Beyond the Standard Model: Theoretical Explorations of Negative Mass Physics

Nathalie Debergh Jean-Pierre Petit

"The mathematician plays a game in which he himself invents the rules while the physicist plays a game in which the rules are provided by Nature, but as time goes on it becomes increasingly evident that the rules which the mathematician finds interesting, are the same as those which Nature has chosen."

P.A.M. Dirac

"The miracle of appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve."

E.P. Wigner

Contents

Introduction **1**

1 Group theory **4**
1.1 What is a group? . . . 4
1.1.1 Example 1 . . . 5
1.1.2 Example 2 . . . 5
1.1.3 Example 3 . . . 6
1.2 What is a Lie group? . . . 7
1.3 What is a Lie algebra? . . . 9
1.4 What is a group representation? . . . 10
1.5 The Lie group $SO(3)$ and the Lie algebra $so(3)$ . . . 12
1.6 What does that have to do with physics? . . . 19
1.6.1 Orbital angular momentum . . . 19
1.6.2 Intrinsic angular momentum or spin . . . 23
1.6.3 Tell me how you transform yourself, and I will tell you who you are. . . . 24
1.7 The Lie group $SO(1,3)$ and the Lie algebra $so(1,3)$ . . . 25
1.7.1 Lorentz boosts . . . 25
1.7.2 Lorentz algebra and its representations . . . 28
1.7.2.1 The representation $D(\frac{1}{2},0)$ or left spinor representation . . . 30
1.7.2.2 The representation $D(0,\frac{1}{2})$ or right spinor representation . . . 31
1.7.2.3 The representation $D(\frac{1}{2},0)\oplus D(0,\frac{1}{2})$ . . . 31
1.7.2.4 The representation $\mathrm{D}(\frac{1}{2},\frac{1}{2})$ or vector representation . . . 31
1.7.3 Connected components of the Lorentz group . . . 33
1.8 The Poincaré group and the Poincaré algebra . . . 35
1.8.1 The Poincaré algebra . . . 35
1.8.2 Casimir operators of the Poincaré algebra . . . 39
1.8.3 Wigner little groups . . . 42
1.8.3.1 Wigner little group for timelike 4-vectors . . . 46
1.8.3.2 Wigner little group for lightlike 4-vectors . . . 46
1.8.3.3 Wigner little group for spacelike 4-vectors . . . 48

1.9 The $su(3,C)$ algebra and the eightfold way . . . 51
1.9.1 The $su(3,C)$ algebra . . . 51
1.9.2 Root system . . . 64
1.9.3 Weights diagram . . . 66
1.9.4 Fundamental representations of $su(3,C)$ . . . 67
1.9.5 What does this have to do with physics? . . . 68
1.9.6 The eightfold way . . . 70
1.10 The complex extension of the Poincaré group . . . 74
1.10.1 The extended real Poincaré group . . . 75
1.10.2 The extended real Poincaré algebra . . . 76
1.10.3 The complex Poincaré algebra . . . 80
1.10.4 The Casimir operators . . . 89
1.10.4.1 The first Casimir operator . . . 89
1.10.4.2 Wigner's little groups . . . 89
1.10.4.3 The second Casimir operator . . . 94

2 The Dirac equation 97
2.1 The Klein-Gordon equation . . . 101
2.2 The Dirac equation . . . 103
2.2.1 The probabilistic interpretation associated with the Dirac equation . . . 107
2.2.2 The relativistic invariance of the Dirac equation . . . 108
2.2.3 How about discussing discrete symmetry operators? . . . 111
2.2.3.1 The parity operator P . . . 112
2.2.3.2 The time-reversal operator T . . . 113
2.2.3.3 The charge conjugation operator C . . . 114
2.2.4 Solutions to the Dirac equation . . . 116
2.2.5 So, should we accept these negative energies or not? . . . 118
2.2.6 The Feynman-Stueckelberg approach . . . 122
2.2.7 We should probably wrap things up, right? . . . 125
2.2.8 Chirality . . . 128

3 The geometric algebra point of view 132
3.1 The foundations . . . 132
3.2 Reflections and rotations . . . 138
3.3 The space algebra . . . 142
3.4 The 4D-space algebra . . . 145
3.5 The spacetime algebra . . . 149
3.6 Some tedious but necessary calculations . . . 153
3.6.1 The conjugation $\widehat{M}$ . . . 153
3.6.2 The conjugation $\widetilde{M}$ . . . 153
3.6.3 The conjugation $\overline{M}$ . . . 154
3.7 Relativistic invariance through the spacetime algebra . . . 155
3.8 The Dirac-Hestenes equation . . . 166
3.9 The real 2x2 Dirac equation . . . 168
3.10 Symmetries of the real 2x2 Dirac equation . . . 172

3.10.1 The even (proper) symmetries . . . 172
3.10.2 The odd (improper) symmetries . . . 175
3.10.3 The discrete symmetries P and T . . . 177
3.10.3.1 The parity operator . . . 178
3.10.3.2 The time-reversal operator . . . 180

4 Perspectives in quantum field theory 183
4.1 The harmonic oscillator . . . 183
4.2 The Klein-Gordon scalar field . . . 185
4.3 The Dirac field . . . 192
4.3.1 Properties of solutions to the Dirac equation . . . 192
4.3.2 The Dirac field . . . 198
4.3.3 Discrete symmetry operators . . . 203
4.3.3.1 The parity operator P . . . 203
4.3.3.2 The time-reversal operator . . . 206
4.3.3.3 The charge conjugation operator . . . 209
4.4 What if we took it a step further? . . . 210
4.4.1 Solutions to the Dirac equation: Would you like to know a little more? . . . 210
4.4.2 Different Dirac fields . . . 213
4.4.3 The Effect of Discrete Symmetry Operators on Solutions . 220
4.4.3.1 The unitary parity operator P_U . . . 220
4.4.3.2 The antiunitary time-reversal operator T_{AU} . . . 221
4.4.3.3 The antiunitary parity operator P_{AU} . . . 222
4.4.3.4 The unitary time-reversal operator T_U . . . 227

Conclusion 231

List of Figures

1.1 Action of the ladder operators of $so(3)$ . . . 17
1.2 Real spherical harmonics for $l = 0, 1, 2$ . . . 22
1.3 Spin up (counterclockwise rotation) and spin down (clockwise rotation) states for the electron . . . 23
1.4 Visualization of Einstein's relation in one spatial dimension . . . 43
1.5 Visualization of Einstein's relation for timelike and spacelike 4-vectors . . . 44
1.6 Action of the ladder operators of $so(2,1)$ . . . 49
1.7 Action of ladder operators of $su(3,C)$ . . . 54
1.8 Representation $D(1,0)$ of $su(3,C)$ . . . 56
1.9 Representation $D(0,1)$ of $su(3,C)$ . . . 58
1.10 Representation $D(1,1)$ of $su(3,C)$ . . . 60
1.11 Root system of $su(3,C)$ . . . 64
1.12 Weights diagram of the representation $D(1,0)$ of $su(3,C)$ . . . 67
1.13 Weights diagram for the representations $D(1,0)$ and $D(0,1)$ of $su(3,C)$ . . . 68
1.14 Mesons octet . . . 72
1.15 Baryons decuplet . . . 73

2.1 Dirac's sea . . . 119
2.2 Anderson's experiment . . . 121
2.3 Feynman diagram . . . 122
2.4 Time reversal in the Feynman diagram . . . 123
2.5 Example of a chiral molecule : CHBrClF or Bromochlorofluoromethane . . . 129

3.1 Two representations of the same bivector $a \wedge b$ . . . 133
3.2 Antisymmetry of the wedge product . . . 134
3.3 Visualization of the geometric product . . . 135
3.4 Contraction of parallel directions-expansion of perpendicular directions . . . 138
3.5 Reflection of a . . . 139
3.6 Timelike, spacelike and lightlike 1-vectors . . . 149

3.7 Actions of the rotors $R1 = \exp(e_1 e_2 \frac{\theta}{2})$ and $R2 = \exp(-e_1 e_2 \frac{\theta}{2})$ on e_2 156
3.8 Actions of the rotors $B1 = \exp(e_0 e_1 \frac{\varphi}{2})$ and $B2 = \exp(-e_0 e_1 \frac{\varphi}{2})$ on e_0 159

4.1 Harmonic oscillator energies 184

Introduction

It is the second half of the 19th century, 1876 to be precise. At that time, the Norwegian mathematician Sophus Lie[1] undertook an in-depth study of infinitesimal transformations, paving the way for new algebraic structures that would later bear his name. These mathematical objects, which Hermann Weyl would name "Lie algebras" nearly sixty years later, appeared at the time as abstract constructs, disconnected from any tangible reality. At first glance, nothing suggested that they would play a fundamental role in our understanding of the physical world.

And yet, the history of science is full of examples where purely theoretical ideas, developed without any immediate practical application in mind, ultimately find profound resonance in the description of nature. Lie's work is a striking illustration of this. For several decades, it remained confined to the realm of mathematics, until certain visionary physicists sensed its potential. As early as 1927, Eugene Wigner[2] emphasized the importance of symmetries in physics and highlighted the role that Lie groups and algebras could play in structuring fundamental laws.

However, it was not until the early 1960s that this intuition truly took shape. At that time, particle physics was in full swing: many particles were being discovered, but their organization remained unclear. It was then that Murray Gell-Mann[3], in the United States, and Yuval Ne'eman[4], in Israel, independently proposed an elegant classification based on internal symmetries. By representing particles using two quantum numbers—isospin and hypercharge—they revealed remarkable geometric structures: hexagons, regular triangles, and ordered configurations that owe nothing to chance. These patterns correspond precisely to the weight diagrams derived from Lie algebra theory.

This correspondence between abstract mathematics and physical reality is striking. Better yet, it does not merely describe what is known; it allows us to

[1] Lie, S. (1876). Theorie der Transformationsgruppen. Abahandlung II. *Archiv for Mathematik og Naturvidenskab,* 1, 152-193

[2] Wigner, E. (1927). Uber nicht kombinieren de Terme in der neueren Quantentheorie. II. Teil. *Zeitschrift für Physik*, 40, 883-892. https://doi.org/10.1007/BF01390906

[3] Gell-Mann, M. (March 15,1961). *The Eightfold Way: A theory of strong interaction symmetry.* Office of Scientific and Technical Information (OSTI). https://doi.org/10.2172/4008239

[4] Ne'eman, Y. (August 1961). Derivation of strong interactions from a gauge invariance. *Nuclear Physics,* 26, 222-229. https://doi.org/10.1016/0029-5582(61)90134-1

predict what is not yet known. By identifying "empty boxes" in these diagrams, Gell-Mann anticipated the existence of new particles. One of them, the Ω^- baryon, was indeed discovered in 1964, spectacularly confirming the predictive power of this approach. A theory developed nearly a century earlier had thus opened a new path for experimental physics.

This success led to a profound rethinking of the very concept of the elementary particle. From now on, a particle is no longer characterized solely by its measurable properties, but also by the way it transforms under the action of a symmetry group. More precisely, it is associated with an irreducible unitary representation of that group. This idea, now central to theoretical physics, places symmetries at the heart of the description of the microscopic world.

Among these symmetry groups, the Poincaré group holds a special place. It describes the fundamental symmetries of spacetime: translations in space and time, rotations, and Lorentz transformations (or "boosts"). The study of its representations reveals essential properties of particles, notably their mass and spin. Remarkably, the only constraint imposed by this structure is that the square of the mass must be real.

This condition, which seems trivial at first glance, actually opens the door to several possibilities. If the square of the mass is positive, the mass itself can be positive or negative. If the square is negative, the mass becomes imaginary. All these situations are mathematically consistent and arise naturally from the analysis of the representations of the Poincaré group. However, within the framework of the Standard Model of particle physics, only the possibility of positive masses is generally accepted, while the others are ruled out for reasons of interpretation or stability.

Although this choice has proven effective in experiments, it raises profound questions. Why should certain mathematical solutions—even though they are permitted by fundamental symmetries—be ruled out a priori? Are there theoretical frameworks in which negative or imaginary masses could have physical meaning? These questions have led some researchers to explore more daring models, in which these concepts have a legitimate place.

This work is part of this broader effort. It aims to examine the interactions between quantum mechanics and particles with negative mass. Such an approach requires a reexamination of the very foundations of theoretical physics, drawing on both the algebraic structures inherited from Lie's work and the fundamental equations of relativistic quantum mechanics.

To carry out this study, the document is organized into four parts.

The first part provides a review of the essential concepts related to Lie groups and Lie algebras, with an emphasis on the symmetries relevant to physics: the rotation group, the Lorentz group, the Poincaré group, as well as certain extensions to more general spacetimes, particularly complex ones.

The second part is devoted to the Dirac equation, which provides the natural framework for describing relativistic fermions. We will see that this equation, from its initial formulation, contains rich structures that allow for various interpretations, particularly regarding solutions with negative mass.

The third part opens the door to a field that is still frequently overlooked

in the scientific community: geometric algebra. We will see that this concept generalizes the traditional approach and leads, in a completely natural way, to the emergence of complex rotations for both position and mass.

Finally, a fourth part opens up new perspectives by exploring current and future research avenues, particularly in the context of quantum field theory, where the concepts of negative masses could find a coherent framework for expression.

Thus, from the abstract rigor of 19th-century mathematics to the most contemporary questions in theoretical physics, this work offers a journey through symmetries, representations, and their implications. The aim is to examine how far these structures can take us, and whether they allow us, beyond established frameworks, to expand our understanding of particles and their fundamental properties.

Chapter 1

Group theory

The philosophy of this book is to present concepts as simply as possible. And we will not deviate from this approach when it comes to groups.

The reason for this is that Lie groups are ubiquitous in quantum physics. Better still, they were what made it possible, for the first time, to predict the existence of an elementary particle. This existence was confirmed experimentally shortly thereafter.

But before we return to this topic, let's go back to the fundamental concepts.

1.1 What is a group?

Group theory was initiated in the early 1830s by French mathematician Evariste Galois, with the aim of establishing a method for solving algebraic equations.

A group $(G, \circ)$ can be seen as a set G of mathematical objects on which you impose a law denoted by $\circ$ that has the following four properties:

- the law is internal: $\forall\, g_j,\, g_k \in G,\; g_j \circ g_k \in G$
- there is a neutral element: $\exists\, e \in G : \forall g_j \in G,\, g_j \circ e = e \circ g_j = g_j$
- there is a symmetric element: $\forall\, g_j \in G, \exists\, g_j^{-1} \in G : g_j \circ g_j^{-1} = g_j^{-1} \circ g_j = e$
- the law is associative: $\forall\, g_j,\, g_k,\, g_l \in G,\; (g_j \circ g_k) \circ g_l = g_j \circ (g_k \circ g_l)$

Moreover, if the law is such that

- $\forall\, g_j,\, g_k \in G,\; g_j \circ g_k = g_k \circ g_j$

then the group is said to be abelian or commutative.

Let's illustrate this definition with a few basic examples.

1.1.1 Example 1

The set of integers equipped with the addition operation is a group named $(\mathbb{Z}, +)$. The neutral element is 0 and the symmetric element is given by the opposite: if $g_j = j$, then $g_j^{-1} = -j$, $(j = 0, \pm 1, \pm 2, \ldots)$.

Due to the commutativity of the addition of integers, this group is abelian.

Note that restricting oneself to positive integers while still choosing addition as the law causes one to lose the notion of a group. Indeed, the existence of the symmetric element is, in this case, ineffective.

1.1.2 Example 2

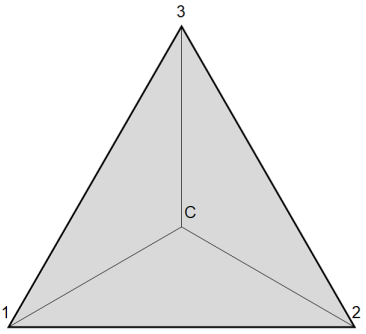
Let's complicate things a little with another example. This time, the mathematical objects are rotations of an equilateral triangle around the triangle's center of gravity C. Consider three rotations:

- g_1=rotation with center C, trigonometric direction, and amplitude 0°
- g_2 = rotation with center C, trigonometric direction, and amplitude 120°
- g_3 = rotation with center C, trigonometric direction, and amplitude 240°

The law $\circ$ will then be the composition of these rotations, i.e., "I perform one first rotation and then a second one".

What is the effect of each of these rotations?

It is easier to answer this question if we draw one of these triangles.



We can see that the triangle as a geometric object remains unchanged under each of these rotations. Only the names of the vertices may change.

Under g_1, the triangle undergoes no transformation; therefore, we have $1 \to 1$, $2 \to 2$, $3 \to 3$ after rotation.

Under g_2, vertices have rotated 120° and are now such that $1 \to 2$, $2 \to 3$, $3 \to 1$.

Under g_3, vertices have rotated 240° and we observe that $1 \to 3$, $2 \to 1$, $3 \to 2$.

What if we combine two rotations? For example, we perform g_2 and then g_3, which we will write as $g_3 \circ g_2$. The transformation g_2 brings about the modifications already mentioned, and if we apply g_3 afterwards, we obtain

$$1 \to 2 \to 1\,;\; 2 \to 3 \to 2\,;\; 3 \to 1 \to 3$$

It's the status quo, marked by rotation g_1. We will write

$$g_3 \circ g_2 = g_1$$

This result clearly shows that a 120° rotation followed by a 240° rotation results in a 360° rotation and therefore 0°.

We have three rotations and we need to consider all possible sequences, i.e., repeat the above reasoning to arrive at the following summary table

$\circ$	g_1	g_2	g_3
g_1	g_1	g_2	g_3
g_2	g_2	g_3	g_1
g_3	g_3	g_1	g_2

This table allows us to see at a glance that we are indeed dealing with a (commutative) group. No cell is empty, and each contains one of the elements g_1, g_2, or g_3. The first property is therefore satisfied. The neutral element is clearly g_1. A look at the table shows us that g_1, g_3, and g_2 are the symmetric elements of g_1, g_2, and g_3, respectively. Associativity takes a little longer to prove, but poses no difficulty.

All three rotations, equipped with the composition law, therefore form a group.

Should this come as a surprise?

You will have noticed that composing two rotations is equivalent to adding their angle amplitudes. Thus $g_2 \circ g_2 = g_3$ because 120°+120°=240°.

It is precisely this very simple property of rotations that leads to a group structure when they are composed.

1.1.3 Example 3

To the three rotations previously considered, we now add three axial symmetries according to the medians of the equilateral triangle.

We therefore have a total of six transformations characterized by

$$\begin{cases} g_1 : 1 \to 1,\ 2 \to 2,\ 3 \to 3 \\ g_2 : 1 \to 2,\ 2 \to 3,\ 3 \to 1 \\ g_3 : 1 \to 3,\ 2 \to 1,\ 3 \to 2 \\ g_4 : 1 \to 2,\ 2 \to 1,\ 3 \to 3 \\ g_5 : 1 \to 1,\ 2 \to 3,\ 3 \to 2 \\ g_6 : 1 \to 3,\ 2 \to 2,\ 3 \to 1 \end{cases}$$

The table now becomes

$\circ$	g_1	g_2	g_3	g_4	g_5	g_6
g_1	g_1	g_2	g_3	g_4	g_5	g_6
g_2	g_2	g_3	g_1	g_5	g_6	g_4
g_3	g_3	g_1	g_2	g_6	g_4	g_5
g_4	g_4	g_6	g_5	g_1	g_3	g_2
g_5	g_5	g_4	g_6	g_2	g_1	g_3
g_6	g_6	g_5	g_4	g_3	g_2	g_1

This group is known as the dihedral group D6. It is a non-abelian group.

