+p^3} \end{pmatrix} \xi \\ \begin{pmatrix} \sqrt{E+p^3} & 0 \\ 0 & \sqrt{E-p^3} \end{pmatrix} \xi \end{pmatrix}. \quad (5.95)$$

In a more condensed, almost cryptic, form

$$u(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \xi \\ \sqrt{p \cdot \bar{\sigma}} \xi \end{pmatrix} \quad \text{and} \quad v(p) = \begin{pmatrix} \sqrt{p \cdot \sigma} \eta \\ -\sqrt{p \cdot \bar{\sigma}} \eta \end{pmatrix}. \quad (5.96)$$

Don't miss the bar over the σ in the lower part of each one. In the case of $p=(E, 0, 0, p^z)$,

$$\begin{aligned} u^1 &= \begin{pmatrix} \sqrt{E-p^z} \\ 0 \\ \sqrt{E+p^z} \\ 0 \end{pmatrix}, & u^2 &= \begin{pmatrix} 0 \\ \sqrt{E+p^z} \\ 0 \\ \sqrt{E-p^z} \end{pmatrix}, \\ v^1 &= \begin{pmatrix} \sqrt{E-p^z} \\ 0 \\ -\sqrt{E+p^z} \\ 0 \end{pmatrix}, & v^2 &= \begin{pmatrix} 0 \\ \sqrt{E+p^z} \\ 0 \\ -\sqrt{E-p^z} \end{pmatrix}. \end{aligned} \quad (5.97)$$

These are the solutions for the spinor part only. The complete solution includes the spinor and spacetime parts, where the spinors of (5.97) are multiplied by the appropriate exponential function as in (5.91).

The inner product of two spinors can be calculated from (5.96):

$$u^\dagger(p)u(p) = 2E\xi^\dagger\xi.$$

But this is not Lorentz invariant, as E depends on the observer. So instead we define the **adjoint** function\

$$\bar{u}(p) = u^\dagger(p)\gamma^0. \quad (5.98)$$

One usually uses a basis $\xi^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\xi^2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, so that $\xi^{\dagger s}\xi^r = \delta^{sr}$. Then

$$\bar{u}^s(p)u^r(p) = 2m\xi^{s\dagger}\xi^r = 2m\delta^{sr}, \quad (5.99)$$

which is the standard normalization for the spinor inner product of Dirac spinors.¹¹³

Now an arbitrary particle solution of the Dirac equation can be represented by a mode expansion composed of an integral over momentum and a sum over spin states, including particles (a terms) and antiparticles (b terms).¹¹⁴ The subscripts on the coefficient operators indicate the corresponding momentum, $a_p^\dagger|0\rangle = |p\rangle$. Thus

$$\psi(x) = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}} \frac{1}{(2E_p)^{\frac{1}{2}}} \sum_{s=1}^2 (a_{sp}u_s(p)e^{-ip \cdot x} + b_{sp}^\dagger v_s(p)e^{ip \cdot x}) \quad (5.100a)$$

and

¹¹³ Schwartz, 191.

¹¹⁴ Seems to be mixture of Blundell and Lancaster, 102, 133 and Klauber, 103. [?]

$$\bar{\psi}(x) = \int \frac{d^3p}{(2\pi)^{\frac{3}{2}}} \frac{1}{(2E_p)^{\frac{1}{2}}} \sum_{s=1}^2 (b_{sp} \bar{v}_s(p) e^{-ip \cdot x} + a_{sp}^\dagger \bar{u}_s(p) e^{ip \cdot x}). \quad (5.100b)$$

The p subscript indicates momentum. The sum over s takes into account the two dimension of the spinors. When we do canonical quantization and consider these fields as operators, then each ψ and each coefficient should get a hat.

5.11.4. The Dirac equation in QFT

There is no classical Lagrangian or Hamiltonian density for spin-1/2 particles. Symmetry and compatibility (with the scalar case) considerations have led to the equation for the Dirac Lagrangian of (5.76):

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi.$$

That this Lagrangian density be valid can be shown by substituting it into Euler-Lagrange equations

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial \mathcal{L}}{\partial \phi_{,\mu}^n} \right) - \frac{\partial \mathcal{L}}{\partial \phi^n} = 0 \quad \text{with} \quad \phi^1 = \bar{\psi}, \phi^2 = \psi \quad (5.101)$$

The conjugate momenta to ψ and $\bar{\psi}$ are then

$$\pi = \frac{\partial \mathcal{L}}{\partial \psi_{,0}} = i\bar{\psi}\gamma^0 = i\psi^\dagger \gamma^0 \gamma^0 = i\psi^\dagger, \quad \bar{\pi} = \frac{\partial \mathcal{L}}{\partial \bar{\psi}_{,0}} = 0. \quad (5.102)$$

So the **Dirac Hamiltonian density** turns out to be¹¹⁵

$$\mathcal{H} = \pi\dot{\psi} + \bar{\pi}\dot{\bar{\psi}} - \mathcal{L} = -i\bar{\psi}\gamma^i \partial_i \psi + m\bar{\psi}\psi. \quad (5.103)$$

Since there are no macroscopic spinor fields, there are no Poisson brackets for them. So the next step of quantization must take place by analogy. It turns out that commutation relations like those of (5.76) don't work for spin-1/2 particles, but the corresponding *anti-commutation* relations do. So with $\{a, b\} = ab + ba$, the anti-commutation relations for spin-1/2 fields are

$$\begin{aligned} \{\hat{\psi}_a(t, \vec{x}), \hat{\psi}_b^\dagger(t, \vec{y})\} &= \delta^3(\vec{x} - \vec{y}) \delta_{ab}, \\ \{\hat{\psi}_a(t, \vec{x}), \hat{\psi}_b(t, \vec{y})\} &= \{\hat{\psi}_a^\dagger(t, \vec{x}), \hat{\psi}_b^\dagger(t, \vec{y})\} = 0. \end{aligned} \quad (5.104)$$

The same relations hold with the adjoint solution $\bar{\psi} = \psi^\dagger \gamma^0$. The hats on the fields indicate that they are considered to be operators. Application of these relations to the spinors of (5.100) shows that the coefficients must also obey the commutation relationship

$$\{\hat{a}_s(p), \hat{a}_r^\dagger(q)\} = \{\hat{b}_s(p), \hat{b}_r^\dagger(q)\} = \delta^3(p - q) \delta_{sr}$$

and so are creation and annihilation operators, as we found for scalar (Klein-Gordon) particles.

The Hamiltonian density can be integrated over all of space and three pages of math¹¹⁶ to give the free Dirac Hamiltonian

$$H = \int d^3p \sum_{s=1}^2 E_p \left(\hat{a}_{sp}^\dagger \hat{a}_{sp} + \hat{b}_{sp}^\dagger \hat{b}_{sp} \right), \quad (5.105)$$

which is the sum of the energies (considered positive) of the particles and antiparticles.¹¹⁷

¹¹⁵ Klauber, 104.

¹¹⁶ Klauber, 105-108.

¹¹⁷ Ignoring any annoying infinities...

As for the Klein-Gordon coefficients (Table 3), \hat{a}^\dagger and \hat{b}^\dagger create a single particle and antiparticle, respectively; \hat{a} and \hat{b} destroy a particle or antiparticle, respectively. Then the ψ of (5.100) destroys particles and creates antiparticles, so it is the **total particle-annihilation operator**, whereas $\bar{\psi}$ creates particles and destroys antiparticles and so is the **total particle-creation operator**.

However, (5.105) shows that particles of both types a and b have positive energies (aside from the vacuum energy), unlike the case of RQM where d particles have negative energy. But we are left with the same infinite-energy problem for the vacuum expectation value (VEV) of the energy we had for the K-G equation.

That is why no more than one spin- $1/2$ particle can occupy the same state. It is due to the anti-commutation rules for such states. Adding a particle to a state which already has one would be¹¹⁸

$$a_r^\dagger(p)|\psi_r p\rangle = a_r^\dagger a_r^\dagger|0\rangle.$$

But since

$$\{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0, \quad (5.106)$$

trying to create a second particle in the same state as the first destroys the state. This is the **Pauli exclusion principle**. It says that fermions make up the stuff around us. And it comes from the anticommutation relations for spinors.

5.11.5. Feynman propagator

Similarly to what was done with bosons, we can (rather laboriously) calculate the Feynman propagator for spin- $1/2$ particles:

$$S_F(x - y) = \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty} \frac{e^{-ip(x-y)} (\not{p} + m)}{p^2 - m^2 + i\epsilon} d^4 p, \quad (5.107)$$

which in the 4-momentum form is

$$S_F(p) = \frac{(\not{p} + m)}{p^2 - m^2 + i\epsilon} = (\not{p} + m)\Delta_F(p). \quad (5.108)$$

5.12. Spin-1 particles – photons

Using the same methods as for spin-0 and spin- $1/2$, we can find the relevant equations for photons. We start directly with relativistic EM because there is no non-relativistic version, EM being already Lorentz-invariant and so the precursor of and inspiration for SR.

From (5.37), the charge-free Lagrangian for the EM field is

$$\mathcal{L}_0^1 = (\partial^\mu A^\nu - \partial^\nu A^\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu). \quad (5.109)$$

The simplest possible QFT Lagrangian density, as proposed by Fermi¹¹⁹, is patterned after the classical, relativistic Lagrangian.

$$\mathcal{L}_0^1 = -\frac{1}{2}(\partial_\nu A_\mu(x)(\partial^\nu A^\mu(x)). \quad (5.110)$$

Putting the Euler-Lagrange equations to work on this leads to

¹¹⁸ Look out, Klauber, 110, uses c and d for fermion operators.

¹¹⁹ “There is actually more than one possible Lagrangian density which leads to the correct classical theory of electromagnetism. The simplest of these, first proposed by Fermi, ...” is the one in (5.110). Klauber, 144.

$$\partial_\alpha \partial^\alpha A^\mu(x) = 0, \quad (5.111)$$

which is the covariant form of Maxwell's equations in the Lorenz gauge. If A^μ is real (it is), it has plane-wave discrete solutions¹²⁰

$$\begin{aligned} A^\mu(x) &= \sum_{r, \vec{k}} \frac{1}{\sqrt{2V\omega_{\vec{k}}}} \mathcal{E}_r^\mu(\vec{k}) a_r(k) e^{-ikx} + \\ &\quad \sum_{r, \vec{k}} \frac{1}{\sqrt{2V\omega_{\vec{k}}}} \mathcal{E}_r^\mu(\vec{k}) a_r^\dagger(k) e^{ikx} \\ &= A^{\mu+} + A^{\mu-} \end{aligned} \quad (5.112)$$

and continuous ones¹²¹

$$\begin{aligned} A^\mu(x) &= \sum_r \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_{\vec{k}}}} \mathcal{E}_r^\mu(\vec{k}) a_r(k) e^{-ikx} + \\ &\quad \sum_r \int \frac{d^3\vec{k}}{\sqrt{2(2\pi)^3\omega_{\vec{k}}}} \mathcal{E}_r^\mu(\vec{k}) a_r^\dagger(k) e^{ikx} \\ &= A^{\mu+} + A^{\mu-} \end{aligned} \quad (5.113)$$

There is no second set of solutions in $b_r(\vec{k})$, as for the K-G equation, because the A^μ are real, since they are the basis of the observables \vec{E} and \vec{B} , the electric and magnetic fields. This is basically because they are their own antiparticles.¹²² In these equations, both r and μ take on integral values from 0 to 3. The solution A^μ is a vector (a 4-vector), so photons are referred to as **vector bosons**.

The \mathcal{E}_r^μ are the components of **polarization vectors** \mathcal{E}_r . Being 4-vectors, each \mathcal{E}_r has 4 components \mathcal{E}_r^μ , labeled by the superscript μ . They are generally taken to be orthogonal. Let's adopt a coordinate system, somewhat simply called the **photon aligned coordinate system**¹²³, in which the 3-vector $\vec{\mathcal{E}}$ components lie along the x, y and z axes, with the 3rd component along the z axis which we also take to be the direction of \vec{k} . Since A^μ is real, we consider only the $r=1$ ($\mathcal{E} = 0, 1, 0, 0$) polarization, so the result of equation (5.112) is¹²⁴

$$A^\mu(x) = \sqrt{\frac{2}{V\omega_{\vec{k}}}} \mathcal{E}_1^\mu(\vec{k}) a_1(k) \cos kx^3$$

where $x^3 = z$ so A^μ is a cosine wave in the x-z plane, an example of **linear polarization**. The magnetic field $\vec{B} = \nabla \times \vec{A}$ is then perpendicular to it in the y-z plane.¹²⁵

Care must be taken in order to avoid confusing spin and polarization. QFT for photons is formulated in terms of polarization; for fermions, of spin. Nevertheless, it is as if each type of particle has what might be called a "pseudo-angular momentum" or spin factor which multiplies its state vector components.. For scalars (K-G) this is of dimension 0, for spinors 1/2 and for photons 1, corresponding to a sort of "spin" in all three cases.

As pointed out in section 5.11.1, SR requires that the spin of a photon, because it is massless, be aligned along

¹²⁰ Klauber, 147.

¹²¹ Blundell and Lancaster, 122.

¹²² Klauber, 141-2, 148.

¹²³ Klauber points out it should be called the "photon-polarization vector-axes aligned" system.

¹²⁴ Klauber, 143.

¹²⁵ For a complete discussion, with diagrams, see Klauber, 142-143.

its direction of motion, either forwards or backwards, and so lies along $+\vec{k}$ or $-\vec{k}$. Polarization vectors, however, have four possible mutually orthogonal states in 4-d Minkowski space.

The equations (5.112) and (5.113) resemble those for the K-G equation, except that:

- photons have no mass ($\mu = 0$ in K-G equation);
- the electric and magnetic fields derived from the 4-vector potential A^μ are real and therefore A^μ is too;
- A^μ is a 4-vector.

In fact, in QM photons are their own anti-particles and this is the meaning of the reality of A^μ .

Second quantization leads unsurprisingly to commutation relations for the coefficients of the photon¹²⁶

$$[a_r(\vec{k}), a_s^\dagger(\vec{k}')] = \zeta_r \delta_{rs} \delta_{\vec{k}\vec{k}'} \quad \text{with} \quad \zeta = (-1, 1, 1, 1) \quad (\text{discrete}) \quad (5.114)$$

$$= \zeta_r \delta_{rs} \delta(\vec{k} - \vec{k}') \quad (\text{continuous}) \quad (5.115)$$

where underlined subscripts are not summed. The factor ζ_r comes from the fact that

$$-\zeta_r \delta_{rs} = \eta_{rs},$$

the Minkowski metric, which arises naturally from EM's already being Lorentz-invariant.

The analogy with the other two particle types does not stop here. Lo, behold, the Hamiltonian is

$$H = \sum_{\vec{k}, r} \omega_{\vec{k}} \left(N_r(\vec{k}) + \frac{1}{2} \right), \quad (5.116)$$

where the **number operator** is

$$N_r(\vec{k}) = \zeta_r a_r^\dagger(\vec{k}) a_r(\vec{k}), \quad (5.117)$$

As expected

- $a_r^\dagger(\vec{k})$ creates a photon with momentum \vec{k} and polarization r, and
- $a_r(\vec{k})$ destroys a photon with momentum \vec{k} and polarization r.

And so forth. Last but not least, the photon propagator is expressed as follows:

$$D_F^{\mu\nu}(x - y) = \frac{-g^{\mu\nu}}{(2\pi)^4} \int \frac{e^{-ik(x-y)}}{k^2 + i\epsilon} d^4k, \quad (5.118)$$

in physical space, and

$$D_F^{\mu\nu}(x - y) = \frac{-g^{\mu\nu}}{k^2 + i\epsilon}, \quad (5.119)$$

in momentum space.

¹²⁶ Klauber, 148.

6. Symmetries and the Standard Model

Now that we've studied the equations and their solutions for free scalars, spinors and vector particles, let's get back to symmetries and their effects. The reason for all this is to understand particle physics. That is expressed nowadays by what is called the Standard Model, which in turn is part of the Core Theory. It's called a "model" because there are lots of parameters necessary for using it, but these are not supplied by the model. The SM is expressed by QFT which depends on QM. OK?

6.1. The standard model (SM)

Particles are either fermions or bosons, according to the type of statistics they obey. ("Type" and "family" are not official terms of the SM.)

Bosons have spin which takes on only integral values – for instance, 0 or 1. They obey Bose-Einstein statistics.

Fermions all have half-integral values of spin – $\frac{1}{2}$, $\frac{3}{2}$ and so on. They obey Fermi-Dirac statistics. Most importantly, they are the basic components of matter, the blocks from which all of us are built. This is due to their being constrained by the Exclusion Principle not to occupy the same state, as we saw by 5.106.

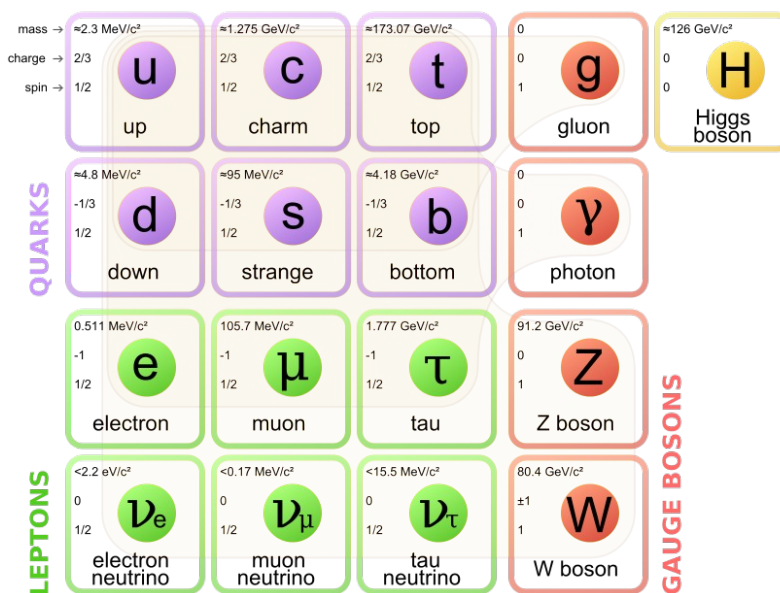


Figure 2: Elements of the standard model¹²⁷

There are two types of basic fermions – **quarks** and **leptons**. They all are shown in Figure 2, the particle zoo, which is composed of six quarks (shown in purple), six leptons (green) and four gauge bosons (orange). Table 4 displays some of their properties more visibly.

¹²⁷ Wikipedia Creative Commons,
https://en.wikipedia.org/?title=Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg.

Family	Type	Generation	Spin (I)	Particles (charge or quarks)
fermions	leptons	1	$\frac{1}{2}$	e (-1), ν_e (0)
		2		μ (-1), ν_μ (0)
		3		τ (-1), ν_τ (0)
	quarks	1	$\frac{1}{2}$	u ($+\frac{2}{3}$), d ($-\frac{1}{3}$)
		2		c ($+\frac{2}{3}$), s ($-\frac{1}{3}$)
		3		t ($+\frac{2}{3}$), b ($-\frac{1}{3}$)
	composite		$\frac{1}{2}$	p (uud), n (udd)
			$\frac{3}{2}$	Δ^{++} (uuu), Δ^{-} (ddd)
bosons	Higgs		0	Higgs
	gauge bosons		1	γ (photon), Z^0 , W^\pm , g_i (gluons)
			2	graviton
	composite (mesons)		0	π^+ ($u\bar{d}$)
			0	K^- ($\bar{u}s$)
			1	ρ^+ ($u\bar{d}$)
			1	K^{*-} ($\bar{u}s$)

Table 4. Particles of the Standard Model

Physics recognizes four forces, three of which are considered by the Standard Model.

The strong force is mediated by **quarks**, so **hadrons**, particles composed of quarks, are subject to it. Hadrons are divided into **mesons** (“middle weight”) and **baryons** (“heavy”, meaning mass \geq the proton mass). The difference between mesons and baryons is not just an imprecise question of mass, but of the number and type of quarks in them. Mesons are composed of two quarks, one quark and the corresponding antiquark; baryons, of three quarks.

Fermions are particles of half-integer spin and include the hadrons, but also the “lighter” **leptons**. Leptons are not composed of quarks and so do not interact via the strong force, although the charged ones are subject to the EM force.

Quarks are distinguished by their **flavor**¹²⁸ – up, down, charm, strange, top or bottom – which obviously are arbitrary terms and have nothing to do with taste or appearance. They are arranged in three columns called **generations**. Quarks have charges $+\frac{2}{3}$ or $-\frac{1}{3}$; leptons, 0 or ± 1 . Each of the six flavors of quark exists in three versions indicated (by analogy) by the colors red, green and blue¹²⁹, for a total of 18 combinations. When forming matter particles, the quarks must group together in such a way that the result is “colorless”, so they occur in the combination R+G+B for **baryons** like the neutron or proton, and C+ \bar{C} , where the bar above the character

128 So called by Murray Gell-Mann, the same guy who found the word “quark” while reading *Finnegans Wake*. Uh-huh...

129 Not the same as the colors on the chart.

indicates an antiparticle, for **bosons**. Color is important because it allows the existence within a nucleon of three quarks which would otherwise be in the same state and so would be forbidden by the Exclusion Principle. Antiquarks have anticolors.

The mass of, for instance, a proton (about $938 \text{ MeV}/c^2$, about 1.67×10^{-24} grams) is far greater than the sum of the masses of its constituent quarks ($3 \times 2.3 \text{ MeV}/c^2 \approx 7 \text{ MeV}/c^2$). The difference, almost 99% of the proton's mass, is potential energy of the strong force which binds the quarks together into the proton. Splitting the proton into quarks requires this amount of energy, which explains why physicists need such powerful particle accelerators.

Matter is made up of atoms with nuclei containing protons and neutrons (together called **nucleons**), with electrons forming a negatively-charged cloud around the nucleus. Such long-lived particles are made up of quarks of the first generation. A proton is composed of two up quarks and a down quark. The former have a charge of $+2/3$, the latter of $-1/3$, so the total is $+1$. A neutron is composed of an up and two downs, for a total charge of $+2/3 - 1/3 - 1/3 = 0$. And so on.

