RELATIVISTIC QUANTUM MECHANICS

C.G. OAKLEY

1. Introduction

The traditional approach to relativistic quantum field theory is unsatisfactory for a number of reasons. Most importantly:

1. There is no reason why the formal "quantization" procedure should work. Quantization takes us from a classical system to a quantum one, and consists in the replacement of Poisson brackets with commutators. We are therefore implying the behaviour of the more comprehensive quantum world from the far less general classical world, which clearly is the wrong way round. In any case the procedure, when applied to quantum fields, only works properly for spin zero.
2. The traditional approach makes extensive use of the interaction picture, which does not exist (Haag's theorem).
3. The result of subtracting infinity from infinity is indeterminate. If one ignores this fact one will produce theories of no scientific value.

There are things, however, that we *can* justify, namely *special relativity* and *quantum mechanics*. Both are on an extremely sound footing both experimentally and theoretically. The tools for combining them have been available for a long time in Wigner's Unitary Irreducible Representations of the Poincaré group. We can see the concepts of *mass* and *spin* arising very naturally, and we are led from here to the notions of Fock space, creation and annihilation operators and finally free quantum fields. Although it covers issues normally associated with quantum field theory, this treatise has the title "relativistic quantum mechanics" on the grounds that the quantum field is not viewed as fundamental here, being derived instead from annihilation and creation operators, which in turn are defined as operators on Fock space. It will be shown that the presence of interactions does not invalidate this analysis. Interactions will be seen, in effect, to be just interference patterns between free field states.

2. Definitions

Relativistic quantum mechanics is the synthesis of quantum mechanics and special relativity. We will therefore start by examining both of these in a formal way.

Quantum mechanics: definitions

I. The states of a physical system are a vector space $\mathcal{V}$ over the complex numbers $\mathcal{C}$.

II. There exists a map $\mathcal{V} \otimes \mathcal{V} \to \mathcal{C}$ called the "inner product" as follows:

(i) It is sesquilinear, i.e. for any $X, Y, Z \in \mathcal{V}$, $a \in \mathcal{C}$

$(X, Y) = (Y, X)^*$
$(X, Y + a.Z) = (X, Y) + a.(X, Z)$

(ii) It is positive definite, i.e.

(a) $(X, X) \geq 0$ for any $X \in \mathcal{V}$
(b) $(X, X) = 0 \Longleftrightarrow X = 0$

We will use Dirac notation where a vector X is shown thus: $|X\rangle$. Dual vectors, which map such vectors to $\mathcal{C}$ are denoted thus: $\langle Y|$. These are defined from vectors through the inner product, i.e.

$$\langle Y|X\rangle = (|Y\rangle, |X\rangle)$$

for any vector $|X\rangle$, defines the dual vector $\langle Y|$ associated with the vector $|Y\rangle$.

III. A unitary representation of the group of space-time displacements and three-dimensional rotations acts non-trivially on $\mathcal{V}$. By *unitary*, we mean that the group action preserves the inner product.
The most general group element is

$$U(\mathbf{d},t,\theta) = \exp(-itH - i\mathbf{d}\cdot\mathbf{P} - i\theta\cdot\mathbf{J})$$

where t is the time displacement, $\mathbf{d}$ is the space displacement, θ is the axis of rotation and $|\theta|$ is the angle. Unitarity implies that the generators are Hermitian operators. The algebra is

$$[H, P_i] = [H, J_i] = 0; \qquad [P_i, P_j] = 0$$
$$[J_i, J_j] = i\epsilon_{ijk}J_k; \qquad [J_i, P_j] = i\epsilon_{ijk}P_k$$

Quantum mechanics associates the time displacement operator with classical energy, the space displacement operator with classical momentum and the rotation operator with classical angular momentum. Conservation of these quantities is then very simply explained as the consequence of the group algebra.

Special Relativity: definitions

The spacetime of special relativity, known as Minkowski space, is an $\mathcal{R}^4$ manifold with a metric of signature -2, whose metric connection has vanishing torsion and curvature. This makes it possible to choose co-ordinate systems in which the metric always takes the form

$$\eta^{ab} = diag.(1, -1, -1, -1) \tag{2.1}$$

Spacetime is then parameterised by a set of co-ordinates

$$x^a = (x^0, x^1, x^2, x^3) = (ct, x, y, z) \tag{2.2}$$

The values x, y and z are Cartesian spatial co-ordinates measured by some observer and t is the time. The constant c gives the fundamental relation between measurements of time and distance. It is also the speed of light in a vacuum.

The group of metric-preserving automorphisms of Minkowski space is known as the Poincaré group. As we will see, this group is ten-dimensional, consisting of space-time displacements, rotations and transformations that set the reference frame in motion.

To satisfy the requirements of special relativity, a representation of the Poincaré group must act non-trivially on the vector space of quantum mechanics. This is how we can embody the principle that the viewpoints of observers related by Poincaré group operations are equivalent. Our study therefore begins with the Poincaré group itself.

3. The Poincaré group

The most general automorphism of Minkowski space that preserves the metric is

$$x^a \to \Lambda^a{}_b x^b + d^a \tag{3.1}$$

for some constants $\Lambda^a{}_b$ and d^a, where

$$\Lambda^a{}_b \Lambda^c{}_d \eta_{ac} = \eta_{bd} \tag{3.2}$$

We can think of Λ and η as matrices, in which case this constraint can be written as

$$\Lambda^T \eta \Lambda = \eta \tag{3.3}$$

If we take determinants we find that

$$\det(\Lambda) = \pm 1 \tag{3.4}$$

so transformations Λ for which $\det(\Lambda) = -1$ are not connected to the identity. Considering the $_{00}$ component of the equation, we see also that

$$(\Lambda^0{}_0)^2 - \sum_k (\Lambda^k{}_0)^2 = 1 \tag{3.5}$$

which implies that either $\Lambda^0{}_0 \geq 1$ or $\Lambda^0{}_0 \leq -1$. Operations for which the latter is true are not connected to the identity. So we can divide the Lorentz group (by which I mean the subgroup of the Poincaré group obtained by putting $d^a = 0$) into four disconnected parts:

$$\begin{array}{lll}
\text{(i)} & \Lambda^0{}_0 \geq 1 & \det(\Lambda) = 1 \\
\text{(ii)} & \Lambda^0{}_0 \geq 1 & \det(\Lambda) = -1 \\
\text{(iii)} & \Lambda^0{}_0 \leq -1 & \det(\Lambda) = -1 \\
\text{(iv)} & \Lambda^0{}_0 \leq -1 & \det(\Lambda) = 1
\end{array} \tag{3.6}$$

(i) alone is called the *restricted* or *identity-connected* Lorentz group. (i)⊕(iv) is called the *proper* Lorentz group and (i)⊕(ii) is called the *orthochronous* Lorentz group.

Two group operations, not connected to the identity, are *parity* or *space inversion*, represented by $\Lambda = I_p$ =diag.$(1, -1, -1, -1)$, and *time reversal*, represented by $\Lambda = I_t$ =diag.$(-1, 1, 1, 1)$. Elements of (ii), (iii) and (iv) can be written as $I_p\Lambda_r$, $I_t\Lambda_r$ and $I_pI_t\Lambda_r$ respectively, where Λ_r is an element of the restricted Lorentz group. So we reduce the study of the Lorentz group to the study of the identity-connected part, combined with the discrete operations of time reversal, parity and spacetime reversal.

The Killing vectors of Minkowski space are given by $\mathcal{L}_K\,\eta = 0$, where $\mathcal{L}$ means the Lie derivative. It is easy to show that the solution to this equation is that K is a linear combination of

$$P_a = i\partial_a \tag{3.7}$$

and

$$M_{ab} = i(x_a\partial_b - x_b\partial_a) \tag{3.8}$$

using a co-ordinate system of the kind defined previously (2.1). The apparently perverse introduction of factors of i is so that generators of a unitary representation will be Hermitian operators. The commutators of the Killing vector fields are

$$[P_a, P_b] = 0 \tag{3.9}$$

$$[M_{ab}, P_c] = -i\left(\eta_{ca}P_b - \eta_{cb}P_a\right) \tag{3.10}$$

$$[M_{ab}, M_{cd}] = -i\left(\eta_{ac}M_{bd} - \eta_{bc}M_{ad} + \eta_{bd}M_{ac} - \eta_{ad}M_{bc}\right) \tag{3.11}$$

which is evidently a Lie algebra —the algebra of the Poincaré group.

The tensor M_{ab} is most readily understood in terms of its component three-vectors. Define

$$J_i = \frac{1}{2}\epsilon_{ijk}M^{jk} \text{ and } N_i = M^{0i}, \text{ where } i, j, k = 1, 2, 3 \tag{3.12}$$

J_i then generate rotations and N_i generate boosts, which are the transformations that set one frame into motion with respect to the other. It is easy to check that

$$[J_i, J_j] = i\epsilon_{ijk}J_k \tag{3.13}$$

$$[J_i, N_j] = i\epsilon_{ijk}N_k \tag{3.14}$$

$$\text{and } [N_i, N_j] = -i\epsilon_{ijk}J_k \tag{3.15}$$

The group element is $\exp(-\frac{i}{2}\omega_{ab}M^{ab}) = \exp\left(-i\theta\cdot\mathbf{J} - i\beta\cdot\mathbf{N}\right)$ where $\beta_i = \omega_{0i}$ and $\theta_i = \frac{1}{2}\epsilon_{ijk}\omega_{jk}$. If $\theta = 0$ then β is the direction of the boost and $c\tanh|\beta|$ is its velocity. If $\beta = 0$ then θ is a vector parallel to the axis of rotation, whose magnitude is the angle of rotation. Note that since there is no non-zero value of β that gives the identity, the Lorentz group is non-compact.