A remarkable fact is that it contains a subset that has the structure of a group. We are, of course, referring to the previously mentioned group of rotations. We will therefore say that this set is a *subgroup* of D6.

1.2 What is a Lie group?

We have just seen that an equilateral triangle is invariant under the action of three rotations, of 0°, 120°, and 240° respectively, forming a group.

Do you need convincing that a square will remain invariant when subjected to four rotations of 0°, 90°, 180°, and 270°, respectively?

You understand what comes next: a regular pentagon will remain unchanged under the action of five rotations of 0°, 72°, 144°, 216°, and 288°, also forming a group.

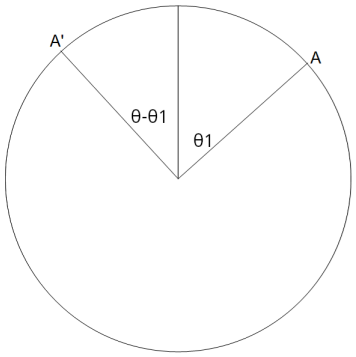
How far can we go?

Up to the ultimate regular flat shape: the disc. This will remain unchanged regardless of the rotation you apply to it. In other words, the angle amplitudes no longer form a discrete set but a continuous one.

This is precisely the principle behind Lie groups.

Let's illustrate what we have just started discussing using a simple disk. The equivalents of the vertices are now the points on the circle. If we rotate them around the center O of the disk, in the trigonometric direction and by an angle θ, they will take the place of other points on the circle.

For example, point A will become $A\prime$:



But what is the relation between the coordinates (x, y) of A and the coordinates $(x\prime, y\prime)$ of $A\prime$?

A little trigonometry is needed.

The angle θ consists of two angles: $\theta 1$, which is the part of the angle included in the first quadrant, and $(\theta - \theta 1)$, which is the part of the angle included in the second quadrant.

We then have

$$x = r\sin\theta 1\,;\; y = r\cos\theta 1$$

but also

$$x\prime = -r\sin(\theta - \theta 1); y\prime = r\cos(\theta - \theta 1)$$

if r is the radius of the disc.

It is easy to deduce that

$$\begin{cases} x\prime = x\cos\theta - y\sin\theta \\ y\prime = x\sin\theta + y\cos\theta \end{cases} \tag{1.1}$$

A very simple way to check that we have not made a mistake is to calculate

$$x\prime^2 + y\prime^2$$

and see that it is equal to

$$x^2 + y^2$$

This clearly ensures the invariance of the circle. In other words, we can say that the norm in the Euclidean plane is preserved.

Eqs. (1.1) can also be written as

$$\begin{pmatrix} x\prime \\ y\prime \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv R(\theta) \begin{pmatrix} x \\ y \end{pmatrix} \tag{1.2}$$

Here we have our first Lie group! The matrix $R(\theta)$ characterizes a Lie group called $SO(2)$.

Why this name?

The reason for this is simple, and we invite you to remember these specificities as they will come up frequently.

- The letter S means that the matrix has a determinant equal to 1.
- The letter O indicates the orthogonality of the matrix, in the sense that the product of this matrix and its transpose gives the identity matrix.
- The number 2 appears in connection with the dimension of the matrix.

However, it is still necessary to verify that the properties of a group are indeed present!

To do this, we can easily calculate that

$$R(\theta)R(\theta\prime) = R(\theta + \theta\prime) \tag{1.3}$$

The internal law is therefore ensured, as is the existence of the neutral element (which corresponds to the zero angle) and the symmetric element ($\theta\prime = -\theta$). Associativity requires more elbow grease but presents no difficulty.

1.3 What is a Lie algebra?

Let's reconsider the disk and, rather than taking into account rotations of any amplitude, let's focus on rotations of very small amplitude... what are known as infinitesimal rotations. After all, all the information is already there: we know the center of rotation, its direction... no need to perform a gigantic rotation for that!

But if we recall the series expansions of cosine and sine:

$$\cos\theta = 1 - \frac{\theta^2}{2} + \frac{\theta^4}{24} - \cdots \tag{1.4}$$

$$\sin\theta = \theta - \frac{\theta^3}{6} + \frac{\theta^5}{120} - \cdots \tag{1.5}$$

we observe that a drastically reduced angle will limit these developments to their initial stages only.

The rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{1.6}$$

can then be written as

$$R(\theta) \simeq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \theta \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \equiv I + \theta J \tag{1.7}$$

The J matrix is precisely the one defined as being associated with the Lie algebra of $SO(2)$. Mathematicians didn't rack their brains too much and called this Lie algebra $so(2)$ (in lowercase letters).

This time

- The letter s means that the matrix has zero trace.
- The letter o indicates the antisymmetry of the matrix.
- The number 2 is always related to the dimension of the matrix.

Lie algebra therefore appears as the infinitesimal version of its group. The two correspond to each other through exponentiation:

$$R(\theta) = \exp(\theta J) \tag{1.8}$$

In summary, the nature of the transformations involves considering one or another type of mathematical structure:

- The transformation is discrete $\rightarrow$ we are dealing with a group.
- The transformation is continuous $\rightarrow$ we are dealing with a Lie group.
- The transformation is infinitesimal $\rightarrow$ we are dealing with a Lie algebra.

Quantum mechanics is not particularly concerned with discrete transformations, except in very specific cases: Lie structures therefore play a key role in this theory.

We will now examine some groups and Lie algebras that are of primary importance in quantum physics. But before that, we must discuss the representations of groups and algebras.

1.4 What is a group representation?

To make it easier to handle a group, we can give it a representation.

A representation T is a mapping of a group G onto a group $G\prime$. Each represented element will therefore be written as $T(g_1), T(g_2), T(g_3)$, ..., in terms of elements of $G\prime$. These can be numbers, but also matrices as well as differential operators... we will come back to this later.

What applies to the elements of group G must apply equally to those of group $G\prime$.

Thus, if $g_2 \circ g_1 = g_3$ in G, we must ensure that $T(g_2) \bullet T(g_1) = T(g_3)$ in $G\prime$. In other words, a representation T is characterized by

$$T(g_2) \bullet T(g_1) = T(g_2 \circ g_1) \tag{1.9}$$

This last relationship translates into saying that T is a group homomorphism. Homomorphism and not isomorphism because there may be a difference in the number of elements.

As an illustration, let us consider two representations of the group generated by g_1, g_2, g_3, the rotations considered previously and leaving the equilateral triangle unchanged.

They must satisfy

$\bullet$	$T(g_1)$	$T(g_2)$	$T(g_3)$
$T(g_1)$	$T(g_1)$	$T(g_2)$	$T(g_3)$
$T(g_2)$	$T(g_2)$	$T(g_3)$	$T(g_1)$
$T(g_3)$	$T(g_3)$	$T(g_1)$	$T(g_2)$

The first representation is trivial:

$$T(g_1) = T(g_2) = T(g_3) = 1$$

The group law $\bullet$ is then simply the multiplication of integers. It is easy to verify that this representation does indeed give the same results as those shown in the table above.

It is immediately clear why we talk about homomorphism: the three elements g_1, g_2, and g_3 are reduced to a single number for this representation.

The second representation is less obvious. It uses 2-dimensional matrices:

$$T(g_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, T(g_2) = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, T(g_3) = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix} \tag{1.10}$$

The group law $\bullet$ is then matrix multiplication.

Group representations can be summarized as follows: the elements are not the same as those in the group, nor are they composed in the same way, but they generate the same table... and they are easier to manipulate because they are concrete!

Not all representations are equally important. Those that are unitary and irreducible are favored in quantum physics. They are also referred to in the literature by the acronym "IUR".

Unitarity is a sufficient condition for ensuring the preservation of the probabilistic interpretation of quantum mechanics, as demonstrated by Wigner[1].

Irreducibility is related to this. A representation T is irreducible if it is not possible to find a unitary matrix U such that

$$U\,T(g)\,U^{-1} = \begin{pmatrix} T1(g) & 0 \\ 0 & T2(g) \end{pmatrix} \tag{1.11}$$

where $T1$ and $T2$ are themselves two representations of the group.

Our second representation (1.10) is not irreducible since the matrix

$$U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

[1] Wigner, E. (1959). *Group Theory and its Applications to the Quantum Mechanics of Atomic Spectra*, Academic Press Inc. (New York)

leaves $T(g_1)$ unchanged but is such that

$$U\,T(g_2)\,U^{-1} = \frac{1}{2}\begin{pmatrix} -1+i\sqrt{3} & 0 \\ 0 & -1-i\sqrt{3} \end{pmatrix}$$

$$U\,T(g_3)\,U^{-1} = \frac{1}{2}\begin{pmatrix} -1-i\sqrt{3} & 0 \\ 0 & -1+i\sqrt{3} \end{pmatrix}$$

So, in this case, we have

$$T1(g_1) = 1,\, T1(g_2) = \frac{1}{2}(-1+i\sqrt{3}),\, T1(g_3) = \frac{1}{2}(-1-i\sqrt{3})$$

$$T2(g_1) = 1,\, T2(g_2) = \frac{1}{2}(-1-i\sqrt{3}),\, T2(g_3) = \frac{1}{2}(-1+i\sqrt{3})$$

It is not difficult to convince oneself that these three complex numbers are also a representation of the group of three rotations related to the equilateral triangle, since they are the cubic roots of unity.

1.5 The Lie group $SO(3)$ and the Lie algebra $so(3)$

Rotations in a plane are all well and good, but we live in a three-dimensional space!

So we need to take the discussion to the next level: the goal is to consider any point on a sphere and rotate it in such a way that its new location is still on that sphere.

In a three-dimensional space, this can be done via three rotations in the respective planes (x, y), (x, z), and (y, z). These three rotations are represented by the matrices[2]

$$R_1(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, R_2(\theta) = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{1.12}$$

$$R_3(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix} \tag{1.13}$$

These three matrices have determinant equal to 1, are orthogonal, and have dimension 3. They generate the Lie group $SO(3)$.

The corresponding algebra is obtained by considering extremely small angles of rotation:

$$R_1(\theta) = \exp(\theta J_1) \simeq I + \theta J_1,\, R_2(\theta) = \exp(\theta J_2) \simeq I + \theta J_2$$

[2]We have used the same notation θ for each of the three rotations for convenience, but it is clear that the three rotations do not require the same amplitude to be set.

$$R_3(\theta) = \exp(\theta J_3) \simeq I + \theta J_3$$

Here, the matrices J_1, J_2, J_3 are explicitly given by

$$J_1 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad (1.14)$$

And that's where the magic of Lie algebras comes in.

It wasn't obvious with $so(2)$ since we only had one matrix, but here we have three. A Lie algebra is, in fact, a vector space equipped with a law called the "Lie bracket" which takes the form of a commutator:

$$[A, B] \equiv AB - BA \quad (1.15)$$

We must therefore apply this law to J_1, J_2, and J_3. Let's do it! We have

$$[J_1, J_2] = J_3,\ [J_3, J_1] = J_2,\ [J_2, J_3] = J_1 \quad (1.16)$$

What you see here is the DNA of Lie algebras: their commutation relations are unique and allow them to be identified at a glance.

First observation: when you calculate the commutator of two elements of a supposed Lie algebra, it must return a linear combination of the elements of that algebra and no new elements. We then say that the algebra is closed.

Second observation: the coefficients of this linear combination perfectly define the algebra. They are called structure constants.

In the case of the $so(3)$ algebra, a remarkable symmetry, a cyclicity, appears. The algebra can then be summarized by

$$[J_j, J_k] = \varepsilon_{jkl}\, J_l \quad (1.17)$$

by considering the antisymmetric Levi-Civita tensor ε:

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1\,;\, \varepsilon_{132} = \varepsilon_{213} = \varepsilon_{321} = -1 \quad (1.18)$$

So far, there's nothing really extraordinary about this... three 3-dimensional matrices... big deal! Yes, but this is where representations come into play. Because we can represent the three generators J_1, J_2, J_3 with matrices of any dimension or even differential operators!

How can these representations be highlighted?
Like this.

First, we redefine the generators of the algebra by complexifying it:

$$J_z \equiv iJ_3; J_\pm \equiv iJ_1 \mp J_2 \quad (1.19)$$

The Lie algebra is then characterized by the following relations:

$$[J_z, J_\pm] = \pm J_\pm, [J_+, J_-] = 2J_z \tag{1.20}$$

The advantage of doing this is that J_z can be seen as a diagonal Hermitian matrix and $J_\pm$ as ladder operators, which are Hermitian conjugates of each other. Patience...it will become clear in a few minutes.

First, the fact that J_z must be diagonal implies

$$J_z \left|m\right> = m \left|m\right> \tag{1.21}$$

whatever $|m>$ may be at this stage. We only know that m is real thanks to the fact that J_z is Hermitian.

Using Eqs. (1.20) and (1.21), we have

$$J_z J_+ \left|m\right> = \left(J_+ J_z + J_+\right) \left|m\right> = \left(m+1\right) J_+ \left|m\right>$$

$$J_z J_- \left|m\right> = \left(J_- J_z - J_-\right) \left|m\right> = \left(m-1\right) J_- \left|m\right>$$

The conclusion is clear: J_+ increases the value of m by one unit, while J_- decreases it by one unit. This is why they are called ladder operators.

We therefore ask for

$$J_+ \left|m\right> = c_+(m) \left|m+1\right>, \; J_- \left|m\right> = c_-(m) \left|m-1\right> \tag{1.22}$$

Now remember that

$$(J_+)^\dagger = J_- \tag{1.23}$$

where the notation $\dagger$ stands for the Hermitian conjugate.

It should also recalled that while ket $|m>$ can be represented by a column matrix, bra $<m\prime|$ is associated with a row matrix such that

$$<m\prime|m> = \delta_{m\prime m} \tag{1.24}$$

This leads to

$$<m+1|J_+|m> = c_+(m) <m+1|m+1> = c_+(m)$$

and

$$(c_+(m))^* = (<m+1|J_+|m>)^* = <m|J_-|m+1> = c_-(m+1) <m|m> = c_-(m+1)$$

Here, notation * refers to complex conjugation.

The equations (1.22) can then be rewritten as

$$J_+ \left|m\right> = c_+(m) \left|m+1\right>, \; J_- \left|m\right> = \left(c_+(m-1)\right)^* \left|m-1\right> \tag{1.25}$$

From these relations, we deduce that

$$J_+J_-|m>=|c_+(m-1)|^2\,|m> \tag{1.26}$$

$$J_-J_+|m>=|c_+(m)|^2\,|m> \tag{1.27}$$

Let us now turn to a crucial point: representations of a Lie algebra associated with a compact Lie group are finite-dimensional. Now, $SO(3)$ is a compact group. Therefore, we must ask that m takes a maximum value as well as a minimum value. We note j as its maximum value. This necessarily implies that

$$J_+|j>=c_+(j)\,|j+1>=0 \Longrightarrow c_+(j)=0 \tag{1.28}$$

Furthermore, it is easy to convince oneself that

$$<m|[J_+,J_-]|m>=<m|J_+J_--J_-J_+|m>=|c_+(m-1)|^2-|c_+(m)|^2 \tag{1.29}$$

We also have

$$<m|[J_+,J_-]|m>=<m|2J_z|m>=2m \tag{1.30}$$

From the last two equations, we deduce the recurrence relation

$$|c_+(m-1)|^2=|c_+(m)|^2+2m \tag{1.31}$$

This is how, by using Eq.(1.28), we have successively

$$|c_+(j-1)|^2=2j$$

$$|c_+(j-2)|^2=2(2j-1)$$

$$|c_+(j-3)|^2=3(2j-2)$$

and, in general

$$|c_+(j-k)|^2=k(2j-k+1) \tag{1.32}$$

Clearly, this result leads to

$$|c_+(m)|^2=(j-m)(j+m+1) \tag{1.33}$$

Obviously, we have $c_+(j)=0$, as expected, but also $c_+(-j-1)=0$. This means that m can only take $(2j+1)$ values, ranging from $-j$ to $+j$. Another implication is that j can only take (positive) integer or (positive) half-integer values.

Given the importance of this number j as developments progress, we will henceforth label the states by $|j,m>$ instead of $|m>$ and summarize the results obtained through the following equations

$$J_z\,|j,m>=m\,|j,m> \tag{1.34}$$

$$J_+\,|j,m>=\sqrt{(j-m)(j+m+1)}\,|j,m+1> \tag{1.35}$$

$$J_-\,|j,m>=\sqrt{(j+m)(j-m+1)}\,|j,m-1> \tag{1.36}$$

$$m=-j,-j+1,\ldots,j-1,j\,;\,j=0,\,\frac{1}{2},\,1,\,\frac{3}{2},\ldots \tag{1.37}$$

We must also mention an extremely important operator in Lie algebras: the Casimir operator.

This operator, a polynomial of the elements of the algebra, is defined by the fact that it commutes with all the generators of the algebra. In the case of $so(3)$, it is given by

$$C=\frac{1}{2}J_+J_-+\frac{1}{2}J_-J_++J_z^2=J_+J_--J_z+J_z^2=J_-J_++J_z+J_z^2 \tag{1.38}$$

Indeed, we have

$$[C,J_z]=[J_+J_-,J_z]=J_+[J_-,J_z]+[J_+,J_z]J_-=J_+J_--J_+J_-=0$$

$$[C,J_+]=J_+[J_-,J_+]-[J_z,J_+]+J_z[J_z,J_+]+[J_z,J_+]J_z=-J_++[J_z,J_+]=0$$

$$[C,J_-]=[J_+,J_-]J_--[J_z,J_-]+J_z[J_z,J_-]+[J_z,J_-]J_z=J_-+[J_z,J_-]=0$$

Using Eqs. (1.34)-(1.36), we can see that

$$C\,|j,m>=j(j+1)\,|j,m> \tag{1.39}$$

Eqs. (1.34)-(1.37) with (1.39) give us what is known as the $D(j)$ representation of $so(3)$.

It can be useful to visualize the actions of the various operators.

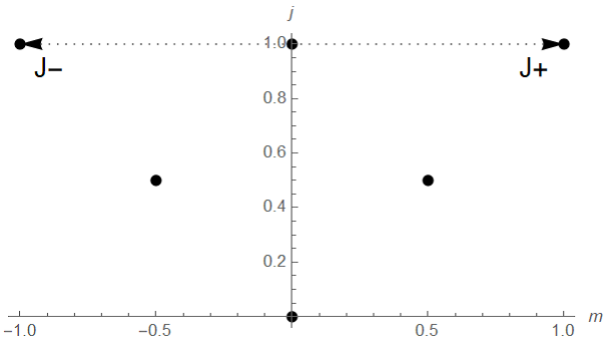


Figure 1.1: Action of the ladder operators of $so(3)$

The number j actually covers what are called the multiplets of $so(3)$, and the number m specifies the states of these multiplets. Thus

- for $j = 0$, there is a single state, or singlet, which corresponds to $(j = 0, m = 0)$;
- for $j = \frac{1}{2}$, there are two states, or doublet, which correspond to $(j = \frac{1}{2}, m = \frac{1}{2})$ and $(j = \frac{1}{2}, m = -\frac{1}{2})$;
- for $j = 1$, there are three states, or triplet, which correspond to $(j = 1, m = 1)$, $(j = 1, m = 0)$, and $(j = 1, m = -1)$;

It is of course understood that, for a fixed value of j, there will be a set of $(2j + 1)$ states (for $m = j, j - 1, j - 2, \ldots, -j + 2, -j + 1, -j$).

These states are represented by dots on the diagram.

The role of ladder operators is to connect the different states within the same multiplet. Thus, J_+ will increase the value of m by one unit, while J_- will connect a state characterized by (j, m) to a state labeled by $(j, m - 1)$.