Obviously, there is some overlap in these categories, as is sketched in the following figure. Although fermions and bosons are distinct, hadrons may be either fermions or bosons, which means that some fermions and some bosons are composed of quarks, those bosons which are not are force-carrying particles. But the particles which constitute matter are always fermions.



Figure 3. Bosons, hadrons and fermions, from Hugo Spinelli via Wikimedia Commons¹³⁰.

To summarize, modern physics recognizes four fundamental forces, or interactions, in nature.

1. Gravity is theorized to be conveyed by a yet-to-be observed particle called the **graviton** (not shown in the figure because never observed and because gravity is not part of the standard model). It is a weak force (the weakest) but works across enormous, interstellar distances and is responsible for no less than the formation of stars, galaxies and planets. Gravity is due to the curvature of space, so there is debate whether it is a force at all. In fact, we should speak not of four fundamental forces, but of four fundamental interactions.¹³¹
2. The electromagnetic force between charges or magnets is conveyed by the **photon**, which is the particle of light. Like gravity, it has infinite range, but is stronger. Since its sources can be either positive or negative charges, the two cancel each other out, making the effective force at large distances weaker than gravity. The quantum theory of the electromagnetic force is **quantum electrodynamics**, or **QED**. We will confront the basic equation for QED in (6.24).
3. The strong force is the strongest, but is very short-ranged. It holds quarks together in the nucleus in spite of the repulsive electric forces between proton charges. It is conveyed by the appropriately-named **gluon**. The quantum theory of strong interactions is called **quantum chromodynamics**, or **QCD**, with which we will deal in Section 6.9.
4. The weak force is weaker than the strong or EM forces, but is still stronger than gravity. It is conveyed by the W and Z bosons. It is a very short-range force. responsible for decays of various radioactive

¹³⁰ Boson-Hadrons-Fermions, <https://commons.wikimedia.org/wiki/File:Bosons-Hadrons-Fermions-RGB.svg>.

¹³¹ Hossenfelder, "The fifth force." <https://backreaction.blogspot.com/2021/07/whats-fifth-force.html>

particles. Such decay is largely responsible for the existence of the Periodic Table and so for the elements of which our Earth and we are made. Since it enables the transmutation of a proton into a neutron and leads thence to the formation of deuterium, it is essential for the “burning” which takes place in the Sun! We will study this force in Sections 6.10 and 6.11.

Weak interactions are more complex than the others. They are the only ones which exhibit parity violation. This occurs because the weak force is asymmetric for left-handed and right-handed particles. There are two kinds of weak interactions: charged interactions, mediated by W^\pm bosons, and neutral ones, mediated by Z^0 bosons. Only the charged weak interactions can change flavor.¹³² For some reason, neutrinos are always left-handed; anti-neutrinos, right-handed.¹³³

So there two infinite-range forces, gravity and EM, and two short-range ones, the strong and weak. By order of strength, from strongest to weakest, they are strong, EM, weak and gravity.

6.2. The Core Theory

It is important to understand that quantum mechanics alone does not explain the world. It is a framework for expressing theories about the world and for doing calculations. It can, for instance, be used to explain atoms and the periodic table, but only by adding information, e.g., that a hydrogen atom is composed of a proton with one electron moving about near it under the sway of the electromagnetic force.

Modern physics considers everything to be made up of fields. This is the idea behind **quantum field theory** (**QFT**). Fields are where the buck stops. Like the bottom turtle, they are not made up of anything else (well, as far as we know). The particles which we see as the constituents of all the stuff around us are vibrations in quantum fields, fermion fields for matter, boson fields for forces. If you have trouble imaging a proton field interacting with an electron field through a vector boson field, you are not alone. It's much easier to imagine them as particles. Is that because our brains are built to comprehend particles better than fields? Who knows? Maybe we can glean a clue to understanding by remembering that Lie groups are abstract thingies which only apply to what we perceive as nature through the construction of representations. Abstract maps to real.

Although the separation of GR and QM keeps us from understanding completely the realm of the infinitely big or the vanishingly tiny, the Big Bang or black holes, most of the time who cares? The stuff around us in our everyday world is just ordinary matter and not composed of black holes, so we can effectively describe the world we live in with the standard model, based on QFT and the four forces, plus GR. This is what some physicists call the **Core theory**.

$$\text{Core Theory} = \text{QFT} + \text{GR}$$

And It works, providing the physical laws underlying chemistry, biology, astrophysics, engineering and much of cosmology (i.e., except for the Big Bang and black holes).

6.3. Noether's theorem and currents

Consider a Lagrangian density $\mathcal{L}(\phi_j, \partial_\mu \phi_j)$ 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 differentiation by parts shows that

$$\delta L = \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \delta\phi + \partial_\mu K^\mu, \quad K^\mu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta\phi.$$

If this field does satisfy the equations of motion, then by the Euler-Lagrange equations. the two-term expression

¹³² Griffiths (2008), 74.

¹³³ Griffiths (2008), 138.

in parentheses is zero by the Euler-Lagrange equations. Then

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right) \delta \phi_i = 0, \quad (6.1)$$

Since it is the action, the integral of the Lagrangian, which is extremized, we can add to the Lagrangian a surface term, a divergence of some function W , which will disappear when integrated over a surface sufficiently far away that fields are all zero. Then the quantity in question is the **Noether current** defined by

$$j^\mu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta \phi - W^\mu. \quad (6.2)$$

Equation (6.1) shows that the current has zero four-divergence $\partial_\mu j^\mu = 0$, which means

$$\partial_\mu J^\mu = \frac{\partial j_0}{\partial t} - \nabla \cdot \vec{j} = 0.$$

Integrated over all of space, so the divergence term disappears, this means the zeroth component, J^0 , is conserved in time.¹³⁴ One can see this as a charge,

$$Q = \int_{all\ space} J^0 d^3x, \quad (6.3)$$

which is constant in time. i.e., it is **conserved**. This is **Noether's theorem**.

Note that in case of several fields ϕ_i , the RHS of equation (6.2) is to be summed over the fields.¹³⁵

Applying this with various transformations leads to beloved conservation laws, as summarized in Table 5.

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

Table 5. Conserved quantities of spacetime transformations

The last boost current in the table depends on t , which can be picked so that the current is zero, which is therefore conserved. Uh... okay.¹³⁶

Let's pause to note that in the case of conserved momentum, for instance, a measurement of momentum is based on the generator of translations and so involves a translation, thus changing the value of the position. This realization offers a qualitative way to understand the **Uncertainty Principle** stating, in this case, that momentum

¹³⁴ Klauber, 173-4; Lancaster and Blundell, 94.

¹³⁵ Klauber, 173-4, 296. Klauber claims that the derivation of this current does not depend on whether ϕ is a constant, so Noether's theorem is equally valid for global and local transformations. But I'm not convinced.

¹³⁶ Schwichtenberg, 105.

and position cannot be simultaneously measured precisely.

Spacetime transformations of fields like rotations or translations are messier because changes come from the fields themselves as well as from the transformed coordinates. The results, though, are consistent with the contents of Table 5. In particular, spacetime translations

$$x^\mu \rightarrow x^\mu - a^\mu$$

lead to field transformations¹³⁷

$$\phi(x^\mu) \rightarrow \phi(x + a) = \phi(x) + a^\mu \partial_\mu \phi(x)$$

and, since it is a scalar, the Lagrangian must transform in similar fashion:

$$\mathcal{L} \rightarrow \mathcal{L} + a^\mu \partial_\mu \mathcal{L} = \mathcal{L} + a^\nu \partial_\mu (\delta^\mu_\nu \mathcal{L}).$$

Putting this together with equation (6.2) leads to a tensor current

$$T^\mu_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} \partial_\nu \phi - \mathcal{L} \delta^\mu_\nu, \quad (6.4)$$

The time-like “charge” which is conserved is

$$P^0 = \int T^{00} d^3x = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \phi)} \partial_0 \phi - \mathcal{L} \right) = \int d^3x \left(\Pi^0 \dot{\phi} - \mathcal{L} \right)$$

is the space integral of the Hamiltonian density, which is the total energy – which is conserved. The spatial “charge” is

$$P^i = \int T^{0i} d^3x = \int \pi \partial_i \phi d^3x$$

which is therefore interpreted as the (physical) momentum of the field, not the canonical momentum, in the i direction, which is also conserved. The current (6.4) is therefore the **energy-momentum tensor**.¹³⁸ The final result is in the conservation of energy and each of the three components of momentum.

We can expand a transformation $U = e^{iq\theta}$ as $U = e^{iq\theta} = 1 + iq\theta + \dots$ so that an infinitesimal transformation is given by

$$\Psi \rightarrow \Psi' = \Psi + iq\theta\Psi\dots$$

and $\delta\Psi = iq\theta\Psi$. Then the Noether current is, from (6.2),

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} \delta\Psi = \frac{\partial(\bar{\Psi} i \gamma^\mu \partial_\mu \Psi)}{\partial(\partial_\mu \Psi)} iq\theta\Psi = -q \bar{\Psi} \gamma^\mu \Psi, \quad (6.5)$$

where we ignore the arbitrary θ term.

We will find the same result for the QED interaction Lagrangian in (6.27).

Table 6 summarizes similar results for fields if equation (6.2) is used to calculate the Noether current due to a U(1) transformation on our three favorite fields.

¹³⁷ Peskin & Schroeder, 18-19.

¹³⁸ Schwichtenber, PS, 110; Blundell and Lancaster, 94-95; Peskin and Schroeder, 18-19.

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 (6.24)	$j^\mu = \bar{\Psi}\gamma^\mu\Psi$	Charge

Table 6. Conserved quantities of Noether currents for internal symmetries

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.

Equation (6.2), when applied to changes only in the fields, not the coordinates, represents internal symmetries of the system.¹³⁹ In particular, for a translation of the field,

$$\phi_i \rightarrow \phi_i + \delta\phi_i,$$

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

$$\pi_\nu = \frac{\partial \mathcal{L}}{\partial(\partial_\nu \phi)}. \quad (6.6)$$

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

6.4. Isospin

It was originally proposed that the strong interaction was symmetric under the exchange of a proton and a neutron. The concept has since been revised to handle the exchange of an up and a down quark. The idea is of a new symmetry like spin and so called **isospin**.¹⁴⁰ Like spin, isospin may be expressed in representations of SU(2) of different dimensions, as shown in Table 1. The mathematical apparatus is the same as that of angular momentum.

Since the symmetry group is SU(2), the transformations behave like rotations from one particle identity into another. The symmetry implies the existence of a Noether current and a conserved Noether charge, the isospin. To the extent that the u and d quarks have approximately the same mass and identical strong coupling to the gluon, isospin is conserved.¹⁴¹ Although the isospin behaves like angular momentum, it is, like spin, in its own internal space and has nothing to do with these other angular-momentum-like properties. Assuming, as usual, a rotation about the 3-axis, we then can talk about a set of n particles with isospin $I^2 = n(n+1)$ and $I_3 = -n \dots + n$ in integral steps.

Some examples are shown in Table 7.

¹³⁹ Klauber, 173.

¹⁴⁰ The term isospin comes from isotopic spin. Nuclear physicists call it isobaric spin. Griffiths, 129.

¹⁴¹ Barr et al., 119.

particles	Isospin I	I_3
p, n or (K^0 , K^+)	$\frac{1}{2}$	n or $K^0 = -\frac{1}{2}$, p or $K^+ = +\frac{1}{2}$
ν_e, e_L^-	$\frac{1}{2}$	$\nu_e = +\frac{1}{2}$, $e^- = -\frac{1}{2}$
e_R	0	0
u, d (quarks)	$\frac{1}{2}$	u = $\frac{1}{2}$, d = $-\frac{1}{2}$
all other quarks	0	0
Λ or Ω	0	0
W^-, Z_0, W^+	1	$W^- = -1$, $Z_0 = 0$, $W^+ = +1$
π mesons	1	$\pi^+ = 1$, $\pi^0 = 0$, $\pi^- = -1$
Δ baryons	2	$\Delta^{++} = 2$, $\Delta^+ = 1$, $\Delta^0 = 0$, $\Delta^- = -1$, $\Delta^{--} = -2$

Table 7. Some isospin examples. Subscripts L or R designate left or right-chiral.

I_3 of hadrons is easily found by summing the respective values for the constituent quarks.¹⁴² For combinations of hadrons, Clebsch-Gordon coefficients are necessary.

6.5. Gauge symmetry

Global symmetries are physical and have physical effects, and they conserve charge (Noether). Gauge invariance is not physical, being "... merely a redundancy of description we introduce to be able to describe the theory with a local Lagrangian."¹⁴³ Gauge symmetries are internal symmetries, functioning in "hidden", unobservable space. They allow changes in the configuration of the underlying, unobservable field(s) that have no distinguishable effects on observable properties. In fact, the internal space may not exist. It is only the symmetry which is a necessary assumption, not the extra dimensions.¹⁴⁴ The standard example is the EM vector potential A , specific changes in which leave the electric and magnetic fields \vec{E} and \vec{B} unchanged. Only \vec{E} and \vec{B} are observable, not A . The unobservable field is called the ***gauge field*** and a ***gauge transformation*** changes *it* from one configuration, or gauge, to another. Examples are the Lorenz or Coulomb gauges for EM.

For instance, if the field in our Lagrangian changes by

$$\phi \rightarrow \phi^{-i\alpha},$$

the Lagrangians themselves remain unchanged (invariant). So phase changes are changes of basis which do not affect the observable variables. Such a phase transformation is a ***global*** one, because α is a constant and so has the same value everywhere. Gauge symmetries, though, can be transformations which are functions of location (if α not constant) and so are ***local*** symmetries. The transformation indicates how the transformed function changes with respect to position (Imagine a rotation angle which depends on the position of the object being rotated.) and so connects together different locations. This local transformation then is said to define ***connections*** between points, usually by participating in the ***covariant derivative***. As in the case of GR, connections can be thought of as measuring the curvature of space caused or measured by the gauge field

¹⁴² "Isospin". www.asc.ohio-state.edu/gan.1/teaching/winter10/Chapter5.pdf

¹⁴³ Schwarz, 130-131.

¹⁴⁴ Maldacena, 9.

which is the connection. So it can be found by parallel transporting the field around a small, closed loop.¹⁴⁵ In fact, it is possible to treat GR as a gauge theory based on symmetries of the Lorentz group $SO(3,1)$.¹⁴⁶

6.6. Interaction Lagrangians – U(1) symmetry and QED

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

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

does not change, for instance, the Dirac Lagrangian. A unitary transformation can be viewed as a “rotation” of a complex state vector in Hilbert space which does not change the magnitude (“length”) of the state vector. It is the same everywhere in spacetime and so is a **global** transformation. But SR and its speed limit make such a global transformation problematic, although this is debated. In any case, let’s try a local transformation by letting α depend on the coordinates. Then

$$\Psi \rightarrow \Psi' = e^{i\alpha(x)} \Psi \quad \text{and} \quad \bar{\Psi} \rightarrow \bar{\Psi}' = \bar{\Psi} e^{-i\alpha(x)}. \quad (6.7)$$

Such conversion of a global symmetry to a local one is referred to as **gauging** the symmetry.

In order to understand why this is important, consider the Dirac Lagrangian

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu \partial_\mu - m)\Psi \quad (6.8)$$

and the local U(1) transformation

$$U = e^{i\alpha(x)}, \quad (6.9)$$

which is a phase rotation through an angle which varies from point to point in space. Under this transformation the mass term in the equation is clearly invariant, but the derivative term is not, since

$$i\partial_\mu e^{i\alpha(x)} \Psi(x) = ie^{i\alpha(x)} \partial_\mu \Psi(x) - \Psi(x) e^{i\alpha(x)} \partial_\mu \alpha(x). \quad (6.10)$$

This should not be surprising, since the derivative measures the rate of change in space and so depends on subtraction of one field from another infinitesimally separated from it. In the direction of a vector n^μ

$$n^\mu \partial_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - \psi(x)]. \quad (6.11)$$

But since the phase transformation varies from x to $x + \epsilon n$, the bracketed subtraction is meaningless. We need something to compensate for this variation. We can obtain this by defining a scalar **comparator** which expresses the difference between points x and y and transforms like this:¹⁴⁷

$$U(y, x) \rightarrow e^{i\alpha(y)} U(y, x) e^{-i\alpha(x)} \quad (6.12)$$

with $U(y, y) = 1$.

Now (6.11) is replaced by the covariant derivative

$$n^\mu D_\mu \psi = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - U(x + \epsilon n, x) \psi(x)]. \quad (6.13)$$

Since this depends on an infinitesimal difference, we can expand the comparator as

$$U(x + \epsilon n, x) = 1 - iq\epsilon n^\mu A_\mu(x) + \mathcal{O}(\epsilon^2), \quad (6.14)$$

where A_μ is called a **connection**, because it connects point x to point $x + \epsilon n$. Then the **covariant derivative**

¹⁴⁵ Schwichtenberg shows an example of this. NNQFT, 129-35.

¹⁴⁶ Carroll (2024), 198.

¹⁴⁷ Peskin and Schroeder, 482-483.

becomes

$$D_\mu \psi(x) = \partial_\mu \psi(x) + iq A_\mu \psi(x), \quad (6.15)$$

a very important equation.

As a consistency check, by inserting (6.14) into (6.12), we find that A_μ must transform as

$$A_\mu(x) \rightarrow A_\mu(x) - \frac{1}{q} \partial_\mu \alpha(x). \quad (6.16)$$

Not only that, but taking into account the transformation of ψ by (6.9) and of A_μ by (6.16) shows that

$$\begin{aligned} D_\mu \psi(x) &\rightarrow [\partial_\mu + iq(A_\mu - \frac{1}{q} \partial_\mu \alpha)] e^{i\alpha(x)} \psi(x) \\ &= e^{i\alpha(x)} (\partial_\mu + iq A_\mu) \psi(x) = e^{i\alpha(x)} D_\mu \psi(x), \end{aligned} \quad (6.17)$$

so the covariant derivative now transforms like the field and the Lagrangian is unchanged. This is why D_μ is called covariant.

We could have done this more simply, beginning with equation (6.10) for the derivative and introducing (produced out a hat) the connection field A_μ into the covariant derivative (6.15). It is then easy to show that this step requires the transformation of A_μ by equation (6.16). But this would have left (6.14) which shows clearly that A_μ really is the element connecting the field at two infinitesimally separated points. In either case, though, we see that the covariant derivative is valid for any derivative of a state transformed by a local U(1) transformation.

Beginning only with the requirement that the Dirac Lagrangian density (6.8) be invariant under the local U(1) transformation (6.9), we have derived the covariant derivative and the transformation rule for the connection A_μ .

There's still more. One can take U(y,x) around a closed square path using (6.14) but keeping the term in ϵ^2 ; or else consider the transformation of the commutation relation

$$[D_\mu, D_\nu] \psi(x) \rightarrow e^{i\alpha(x)} [D_\mu, D_\nu] \psi(x).$$

Both methods lead to the invariant quantity

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \quad (6.18)$$

This of course is the familiar EM field tensor and (6.16) is the gauge transformation of Maxwellian EM. So local U(1) symmetry gives us not only the covariant derivative and the transformation rule for the connection, the gauge field, but also the EM field tensor and thus all the ingredients for the QED Lagrangian.

That (6.16) be a U(1) symmetry also can be shown by considering the transformation

$$T_\alpha(A^\mu) = A^\mu(x) - \partial^\mu \alpha(x), \quad (6.19)$$

with the map Φ from U(1) to T_α

$$e^{i\alpha(x)} \rightarrow T_\alpha. \quad (6.20)$$

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

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

and

$$\begin{aligned} T_{\alpha+\beta}(A^\mu) &= (A^\mu - \partial^\mu(\alpha(x) + \beta(x))) = (A^\mu - \partial^\mu \alpha(x)) - \partial^\mu \beta(x) = T_\beta(T_\alpha(A^\mu)) \\ &= (T_\beta \otimes T_\alpha)(A^\mu) \end{aligned} \quad (6.21)$$

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 (6.21) shows that (6.16) is a homomorphism to U(1) and so a representation of the U(1) group.¹⁴⁸

A few words on the covariant derivative. As seen, the derivative of a function is used to pass from its value at one point in spacetime to a point at an infinitesimal distance from there. If a field is submitted to a transformation which varies locally, something is needed to keep track of that change from point to point. Such objects are called **connections**. They should not be confused with the Christoffel connections of GR, although those have a logically similar function, serving to keep track of the changes of the basis vectors in terms of the metric.¹⁴⁹ We have expressed the connection as an additional term to the derivative, which then becomes the covariant derivative of (6.15). At the same time, the connection introduces the vector field and in a sense is the vector field, give or take a proportionality constant. The vector field, in this case, is the photon. *The connection necessary to take into account the local U(1) gauge transformation of the Dirac Lagrangian is the photon.*

The requirement of a connection introduces a new field, a **gauge field**, into the **covariant derivative** (6.15). The change in the transformation itself is expressed in the transformation of the gauge field, as in (6.16). Taking these factors into account makes the Lagrangian locally symmetric.

Equation (6.15) should be called a **gauge covariant derivative**, to distinguish it from the GR covariant derivative, but this appellation is rarely used. On the one hand, this is a notational device to make the equation look simpler. On the other, it tells us we need to know more than just the behavior of the free Lagrangian.