3.1 Finite-dimensional representations of the Lorentz group. The group SL(2,C).

The vector space of quantum mechanics carries an infinite-dimensional unitary representation of the Poincaré group. In analysing this, we will come across the notion of spin, which uses finite-dimensional representations of the Lorentz sub-group. Let us examine the latter now. Our first observation is that the representation cannot be unitary, since the group is non-compact. Define

$$K_i^{(\pm)} = \frac{1}{2}\left(J_i \pm iN_i\right) \tag{3.16}$$

from which it follows that

$$\left[K_i^{(\pm)}, K_j^{(\pm)}\right] = i\epsilon_{ijk}K_k^{(\pm)} \tag{3.17}$$

$$\text{and } \left[K_i^{(+)}, K_j^{(-)}\right] = 0 \tag{3.18}$$

The Casimir operators are

$$\mathbf{K}^{(\pm)2} = \frac{1}{4}\left(\mathbf{J} \pm i\mathbf{N}\right)\cdot\left(\mathbf{J} \pm i\mathbf{N}\right) \tag{3.19}$$

$$= \frac{1}{4}\left(\mathbf{J}^2 - \mathbf{N}^2 \pm 2i\mathbf{J}\cdot\mathbf{N}\right) \tag{3.20}$$

$$= \frac{1}{8}M_{ab}M^{ab} \pm \frac{i}{16}\epsilon_{abcd}M^{ab}M^{cd} \tag{3.21}$$

So in representations where $K_i^{(\pm)}$ are Hermitian, we have two commuting $SU(2)$ algebras. We know how to construct finite-dimensional representations of this group. Labelling the representation with the ordered pair $\left(j^{(+)}, j^{(-)}\right)$ the representation will be $\left(2j^{(+)}+1\right)\left(2j^{(-)}+1\right)$-dimensional and the Casimir operators will have eigenvalues $j^{(\pm)}\left(j^{(\pm)}+1\right)$.

The displacements generated by the Lorentz transformation Killing vectors are given by

$$\exp\left(-\frac{i}{2}\omega_{cd}M^{cd}\right) \ : \ x^a \to \Lambda^a{}_b\, x^b \tag{3.22}$$

We can write

$$\Lambda^a{}_b = \exp\left(-\frac{i}{2}\omega_{cd}S^{cd}\right)^a{}_b \tag{3.23}$$

$$\text{with } \left(S^{cd}\right)^a{}_b = i\left(\eta^{ca}\delta_b^d - \eta^{da}\delta_b^c\right) \tag{3.24}$$

being matrices which have the same algebra as the M's.

We can think of these matrices as generating a representation of the Lorentz group in an abstract kind of way —the *four vector* representation. This can then be classified in terms of its $\mathbf{K}^{(\pm)2}$ eigenvalues. A quick calculation gives

$$\left(\mathbf{K}^{(\pm)2}\right)^a{}_b = \left(\frac{1}{8}S^{cd}S_{cd} \pm \frac{i}{16}\epsilon_{cdef}S^{cd}S^{ef}\right)^a{}_b$$

$$= \frac{3}{4}\delta_b^a \tag{3.25}$$

So the four vector belongs to the $\left(\frac{1}{2},\frac{1}{2}\right)$ representation of the Lorentz group.

Let us now construct the $\left(\frac{1}{2},0\right)$ representation.

The vector space is two-dimensional and complex, and an appropriate choice of generators of the Lorentz group is given by

$$\mathbf{K}^{(+)} = \frac{1}{2}\tau \quad ; \quad \mathbf{K}^{(-)} = 0 \tag{3.26}$$

where τ_i are the Pauli matrices. We write the vectors —usually called *spinors* —as ψ_A where $A = 1,2$. The group action is a linear transformation with coefficients

$$M_A{}^B = \exp\left(-\frac{i}{2}\alpha\cdot\tau\right)_A{}^B \tag{3.27}$$

where $\alpha = \theta - i\beta$. Now $-\frac{i}{2}\alpha\cdot\tau$ is the most general traceless complex 2×2 matrix, any matrix of this kind being formed by making suitable choices for θ and β, so we conclude that M is the most general complex unimodular matrix (although it should be stated that there are some pathological cases of unimodular 2×2 matrices that cannot be written as an exponential).

Hence the $\left(\frac{1}{2},0\right)$ representation of the Lorentz group is also the defining representation of the group SL(2,C).

Although there is clearly a local isomorphism here, globally we find that SL(2,C) covers the identity-connected part of the Lorentz group 2 to 1. We see this if we form a pure rotation

$$R = \exp(-\frac{i}{2}\theta\cdot\tau) = \cos(\frac{1}{2}\theta\cdot\tau) - i\sin(\frac{1}{2}\theta\cdot\tau) \tag{3.28}$$

Using $(\theta\cdot\tau)(\theta\cdot\tau) = \theta^2$ it follows that

$$R = \cos\left(\frac{|\theta|}{2}\right) - i\hat{\theta}\cdot\tau\sin\left(\frac{|\theta|}{2}\right) \tag{3.29}$$

Where $\hat{\theta}$ is the unit vector in the θ direction. Note that if $|\theta| = 2\pi$, then $R = -1$. A 2π rotation changes the sign, but a 4π rotation is required to get back to the identity.

The representation $\left(0,\frac{1}{2}\right)$ is obtained by choosing generators such that

$$\mathbf{K}^{(+)} = 0 \quad ; \quad \mathbf{K}^{(-)} = -\frac{1}{2}\tau^* \tag{3.30}$$

The transformation law is given by

$$M'^{\,B'}_{A} = \exp\left(\frac{i}{2}\alpha^*\cdot\tau^*\right)_{A'}{}^{B'} \tag{3.31}$$

which is the complex conjugate of the matrix M defined previously. Vectors (spinors) are written as (e.g.) $\chi_{A'}$, the primed index being used to make it clear that the transformation property is different. The choice of generators in the $\left(0,\frac{1}{2}\right)$ case is such that $\chi_{A'} = (\psi_A)^*$ is an invariant equation. We can use this fact to define a vector transforming as the $\left(0,\frac{1}{2}\right)$, $\bar{\psi}$, from one of $\left(\frac{1}{2},0\right)$, ψ, i.e.

$$\bar{\psi}_{A'} = (\psi_A)^* \tag{3.32}$$

The antisymmetric tensors ϵ_{AB}, ϵ^{AB}, $\epsilon_{A'B'}$ and $\epsilon^{A'B'}$ are by definition invariants of SL(2,C) (SL(2,C) is unimodular, so volumes are preserved, hence ϵ is preserved). These can be used to raise and lower indices. If we define index raising by

$$\psi^A = \epsilon^{AB}\psi_B \tag{3.33}$$

then the lowering operation must be defined by

$$\psi_B = \psi^A\epsilon_{AB} \tag{3.34}$$

in order that we should get back to the same spinor. This is necessary because $\epsilon_{AB} = \epsilon^{AB}$ as may be discovered by using ϵ to raise and lower its own indices. I shall use the convention $\epsilon_{12} = \epsilon^{12} = \epsilon_{1'2'} = \epsilon^{1'2'} = -1$. Similarly, we have

$$\bar{\psi}^{A'} = \epsilon^{A'B'}\bar{\psi}_{B'} \quad \text{and} \quad \bar{\psi}_{B'} = \bar{\psi}^{A'}\epsilon_{A'B'} \tag{3.35}$$

for raising and lowering primed indices. ϵ has the properties

$$\epsilon_{AB}\,\epsilon^{AC} = \epsilon_{BA}\,\epsilon^{CA} = \delta^C_B \tag{3.36}$$

$$\text{and} \quad \epsilon_{AB}\,\epsilon^{CD} = \delta^C_A\,\delta^D_B - \delta^D_A\,\delta^C_B \tag{3.37}$$

and similarly for primed indices. I shall refer to unprimed indices as *left-handed* (LH) and primed indices as *right-handed* (RH). This is because massless particle states with these transformation properties have helicities which are respectively left-handed and right-handed).

Any quantity transforming as the representation $\left(\frac{m}{2}, \frac{n}{2}\right)$ of SL(2,C) can be written as a symmetric SL(2,C) spinor

$$\psi_{(ABC\ldots)(A'B'C'\ldots)} \tag{3.38}$$

which has m LH and n RH indices (we just apply the laws of coupling angular momenta to the pseudo-angular momenta $\mathbf{K}^{(\pm)}$).

[Note: we use brackets to indicate symmetrisation or antisymmetrisation of indices:

$$\psi_{(\alpha_1\alpha_2\ldots\alpha_n)} = \frac{1}{n!}\left(\psi_{\alpha_1\alpha_2\ldots\alpha_n} + \psi_{\alpha_2\alpha_1\ldots\alpha_n} + \text{permutations}\right) \tag{3.39}$$

$$\psi_{[\alpha_1\alpha_2\ldots\alpha_n]} = \frac{1}{n!}\left(\psi_{\alpha_1\alpha_2\ldots\alpha_n} - \psi_{\alpha_2\alpha_1\ldots\alpha_n} + \text{even perms.} - \text{odd perms.}\right) \tag{3.40}$$

Vertical bars ("|") are used to exempt indices from (anti) symmetrisation. If there are different types of index in the field then the (anti) symmetrisation is over the type given by the first one in the field.]