All this may seem rather indigestible, but really, once you have the results, it becomes child's play!

Want a representation with 2-dimensional matrices? Just take

$$j = \frac{1}{2}; m = -\frac{1}{2}, \frac{1}{2}$$

Both states of the doublet can be achieved via

$$|\frac{1}{2}, \frac{1}{2}>= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |\frac{1}{2}, -\frac{1}{2}>= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{1.40}$$

The diagonal matrix J_z can then only be

$$J_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.41}$$

while the ladder matrices correspond to

$$J_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; J_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{1.42}$$

The Casimir operator is, of course, a multiple of the identity operator:

$$C = \frac{3}{4} I \tag{1.43}$$

Do you prefer a representation with 3-dimensional matrices? In that case, you need to consider

$$j = 1 \, ; \, m = -1, \, 0, \, 1$$

and the states

$$|1,1> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; |1,0> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \; |1,-1> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.44}$$

You then have

$$J_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, J_+ = \sqrt{2} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, J_- = \sqrt{2} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{1.45}$$

and

$$C = 2 \, I \tag{1.46}$$

The perceptive reader will have noticed, and perhaps been concerned about, the fact that these three matrices are not the same as those obtained from Eq. (1.14):

$$J_z = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, J_+ = \begin{pmatrix} 0 & -i & 1 \\ i & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J_- = \begin{pmatrix} 0 & -i & -1 \\ i & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{1.47}$$

Don't panic! If we calculate the Casimir operator in this case, we confirm that $C = 2\,I$, meaning that Eqs. (1.45) and (1.47) are two "different" (actually, unitarily equivalent) realizations of the same representation $D(1)$ of $so(3)$!

We can also see this directly by considering the unitary matrix

$$U = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & -i\sqrt{2} & 0 \\ 1 & 0 & 1 \\ i & 0 & -i \end{pmatrix}$$

which allows to switch between them.

So why choose one over the other? To facilitate the link with physics. In the $D(1)$ representation adopted in physics literature, the matrix J_z is diagonal and the matrices $J_\pm$ are Hermitian conjugates of each other, which is much easier to manipulate in quantum theory.

It should also be noted that this constraint is only feasible because the original $SO(3)$ group is compact. In other words, its parameters θ lie within a finite interval (between 0 and 2π, in this case). We will see that this constraint of having Hermitian conjugate operators will no longer apply once the group is no longer compact.

1.6 What does that have to do with physics?

The $SO(3)$ group and the $so(3)$ algebra are mathematical concepts related to a geometric transformation (rotation) and a physical quantity: angular momentum.

More specifically, some representations are associated with orbital angular momentum, while other representations correspond to what is known as intrinsic angular momentum or spin.

1.6.1 Orbital angular momentum

In classical physics, rotation refers to kinetic momentum or orbital angular momentum. This is defined by

$$\overrightarrow{L} = \overrightarrow{r} \wedge \overrightarrow{p} \tag{1.48}$$

which is the vector product of the position vector and the momentum vector, itself the product of mass and velocity.

Quantum theory retains this notion but adapts it: from now on, it will be an operator acting on a wavefunction. Let us recall that

$$\overrightarrow{r}_{op}\,\psi(t,\overrightarrow{r}) = \overrightarrow{r}\,\psi(t,\overrightarrow{r})\;;\;\overrightarrow{p}_{op}\,\psi(t,\overrightarrow{r}) = -i\,\hbar\,\overrightarrow{\nabla}\,\psi(t,\overrightarrow{r}) \tag{1.49}$$

As a consequence, we get

$$L_1\,\psi(t,\overrightarrow{r}) = -i\,\hbar\,(y\frac{\partial\psi}{\partial z} - z\frac{\partial\psi}{\partial y}),\; L_2\,\psi(t,\overrightarrow{r}) = -i\,\hbar\,(z\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial z}) \tag{1.50}$$

$$L_3\,\psi(t,\overrightarrow{r}) = -i\,\hbar\,(x\frac{\partial\psi}{\partial y} - y\frac{\partial\psi}{\partial x}) \tag{1.51}$$

It is easy to convince oneself that

$$[L_1,L_2] = i\,\hbar\,L_3,\;[L_3,L_1] = i\,\hbar\,L_2,\;[L_2,L_3] = i\,\hbar\,L_1 \tag{1.52}$$

when acting on $\psi(t,\overrightarrow{r})$.

Cyclicity requires that this be the algebra $so(3)$. And, indeed, by defining,

$$J_1 = \frac{1}{i\hbar}L_1,\ J_2 = \frac{1}{i\hbar}L_2,\ J_3 = \frac{1}{i\hbar}L_3 \tag{1.53}$$

we recover relations (1.16).

Since orbital angular momentum involves position and momentum, whose values are not limited, it is advisable to use differential operators and functions instead of matrices and column matrices to symbolize operators and states.

What are these functions?

To highlight them, we must leave the cartesian coordinates and consider the spherical ones:

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases} \tag{1.54}$$

$$r = \sqrt{x^2+y^2+z^2}, \varphi = \arctan\frac{y}{x}, \theta = \arccos\frac{z}{\sqrt{x^2+y^2+z^2}} \tag{1.55}$$

This change of variables has implications for the derivatives:

$$\begin{cases} \frac{\partial}{\partial x} = \sin\theta\cos\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\cos\varphi\frac{\partial}{\partial\theta} - \frac{1}{r\sin\theta}\sin\varphi\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial y} = \sin\theta\sin\varphi\frac{\partial}{\partial r} + \frac{1}{r}\cos\theta\sin\varphi\frac{\partial}{\partial\theta} + \frac{1}{r\sin\theta}\cos\varphi\frac{\partial}{\partial\varphi} \\ \frac{\partial}{\partial z} = \cos\theta\frac{\partial}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial}{\partial\theta} \end{cases} \tag{1.56}$$

We then have

$$J_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z} = \sin\varphi\frac{\partial}{\partial\theta} + \frac{\cos\varphi}{\tan\theta}\frac{\partial}{\partial\varphi} \tag{1.57}$$

$$J_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x} = -\cos\varphi\frac{\partial}{\partial\theta} + \frac{\sin\varphi}{\tan\theta}\frac{\partial}{\partial\varphi} \tag{1.58}$$

$$J_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} = -\frac{\partial}{\partial\varphi} \tag{1.59}$$

In total, this gives

$$J_\pm = \pm\exp(\pm i\varphi)\,(\frac{\partial}{\partial\theta} \pm \frac{i}{\tan\theta}\frac{\partial}{\partial\varphi}) \tag{1.60}$$

$$J_z = -i\frac{\partial}{\partial\varphi} \tag{1.61}$$

The Casimir operator then takes the form

$$C = -\frac{\partial^2}{\partial\theta^2} - \frac{\cos\theta}{\sin\theta}\frac{\partial}{\partial\theta} - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} \tag{1.62}$$

Note these very pleasing results: the $so(3)$ generators now depend only on the two variables θ and φ. Their eigenstates will therefore be functions of these two variables alone.

These functions are the so-called spherical harmonics:

$$Y_{l,m}(\theta,\varphi) \equiv \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}\, \exp(im\varphi)\, P_{l,m}(\cos\theta);\ m = -l, -l+1, \cdots, l-1, l \tag{1.63}$$

In this definition, the functions $P_{l,m}(\cos\theta)$ are the associated Legendre polynomials, i.e.

$$P_{l,m}(x) \equiv \frac{(-1)^m}{2^l l!}(1-x^2)^{\frac{m}{2}} \frac{d^{l+m}}{dx^{l+m}}(x^2-1)^l \tag{1.64}$$

Let's take a specific example to clarify this and consider $l = 1$.

We then have

$$Y_{1,1}(\theta,\varphi) = -\sqrt{\frac{3}{8\pi}}\exp(i\varphi)\sin\theta \equiv |1,1> \tag{1.65}$$

$$Y_{1,0}(\theta,\varphi) = \sqrt{\frac{3}{4\pi}}\cos\theta \equiv |1,0> \tag{1.66}$$

$$Y_{1,-1}(\theta,\varphi) = \sqrt{\frac{3}{8\pi}}\exp(-i\varphi)\sin\theta \equiv |1,-1> \tag{1.67}$$

It is then easy to verify that the typical relations of $so(3)$

$$J_z|1,1> = |1,1>, J_z|1,0> = 0, J_z|1,-1> = -|1,-1> \tag{1.68}$$

$$J_+|1,1> = 0, J_+|1,0> = \sqrt{2}|1,1>, J_+|1,-1> = \sqrt{2}|1,0> \tag{1.69}$$

$$J_-|1,1> = \sqrt{2}|1,0>, J_-|1,0> = \sqrt{2}|1,-1>, J_-|1,-1> = 0 \tag{1.70}$$

are satisfied.

The Casimir operator returns, of course, each of these functions twice.

So once again, it all seems abstract... but the fact remains: spherical harmonics are solutions to Schrödinger's equation when studying an electron. The numbers l and m therefore have very specific interpretations. These interpretations take place within the framework of orbitals, i.e., areas of space around the nucleus of an atom where the electron is most likely to be found. The number l, or azimuthal quantum number, describes the shape of the orbitals, while the

number m, or magnetic quantum number, describes the orientation of these orbitals.

An illustration would be welcome! We consider the cases $l = 0, 1, 2$ but moving towards real spherical harmonics[3]:

$$Y^R_{l,m} \equiv \begin{cases} \frac{i}{\sqrt{2}}(Y_{l,m} - (-1)^m Y_{l,-m}) \; if \; m < 0 \\ Y_{l,0} \; if \; m = 0 \\ \frac{1}{\sqrt{2}}(Y_{l,-m} + (-1)^m Y_{l,m}) \; if \; m > 0 \end{cases} \tag{1.71}$$

We obtain:

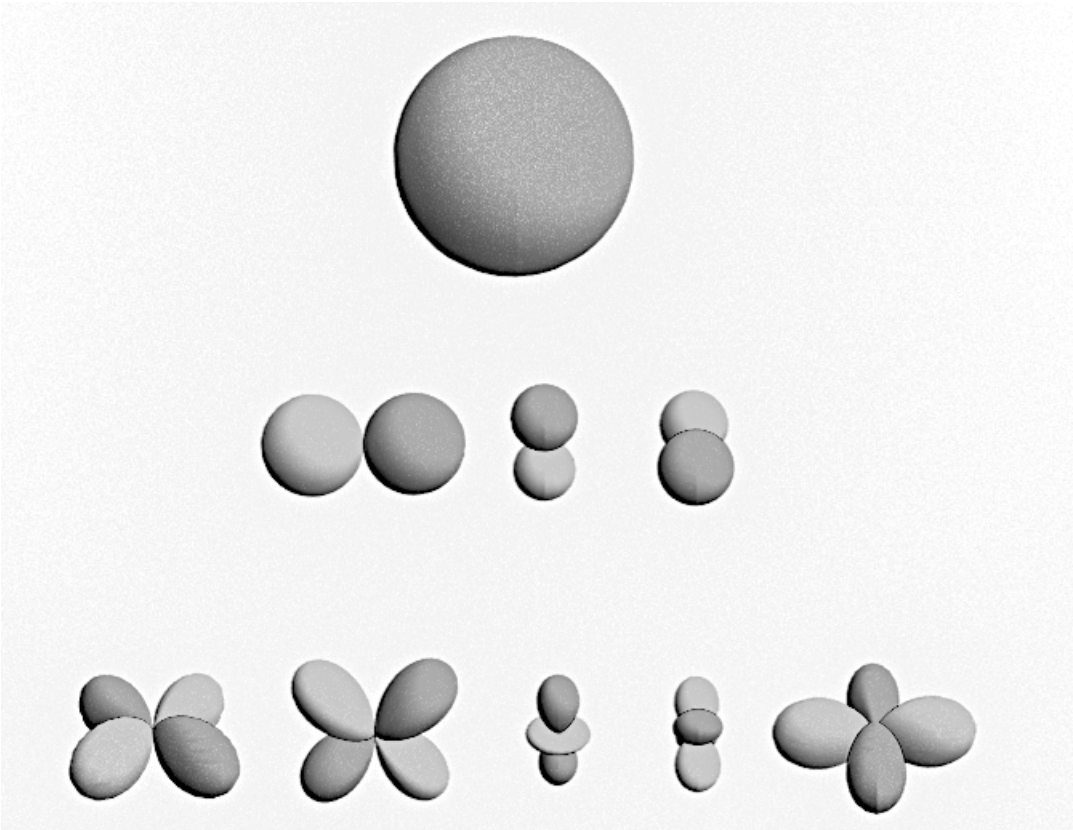


Figure 1.2: Real spherical harmonics for $l = 0, 1, 2$

Mathematics is beautiful, isn't it?

Because you have to realize that the Norwegian mathematician Sophus Lie developed the algebras that bear his name at the end of the 19th century. When

[3] Just a graphic requirement...

he died in 1899, it would be another quarter of a century before Schrödinger's equation saw the light of day! In other words, Lie algebras were completely unknown to most physicists at the time. And there were probably lots of people who thought that "this stuff will never be of any use"!

However, even at this stage, using one of the simplest algebras possible, we can already see how little vision these people had...

And that's not all!

The most attentive among you will have noticed that we mentioned the number l and not j in the representations of $so(3)$ dealing with the angular momentum... The difference? The number l is a positive integer, while j also takes the values 1/2, 3/2, 5/2, etc.

Where did these values go?

To another angular momentum that we will discuss right now.

1.6.2 Intrinsic angular momentum or spin

In 1927, German physicists Otto Stern and Walther Gerlach devised an experiment that would revolutionize the world of quantum physics.

A chamber was heated to a temperature of around 1200 K. Silver atoms were then expelled from it. They moved toward slits (each 1 mm long) designed to select atoms traveling in a certain direction. The selected atoms then arrived at the air gap of an electromagnet, whose shape was chosen so that the magnetic field created was not uniform. The atomic jet was then deflected by this field and arrived at a screen.

It was expected that the atoms would be evenly distributed across the screen. This was not the case: Stern and Gerlach actually observed two spots, resulting from the impacts of the atoms.

This is typical of quantum theory: instead of continuous values, only a finite number of values, two in this case, are obtained.

To explain this experimental result, American physicists George Uhlenbeck and Samuel Goudsmit proposed defining an intrinsic angular momentum or spin (from the verb "to spin").

The spin of an electron is often imagined as an arrow linked to a ball rotating clockwise or counterclockwise:

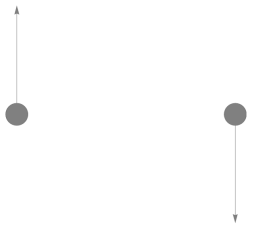


Figure 1.3: Spin up (counterclockwise rotation) and spin down (clockwise rotation) states for the electron

While this representation is convenient, it is, however, very simplistic: positions are only marked by probabilities, and Heisenberg's uncertainty principle destroys the notion of a vector as an oriented segment. But never mind... we can keep it in what follows.

These two possibilities of rotation, and therefore orientation of the spin "arrow", are found in Stern and Gerlach's experiment. We will refer to "spin up" (arrow pointing up) and "spin down" (arrow pointing down).

How can this be translated from a "Lie algebra" perspective?

Rotation means $so(3)$. And two states are needed to mark spin up and spin down. So we need to use $j = \frac{1}{2}$.

The representation $D(\frac{1}{2})$ of $so(3)$, which, as you will recall, highlights two states, which are realized by Eqs.(1.40) while the matrices of the algebra are given by Eqs. (1.41)-(1.42).

The first state $|\frac{1}{2}, \frac{1}{2}>$, eigenvector of J_z with eigenvalue $\frac{1}{2}$, will be referred to as "spin up state". The second state $|\frac{1}{2}, -\frac{1}{2}>$, eigenvector of J_z with eigenvalue $-\frac{1}{2}$, will be referred to as "spin down state".

We should also note an unexpected and immediate consequence: if we want to describe a quantum object, a particle, with spin, we must assign it a matrix equation.

This fact annoyed quite a few physicists at the time, since Schrödinger's equation does not have this matrix character... We will come back to this later.

1.6.3 Tell me how you transform yourself, and I will tell you who you are.

This is a curious saying, yet it couldn't be more true in quantum mechanics.

For example, let's focus on the representation $D(\frac{1}{2})$ of $so(3)$. The rotation around $\overrightarrow{e_z}$ can be written as

$$J_3 = -\frac{i}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1.72}$$

and his transition to the group is marked by exponential growth:

$$R_3 = I + \theta J_3 + \frac{1}{2}\theta^2 J_3^2 + \cdots = \begin{pmatrix} \exp(-\frac{i}{2}\theta) & 0 \\ 0 & \exp(\frac{i}{2}\theta) \end{pmatrix} \tag{1.73}$$

We then have

$$R_3\,|\frac{1}{2}, \frac{1}{2}>= \exp(-\frac{i}{2}\theta)\,|\frac{1}{2}, \frac{1}{2}> \tag{1.74}$$

Let's imagine the situation: we make a complete turn, which corresponds to $\theta = 2\pi$... what happens? The state $|\frac{1}{2}, \frac{1}{2}>$ does not return to itself but to its opposite! It takes two turns and $\theta = 4\pi$ to return to its original state.

This is also true for the spin down state, since

$$R_3 \,|\frac{1}{2}, -\frac{1}{2}> = \exp(\frac{i}{2}\theta)\,|\frac{1}{2}, -\frac{1}{2}> \tag{1.75}$$

A single turn still returns to the opposite state, and only after two turns does it return to its original state.

It is strange behavior to have to make two turns to return to its starting state... we don't often notice this in our daily lives!

And this is what distinguishes the states of representation $D(\frac{1}{2})$ from those of representation $D(1)$.

Indeed, for $D(1)$, moving from algebra to group and focusing again on rotation with respect to $\overrightarrow{e_z}$, we have

$$J_3 = -i \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \Leftrightarrow R_3 = \begin{pmatrix} \exp(-i\theta) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \exp(i\theta) \end{pmatrix} \tag{1.76}$$

So, for instance, we get

$$R_3 \,|1,1> = \exp(-i\theta)\,|1,1> \tag{1.77}$$

It is clear that, for a complete turn ($\theta = 2\pi$), the state $|1,1>$ transforms into itself. And this is what we expect in our three-dimensional space: a vector makes a complete turn on itself to return to its initial position.

To distinguish between these different behaviors, we will use different names.

We will say that the states associated with the representation $D(\frac{1}{2})$ of $so(3)$ are called *spinors*, while those corresponding to the representation $D(1)$ are called *vectors*.

1.7 The Lie group $SO(1,3)$ and the Lie algebra $so(1,3)$

We have seen that rotations preserve the norm $x^2 + y^2 + z^2$ of 3D-Euclidean space. Let us now consider transformations that preserve the pseudo-norm $t^2 - x^2 - y^2 - z^2$ of spacetime. The rotations seen previously clearly do the job. But we need to add three other transformations, which we present hereafter.

1.7.1 Lorentz boosts

Let us consider a reference frame and a second reference frame moving in uniform linear motion at velocity $\overrightarrow{v_0}$ relative to the first one.

The spacetime coordinates of a macroscopic object will differ depending on whether the object is placed in the first reference frame or the second. More precisely, if $(t, \overrightarrow{r})$ are the coordinates of this object in the first reference frame and $(t\prime, \overrightarrow{r\prime})$ are those of the same object in the second reference frame, we have

$$\begin{cases} t\prime = t \\ \overrightarrow{r\prime} = \overrightarrow{r} - \overrightarrow{v_0}t \end{cases} \tag{1.78}$$

These equations are known as the Galilean transformations. They have the consequence that

$$\overrightarrow{v\prime} = \overrightarrow{v} - \overrightarrow{v_0} \tag{1.79}$$

Therefore, the speed of this object is not the same from one frame of reference to another. This is not important as long as the object is macroscopic, but it becomes much more delicate if the speed of the object, this time microscopic, is close to that of light. Indeed, in 1905, Albert Einstein established the theory of special relativity by postulating that the speed of light must be constant, regardless of the reference frame. So there is a problem!