Meanwhile, back at the electron, we have a Lagrangian which is locally invariant under U(1) transformations:

$$\mathcal{L} = \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi - q\bar{\psi}\gamma^{\mu}A_{\mu}\psi. \quad (6.22)$$

There is a problem, though. The equation now contains an electron and an interaction term, but nothing for the electron to interact with. So we have to add in the photon field itself and the way to do this is to add to the Lagrangian its own massless Proca term (4.7), which we just have found to be invariant under the transformation. This addition makes the A_{μ} dynamic. We could just as well (and perhaps more logically) have started with a Lagrangian for two free particles, one Dirac (6.8) and one Proca, which would have avoided having to add the Proca term here. In either case, the final result for this interaction is:

$$\begin{aligned} \mathcal{L}_{Dirac+Proca+int} &= \bar{\psi}(i\gamma^{\mu}\partial_{\mu} - m)\psi - q\bar{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{2}(\partial^{\mu}A^{\nu}\partial_{\mu}A_{\nu} - \partial^{\mu}A^{\nu}\partial_{\nu}A_{\mu}) \\ &= \mathcal{L}_0^{\frac{1}{2}} + \mathcal{L}_I^{\frac{1}{2},1} + \mathcal{L}_0^1 \end{aligned} \quad (6.23)$$

Now the Lagrangian is invariant under local SU(1) transformation, which requires transforming all three, the Dirac and Proca terms and the interaction term, according to (6.9) and (6.16). The first term in (6.23) is the Dirac Lagrangian; the second, the gauge (vector) potential multiplied by the Dirac 4-current (6.5) and the third, the Maxwell Lagrangian for a massless particle.

Note that the interaction term \mathcal{L}_I contains elements of both the Dirac and Proca (vector) Lagrangian, which seems right for an interaction between the two particles. It also shows what it means for a particle to have a **charge**, since $q=0$ removes the interaction term.

This technique (or recipe) of adding together free fields and then using the covariant derivative to bring out their interaction is called **minimal coupling**, since it ignores such things as magnetic moments. A theory which introduces a field, A^{μ} , to bring about local invariance is a **gauge theory**, so the field A^{μ} is the **gauge field**. As

¹⁴⁸ 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.

¹⁴⁹ According to Sean Carroll, the connections can be calculated, as in GR, by considering the parallel transport of a vector in the space. <https://www.preposterousuniverse.com/blog/2020/06/30/the-biggest-ideas-in-the-universe-15-gauge-theory/>

one book puts it, the gauge field is “... designed to roll around spacetime cancelling [*sic*] out terms that stop the theory from being invariant.”¹⁵⁰

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

$$\begin{aligned}\mathcal{L}_{Dirac+Proca+int} &= \bar{\Psi}(i\gamma_\mu D^\mu - m)\Psi - \frac{1}{2}(\partial^\mu A^\nu \partial_\mu A_\nu - \partial^\mu A^\nu \partial_\nu A_\mu) \\ &= \bar{\Psi}(i\gamma_\mu D^\mu - m)\Psi - \frac{1}{4}(F^{\mu\nu} F_{\mu\nu}).\end{aligned}\quad (6.24)$$

Adding local phase invariance to the Dirac field requires the inclusion of a massless vector (gauge) field, the electromagnetic field. Repeat: The A_μ field is the direct result of requiring local symmetry. All QED (as well as classical EM, of course) comes from Lorentz and local U(1) invariance + a bit of imagination. If you don't find this amazing, you should.

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 (6.23) and both variants of (6.24) represent the same physical system.

Then the Noether current

$$J^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi_i)} \delta \Psi = -g \bar{\Psi} \gamma^\mu \Psi, \quad (6.25)$$

is the **electric four-current**. The zeroth component of this is the electric charge density, so the total charge is the integral of this quantity:

$$Q = \int d^3x J^0 = -g \int d^3x \bar{\Psi} \gamma^0 \Psi = -g,$$

because of normalization. So by Noether's theorem, global U(1) symmetry means *electric charge* is conserved.

Recap:

1. Start with a Lagrangian for a spinor, which is invariant under a global U(1) transformation.
2. Gauge it, i.e., make the transformation local.
3. In order for the Lagrangian to be invariant under the local U(1) transformation, use the covariant derivative (6.11) to introduce the vector field A_μ , and
4. require that A_μ transform like (6.9), i.e., like a spin-1 particle under the same local U(1) transformation. Now the Lagrangian is locally invariant under SU(1).
5. Add in a Proca Lagrangian for the spin-1 particle so that the Noether charge will be non-zero and A_μ dynamic.¹⁵¹

There is another way to understand the interaction term. Start with Maxwell's equations in tensor form and the Lorenz gauge (i.e., with currents) and include the electron charge¹⁵²,

$$\partial_\alpha \partial^\alpha A^\mu(x) = -e j^\mu(x). \quad (6.26)$$

We have already seen by (6.5) and (6.1) that the current for a Dirac *electron* is given by

$$j^\mu = (\rho, \vec{j}) = \bar{\psi} \gamma^\mu \psi \quad \text{with} \quad \partial_\mu j^\mu = 0,$$

¹⁵⁰ Lancaster and Blundell, 128.

¹⁵¹ Robinson's section 4.5.3 is an excellent and concise presentation of this gauging of the spinor field.

¹⁵² Following Klauber.

so from (6.26) for a *photon*, we can write

$$\partial_\alpha \partial^\alpha A^\mu(x) = -e\bar{\psi}\gamma^\mu\psi. \quad (6.27)$$

We have just linked Maxwell and Dirac by using the latter's electron current density from (6.5) in the former's continuity equation. In other words, we have assumed that *the Maxwell current is composed of Dirac electrons*, by which we have linked the two and showed the interaction between them. The result is that, in general, A_μ and ψ are no longer independent of each other.

From the Lagrangian (6.24), the Euler-Lagrange equation for $\bar{\psi}$ leads us to the interaction form of the Dirac equation.

$$(i\gamma^\mu\delta_\mu - m)\psi = -e\gamma^\mu\psi A_\mu. \quad (6.28)$$

This has been applied to the hydrogen atom by assuming no magnetic field and pure Coulomb charge, so that

$$A^\mu = (\Phi, 0), \quad \text{with} \quad \Phi = \frac{e}{4\pi r}.$$

The result of the analysis gives a relativistic fine-structure that correctly describes the observed spectrum.¹⁵³

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.¹⁵⁴ So locally gauge-invariant Lagrangians are available for the interaction of a massive scalar, Dirac or vector field with a massless vector field.

In fact, as Robinson puts it: "Starting with a a non-interacting Lagrangian that is invariant under the global SU(N), we can gauge the SU(N) to create a theory with a gauge field (or synonymously a 'force carrying field') A^μ , which is an $N \times N$ matrix. Hence, every Lie group gives rise to a particular gauge field (which is a force carrying particle, like the photon), and therefore a particular force. For this reason, we discuss forces in terms of Lie groups, or synonymously **gauge Groups**."

We will see shortly (in section 6.11) how breaking a local symmetry leads to a gauge field's taking on mass.

That is the importance of these concepts in physics:

*Group theory and gauge transformations (or **gauge theory**) explain all the forces of nature.*¹⁵⁵

6.7. SU(n) gauge invariance

In order to ensure the local gauge invariance under U(1) symmetry for a spin-1/2 particle, three things were required:¹⁵⁶

- introduction of a massless vector (spin-1) field, which we take to represent a photon, including its free Lagrangian;
- addition of an interaction term depending on both types of fields (through the use of the the covariant derivative); and
- taking into account the different transformation representations of the spinor and vector fields (the latter being the same transformations used for the vector and scalar potentials of classical electromagnetism).

This was expressed in equations (6.9), (6.15) and (6.24).

¹⁵³ Klauber, 185-6.

¹⁵⁴ Schwichtenberg, PS, 142-3.

¹⁵⁵ In principle, this includes gravity, but I've never seen it demonstrated.

¹⁵⁶ Griffiths, Elementary particles, 360.

Equations (6.9) and (6.15) can be generalized to arbitrary dimensions.¹⁵⁷ For any group SU(n), composed of $n \times n$ matrices, with $n^2 - 1$ generators and of rank $n - 1$, a general transformation will be $\phi_i \rightarrow U_{ij}\phi_j$ ($i, j = 1 \dots n$) and the gauge fields transform like

$$A_\mu \rightarrow U(x)A_\mu U^\dagger(x) + \frac{i}{g}U(x)\partial_\mu U^\dagger(x). \quad (6.29)$$

which reduces to (6.16) for $U(x) = e^{i\alpha(x)}$. If we write the transformation (summed over a) as

$$U(x) = e^{-igA_\mu^a T^a},$$

where A_μ^a are the $n^2 - 1$ real parameters and T^a are the generators, the covariant derivative will be¹⁵⁸

$$D_\mu = \partial_\mu - igA_\mu^a T^a. \quad (6.30)$$

Remember that for $n > 1$, the generators are matrices and may not commute, meaning we are dealing with **non-Abelian** groups. In this case, the field tensor is

$$F_{\mu\nu}(x) = \frac{i}{g}[D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu], \quad (6.31)$$

where the final commutator term vanishes for U(1). From these last equations, we see that a Lie group implies the existence of gauge fields (6.29) which represent forces. This is summarized in Table 9 on page 98.¹⁵⁹

We can go further, although the calculations are more laborious and the results, more complicated. This is because the SU(2) and SU(3) groups, being matrices and not simple exponentials, are **non-Abelian**. Gauge theory of such fields is **Yang-Mills theory**.

6.8. SU(2) gauge invariance

We can study *local* SU(2) gauge invariance for two equal-mass Dirac fields by expressing them as one two-component column vector, or **doublet**.¹⁶⁰ A unitary matrix then can be expressed in the form

$$U = e^{iH},$$

where H is Hermitian. We can use the Pauli matrices, $\vec{\sigma}$, to form a basis, so that

$$H = \theta \vec{1} + \frac{1}{2} \vec{\sigma} \cdot \vec{a}.$$

For local SU(2) symmetry, the a s in the second term are functions of position. In this case, the derivative of the SU(2) unitary operator will bring down a factor proportional to a linear combination of the three Pauli matrices, σ , and so the calculation requires not one, but three massless vector fields in order to guarantee invariance. The covariant derivative then can take on the form of (6.30):

$$D_\mu = \partial_\mu - \frac{i}{2}g\vec{\sigma} \cdot \vec{W}_\mu. \quad (6.32)$$

In order to make the Lagrangian invariant under local SU(2) transformations, the vector fields must transform like

$$\sigma \cdot W_\mu \rightarrow \sigma \cdot W_\mu + \frac{i}{g}\sigma \cdot (\partial_\mu \alpha) - \sigma \cdot (\alpha \times W_\mu). \quad (6.33)$$

¹⁵⁷ After Robinson, 240-1; Maggiore, 244-6..

¹⁵⁸ Maggiore, 245-6.

¹⁵⁹ After Schwichtenberg, PS, 133; Maggiore, 244 (in somewhat different notation).

¹⁶⁰ Using notation of Blundell and Lancaster, 425-9. See also Griffiths, op. cit., 361-6

To complete the comparison with the SU(1) Lagrangian of (6.24), we need the force tensor

$$G_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + g(W_\mu \times W_\nu), \quad (6.34)$$

Then the locally invariant SU(2) Lagrangian is

$$\mathcal{L} = \bar{\Psi}(i\gamma^\mu D_\mu - m)\Psi - \frac{1}{4}G_{\mu\nu} \cdot G^{\mu\nu}, \quad (6.35)$$

Remember, this is for two equal-mass Dirac fields interacting with three zero-mass vector gauge fields. But no such two equal-mass Dirac fields are known. And the best candidates for the gauge fields, the W and Z bosons, have masses around 100 GeV, nowhere near zero. To get out of this dilemma, we will need the notion of symmetry breaking, coming soon in Section 6.11.

6.9. SU(3) symmetry and QCD

Requiring local gauge invariance for three equal-mass Dirac fields, expressed as a three-component column vector and so under U(3) symmetry, will require the addition of $3^2 - 1 = 8$ massless vector fields. Here the three Dirac particles are identified with 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.¹⁶¹ This set of particles forms the basis for **quantum chromodynamics** or **QCD**.

Note that all these cases require the existence of vector bosons, or **gauge bosons**. *Gauge bosons are always vector bosons.*¹⁶² In these examples, they are massless, but one can invent (unrealistic?) massive examples.¹⁶³

Now the state vector¹⁶⁴ representing three equal-mass particles is

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

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

$$U = e^{iH}$$

where H may be expressed in terms of nine real numbers a_i and θ . The *eight* 3x3 **Gell-Mann matrices** λ , the SU(3) equivalent of the three Pauli matrices in SU(2), fit the bill nicely:

$$\lambda^1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

¹⁶¹ Griffiths, 366-9. The following triplet example is inspired by these pages.

¹⁶² Quora, www.quora.com/How-do-gauge-bosons-and-vector-bosons-differ?share=1

¹⁶³ Schwichtenber, 150, presents the example of two massless Dirac fields (What would that be?), which requires three *massive* gauge bosons.

¹⁶⁴ Or whatever it is called...

$$\lambda^4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\lambda^7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}. \quad (6.36)$$

Note that, on the one hand, they are Hermitian and, on the other, that only λ^3 and λ^8 are diagonal, Cartan operators, meaning that the **rank** of SU(3) is 2. The number of generators is indeed $n^2 - 1 = 8$. With these matrices, H may be expressed as

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

so

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

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.

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

Following the U(1) case of (6.11) and the SU(2) of (6.32), use a covariant derivative

$$\mathcal{D}_\mu \equiv \partial_\mu + iq\lambda \cdot A_\mu \quad (6.37)$$

where there are now eight gauge fields A_μ . We want the transformation to function as

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

In the infinitesimal case, this yields a formula equivalent to (6.9) or (6.33)

$$A'_\mu \cong A_\mu + \partial_\mu \phi + 2q(\phi \times A_\mu) \quad (6.38)$$

with the cross product being

$$(B \times C) = \sum_{j,k=1}^8 f_{ijk} B_j C_k \quad (6.39)$$

summed over all eight vector fields, where f_{ijk} are the structure constants of SU(3). Finally, the **complete Lagrangian for chromodynamics** is

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi - \frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} - (q\bar{\psi}\gamma^\mu \lambda \psi) \cdot A_\mu. \quad (6.40)$$

This just comes from requiring that the free-particle Lagrangian be locally invariant under SU(3). Comparing (6.40) to (6.23) and using (6.25), we see that the Dirac fields provide eight (one for each λ matrix) color currents”

$$J^\mu \equiv q(\bar{\psi}\gamma^\mu \lambda \psi) \quad (6.41)$$

which constitute sources for the color fields A_μ .

We have therefore described a state of three equal-mass Dirac particles, taken to be the three color states of a given flavor of **quark**, interacting with eight massless vector fields, the **gluons**. Equation (6.40) is the correct one for the strong interaction, one for each of the six quark flavors (red, blue, green and their anti-colors).¹⁶⁵

¹⁶⁵ All this derivation based on – almost copied from – Griffiths, EP, 366-369. Essentially the same derivation, some

The symmetries of this system are under rotations in the 3-dimensional **color space**. In general, a quark is not in a specific color state, the triplet being able to take on any direction in the space which sums the colors to white.

As in the SU(2) case (6.31), $F^{\mu\nu}$ now is not as simple as in QED, but contains a cross product of the vector field with itself:

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu - 2q(A^\mu \times A^\nu). \quad (6.42)$$

The cross product, defined as in (6.39), allows interactions to take place between one A_μ field and another. Remember that we interpret field excitations as particles, via the creation and annihilation operators resulting from second quantization. So what happens here is that one field's particle excitations may interact with another's and "create" more particles. Such a field can be a source of itself, as in the case where a gluon may decay into a pair of gluons.¹⁶⁶ So as we measure the field farther and farther away from a quark, the color charge we see actually increases. This is the inverse of the QED case, wherein virtual e^-e^+ pairs form dipoles resulting in **vacuum polarization** which screens the electron charge, making it seem smaller at a distance.¹⁶⁷

We can interpret this interaction among quarks as being due to the fact that they carry color. This is contrary to the case of photons, which do not carry EM charge and so do not interact with each other by the EM force.¹⁶⁸

The color force acts only between quarks and is responsible for binding them tightly one to another. But deep probes (inelastic scattering) show weaker coupling between them at higher energies, in agreement with the behavior of the preceding paragraph. Close together, they can act almost as if they were free, hence the name **asymptotic freedom** given to this phenomenon. It is the freedom of the quarks in the limit as the distance between them goes to zero. But as we try to pull them away from each other, the binding force becomes stronger as more gluons are created. This fact accounts for our inability to observe separate quarks.

Since SU(3) non-Cartan generators describe color-changing rotations in color space, gluons must carry color from one quark to another. Therefore each gluon must correspond to a color and an anti-color, the latter taking away what the former donates, effectively exchanging one color for another. For instance, a "blue" quark may convert to a red quark (of the same flavor) by emission of a (b, \bar{r}) gluon, so color is conserved.

The matrix representation of gluons has the form¹⁶⁹

$$g_\alpha^\beta = \begin{pmatrix} r\bar{r} & r\bar{g} & r\bar{b} \\ g\bar{r} & g\bar{g} & g\bar{b} \\ b\bar{r} & b\bar{g} & b\bar{b} \end{pmatrix},$$

so each gluon is bicolored, with one positive unit of color and one negative unit. Consider then an $r\bar{g}$ gluon:¹⁷⁰

$$g_r^{\bar{g}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and three quarks

constants aside, is done by Lancaster & Blundell, 242-244, 307-308, but I can't find that anymore.

166 For details, such as (6.41), not all of which I can claim to understand, see Lancaster & Blundell, 428-429.

167 Griffiths, 68-69.

168 They might do so at very high energies by mutating into an electron-positron pair. I don't think this has been observed yet.

169 Robinson, op. cit., 265

170 Example from Robinson, 254.

$$q_r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad q_g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad q_b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then the possible interactions are

$$g_r^{\bar{g}} q_r = 0, \quad g_r^{\bar{g}} q_g = q_r, \quad g_r^{\bar{g}} q_b = 0,$$

so the anti-green gluon only interacts with a green quark, which it converts to a red one.

6.10. Physical results of gauge symmetry

In all three cases of U(1), U(2) or U(3), in order to guarantee local symmetry, one can either write the interaction term or use a covariant derivative with an additional connection term. Note that for U(1) and SU(3), EM and QCD, the vector fields – photons and gluons – are indeed massless. Such is not the case for the W and Z bosons of the weak interactions, nor for an initial doublet (of one generation) of an electron and an electron neutrino, which otherwise could correspond to the SU(2) case.

Consider the general schema of what happens in the last cases.

- Take one, two or three spinors and write the globally symmetric free-field (Dirac) Lagrangian.
- Make the global symmetry local by U(1), SU(2) or SU(3), which calls for an extra term in the Lagrangian.
- Include the extra term by using a covariant derivative with a connection which represents one or more vector particles, the gauge bosons. This also requires taking into account their transformation properties.
- For overall symmetry and non-zero Noether current, add in the free (Proca) Lagrangian for the gauge bosons.

You now have a locally symmetric Lagrangian for the original spinors plus a number (1, 3 or 8) of massless gauge bosons, including the interactions terms in their minimal coupling version. It only remains to identify this system with some particles in nature and their properties.

Klauber sums up an amazing general rule for QFT:¹⁷¹

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.

As Sean Carroll points out, gauge theories not only give rise to forces, but impose constraints on interactions (through conservation laws due to Noether's theorem).¹⁷²

On top of all that, gauge invariance of the QED Lagrangian requires the photon to have zero mass.

But there is more...

6.11. Spontaneous symmetric breaking and the Higgs mechanism

A magnet heated above a certain critical temperature, T_c , loses its magnetism as the individually moving dipoles come to point randomly in different directions, making the system symmetric: Turning it through any angle

¹⁷¹ Klauber, 296.

¹⁷² Carroll, video, gauge theory. <https://www.preposterousuniverse.com/blog/2020/06/30/the-biggest-ideas-in-the-universe-15-gauge-theory/>. Carroll seems to be saying that multiple quarks are created, not gluons, but I doubt that.

changes nothing observable. As the magnet cools back down through T_c , a phase transition takes place and it becomes re-magnetized with the dipoles eventually all pointing in only one direction. Cooling, by forcing the dipoles to align, has caused the system to lose its symmetry, or to “hide” it. This kind of phenomenon is referred to as **spontaneous symmetry breaking**.¹⁷³

A similar event is supposed to have occurred around 10^{-35} seconds after the Big Bang, when a symmetric state of the Universe went through a phase change into a configuration which was no longer symmetric.

In order to show how this can take place, we must modify our Lagrangian. The requirement that it be a scalar excludes odd powers of the field, so the next simplest equation to the Klein-Gordon Lagrangian for scalar fields is obtained by addition of a 4th-power term:

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{\lambda}{4!}\phi^4$$

This Lagrangian has been used in the analysis of ferromagnetism and is known as the ϕ^4 theory, or **phi-4 theory**. If the mass of the magnetic elements varies with temperature, we can suppose that, near the critical temperature T_C , a Taylor series will give $m^2T = c(T - T_C)$ for some constant c .¹⁷⁴ Then for $T > T_C$ the mass term in ϕ^2 has the sign of an ordinary scalar Lagrangian. However, below T_C the “mass” term is negative, corresponding to an imaginary mass, and so looks like part of a potential. In this case, the extremum at $\phi = 0$ is a local maximum and is unstable. This motivates changing the sign of the mass term. Then the minimum of the potential is not at $\phi = 0$, but at $\phi = \pm\sqrt{\frac{6m^2}{\lambda}}$. Expanding the Lagrangian around one of these values leads to a term for a real value of mass. Following these considerations, we will use as an example a complex version of the same ϕ^4 theory with an inverted sign of the mass term.¹⁷⁵

Remember that SU(2) gauge invariance for a doublet of two equal-mass Dirac spinors (section 6.8)? It requires the existence of three massless vector fields, but no such particles are known, nor are the initial equal-mass spinors. Here is where symmetry breaking comes to the rescue.

In order to get an idea of what happens, consider a simpler two-scalar configuration and U(1) symmetry, which is Abelian.¹⁷⁶ We will complexify the system by taking linear combinations of the two fields, ϕ_1 and ϕ_2 .¹⁷⁷

$$\phi \equiv \phi_1 + i\phi_2, \quad \text{so} \quad \phi^* \equiv \phi_1 - i\phi_2 \quad (6.43)$$

so that

$$\phi^* \phi = \phi_1^2 + \phi_2^2.$$

The system now consists of a complex scalar. Then a ϕ^4 Lagrangian with inverted mass-term sign¹⁷⁸ for the two fields can be written

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2. \quad (6.44)$$

The inverted sign on the mass (μ^2) term looks like the mass is imaginary, which is not physical.¹⁷⁹ Or, as already mentioned, one can say it looks like part of the potential energy. Now **global** rotational symmetry of this system is U(1) symmetry, which is Abelian. The minimum of the potential, the two right-hand terms, is not at $\phi = 0$, but on a circle at

¹⁷³ Example from Griffiths, 376.