Since the four vector is the $\left(\frac{1}{2}, \frac{1}{2}\right)$ of SL(2,C), we can write

$$V_{AA'} = \sigma^a_{AA'}V_a \tag{3.41}$$

where $\sigma^a_{AA'}$ are a set of numerical coefficients we use for setting up the bispinor in terms of the four vector. These are often called the Infeld-van der Waarden symbols. The coefficients must obey

$$\sigma^a_{AA'} = M_A{}^B M^*_{A'}{}^{B'}\Lambda^a{}_b\sigma^b_{BB'} \tag{3.42}$$

where M and Λ refer to the same Lorentz transformation, i.e.

$$\Lambda^a{}_b = \exp\left(-\frac{i}{2}\omega_{cd}S^{cd}\right)^a{}_b \quad \text{and} \quad M_A{}^B = \exp\left(-\frac{i}{2}\omega_{cd}J^{cd}\right)_A{}^B \tag{3.43}$$

where $J^{0i} = -\frac{i}{2}\tau_i$ and $J^{ij} = \frac{1}{2}\epsilon_{ijk}\tau_k$ are the generators of the spinor representation. One can solve this to obtain

$$\sigma^0_{AA'} = k\,\delta_{AA'} \;\; ; \;\; \sigma^i_{AA'} = k\,\tau^i_{AA'} \tag{3.44}$$

where k is arbitrary. We will choose it to be real so that a real four vector corresponds to a Hermitian bispinor, and will further choose $k = 1/\sqrt{2}$, i.e.

$$\sigma^a_{AA'} = \frac{1}{\sqrt{2}}\,(1,\tau)^a_{AA'} \tag{3.45}$$

This is because we will now get the relation

$$\eta_{ab}\,\sigma^a_{AA'}\,\sigma^b_{BB'} = \epsilon_{AB}\,\epsilon_{A'B'} \tag{3.46}$$

which enables us to instantly translate expressions involving Lorentz indices into their spinor equivalents, without requiring extra numerical factors. E.g.

$$V_{AA'}W^{AA'} = \epsilon_{AB}\,\epsilon_{A'B'}V^{AA'}W^{BB'} = \eta_{ab}\sigma^a_{AA'}\sigma^b_{BB'}V^{AA'}W^{BB'} \tag{3.47}$$

$$= \eta_{ab}V^aW^b = V_aW^a \tag{3.48}$$

The rule is that each Lorentz index is replaced by a pair of spinor indices. We can introduce an "abstract index notation" wherein a mismatch of types of indices implies the presence of the σ symbols. E.g.

$$V_{ab} = W_{AA'BB'} \tag{3.49}$$

really means

$$\sigma^a_{AA'}\sigma^b_{BB'}V_{ab} = W_{AA'BB'} \tag{3.50}$$

With this notation, we can discover that

$$\epsilon_{ab}{}^{cd} = i\left(\delta^C_A\delta^D_B\delta^{D'}_{A'}\delta^{C'}_{B'} - \delta^D_A\delta^C_B\delta^{C'}_{A'}\delta^{D'}_{B'}\right) \tag{3.51}$$

(our ϵ convention is that $\epsilon_{0123} = 1$. The indices are raised and lowered using η). Evidently every equation which involves Lorentz indices has its spinor equivalent (although the converse is not true), and any irreducible representation of the Lorentz group can be classified in terms of SL(2,C). One important one is $(1,0)$.

In O(3,1) (Lorentz group) language, this representation is the one that self-dual antisymmetric tensors belong to, i.e. tensors $T_{[ab]}$ satisfying

$$T_{ab} = \frac{i}{2}\epsilon_{abcd}T^{cd} \tag{3.52}$$

This tensor has its spinor equivalent $\phi_{(AB)}$ given by

$$T_{ab} = \phi_{(AB)}\epsilon_{A'B'} \tag{3.53}$$

We discover this by considering $T_{ab} = T_{ABA'B'}$

$$= T_{(AB)(A'B')} + T_{[AB][A'B']} + T_{[AB](A'B')} + T_{(AB)[A'B']} \tag{3.54}$$

$$= T_{[AB](A'B')} + T_{(AB)[A'B']} \tag{3.55}$$

since this is antisymmetric with respect to the exchange of a and b.

$$= \bar{\chi}_{(A'B')}\epsilon_{AB} + \phi_{(AB)}\epsilon_{A'B'} \tag{3.56}$$

for some tensors $\bar{\chi}$ and ϕ, where we have used the fact that any quantity of the kind $\psi_{[AB]}$ is proportional to ϵ_{AB}.

Using the formula for ϵ_{abcd} in terms of two-component spinors, we find that

$${}^*T_{ab} = \frac{1}{2}\epsilon_{abcd}T^{cd} = i\bar{\chi}_{(A'B')}\epsilon_{AB} - i\phi_{(AB)}\epsilon_{A'B'} \tag{3.57}$$

Hence

$$T_{ab} = i\,{}^*T_{ab} \;\Rightarrow\; T_{ab} = \phi_{(AB)}\epsilon_{A'B'}. \tag{3.58}$$

If we define the symbols $\sigma^{ab}{}_{CD} = \sigma^{[ab]}{}_{CD} = \sigma^{ab}{}_{(CD)}$ by

$$\phi_{CD} = \frac{1}{2}\sigma^{ab}{}_{CD}\,T_{ab} \tag{3.59}$$

then we find that $J^{ab}{}_C{}^D = i\sigma^{ab}{}_C{}^D$ where $J^{ab}{}_C{}^D$ are the generators of the group SL(2,C). The adjoint representation of the Lorentz group or SL(2,C) is classifiable as $(1,0)\oplus(0,1)$ although each part of this direct sum contains all the information since they are related by Hermitian conjugation.

It is easy to establish that

$$\sigma^{ab}{}_C{}^D = \frac{1}{2}\left(\sigma^a_{CC'}\sigma^{bDC'} - \sigma^b_{CC'}\sigma^{aDC'}\right) \tag{3.60}$$

Also, using "abstract index notation" :

$$(\sigma_{ab})_C{}^D = \frac{1}{2}\epsilon_{A'B'}\left(\epsilon_{AC}\,\delta^D_B + \epsilon_{BC}\,\delta^D_A\right) \tag{3.61}$$

This is one of many identities that can be derived from the σ symbols. Another useful one is

$$\eta^{ab}\delta^C_D = \sigma^a_{DD'}\sigma^{bCD'} + \sigma^a_{DD'}\sigma^{bCD'} \tag{3.62}$$

The SL(2,C) ϵ conventions here follow Penrose. A different set of conventions is, however, in more widespread use amongst physicists, but these should be avoided for reasons given here.

3.2 Parity, time reversal and spacetime reversal

The group SL(2,C) is connected and so does not contain the discrete operations of parity, time reversal and spacetime reversal which are part of the full Lorentz group. However, it is possible to have representations of SL(2,C) where these operations are defined. Consider the parity operation: from the group theory

$$P\,\mathbf{K}^{(\pm)}P = \mathbf{K}^{(\mp)} \tag{3.63}$$

Therefore the effect of P on a representation of the type $\left(j^{(+)}, j^{(-)}\right)$ is to turn it into one of type $\left(j^{(-)}, j^{(+)}\right)$. Thus representations of the type (j,j) are the only irreducible representations which may contain the parity operation. Such a representation is the symmetric, traceless part of $2j$ four-vector representations and so it is easy to see that parity (and indeed all the discrete operations) is implemented. Otherwise we need to abandon irreducibility when we include P, such as in the Dirac spinor representation, which is the direct sum $\left(\frac{1}{2},0\right)\oplus\left(0,\frac{1}{2}\right)$.

3.3 The unitary irreducible representations of the Poincaré group

As discovered earlier (3.9)-(3.11), the algebra of the Poincaré group is

$$\left[M^{ab}, M_{cd}\right] = 4i\delta^{[a}{}_{[c}M_{d]}{}^{b]} \tag{3.64}$$

$$\left[M^{ab}, P^c\right] = -2i\eta^{c[a}P^{b]} \tag{3.65}$$

$$\left[P^a, P^b\right] = 0 \tag{3.66}$$

In a unitary representation, these will be Hermitian operators. This enables us to diagonalise as many of these as will commute with each other. Thus we can make all of P^a diagonal. The eigenvalues p^a then are continuous and unbounded so the eigenvectors belong to a (non-denumerably) infinite-dimensional vector space.

Having done this, we notice that under a displacement

$$M^{ab} \to e^{-id\cdot P}M^{ab}e^{id\cdot P} = M^{ab} + d^aP^b - d^bP^a \tag{3.67}$$

whereas something that commutes with P^a must be displacement invariant. So we project out the displacement-invariant part $P^{[a}M^{bc]}$ which is conveniently embodied in the Pauli-Lubanski vector

$$W_a = \frac{1}{2}\epsilon_{abcd}P^bM^{cd} \tag{3.68}$$

Thus

$$\left[P^a, P^b\right] = \left[P^a, W^b\right] = 0. \tag{3.69}$$

However the components of W^a do not commute with each other. We find that

$$[W_a, W_b] = i\epsilon_{abcd} P^c W^d \tag{3.70}$$

which will prevent us from making all the components diagonal. W_a are the generators of the so-called "little group", which is the group of Lorentz transformations and rotations that preserve p^a. Since $P_a W^a = 0$, only three are independent.

To form irreducible representations we use the labels given by the eigenvalues of $P^2 = P_a P^a$ and $W^2 = W_a W^a$ which are Casimir operators of the Poincaré group. As we will see, the spectrum of eigenvalues of W^2 depends on the eigenvalue of P^2.