It was Dutch physicist Hendrik Lorentz who provided the solution to this problem.

The speed of movement of the second reference frame is fixed along the x-axis to simplify the calculations:

$$\overrightarrow{v_0} = v_0 \overrightarrow{e_x} \tag{1.80}$$

Eqs (1.78) are modified by Lorentz to become

$$\begin{cases} ct\prime = \gamma(ct - \beta x) \\ x\prime = \gamma(x - \beta ct) \end{cases} \tag{1.81}$$

Here c denotes, as usual, the speed of light while β and γ are defined by

$$\beta \equiv \frac{v}{c} \; ; \; \gamma \equiv \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{1.82}$$

The number γ was, in fact, set to preserve the pseudo-norm:

$$c^2t\prime^2 - x\prime^2 = \gamma^2(1-\beta^2)(c^2t^2 - x^2) \Rightarrow \gamma^2(1-\beta^2) = 1 \tag{1.83}$$

The relation $\gamma^2(1-\beta^2) = 1$ suggests setting

$$\gamma \equiv \cosh\varphi \; ; \; \beta\gamma \equiv -\sinh\varphi \tag{1.84}$$

Lorentz boosts are then finally written as

$$\begin{cases} ct\prime = (\cosh\varphi)ct + (\sinh\varphi)x \\ x\prime = (\sinh\varphi)ct + (\cosh\varphi)x \end{cases} \tag{1.85}$$

When we leave the plane (ct, x) to move to the entirety of spacetime, setting $c = 1$, we ultimately arrive at the three boosts, to which we must add the three rotations already mentioned, giving us six Lorentz transformations:

- Three rotations

$$R_1(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix}, R_2(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & \sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{1.86}$$

$$R_3(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.87}$$

- Three boosts

$$B_1(\varphi) = \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, B_2(\varphi) = \begin{pmatrix} \cosh\varphi & 0 & \sinh\varphi & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\varphi & 0 & \cosh\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.88}$$

$$B_3(\varphi) = \begin{pmatrix} \cosh\varphi & 0 & 0 & \sinh\varphi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\varphi & 0 & 0 & \cosh\varphi \end{pmatrix} \tag{1.89}$$

The pseudo-norm of a 4-vector in spacetime

$$r = \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \tag{1.90}$$

being given by

$$r^T g\, r \;;\; g = diag(1,-1,-1,-1) \tag{1.91}$$

conserving it requests a transformation involving spacetime coordinates:

$$r\prime = \Lambda\, r \tag{1.92}$$

This transformation is such that

$$r\prime^T g\, r\prime = r^T \Lambda^T g\, \Lambda\, r = r^T g\, r$$

In other words, we must impose

$$\boxed{\Lambda^T g\, \Lambda = g} \tag{1.93}$$

It can be easily verified that the six matrices (1.86)-(1.89) have this characteristic in addition to having a determinant equal to 1. They are said to form the $SO(1,3)$ group or Lorentz group[4].

The group law is obtained by considering a sequence of two transformations:

$$r'' = \Lambda_1 r' = \Lambda_1 \Lambda_2 r$$

We have indeed

$$(\Lambda_1 \Lambda_2)^T g \, (\Lambda_1 \Lambda_2) = \Lambda_2^T \Lambda_1^T g \, \Lambda_1 \Lambda_2 = \Lambda_2^T g \, \Lambda_2 = g$$

The group law therefore corresponds to simple multiplication

$$\Lambda_1 \Lambda_2 \tag{1.94}$$

The existence of the neutral element and the symmetric element are ensured via the identity I and the inverse Λ^{-1}, respectively. Associativity is not a problem either.

We are therefore indeed dealing with a group.

1.7.2 Lorentz algebra and its representations

The six generators of the algebra corresponding to the Lorentz group, which is obviously called the Lorentz algebra or $so(1,3)$, are obtained from the matrices of the group, considering extremely small angles:

$$J_1 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.95}$$

$$K_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{1.96}$$

Note: some of these matrices are Hermitian and others are not; if we multiply them by the imaginary number i as recommended by physicists, nothing changes except that this time the rotations are Hermitian while the boosts are not. This observation is related to the fact that the Lorentz group is not compact, since there is no limitation on the values of the boost parameters. In other words, there is no finite-dimensional IUR of the Lorentz group. We must therefore expect a mixture of Hermitian and anti-Hermitian matrices at the algebra level.

[4]The denomination $(1,3)$ is directly related to the matrix g: one $+$ sign and three $-$ signs.

The non-zero commutation relations are then

$$[J_1, J_2] = J_3\ ,\ [J_3, J_1] = J_2\ ,\ [J_2, J_3] = J_1 \tag{1.97}$$

$$[J_1, K_2] = K_3\ ,\ [J_1, K_3] = -K_2\ ,\ [J_2, K_1] = -K_3 \tag{1.98}$$

$$[J_2, K_3] = K_1\ ,\ [J_3, K_1] = K_2\ ,\ [J_3, K_2] = -K_1 \tag{1.99}$$

$$[K_1, K_2] = -J_3\ ,\ [K_3, K_1] = -J_2\ ,\ [K_2, K_3] = -J_1 \tag{1.100}$$

Two observations at this stage.

The first is that algebra is invariant under changes

$$J_j \to J_j\ ;\ K_j \to -K_j\ ;\ j = 1,\ 2,\ 3 \tag{1.101}$$

The second is that we recognize, in the relations (1.97), the algebra $so(3)$. This suggests defining

$$A_j \equiv \frac{1}{2}(J_j + iK_j)\ ;\ j = 1,\ 2,\ 3 \tag{1.102}$$

$$B_j \equiv \frac{1}{2}(J_j - iK_j)\ ;\ j = 1,\ 2,\ 3 \tag{1.103}$$

It is therefore not difficult to convince oneself that

$$[A_j, A_k] = \varepsilon_{jkl}\, A_l\ ;\ [B_j, B_k] = \varepsilon_{jkl}\, B_l\ ;\ j,\ k,\ l = 1,\ 2,\ 3 \tag{1.104}$$

$$[A_j, B_k] = 0, \forall j, k = 1, 2, 3 \tag{1.105}$$

And this is a rather nice result because it means that we have two algebras $so(3)$ in direct sum (which is marked by Eq. (1.105)).

Simply proceed as before when we were only dealing with rotations and a single $so(3)$ and move on to complexification:

$$A_+ \equiv iA_1 - A_2; A_- \equiv iA_1 + A_2; A_z \equiv iA_3 \tag{1.106}$$

$$B_+ \equiv iB_1 - B_2; B_- \equiv iB_1 + B_2; B_z \equiv iB_3 \tag{1.107}$$

Each of these sets leads to a representation $D(j)$ of $so(3)$...the two sets will therefore lead to $D(j1, j2)$, which is of dimension $(2j1 + 1)(2j2 + 1)$.

The explicit relationships are obviously given by

$$A_+|j1, j2, m1, m2 >= \sqrt{(j1 - m1)(j1 + m1 + 1)}|j1, j2, m1 + 1, m2 > \tag{1.108}$$

$$A_-|j1,j2,m1,m2>=\sqrt{(j1+m1)(j1-m1+1)}|j1,j2,m1-1,m2> \quad (1.109)$$

$$A_z|j1,j2,m1,m2>=m1|j1,j2,m1,m2> \quad (1.110)$$

$$B_+|j1,j2,m1,m2>=\sqrt{(j2-m2)(j2+m2+1)}|j1,j2,m1,m2+1> \quad (1.111)$$

$$B_-|j1,j2,m1,m2>=\sqrt{(j2+m2)(j2-m2+1)}|j1,j2,m1,m2-1> \quad (1.112)$$

$$B_z|j1,j2,m1,m2>=m2|j1,j2,m1,m2> \quad (1.113)$$

Let's focus on a few specific representations.

1.7.2.1 The representation $D(\frac{1}{2},0)$ or left spinor representation

It is of dimension 2 and is characterized by

$$A_+=\begin{pmatrix}0&1\\0&0\end{pmatrix}, A_-=\begin{pmatrix}0&0\\1&0\end{pmatrix}, A_z=\frac{1}{2}\begin{pmatrix}1&0\\0&-1\end{pmatrix} \quad (1.114)$$

the generators $B_\pm$ and B_z reducing to zero. This leads to

$$\overrightarrow{J}=-\frac{i}{2}\overrightarrow{\sigma}\ ;\ \overrightarrow{K}=-\frac{1}{2}\overrightarrow{\sigma} \quad (1.115)$$

where the notation $\overrightarrow{\sigma}$ stands for the Pauli matrices.

At the group level, you need to move to exponentiation:

$$\overrightarrow{R}=\exp(\theta\overrightarrow{J})\ ;\ \overrightarrow{B}=\exp(\varphi\overrightarrow{K}) \quad (1.116)$$

In detail, this gives

$$R_1=\begin{pmatrix}\cos\frac{\theta}{2}&-i\sin\frac{\theta}{2}\\-i\sin\frac{\theta}{2}&\cos\frac{\theta}{2}\end{pmatrix}, R_2=\begin{pmatrix}\cos\frac{\theta}{2}&-\sin\frac{\theta}{2}\\\sin\frac{\theta}{2}&\cos\frac{\theta}{2}\end{pmatrix} \quad (1.117)$$

$$R_3=\begin{pmatrix}\exp(-\frac{i}{2}\theta)&0\\0&\exp(\frac{i}{2}\theta)\end{pmatrix} \quad (1.118)$$

$$B_1=\begin{pmatrix}\cosh\frac{\varphi}{2}&-\sinh\frac{\varphi}{2}\\-\sinh\frac{\varphi}{2}&\cosh\frac{\varphi}{2}\end{pmatrix}, B_2=\begin{pmatrix}\cosh\frac{\varphi}{2}&i\sinh\frac{\varphi}{2}\\-i\sinh\frac{\varphi}{2}&\cosh\frac{\varphi}{2}\end{pmatrix} \quad (1.119)$$

$$B_3=\begin{pmatrix}\exp(-\frac{1}{2}\varphi)&0\\0&\exp(\frac{1}{2}\varphi)\end{pmatrix} \quad (1.120)$$

These matrices are those of the group $SL(2,C)$, i.e., the set of 2×2 matrices with complex entries and determinant equal to 1.

The representation $D(\frac{1}{2},0)$ is called left, and the corresponding matrices of $SL(2,C)$ act on spinors that are described as *left-handed*.

1.7.2.2 The representation $D(0,\frac{1}{2})$ or right spinor representation

It is also of dimension 2 and is characterized by

$$\overrightarrow{J} = -\frac{i}{2}\overrightarrow{\sigma}\ ;\ \overrightarrow{K} = \frac{1}{2}\overrightarrow{\sigma} \tag{1.121}$$

Although it corresponds to the previous one by a simple change of sign in $\overrightarrow{K}$ and a simple change of sign in φ for the group, this representation is different from the representation $D(\frac{1}{2},0)$.

It is said to be right-handed, and the associated matrices of $SL(2,C)$ act on spinors that are described as *right-handed.*

1.7.2.3 The representation $D(\frac{1}{2},0)\oplus D(0,\frac{1}{2})$

It is simply the superposition of the two previous ones:

$$\overrightarrow{J} = -\frac{i}{2}\begin{pmatrix} \overrightarrow{\sigma} & 0 \\ 0 & \overrightarrow{\sigma} \end{pmatrix}\ ;\ \overrightarrow{K} = \frac{1}{2}\begin{pmatrix} -\overrightarrow{\sigma} & 0 \\ 0 & \overrightarrow{\sigma} \end{pmatrix} \tag{1.122}$$

It is, of course, 4-dimensional.

Another representation of dimension 4, not equivalent to the previous one, is as follows.

1.7.2.4 The representation $\mathbf{D}(\frac{1}{2},\frac{1}{2})$ or vector representation

States can be generated according to

$$|\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2}>= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |\frac{1}{2},\frac{1}{2},-\frac{1}{2},\frac{1}{2}>= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{1.123}$$

$$|\frac{1}{2},\frac{1}{2},\frac{1}{2},-\frac{1}{2}>= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |\frac{1}{2},\frac{1}{2},-\frac{1}{2},-\frac{1}{2}>= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.124}$$

Matrices of the Lorentz algebra are then given by

$$J_1 = -\frac{i}{2}\begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, J_2 = \frac{1}{2}\begin{pmatrix} 0 & -1 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \tag{1.125}$$

$$J_3 = \begin{pmatrix} -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \end{pmatrix}, K_1 = \frac{1}{2}\begin{pmatrix} 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{pmatrix} \tag{1.126}$$

$$K_2 = \frac{i}{2}\begin{pmatrix} 0 & 1 & -1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{pmatrix}, K_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.127}$$

That's for matrix realizations... but, as we mentioned for $so(3)$ and the orbital angular momemtum, there are also realizations based on differential operators:

$$J_1 = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z},\ J_2 = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x},\ J_3 = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y} \tag{1.128}$$

$$K_1 = t\frac{\partial}{\partial x} + x\frac{\partial}{\partial t},\ K_2 = t\frac{\partial}{\partial y} + y\frac{\partial}{\partial t},\ K_3 = t\frac{\partial}{\partial z} + z\frac{\partial}{\partial t} \tag{1.129}$$

It is a representation of infinite dimension.

There is a more elegant way to write these generators. It uses relativistic standards:

$$x^\mu \to (t, x, y, z)\ ;\ x_\mu \to (t, -x - y - z) \tag{1.130}$$

$$\partial^\mu \to (\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z})\ ;\ \partial_\mu \to (\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})\ ;\ \mu = 0, 1, 2, 3 \tag{1.131}$$

All operators are then written uniformly as

$$\boxed{J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu)} \tag{1.132}$$

having defined

$$J^{23} \equiv iJ_1,\ J^{31} \equiv iJ_2,\ J^{12} \equiv iJ_3,\ J^{01} \equiv -iK_1,\ J^{02} \equiv -iK_2,\ J^{03} \equiv -iK_3 \tag{1.133}$$

The presence of imaginary i clearly calls for the complexity of Lorentz algebra, but it is necessary if we want to consider Hermitian generators.

The commutation relations of the algebra can then be summarized by

$$\boxed{[J^{\mu\nu}, J^{\varrho\sigma}] = i(g^{\nu\varrho}J^{\mu\sigma} + g^{\mu\sigma}J^{\nu\varrho} + g^{\mu\varrho}J^{\sigma\nu} + g^{\nu\sigma}J^{\varrho\mu})} \tag{1.134}$$

where g is, of course, the metric (1.91).

1.7.3 Connected components of the Lorentz group

As already mentioned, Lorentz transformations are characterized by

$$r\prime = \Lambda\, r \;;\; \Lambda^T g\, \Lambda = g \tag{1.135}$$

There are, in fact, four connected components depending on whether

$$det\, \Lambda = 1 \;;\; \Lambda_{00} \geq 1 \tag{1.136}$$

$$det\, \Lambda = -1 \;;\; \Lambda_{00} \geq 1 \tag{1.137}$$

$$det\, \Lambda = 1 \;;\; \Lambda_{00} \leq -1 \tag{1.138}$$

$$det\, \Lambda = -1 \;;\; \Lambda_{00} \leq -1 \tag{1.139}$$

The sign of the determinant leads to transformations that are either proper ($det\, \Lambda = 1$) or improper ($det\, \Lambda = -1$); the sign of the element at the intersection of the first row and the first column leads to transformations that are either orthochronous ($\Lambda_{00} \geq 1$) or retrochronous ($\Lambda_{00} \leq -1$).

These components come from the definition of the group. For example, it is clear from the definition (1.93) that

$$det(\Lambda^T g\, \Lambda) = det(g) \Leftrightarrow det(\Lambda^T) det(g) det(\Lambda) = det(g) \Leftrightarrow (det(\Lambda))^2 = 1$$

leading to both possibilities of sign.

The two inequalities for Λ_{00} also result from the definition (1.93) of the Lorentz group. Indeed, it can be written, successively, as

$$(\Lambda^T g\, \Lambda)_{\alpha\beta} = g_{\alpha\beta}$$

$$\Leftrightarrow (\Lambda^T g)_{\alpha\nu} \Lambda^\nu_\beta = g_{\alpha\beta}$$

$$\Leftrightarrow (\Lambda^T)^\mu_\alpha g_{\mu\nu} \Lambda^\nu_\beta = g_{\alpha\beta}$$

$$\Leftrightarrow g_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta = g_{\alpha\beta}$$

$$\Leftrightarrow \Lambda_{\nu\alpha} \Lambda^\nu_\beta = g_{\alpha\beta}$$

Equivalently, by "raising" the index α, we have

$$\Lambda^\alpha_\nu \Lambda^\nu_\beta = g^\alpha_\beta = \delta_{\alpha\beta} \tag{1.140}$$

And, taking $\alpha = \beta = 0$, this implies that

$$(\Lambda_{00})^2 - \sum (\Lambda_{j0})^2 = 1 \Leftrightarrow (\Lambda_{00})^2 = 1 + \sum (\Lambda_{j0})^2 \geq 1$$

This will convince you of the two possibilities: $\Lambda_{00} \geq 1$ and $\Lambda_{00} \leq -1$.

The four connected components are linked together by discrete symmetry operators. We haven't talked much about these operators so far, but they will play a crucial role later on. They represent only two possible positions, like being on the right or on the left, with no intermediate position possible.

So let us consider the parity operator P, which affects the position 4-vector as follows

$$P \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} t \\ -x \\ -y \\ -z \end{pmatrix} \tag{1.141}$$

and which can therefore be achieved by

$$P = diag(1, -1, -1, -1) \tag{1.142}$$

We see that its action on the group matrices is to change the sign of the determinant, without changing that of Λ_{00}, i.e., to give rise to improper and orthochronous transformations.

A second operator is the time-reversal operator T:

$$T \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -t \\ x \\ y \\ z \end{pmatrix} \tag{1.143}$$

which corresponds to

$$T = diag(-1, 1, 1, 1) \tag{1.144}$$

It leads to improper and retrochronous transformations.

Finally, the combination of the two operators P and T leads to the fourth connected component of the Lorentz group.

Of course, these operators P and T are also part of the Lorentz group since

$$P^T g\, P = g \; ; \; T^T g\, T = g \tag{1.145}$$

We will have the opportunity to return to these operators later, but for now, let us remember this:

	Orthochronous	Retrochronous
Proper	Λ	$PT\Lambda$
Improper	$P\Lambda$	$T\Lambda$

Table 1.1: Connected components of the Lorentz group

1.8 The Poincaré group and the Poincaré algebra

Poincaré's group is interested in proper orthochronous transformations and adds translations to them. There are four translations, since there are four spacetime coordinates. Their action is as follows:

$$r\prime = r + a \tag{1.146}$$

Combined with Lorentz transformations, translations lead to the transformation law

$$\boxed{r\prime = \Lambda\, r + a} \tag{1.147}$$

This law effectively leads to a group since

$$r\prime\prime = \Lambda\prime\, r\prime + a\prime = \Lambda\prime\,(\Lambda\, r + a) + a\prime = \Lambda\prime\prime\, r + a\prime\prime \tag{1.148}$$

The group law is thus given by

$$\boxed{\left\{ \begin{array}{c} \Lambda\prime\prime = \Lambda\prime\,\Lambda \\ a\prime\prime = \Lambda\prime\, a + a\prime \end{array} \right.} \tag{1.149}$$

The existence of a neutral is ensured via

$$\Lambda = I \;;\; a = 0 \tag{1.150}$$

and the existence of a symmetric element as well, with

$$\Lambda\prime = \Lambda^{-1} \;;\; a\prime = -\Lambda^{-1} a \tag{1.151}$$

Associativity can also be demonstrated without any difficulty. We are therefore indeed dealing with a group.