¹⁷⁴ Schwartz, 562.

¹⁷⁵ That was a hand-waving defense of equation (6.44). Another defense is the fact that it works

¹⁷⁶ Blundell and Lancaster, 430-431, treat the same problem with a non-Abelian field.

¹⁷⁷ Notation from Griffiths, 376.

¹⁷⁸ Grabbed out of a convenient hat...because it works. But we see only in a moment that it represents a scalar. Defended by Schwartz, 562.

¹⁷⁹ Griffiths, 373.

$$\phi_{1min}^2 + \phi_{2min}^2 = \mu^2 / \lambda^2,$$

corresponding to the inside of the brim of the sombrero in the image, the so-called “Mexican hat potential”. (It has also been, more accurately in my opinion, compared to the bottom of a wine bottle.¹⁸⁰) Since any point on the circle represents the same vacuum energy, the vacuum is *degenerate*.

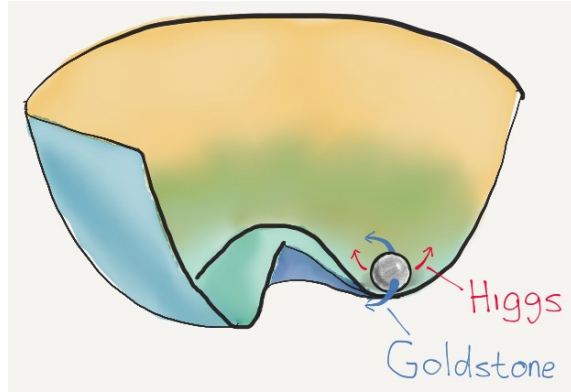


Figure 4: “Mexican hat potential”, from Free Thought Blogs¹⁸¹

We are in a situation like that of the magnet at $T > T_c$. In case of a “phase transition”, the system can fall into any state on this circle – but only one. The system still possesses symmetry, but it is obliged to choose only one state and this selection of one state spontaneously breaks the symmetry.

Choice of a state of minimum, or **vacuum**, potential puts the system at a specific point on the inverted brim of the hat, where a particle can move in one of two perpendicular directions:

- perpendicular to the rim, in which case it must roll uphill, against a force (the gradient of the potential), thus behaving as if it had mass;
- along the rim, in which case the potential does not change and the lack of any force on the particle is interpreted as it's being massless.

Global U(1) symmetry allows us to choose a gauge in which the vacuum is at some real value of ϕ , Φ .¹⁸² Nevertheless, the new field will be complex, depending on two real parameters:

$$\phi = \Phi + \alpha + i\beta \quad \text{and} \quad \phi^\dagger = \Phi + \alpha - i\beta.$$

Expanding the Lagrangian leads to

$$\mathcal{L} = \left[-\frac{1}{2} \partial^\mu \alpha \partial_\mu \alpha - \frac{1}{2} 4\lambda\mu^2 \Phi^2 \alpha^2 - \frac{1}{2} \partial^\mu \beta \partial_\mu \beta \right] \dots + \text{interaction terms}.$$

This now looks like a real massive scalar α , with $m = \sqrt{4\lambda^2 \Phi^2}$, plus a massless real scalar field β , plus some interaction terms. This is a general result called **Goldstone's theorem**: Global symmetry breaking always leads to the existence of a massless boson, called a **Goldstone boson**.

That was for breaking of global symmetry. In order for the Lagrangian to be invariant under *local* U(1) gauge transformations

$$\phi \rightarrow e^{i\theta(x)} \phi, \tag{6.45}$$

¹⁸⁰ I have much more experience with wine bottles than with sombreros.

¹⁸¹ Illustrating the Higgs mechanism, Part 1. <https://freethoughtblogs.com/atrivialknot/2016/06/27/illustrating-the-higgs-mechanism-part-1/>

¹⁸² This example from Robinson, 236-8.

we must employ the usual trick of introducing a massless vector gauge field A^μ and a covariant derivative \mathcal{D}_μ , such that¹⁸³

$$\mathcal{D}_\mu = \partial_\mu + iqA_\mu \quad (6.46)$$

and

$$\mathcal{L} = \frac{1}{2}(\mathcal{D}_\mu \phi^*)(\mathcal{D}^\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2 - \frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu}. \quad (6.47)$$

Let's choose a particular state, thus breaking the symmetry, by changing coordinates so that

$$\eta \equiv \phi_1 - \frac{\mu}{\lambda} \quad \text{and} \quad \xi \equiv \phi_2. \quad (6.48)$$

Then the Lagrangian expands to¹⁸⁴

$$\begin{aligned} \mathcal{L} = & \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2 \right] + \left[\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi) \right] \\ & + \left[-\frac{1}{16\pi}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2}\left(\frac{q\mu}{\lambda}\right)^2 A_\mu A^\mu \right] + \dots + \left(\frac{q\mu}{\lambda}(\partial_\mu \xi)A^\mu\right). \end{aligned} \quad (6.49)$$

This Lagrangian represents the same state as (6.47), only now it contains the following terms (separated by brackets):

- One term which looks like the K-G Lagrangian for a scalar particle η of mass μ . (Spoiler: This is the Higgs field.)
- A second term which corresponds to a massless scalar Goldstone boson.
- A Proca term for the vector field A^μ which – lo, behold! – now has a mass

$$m_A = 2\sqrt{\pi}\frac{q\mu}{\lambda},$$

which is proportional to the shift $\frac{\mu}{\lambda}$ in ϕ_1 and so is also due to the potential minimum's not being at symmetric zero.

- An extraordinarily messy term coupling the three fields η , ξ and A^μ (not shown).
- A term in $(\partial_\mu \xi)A^\mu$, which looks inconveniently like a ξ turning into an A^μ .

So we can say that local U(1) symmetry and subsequent symmetry breaking by the translation of the field zero, which we will see is its vacuum state, to the minimum of the potential has conferred mass on two fields, scalar and vector, and produced a third – unwanted – one.

Both these problem terms involve $\xi = \phi_2$. That can be fixed. We can specify a **particular gauge** for the U(1) transformation (6.45) so that¹⁸⁵

$$\theta = -\tan^{-1}(\phi_2/\phi_1). \quad (6.50)$$

Then

$$\phi \rightarrow e^{i\theta x} \phi = (\cos\theta + i\sin\theta)(\phi_1 + i\phi_2)$$

becomes real and $\xi = 0$, which eliminates both the Goldstone boson and the unfortunate $\xi - A^\mu$ interaction term.¹⁸⁶ Finally, we have

183 Griffiths, 378-80.

184 Griffiths, 379.

185 Robinson, 238-9, applies such a unitary gauge earlier, which simplifies the series of equations. A different path yet is taken by Blundell and Lancaster, 239-242.

186 Griffiths, 380.

$$\mathcal{L} = \left[\frac{1}{2} (\partial_\mu \eta) (\partial^\mu \eta) - \mu^2 \eta^2 \right] + \left[-\frac{1}{16\pi} F^{\mu\nu} F_{\mu\nu} + \frac{1}{2} \left(\frac{q\mu}{\lambda} \right)^2 A_\mu A^\mu \right] + \text{interaction terms and a constant.} \quad (6.51)$$

We started with one complex scalar field and a massless vector field. From that, **local** symmetry breaking has led to a single, massive (vector) **gauge field** A^μ and a single massive scalar field η , called the **Higgs field**. The mechanism whereby the scalar field is conferred a mass by shifting the field zero is called the **Higgs mechanism**.

Recap:

- Break symmetry by expressing the Lagrangian relative to the vacuum, at a non-zero minimum value of the potential.
- Choose a gauge to eliminate Goldstone bosons and give mass to a scalar.

So we have made the following discoveries:

- Breaking a global symmetry always leads to the creation of a massless Goldstone boson.
- Local symmetry requires the presence of a gauge field, and breaking the symmetry adds mass to that field.

We have seen this in the case of global and local U(1)-invariant Lagrangians. Since the quantities involved are scalars, they commute and this is an Abelian gauge theory.

Note well that in equations (6.45) through (6.51), the force-carrying vector boson field A_μ is introduced in order to assure local U(1) symmetry. The Higgs only shows up as a result of symmetry breaking. Sabine Hossenfelder claims that, for this reason, most physicists don't call the Higgs a force. "The reason is that the exchange particles of electromagnetism, the strong and weak nuclear force, and even gravity, hypothetically, all come out of symmetry requirements. The Higgs-boso [sic] doesn't. That may not be a particularly good reason to not call it a force carrier, but that's the common terminology. Four fundamental forces, among them is gravity, which isn't a force, but not the Higgs-exchange, which is a force. Yes, it's confusing."¹⁸⁷

We shall see shortly that the Feynman method of adding possible states is a perturbation theory around a ground state (or vacuum) which is a state of minimum energy, hence the necessity of a choice of vacuum (6.48) which brings about spontaneous symmetry breaking. Then a suitable gauge choice (6.50) causes the gauge field to "eat" the Goldstone boson and gain a mass. At the same time, it acquires a third polarization state, whereas massless particles only have two.¹⁸⁸ So we can say:

Local gauge invariance + spontaneous symmetric breaking \implies **Higgs mechanism**.

This is the process which is supposed to have taken place when the Universe was very, very young and which brought about creation of mass for the weak-interaction gauge bosons W^\pm and Z^0 .

The Higgs mechanism is responsible for the masses not only of vector bosons but of fermions (e^- , τ , ...) and even quarks. A more nuanced statement of this is "...the Higgs field is indirectly responsible for the fermion masses... It just allows you to have consistent extra terms in your initial Lagrangian, which would give you the mass terms result..."¹⁸⁹ Proton and neutron masses, though, come not from Higgs but from quark binding energy.

¹⁸⁷ "The fifth force", [https://backreaction.blogspot.com/2021/07/whats-fifth-force.html?](https://backreaction.blogspot.com/2021/07/whats-fifth-force.html?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+blogspot%2Fermku+%28Backreaction%29)

[utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+blogspot%2Fermku+%28Backreaction%29](https://backreaction.blogspot.com/2021/07/whats-fifth-force.html?utm_source=feedburner&utm_medium=email&utm_campaign=Feed%3A+blogspot%2Fermku+%28Backreaction%29).

¹⁸⁸ Griffiths, 380-1.

¹⁸⁹ Higgs field mass of fermions. Physics Forums. <https://www.physicsforums.com/threads/higgs-field-mass-of->

For once, Wikipedia explains it simply and understandably:

“The simplest description of the [Higgs] mechanism adds a quantum field (the Higgs field) that permeates all space to the Standard Model. Below some extremely high temperature, the field causes spontaneous symmetry breaking during interactions. The breaking of symmetry triggers the Higgs mechanism, causing the bosons it interacts with to have mass.”¹⁹⁰

6.12. Electroweak interactions

It is supposed that when the Universe was less than 10^{-12} seconds of age, the weak and EM forces were parts of an **electroweak force** which unified the two into one. This force was symmetric before a phase change resulted in “spontaneous” symmetry breaking and bestowed masses on the gauge bosons.

In this document, we have not yet reached understanding of the weak interactions, much less the electroweak. Starting with two equal-mass Dirac fields got us nowhere, which is understandable, since we don’t know what such initial state would be (section 6.8). And the predicted gauge bosons were massless, which the W^\pm and the Z^0 are not. The “simple” version of the Higgs mechanism we have considered started from two scalars, but what could they be? We need to take into account the experimental result that although electrons may show left or right chirality, neutrinos only have left-handed chirality. All are excitations of Fermi fields.

6.12.1. La recette

Here’s the recipe for what we will do, with references in square brackets to following explanations:

- 1 triplet of massless ν_e , e_L and e_R [1]
- 1 local U(1) transformation [2], a function of weak hypercharge Y , with its associated gauge field, B_μ
- 1 local SU(2) transformation [3], a function of isospin, I , with its associated gauge field, W_μ
- 1 covariant derivative on the wave function Ψ for each of the two local transformations [4]
 - Define the values of I , I_3 and Y for the triplet particles (Table 8)
 - Separate left and right-chiral parts in the triplet [5]
- 1 four-component complex Higgs field ϕ [6] of non-zero mass m_h with its U(1) and SU(2) covariant derivatives
- 1 term for the interaction between the Higgs and the fermions [7]
 - Add ingredients together [8] to get the Weinberg-Salam electroweak (EW) Lagrangian
 - Break the symmetry [9] by defining a simple ground state using a unitary gauge
 - [10] Note that the EM transformation looks like what we did for the EM field, so all that is still good and Maxwell’s equations are valid.
 - [11] Fold together, let rise (expand) and identify terms to get non-zero masses [12] for the electron, W^\pm and Z^0 .

So let’s get cooking.

fermions.752061/

190 Higgs mechanism, https://en.wikipedia.org/wiki/Higgs_mechanism.

6.12.2. La cuisine

[1] We need a triplet with left and right-chiral electrons plus left-chiral neutrinos.¹⁹¹ In line with what we just said about spontaneous symmetry breaking, we will take them all to be *massless*. We are then led to a column vector which is significantly more complex than it looks here, because the elements contain Dirac spinors:

$$\Psi(x) = \begin{pmatrix} \nu_e(x) \\ e_L(x) \\ e_R(x) \end{pmatrix}.$$

At a moment just after the Big Bang, we assume the system possessed local U(1) and SU(2) symmetries.

[2] Define the local U(1) transformation by

$$U = e^{\frac{i}{2}Y\beta(x)} \quad (6.52)$$

with a required gauge field-theory

$$B_\mu \rightarrow B_\mu + \frac{1}{g_1}\partial_\mu\beta. \quad (6.53)$$

The factor g_1 is the strength of the coupling of particles to the **weak hypercharge**, Y. Note that it is not a parameter in the transformation (6.52) but comes from the transformation of B_μ (6.21). In fact, to say that a particle is **charged** means that such a term occurs in the covariant derivative.¹⁹²

[3] Similarly, we can write the effect of the local SU(2) symmetry by the transformation (6.32-6.34)

$$\begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \rightarrow e^{\frac{1}{2}\sigma\cdot\alpha(x)} \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \quad e_R \rightarrow e_R$$

and

$$\sigma \cdot W_\mu \rightarrow \sigma \cdot W_\mu + \frac{1}{g_2}\sigma \cdot (\partial_\mu\alpha) - \sigma \cdot (\alpha \times W_\mu).$$

[4] The corresponding covariant derivatives are then

$$\text{U(1):} \quad D_\mu\Psi = \partial_\mu\Psi - \frac{i}{2}g_1YB_\mu(x)\Psi$$

and

$$\text{SU(2):} \quad D_\mu\Psi = \partial_\mu\Psi - ig_2I\vec{\sigma} \cdot \vec{W}_\mu\Psi,$$

where I is the isospin. I have picked g_1 for the U(1) part and g_2 for the SU(2),¹⁹³ although each author uses something different. The values of Y and I for these leptons is shown in Table 8.

¹⁹¹ I am following here the method of Lancaster & Blundell, chap. 47. Numbers in bracket are mine, steps in the recipe.

¹⁹² Robinson, 202.

¹⁹³ Like Robinson.

Field	ν_e	e_L	e_R	Higgs boson	Symmetry group
Y	-1	-1	-2	+1	U(1)
I	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	SU(2)
I_3	$\frac{1}{2}$	$-\frac{1}{2}$	0		

Table 8. Weak hypercharge Y and isospin I of leptons

The Lagrangians and covariant derivatives can be written in terms of left (ν_e, e_L) and right (e_R) parts, which simplifies the equations somewhat.

[5] We can simplify matters a bit by splitting the Dirac Lagrangian into left-chiral and right-chiral parts by writing

$$\psi = \begin{pmatrix} L \\ R \end{pmatrix},$$

with

$$L = \begin{pmatrix} \nu_e \\ e_L \end{pmatrix}, \quad R = e_R.$$

[6] Since we know these particles are not at all massless, this is where we need the Higgs mechanism. It is a four-component field with hypercharge and isospin defined as in Table 8. Its covariant derivative then looks like

$$D_\mu \phi = \partial_\mu \phi - \frac{i}{2} g_2 \vec{\sigma} \cdot \vec{W}_\mu \phi - \frac{i}{2} g_1 B_\mu \phi. \quad (6.54)$$

As before (6.44), the Higgs contribution to the Lagrangian is a phi-4 potential with positive mass term:

$$\mathcal{L}_\phi = (D^\mu \phi)^\dagger (D_\mu \phi) + \frac{m_h^2}{2} \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2. \quad (6.55)$$

[7] Since we want to understand the interactions among these particle, we need to add a term for the interaction between the Higgs and the electron/neutrino fields. We use

$$-G_e (\bar{L} \phi R + \bar{R} \phi^\dagger L), \quad (6.56)$$

where G_e is the coupling strength. This is analogous to the EM interaction term, $-q A_\mu \bar{\psi} \gamma^\mu \psi$ from (6.22).

[8] When this is done, we get the Lagrangian for the **Weinberg-Salam electroweak model** in all its glory.

$$\begin{aligned} \mathcal{L} = & \bar{L} i \gamma^\mu D_\mu L + \bar{R} i \gamma^\mu D_\mu R \\ & + (D^\mu \phi)^\dagger (D_\mu \phi) + \frac{m_h^2}{2} \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2 \\ & - G_e (\bar{L} \phi R + \bar{R} \phi^\dagger L) \\ & - \frac{1}{4} G_{\mu\nu}^{(W)} \cdot G^{(W)\mu\nu} - \frac{1}{4} F_{\mu\nu}^{(B)} F^{(B)\mu\nu}. \end{aligned} \quad (6.57)$$

The first line contains terms for the massless, left and right-chiral Dirac particles (fermions); the second, the Higgs ϕ^4 potential; the third the fermion-Higgs interaction term; the last the boson fields for the B_μ and W_μ .

[9] Starting with the total Lagrangian for the three massless fermions and the Higgs, including the interaction, we break the symmetry by assigning a particular ground state to this last field, in the “gutter” of the sombrero-shaped potential, such that

$$(\phi^\dagger \phi)_0 \equiv v^2 = \left(\frac{m_h^2}{\lambda}\right).$$

Then we can pick a gauge transformation of such form that the ground state is

$$(\phi)_0 = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (6.58)$$

This is the form of a **unitary gauge**. Then

$$\phi = \begin{pmatrix} 0 \\ v + \frac{h(x)}{\sqrt{2}} \end{pmatrix}, \quad (6.59)$$

where $h(x)$ is the constant background value of the Higgs field.¹⁹⁴

[10] First, notice that the total local transformation from U(1) and SU(2) is

$$U = e^{i(\frac{Y}{2} + I_3 \sigma^3) \alpha(x)},$$

which is equivalent to

$$U = e^{iQ\alpha(x)},$$

with

$$Q = \frac{Y}{2} + I_3, \quad (6.60)$$

the Gell-Mann-Nishijima relation. This is the same U(1) transformation which led us to QED in section 6.6, and so the electroweak theory is consistent with Maxwell's equations. Whew!

[11] Inserting (6.59) into the Higgs-lepton interaction term (6.56) in the electroweak Lagrangian leads to the electron mass in terms of the coupling strength of this interaction, G_e , and the Higgs ground-state parameter v :

$$v^2 = \left(\frac{m_h^2}{\lambda}\right), \quad (6.61)$$

$$m_e = G_e v. \quad (6.62)$$

So the Higgs field is responsible for the mass of the electron. Since the neutrino does not occur in the equation, its mass remains at zero, fortunately. Looked at the other way, the Higgs mass is given by $m_h^2 = \lambda v^2$ and has been found experimentally to be 125.10 ± 0.14 .¹⁹⁵

[12] But there's more. We can calculate the value of

$$\begin{aligned} (D_\mu \phi)^\dagger (D^\mu \phi) &= \frac{1}{2} [\partial_\mu h(x)]^2 + \frac{g_2^2 v^2}{4} (W_\mu^1)^2 + \frac{g_2 v^2}{4} (W_\mu^2)^2 \\ &\quad + \frac{v^2}{4} [(g_2 W_\mu^3 - g_1 B_\mu)^2] + \text{higher order terms} \end{aligned} \quad (6.63)$$

Looking for boson mass terms of the form $\frac{1}{2} (\text{mass})^2 \times \{\text{field}\}^2$. we find massive vector particles W_μ^1 and

W_μ^2 with mass squares $M_W^2 = \frac{g_2^2 v^2}{4}$ and a linear combination $(g_2 W_\mu^3 - g_1 B_\mu)$ whose mass depends on the coupling constants. We then use a shorthand notation in terms of something called the **weak mixing angle** (or

¹⁹⁴ Schwichtenberg (2015), 155,

¹⁹⁵ Particle physics data group, <https://pdg.lbl.gov/2020/listings/rpp2020-list-higgs-boson.pdf>. There is a factor of 2 difference with Robinson due to his definition of the ground state $\frac{v}{\sqrt{2}}$.

Weinberg angle):

$$\theta_W = \tan^{-1}\left(\frac{g_1}{g_2}\right),$$

so that

$$s_W = \sin\theta_W = \frac{g_1}{\sqrt{g_1^2 + g_2^2}}, \quad c_W = \cos\theta_W = \frac{g_2}{\sqrt{g_1^2 + g_2^2}}.$$

From this we define two fields as follows:

$$Z_\mu = c_W W_\mu^3 - s_W B_\mu, \quad A_\mu = s_W W_\mu^3 + c_W B_\mu. \quad (6.64)$$

Then the last, mixed term in (6.63) becomes

$$+ \frac{g_2^2 v^2}{4c_W^2} Z_\mu^2$$

and we see that the mass of the Z^0 is $\frac{M_W}{c_W}$.