In the following we will only consider the case of finite-dimensional spin, i.e. that where the states of given p^a are not infinitely degenerate.

Case (i): $p^2 = m^2 > 0$

Without loss of generality, we can choose a basis where the eigenvalues of P^a are

$$p^a = (m, 0, 0, 0) \tag{3.71}$$

since this will always be accessible from another by a linear transformation, namely a Lorentz transformation. In this case $W_0 = 0$ but

$$\left[W^i, W^j\right] = im\,\epsilon^{ijk} W^k \tag{3.72}$$

so the W's generate the group SU(2). Irreducibility requires $W^2 = -W^i W^i$ to be definite, so it must have the value $-m^2 s(s+1)$ in which case the eigenvectors of a given p^a form a $(2s+1)$-dimensional subspace which can be labelled with $J_3 = -W^3/m$ which has eigenvalues $-s, -s+1, \cdots, s-1, s$. Evidently $2s$ is an integer, and s is called the *spin*.

The spin vector can also be written as a symmetric $2s$-index SU(2) tensor, i.e.

$$|p;{}_{\alpha\beta\gamma\cdots}\rangle = |p;{}_{(\alpha\beta\gamma\cdots)}\rangle \qquad \alpha, \beta, \gamma, \cdots = 1, 2 \tag{3.73}$$

Under a Lorentz transformation

$$U(\Lambda)|p;{}_{\alpha\beta\gamma\cdots}\rangle = |\Lambda p;{}_{\delta\epsilon\zeta\cdots}\rangle D^\delta{}_\alpha(\Lambda) D^\epsilon{}_\beta(\Lambda) D^\zeta{}_\gamma(\Lambda) \cdots \tag{3.74}$$

Evidently the matrices $D^\alpha{}_\beta(\Lambda)$ satisfy

$$D(\Lambda) D(\Lambda') = D(\Lambda\Lambda') \tag{3.75}$$

In other words, the spinor indices form a representation not merely of the little group, but also of the Lorentz group as a whole. The rotation generators are given by

$$\mathbf{J} = \frac{1}{2}\tau \tag{3.76}$$

The boost generators are then to be obtained by solving (3.14) and (3.15). This gives

$$\mathbf{N} = \pm\frac{i}{2}\tau \tag{3.77}$$

If we choose the top sign, the Lorentz transformation is

$$\exp\Bigl(-\frac{1}{2} i(\theta - i\beta)\cdot\tau\Bigr) = \exp\Bigl(-\frac{i}{2}\alpha\cdot\tau\Bigr) \tag{3.78}$$

which means that the SU(2) index is also a covariant LH index of SL(2,C). With the top sign, we have

$$\exp\Big(-\frac{1}{2}i(\theta+i\beta).\tau\Big)=\exp\Big(-\frac{1}{2}i\alpha^*\cdot\tau\Big)$$
$$=\Big[\exp(-\frac{1}{2}i\alpha^*\cdot\tau^T)\Big]^T=\Big[\Big(\exp\Big(\frac{i}{2}\alpha^*\cdot\tau^T\Big)\Big)^{-1}\Big]^T$$
$$=\Big(\Big(\exp\Big(-\frac{1}{2}i\alpha\cdot\tau\Big)^*\Big)^{-1}\Big)^T \tag{3.79}$$

which is the transformation of a contravariant RH index.

A state of 3-momentum $\mathbf{p}$ can be accessed from the rest frame with the boost

$$M=\exp\Big(-i\tanh^{-1}\Big(\frac{p}{E}\Big)\hat{\mathbf{p}}\cdot\mathbf{N}\Big) \tag{3.80}$$

where $E=\sqrt{\mathbf{p}^2+m^2}$ and $\hat{\mathbf{p}}$ is a unit vector in the $\mathbf{p}$ direction. Calculating this for the fundamental representation of SL(2,C) gives

$$M=\frac{1}{\sqrt{2m(E+m)}}(E+m-\mathbf{p}\cdot\tau) \tag{3.81}$$

We may choose normalisations in the rest frame that are SU(2) covariant, i.e.:

$$\langle{}^\beta|{}_\alpha\rangle\ \propto\ \delta^\beta_\alpha \tag{3.82}$$

The vector $\langle{}^\beta|$ transforms as the complex conjugate, which is also the dual for SU(2), but not for SL(2,C). Applying the above boost out of the rest frame, we have

$$\langle{}^\beta|{}_\alpha\rangle\ \to M^{*\beta}{}_\gamma M_\alpha{}^\delta\langle{}^\gamma|{}_\delta\rangle$$
$$\propto M_\alpha{}^\gamma M^{*\beta}{}_\gamma=(MM^\dagger)_\alpha{}^\beta \tag{3.83}$$

Now, $MM^\dagger=\frac{\sqrt{2}}{m}p^a(\sigma_a)_{AA'}$. In general, therefore, the normalisation must be

$$\langle{}_{A'}|{}_A\rangle\propto p_{AA'} \tag{3.84}$$

Which for spin s is

$$\langle{}_{A'B'C'\cdots}|{}_{ABC\cdots}\rangle\propto p_{(AA'}p_{BB'}p_{CC'}\cdots{}_{)} \tag{3.85}$$

with $2s$ indices of each type.

The case of the covariant SU(2) index also being a contravariant RH SL(2,C) index is dealt with by noting that this can be written as

$$|{}^{A'}\rangle=\frac{\sqrt{2}}{m}p^{AA'}|{}_A\rangle \tag{3.86}$$

since, in the rest frame, $\frac{\sqrt{2}}{m}p^{AA'}$ is the Kronecker delta. The tensor $p^{AA'}$ can thus be used to exchange LH and RH indices, and the requirement that the purely LH or purely RH tensor is symmetric translates into the requirement that

$$p^{AA'}|{}_{ABC\cdots A'B'C'\cdots}\rangle=0 \tag{3.87}$$

for a mixed tensor. The most general spin s tensor is thus a $2s$-index SL(2,C) tensor, symmetric in LH and RH indices, and subject to (3.87) when both LH and RH indices appear. These tensors are all equivalent as one can use the momentum matrix to switch between them.

Case (ii): $p^2=0$

We cannot do as in (i) because the basis required would have $p^a=0$ which cannot be accessed from $p^a\neq 0$ by a Lorentz transformation, which is necessarily invertible. So let us take

$$p^a=(\rho,0,0,\rho) \tag{3.88}$$

($\rho \neq 0$) which will follow from an appropriate linear transformation. Then W_1 and W_2 are non-compact generators (they contain boosts) and $W_0 = -W_3 = \rho J = \rho M^{12}$ with the algebra

$$[J, W_1] = iW_2 \ ; \ [J, W_2] = -iW_1 \ ; \ [W_1, W_2] = 0 \tag{3.89}$$

The eigenvalues of W_1 and W_2 are continuous and unbounded, so finite-dimensional spin and unitarity imply that their action on the vector space is zero. J on the other hand is a rotation operator and so generates a compact U(1) group. Thus we may have vectors which satisfy

$$J\,|p;s\rangle = s\,|p;s\rangle \tag{3.90}$$

where $2s$ is an integer. s is then called the *helicity.* Half-integral values are allowed because J, with M^{23} and M^{31}, generates the group SU(2). Note that, unlike the massive case, there are no extra degrees of freedom.

Now $W_a = (\rho J, 0, 0, -\rho J)$ when acting on the representation space so a covariant statement of the helicity constraint is

$$(W^a - sP^a)\,|p;s\rangle = 0 \tag{3.91}$$

Let us try to see this in terms of SL(2,C) indices, assigning the most general irreducible form of a tensor with N_L LH indices and N_R RH ones, symmetric in each set. The Pauli-Lubanski vector acting on a single LH index takes the form

$$\begin{aligned}(W_a)_X{}^Y &= \frac{1}{2}\epsilon_{abcd}P^b\left(M^{cd}\right)_X{}^Y = \frac{i}{2}\epsilon_{abcd}P^b\left(\sigma^{cd}\right)_X{}^Y \\ &= P_b\left(\sigma_a{}^b\right)_X{}^Y \\ &= -p_{XA'}\delta^Y_A + \frac{1}{2}p_{AA'}\delta^Y_X\end{aligned} \tag{3.92}$$

where we have used the self-duality of σ, replaced P_a with its eigenvalue and used the identity

$$\left(\sigma_a{}^b\right)_X{}^Y = \delta^{B'}_{A'}\left(-\delta^Y_A\delta^B_X + \frac{1}{2}\delta^B_A\delta^Y_X\right) \tag{3.93}$$

Using the analogous expression for RH indices, the effect on a symmetric tensor with N_L left-handed SL(2,C) indices and N_R right-handed ones is thus:

$$\begin{aligned}W_m|_{ABC\cdots D'E'F'\cdots}\rangle = &-N_L p_{(AM'}|_{BC\cdots)MD'E'F'\cdots}\rangle + N_R p_{M(D'}|_{ABC\cdots E'F'\cdots)M'}\rangle + \\ &\frac{1}{2}\left(N_L - N_R\right)p_m|_{ABC\cdots D'E'F'\cdots}\rangle\end{aligned} \tag{3.94}$$

Which gives

$$-N_L\,p_{(AM'}|_{BC\cdots)MD'E'F'\cdots}\rangle + N_R\,p_{M(D'}|_{ABC\cdots E'F'\cdots)M'}\rangle + \left(\frac{1}{2}\left(N_L - N_R\right) - s\right)p_m|_{ABC\cdots D'E'F'\cdots}\rangle = 0 \tag{3.95}$$

We may now use the fact that if p_m is a null vector then $p_m = \bar{\phi}_{M'}\phi_M$ for some spinor ϕ_M. Applying ϕ^M to the above equation, and using the fact that

$$\phi_{(A}T_{BC\cdots)} = 0 \tag{3.96}$$

for non-zero ϕ_A implies that

$$T_{(BC\cdots)} = 0 \tag{3.97}$$

We obtain

$$\phi^A|_{ABC\cdots D'E'F'\cdots}\rangle = 0 \tag{3.98}$$

The contraction is equivalent to antisymmetrisation in the two-component world, to which the only solution is that the LH part of the tensor is proportional to a product of ϕ_A spinors. A contraction of (3.95) with $\bar{\phi}^{M'}$ leads to a similar conclusion for RH indices. Thus

$$|_{ABC\cdots D'E'F'\cdots}\rangle = \phi_A \phi_B \phi_C \cdots \bar{\phi}_{D'} \bar{\phi}_{E'} \bar{\phi}_{F'} \cdots |\,\rangle \tag{3.99}$$

which may be substituted back into (3.95) to give the helicity in terms of number of LH and RH indices:

$$s = \frac{1}{2}\left(N_R - N_L\right) \tag{3.100}$$

Note that in the massless case, mixed tensors are just momentum matrices multiplied by purely left- or right-handed tensors. The normalisation of LH tensors therefore must be the same as in the massive case (3.85).