1.8.1 The Poincaré algebra

What interests us are the irreducible unitary representations of this group. The Lorentz group was not compact, so there is no chance that Poincaré's group will be either. Consequently, the IURs of the Poincaré group, which we will denote $U(\Lambda, a)$, are infinite-dimensional. Nevertheless, it may be useful to view them in the following matrix form

$$U(\Lambda, a) = \left(\begin{array}{cc} \Lambda & a \\ 0 & 1 \end{array} \right) \tag{1.152}$$

Thus, group law is manifested through

$$U(\Lambda\prime, a\prime)\, U(\Lambda, a) = \left(\begin{array}{cc} \Lambda\prime & a\prime \\ 0 & 1 \end{array} \right) \left(\begin{array}{cc} \Lambda & a \\ 0 & 1 \end{array} \right) = \left(\begin{array}{cc} \Lambda\prime\Lambda & \Lambda\prime a + a\prime \\ 0 & 1 \end{array} \right) = U(\Lambda\prime\Lambda, \Lambda\prime a + a\prime)$$

Any irreducible unitary representation $U(\Lambda, a)$ can therefore be decomposed into

$$U(\Lambda, a) = U(I, a)\, U(\Lambda, 0) \equiv U(a)\, U(\Lambda) = \begin{pmatrix} I & a \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Lambda & 0 \\ 0 & 1 \end{pmatrix} \tag{1.153}$$

The advantage is, of course, that translations can be treated independently of other transformations.

In this regard, it should be noted that

$$U(a)\, U(a\prime) = U(I, a)\, U(I, a\prime) = U(I, a + a\prime) \tag{1.154}$$

At the last equality, we applied the group law. To continue, it suffices to note that since addition of numbers is commutative, we have

$$U(a)\, U(a\prime) = U(I, a + a\prime) = U(I, a\prime + a) = U(I, a\prime)\, U(I, a) = U(a\prime)\, U(a) \tag{1.155}$$

The operators associated with translations therefore commute.

If we move on to algebra via exponentiation

$$U(a) = \exp(-i a_\mu P^\mu) \tag{1.156}$$

we thus obtain the commutation relations that govern the generators of infinitesimal translations:

$$[P^\mu, P^\nu] = 0 \tag{1.157}$$

Let's move on to the next part.

We have

$$U(\Lambda, a)\, U(\Lambda, a) = U(\Lambda^2, \Lambda a + a) \tag{1.158}$$

by the group law. As we have seen in (1.153), the second member of this equality can be factorized

$$U(\Lambda, a)\, U(\Lambda, a) = U(\Lambda a + a)\, U(\Lambda^2) \tag{1.159}$$

This is also true for translations, as we have just demonstrated:

$$U(\Lambda, a)\, U(\Lambda, a) = U(\Lambda a)\, U(a)\, U(\Lambda^2) \tag{1.160}$$

Still according to the group law, it is also possible to split the "Lorentz" part of the second member:

$$U(\Lambda, a)\, U(\Lambda, a) = U(\Lambda a)\, U(a)\, U(\Lambda)\, U(\Lambda) \tag{1.161}$$

But it's not just the second member! The first member can also be broken down according to the conventions adopted. As a result, we get

$$U(a)\,U(\Lambda)\,U(a)\,U(\Lambda) = U(a)\,U(\Lambda a)\,U(\Lambda)\,U(\Lambda) \tag{1.162}$$

The modification of the second term comes from the commutativity of translations.

Now remove the extremes, on the right and left, by applying their respective inverses, and you will obtain

$$U(\Lambda)\,U(a) = U(\Lambda a)\,U(\Lambda) \tag{1.163}$$

or

$$U(\Lambda)\,U(a)\,U^{-1}(\Lambda) = U(\Lambda a) \tag{1.164}$$

This result clearly indicates that translation operators do not commute with Lorentz transformation operators.

Let us now move on to the infinitesimal form to clarify this observation. We introduce parameters ω and ε that are negligible to second order:

$$\Lambda^{\mu}_{\ \nu} = \delta^{\mu}_{\nu} + \omega^{\mu}_{\ \nu}\ ;\ a^{\mu} = \varepsilon^{\mu} \tag{1.165}$$

Next, let us consider

$$U(\Lambda, a)\,U(1+\omega, \varepsilon)\,U^{-1}(\Lambda, a)$$

With the inverse calculation that was performed in (1.151), this gives

$$U(\Lambda, a)\,U(1+\omega, \varepsilon)\,U(\Lambda^{-1}, -\Lambda^{-1}a)$$

and, by applying the group law twice, we are led to

$$U(\Lambda(1+\omega), \Lambda\varepsilon+a)\,U(\Lambda^{-1}, -\Lambda^{-1}a) = U(\Lambda(1+\omega)\Lambda^{-1}, \Lambda(1+\omega)(-\Lambda^{-1}a)+\Lambda\varepsilon+a)$$

In total, we therefore have

$$U(\Lambda, a)\,U(1+\omega, \varepsilon)\,U^{-1}(\Lambda, a) = U(1+\Lambda\omega\Lambda^{-1}, -\Lambda\omega\Lambda^{-1}a+\Lambda\varepsilon) \tag{1.166}$$

If we use the operators of algebra in infinitesimal form

$$U(1+\omega, \varepsilon) = 1 - \frac{i}{2}\,\omega_{\varrho\sigma}\,J^{\varrho\sigma} - i\,\varepsilon_{\varrho}\,P^{\varrho} \tag{1.167}$$

this last result is written as

$$U(\Lambda, a)(\frac{1}{2}\omega_{\varrho\sigma}J^{\varrho\sigma} + \varepsilon_{\varrho}P^{\varrho})U^{-1}(\Lambda, a) = \frac{1}{2}(\Lambda\omega\Lambda^{-1})_{\mu\nu}J^{\mu\nu} + (-\Lambda\omega\Lambda^{-1}a+\Lambda\varepsilon)_{\mu}P^{\mu}$$

The terms dependent on ε are therefore

$$U(\Lambda, a)\, P^{\varrho}\, U^{-1}(\Lambda, a) = \Lambda_{\mu}^{\ \varrho}\, P^{\mu} \tag{1.168}$$

while those dependent on ω lead to

$$U(\Lambda, a)\, J^{\varrho\sigma}\, U^{-1}(\Lambda, a) = \Lambda_{\mu}^{\ \varrho}\, \Lambda_{\nu}^{\ \sigma}\, (J^{\mu\nu} + a^{\mu} P^{\nu} - a^{\nu} P^{\mu}) \tag{1.169}$$

The last relation comes from (1.140) compared with

$$\Lambda^{-1}\Lambda = I \Leftrightarrow (\Lambda^{-1})_{\nu}^{\alpha}\Lambda_{\beta}^{\nu} = \delta_{\alpha\beta} \tag{1.170}$$

giving then

$$\Lambda_{\nu}^{\alpha} = (\Lambda^{-1})_{\nu}^{\alpha} \tag{1.171}$$

We then have successively

$$(\Lambda\omega\Lambda^{-1})_{\mu\nu}J^{\mu\nu} = (\Lambda\omega)_{\mu}^{\sigma}(\Lambda^{-1})_{\sigma\nu}J^{\mu\nu} = \Lambda_{\mu}^{\varrho}\omega_{\varrho}^{\ \sigma}(\Lambda^{-1})_{\sigma\nu}J^{\mu\nu} = \Lambda_{\mu}^{\varrho}\omega_{\varrho\sigma}(\Lambda^{-1})_{\nu}^{\sigma}J^{\mu\nu}$$

and therefore

$$(\Lambda\omega\Lambda^{-1})_{\mu\nu}J^{\mu\nu} = \omega_{\varrho\sigma}\, \Lambda_{\mu}^{\ \varrho}\, \Lambda_{\nu}^{\ \sigma}\, J^{\mu\nu}$$

The other term of the second member is proved in the same way.

We are therefore faced with the system

$$\begin{cases} U(\Lambda, a) P^{\varrho} U^{-1}(\Lambda, a) = \Lambda_{\mu}^{\ \varrho} P^{\mu} \\ U(\Lambda, a) J^{\varrho\sigma} U^{-1}(\Lambda, a) = \Lambda_{\mu}^{\varrho}\Lambda_{\nu}^{\sigma}(J^{\mu\nu} + a^{\mu}P^{\nu} - a^{\nu}P^{\mu}) \end{cases} \tag{1.172}$$

We still need to consider the infinitesimal form for $U(\Lambda, a)$ as for $U^{-1}(\Lambda, a)$, remembering that these are unitary transformations. The inverse will therefore be marked by a change of sign in the argument and ultimately lead to a commutator.

The first equation of (1.172) is written as follows

$$-i\,[\frac{1}{2}\omega_{\mu\nu}J^{\mu\nu} + \varepsilon_{\mu}P^{\mu}, P^{\varrho}] = \omega_{\mu}^{\ \varrho}\, P^{\mu}$$

Recalling that operators associated with translations commute, this leads to

$$[J^{\mu\nu}, P^{\varrho}] = -i\,(g^{\mu\varrho}P^{\nu} - g^{\nu\varrho}P^{\mu})$$

The second equation of (1.172) becomes, for its part,

$$-i\,[\frac{1}{2}\omega_{\mu\nu}J^{\mu\nu} + \varepsilon_{\mu}P^{\mu}, J^{\varrho\sigma}] = \omega_{\mu}^{\ \varrho}J^{\mu\sigma} + \omega_{\nu}^{\ \sigma}J^{\varrho\nu} + \varepsilon^{\varrho}P^{\sigma} - \varepsilon^{\sigma}P^{\varrho}$$

since we eliminate what is of second order and more in terms of ε and ω.

We then take into account the commutation relations between translations and Lorentz transformations so that we are left with only

$$\frac{i}{2}\,\omega_{\mu\nu}\,[J^{\mu\nu}, J^{\varrho\sigma}] = -\omega_{\mu}^{\ \varrho} J^{\mu\sigma} - \omega_{\nu}^{\ \sigma} J^{\varrho\nu}$$

which leads to

$$[J^{\mu\nu}, J^{\varrho\sigma}] = i\,(g^{\nu\varrho} J^{\mu\sigma} + g^{\mu\sigma} J^{\nu\varrho} + g^{\mu\varrho} J^{\sigma\nu} + g^{\sigma\nu} J^{\varrho\mu})$$

We thus obtain all the relations characterizing Poincaré algebra:

$$[P^{\mu}, P^{\nu}] = 0 \tag{1.173}$$

$$[J^{\mu\nu}, P^{\varrho}] = -i\,(g^{\mu\varrho} P^{\nu} - g^{\nu\varrho} P^{\mu}) \tag{1.174}$$

$$[J^{\mu\nu}, J^{\varrho\sigma}] = i\,(g^{\nu\varrho} J^{\mu\sigma} + g^{\mu\sigma} J^{\nu\varrho} + g^{\mu\varrho} J^{\sigma\nu} + g^{\sigma\nu} J^{\varrho\mu}) \tag{1.175}$$

A very simple way to satisfy these relations is to consider the standard realization of quantum mechanics:

$$P^{\mu} = i\,\partial^{\mu} \tag{1.176}$$

$$J^{\mu\nu} = i\,(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu}) \tag{1.177}$$

1.8.2 Casimir operators of the Poincaré algebra

The next step is to look for Casimir operators. Yes... "operatorS" because there are two of them.

The first one is pretty obvious:

$$C1 = P^{\mu} P_{\mu} \tag{1.178}$$

It commutes, obviously, with translations. With regard to the generators of the Lorentz algebra, we have

$$[J^{\mu\nu}, P^{\varrho} P_{\varrho}] = i\,(g^{\nu\varrho} P^{\mu} P_{\varrho} - g^{\mu\varrho} P^{\nu} P_{\varrho}) + i\,(g^{\nu}_{\varrho} P^{\varrho} P^{\mu} - g^{\mu}_{\varrho} P^{\varrho} P^{\nu}) \tag{1.179}$$

And the metric tensor does its job:

$$[J^{\mu\nu}, P^{\varrho} P_{\varrho}] = i\,(P^{\mu} P^{\nu} - P^{\nu} P^{\mu}) + i\,(P^{\nu} P^{\mu} - P^{\mu} P^{\nu}) = 0$$

The second Casimir operator is much less obvious. It involves the Pauli-Lubanski pseudo-vector, defined by

$$W_\mu \equiv \frac{1}{2}\,\varepsilon_{\mu\nu\varrho\sigma}\,J^{\nu\varrho}\,P^\sigma \tag{1.180}$$

In detail, this gives

$$W_0 = J^{12}P^3 + J^{31}P^2 + J^{23}P^1 \tag{1.181}$$

$$W_1 = -J^{02}P^3 + J^{03}P^2 - J^{23}P^0 \tag{1.182}$$

$$W_2 = J^{01}P^3 - J^{03}P^1 - J^{31}P^0 \tag{1.183}$$

$$W_3 = -J^{01}P^2 + J^{02}P^1 - J^{12}P^0 \tag{1.184}$$

We verify that

$$[P^\mu, W^\nu] = \frac{1}{2}\,\varepsilon^{\nu}_{\varrho\sigma\alpha}\,[P^\mu, J^{\varrho\sigma}]\,P^\alpha = -\frac{i}{2}\,\varepsilon^{\nu}_{\varrho\sigma\alpha}\,(g^{\sigma\mu}P^\varrho - g^{\varrho\mu}P^\sigma)\,P^\alpha = 0 \tag{1.185}$$

since we have both symmetry and antisymmetry on the same indices.

We have also successively

$$[J^{\mu\nu}, W^\varrho] = \frac{1}{2}\,\varepsilon^{\varrho}_{\alpha\beta\sigma}\,([J^{\mu\nu}, J^{\alpha\beta}]P^\sigma + J^{\alpha\beta}[J^{\mu\nu}, P^\sigma])$$

$$= \frac{i}{2}\,\varepsilon^{\varrho}_{\alpha\beta\sigma}\,(g^{\nu\alpha}J^{\mu\beta}P^\sigma + g^{\mu\beta}J^{\nu\alpha}P^\sigma + g^{\mu\alpha}J^{\beta\nu}P^\sigma + g^{\beta\nu}J^{\alpha\mu}P^\sigma + g^{\nu\sigma}J^{\alpha\beta}P^\mu - g^{\mu\sigma}J^{\alpha\beta}P^\nu)$$

$$= i\,\varepsilon^{\varrho}_{\alpha\beta\mu}\,J^{\alpha\nu}P^\beta - i\,\varepsilon^{\varrho}_{\alpha\beta\nu}\,J^{\alpha\mu}P^\beta + \frac{i}{2}\,\varepsilon^{\varrho}_{\alpha\beta\nu}\,J^{\alpha\beta}P^\mu - \frac{i}{2}\,\varepsilon^{\varrho}_{\alpha\beta\mu}\,J^{\alpha\beta}P^\nu$$

Here, it is worth lingering a little...

If the indices μ, ν, ρ are all different, the antisymmetry of ε and J implies that α can take only one value in the first two terms of the second member. This also results in a unique value for β. On the other hand, α, β can take two values in the last two terms... but the antisymmetry leads to the elimination of the factor $1/2$.

An example will probably make this clearer:

$$[J^{01}, W^2] = i\,\varepsilon^2_{310}J^{31}P^1 - i\,\varepsilon^2_{301}J^{30}P^0 + i\,\varepsilon^2_{031}J^{03}P^0 - i\,\varepsilon^2_{310}J^{31}P^1 = 0$$

We can therefore conclude that if the three initial indices are different, the result can only be zero.

If $\varrho = \nu$, we will have

$$[J^{\mu\nu}, W^{\nu}] = i\,\varepsilon^{\nu}_{\alpha\beta\mu}\,J^{\alpha\nu}P^{\beta} - \frac{i}{2}\,\varepsilon^{\nu}_{\alpha\beta\mu}\,J^{\alpha\beta}P^{\nu} \tag{1.186}$$

(without summation on ν!).

For example,

$$[J^{01}, W^{1}] = i\,\varepsilon^{1}_{\alpha\beta 0}\,J^{\alpha 1}P^{\beta} - \frac{i}{2}\,\varepsilon^{1}_{\alpha\beta 0}\,J^{\alpha\beta}P^{1}$$

$$= i\,\varepsilon^{1}_{230}\,(J^{21}P^{3} - J^{31}P^{2} - J^{23}P^{1}) = -iW^{0}$$

If $\varrho = \mu$, we will have

$$[J^{\mu\nu}, W^{\mu}] = -i\,\varepsilon^{\mu}_{\alpha\beta\nu}\,J^{\alpha\mu}P^{\beta} + \frac{i}{2}\,\varepsilon^{\mu}_{\alpha\beta\nu}\,J^{\alpha\beta}P^{\mu} \tag{1.187}$$

(without summation on μ!).

For example,

$$[J^{01}, W^{0}] = -i\,\varepsilon^{0}_{\alpha\beta 1}\,J^{\alpha 0}P^{\beta} + \frac{i}{2}\,\varepsilon^{0}_{\alpha\beta 1}\,J^{\alpha\beta}P^{0}$$

$$= i\,\varepsilon^{0}_{231}\,(J^{02}P^{3} - J^{03}P^{2} + J^{23}P^{0}) = -iW^{1}$$

In general, we will obtain

$$[J^{\mu\nu}, W^{\varrho}] = i\,(g^{\varrho\nu}W^{\mu} - g^{\varrho\mu}W^{\nu}) \tag{1.188}$$

The second Casimir vector is then constructed from this Pauli-Lubanski pseudo-vector:

$$C2 \equiv W^{\mu}W_{\mu} \tag{1.189}$$

Indeed, we have

$$[J^{\mu\nu}, W^{\varrho}W_{\varrho}] = i\,(g^{\varrho\nu}W^{\mu}W_{\varrho} - g^{\varrho\mu}W^{\nu}W_{\varrho} + g^{\nu}_{\varrho}W^{\varrho}W^{\mu} - g^{\mu}_{\varrho}W^{\varrho}W^{\nu}) \tag{1.190}$$

$$\Leftrightarrow [J^{\mu\nu}, W^{\varrho}W_{\varrho}] = i\,(W^{\mu}W^{\nu} - W^{\nu}W^{\mu} + W^{\nu}W^{\mu} - W^{\mu}W^{\nu}) = 0 \tag{1.191}$$

But we still need to give a physical interpretation of these two Casimir operators.

1.8.3 Wigner little groups

Let us recall that Casimir's operator for $so(3)$ allowed us to define either the azimuthal quantum number l or the spin s.

What are the quantum numbers involved here?

Before answering this question, we bring in the kets $|p, \alpha>$.

While the designation α will remain mysterious for a moment, the "p" refers to the 4-vector formed by the eigenvalues p^μ of translation operators

$$P^\mu |p, \alpha >= p^\mu |p, \alpha > \tag{1.192}$$

These are the same eigenvectors because the operators associated with translations commute. They are therefore simultaneously diagonalizable.

We will thus have

$$U(a) |p, \alpha >= \exp(-ia_\mu P^\mu) |p, \alpha >= \exp(-iap) |p, \alpha > \tag{1.193}$$

But what we want to highlight is not the action of $U(a)$ alone, but that of $U(\Lambda, a)$.