Finally, after a fair amount of math, we can express the physical particles as sums of the W and B fields.

$$\begin{aligned} W_\mu^+ &= \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2), & W_\mu^- &= \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2), \\ Z_\mu &= c_W W_\mu^3 - s_W B_\mu, & A_\mu &= s_W W_\mu^3 + c_W B_\mu. \end{aligned} \quad (6.65)$$

So, once again, symmetry breaking has contributed a mass to the W and Z fields but leaves the A field massless, as desired. The W and Z vector bosons are the force-carrying fields of the **weak force**.

Again, the W^\pm are sums of non-Cartan generators of $SU(2)$ and so are ladder operators, capable of raising or lowering the charge. They constitute what is referred to as a **weak charged current**, responsible for events like the decay of a neutron into a proton, an electron and an anti-electron-neutrino. On the other hand, Z^0 and A are sums of Cartan generators of $SU(2)$ and $U(1)_Y$, respectively, and will not change the charge. Since the basis vectors are eigenvectors of the Cartan operators, the result of operating with them just gives the eigenvalue multiplied by the same basis vector. The Z^0 is thus responsible for the **weak neutral current**, accounting, for instance, for neutrino-electron scattering. Back in those very brief times, the good ol' massless EM photon was not just a descendant of the EM gauge boson but a linear combination of that and of the Cartan generator of $SU(2)$. The A is massless and a single $U(1)$ symmetry remains unbroken. These are the field and gauge group of EM. The W and Z masses are quite large, making the weak force effective only over a very short range.

This result was obtained by symmetry breaking. Before the symmetry broke, there was only a Higgs complex vector field with three massless vector boson gauge fields, each behaving generally like a photon. The low-energy, "broken" theory is based on four linear combinations of the original fields, three of which have gained mass. The original high-energy state is the content of **electroweak theory**. The original three photon-like fields break to form two distinct forces, the broken weak and the unbroken EM. Looked at the other way around, the merging of two states into one at high energies is called **unification**.

Putting everything together, one can calculate a Lagrangian for the electroweak gauge fields and the Higgs boson in unitary gauge. The resulting four-line equation (6.57) leads to the following conclusions.¹⁹⁶

- The photon field A_μ does not couple to the Higgs field, as there is no term with the product of the two. So the photon remains massless.

¹⁹⁶ Srednicki, 530; Robinson, 258-9.

- The W^\pm fields do interact with the Higgs field, there being a term proportional to W^+W^-h . This gives mass to these bosons and also accounts for the decay of the Higgs particle into pairs of W (and Z) bosons, which allows us to detect it.

6.13. Summary – Forces and particles

The notion of symmetry is based on a transformation after which a system is left in a different state but behaves the same. Such transformations are seen as changes resulting from the application of forces. Since each symmetry of the Lagrangian is described by a **Lie group**, we can take a **representation** of the Lie group in order to describe the results of applying a force to the system described by the Lagrangian.

Valid Lagrangians possess symmetries. Indeed, what constitutes a valid Lagrangian is imposed by the requirement of Lorentz invariance – symmetry under Lorentz transformations. Noether's theorem tells us that for each symmetry, a calculable quantity is conserved -- energy or momentum or angular momentum, for instance, but also spin, hypercharge or color. This quantity will be the **generator** of the symmetry. (Think j_z .)

Lie groups tell how rotations, Lorentz boosts and changes of phase (unitary transformations), described respectively by SU(2), SU(2) and U(1), modify vectors within the vector space of the representation of the group. In many cases (spin, hypercharge, isospin), these rotations take place not in the 4-d space of SR, but in internal states of the system.

A representation is described by means of a **Lie algebra**, which is made up of the **generators** of a representation. Although the dimensions of representations vary, the number of generators of all those of a given group is always the same and equal to the **order**, the number of the group's parameters.

Some of the generators are diagonal and therefore correspond to eigenvectors, which can be used as a basis for the vector space of the representation. Such generators constitute the **Cartan subalgebra** of the Lie algebra and their number is the **rank** of the algebra. The eigenvalues are the physically measurable charges of the particle on which the force acts, and are equal in number to the dimension of the fundamental representation, which in the case of SU(n) or SO(n) consists of $n \times n$ matrices. In other words, an arbitrary vector will have components along different eigenvectors of the Cartan generators. Transformations of the group then will rotate one vector into another, changing the eigenvalues to which it corresponds, as will a rotation on J_3 . The "rotations" may be in real space, in Minkowski spacetime or in some internal space, such as spin or color space.

An arbitrary SU(n) group always has $n^2 - 1$ generators and is of rank $n - 1$, whereas an arbitrary SO(n) group has $\frac{n(n-1)}{2}$ generators.

One more time!

- A fundamental force – strong, weak, electromagnetic or, hopefully one day, gravitational – is described by a representation. We therefore describe a physical system by means of a Lie algebra.
- The eigenvectors of the diagonal Cartan generators describe the particles which are acted on by the force.
- The eigenvalues are the physically measurable charges (such as spin) of the particles.
- The number of charges fixes the dimension of the fundamental representation.
- The generators of the group, which parameterize the infinitesimal changes wrought by the transformation on the vector space of the representation, correspond to the gauge bosons, the force-carrying particles – photons, gluons, W and Z bosons – which act on the physical particles. There are two sorts.

- Cartan generators represent forces which act on all charged particles in the group by transferring energy-momentum.
- Non-Cartan generators not only can transfer energy-momentum but also can be combined to constitute creation/annihilation (ladder) operators which change the charge, interpreted as the number of particles.

But symmetry and QFT bring more than the understanding of forces and how they arise.

For instance, in $SU(2)$, there are three generators, J_{1-3} , of which only one is a Cartan generator, by convention J_3 . Its eigenvectors form a basis for the space. In the 2-dimensional representation, for $j = \frac{1}{2}$, the eigenvectors represent particles of positive and negative charge (spin), $j_z = \pm \frac{1}{2}$. The non-Cartan generators, J_1 and J_2 , are used to construct ladder (raising and lowering) operators which change the charge, in this case, the spin.

In the case of the strong force, the 3-dimensional fundamental representation of $SU(3)$ gives rise to the three eigenvalues which are the three quark colors. The gauge bosons describe how the colors are changed.

The Cartan generators (photon, Z^0) represent force-carrying particles which can modify, say, position or momentum, but without any change in charge. The non-Cartan generators (W bosons, gluons) are those which comprise the QM raising and lowering operators, so it is not a surprise that they can also bring about changes in charge.¹⁹⁷

6.14. So what's the big deal?

Quantum field theory is at the heart of modern physics, explaining QED, QCD, the electroweak force and much of the Standard Model. It is based on Lagrangian mechanics and takes into account the requirement of Special Relativity that the Lagrangian must be invariant under Lorentz transformations. As we have seen, this means it must be invariant under $SU(2)$, the covering group for both $SO(3)$ and the restricted Lorentz group Λ_+^\uparrow . This limits possible forms for the Lagrangian, as we saw in section 4.

Deal 1: Invariance under Lorentz transformation leads to the form of the Lagrangian for scalar, spinor and vector particles and fields.

Applying local gauge invariance to free-field equations for various numbers of spinors leads to an interaction term, often expressed as an additional term in the covariant derivative, and to the existence of additional massless spin-1 particles or fields (bosons). This is summarized in Table 9.

¹⁹⁷ Robinson, 116-117.

Field/ force	Spinors	Lie group	Order	Fundamental representation	Cartan ops	Eigen- vectors	Gauge fields/particles	Conserved charge
EM	1	U(1)	1		1	$2(\pm\frac{1}{2})$	1 massless photon	electric charge
Weak	2	SU(2)	3	$\tau = \frac{1}{2}\sigma, \sigma = \text{Pauli matrices}$	1	$3(0, \pm 1)$	3 massive (with Higgs) bosons (W and Z)	isospin
QCD	3	SU(3)	8	$\frac{1}{2}\lambda, \lambda = \text{Gell-Mann matrices}$	2	3 x 2 e- vals each	8 massless gluons	color charge

Table 9. Effects of SU(n) transformations on spinors.

Deal 2: Invariance under U(1) and SU(3) leads to EM and QCD.

In particular, invariance under a gauge transformation – meaning local changes in charges – requires the existence of a point-to-point connection which is identically the EM field that obeys Maxwell's equations.

U(1) and SU(3) symmetries thus give rise to photons and gluons, the massless vector “particles” of EM and strong interactions, respectively. In the case of the weak interactions, the W and Z particles concerned are most definitely not massless, so we must draw on “spontaneous” symmetry breaking¹⁹⁸ and the Higgs mechanism in order to bestow masses on those gauge particles (W^\pm, Z^0).

Deal 3: Invariance under SU(2) and U(1)_Y leads to the electroweak theory, complete with masses for the electron, W^\pm and Z^0 , with EM as an invariant subset which retains its U(1) symmetry and massless photon.

6.15. Construction of SU(3) multiplets – the Eightfold Way

SU(3) has $(n^2 - 1) = 8$ parameters and so 8 generators *in every representation*, whatever be the dimension of the representation.¹⁹⁹ In its *fundamental representation* of 3x3 matrices, the standard generators are taken to be $\frac{1}{2}$ the Gell-Mann matrices, $\lambda^i, i = 1 - 8$:

$$\begin{aligned}
\lambda^1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda^4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda^5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \lambda^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \\
\lambda^7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda^8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{6.66}$$

¹⁹⁸ I object to the word “spontaneous” here, which indicates the symmetry breaking is not brought about by an external agency, to be confusing to the point of being incorrect. But I have no better candidate.

¹⁹⁹ Robinson, 111-116.

Only two of the generators are diagonal and so Cartan generators, λ^3 and λ^8 , meaning that the **rank** of SU(3) is 2. All the generators are Hermitian.

The upper-left quadrangles of $\lambda_1 - \lambda_3$ are the Pauli matrices. So the first three matrices are a closed SU(2) subgroup of SU(3).

Being a 3-d representation, this one has three eigenvectors, each one with a pair of eigenvalues, which in turn means three basis vectors.

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

(We could perhaps better call them u, d and s without waiting, but....) Pairs of eigenvalues for λ^3 and λ^8 can be represented by so-called **weight vectors**, column doublets in this case, of pairs of eigenvalues for a given eigenvector. They are

$$t_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2\sqrt{3}} \end{pmatrix}, \quad t_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{1}{2\sqrt{3}} \end{pmatrix} \quad \text{and} \quad t_3 = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}, \quad (6.67)$$

corresponding to the basis vectors v_1 , v_2 and v_3 .

Remembering the happy results of defining ladder operators in SU(2) and SO(3), let's try this for SU(3). First, define the *generators* by

$$F_i = \frac{1}{2}\lambda_i. \quad ^{200} \quad (6.68)$$

Then take what is called the spherical representation of the F operators.²⁰¹

$$\begin{aligned} T_{\pm} &= F_1 \pm iF_2, & T_3 &= F_3, \\ V_{\pm} &= F_4 \pm iF_5, \\ U_{\pm} &= F_6 \pm iF_7, & Y &= \frac{2}{\sqrt{3}}F_8. \end{aligned} \quad (6.69)$$

One can work out the commutation properties of these operators, including definitions of U_3 and V_3 by analogy with T_3 .²⁰²

$$\begin{aligned} [T_3, T_{\pm}] &= \pm T_{\pm}, & [T_+, T_-] &= 2T_3, \\ [T_3, U_{\pm}] &= \mp \frac{1}{2}U_{\pm}, & [U_+, U_-] &= \frac{3}{2}Y - T_3 := 2U_3, \\ [T_3, V_{\pm}] &= \pm \frac{1}{2}V_{\pm}, & [V_+, V_-] &= \frac{3}{2}Y + T_3 := 2V_3. \end{aligned} \quad (6.70)$$

We have already noted that the T's form a closed SU(2) subset. The first line of (6.70) confirms that the operators $T_i (i = +, -, 3)$ are consistent with the Lie algebra for SU(2) of (3.12). The same thing can be shown to be true for the U and V operators:

$$\begin{aligned} [U_3, U_{\pm}] &= \pm U_{\pm}, & [U_+, U_-] &:= 2U_3 \\ [V_3, V_{\pm}] &= \pm V_{\pm}, & [V_+, V_-] &:= 2V_3. \end{aligned}$$

All three subalgebras of SU(3) – call them **T-spin**, **U-spin** and **V-spin** – are isomorphic to SU(2). (The whole thing is sometimes referred to as **F-spin**.) Better yet, T_{\pm} , U_{\pm} and V_{\pm} are ladder operators. Call our two Cartan

²⁰⁰ Apologies for not distinguishing carefully between superscripts and subscripts in this context.

²⁰¹ I am ignoring the hats which should adorn the F and λ operators.

²⁰² Greiner, 203-204, 212-215.

operators T_3 and Y (isospin 3-component and strong hypercharge²⁰³), and our eigenvectors $|t_3, y\rangle$, so that (denoting eigenvectors by lower case)

$$T_3|t_3, y\rangle = t_3|t_3, y\rangle, \quad Y|t_3, y\rangle = y|t_3, y\rangle.$$

Then

$$(T_3 V_{\pm} - V_{\pm} T_3)|t_3, y\rangle = (T_3 V_{\pm} - t_3 V_{\pm})|t_3, y\rangle = \pm \frac{1}{2} V_{\pm}|t_3, y\rangle,$$

by (6.70), which tells us that

$$T_3 V_{\pm}|t_3, y\rangle = (t_3 \pm \frac{1}{2}) V_{\pm}|t_3, y\rangle,$$

i.e., V_{\pm} is a couple of ladder operators which raise/lower t_3 by $\frac{1}{2}$.²⁰⁴ In a similar manner, one can show that

U_{\pm} lowers/raises t_3 by $\frac{1}{2}$, i.e., in the opposite direction to V_{\pm} . Then

$$[Y, V_{\pm}] = \pm V_{\pm}$$

may be used to show that

$$U|(V_{\pm}|t_3, y\rangle) = (y \pm 1)(V_{\pm}|t_3, y\rangle),$$

so V_{\pm} raises/lowers y by 1. The same is true for U_{\pm} .

So, more succinctly, including as-yet uncalculated normalization constants:

$$T_{\pm}|t_3, y\rangle = \alpha|t_3 \pm 1, y\rangle,$$

$$U_{\pm}|t_3, y\rangle = \beta|t_3 \mp \frac{1}{2}, y \pm 1\rangle,$$

$$V_{\pm}|t_3, y\rangle = \gamma|t_3 \pm \frac{1}{2}, y \pm 1\rangle. \quad (6.71)$$

From these equations, we can diagram the actions of the ladder operators in the $t_3 - Y$ plane.

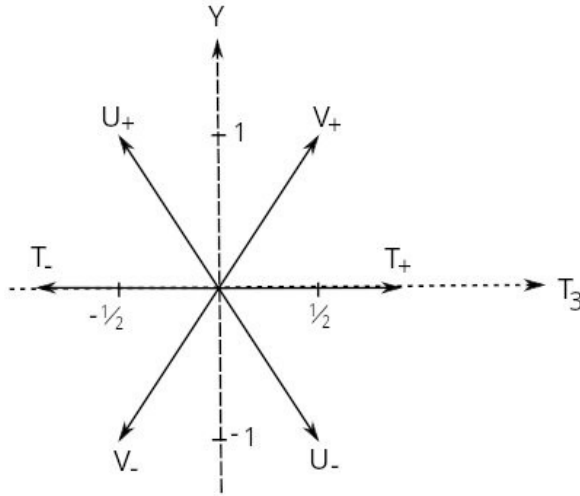


Figure 5. Action of ladder operators in $t_3 - Y$ plane of the fundamental representation of $SU(3)$.

Now we can understand some interesting points.

The $SU(3)$ algebra includes the three T , U and V subalgebras, each of which is isomorphic to the algebra of

²⁰³ The strong hypercharge is not the weak hypercharge we met in section 6.12.

²⁰⁴ We are also not worrying about normalization factors here.

SU(2).²⁰⁵ So SU(3) multiplets can be constructed from coupled T, U and V multiplets.

Since U_3 and V_3 are linear combinations of T_3 and Y by (6.70), all four may be simultaneously diagonalized, with eigenvalues

$$t_3, y, U_3 = \frac{1}{2}(\frac{3}{2}Y - T_3) \text{ and } V_3 = \frac{1}{2}(\frac{3}{2}Y + T_3).$$

The ladder operators shift the SU(3) multiplet states as in Figure 5, where the operator end-points are on a regular hexagon. So an SU(3) multiplet will be constructed from a T multiplet parallel to the T_3 axis and U and V multiplets parallel to the U and V axes in Figure 5. They are coupled by the non-zero commutation relations among them, e.g., $[T_+, V_-] = -U_-$.

The symmetry among the T, U and V algebras requires symmetry with respect to the $T_3 = 0$, $U_3 = 0 = \frac{3}{2}(Y - T_3)$ and $V_3 = 0 = \frac{3}{2}(Y + T_3)$ with an angle of 120° between any two. Normally, the SU(3) multiplet is centered on the origin $T_3 = Y = 0$. Further symmetry considerations show that an SU(3) multiplet must be a *regular triangular or hexagonal structure*.

Within the F-spin algebra, the T, U and V-spin subalgebras appear symmetrically, so the simplest non-trivial SU(3) multiplet will contain doublets of all three.²⁰⁶ This leads us to the diagrams of Figure 6. Each diagram contains an isodoublet $T = \frac{1}{2}$ and an isosinglet $T = 0$. For the representation $[3]$, the isodoublet is composed of the states

$$\psi_1 = |\frac{1}{2}, y\rangle \quad \text{and} \quad \psi_2 = |-\frac{1}{2}, y\rangle$$

and the isosinglet

$$\psi_3 = |0, y\rangle.$$

The values of the hypercharge Y may now be determined. From the diagram, ψ_2 and ψ_3 form the U-spin doublet, so the singlet must be ψ_1 .

$$U_3\psi_1 = 0.$$

Then the definitions of U_3 and V_3 may be used to find the eigenvalues of Y:

$$Y\psi_1 = \frac{1}{3}\psi_1, \quad Y\psi_2 = \frac{1}{3}\psi_2, \quad Y\psi_3 = -\frac{2}{3}\psi_3$$

and the corresponding negative values for the antiparticles. The charge is given by the Gell-Mann-Nishijima formula

$$q = t_3 + \frac{1}{2}y. \tag{6.72}$$

The charges are then

$$Q\psi_1 = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} = \frac{2}{3}, \quad Q\psi_2 = -\frac{1}{2} + \frac{1}{2} \cdot \frac{1}{3} = -\frac{1}{3}, \quad Q\psi_3 = 0 + \frac{1}{2} \cdot (-\frac{2}{3}) = -\frac{1}{3}.$$

So the representation is of three particles with fractional charges – the **quarks** (i.e., the lightest three quarks) u, d and s (ψ_1 , ψ_2 and ψ_3).

²⁰⁵ Greiner, 212.

²⁰⁶ This overview of the subject must be qualified as “quick and dirty”. For a more complete version, see Greiner, 231-3.

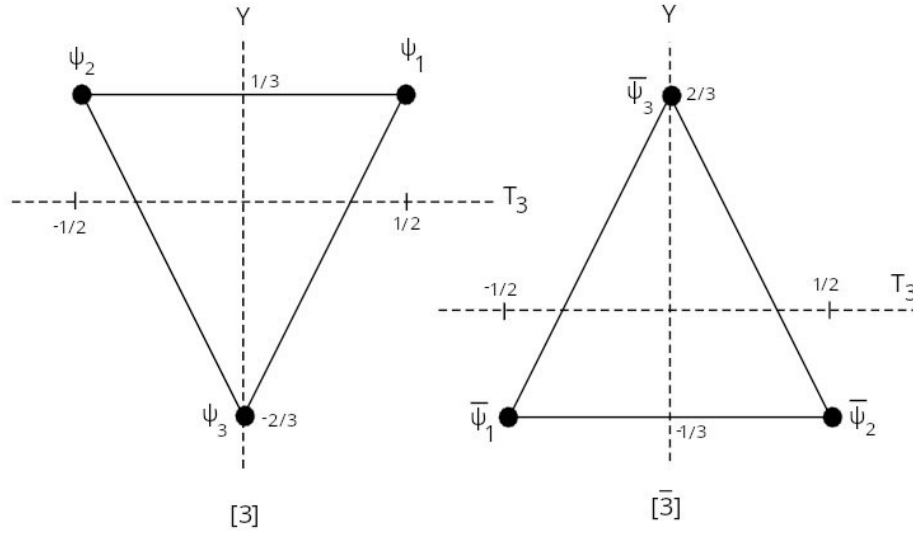


Figure 6. The smallest non-trivial representation of SU(3)

The values for $[\bar{3}]$ are the negative of these, those of the corresponding antiquarks. It is clear from the eigenvalues that the $[3]$ diagram represents particles, the $[\bar{3}]$ diagram, antiparticles. In support of this, the Gell-mann-Nishijima formula shows that

$$Q\psi_\nu = q_\nu\psi_\nu \implies Q\bar{\psi}_\nu = -q_\nu\psi_\nu.$$

They are the *two fundamental representations of SU(3)*. “In principle the construction of all irreducible representations of the *[sic]* SU(3) requires only one of the two fundamental representations $[3]$ and $[\bar{3}]$ However, for reasons of physics, one needs both fundamental representations, because quarks (represented by $[3]$) and antiquarks (represented by $[\bar{3}]$) differ by their baryon number ($B=1/3$ for quarks, $B=-1/3$ for antiquarks) and charge.”²⁰⁷ With these baryon numbers, the Gell-mann-Nishijima relation works out consistently for the charges.

Remember that SU(2) representations (multiplets) can be constructed from the fact that each corresponds to a value $j = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots$ with dimension $(2j + 1)$ and states $|jm\rangle = +j, \dots -j$. Alternatively, they could be constructed by successive coupling of the fundamental doublets $|j = \frac{1}{2}, m = \frac{1}{2}\rangle = \frac{1}{2}$. This latter method will be used to construct SU(3) multiplets from the two fundamental multiplets $[3]$ and $[\bar{3}]$.