When $s \neq 0$ a rotation around the momentum axis will change the phase of the state, which seems to be in conflict with equation (3.99) where the tensor is expressed in terms of constant ϕ_A spinors and a Lorentz scalar until one realises that these spinors are arbitrary to the extent of a phase. To properly capture the behaviour under a Lorentz transformation one should therefore avoid these, writing (e.g.) the massless LH tensor simply as a symmetric $2s$-index tensor satisfying

$$p^{AA'}|_{ABC\cdots}\rangle = 0 \tag{3.101}$$

Case (iii): $p^2 = -\rho^2 < 0$.

Transform to a frame where the eigenvalues of P_a are

$$p^a = (0, 0, 0, \rho) \tag{3.102}$$

Then we find that

$$W^a = (\rho J_3, \rho N_2, -\rho N_1, 0) \tag{3.103}$$

The algebra is

$$[W_0, W_1] = i\rho W_2 \tag{3.104}$$
$$[W_0, W_2] = -i\rho W_1 \tag{3.105}$$
$$[W_1, W_2] = -i\rho W_0 \tag{3.106}$$

W_1 and W_2 are boosts in a finite-dimensional unitary representation of the Poincaré group and so must act trivially on the vector space. Equation (3.106) then determines that the helicity generator must also act trivially. Thus, tachyonic states can have no spin.

3.4 Normalisations of irreducible representations

In this representation, the four-momentum operator is Hermitian. Eigenvectors with different eigenvalues must be therefore be orthogonal. We see this as follows:

$$\langle q; s'|P_a|p; s\rangle = q_a \langle q; s'|p; s\rangle = p_a \langle q; s'|p; s\rangle \tag{3.107}$$

Applying P_a both to the left and the right. So

$$\left(q_a - p_a\right)\langle q; s'|p; s\rangle = 0 \tag{3.108}$$

We can now use the Dirac result

$$x\, f\left(x\right) = 0 \Longleftrightarrow f\left(x\right) = \lambda\, \delta\left(x\right) \tag{3.109}$$

(for some constant λ) to obtain

$$\langle q; s'|p; s\rangle = \delta\left(q-p\right)\, f_0\left(p, s, s'\right) \tag{3.110}$$

For some function f_0. We may also note that

$$\langle q; s'|\left(P^2-m^2\right)|p; s\rangle = \left(p^2-m^2\right)\langle q; s'|p; s\rangle = 0 \tag{3.111}$$

From which we may further specify

$$\langle q; s'|p; s\rangle = \delta\left(p^2-m^2\right)\delta\left(q-p\right)\, f\left(p, s, s'\right) \tag{3.112}$$

for some arbitrary function $f\left(p, s, s'\right)$.

The spin labels s and s' are in fact SL(2,C) tensor structures.

Thus the normalisation condition, which applies to all finite-dimensional spin unitary representations of the Poincaré group is:

$$\langle q;_{A'B'C'\cdots}|p;_{ABC\cdots}\rangle = \delta\left(p-q\right)\delta\left(p^2-m^2\right)p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)}\, f\left(p\right) \tag{3.113}$$

In the case of tachyonic states, there can be no SL(2,C) tensor structure, and $f(p)$ must be positive definite.

In the case of non-tachyonic states (i.e. states with positive or zero m^2), if they have negative energy and half-integral spin then $f(p)$ must be negative definite. Otherwise it is positive definite. Lorentz invariance requires it to be an invariant function, and although there is no requirement to choose either 1 or -1, it is convenient to do so. In other words

$$\langle q;_{A'B'C'\cdots}|p;_{ABC\cdots}\rangle = (-1)^{2s\cdot\theta(-p_0)}\ \delta(p-q)\delta(p^2-m^2)p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)} \tag{3.114}$$

where s is the spin and the Heaviside step function is defined by

$$\theta(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases} \tag{3.115}$$

defines the normalisation of a standard set of basis of vectors in an irreducible unitary representation of the Poincaré group. It must be stressed that we have said nothing about these vectors other than their normalisation and their properties under the action of the Poincaré group. They may be fundamental particle states, or they may be composites. If the whole universe, for example, was characterised by a definite mass and spin then it too could be represented as one of these vectors.

One of our requirements was that the states should have positive norm. That this is always the case when using the Wigner rotation is clear, but seeing it here requires a little more effort. For $s > 0$ a necessary and sufficient condition is that the eigenvalues of the momentum matrix $p_{AA'}$ should have the same sign. If this is the case, then states of positive norm can always be constructed (both-negative eigenvalues can of course be made positive by an overall sign). The eigenvalues of the momentum matrix are $\frac{1}{\sqrt{2}}(p^0 \pm |\mathbf{p}|)$, so we see that in the case of $p^2 < 0$ we have $|p^0| < |\mathbf{p}|$ forcing at least one of the corresponding vectors to have negative norm. This rules out spin for so-called "tachyonic" states. In the case of $p^2 = 0$ one of the eigenvalues will be zero. This does not matter as, owing to the additional helicity constraint, there is no corresponding vector. In the massive case we may transform to the rest frame, where $p_{AA'} = \frac{m}{\sqrt{2}}\delta_{AA'}$, which demonstrates positive norm explicitly.

The mass-shell delta function can be extracted, so for $m^2 \geq 0$ we can write

$$|p;_{ABC\cdots}\rangle = \theta(\pm p_0)\delta(p^2-m^2)|p;_{ABC\cdots}\rangle' \tag{3.116}$$

where p is implicitly on shell in the primed states. The normalisation of these states is then

$$\langle q;_{A'B'C'\cdots}{}'|p;_{ABC\cdots}\rangle' = \delta_m(p,q)p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)}|_{p^0\text{ on shell}} \tag{3.117}$$

where

$$\delta_m(p,q) = 2\sqrt{\mathbf{p}^2 + m^2}\ \delta(\mathbf{p} - \mathbf{q}) \tag{3.118}$$

From this follows the "completeness" relation:

$$1 = \left(\frac{2}{m^2}\right)^N \int dm(p) |p;_{ABC\cdots}\rangle' p^{AA'} p^{BB'} p^{CC'} \cdots \langle p;_{A'B'C'\cdots}|' \tag{3.119}$$

where

$$dm(p) = \frac{d^3\mathbf{p}}{2\sqrt{\mathbf{p}^2 + m^2}} \tag{3.120}$$

is an invariant measure. The case of $m = 0$ is treated by first factoring out ϕ_A spinors and then using the same arguments as for spin zero.

It is often convenient to use a single index to enumerate the $N + 1$ independent components of the symmetric SL(2,C) spin tensor. Accordingly, we may define an index α which is such that $\alpha = 1, 2, 3, \cdots$ corresponds to $ABC \cdots = 222\cdots, 122\cdots, 112\cdots, \cdots$. We then can write

$$\langle q;_{\alpha'}{}' | p;_{\alpha}\rangle' = \delta_m(p,q) N_{\alpha\alpha'} \tag{3.121}$$

and

$$1 = \int dm(p) |p;_{\alpha}\rangle' T^{\alpha\alpha'} \langle p;_{\alpha'}|' \tag{3.122}$$

where $N_{\alpha\alpha'}$ and $T^{\alpha\alpha'}$ are representations of the symmetrised products of momentum tensors. These will have the property

$$N_{\alpha\beta'} T^{\alpha\alpha'} = \delta^{\alpha'}_{\beta'} \tag{3.123}$$

4. Fock space

The vectors defined as unitary irreducible representations of the Poincaré group are the basis of relativistic quantum mechanics. Multi-particle states are formed as tensor products of these vectors. Such composite vector spaces may include different species of particle, multiple particles of the same kind, or some combination.

In the case of identical particles, we will wish to incorporate the principle of *indistinguishability.*

If we characterise a vector in the product space of two identical particles as $|12\rangle$, where "1" and "2" are shorthand for the momentum, spin and other labels, then the principle of indistinguishability requires that the state $|21\rangle$ gets treated in the same way as $|12\rangle$.