First, we have

$$U(\Lambda, a) |p, \alpha >= U(a)U(\Lambda) |p, \alpha > \tag{1.194}$$

By making use of (1.164) and replacing in it Λ by Λ^{-1}, we obtain

$$U^{-1}(\Lambda)U(a)U(\Lambda) = U(\Lambda^{-1}a) \tag{1.195}$$

Thus, we get

$$U(\Lambda, a) |p, \alpha >= U(\Lambda)U(\Lambda^{-1}a) |p, \alpha >= \exp(-i\Lambda^{-1}ap)U(\Lambda) |p, \alpha > \tag{1.196}$$

For the last equality, we used the result (1.193).

Second, we proceed as follows

$$\Lambda^{-1}ap = (\Lambda^{-1}a)_\nu p^\nu = (\Lambda^{-1})^\mu_\nu a_\mu p^\nu = \Lambda^\mu_\nu a_\mu p^\nu = a_\mu(\Lambda p)^\mu = a\,\Lambda p \equiv a\,p\prime \tag{1.197}$$

As a result, we are led to

$$U(\Lambda, a) |p, \alpha >= \exp(-i\Lambda^{-1}ap)U(\Lambda) |p, \alpha >= \exp(-iap\prime)U(\Lambda) |p, \alpha > \tag{1.198}$$

This is where a crucial result, demonstrated by Wigner[5], comes into play.

He asserted that a particular case k of $\{P^\mu\}$ from one of the following three classes provides complete knowledge of the representations. Here are the classes:

[5]Wigner, E. (1939). On unitary representations of the inhomogeneous Lorentz group. *Ann. Math.* 40,149-204. http://dx.doi.org/10.2307/1968551

1. Timelike 4-vectors (i.e., positive pseudo-norm): $k = (\pm m, 0, 0, 0)^T$
2. Lightlike 4-vectors (i.e., pseudo-norm equal to 0): $k = (\pm\kappa, 0, 0, \kappa)^T$
3. Spacelike 4-vectors (i.e., negative pseudo-norm): $k = (0, 0, 0, \xi)^T$

Before continuing, a word about these names. Why these classes?

They are in fact related to the first Casimir operator of the Poincaré group. For timelike 4-vectors, this operator will be such that

$$P^\mu P_\mu = m^2 \tag{1.199}$$

In other words, with $P^0 = E$ while the other components of P^μ are assimilated to those of the momentum $\overrightarrow{p}$, it is Einstein's relation

$$E^2 = \overrightarrow{p}^2 + m^2 \tag{1.200}$$

that is implied. This relation will form the basis of any relativistic quantum theory. We will have the opportunity to discuss this in greater detail later on.

In one spatial dimension, it is possible to represent this relation graphically. Simply fix the mass and consider the unique component of the momentum, denoted p, as the x-axis and the corresponding energy E as the y-axis.

For example, for $m^2 = 16$, this gives rise to the following equilateral hyperbola

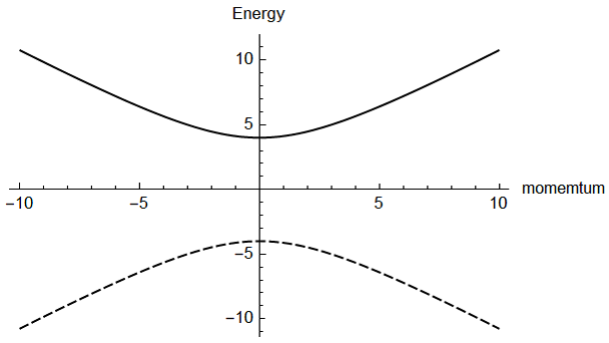


Figure 1.4: Visualization of Einstein's relation in one spatial dimension

It should be noted that the positive energy branch reaches its minimum at $p = 0$, $m = 4$, while the negative energy branch reaches its maximum at $p = 0$, $m = -4$.

Only the positive energy branch is currently considered by the vast majority of physicists, as the experimentally confirmed particles relate to it.

For lightlike 4-vectors, the Casimir operator will be such that

$$P^\mu P_\mu = 0 \tag{1.201}$$

It is the light cone, specific to special relativity, that underlies this case.

For spacelike 4-vectors, the Casimir operator will be such that

$$P^\mu P_\mu = -\xi^2 \tag{1.202}$$

Einstein's relation is then modified and written as

$$E^2 = \overrightarrow{p}^2 - \xi^2 \tag{1.203}$$

Here is the completed graphic:

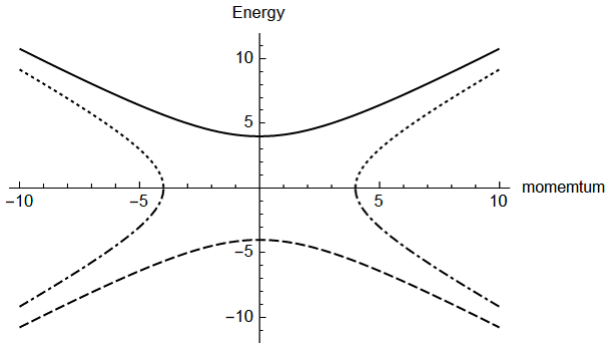


Figure 1.5: Visualization of Einstein's relation for timelike and spacelike 4-vectors

With that clarification made, let us return to the IURs of the Poincaré group.

Any 4-vector p other than the chosen representative k will be related to the latter via a specific Lorentz transformation that we denote by L_p:

$$L_p\, k = p \tag{1.204}$$

This relation will also be written later on as

$$k = L_p^{-1}\, p \;;\; \forall p \tag{1.205}$$

The two 4-vectors p and k belong to the same class since it is the job of Lorentz transformations to preserve the pseudo-norm.

A Wigner rotation is then defined by

$$R_W \equiv L_{\Lambda p}^{-1}\, \Lambda\, L_p \tag{1.206}$$

What does this operator do? Here it is:

$$R_W\, k = L^{-1}_{\Lambda p}\, \Lambda\, L_p\, k = L^{-1}_{\Lambda p}\, \Lambda p = k \tag{1.207}$$

Wigner rotations therefore have the characteristic of stabilizing the chosen representative of each of the three classes. All of these rotations together form what is called a Wigner little group.

Group because there is also a law associated with it:

$$R_W(\Lambda, p) R_W(\Lambda\prime, p\prime) = L^{-1}_{\Lambda p}\, \Lambda\, L_p\, L^{-1}_{\Lambda\prime p\prime}\, \Lambda\prime\, L_{p\prime} = R_W(\Lambda\prime\prime, p\prime\prime) = L^{-1}_{\Lambda\prime\prime p\prime\prime}\, \Lambda\prime\prime\, L_{p\prime\prime} \tag{1.208}$$

with

$$p\prime\prime = p\prime\ ;\ \Lambda\prime\prime = \Lambda\, L_p\, L^{-1}_{\Lambda\prime p\prime}\, \Lambda\prime \tag{1.209}$$

Indeed, we have

$$\Lambda\prime\prime p\prime\prime = \Lambda\, L_p\, L^{-1}_{\Lambda\prime p\prime}\, \Lambda\prime p\prime = \Lambda\, L_p\, k = \Lambda p \tag{1.210}$$

We now use these Wigner rotations to complete the specification of the states $|p, \alpha>$.

We obtain successively

$$U(\Lambda)\, |p, \alpha> = U(\Lambda) U(L_p)\, |k, \alpha>$$

$$\Leftrightarrow U(\Lambda)\, |p, \alpha> = U(L_{\Lambda p}) U(L^{-1}_{\Lambda p}) U(\Lambda) U(L_p)\, |k, \alpha>$$

$$\Leftrightarrow U(\Lambda)\, |p, \alpha> = U(L_{\Lambda p}) U(R_W)\, |k, \alpha>$$

If we know (and we will know!) the IURs of Wigner's little groups, this last result leads to

$$U(\Lambda)\, |p, \alpha> = U(L_{\Lambda p}) D_{\alpha\beta}(R_W)\, |k, \beta> = D_{\alpha\beta}(R_W)\, |\Lambda p, \beta> \tag{1.211}$$

In total, we will therefore have the complete action of the Lorentz group on the basis states:

$$U(\Lambda, a)\, |p, \alpha> = \exp(-ia\Lambda p)\, D_{\alpha\beta}(R_W)\, |\Lambda p, \beta> \tag{1.212}$$

The moral of the story: it is the IURs of little groups that enable us to determine the IURs of the Poincaré group!

This will therefore be our ultimate task: to highlight these little Wigner groups.

1.8.3.1 Wigner little group for timelike 4-vectors

Which operators among (1.86)-(1.89) preserve the chosen representative $k = (\pm m, 0, 0, 0)^T$?

It is clear that R_1, R_2, R_3 do:

$$R_1 k = R_2 k = R_3 k = k \tag{1.213}$$

The Wigner little group of this class is therefore $SO(3)$.

We also have

$$C1 = m^2 \, I \tag{1.214}$$

and

$$W_0 = 0, \, W_1 = \mp m \, J^{23}, \, W_2 = \mp m \, J^{31}, \, W_3 = \mp m \, J^{12} \tag{1.215}$$

As a consequence, we obtain

$$C2 = -m^2 \, ((J^{23})^2 + (J^{31})^2 + (J^{12})^2) = -m^2 \, s(s+1) \, I \tag{1.216}$$

The physical quantities involved in the unitary irreducible representations of Poincaré's algebra (or group) for the timelike case are therefore (real) mass and spin.

Let us emphasize: from a mathematical point of view, mass can be either positive or negative!

Lie algebras have proven themselves to be particularly efficient: everything they predicted has always been correct from an experimental point of view, even if it took decades to confirm! From there, it is only a step to conjecture the experimental evidence of negative mass particles, which is precisely what more and more models do.

Of course, in standard physics literature, the only interest is limited to the case $m > 0$.

1.8.3.2 Wigner little group for lightlike 4-vectors

The rotation R_3 clearly ensures the invariance of $k = (\pm\kappa, 0, 0, \kappa)^T$.

The two other operators that allow the same result are a little more difficult to highlight. Indeed, one must consider

$$[J^{0j}, P^\mu] = i \, g^{j\mu} \, P^0 \; ; \; [J^{3j}, P^\mu] = i \, g^{j\mu} \, P^3 \tag{1.217}$$

The fact that $P^0 = P^3$ thus implies considering $J^{01} - J^{31}$, $J^{02} + J^{23}$ as well as J^{12} which commute with P^0 and P^3.

At the level of Casimir's first operator, we have

$$C1 = 0 \tag{1.218}$$

Therefore, particles with zero mass are involved.

This is a result that we also find with the Pauli-Lubanski pseudo-vector since we have

$$W_0 = -W_3 = \pm\kappa\, J^{12}\ ;\ W_1 = \mp\kappa\,(J^{02} + J^{23})\ ;\ W_2 = \pm\kappa\,(J^{01} - J^{31}) \quad (1.219)$$

We then check that

$$[W_0, W_1] = \pm i\,\kappa\, W_2\ ;\ [W_2, W_0] = \pm i\,\kappa\, W_1\ ;\ [W_1, W_2] = 0 \quad (1.220)$$

This concerns the algebra related to the group formed by two translations and one rotation, namely the two-dimensional Euclidean group. This group is non-compact, a fact due to the presence of the translations. The direct consequence of this is that there is no irreducible unitary representation of finite dimension.

This can be seen directly through the explicit realization of the infinitesimal form of Wigner's rotation:

$$W_R(a, b, \theta) = I + a\,(J^{01} - J^{31}) + b\,(J^{02} + J^{23}) + \theta\, J^{12} \quad (1.221)$$

$$\Leftrightarrow W_R(a, b, \theta) = \begin{pmatrix} 1 & a & b & 0 \\ a & 1 & -\theta & -a \\ b & \theta & 1 & -b \\ 0 & a & b & 1 \end{pmatrix} \quad (1.222)$$

The states of the representation are first labeled by k and the parameters a, b. We have, of course,

$$(J^{01} - J^{31})\,|k, a, b> = a\,|k, a, b>\ ;\ (J^{02} + J^{23})\,|k, a, b> = b\,|k, a, b> \quad (1.223)$$

If we define

$$|k, a, b, \theta> \equiv \exp(-i\,\theta\, J^{12})\,|k, a, b> \quad (1.224)$$

we obtain

$$(J^{01} - J^{31})\,|k, a, b, \theta> = \exp(-i\,\theta\, J^{12})\exp(i\,\theta\, J^{12})(J^{01} - J^{31})\exp(-i\,\theta\, J^{12})|k, a, b> \quad (1.225)$$

We also have

$$\exp(i\,\theta\, J^{12})(J^{01} - J^{31})\exp(-i\,\theta\, J^{12})$$

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

implying

$$\exp(i\,\theta\,J^{12})(J^{01}-J^{31})\exp(-i\,\theta\,J^{12})=\cos\theta\,(J^{01}-J^{31})+\sin\theta\,(J^{02}+J^{23}) \tag{1.226}$$

This leads to

$$(J^{01}-J^{31})\,|k,a,b,\theta>=(a\cos\theta+b\sin\theta)\,|k,a,b,\theta> \tag{1.227}$$

and, consequently, an infinite number of states.

The only way to circumvent this is to not consider translations by setting a and b to zero, and, consequently, to limit ourselves to the rotation associated with J^{12}.

What is the quantum number associated with this operator?

In this case, it is helicity, defined by

$$\frac{\overrightarrow{J}.\overrightarrow{p}}{||\overrightarrow{p}||} \tag{1.228}$$

The only non-zero component of $\overrightarrow{p}$ is the one along the z-axis, which will select J^{12} only.

The physical quantities that appear in the unitary irreducible representations of Poincaré algebra (or group) for the lightlike case are therefore (zero) mass and helicity.

One possible realization of the set of generators of the algebra associated with the little group in question is

$$J^{01}-J^{31}=-ix\frac{\partial}{\partial z}\;;\;J^{02}+J^{23}=-iy\frac{\partial}{\partial z}\;;\;J^{12}=-ix\frac{\partial}{\partial y}+iy\frac{\partial}{\partial x} \tag{1.229}$$

1.8.3.3 Wigner little group for spacelike 4-vectors

The 4-vector $k=(0,0,0,\xi)^T$ is left invariant unde the actions of R_3, B_1 and B_2.

The first Casimir operator is such that

$$C1=-\xi^2 \tag{1.230}$$

Thus, it is particles of imaginary mass that underlie this choice. Of course, none of them have been observed. But, again, Poincaré algebra does not rule them out.

From the Pauli-Lubanski pseudo-vector, there remains

$$W_0=\xi J^{12}\,,\,W_1=-\xi J^{02}\,,\,W_2=\xi J^{01} \tag{1.231}$$

which generate

$$[W_0, W_1] = i\xi W_2 \; ; \; [W_2, W_0] = i\xi W_1 \; ; \; [W_1, W_2] = -i\xi W_0 \tag{1.232}$$

With one difference between a sign and the algebra $so(3)$, we recognize (the complexification of) the algebra $so(2,1)$. The corresponding Wigner little group is therefore $SO(2,1)$.

This will surprise no one: the IURs of $so(2,1)$ are very close to those of $so(3)$. Thus, it is possible to define

$$K_+ \equiv J^{01} + J^{02} \; ; \; K_- \equiv -J^{01} + J^{02} \; ; \; K_z \equiv iJ^{12} \tag{1.233}$$

and to convince oneself that

$$[K_z, K_\pm] = \pm K_\pm \; ; \; [K_+, K_-] = -2K_z \tag{1.234}$$

So, the same relations as those of $so(3)$ arise, except for the change of sign in the last commutation relation.

The Casimir operator also undergoes a sign change since it is equal to

$$C = -\frac{1}{2}K_+K_- - \frac{1}{2}K_-K_+ + K_z^2 \tag{1.235}$$

There are four types of representations associated with the algebra $so(2,1)$.

We only present results similar to those of $so(3)$.

A representation uses the same states $|j, m>$ as for $so(3)$ but with the important difference that the ladder operators will no longer act on m but on j... and since there is no limitation on this quantum number, this leads to an infinite number of states... the fault of the boosts, of course!

Nothing beats a simple graphic to understand the situation:

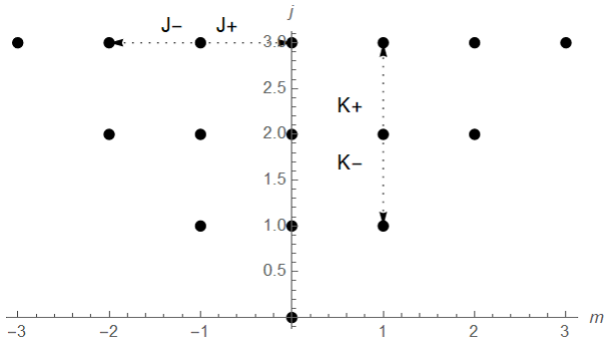


Figure 1.6: Action of the ladder operators of $so(2,1)$

The ladder operators $J_\pm$ of $so(3)$ connect the states $|j, m>$ to the states $|j, m \pm 1>$ for a fixed j; they work on a horizontal line, of length $2j + 1$.

The ladder operators $K_\pm$ of $so(2,1)$ connect the states $|j,m>$ to the states $|j\pm 1,m>$ for a fixed m; they work on a vertical line, of unlimited length.

Specifically we have

$$K_z\,|j,m>=(j+\frac{1}{2})\,|j,m> \tag{1.236}$$

$$K_+\,|j,m>=\sqrt{(j-m+1)(j+m+1)}\,|j+1,m> \tag{1.237}$$

$$K_-\,|j,m>=\sqrt{(j-m)(j+m)}\,|j-1,m> \tag{1.238}$$

$$C\,|j,m>=(m^2-\frac{1}{4})\,|j,m> \tag{1.239}$$

There's no point in dreaming about finite-dimensional matrices here. We must therefore move towards states of the spherical harmonic type, as we did for $so(3)$. But are spherical harmonics really necessary here? Well, no! Indeed, the generators of $so(2,1)$ only affect the number j, which limits us to the associated Legendre polynomials. An immediate consequence is that we must restrict the values of j to integers, which is why we will write it as l in the following.

The ladder operators of $so(2,1)$ are realized via

$$K_+=\sin\theta\,\frac{d}{d\theta}+(l+1)\,\cos\theta \tag{1.240}$$

$$K_-=-\sin\theta\,\frac{d}{d\theta}+l\,\cos\theta \tag{1.241}$$

States are then written as

$$|l,m>=\sqrt{\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta) \tag{1.242}$$

In detail, and for the first states, we have:

$$|0,0>=1 \tag{1.243}$$

$$|1,1>=-\frac{1}{\sqrt{2}}\sin\theta\;;\;|1,0>=\cos\theta\;;\;|1,-1>=\frac{1}{\sqrt{2}}\sin\theta \tag{1.244}$$

$$|2,2>=|2,-2>=\sqrt{\frac{3}{8}}\sin^2\theta \tag{1.245}$$

$$|2,1>=-|2,-1>=-\sqrt{\frac{3}{2}}\sin\theta\cos\theta\;;\;|2,0>=1-\frac{3}{2}\sin^2\theta \tag{1.246}$$

1.9 The $su(3, C)$ algebra and the eightfold way

Other algebras than the ones already mentioned appear in quantum theory, notably the $su(n, C)$ algebras.

At first glance, these algebras no longer have anything to do with coordinate transformations such as rotations, but everything to do with what are called internal symmetries, i.e., those that affect the quantum numbers specific to elementary particles.

A priori... because we will see that these algebras also concern transformations involving complex coordinates.