An SU(3) multiplet expressed as a combination of p quarks and q antiquarks is denoted $D(p,q)$.²⁰⁸ A state deemed of maximal weight is constructed starting with p quarks of maximal weight state $|\frac{1}{2}, \frac{1}{3}\rangle$ and q antiquarks of maximal weight state $|\frac{1}{2}, -\frac{1}{3}\rangle$, so the state has the following values of T_3 and Y for this maximal state:

$$(T_3)_{max} = \frac{p+q}{2} \quad \text{and} \quad (Y)_{max} = \frac{p-q}{3}. \quad (6.73)$$

This state will be the farthest to the right on the diagram. Other states are then constructed from it by use of the ladder (shift) operators V_\pm and U_\pm . Then the states $[3]$ and $[\bar{3}]$ of Figure 6 are the single-quark state $D(1,0)$ and the single antiquark state $D(0,1)$.

²⁰⁷ Greiner, 242.

²⁰⁸ This is, again, a quick version of what is explained in more detail by Greiner, 241-6.

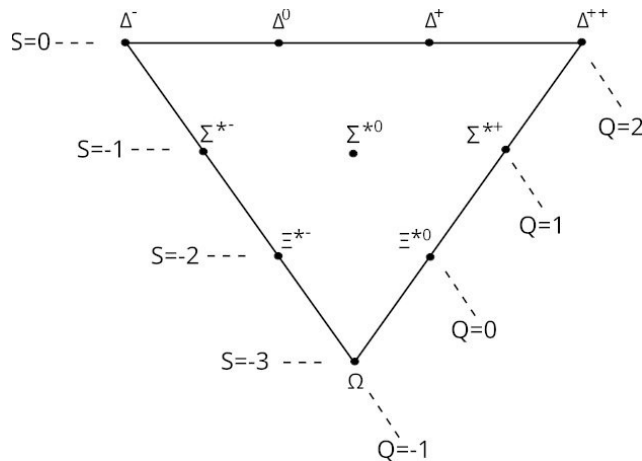


Figure 7. The baryon decuplet.

Take the example of the baryon decuplet of Figure 7. Using the $D(3,0)$ representation, we know from (6.73) that the maximum point is at $(\frac{3}{2}, 1)$ with charge $q=+2$, from (6.72). It therefore may correspond to the Δ^{++} , composed of three up quarks, uuu . Now use V_- to construct the other three baryons on the right hand side, at each step reducing t_3 and Y by $\frac{1}{2}$ and 1, respectively. This is equivalent to transforming an up quark to a strange one at each step and gives us particles successively at $(1, 0)$, $(\frac{1}{2}, -1)$ and $(0, -2)$. From each of these three, we may use U_+ change the coordinates each time by $(-\frac{2}{3}, 1)$, changing a strange quark to a down one each time, and this will generate the other particles. Or we could use T_- , to get the equivalent diagram by changing up quarks one at a time to down ones.

These diagrams are representative of what is called the **Eightfold Way** of particle physics.

The meson octet of Figure 7 is the $D(1,1)=[8]$ representations formed of a quark and antiquark. These are the correct components of baryons and mesons.

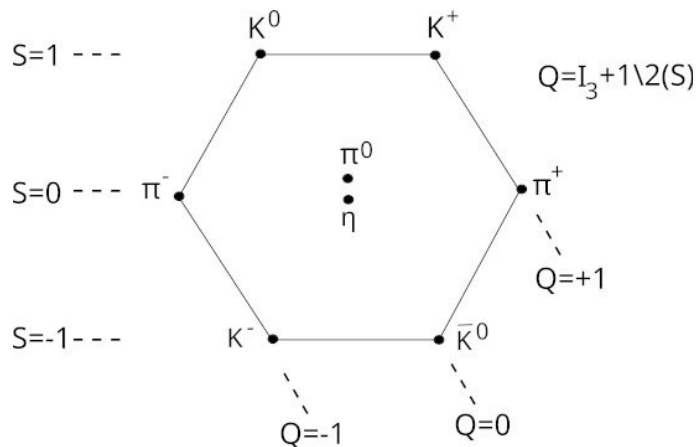


Figure 8. The meson octet.

The baryon octet of Figure 9 is more complicated and necessitates a digression.

QFT proves²⁰⁹ that bosons must have symmetric wave functions and fermions, antisymmetric. This of course leads to the Exclusion Principle – on which the existence of chemistry and so everything else is based. The symmetry of a quark wave function has four parts – space, spin, flavor and color. The law requiring hadrons to be color neutral is in fact an example of a more fundamental law which states that “every naturally occurring particle

²⁰⁹ Where?

is a color singlet.²¹⁰ So the color part is antisymmetric, which means the rest must be symmetric. The ground state we are assuming must be spatially symmetric, so the product of the spin and flavor parts must be symmetric. The 27 (3^3) possible combinations of three quarks can be grouped into symmetric, antisymmetric or mixed states – a completely symmetric decuplet, a completely asymmetric singlet and two octets of mixed symmetry (antisymmetric in either particles 1 and 2 or 2 and 3).²¹¹ In group-theory speak, we can see this as the decomposition of the product of three fundamental states

$$3 \otimes 3 \otimes 3 = 10 \oplus 8 \oplus 8 \oplus 1.$$

Now the baryon decuplet is composed of completely symmetric states, since, e.g., what is called Δ^+ in Figure 7 can be considered as a symmetric combination $(uud + udu + duu)\sqrt{3}$, and so on for the other states. So this decuplet must go along with a symmetric spin state. The baryon octet is asymmetric in two quarks, say 1 and 2. But the corner quarks would be uuu , ddd and sss , all highly symmetric. So they must go. Such is a somewhat hand-wavy argument for why the baryon octet looks like a triply truncated triangle.

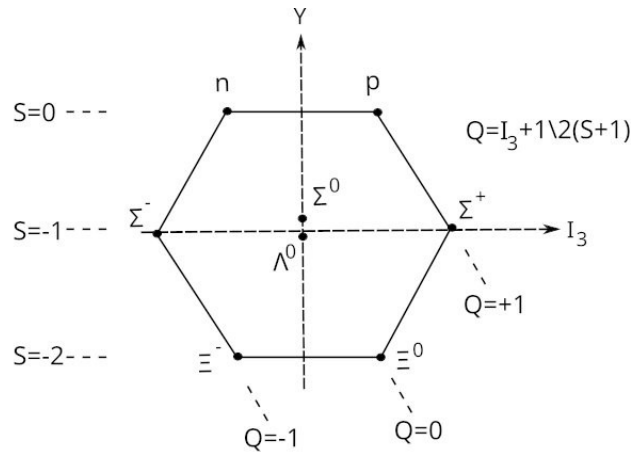


Figure 9. The baryon octet.

Table 9 shows the most common particles, with their spin, quark composition and charge.

An important word on hypercharge. We have called the hypercharge Y and used it and the 3rd component of the isospin in the Gell-Mann-Nishijima formula for the charge:

$$Q = \frac{1}{2}Y + I_3. \quad (6.60)$$

This is the **strong hypercharge** and has nothing to do with the weak hypercharge we will discuss in section 6.12. It is the sum of the baryon (atomic) number A and the strangeness S and one often sees eightfold-way diagrams in terms of S . It is also twice the average charge, Q , since the average of I_3 is zero. This is an old concept, developed in the 1960s and today the strong hypercharge is best represented as

$$Y = \frac{1}{3}n_u + \frac{1}{3}n_d - \frac{2}{3}n_s + \frac{4}{3}n_c - \frac{2}{3}n_b + \frac{4}{3}n_t, \quad (6.74)$$

in terms of the number of up, down, strange, charmed, bottom and top quarks.

²¹⁰ Griffiths, 187.

²¹¹ Griffiths, 185-6.

Baryons	Spin	Particle	Quark content	Charge
	1/2	p	uud	1
		n	udd	0
		Λ	uds	0
		Σ^+	uus	1
		Σ^0	uds	0
		Σ^-	dds	-1
		Ξ^0	uss	0
		Ξ^-	dss	-1
		Λ_c^+	udc	1
	3/2	Δ	uuu, uud, udd, ddd	2, 1, 0, -1
		Σ^*	uus, uds, dds	1, 0, -1
		Ξ^*	uss, dss	0, -1
		Ω^-	sss	-1

Table 10. Properties of most common baryons

	Spin	Particle	Quark content	Charge
Pseudoscalar mesons	0	π^\pm	$u\bar{d}, d\bar{u}$	1, -1
		π^0	$(u\bar{u} - d\bar{d})/\sqrt{2}$	0
		K^\pm	$u\bar{s}, s\bar{u}$	1, -1
		K^0, \bar{K}^0	$d\bar{s}, s\bar{d}$	0
		η	$(u\bar{u} + d\bar{d} - 2s\bar{s})/\sqrt{6}$	0
		η'	$(u\bar{u} + d\bar{d} + s\bar{s})/\sqrt{3}$	0
		D^\pm	$c\bar{d}, d\bar{c}$	1, -1
		D^0, \bar{D}^0	$c\bar{u}, u\bar{c}$	0
		D_s^\pm	$c\bar{s}, s\bar{c}$	1, -1
		B^\pm	$u\bar{b}, b\bar{u}$	1, -1
		B^0, \bar{B}^0	$d\bar{b}, b\bar{d}$	0
Vector mesons	1	ρ	$u\bar{d}, (uu - d\bar{d})/\sqrt{2}, d\bar{u}$	1, 0, -1
		K^*	$u\bar{s}, d\bar{s}, s\bar{d}, s\bar{u}$	1, 0, -1
		ω	$(uu + d\bar{d})/\sqrt{2}$	0
		ψ	$c\bar{c}$	0
		D^*	$c\bar{d}, c\bar{u}, u\bar{c}, d\bar{c}$	1, 0, -1
		Υ	$b\bar{b}$	0

Table 11. Properties of common mesons

7. Scattering

Before studying scattering, meaning transitions from one state to another, we need another picture.

7.1. The interaction picture

Back for the moment, to the Schrödinger picture (SP). From the equation (6.24) for the total Lagrangian density, we can write the Hamiltonian (total, not density) as

$$H := H^S = H^{\frac{1}{2},1} = \int (\pi_r \dot{\phi}^r - \mathcal{L}_0^1 - \mathcal{L}_0^{\frac{1}{2}}) d^3x - \int \mathcal{L}_1^{\frac{1}{2},1} d^3x \quad (7.1)$$

$$(H_0^S = H_0) \quad (H_I^S)$$

which is just

$$H = H_0 + H_I^S, \quad (7.2)$$

the free part plus the interaction part. In this equation, ϕ^r represents any quantum field. The unitary time evolution operator is then

$$U(t) = e^{-i(H_0 + H_I^S)t} = U_0(t)U_I(t).$$

Let's go to a special Heisenberg picture (HP) which is called the **interaction picture** (IP), one where the unitary transformation uses only the free part of the Hamiltonian.²¹²

$$U_0 = e^{-iH_0 t}. \quad (7.3)$$

Since we have defined

$$L^H(t) = \mathbf{U}^\dagger(t) |L^S| \mathbf{U}(t), \quad (5.44)$$

the transformation from the SP to the IP is defined by

$$|\Psi\rangle_I = U_0^\dagger |\Psi\rangle_S$$

as in equation (5.45) for the Heisenberg picture, so that

$$\mathcal{O}^I = U_0^\dagger \mathcal{O}^S U_0. \quad (7.4)$$

Then, because H_0 commutes with itself, the free part of the Hamiltonian $H_0 = H_0^S$

$$H_0^I = U_0^\dagger H_0^S U_0 = H_0, \quad (7.5)$$

but the same is not generally true of the interaction part, H_I^I . So from (7.2) and (7.5), the **interaction-picture Hamiltonian** is

$$H^I = H_0 + H_I^I. \quad (7.6)$$

This notation can be confusing:

- The subscript I refers to the *interaction energy* of the EM field and the Dirac electron;
- the superscript I indicates the *interaction picture*.

Remember that

$$H_0 = H_0^S = H_0^I. \quad (7.7)$$

The **equation of motion** for an operator in the IP is then

$$\frac{d\mathcal{O}^I}{dt} = \frac{d}{dt}(U_0^\dagger \mathcal{O}^S U_0) = \dots = -i[\mathcal{O}^I, H_0] + \frac{\partial \mathcal{O}^I}{\partial t}, \quad (7.8)$$

where the last quantity

$$\frac{\partial \mathcal{O}^I}{\partial t} := U_0^\dagger \frac{\partial \mathcal{O}^S}{\partial t} U_0 \quad (7.9)$$

is for our purposes always 0. It can also be shown that²¹³

$$i \frac{d}{dt} |\Psi\rangle_I = H_I^I |\Psi\rangle_I. \quad (7.10)$$

²¹² This is the version of Klauber, of Lancaster and Blundell and of Schwichtenber PS, but he uses a different method in NNQFT.

²¹³ Klauber, Solution, problem 7.8, 7-5.

The equations of motion for states in the IP depend only on the interaction part of the Hamiltonian.

This is worth repeating:

- The equations of motion (7.8) for *operators* in the IP depend only on the *free* part of the Hamiltonian.
- The equations of motion (7.10) for *states* in the IP depend only on the *interaction* part of the Hamiltonian.

The first point means that all the results we have studied above for free fields in the HP apply also to operators in the IP. Since that applies to HP field operators like ϕ , ψ or A^μ . this means that we can use the K-G equation for scalar fields, the Dirac equation for spinors and the Maxwell equation for photons. Not only that, but we can use the free-field

- operator solutions,
- creation, annihilation and number operators,
- observables operators, and
- Feynman propagators.²¹⁴

There remains “only” to solve equation (7.10) for the states.

The following table summarizes the situation for the three picturex.

Picture	Schrödinger	Heisenberg	Interaction
states	$\Psi(t)$	-	H_1
operators	-	$\hat{O}_H(t)$	H_0

For expectation values.

$$\frac{d\langle O \rangle}{dt} = {}_I\langle \Psi | \left(-i[\mathcal{O}^I, H^I] + \frac{\partial \mathcal{O}^I}{\partial t} \right) | \Psi \rangle_I \quad (7.11)$$

This is the same equation as that for the expectation value in the SP, even though operators in that picture do not vary in time.

Note that both methods of deriving the IP agree in their results for the time-evolution dependence of the operators and state vectors.

7.2. The S-matrix and ordering

We'll use the word “scattering” rather largely.

Consider the implications of the division of the Hamiltonian into a free part and an interaction part. To calculate the probability that a particular reaction or scattering process takes place, we want the real amplitude between the initial and final states

$$A = {}_r\langle q_1 q_2 | p_2 p_1 \rangle_r,$$

corresponding to the entire Hamiltonian, free plus interaction. But the states we see around us are free ones, corresponding to H_0 . They are the same as the true, or real, states at $t = \pm\infty$, but not necessarily (or usually) at any other time.

²¹⁴ Klauber, 194.

$$\lim_{t_1 \rightarrow -\infty, t_2 \rightarrow +\infty} |p_2(t_2)p_1(t_1)\rangle_r = |p_2p_1\rangle_0. \quad (7.12)$$

Rather than defining one of the two types of states in terms of the other, we define the real amplitude in terms of the free states and the **S-matrix**.²¹⁵

$$A = {}_r\langle q_1q_2|p_2p_1\rangle_r := {}_0\langle q_1q_2|\hat{S}|p_2p_1\rangle_0. \quad (7.13)$$

So to calculate scattering amplitudes, we need the simple states and a way to calculate \hat{S} . We will use the eigenstates of H_0 and will henceforth drop the 0 subscript on the bras and kets.²¹⁶

Using the *interaction picture*, the initial and final states can be taken as free states, so we can construct them from the vacuum state by using the creation and annihilation operators we have already studied.

As for the time-evolution operator, we know that the states in the IP evolve by the interactive part of the Hamiltonian, and so

$$i\frac{d}{dt}|\Psi(t)\rangle^I = H_I^I|\Psi(t)\rangle^I,$$

which leads to the equation for the unitary time evolution operator $U(t_2, t_1)$

$$i\frac{d}{dt_2}\hat{U}(t_2, t_1) = H_I^I(t_2)\hat{U}(t_2, t_1). \quad (7.14)$$

Unfortunately, since $H_I^I(t)$ may not commute with itself at different times, the “simple” solution

$$\hat{U}(t_2, t_1) = e^{-i\int_{t_1}^{t_2} dt \hat{H}_I^I(t)}$$

is shown to be incorrect when this is expanded in a power series, as we must do in order to solve it.²¹⁷ Instead, we must use the time-ordered solution

$$\hat{U}(t_2, t_1) = T[e^{-i\int_{t_1}^{t_2} dt \hat{H}_I^I(t)}], \quad (7.15)$$

where the **time-ordering operator** (It's not really an operator, but a specification of method) arranges the elements of each term of the power-series expansion of the formula from right to left in order of increasing time. So now we have learned of two methods of ordering operators:

- **normal ordering** puts creation operators on the left, annihilation operators on the right;
- **time ordering** puts operators in order of increasing time from right to left, earliest to the right, latest to the left.

Merging (7.13) and (7.15) gives us the **S-matrix equation**

$$\hat{S} = T[e^{-i\int_{t_1}^{t_2} dt \hat{H}_I^I(t)}]. \quad (7.16)$$

7.3. Contractions and Wick's theorem

Equation (7.16) for the S-matrix uses time ordering, but normal ordering is easier to work with, as the annihilation operators on the right destroy a vacuum state, as do creation operators on the left, thus simplifying the equation. Here is where we need contractions, or Wick contractions. A field operator such as those we have found in, for instance equations (5.58) are sums of creation and annihilation parts. Note that we are now using the opposite

²¹⁵ One more genial idea of John Wheeler.

²¹⁶ I rather like this introduction to the S matrix, as given by Blundell and Lancaster, 166-7, on which this one is based.

²¹⁷ Schwichtenberg, NNQFT, 388-9.

notation to that of section 5²¹⁸, so A^+ now denotes the creation part and A^- the annihilation, which we therefore want on the right in normal ordering. We could consider this for two operators \hat{A} and \hat{B} such that

$$\hat{A}\hat{B} = (A^+ + A^-)(B^+ + B^-).$$

where this pair of operators differs from normal ordering $N[\hat{A}\hat{B}]$, often delimited by colons as $:\hat{A}\hat{B}:$, by the commutator of A and B,

$$\hat{A}\hat{B} = N[\hat{A}\hat{B}] + [\hat{A}^-, \hat{B}^+].$$

Or we could consider the same operator evaluated at different spacetime moments, using the notation of (5.59).

$$\hat{\phi}(x) = \hat{\phi}^+(x) + \hat{\phi}^-(x),$$

where the + and – superscripts now again refer to creation and annihilation terms, respectively.²¹⁹ Then

$$\hat{\phi}(x)\hat{\phi}(y) = N[\hat{\phi}(x)\hat{\phi}(y)] + [\hat{\phi}^-(x), \hat{\phi}^+(y)]. \quad (7.17)$$

In either case, this tells us that we can normal-order operators as long as we include the appropriate commutator terms. It looks interesting because a commutator either is zero or has a value like some multiple of \hbar and maybe a delta function. In other words, it's just a number, perhaps complex, called a c-number. Since time ordering of the same two operators gives

$$T[\hat{A}(x)\hat{B}(y)] = \begin{cases} \hat{A}(x)\hat{B}(y), & x^0 > y^0 < 0 \\ \hat{B}(y)\hat{A}(x), & x^0 < y^0 \geq 0 \end{cases}$$

($x^0 = t$) then the difference between the time-ordered pair and the normal-ordered one

$$\underline{AB} = T[\hat{A}(x)\hat{B}(y)] - N[\hat{A}(x)\hat{B}(y)] = \begin{cases} [\hat{A}^-(x), \hat{B}^+(y)], & x^0 > y^0 \\ [\hat{B}^-(y), \hat{A}^+(x)], & x^0 < y^0 \end{cases} \quad (7.18)$$

is so useful that it is defined to be the **contraction** or the **Wick contraction**. The first term in the commutator is time-ordered, since it is originates from the time-ordered sequence, with the creation term on the right, as that is the one left over after subtraction of the normal-ordered sequence. In the case of two operators, we can find their contraction by calculating the result of their time-ordered product between $\langle 0|$ and $|0\rangle$, since the normal-ordered part will give zero and the contraction, by (7.18), is simply a commutator, a complex number.²²⁰

$$\underline{\phi(x)\phi(y)} = \langle 0|T\phi(x)\phi(y)|0\rangle. \quad (7.19)$$

Since fermions do not commute but anticommute, a factor of -1 is introduced for each contraction of a pair of fermions.

(Note on notation: A contraction symbol is most often like a square bracket lying on its side, either above or below the characters involved, with its two ends pointing at the operators to be contracted. Not finding this in TexMaths for LibreOffice, I am obliged to settle for an underline, which will not always be adequate, in which case I will be forced to fall back on words. Another possibility, used by Wikipedia, is with bullet superscripts, where pairs are linked by an identical number of bullets: E.g., $N[\hat{A}^\bullet \hat{B}^{\bullet\bullet} \hat{C}^{\bullet\bullet} \hat{D}^\bullet]$, where AD is a contracted pair and BC is another. Then $\underline{AB} \equiv A^\bullet B^\bullet$.)

Equation (7.18) is a special case of **Wick's theorem**, for a pair of operators. The theorem states, more generally, that a time-ordered product of creation and annihilation operators can be expressed as the normal ordered

²¹⁸ Klauber's notation.