To deal with this, we may define the *exchange operator* X_{12} as the operator that switches the identities of the particles. It is easy to see that the operator can be defined to be both linear and Hermitian. Also, since $X_{12}^2 = 1$ it follows that the eigenvalues are +1 and -1. The eigenvectors are then

$$\begin{aligned} |S\rangle &= |12\rangle + |21\rangle \\ |A\rangle &= |12\rangle - |21\rangle \end{aligned} \tag{4.1}$$

The state $|S\rangle$, which is invariant under particle exchange, demonstrates Bose-Einstein statistics, whereas $|A\rangle$, which changes sign as a result of particle exchange, demonstrates Fermi-Dirac statistics.

Extension of these arguments to three particles reveals that mixed statistics are impossible. Consider $X_{12}X_{23}$. This performs a single cyclic permutation of three particles, so three applications should bring us back to our starting point. In other words,

$$X_{12}X_{23}X_{12}X_{23}X_{12}X_{23} \tag{4.2}$$

is the identity. A state that is antisymmetric under $1 \leftrightarrow 2$ exchange, but symmetric under $2 \leftrightarrow 3$ exchange however will change sign under the action of the operator ($-1 = -1 \times 1 \times -1 \times 1 \times -1 \times 1$). Since this contradicts the need for the state to remain the same, we are forced to conclude that no such state can exist.

Thus, three-particle eigenstates of the exchange operators are either totally symmetric or totally anti-symmetric, a result that extends to all particle numbers. Note that when we have more than two particles, restricting ourselves to exchange-symmetric eigenstates also means ignoring parts of the vector space. For example, with three particles, we may form only two exchange eigenvectors from the six states generated by particle exchanges.

The direct sum of single-particle product spaces for identical particles is known as a *Fock space*.

To understand Fock space, let us extend (3.117) to an n-particle product space. For the time being, we will consider only spin zero.

$$\langle q_1 q_2 \cdots q_n{}' | p_1 p_2 \cdots p_n \rangle' = \delta_m(p_1, q_1)\delta_m(p_2, q_2) \cdots \delta_m(p_n, q_n) \tag{4.3}$$

A state comprised of identical particles with Fermi-Dirac or Bose-Einstein statistics is formed as the sum of such vectors over all permutations of the momentum labels, i.e.

$$|p_1 p_2 \cdots p_n; \pm\rangle = \frac{1}{\sqrt{n!}} \sum_{\substack{\text{perms} \\ i_1, \cdots, i_n \in \\ 1, \cdots, n}} (\pm 1)^P \; |p_{i_1} p_{i_2} \cdots p_{i_n}\rangle' \tag{4.4}$$

(the normalisation factor here is chosen for later convenience). The top sign is for the case of Bose-Einstein statistics and the bottom sign for Fermi-Dirac. Here P is the number of exchanges required to get to the given permutation from the the original one (where the labels are in ascending order).

Now consider

$$\langle q_1 q_2 \cdots q_n; \pm | p_1 p_2 \cdots p_n; \pm\rangle$$

Each of the $n!$ permutations in $\langle q_1 q_2 \cdots q_n; \pm|$ will give the same contribution (one just permutes the labels $q_1, q_2 \cdots q_n$ and $p_1, p_2 \cdots p_n$ simultaneously), so we get

$$\sqrt{n!}\; \langle q_1 q_2 \cdots q_n|' \frac{1}{\sqrt{n!}} \sum_{\substack{\text{perms} \\ i_1 \cdots i_n \in \\ 1 \cdots n}} (\pm)^P |p_{i_1} p_{i_2} \cdots p_{i_n}\rangle' = \sum_{\substack{\text{perms} \\ i_1 \cdots i_n \in \\ 1 \cdots n}} (\pm 1)^P \; \delta_m(q_1, p_{i_1})\delta_m(q_2, p_{i_2}) \cdots \delta_m(q_n, p_{i_n}) \tag{4.5}$$

We are now in a position to define the *creation* operator:

$$\begin{aligned} a^\dagger(p) &= \sum_{n=0}^{\infty} \frac{1}{n!} \int |p p_1 \cdots p_n; \pm\rangle dm(p_1) \cdots dm(p_n) \langle p_1 \cdots p_n; \pm| \\ &= |p\rangle\langle| + \int |p p_1; \pm\rangle dm(p_1) \langle p_1| + \frac{1}{2} \int |p p_1 p_2; \pm\rangle dm(p_1) dm(p_2) \langle p_1 p_2; \pm| + \cdots \end{aligned} \tag{4.6}$$

As the name suggests, this operator has the effect of adding a particle of momentum p to the state. This is done in such a way as to preserve the symmetry or antisymmetry. From this definition, it follows that

$$a^\dagger(p)|p_1 \cdots p_n; \pm\rangle = |p p_1 \cdots p_n; \pm\rangle \tag{4.7}$$

Note that

$$a^\dagger(p)a^\dagger(q)|p_1 \cdots p_n; \pm\rangle = |p\; q\; p_1 \cdots p_n; \pm\rangle = \pm|q\; p\; p_1 \cdots p_n; \pm\rangle = \pm a^\dagger(q)a^\dagger(p)|p_1 \cdots p_n; \pm\rangle \tag{4.8}$$

Since this applies to any state $|p_1 \cdots p_n; \pm\rangle$, it follows that

$$[a^\dagger(p), a^\dagger(q)]_\mp = 0 \tag{4.9}$$

where we use the notation $[A, B]_- = [A, B] = AB - BA$ and $[A, B]_+ = \{A, B\} = AB + BA$.

The Hermitian adjoint is

$$a(p) = \sum_{n=0}^{\infty} \frac{1}{n!} \int |p_1 \cdots p_n; \pm\rangle dm(p_1) \cdots dm(p_n) \langle p p_1 \cdots p_n; \pm| \tag{4.10}$$

This is known as the *annihilation* operator as it will reduce the particle number of the state by one, giving zero in the case of the zero-particle (vacuum) state.

Consider

$$a(p)|p_1 \cdots p_n; \pm\rangle$$

Evidently the only term in the series not orthogonal to the state is that where the dual vector on the right has n particles. We therefore get

$$a(p)|p_1 \cdots p_n; \pm\rangle = \frac{1}{(n-1)!} \int |q_1 \cdots q_{n-1}; \pm\rangle dm(q_1) \cdots dm(q_{n-1}) \langle p\, q_1 \cdots q_{n-1}; \pm|p_1 \cdots p_n; \pm\rangle \tag{4.11}$$

Substituting (4.5) we then have

$$\begin{aligned}
&= \frac{1}{(n-1)!} \int |q_1 \cdots q_{n-1}; \pm\rangle dm(q_1) \cdots dm(q_{n-1}) \sum_{\substack{\text{perms} \\ i_1 \cdots i_n}} (\pm 1)^P \delta_m(p, p_{i_1}) \delta_m(q_1, p_{i_2}) \delta_m(q_2, p_{i_3}) \cdots \delta_m(q_{n-1}, p_{i_n}) \\
&= \frac{1}{(n-1)!} \sum_{\substack{\text{perms} \\ i_1 \cdots i_n}} (\pm 1)^P \delta_m(p, p_{i_1}) |p_{i_2} p_{i_3} \cdots p_{i_n}; \pm\rangle \\
&= \sum_{i=1}^{n} (\pm 1)^{i-1} \delta_m(p, p_i) |p_1 p_2 \cdots p_{i-1} p_{i+1} \cdots p_n; \pm\rangle
\end{aligned} \tag{4.12}$$

In the final step we are collecting together the $(n-1)!$ permutations associated with each distinct value of i_1. The factor of $(\pm)^{i-1}$ arises because $i-1$ exchanges are needed to get to the sequence $p_1, p_2, \cdots, p_n$ from $p_i, p_1, p_2, \cdots p_{i-1}, p_{i+1}, \cdots, p_n$.

Now consider

$$a^\dagger(q) a(p) |p_1 \cdots p_n; \pm\rangle = \sum_{i=1}^{n} (\pm 1)^{i-1} \delta_m(p, p_i) |q\, p_1 \cdots p_{i-1} p_{i+1} \cdots p_n\rangle \tag{4.13}$$

We also have

$$\begin{aligned}
a(p) a^\dagger(q) |p_1 \cdots p_n; \pm\rangle &= a(p) |q\, p_1 \cdots p_n; \pm\rangle \\
&= \delta_m(p, q) |p_1 \cdots p_n\rangle \pm \sum_{i=1}^{n} (\pm)^{i-1} \delta_m(p, p_i) |q\, p_1 \cdots p_{i-1} p_{i+1} \cdots p_n\rangle
\end{aligned} \tag{4.14}$$

Thus

$$[a(p), a^\dagger(q)]_\mp |p_1 \cdots p_n; \pm\rangle = \delta_m(p, q) |p_1 \cdots p_n; \pm\rangle \tag{4.15}$$

This applies to all states, hence

$$[a(p), a^\dagger(q)]_\mp = \delta_m(p, q) \tag{4.16}$$

The same arguments apply when the spin is greater than zero. The arguments are most easily developed using the single-index spin label ((3.121)-(3.123)), and we find that (4.9) and (4.16) generalise to give

$$[a_{A'B'C'\cdots}(p), a_{K'L'M'\cdots}(q)]_\mp = 0 \tag{4.17}$$

and

$$[a_{A'B'C'\cdots}(p), a^\dagger_{ABC\cdots}(q)]_\mp = \delta_m(p, q) p_{A(A'} p_{BB'} p_{CC'} \cdots{}_{)} \tag{4.18}$$

We may define

$$\Phi_{ABC\cdots}(p) = \theta(p_0)\delta(p^2 - m^2)a^\dagger_{ABC\cdots}(p) \tag{4.19}$$