But we are not there yet, and we first propose to get to know these unitary algebras a little better.

By definition, $su(n, C)$ algebras are generated by n-dimensional Hermitian (referring to the letter u) matrices with complex elements and having zero trace (referring to the letter s).

1.9.1 The $su(3, C)$ algebra

The algebra we are interested in[6] is $su(3, C)$.

It is generated by eight basic matrices, known as Gell-Mann matrices:

$$F_1 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; F_2 = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; F_3 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.247}$$

$$F_4 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \; F_5 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, \; F_6 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{1.248}$$

$$F_7 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \; F_8 = \frac{1}{2\sqrt{3}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{1.249}$$

These matrices are indeed zero-trace, Hermitian, and of dimension 3.

Since there are eight of them, it is more tedious to calculate the commutation relations. But if you are brave enough or have calculation software at hand, you can confirm the following results:

$$[F_A, F_B] = i\, f_{ABC}\, F_C, \; A, B, C = 1, 2, \ldots, 8 \tag{1.250}$$

The structure constants, which are completely antisymmetric, are given in this table:

[6] All details concerning this algebra can be found, for example, in Greiner, W. and Müller, B. (1989). *Quantum Mechanics-Symmetries.* Springer-Verlag

ABC	123	147	156	246	257	345	367	458	678
f_{ABC}	1	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{2}$

We note the presence of two diagonal generators. We know that they play an important role, along with the Casimir operators. Yes...the Casimir operators, because here too, there are two of them!

The first one is given by

$$C1 \equiv \sum_A F_A^2 \tag{1.251}$$

In the specific case of Gell-Mann matrices, it is equal to

$$C1 = \frac{4}{3} I \tag{1.252}$$

The second one is cubic:

$$C2 \equiv \sum_{A,B,C} d_{ABC} F_A F_B F_C \tag{1.253}$$

The coefficients d_{ABC}, which are completely symmetric, are determined via anticommutation relations:

$$\{F_A, F_B\} \equiv F_A F_B + F_B F_A = \frac{1}{3}\delta_{AB} + d_{ABC} F_C \tag{1.254}$$

Once again, with a little elbow grease or computer assistance, it is not difficult to convince oneself that

ABC	118	146	157	228	247	256	338	344
d_{ABC}	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{\sqrt{3}}$	$\frac{1}{2}$

ABC	355	366	377	448	558	668	778	888
d_{ABC}	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{2\sqrt{3}}$	$-\frac{1}{\sqrt{3}}$

The second Casimir operator can thus be evaluated and, in the case of Gell-Mann matrices, gives

$$C2 = \frac{10}{9} I \tag{1.255}$$

You have seen the power of ladder operators in the case of $so(3)$ or $so(1,3)\ldots$ we will use them here as well.

To do this, we define

$$T_\pm \equiv F_1 \pm iF_2,\ V_\pm \equiv F_4 \pm iF_5,\ U_\pm \equiv F_6 \pm iF_7 \tag{1.256}$$

We then have

$$[T_z, T_\pm] = \pm T_\pm,\ [T_+, T_-] = 2T_z\ ;\ T_z \equiv F_3 \tag{1.257}$$

$$[V_z, V_\pm] = \pm V_\pm,\ [V_+, V_-] = 2V_z\ ;\ V_z \equiv \frac{1}{2}F_3 + \frac{\sqrt{3}}{2}F_8 \tag{1.258}$$

$$[U_z, U_\pm] = \pm U_\pm,\ [U_+, U_-] = 2U_z\ ;\ U_z \equiv -\frac{1}{2}F_3 + \frac{\sqrt{3}}{2}F_8 \tag{1.259}$$

In order to avoid the awkwardness that you will no doubt have noticed, namely the recurring appearance of $\sqrt{3}$, we define a new operator

$$Y \equiv \frac{2}{\sqrt{3}}F_8 \tag{1.260}$$

The states of the representations of $su(3,C)$ will therefore be labeled by t_z and y, the eigenvalues of the two diagonal operators in question:

$$T_z\,|t_z, y> = t_z\,|t_z, y> \tag{1.261}$$

$$Y\,|t_z, y> = y\,|t_z, y> \tag{1.262}$$

How do ladder operators modify these values? Simply examine the corresponding commutation relations:

$$[T_z, T_\pm] = \pm T_\pm\ ;\ [Y, T_\pm] = 0$$

The operators $T_\pm$ therefore have the following actions:

$$t_z \to t_z \pm 1\ ;\ y \to y$$

We also have

$$[T_z, V_\pm] = \pm\frac{1}{2}V_\pm\ ;\ [Y, V_\pm] = \pm V_\pm$$

implying that $V_\pm$ are such that

$$t_z \to t_z \pm \frac{1}{2}\ ;\ y \to y \pm 1$$

Finally, we notice

$$[T_z, U_\pm] = \mp\frac{1}{2}U_\pm\ ;\ [Y, U_\pm] = \pm U_\pm$$

leading to the following actions in what concerns $U_{\pm}$:

$$t_z \to t_z \mp \frac{1}{2} \; ; \; y \to y \pm 1$$

As with $so(3)$, a diagram will undoubtedly be useful for visualizing the actions of each of these ladder operators:

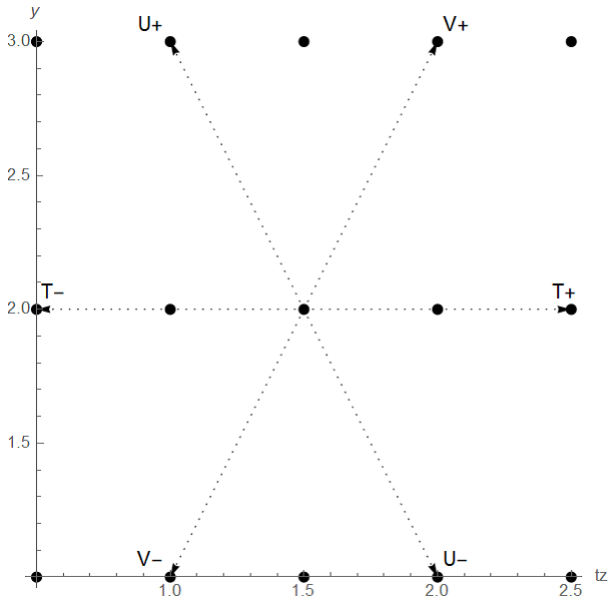


Figure 1.7: Action of ladder operators of $su(3, C)$

This diagram will also help us understand what follows.

There is a maximum value for t_z. And, inevitably, the corresponding state, which we denote $|t_{zmax}, y>$, must be annihilated by T_+, V_+, and U_-, whose precise purpose is to increase this value.

However, with other scale operators, significant results can be obtained.

This is how the positive integers p and q are defined by

$$(V_-)^{p+1} |t_{zmax}, y >= 0, \; (T_-)^{q+1} (V_-)^p |t_{zmax}, y >= 0 \tag{1.263}$$

The representations involved will then be noted

$$D(p, q) \tag{1.264}$$

All this may seem a little abstract... so let's see what it means with Gell-Mann matrices!

We have

$$T_z = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Y = \frac{1}{3}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{1.265}$$

$$T_+ = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_+ = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, U_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \tag{1.266}$$

$$T_- = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, U_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.267}$$

As a result,

$$|t_{zmax}, y >= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{1.268}$$

is canceled by T_+, V_+ and U_-.

This state is an eigenvector of T_z, with eigenvalue $\frac{1}{2}$, and an eigenvector of Y, with eigenvalue $\frac{1}{3}$.

We can therefore write

$$|t_{zmax}, y > \equiv |\frac{1}{2}, \frac{1}{3}> \tag{1.269}$$

The action of T_- is

$$T_-|\frac{1}{2}, \frac{1}{3}>= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv |-\frac{1}{2}, \frac{1}{3}> \tag{1.270}$$

That of V_- is given by

$$V_-|\frac{1}{2},\frac{1}{3}>=\begin{pmatrix}0&0&0\\0&0&0\\1&0&0\end{pmatrix}\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}0\\0\\1\end{pmatrix}\equiv|0,-\frac{2}{3}> \tag{1.271}$$

Finally we have

$$V_-^2|\frac{1}{2},\frac{1}{3}>=0,\ T_-V_-|\frac{1}{2},\frac{1}{3}>=0 \tag{1.272}$$

leading to

$$p=1,\ q=0 \tag{1.273}$$

We will say that the representation underlying Gell-Mann's matrices is $D(1,0)$. Here is a diagram of this representation of $su(3,C)$:

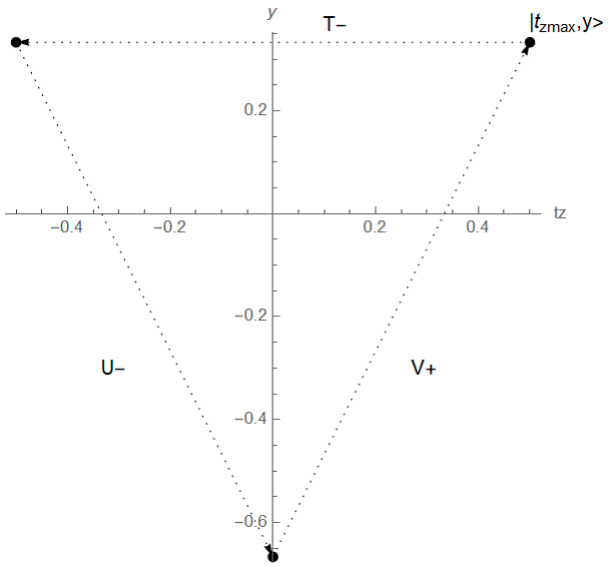


Figure 1.8: Representation $D(1,0)$ of $su(3,C)$

There is another 3-dimensional representation of $su(3,C)$. It is given by

$$F_A\rightarrow -F_A^* \tag{1.274}$$

In this case, the matrices become

$$T_z = \frac{1}{2}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Y = \frac{1}{3}\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \tag{1.275}$$

$$T_+ = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, U_- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{pmatrix} \tag{1.276}$$

$$T_- = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, V_- = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, U_+ = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix} \tag{1.277}$$

This time, that means that

$$|t_{zmax}, y >= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \equiv |\frac{1}{2}, -\frac{1}{3}> \tag{1.278}$$

$$T_-|\frac{1}{2}, -\frac{1}{3}>= \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \equiv |-\frac{1}{2}, -\frac{1}{3}> \tag{1.279}$$

$$U_+|\frac{1}{2}, -\frac{1}{3}>= \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} \equiv |0, \frac{2}{3}> \tag{1.280}$$

This representation is very similar to the previous one, but not identical! In fact, it is such that $p = 0$, $q = 1$, and therefore it is $D(0, 1)$.

The diagram is also different from the previous one; it is its symmetrical counterpart:

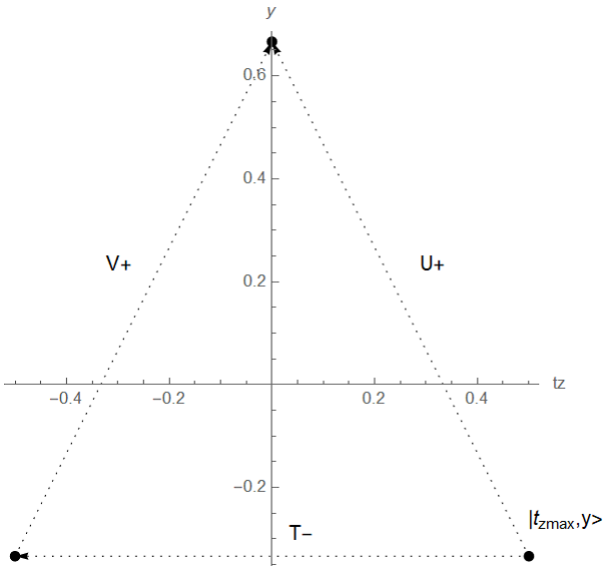


Figure 1.9: Representation $D(0,1)$ of $su(3,C)$

Let's consider one last representation: the adjoint representation.
It is characterized by two specific features:

- The dimension of its matrices is equal to the number of its matrices (i.e., 8, in this case)
- The matrices in question are given by

$$(F_A)_{BC} = -if_{ABC} \tag{1.281}$$

All you need to know are the constants of structure to highlight them.
Here are the simultaneously diagonalizable matrices:

$$T_z = \frac{i}{2}\begin{pmatrix} 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, Y = i\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.282}$$

At this stage, we can already find the eigenvalues of these two matrices and obtain

$$\left\{-1, 1, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, 0, 0\right\} \tag{1.283}$$

for T_z and

$$\{-1, -1, 1, 1, 0, 0, 0, 0\} \tag{1.284}$$

for Y.

We will therefore have to consider the states

$$|1, 0>, |0, 0>, |0, 0>, |-1, 0> \tag{1.285}$$

$$|\frac{1}{2}, 1>, |-\frac{1}{2}, 1>, |-\frac{1}{2}, -1>, |\frac{1}{2}, -1> \tag{1.286}$$

The state $|1, 0>$ is, of course, the one that must be assimilated to $|t_{zmax}, y>$. It will be connected to the other states according to the diagram:

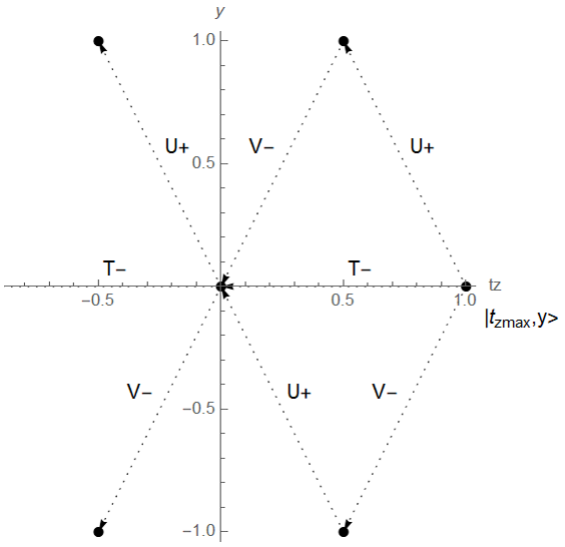


Figure 1.10: Representation $D(1,1)$ of $su(3,C)$

If we continue with the matrix implementation by focusing on ladder operators, we have

$$T_+ = \frac{1}{2}\begin{pmatrix} 0 & 0 & -2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2i & 0 & 0 & 0 & 0 & 0 \\ 2 & 2i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 1 & 0 \\ 0 & 0 & 0 & -1 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$T_- = \frac{1}{2}\begin{pmatrix} 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2i & 0 & 0 & 0 & 0 & 0 \\ -2 & 2i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & -1 & 0 \\ 0 & 0 & 0 & 1 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$V_+ = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & -1 & 0 \\ 0 & 0 & 0 & 1 & i & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -\sqrt{3} \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & -i\sqrt{3} \\ -1 & -i & 0 & 0 & 0 & 0 & 0 & 0 \\ -i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & i\sqrt{3} & 0 & 0 & 0 \end{pmatrix}$$

$$V_- = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & i & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 1 & 0 \\ 0 & 0 & 0 & -1 & i & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & -i\sqrt{3} \\ 1 & -i & 0 & 0 & 0 & 0 & 0 & 0 \\ -i & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sqrt{3} & i\sqrt{3} & 0 & 0 & 0 \end{pmatrix}$$

$$U_+ = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 1 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & -i & 0 \\ -1 & i & 0 & 0 & 0 & 0 & 0 & 0 \\ -i & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & -\sqrt{3} \\ 0 & 0 & i & 0 & 0 & 0 & 0 & -i\sqrt{3} \\ 0 & 0 & 0 & 0 & 0 & \sqrt{3} & i\sqrt{3} & 0 \end{pmatrix}$$

$$U_- = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -1 & i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -i & 0 \\ 1 & i & 0 & 0 & 0 & 0 & 0 & 0 \\ -i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & \sqrt{3} \\ 0 & 0 & i & 0 & 0 & 0 & 0 & -i\sqrt{3} \\ 0 & 0 & 0 & 0 & 0 & -\sqrt{3} & i\sqrt{3} & 0 \end{pmatrix}$$

This highlights the matrix realization of all the incriminated states.
Up to a normalization factor, they are given by

$$|1,0>= \begin{pmatrix} 1 \\ i \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |-1,0>= \begin{pmatrix} 1 \\ -i \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{1.287}$$

$$|\frac{1}{2},1>=\begin{pmatrix}0\\0\\0\\1\\i\\0\\0\\0\end{pmatrix},\ |\frac{1}{2},-1>=\begin{pmatrix}0\\0\\0\\0\\0\\1\\-i\\0\end{pmatrix},\ |-\frac{1}{2},-1>=\begin{pmatrix}0\\0\\0\\1\\-i\\0\\0\\0\end{pmatrix},\ |-\frac{1}{2},1>=\begin{pmatrix}0\\0\\0\\0\\0\\1\\i\\0\end{pmatrix} \tag{1.288}$$

There remains the thorny issue of the state $|0,0>$ or rather the two states $|0,0>$!

In fact, they can be obtained from $|1,0>$ in three different ways (take another look at the Figure (1.10)!):

$$|0,0>_I = T_-|1,0> \tag{1.289}$$

$$|0,0>_{II} = V_-U_+|1,0> \tag{1.290}$$

$$|0,0>_{III} = U_+V_-|1,0> \tag{1.291}$$

Fortunately, we have a linear dependency:

$$|0,0>_{III} = |0,0>_I + |0,0>_{II} \tag{1.292}$$

So there remains

$$|0,0>_I=\begin{pmatrix}0\\0\\1\\0\\0\\0\\0\\0\end{pmatrix},\ |0,0>_{II}=\begin{pmatrix}0\\0\\1\\0\\0\\0\\0\\\sqrt{3}\end{pmatrix} \tag{1.293}$$

We do indeed have eight states!

The adjoint representation of $su(3, C)$ is also denoted $D(1, 1)$ according to the conventions specified above.

In general, the representation $D(p, q)$ of $su(3, C)$ has dimension $\frac{1}{2}(p+1)(q+1)(p+q+2)$.

The Casimir operator $C1$ has eigenvalues given by

$$C1 = \frac{1}{3}(p^2 + q^2 + 3p + 3q + pq)\, I \tag{1.294}$$

Clearly, this confirms the value $\frac{4}{3}$ for representations $D(1, 0)$ and $D(0, 1)$ and the value 3 for representation $D(1, 1)$.

The eigenvalues of the second Casimir operator are

$$C2 = \frac{1}{18}(p - q)(3 + p + 2q)(3 + q + 2p)\, I \tag{1.295}$$

We indeed find the value $\pm\frac{10}{9}$ for the representations $D(1, 0)$ and $D(0, 1)$. However, the adjoint representation sees the cancellation of this Casimir operator.

1.9.2 Root system

To this algebra $su(3, C)$, as we could do for all Lie algebras, we will associate a root system. This is, in a way, the graphical representation of a Lie algebra. It is also its schematized DNA: as soon as we examine the drawing, we should be able to say to ourselves, "Good heavens! But of course, it's $su(3, C)$!" And this drawing is as follows: a regular hexagon:

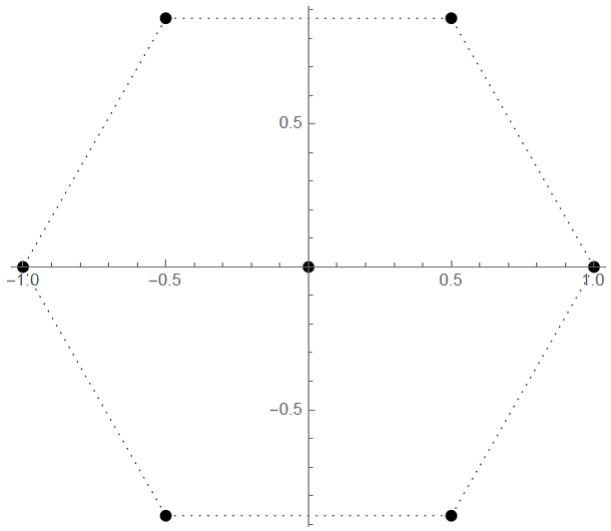


Figure 1.11: Root system of $su(3, C)$

How can we understand this? By using what is known as the Cartan-Weyl basis.