²¹⁹ Be aware that some authors, such as Klauber or Peskin and Schroeder, use the + and – superscripts in the opposite sense, with ϕ^+ being the annihilation operator. *Comprene qui pourra.*

²²⁰ Lancaster and Blundell, 172; Schwichtenberg, QFT, 427.

product of the fields plus the *normal-ordered sum* of all possible contractions of the fields.²²¹

For an arbitrary number of fields, this is:²²²

$$T\{\phi(x_1)\phi(x_2)\dots\phi(x_n)\} = N\{\phi(x_1)\phi(x_2)\dots\phi(x_n) + \text{all possible contractions}\}.$$

Note that the entire expression on the right is normal-ordered, including the contractions; i.e., each term is normal ordered. The theorem can be proved by induction.²²³ The really useful fact here is that when operating on the vacuum, the normal-ordered term will annihilate the state vector, on the right or on the left ($\langle 0|A^+ = 0$), and so gives zero. So nothing is left but the contractions. *Only completely contracted terms will give non-zero results in a transition amplitude between initial and final vacuum states.*

An example of a more complicated case would be that of four operators, in which individual contractions of operators as well as of pairs of contractions must be taken into account.

$$\begin{aligned} T[\hat{A}\hat{B}\hat{C}\hat{D}] &= N[\hat{A}\hat{B}\hat{C}\hat{D}] + N[\hat{A}\hat{\bullet}\hat{B}\hat{\bullet}\hat{C}\hat{D}] + N[\hat{A}\hat{\bullet}\hat{B}\hat{C}\hat{\bullet}\hat{D}] + N[\hat{A}\hat{\bullet}\hat{B}\hat{C}\hat{D}\hat{\bullet}] \\ &+ N[\hat{A}\hat{B}\hat{\bullet}\hat{C}\hat{\bullet}\hat{D}] + N[\hat{A}\hat{B}\hat{\bullet}\hat{C}\hat{D}\hat{\bullet}] + N[\hat{A}\hat{B}\hat{C}\hat{\bullet}\hat{D}\hat{\bullet}] \\ &+ N[\hat{A}\hat{\bullet}\hat{B}\hat{\bullet}\hat{C}\hat{\bullet}\hat{D}\hat{\bullet}] + N[\hat{A}\hat{\bullet}\hat{B}\hat{\bullet}\hat{C}\hat{D}\hat{\bullet}\hat{\bullet}] + N[\hat{A}\hat{\bullet}\hat{B}\hat{\bullet}\hat{C}\hat{\bullet}\hat{D}\hat{\bullet}\hat{\bullet}]. \end{aligned}$$

If A and B are free scalar fields $\phi(x)$ and $\phi^\dagger(y)$, then the quantity

$$\langle 0|T\{\phi(x)\phi^\dagger(y)\}|0\rangle = \langle 0|\phi(x)\phi^\dagger(y)|0\rangle \equiv i\Delta_F(x-y)$$

is the **Feynman propagator** of equation (5.73).

Now we can prepare a summary of important contractions for later use.²²⁴ (Beware of slightly different notations or factors of i among authors, especially in the propagator terms.)

$$\underline{\phi(x), \phi(y)} = D_F(x, y), \quad \text{the Feynman propagator} \quad (7.20)$$

$$\underline{\phi(x)|1_k\rangle} = \underline{\phi(x)a_k^\dagger|0\rangle} = e^{-i(kx)}|0\rangle \quad (7.21)$$

$$\langle 1_k|\phi(x) = \langle 0|a_k\phi(x) = \langle 0|e^{i(kx)} \quad (7.22)$$

$$\langle 1_q|\dots|1_k\rangle = \sqrt{2\omega_q}\sqrt{2\omega_k}(2\pi)^3\delta(\vec{q}-\vec{k}) \quad (7.23)$$

$$|1_q, 1_k\rangle = 0 \quad (7.24)$$

$$\langle 1_q, 1_k| = 0 \quad (7.25)$$

While we're at it, let's write the contraction formula (7.18) in terms of our ϕ variables:

$$\underline{\phi(x), \phi(y)} = \begin{cases} [\phi_-(x), \phi_+(y)], & x^0 > y^0 < 0 \\ [\phi_-(y), \phi_+(x)], & x^0 < y^0 \geq 0 \end{cases} \quad (7.26)$$

Last but not least, Wick's theorem

$$T\{\phi(x_1)\phi(x_2)\dots\phi(x_n)\} = N\{\phi(x_1)\phi(x_2)\dots\phi(x_n) + \text{all possible contractions}\} \quad (7.27)$$

221 Schwichtenberg, op. cit., 428.

222 Lancaster and Blundell, 172; Schwichtenberg, QFT, 428.

223 Peskin and Schroeder, 89-90.

224 Schwichtenberg, QFT, 430-6, 461.

7.4. Putting it all together – the Dyson series

So, now, we have Lagrangians for free scalars, spinors and vectors. Requiring (local) gauge invariance under our favorite Lie groups gives us minimal-coupling interactions, and we can use those to build the S-matrix for particle interactions plus creation and destruction of particles or energy quanta. And we can use operator commutation relations to simplify the search for solutions. So what does that give us?

The time for even the limited amount of detail we have been giving is past. We must accelerate. We will merely summarize the rest of this section.²²⁵

We want to

- start from nothing, the vacuum state,
- build some initial and final states, and
- investigate the ways in which one can transform into the other.

Starting from the transition amplitude of equation (7.13) and based on the vacuum state, we want to consider

$$\langle 0|S|0\rangle,$$

where now S must not only create the necessary particles to give us our initial and final states but also represent the time evolution of the initial state into the final one. The classic example used is the ϕ^4 theory with scalar Lagrangian (6.44), i.e.,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^*(\partial^\mu \phi) + \frac{1}{2}\mu^2(\phi^* \phi) - \frac{1}{4}\lambda^2(\phi^* \phi)^2, \quad (6.44)$$

which leads to the Hamiltonian

$$H = \int d^3x \left(\frac{1}{2}(\partial_\mu \phi)^2 + \frac{1}{2}m^2\phi^2 + \frac{1}{4!}\lambda\phi^4 \right) = H_0 + H_I, \quad (7.28)$$

containing a “free” and an interaction part.²²⁶ Hopefully, λ is a small quantity so we can ignore higher orders of it in our calculations. We know from equation (7.16) that we must use the time-ordered evolution operator so the scattering operator, the S-matrix, is

$$\hat{S} = T[e^{-i \int_{-\infty}^{\infty} dt' \hat{H}_I(t')}],$$

using the interaction part of the Hamiltonian.²²⁷ This can be expanded to give the transition amplitudes

$$\begin{aligned} A(i \rightarrow f) &= \langle f|\hat{S}|i\rangle \\ &= \langle f|i\rangle - i\langle f|\int_{-\infty}^{\infty} dt_1 H_I(t_1)|i\rangle \\ &\quad - \frac{1}{2!}\langle f|T\left(\int_{-\infty}^{\infty} dt_1 H_I(t_1)\right)\left(\int_{-\infty}^{\infty} dt_2 H_I(t_2)\right)|i\rangle + \dots \\ &= A^{(0)} + A^{(1)} + A^{(2)} + \dots \end{aligned} \quad (7.29)$$

This is the **Dyson series**. Note that we are in the interaction picture (IP), so we use the interaction part of the Hamiltonian, as was calculated in equation (7.14). Note also the presence of the time-ordering operator T in all terms from $A^{(1)}$ on.

²²⁵ Summary based on Schwichtenberg, NNQFT, section 11.7.

²²⁶ Schwichtenberg, QFT, 407.

²²⁷ Schwichtenberg, QFT, 393.

As an example, let's consider 2-pion scattering with the ϕ^4 potential of equation (Error: Reference source not found). Our initial and final states are then

$$\langle f | = \langle \pi_{k_3}^0, \pi_{k_4}^0 | \quad \text{and} \quad | i \rangle = | \pi_{k_1}^0, \pi_{k_2}^0 \rangle$$

For simplicity, we will consider only initial and final states where $(\vec{k}_3, \vec{k}_4) = (\vec{k}_1, \vec{k}_2)$.

Various techniques can be used to evaluate the zero-th order term to 1 in the case of equal momenta for the two pions before and after “interaction”²²⁸, otherwise zero. This is the case of nothing happening, and that is not interesting. So we will skip on to the first-order term.

The formulae (7.20) through (7.27) can be used to evaluate the first and second order terms, which we will simply cite here.

$$A^{(1)} = -i\lambda V$$

where V is the volume integrated and in principle will disappear at a later moment.

7.5. The final step – Feynman diagrams

Now we have a power series (7.29) to represent the probability amplitudes we want to calculate. Each term is a product of contractions. The evaluation of the terms in the series is facilitated greatly by the use of Feynman diagrams. Here's how it's done.

- The total transition amplitude is represented by a power series, each term of which is represented by a diagram.
- Just as each term in the series is a product of contractions, so will each diagram consist of a selection of factors corresponding to the contractions. Their value is given by the so-called Feynman rules, which depend on the interaction type. One set of Feynman rules is just the equations (7.20) - (7.25).

For our examples, Feynman diagrams are constructed according to the **Feynman rules for the ϕ^4 model**, which are resumed in Table 12. The rules give the diagrammatic representation and the amplitude factor for each possible type of contraction.

Each term in the series will include a power of the **coupling constant** for the basic interaction, λ in our ϕ^4 examples. More complicated terms (diagrams) lead to higher powers of the coupling constant (cc). If the cc is small, as for QED where its standard value is approximately $\alpha = \frac{1}{137}$, higher powers will be so small they can be neglected and this simplifies the calculation. For the strong force, this is complicated by the creation of virtual gluons referred to in section 6.9 which make the cc at a distance seem much larger than 1. Fortunately, at closer distances the cc becomes smaller and calculations are possible.²²⁹

228 Leaving out a whole lot here. See Schwichtenberg, op. cit. 409-11.

229 Griffiths, 70.

Contraction of	represents	Amplitude factor
a field with a particle in the initial state, $\phi(x) 1_k\rangle$	an incoming external line coming from outside and leading to x	e^{-ipx}
a field with a particle in the final state, $\langle 1_k \phi(x)$	an outgoing external line starting at x and ending outside the diagram	e^{ipx}
a field with itself, $\phi(x)\phi(y)$	an internal line running from x to y	$D_F(x, y)$ (propagator)
a point where lines meet	a vertex , denoted by a dot	$-i\lambda$ for each vertex
a line entering and leaving without interaction		$\sqrt{2\omega_q}\sqrt{2\omega_k}(2\pi)^3\delta(\vec{q}-\vec{k})$

Table 12. Feynman rules for ϕ^4 potential.

The first three of these configurations are just restatements of equations (7.20) - (7.25).

It is Wick's theorem that allows us to write any expression of the form

$$\langle 0|T\{\phi(x_1)\phi(x_2)\dots\phi(x_n)\}|0\rangle$$

as a sum of products of Feynman propagators. Let's go back and consider the four-field case. simplifying the notation some from the second part on.²³⁰

$$T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\} = N\{\phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 \\ + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 \\ + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4 + \phi_1\phi_2\phi_3\phi_4\}.$$

When we sandwich this between two vacuum states, with the help of (7.20), the result is

$$\langle 0|T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\}|0\rangle = D_F(x_1 - x_2)D_F(x_3 - x_4) \\ + D_F(x_1 - x_3)D_F(x_2 - x_4) + D_F(x_1 - x_4)D_F(x_2 - x_3). \quad (7.30)$$

since the normal-ordered term and any un-contracted terms result in 0. In terms of Feynman diagrams, this case gives the following three ways of connecting points 1-4.

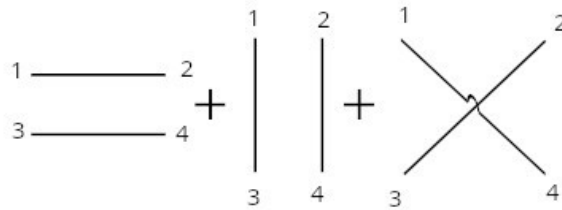


Figure . Feynman diagrams for equation (7.30).

²³⁰ Peskin and Schroeder, 89-90.

More interesting examples include these for two-particle scattering.

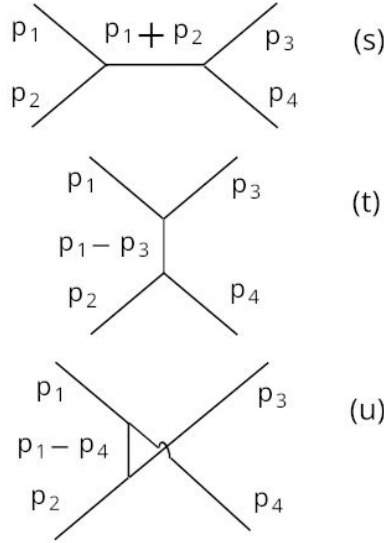


Figure B. Feynman diagrams for s, t and u-channel scattering.

The amplitude matrices for these diagrams, assuming coupling constant g at both vertices, are as follows:

$$\mathcal{M}_s = (-ig) \frac{i}{(p_1 + p_2)^2 - m^2 + i\epsilon} (-ig) = \frac{-ig^2}{s - m^2 + i\epsilon},$$

for the s-channel,

$$\mathcal{M}_t = (-ig) \frac{i}{(p_1 - p_3)^2 - m^2 + i\epsilon} (-ig) = \frac{-ig^2}{t - m^2 + i\epsilon},$$

for the t-channel, and

$$\mathcal{M}_u = (-ig) \frac{i}{(p_1 - p_4)^2 - m^2 + i\epsilon} (-ig) = \frac{-ig^2}{u - m^2 + i\epsilon}$$

for the u-channel. The quantities s , t and u are the **Mandelstam variables**. They satisfy

$$s + t + u = \sum m_j^2,$$

the sum of the invariant masses of the four particles involved. Note that the s-channel represents an annihilation, whereas the t and u channels are scattering processes.

8. Path integral formulation

The path-integral formulation of QFT, usually attributed to Feynman but in fact invented years before by Dirac, takes off from the double-slit experiment. It assumes multiplication not only of the number of slits but even of the screens containing them until there is nothing left but space. We must integrate over all possible paths between the light source and the chosen destination point on the detecting screen in order to find the result of the superposition of all those waves. So the integral is not just over the paths due to the two slits, but over all possible paths in space between the source and the destination. Hence the name, “**path integral**”.

We will look at this in some (not much) detail in order to get a good idea of what the final result means. Spoiler: It's all about phases.

The amplitude for a particle to start at spacetime point q_1, t_1 and propagate to q_2, t_2 can be calculated by using the unitary time-evolution operator:

$$\langle q_2, t_2 | q_1, t_1 \rangle = \langle q_2 | e^{-i\hat{H}(t_2-t_1)} | q_1 \rangle.$$

In order to integrate over all possible paths, we first break the time interval into N segments, each of length $\Delta t = \frac{t_2-t_1}{N}$.²³¹ Since the time-evolution operator is unitary, $U(t_b - t_a) = U(t_b - t_x)U(t_x - t_a)$ and we can use time slicing:

$$\langle q_2, t_2 | q_1, t_1 \rangle = \langle q_2 | (e^{-i\hat{H}\Delta t})^N | q_1 \rangle.$$

In order to get a product of amplitudes between neighboring states, we insert $N-1$ identity operators, each of which is curiously called a **fat unity**,²³²

$$\int dq_n |q_n\rangle \langle q_n| = 1.$$

Collect the integral business at the beginning and rename the initial and final states a and b to avoid confusion:

$$\langle q_b, t_b | q_a, t_a \rangle = \int dq_1 dq_2 \dots dq_N \langle q_a | e^{-i\hat{H}\Delta t} | dq_N \rangle \dots \langle q_{n+1} | e^{-i\hat{H}\Delta t} | dq_n \rangle \dots \langle q_1 | e^{-i\hat{H}\Delta t} | dq_a \rangle,$$

which contains a product of mini-propagators. We then can use a classical Hamiltonian operator like $\hat{H} = \frac{\hat{p}^2}{2m} + \hat{V}(q)$, but the q states are not eigenstates of \hat{p} . So we jump over to a p representation by inserting fat unities for p , thanks to Mr Fourier:

$$|q_n\rangle = \int dp |p\rangle \langle p | q_n \rangle = \frac{1}{\sqrt{2\pi}} \int dp |p\rangle e^{-ipq_n}.$$

An individual mini-propagator then looks like

$$G_n = \int \frac{dp}{\sqrt{2\pi}} \langle q_{n+1} | e^{-i\frac{p^2}{2m}\Delta t} | p \rangle e^{-ipq_n} e^{-iV(q_n)\Delta t},$$

where the three exponential factors are the kinetic energy, the Fourier factor and the potential energy. Having replaced the operators by their eigenvalues, we then go back to x via

$$\langle q | p \rangle = \frac{1}{2\pi} e^{\frac{ipq}{\hbar}}, \quad (8.1)$$

to get

$$G_n = \int \frac{dp}{2\pi} e^{-i\frac{p^2}{2m}\Delta t + ip(q_{n+1}-q_n)} e^{-iV(q_n)\Delta t},$$

which – believe it or not – can be integrated (Gaussian integral). Doing so and putting everything back together gives

$$\langle q_b, t_b | q_a, t_a \rangle = \prod_{n=1}^{N-1} \int \frac{dq_n}{\xi} e^{\frac{im}{2} \frac{(q_{n+1}-q_n)^2}{(\Delta t)^2} \Delta t} e^{-iV(q_n)\Delta t}, \quad (8.2)$$

²³¹ This particular explanation is from Blundell and Lancaster, 210-212. I think it is the clearest. See also Robinson, 219-222 or Townsend, 218-226 (the most complete/).

²³² Blundell and Lancaster, 211.

where $\xi^{-1} = (\frac{-im}{2\pi\Delta t})^{\frac{1}{2}}$ absorbs some constants. Now let $N \rightarrow \infty$, which means $\Delta t \rightarrow 0$ and our sum becomes an integral, so that, finally,

$$\langle q_b, t_b | q_a, t_a \rangle = \int Dq e^{i \int_{t_a}^{t_b} [\frac{m\dot{q}^2}{2} - V(q)] \Delta t} = \int Dq e^{i \int_{t_a}^{t_b} L dt} = \int Dq e^{i S[q(t)]}, \quad (8.3)$$

where L is the Lagrangian, S the action

$$S[q(t)] = \int dt L(q, \dot{q}) \quad (8.4)$$

and

$$Dq = \prod_{i=1}^{\infty} dq_i. \quad (8.5)$$

Equation (8.3) is a **sum over paths** with each path weighted statistically by $e^{iS[x(t)]}$.

To understand this, note that a set of $q_i, i = a, \dots, b$ represents a jagged line between fixed points a and b , a sequence of short straight-line segments with each segment starting from the spatial end of the preceding one and at a time equal to that of the preceding one plus Δt . As $N \rightarrow \infty$ and $\Delta t \rightarrow 0$, the string in the exponent becomes an integral over t , from the beginning point q_a, t_a to the end point q_b, t_b . A set of q_i then defines a path from a to b . The exponential integral over $L(q, \dot{q})$ defines a functional, the action $S[q(t)]$, which is calculated using that path. So each path is weighted by a path-dependent factor which changes only its phase.

Note that the result (8.3) is essentially a product of exponentials, each one of which contributes a phase. One can expand the difference between the phase at the minimum of the action and that at a point infinitesimally close to it in a Taylor's series. If the result is evaluated at a point close to the minimum of the action, its first derivative will disappear, so the difference will be a minimum and the phases almost the same. They then will add constructively and the probability amplitude at this point will be large. So the probability will be greatest at the point of minimum action, in accordance with classical mechanics. Although this may not be important for an individual event, the statistical sum of such events will show a strong probability for the classical result at a minimum of the action.

This derivation does not depend on use of SR or of the commutation relations of second quantization, making the path-integral formulation an *alternative* to second quantization. It can even be derived without mentioning operators at all.²³³

In order to actually use this formulation to calculate probability amplitudes, we introduce the **generating functional**²³⁴

$$Z[J] = \int D\phi \exp \left\{ iS[\phi] + i \int d^4x J(x)\phi(x) \right\}, \quad (8.6)$$

where

$$S[\phi] = \int d^4x \mathcal{L}$$

and we introduce an **external source** $J(x)$. Note the square brackets to indicate that $Z[J]$ is a functional of J , not a function. We also need the **functional derivative**, or **generating functional**, defined by analogy with $(\partial/\partial x_i)x_j = \delta_{ij}$ by

²³³ Townsend, 224.

²³⁴ Lancaster and Blundell, 201-203. I'm not sure why I included this. Could just skip it.

$$\frac{\partial J(x)}{\partial J(y)} = \delta^4(x - y) \quad (8.7)$$

so that

$$J(y) = \int d^4x \delta(x - y) J(x).$$

Then

$$-i \frac{\partial Z}{\partial J(x_1)} = \int D\phi \exp \left\{ iS[\phi] + i \int d^4x J(x) \phi(x) \right\} \phi(x_1)$$

brings “down” ϕ evaluated at the point x_1 and

$$-i \frac{1}{Z(0)} \frac{\partial Z}{\partial J(x_1)} \Big|_{J=0} = \frac{\int D\phi \exp \{ iS[\phi] \} \phi(x_1)}{\int D\phi e^{i \int d^4x \mathcal{L}(\phi)}} = \langle \Omega | \hat{\phi}(x_1) | \Omega \rangle,$$

where $|\Omega\rangle$ is the vacuum state. Better yet, applying this derivative n times gives

$$(-i)^n \frac{1}{Z(0)} \frac{\partial^n Z}{\partial J(x_1) \dots \partial J(x_n)} \Big|_{J=0} = \langle \Omega | T \{ \hat{\phi}(x_1) \dots \hat{\phi}(x_n) \} | \Omega \rangle,$$

where the operators on the right are automatically time-ordered. (What more can we ask?) So the generating functional (functional derivative) can be used to calculate *any* amplitude. Find it and your problem is solved. Right, as John Cleese would say!

Indeed, for the classical scalar Lagrangian density (5.52), this can be solved²³⁵ and the resulting amplitude is the Feynman propagator of (5.75). Although the result lacks the term $i\epsilon$ in the denominator, this will be recovered later in the boundary conditions for S-matrix calculation, but we won't go there.²³⁶

9. Renormalization

Anytime there is a loop in a Feynman diagram, there is the possibility of an infinite result. Consider the diagram in Figure 10, a possible Feynman diagram.²³⁷

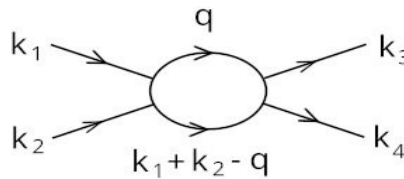


Figure 10. Possible second-approximation term (s-channel)

It is clear that all we can say from conservation of energy is that $k_1 + k_2 = k_3 + k_4$. The quantity q is undetermined and so integrating over all possible values of it goes like

$$\int_0^\infty d^4q \frac{1}{q^4} \sim \ln(q) \Big|_0^\infty \sim \infty \quad (9.1)$$

²³⁵ See Schwartz, 262 or Robinson, 225-227.