In terms of these operators, the commutation relations are

$$[\Phi_{ABC\cdots}(p), \Phi_{KLM\cdots}(q)]_\mp = 0 \tag{4.20}$$

and

$$[\Phi^\dagger_{A'B'C'\cdots}(p), \Phi_{ABC\cdots}(q)]_\mp = \delta(p-q)\theta(p_0)\delta(p^2 - m^2)p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)} \tag{4.21}$$

5. The free relativistic particle

A vector with definite four-position x is going to have the transformation property

$$U(d, 0)|x\rangle = e^{-id.P}|x\rangle = |x + d\rangle \tag{5.1}$$

under spacetime displacement. The four-dimensional Fourier transform

$$|p\rangle = (2\pi)^{-4}\int d^4x\; e^{ip\cdot x}|x\rangle \tag{5.2}$$

will therefore be an eigenvector of the four-momentum operator with eigenvalue p, something we can identify with one of the vectors in a unitary irreducible representation of the Poincaré group developed earlier. So, allowing for a possible scaling function $f(p)$, and inverting the Fourier transform, we have

$$|x\rangle = \int d^4p\; f(p)\; e^{-ip\cdot x}|p\rangle \tag{5.3}$$

Considering only states within the forward light cone, the normalisation is then

$$\langle x'|x\rangle = \int d^4p\; \theta(p_0)\delta(p^2 - m^2)e^{-ip\cdot(x-x')}|f(p)|^2 = \int \frac{d^3\mathbf{p}}{2p_0}e^{-ip\cdot(x-x')}|f(p)|^2\bigg|_{p_0=E(\mathbf{p})} \tag{5.4}$$

where $E(\mathbf{p}) = \sqrt{\mathbf{p}^2 + m^2}$. The other requirement is the Lorentz transformation property

$$U(0, \Lambda)|x\rangle = e^{-\frac{i}{2}\omega_{ab}M^{ab}}|x\rangle = |\Lambda x\rangle \tag{5.5}$$

Applying this to (5.3), we find that $f(p) = f(\Lambda p)$, in other words, $f(p)$ must be an invariant function. The the only invariant we have to form this is p^2, which is a constant, so we are free to choose $f(p) = 1$ and the state of definite four-position is

$$|x\rangle = \int d^4p\; e^{-ip\cdot x}\;|p\rangle \tag{5.6}$$

which is unique up to a scaling factor. The points to note are, firstly, that the matrix element (5.4) is not necessarily zero when $x \neq x'$. There cannot therefore be a Hermitian four-position operator, since this requires eigenvectors with different eigenvalues to be orthogonal. This is as it should be, as having non-zero matrix elements for states at different times is what gives us a basis for calculating amplitudes for scattering and other processes in quantum mechanics. Secondly, the four-position states obey the Klein-Gordon equation. The group velocity of wave packets here is

$$\mathbf{v}_g = \frac{\partial E(\mathbf{p})}{\partial \mathbf{p}} = \frac{\mathbf{p}}{E(\mathbf{p})} \tag{5.7}$$

from which we may derive the expressions for the momentum and energy for the free relativistic particle in terms of its velocity.

Although there is no Hermitian four-position operator, we can nevertheless construct a Hermitian 3-position operator for a particular Lorentz frame by choosing $f(p) = (2\pi)^{-3/2}\sqrt{2E(\mathbf{p})}$ in (5.4). This gives

$$\langle \mathbf{x}'; t' | \mathbf{x}; t\rangle = \int \frac{d^3\mathbf{p}}{(2\pi)^3} e^{-iE(\mathbf{p})(t-t')+i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \tag{5.8}$$

At equal time, this is

$$\langle \mathbf{x}'; t | \mathbf{x}; t\rangle = \delta(\mathbf{x}-\mathbf{x}') \tag{5.9}$$

The 3-position operator is then

$$\hat{\mathbf{x}} = \int d^3\mathbf{x} |\mathbf{x}\rangle \mathbf{x} \langle \mathbf{x}| \tag{5.10}$$

which leads to the physical interpretation that for state $|\psi\rangle$

$$|\langle \mathbf{x}; t|\psi\rangle|^2 d^3\mathbf{x} \tag{5.11}$$

is the probability of finding the particle in the region $d^3\mathbf{x}$ around $\mathbf{x}$ at time t.

6. Quantum field theory; the spin-statistics theorem

Quantization is the process whereby a classical theory is converted into a quantum theory by the replacing of classical Poisson brackets with quantum commutators. Since Poisson brackets are an artefact of Lagrangian and Hamiltonian dynamics, the first task is to express the dynamical system in this form. This means identifying dynamical variables and then constructing a Lagrangian in terms of these such that the variational equations become the equation of motion. For a single particle, the dynamical variables will normally be the position and velocity at a given time, but for a classical field, there will be an infinite number, being the field amplitudes (and possibly also the first time derivative) at every point in space at a given time. The definition requires that the Poisson bracket of a dynamical variable with another at the same time is zero, a stricture on quantum field theory before we even begin to examine equations of motion, i.e.

$$[\Phi(t,\mathbf{x}), \Phi(t,\mathbf{x}')] = 0 \qquad \text{when} \qquad \mathbf{x} \neq \mathbf{x}' \tag{6.1}$$

Since we could choose any relativistic frame of reference to quantize, our requirement is therefore that field operators must commute for spacelike intervals.

The creation operators can be made functions of the spacetime co-ordinate in the same way as was done for states in equation (5.3), i.e. by Fourier transform. Using the operators of (4.19), we have

$$\Phi_{ABC\cdots}(x) = \int d^4p\, e^{-ip.x} \Phi_{ABC\cdots}(p) \tag{6.2}$$

As with the states, the construction is unique, apart from a scaling factor. The (anti)commutators of this with itself and its Hermitian adjoint are respectively

$$[\Phi_{ABC\cdots}(x), \Phi_{KLM\cdots}(x')]_{\pm} = 0 \tag{6.3}$$

and

$$[\Phi^{\dagger}_{A'B'C'\cdots}(x), \Phi_{ABC\cdots}(x')]_{\pm} = \int d^4p\, \theta(p_0)\delta(p^2-m^2) e^{ip\cdot(x-x')} p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)} \tag{6.4}$$

Since the latter does not in general vanish for spacelike intervals, we cannot use Φ as a quantum field. We may, however, form linear combinations of the fields and their Hermitian adjoints which commute for spacelike intervals.

Let us consider the massless case first. Write

$$\Phi_{ABC\cdots}(p) = \phi_A\phi_B\phi_C\cdots\Phi(p) \quad \text{and} \quad \Phi'_{A'B'C'\cdots}(p) = \phi^*_{A'}\phi^*_{B'}\phi^*_{C'}\cdots\Phi(p) \tag{6.5}$$

where $p_{AA'} = \phi_A\phi^*_{A'}$ and $\Phi(p)$ is a spinless field. Now form the combination

$$\tilde{\Phi}_{ABC\cdots}(x) = \Phi_{ABC\cdots}(x) + k\Phi'^{\dagger}_{ABC\cdots}(x) \tag{6.6}$$

for some constant k. Evidently

$$[\tilde{\Phi}^{\dagger}_{A'B'C'\cdots}(x), \tilde{\Phi}_{ABC\cdots}(x')]_{\pm} = \int d^4p\, e^{-ip\cdot(x-x')}\delta(p^2)p_{AA'}p_{BB'}p_{CC'}\cdots\{\theta(-p_0)(-1)^N \pm |k|^2\theta(p_0)\}$$

$$= (i\partial_{AA'})(i\partial_{BB'})(i\partial_{CC'})\cdots\int \frac{d^3\mathbf{p}}{2|\mathbf{p}|}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}\{(-1)^N e^{i|\mathbf{p}|(x_0-x'_0)} \pm |k|^2 e^{-i|\mathbf{p}|(x_0-x'_0)}\} \tag{6.7}$$

When $x_0 = x'_0$ this will vanish provided that $|k| = 1$ and $(-1)^N = \mp 1$. For the *commutator* to vanish, we therefore require N to be even, i.e. the field must have integral spin. This allows us to proceed to the next stage in field quantization, with the particles obeying Bose-Einstein statistics. There is, however, another possibility. If we extend the notion of quantization to permit Poisson brackets to become *anti*-commutators as well as commutators, then for odd N, and therefore half-integral spin, we can have particles with Fermi-Dirac statistics as well. This connection between spin and statistics, the *spin-statistics theorem*, which is known to obtain experimentally, is strong encouragement that relativistic quantum theory is on the right track.

For the massive case we can form the combination

$$\tilde{\Phi}_{ABC\cdots}(x) = \Phi_{ABC\cdots}(x) + k\partial_A^{A'}\partial_B^{B'}\partial_C^{C'}\cdots\Phi^{\dagger}_{A'B'C'\cdots}(x) \tag{6.8}$$

For some k. It then may be demonstrated that

$$[\tilde{\Phi}_{ABC\cdots}(x), \tilde{\Phi}^{\dagger}_{A'B'C'\cdots}(x')]_{\pm} = \pm\int d^4p\, \delta(p^2-m^2)e^{-ip\cdot(x-x')}p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)}\times$$
$$\left(\theta(p_0) \pm \theta(-p_0)(-\tfrac{1}{2}m^2)^N|k|^2\right)$$

$$= \int \frac{d^3\mathbf{p}}{2p_0}e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')}p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)}\left(e^{-ip_0(x_0-x'_0)} \pm e^{ip_0(x_0-x'_0)}(-\tfrac{1}{2}m^2)^N|k|^2\right)\Bigg|_{p_0=\sqrt{\mathbf{p}^2+m^2}} \tag{6.9}$$

When $x_0 = x'_0$ this will vanish provided that $|k| = (\sqrt{2}/m)^N$ and $(-1)^N = \mp 1$. The spin-statistics theorem connection therefore applies here as well.

We then find that

$$[\tilde{\Phi}_{ABC\cdots}(x), \tilde{\Phi}^{\dagger}_{A'B'C'\cdots}(x')]_{\pm} = \pm\int d^4p\, \delta(p^2-m^2)\epsilon(p_0)e^{-ip\cdot(x-x')}p_{A(A'}p_{BB'}p_{CC'}\cdots{}_{)} \tag{6.10}$$

and

$$[\tilde{\Phi}_{ABC\cdots}(x), \tilde{\Phi}^{KLM\cdots}(x')]_{\pm} = \pm k(-\tfrac{i}{2}m^2)^N\int d^4p\, \delta(p^2-m^2)\epsilon(p_0)e^{-ip\cdot(x-x')}\delta^{KLM\cdots}_{(ABC\cdots)} \tag{6.11}$$

where $\epsilon(x) = -1$ when $x < 0$, $\epsilon(0) = 0$ and $\epsilon(x) = 1$ when $x > 0$. We use the notation

$$\delta^{ABC\cdots}_{KLM\cdots} = \delta^A_K\delta^B_L\delta^C_M\cdots \tag{6.12}$$

Equation (6.10) applies to massless fields as well. For the massive case, note that

$$\tilde{\Phi}^{\dagger}_{A'B'C'\cdots}(x) = k^*\partial^A_{A'}\partial^B_{B'}\partial^C_{C'}\cdots\tilde{\Phi}_{ABC\cdots}(x) \tag{6.13}$$

If N is even we can form the $N/2$-index Lorentz tensor field

$$A_{abc\cdots}(x) = \partial^K_{A'}\partial^L_{B'}\partial^M_{C'}\cdots\tilde{\Phi}_{ABC\cdots KLM\cdots}(x) \tag{6.14}$$

This is symmetric and traceless, and the contraction with ∂^a will vanish. It can then be shown that if we choose $k = (\sqrt{2}/m)^N$ then condition (6.13) means that $A_{abc\cdots}(x)$ is also Hermitian. The commutator function of this field is then

$$[A^{abc\cdots}(x), A_{klm\cdots}(x')] = -(m/\sqrt{2})^N \int d^4p\ \delta(p^2 - m^2)\epsilon(p_0)e^{-ip.(x-x')}\ T^{abc\cdots}_{klm\cdots}(p) \tag{6.15}$$

where

$$T^{abc\cdots}_{klm\cdots}(p) = p^{N'}_K p^{O'}_L p^{P'}_M \cdots p^A_{D'} p^B_{E'} p^C_{F'} \cdots \delta^{A'B'C'\cdots D'E'F'\cdots}_{(K'L'M'\cdots N'O'P'\cdots)} \tag{6.16}$$

If N is odd, we may form a $\frac{1}{2}(N-1)$-index Lorentz tensor-spinor

$$A_{abc\cdots Q}(x) = \partial^K_{A'}\partial^L_{B'}\partial^M_{C'} \cdots \tilde{\Phi}_{ABC\cdots KLM\cdots Q}(x) \tag{6.17}$$

This is symmetric and traceless in any pair of Lorentz indices, and contractions with ∂_a will vanish. Choosing $k = -i(\sqrt{2}/m)^N$ we find that (6.13) leads to a constraint on A as follows:

$$A^{\dagger}_{abc\cdots Q'}(x) = \frac{\sqrt{2}}{m} i\partial^Q_{Q'} A_{abc\cdots Q}(x) \tag{6.18}$$

For the case of spin $\frac{1}{2}$, this will be recognised as the Dirac equation for a Majorana spinor. The anticommutator is then

$$\{A^{abc\cdots Q}(x), A_{klm\cdots S}(x')\} = -(m/\sqrt{2})^{N-2} \int d^4p\ \delta(p^2 - m^2)\epsilon(p_0)e^{-ip.(x-x')}\ T^{abc\cdots Q}_{klm\cdots S}(p) \tag{6.19}$$

where

$$T^{abc\cdots Q}_{klm\cdots S}(p) = p^{U'}_K p^{V'}_L p^{W'}_M \cdots p^{S'}_S\ p^A_{D'} p^B_{E'} p^C_{F'} \cdots p^Q_{Q'}\ \delta^{A'B'C'\cdots D'E'F'\cdots Q'}_{(K'L'M'\cdots U'V'W'\cdots S')} \tag{6.20}$$

7. Interactions

7.1 Quantization of classical electrodynamics; Haag's theorem

The fundamental mathematical structure of quantum mechanics (i.e. that the states of a physical system are an infinite-dimensional complex linear vector space with a sesquilinear inner product that carries a non-trivial, inner-product-preserving representation of the spacetime and other symmetry groups) was stated in section 2.

The remainder of what one needs to know in order to turn quantum mechanics into a calculational tool cannot be stated as elegantly. It is based on applying the formal "quantization" procedure to classical mechanics and electrodynamics. Since defining the microscopic behaviour of a system from its macroscopic behaviour cannot be expected to be reliable, we need not be too surprised when we find problems.

The first problem is that if, as would appear to be necessary, we use the four-vector potential A_a as dynamical variables for the electromagnetic field, we find that the conjugate momentum to A_0 vanishes identically, preventing us from forming Poisson brackets. We rescue this situation non-axiomatically by choosing a gauge, and then treating A_0 as a derived, rather than fundamental variable. Quantization can then proceed, leaving a mutual interaction between fermions via a static $1/r$ potential.

The Hamiltonian, which in the quantum world is the time displacement operator, then is the sum of the free Hamiltonians for fermions and photons, plus a Coulomb interaction between the fermions, and a three-point fermion-photon interaction. If one adds couplings arising from the half-integral spin of the fermions (which have no classical analogue) one then has a theory that, *inter alia* can (a) accurately generate energy levels of fermion bound states and (b) correctly accounts, at least to first order in perturbation theory, for the absorption and emission of radiation by free and bound fermions.

One does, however, run into the very serious problem that, to second order in perturbation theory, the fermion-photon interaction— the so-called fermion *self-energy*—is infinite.

Less serious, but still worrying, is the lack of explicit covariance. It seems wrong that even with fully relativistic equations of motion we can end up with a theory that does not look relativistic at all. In particular, the notion that the Hamiltonian can be written in the form

$$H = H_0 + V \tag{7.1}$$

where H_0 is the "free" Hamiltonian for the particles, and V is the "interaction", can be demonstrated to be incompatible with special relativity as follows.

Form a unitary operator $U(t)$ thus:

$$U(t) = e^{iH_0 t} e^{-iHt} \tag{7.2}$$

This can be used to transform an interacting particle at position $\mathbf{x}$ at time x_0 to a free one:

$$|x\rangle_{\text{free}} = U(n \cdot x)|x\rangle_{\text{int}} \tag{7.3}$$

where $x = (x_0, \mathbf{x})$ and $n_a = (1, 0, 0, 0)$. The free and interacting states are the same when $x_0 = 0$. Now, we are requiring that both free and interacting theories are covariant. Hence, applying a Lorentz transformation

$$U(0, \Lambda)|x\rangle_{\text{free}} = |\Lambda x\rangle_{\text{free}} \qquad \text{and} \qquad U(0, \Lambda)|x\rangle_{\text{int}} = |\Lambda x\rangle_{\text{int}} \tag{7.4}$$

Applying this Lorentz transformation to (7.3), rearranging, and replacing Λx with x, we then find

$$|x\rangle_{\text{free}} = U(0, \Lambda)U(\Lambda n \cdot x)U(0, \Lambda)^{\dagger}|x\rangle_{\text{int}} \tag{7.5}$$

This shows that for all x on the spacelike hyperplane $n'.x = 0$ where $n' = \Lambda n$, the free state is the same as the interacting state. Since Λ can be any Lorentz transformation, we conclude that for *any* spacelike x the free and interacting states are the same. Now, for a later time, where x is timelike and future-pointing, (7.5) shows that the unitary transformation that relates the free and interacting states here, although not necessarily the identity, must be the same as that connecting the states for points separated from here by a spacelike interval. Some of these points will also be separated from (0, 0, 0, 0) by a spacelike interval, for which the unitary transformation is known to be one. Hence the unitary transformation for *all* points is one. Thus $H_0 = H$ and the interaction V is trivial.

The notion that a relativistic field theory that is related to a free field theory by a unitary transformation $U(t)$ must itself be a free field theory, is known as *Haag's theorem.*

Haag's theorem is a valid result that can be demonstrated in a number of different ways. It is also an important result. The fact that it is almost completely ignored by writers of text books on quantum field theory is therefore a source of puzzlement to the author.

7.2 Local field equations

Local field equations in general are those where a disturbance cannot propagate at infinite speed. In a relativistic system, one further requires that the disturbance does not propagate faster than the speed of light. This definition assumes that one may easily disentangle cause and effect: something that is much harder to do in a quantum system than a classical one, so in practise locality is just taken to mean that the equations of motion are of the form

$$K(\partial)\Phi_i(x) = P(\partial, \Phi_j(x)) \tag{7.6}$$

where K is some finite-order differential kernel and P is some finite-order polynomial in the fields and spacetime derivatives, such that covariance is respected. One may see that the operator $\exp(a.\partial)$, which has the property $\exp(a.\partial)f(x) = f(x + a)$, is an example of one that both is non-local and contains derivatives up to infinite order. Whether derivatives up to infinite order *always* imply non-locality in the original sense is, however, not so clear.

To be continued ...