²³⁶ But Schwartz does, 264-266.

²³⁷ The loop could just as well be represented as a square.

and diverges logarithmically.

The classical way of handling this problem was the method of Feynman, Schwinger, Tomonaga and Dyson, to “renormalize” QFT by canceling so-called infinite counterterms. This works, as “QED is the most precisely tested theory of all time.”²³⁸ The more modern way of doing this is due to Ken Wilson and is called **Effective Field Theory**, or **EFT**. This method does not obsolete the older one, as different cases are easier to solve with one or the other.

In EFT, we admit we do not know what is going on above some energy called the **cutoff energy** and denoted here by E_* . Above this energy is the so-called “ultraviolet (UV)” zone of higher energies; below, the “infrared (IR)” zone of lower energies. EFT supposes that it is possible to do physics on the low-energy level without worrying about the higher-energy stuff, since we can’t observe it anyway. In fact, this is what we do all the time, since we really don’t know what is going on up there. This being the case, our integrals probably should not extend into this unknown region just in order to satisfy convenient boundary conditions. Specifically, EFT proposes handling Feynman diagrams with loops, which give infinite results, as sums of diagrams below E_* and then the result is finite – sometimes.

Dimensional analysis is helpful for understanding. Assuming $c = \hbar = 1$:

$$E = mc^2 \Rightarrow [E] = [M][V]^2 = [M][L]^2/[T]^2, \quad \text{and} \quad c = 1 = [L]/[T]$$

along with

$$\hbar = 1 = [M][L]^2/[T]$$

give us

$$[L] = 1/[E] = [T].$$

Then the scalar field-theory Lagrangian goes like

$$\frac{K.E.}{\text{space volume}} \sim \frac{[E]}{[L]^3} \sim [E^4] \quad \text{and} \quad \frac{1}{2} \left(\frac{d\phi}{dt} \right)^2 \sim \frac{[\phi]^2}{[T]^2} \sim [\phi]^2 [E]^2.$$

This tells us that $[\phi] \sim [E]$, so $\mathcal{L} \sim [E]^4$. Now expand the interaction \mathcal{L} in powers of ϕ ,

$$\mathcal{L} \sim c_3 \phi^3 + c_4 \phi^4 + c_5 \phi^5 + \dots$$

If we assume a cutoff energy E_* , we can guess that $c_3 \sim [E_*]$, $c_4 \sim 1$ and $c_5 \sim [E_*]^{-1}$. We then work at energies well below E_* , where c_5 should be irrelevant. The terminology is

$c_2 \phi^2, c_3 \phi^3$	are	<i>relevant</i> interactions
$c_4 \phi^4$	are	<i>marginal</i> interactions
$c_5 \phi^5, c_6 \phi^6 \dots$	are	<i>irrelevant</i> interactions

With luck, there will only be a small number of relevant or marginal terms, and this usually turns out to be the case. But what is really advantageous is that *the physical results of EFT calculations do not depend on the cutoff*. This is because the coupling constants, c_n , evolve with changing E_* . There is a specific functional form for the dependence of the c_n on E_* . The coupling “constants” – which are not at all constant – are said to “flow” or “run” along a trajectory in the space of theories.

This method of calculating is called the **renormalization group**, although it is not at all a group in the group-

238 Carroll

theory sense. Referring to E_* as Λ , “In a finite theory with a UV cutoff Λ , physics at energies $E \ll \Lambda$ is independent of the precise value of Λ . Changing Λ changes the couplings in the theory so that observables remain the same.”²³⁹

There is an excellent analogy which considers firing a cannon ... underwater. The cannonball now undergoes drag from the water. Nevertheless, its *effective mass* can be determined by shaking it to and fro, and this measured value then can be used in calculations, otherwise ignoring the influence of the liquid environment. But the effective mass varies with the velocity, since as the latter approaches 0, the effects of the water become negligible. “In other words, the presence of a medium can introduce a scale-dependent effective mass. We say that the effective mass is ‘renormalized’ by the medium. In quantum physics, every particle moves through a ‘medium’ consisting of the quantum fluctuations of all particles present in the theory. We again take into account this medium by ignoring it but changing the values of our parameters to scale-dependent ‘effective’ values.”²⁴⁰

The Wilsonian renormalization group is only one of several methods for taming divergent amplitudes in physics. This is a vast subject and, to my mind, somewhat less than passionately interesting – especially in the details – in spite of its importance. After all, if we can’t calculate it, we don’t really know it – at least not completely.

Let’s look at one aspect of this in somewhat more detail, at the risk of repeating things that have already been stated.

Often an integral which diverges can be “tamed” by putting a cutoff value on the variable of integration, usually denoted Λ . It may even come about that Λ disappears in the final equation, even if we let it go to ∞ . The step of imposing a cutoff value is called **regularization**. Often, the integral in question is constructed in an unnatural way, with infinite limits to satisfy some idealized boundary conditions. Also, it is not certain – nor agreed upon by all physicists – that the SM is valid at very high energies.

Consider the divergence discovered in the diagram of Figure 10. We will **regularize** the integral in (9.1) by using a cutoff Λ so the amplitude of order 2 can be written as²⁴¹

$$\mathcal{M}_2 \sim \lambda^2 \int_{\Lambda}^{\Lambda} d^4q \frac{1}{q^2 - m^2 + i\epsilon} \frac{1}{(k_1 + k_2 - q)^2 + i\epsilon},$$

which leads to

$$\mathcal{M} = -\lambda - \lambda^2 C \ln \frac{s}{\Lambda^2}, \quad (9.2)$$

where s is the energy scale $s = (k_1 + k_2)^2$ and C is a constant.²⁴² We then compare the amplitude at two values of the energy scale and find

$$\mathcal{M}(s_1) - \mathcal{M}(s_2) = -\lambda^2 C \ln \frac{s_2}{s_1}. \quad (9.3)$$

Now we are getting somewhere. Define the **renormalized coupling constant** λ_R by

$$\lambda_R \equiv -\mathcal{M} = \lambda + \lambda^2 C \ln \frac{s_0}{\Lambda^2}. \quad (9.4)$$

Then expand λ about λ_R using $\lambda = \lambda_R + a\lambda_R^2 + \dots$ and collect all this to get

$$\mathcal{M} = -\lambda_R - \lambda_R^2 C \ln \frac{s}{s_0}. \quad (9.5)$$

²³⁹ Schwartz, 417.

²⁴⁰ Nelson, P. *American Scientist*, 73, 66 (1985). Quoted by Griffiths, 220.

²⁴¹ Schwartz, 297.

²⁴² Scharz finds $C = \frac{1}{32\pi^2}$.

This is great! The renormalized coupling constant λ_R is measurable at a given energy scale s_0 . Then the value of these two parameters can be used to predict the amplitude at any other energy scale s by (9.4). In fact, the renormalized coupling constant is the only one we can measure.

This result can also be derived from the Lagrangian by assuming an extra, infinite term, supposed of the order of λ_R^2 :

$$\mathcal{L} = -\frac{1}{2}\phi\Box\phi - \frac{\lambda_R}{4!}\phi^4 - \frac{\delta_\lambda}{4!}\phi^4, \quad (9.6)$$

for a real scalar field. Then from (9.2)

$$\mathcal{M}(s) = -\lambda_R - \delta_\lambda - \lambda_R^2 C \ln \frac{s}{\Lambda^2} + \mathcal{O}(\lambda_R^4).$$

Choosing

$$\delta_\lambda = -\lambda_R^2 C \ln \frac{s_0}{\Lambda^2}$$

takes us right back to equation (9.4). The 3rd term in the Lagrangian (9.4) is called a **counterterm**. Such terms contribute an infinity to the Lagrangian but drop out when *physical quantities* are computed.

In résumé, we can quote another book:²⁴³

A renormalizable theory is one in which infinities to all orders of perturbation theory can be absorbed by the redefinition of a finite number of the parameters of the theory, such as masses and coupling constants. These parameters are then fixed from experimental observation.

In fact, it has been proven that all gauge theories, including those we have studied, are renormalizable.²⁴⁴

10. Physics and finance

It is interesting as well as instructive to consider the parallel between physics and finance.²⁴⁵

In a toy model of economic exchanges, start by defining the exchange rate

$$R_i(\vec{n}) \equiv e^{A_i(\vec{n})}. \quad (10.1)$$

Here \vec{n} indicates the start point (country) and i the destination, taken as one unit vector away ($i, j = x, y$, for instance). Since each “country” can adjust its currency at will, define the currency re-scaling factor, which is the *gauge transformation*,

$$f(\vec{n}) \equiv e^{\epsilon(\vec{n})}. \quad (10.2)$$

It is a gauge symmetry, since nothing material in the real world changes as a result of it. Then if the neighboring i country rescales its currency, the exchange rate $R_i(\vec{n})$ changes by

$$R_i(\vec{n}) \rightarrow \frac{f(\vec{n} + \vec{e}_i)}{f(\vec{n})} R_i(\vec{n}) \quad (10.3)$$

and these lead to

$$A_i(\vec{n}) \rightarrow A_i(\vec{n}) + \epsilon(\vec{n} + \vec{e}_i) - \epsilon(\vec{n})$$

²⁴³ Barr et al., 3.

²⁴⁴

²⁴⁵ This section inspired by Schwichtenberg, PF, and Maldacena.

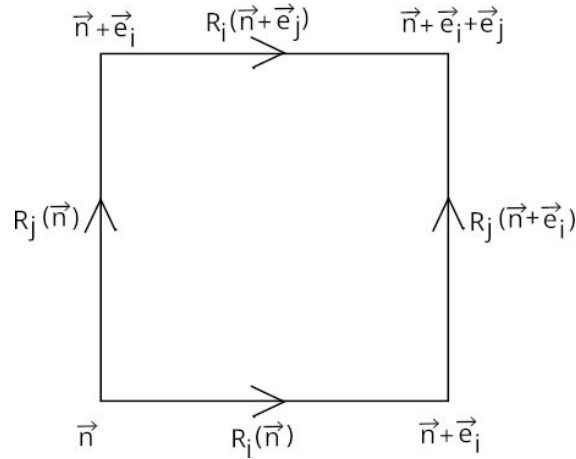
which in the continuum limit becomes

$$A_\mu \rightarrow A_\mu + \partial_\mu \epsilon. \quad (10.4)$$

Comparison with EM shows that ϵ is the gauge field, give or take a factor of i .

Considering only circulation of currency (no charges or commodities), the currency gain factor

$$G(\vec{n}) \equiv e^{F_{ij}(\vec{n})}. \quad (10.5)$$



This equation assumes taking R around a closed loop from (\vec{n}) to $(\vec{n} + \vec{e}_i)$ to $(\vec{n} + \vec{e}_i + \vec{e}_j)$ to $(\vec{n} + \vec{e}_j)$ and back to (\vec{n}) ,

$$G(\vec{n}) = R_i(\vec{n})R_j(\vec{n} + \vec{e}_i) \frac{1}{R_i(\vec{n} + \vec{e}_j)R_j(\vec{n})}. \quad (10.6)$$

This leads to

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (10.7)$$

which looks pretty much like the EM field tensor. $G(\vec{n}) \neq 1$ means traders can gain money, which can only occur if $F_{\mu\nu} \neq 0$ so $A_\mu \neq 0$. In words, if an EM field is present.

Now introduce commodities, which we'll take to be copper (Cu). Define the commodity-exchange gain factor

$$g_i(\vec{n}) \equiv e^{H_i(\vec{n})} \quad (10.8)$$

and the local commodity price

$$p(\vec{n}) \equiv e^{\phi(\vec{n})}. \quad (10.9)$$

The gain factor across one border is

$$g_i(\vec{n}) = \frac{p(\vec{n} + \vec{e}_i)}{p(\vec{n})R_i(\vec{n})}$$

which leads to

$$H_i(\vec{n}) = \phi(\vec{n} + \vec{e}_i) - \phi(\vec{n}) - A_i(\vec{n}).$$

(This looks suspiciously like a covariant derivative.) We can define a current in terms of $H_i(\vec{n})$ which in the continuum limit leads to

$$J_\mu(t, \vec{x}) = qH_i(\vec{n}) = q(\partial_\mu\phi(t, \vec{x}) - A_\mu(t, \vec{x})). \quad (10.10)$$

Equation (10.10) shows the current as a sum of two parts, the first dependent on changes in the price of Cu, the second on changes of the exchange rate. We can call it a current because J_i measures the exchange of Cu between two countries and J_0 , in time within a country. Putting (10.3) into (10.6) shows that G and $F_{\mu\nu}$ are independent of price changes. Considerations of symmetry can inform us which quantities are independent of local conventions.

All these exponential functions are really talking about Lie algebras in terms of their generators, which are the quantities we have considered.

One can also jump a mathematical level up and consider fiber bundles. A fiber bundle is helpful for representing a product of two spaces, as we have done for spacetime and the charge space A_μ . At each point of the base space, spacetime for us, a fiber is added, the internal charge space. A fiber bundle is a generalization of the direct product of two spaces.²⁴⁶

11. Review and summing up

Here is where we have been. We started this document with the following paragraph.

11.1. Really short summary

For physicists, 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. Such groups are abstract objects, but we can imagine concrete examples in some space (for instance, spacetime or spinor space); these instantiations of the group are called **representations**. Since the Lie group and therefore the representation are continuous, the latter can be generated incrementally from the identity operator by using operators called **generators**. The generators are derivatives of Lie group elements (in a representation) and form a vector space called a **Lie algebra**. The generators are especially interesting to physicists because they may represent physical observables such as angular momentum or spin. They thus provide links between transformations and observables. This allows us to deduce properties of the world around us from properties of mathematical groups, which are thus a window onto the world.²⁴⁷

11.2. Lie groups and algebras – Forces and particles

The notion of symmetry in physics is based on a transformation after which a system is left in a different state but behaves the same. Such transformations are seen as changes resulting from the application of forces. Each symmetry of the Lagrangian is described by a **Lie group**, which allows us to use a **representation** of the Lie group in order to describe the results of applying a force to the system described by the Lagrangian.

A representation is described by means of a **Lie algebra**, which is made up of the **generators** of a representation. Although the dimensions of representations vary, the number of generators of all those of a given group is always the same and is equal to the **order**, the number of parameters describing the group transformations.

Valid Lagrangians possess symmetries. Indeed, what constitutes a valid Lagrangian is imposed by the requirement of Lorentz invariance – symmetry under Lorentz transformations. Noether's theorem tells us that for each symmetry, a calculable quantity is conserved -- energy or momentum or angular momentum, for instance, but also spin, hypercharge or color.

²⁴⁶ More in Schwichtenberg, PF, 163ff.

²⁴⁷ I think Pythagoras would have loved this!

Lie groups tell how rotations, Lorentz boosts and changes of phase (unitary transformations), described respectively by SU(2), SU(2) and U(1), modify vectors within the vector space of the representation of the group.

Some of the generators are diagonal and therefore correspond to eigenvectors, which can be used as a basis for the vector space of the representation. They constitute the **Cartan subalgebra** of the Lie algebra and their number is the **rank** of the algebra. The eigenvalues are the physically measurable charges of the particle on which the force acts, and are equal in number to the dimension of the fundamental representation, which in the case of SU(n) or SO(n) consists of $n \times n$ matrices. In other words, an arbitrary vector will have components along different eigenvectors of the Cartan generators. Transformations of the group then will rotate one vector into another, changing the eigenvalues to which it corresponds, as will a rotation on j_z . The “rotations” may be in real space, in Minkowski spacetime or in some internal space, such as spin or color space.

An arbitrary SU(n) group always has $n^2 - 1$ generators and is of rank $n - 1$, whereas an arbitrary SO(n) group has $\frac{n(n-1)}{2}$ generators.

11.3. Examples

Symmetry and QFT bring more than the understanding of forces and how they arise.

For instance, in SU(2), there are three generators, J_{1-3} , of which only one is a Cartan generator, by convention J_3 . Its eigenvectors form a basis for the space. In the 2-dimensional representation, for $j = \frac{1}{2}$, the eigenvectors represent particles of positive and negative charge (spin), $j_z = \pm \frac{1}{2}$. The non-Cartan generators, J_1 and J_2 , are used to construct ladder (raising and lowering) operators which change the charge, in this case, the spin.

In the case of the strong force, the 3-dimensional fundamental representation of SU(3) gives rise to the three eigenvalues which are the three quark colors. The gauge bosons (gluons) describe how the colors are changed.

The Cartan generators (photon, Z^0) represent force-carrying particles which can modify, say, position or momentum, but without any change in charge. The non-Cartan generators (W bosons, gluons) are those which comprise the QM raising and lowering operators, so it is not a surprise that they can also bring about changes in charge.²⁴⁸

11.4. QFT

The requirement of SR that equations be invariant under Lorentz transformations leads to the Lagrangians for particles of spin 0, $\frac{1}{2}$ and 1, our old friends the Klein-Gordon, Dirac and Proca Lagrangians. From here, we can proceed along three (or four) different paths.

1. Use of the Euler-Lagrange equations on these Lagrangians furnishes the equations of motion for single free particles. The equations may be solved in terms of Fourier expansions of the field. We then employ second quantization: We quantify the fields themselves, making them operators by fixing their commutation relations à la QM. Then the coefficients of the terms are also operators and their commutation relations can be found. The result is that they are like the QM solution to the harmonic-oscillator problem and can be combined to form ladder operators which “create” and “destroy” equal chunks of energy in the vacuum state. We may identify these “chunks” as particles. So use of second quantization leads to the existence of particles of the field. The solutions to the equations of motion also may be used to calculate the Feynman propagators, mathematical representations of a virtual particle destroyed at one point in space and time in the vacuum and (re)created at a later time at another such point.
2. The equations of motion are invariant under global transformations. Also requiring invariance under

²⁴⁸ Robinson, 116-117.

local transformations gives extra terms (connections) which can be included in a covariant derivative and require addition of another field, representing one or more massless gauge bosons which are the force-carrying particle for the field.

- For a single spinor, imposing local $U(1)$ symmetry leads to the interaction between an electron and a gauge boson (photon) and to Maxwell's equations – QED.
 - For a doublet of two equal-mass Dirac fields, imposing local $SU(2)$ invariance leads to the requirement of interactions with three zero-mass gauge fields. This is because the 2-dimensional representation of $SU(2)$ is expressed by the three Pauli matrices and each one contributes an interaction term. The problem now is that the original fields have equal mass and that the gauge bosons have zero mass. The latter situation is obviated by symmetry breaking through the Higgs mechanism. The result then applies to the weak interaction and the W and Z bosons..
 - A triplet of three equal-mass Dirac fields (quarks!) can be subjected to local $SU(3)$ invariance which requires eight massless gauge fields, easily identified as gluons, the force-carrying particles of the strong interaction. Voila QCD.
3. One can use the S matrix approach in order to define scattering terms, leading to the Dyson series and Feynman diagrams. A more modern approach uses Feynman's path-integral approach to construct terms of the interaction which correspond to Feynman diagrams. However we get there, we can use the diagrams to calculate interaction probabilities, some of which can use the Feynman propagators already mentioned.

The principal ingredients used in this cuisine are Lorentz invariance, mathematical Lie groups and Lie algebras, and the commutation relations of QM operators.

11.5. Why second quantization?

First quantization takes “ordinary” quantities, position and momentum, and quantizes them, which results in an equation for new entities -- probability amplitudes (wave functions) which entail discrete modes for the energy of a system like the hydrogen atom or a particle in a box. Second quantization quantizes the resulting wave functions into fields which, with their conjugate momenta, produce integral numbers of excitations of each such mode. They also result in creation, annihilation and counting operators to produce such states.

As Schwartz says: “At the risk of oversimplifying things a little, that [second quantization] is all there is to quantum field theory. The rest is just quantum mechanics.”²⁴⁹ In his analysis, QFT brings two great advantages:

- Because of the creation/annihilation operators, it can account for multi-particle states.
- The n^{th} excitation state of a system is considered to represent n particles.

For instance, applying a creation operator twice to the vacuum $|0\rangle$ creates a two-particle state (for bosons only).

12. Annex: Math symbols

Symbol	Means
\forall	for all, for each
\in	in, a member of
\exists	there exists
$ $ or $:$	such that

²⁴⁹ Schwartz, 20.

13. Bibliography

Barr, G., Devenish, R., Walczak, R. Weidberg, T. *Particle physics in the LHC era*. Oxford: Oxford University Press, 2016. Print.

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

Feynman, Richard. *QED: The strange theory of light and matter*. London: Penguin, 1990. Print.

Georgi, Howard. *Lie algebras in particle physics*. Second edition. Westview Press, 1999. Pdf.

Greiner, Walter and Müller, Berndt. *Quantum mechanics symmetries*. 2nd edition. Berlin: Springer, 1994. Pdf.

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.

Maggiore, Michele. *A modern introduction to quantum field theory*. Oxford: Oxford University Press, 2009. Print.

Maldacena, Juan. *The symmetry and simplicity of the laws of physics and the Higgs boson*. On-line. arxiv.org/abs/1410.6753.

Peskin, Michael E. and Schroeder, Daniel V. *An introduction to quantum field theory*. Westview Press, 1995. Print.

Robinson, Matthew. *Symmetry and the standard model*. New York: Springer, 2011. Print.

Schwartz, Matthew D. *Quantum field theory and the standard model*. Cambridge: Cambridge University Press, 2014. 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.

Srednicki, Mark. *Quantum field theory*. Cambridge: Cambridge University Press, 2007. Draft version (2006) online, www.physics.ucsb.edu/~mark/qft.html.

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

Teller, Paul. *An interpretive introduction to quantum field theory*. Princeton: Princeton University Press, 1995. Print.

Weinberg, Steven. *The quantum theory of fields. Volume I. Foundations*. Cambridge: Cambridge University Press, 1995. Print.