It highlights, for $su(3,C)$, two diagonal elements (H^1, H^2) and six elements $(E_{\pm\alpha^1}, E_{\pm\alpha^2}, E_{\pm(\alpha^1+\alpha^2)})$ that are not diagonal. The relations between these elements are

$$[H^1, H^2] = 0 \tag{1.296}$$

$$[H^1, E_{\pm\alpha^1}] = \pm E_{\pm\alpha^1},\ [H^2, E_{\pm\alpha^1}] = 0 \tag{1.297}$$

$$[H^1, E_{\pm\alpha^2}] = \mp\frac{1}{2}E_{\pm\alpha^2},\ [H^2, E_{\pm\alpha^2}] = \pm\frac{\sqrt{3}}{2}E_{\pm\alpha^2} \tag{1.298}$$

$$[H^1, E_{\pm(\alpha^1+\alpha^2)}] = \pm\frac{1}{2}E_{\pm(\alpha^1+\alpha^2)},\ [H^2, E_{\pm(\alpha^1+\alpha^2)}] = \pm\frac{\sqrt{3}}{2}E_{\pm(\alpha^1+\alpha^2)} \tag{1.299}$$

What you see appearing here are the so-called roots:

$$\alpha^1 = (1,0), -\alpha^1, \alpha^2 = (-\frac{1}{2}, \frac{\sqrt{3}}{2}), -\alpha^2, \alpha^1+\alpha^2 = (\frac{1}{2}, \frac{\sqrt{3}}{2}), -\alpha^1-\alpha^2 \tag{1.300}$$

They are determined by analyzing the second member of the commutation relations between the diagonals and the non-diagonals. The result involving H^1 will give the first component, while the relation with H^2 will provide the second component.

Here, they correspond to vectors in the plane starting from the origin and ending at the six vertices of the hexagon. Two zero roots must be added, positioned at the center of the hexagon and linked to the two diagonal elements (which obviously commute with themselves).

So $su(3,C)$ is this: six vertices placed on a regular hexagon and two points at the center of this polygon.

With regard to the Cartan-Weyl basis, we must also add the relations

$$[E_{\alpha^1}, E_{-\alpha^1}] = 2H^1, [E_{\alpha^2}, E_{-\alpha^2}] = -H^1+\sqrt{3}H^2, [E_{(\alpha^1+\alpha^2)}, E_{-(\alpha^1+\alpha^2)}] = H^1+\sqrt{3}H^2 \tag{1.301}$$

There, we take the linear combination of H^1, H^2 with, to within a factor of 2, each of the components of the roots.

Finally, we must also impose

$$[E_{\alpha^1}, E_{\alpha^2}] = E_{(\alpha^1+\alpha^2)},\ [E_{\alpha^1}, E_{-(\alpha^1+\alpha^2)}] = -E_{-\alpha^2} \tag{1.302}$$

$$[E_{-\alpha^1}, E_{-\alpha^2}] = -E_{-(\alpha^1+\alpha^2)},\ [E_{-\alpha^1}, E_{(\alpha^1+\alpha^2)}] = E_{\alpha^2} \tag{1.303}$$

$$[E_{\alpha^2}, E_{-(\alpha^1+\alpha^2)}] = E_{-\alpha^1},\ [E_{-\alpha^2}, E_{(\alpha^1+\alpha^2)}] = -E_{\alpha^1} \tag{1.304}$$

These relations indicate that the commutator of two non-diagonals related to different roots produces a non-diagonal related to the sum of these two roots.

All other commutation relations are zero.

All this is typical of the algebra and valid for all representations of $su(3, C)$.

The connection with Gell-Mann's realization is as follows:

$$H^1 = F_3,\ H^2 = F_8 \tag{1.305}$$

$$E_{\pm\alpha^1} = F_1 \pm iF_2,\ E_{\pm\alpha^2} = F_6 \pm iF_7,\ E_{\pm(\alpha^1+\alpha^2)} = F_4 \pm iF_5 \tag{1.306}$$

1.9.3 Weights diagram

But how can we distinguish between representations of the same algebra? By using a diagram similar to the one of the root system, which in the case of a representation is called a weights diagram. So, be careful:

Algebra	$\longleftrightarrow$	Root system
Representation	$\longleftrightarrow$	Weights diagram

The weights are obtained by finding the eigenvalues and eigenvectors of the diagonal elements H^1, H^2

$$H^j \left|\psi\right> = m^j \left|\psi\right> \tag{1.307}$$

The weights are then the vectors

$$\overrightarrow{m} = (m^1, m^2) \tag{1.308}$$

In Gell-Mann's formulation, it is particularly simple: the eigenvectors are

$$\left|\psi 1\right> = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ \left|\psi 2\right> = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ \left|\psi 3\right> = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{1.309}$$

They give rise to the respective weights

$$\overrightarrow{m_1} = \left(\frac{1}{2}, \frac{1}{2\sqrt{3}}\right),\ \overrightarrow{m_2} = \left(-\frac{1}{2}, \frac{1}{2\sqrt{3}}\right),\ \overrightarrow{m_3} = \left(0, -\frac{1}{\sqrt{3}}\right) \tag{1.310}$$

If we plot their endpoints on a plane, we find the three vertices of an equilateral triangle, as discussed earlier:

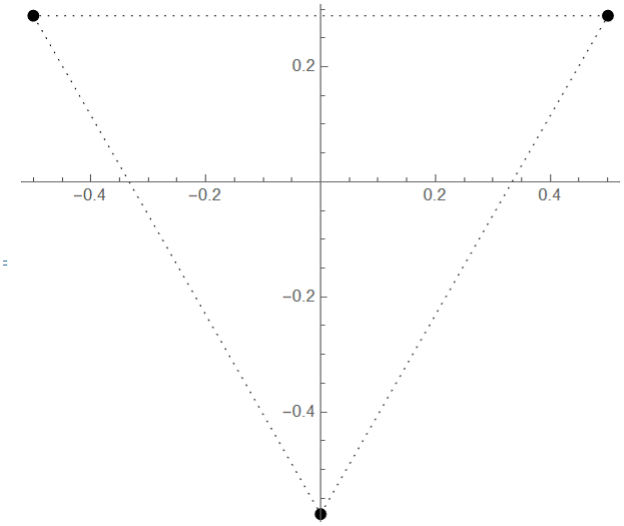


Figure 1.12: Weights diagram of the representation $D(1,0)$ of $su(3,C)$

This representation is known in physics literature as the representation **3**. We have also referred to it as $D(1,0)$.

1.9.4 Fundamental representations of $su(3,C)$

We know that there is another three-dimensional representation and that it is not equivalent to representation **3**. It is the representation $D(0,1)$.

Being given by Eq. (1.274), the consequences are

$$H^1 \to -H^1 \; ; \; H^2 \to -H^2 \tag{1.311}$$

$$E_{\pm\alpha^1} \to -E_{\mp\alpha^1} \; ; \; E_{\pm\alpha^2} \to -E_{\mp\alpha^2} \; ; \; E_{\pm(\alpha^1+\alpha^2)} \to -E_{\mp(\alpha^1+\alpha^2)} \tag{1.312}$$

And this does not change the commutation relations of the Cartan–Weyl basis!

In terms of weights, however, this representation is characterized by

$$\overrightarrow{m_1} = \left(-\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right), \overrightarrow{m_2} = \left(\frac{1}{2}, -\frac{1}{2\sqrt{3}}\right), \overrightarrow{m_3} = \left(0, \frac{1}{\sqrt{3}}\right) \tag{1.313}$$

It is clear that these weights are opposite to those in representation **3**; we thus confirm that we are dealing with another representation of $su(3,C)$, which is denoted **3*** or $D(0,1)$ to distinguish it from the first.

Graphically, this representation still corresponds to an equilateral triangle, but with the positions reversed:

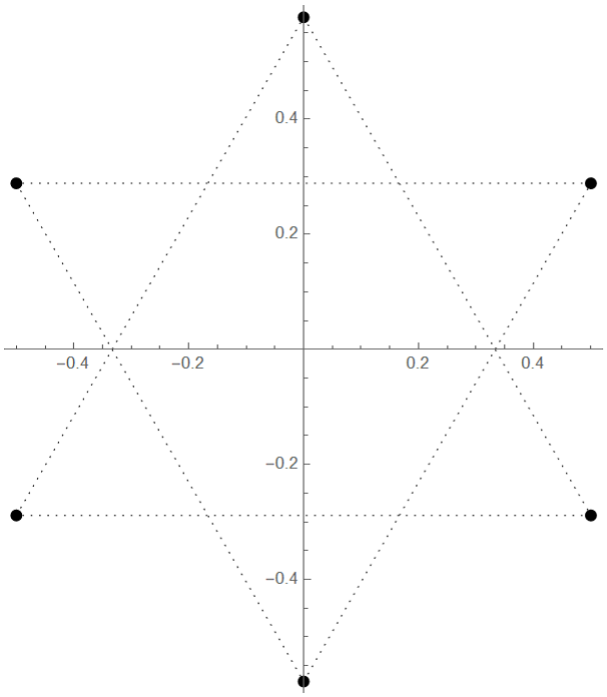


Figure 1.13: Weights diagram for the representations $D(1,0)$ and $D(0,1)$ of $su(3,C)$

The fact that the weights diagrams differ for the two representations clearly indicates that the two representations are different, even though they appear extremely similar at first glance.

1.9.5 What does this have to do with physics?

That's all very nice, but once again: what does this have to do with physics?

We're getting to that.

Since the discovery of spin, time has passed, bringing with it a host of discoveries regarding both particles and quantum numbers.

Thus, in 1947, the existence of particles called kaons was demonstrated. They differed from the other particles already cataloged to such an extent that it became necessary to introduce the quantum number S (for "strangeness"). Together with the baryon number B, which takes the value 1 for baryons (protons and neutrons, in particular) and 0 for mesons (π mesons, for example),

they form what is known as hypercharge:

$$Y \equiv B + S \tag{1.314}$$

$Y\ldots$ doesn't that ring a bell?

To this quantum number, we must also add the isospin I proposed by Heisenberg to distinguish the proton from the neutron, which were then viewed as two different states of the nucleon. Isospin is in fact related to the $so(3)$ algebra because the vector corresponding to it has three components satisfying the relations of this algebra. In particular, the third component I_3 is diagonal. The electric charge is then given by

$$Q = I_3 + \frac{1}{2}Y \tag{1.315}$$

Gell-Mann's idea emerged in the wake of all these discoveries, and on the surface, it is very simple:

What if all the points that appear in the $su(3, C)$ weights diagrams actually corresponded to elementary particles?

Let's consider the three quarks known at that time: u, d, s. Here are their quantum numbers:

	I_3	Y	B	S	Q
u	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$\frac{2}{3}$
s	$-\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{3}$	0	$-\frac{1}{3}$
d	0	$-\frac{2}{3}$	$\frac{1}{3}$	-1	$-\frac{1}{3}$

Now, recall the weights from representation **3** given by Eq. (1.310). It makes sense, doesn't it?

We have

$$\overrightarrow{m} = \left(I_3, \frac{\sqrt{3}}{2}Y\right) \tag{1.316}$$

The three quarks are thus placed on the vertices of the weights diagram for representation **3** of $su(3, C)$.

But we have another fundamental representation: **3***.

It is characterized by the fact that all the weights change sign. Therefore, the isospin and hypercharge quantum numbers must also change sign. Which particles have this property? The antiquarks $\overline{u}$, $\overline{d}$, $\overline{s}$!

The process is set in motion...

1.9.6 The eightfold way

It is 1961. Gell-Mann pursues his idea. This idea will be confirmed, independently, by Ne'eman that same year.

Both know that mesons are composed of a quark and an antiquark. From the perspective of Lie algebras, this leads us to consider the representation $3 \otimes 3^*$.

Makes sense, right?

The tensor product $\otimes$ is defined as follows

$$\begin{pmatrix} a1 & a2 \\ b1 & b2 \end{pmatrix} \otimes \begin{pmatrix} A1 & A2 \\ B1 & B2 \end{pmatrix} = \begin{pmatrix} a1A1 & a2A1 & a1A2 & a2A2 \\ b1A1 & b2A1 & b1A2 & b2A2 \\ a1B1 & a2B1 & a1B2 & a2B2 \\ b1B1 & b2B1 & b1B2 & b2B2 \end{pmatrix} \tag{1.317}$$

As you can see, with the representation $3 \otimes 3^*$, we obtain 9-dimensional matrices that are unitarily equivalent to other 9-dimensional matrices consisting of two diagonal blocks of dimensions 1 and 8, respectively (with zeros outside these blocks). We will then write

$$3 \otimes 3^* = 1 \oplus 8 \tag{1.318}$$

We are not interested in representation **1** because it is trivial, reducing to the number 0.

On the other hand, representation **8**... yes, indeed: this is the adjoint representation!

We have already considered it through Eqs. (1.282) in particular giving rise to

$$H^1 = T_z = \frac{i}{2} \begin{pmatrix} 0 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.319}$$

$$H^2 = \frac{\sqrt{3}}{2}Y = i\frac{\sqrt{3}}{2}\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.320}$$

Both of these matrices have a zero eigenvalue with a multiplicity of two for the eigenvectors

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T \; ; \; \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}^T \tag{1.321}$$

The vectors

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \pm i & 1 & 0 \end{pmatrix}^T \; ; \; \begin{pmatrix} 0 & 0 & 0 & \pm i & 1 & 0 & 0 & 0 \end{pmatrix}^T \tag{1.322}$$

are also simultaneous eigenvectors of H^1 and H^2, with respective eigenvalues

$$\left\{\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}\right\} \tag{1.323}$$

for H^1 and

$$\left\{-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}, -\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}\right\} \tag{1.324}$$

for H^2.

Finally, the vectors

$$\begin{pmatrix} \pm i & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}^T \tag{1.325}$$

are simultaneous eigenvectors of H^1 and H^2, with respective eigenvalues ∓ 1 and 0.

Let's tally up the weights! They are

$$(0,0)\,;\,(0,0)\,;\,(\frac{1}{2},-\frac{\sqrt{3}}{2})\,;\,(-\frac{1}{2},\frac{\sqrt{3}}{2})\,;\,(-\frac{1}{2},-\frac{\sqrt{3}}{2})\,;\,(\frac{1}{2},\frac{\sqrt{3}}{2})\,;\,(-1,0)\,;\,(1,0) \tag{1.326}$$

In other words, they coincide with the roots!

This is a result that is generally encountered in the study of Lie algebras: the weights diagram is identical to the roots diagram for the adjoint representation.

Graphically, we are once again looking at the six vertices of a regular hexagon and the center of that hexagon.

But what physical particles should we place there? Naturally, mesons.

This is 1961, and the number of known mesons is limited. Nevertheless, they can be accounted for by knowing their isospin (the first component of the weights) and their hypercharge (the second component of the weights, to within a factor of $\frac{\sqrt{3}}{2}$):

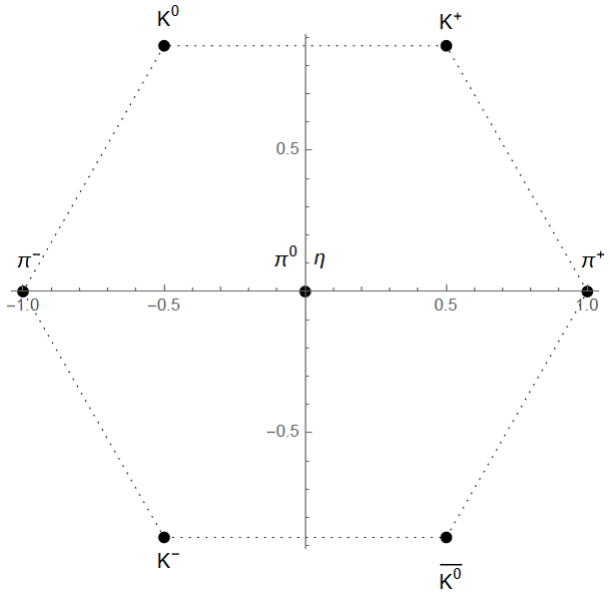


Figure 1.14: Mesons octet

The three π mesons and the four K mesons are in place. Gell-Mann completes the weights diagram by placing the particle called η, in the same position as π^0.

Not bad for a theory developed by a Norwegian mathematician nearly seventy years ago—one that no one would have bet a penny on having any future

use in physics...

It's all the more remarkable because the story doesn't end there!

Because in addition to mesons, there are also baryons. And since these particles are made up of three quarks, we'll have to consider the representation

$$3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10 \tag{1.327}$$

And here, a new representation emerges involving 10-dimensional matrices.

Its weights diagram takes the form of a triangle with four particles $(\Delta^-, \Delta^0, \Delta^+, \Delta^{++})$ on the top base, three particles $(\Sigma^{*-}, \Sigma^{*0}, \Sigma^{*+})$ on the level below, two particles (Ξ^{*-}, Ξ^{*0}) on the level below that, and finally the apex... empty.

Using his knowledge of the algebra $su(3, C)$, Gell-Mann was able to predict the quantum numbers of this new particle. It was confirmed experimentally shortly thereafter: it is the Ω^- baryon:

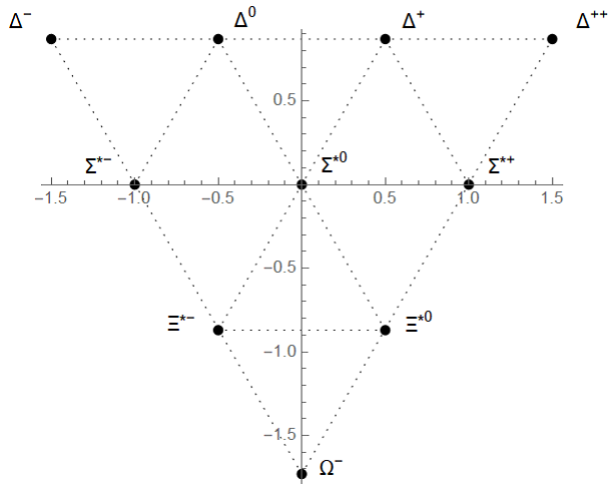


Figure 1.15: Baryons decuplet

We can see this confirmed by the example of $su(3, C)$: Lie algebras are extraordinarily useful in quantum physics. They have made predictions that, so far, have all been confirmed by experiment! And remember they predict the existence of particles with negative mass, based on the representations of the Poincaré algebra...

Let us continue by proposing a generalization of the latter.

1.10 The complex extension of the Poincaré group

Up until now, like the vast majority of physicists, we have worked with real spacetime coordinates.

There's no reason to change that, at first glance.

Well, actually, there is.

The theoretical effect is well known: light passing through a material can be slowed down depending on the number of interactions it has with the atoms in that object. Its speed is therefore no longer truly constant, and a time shift occurs. The catch is that this shift is thought to be imaginary.

Imaginary time is a strange thing. Instead of flowing along a real line—mostly from the past to the future, or in the opposite direction if we want to avoid negative energies, as we shall see—it is, in a sense, frozen, and in any case, orthogonal to real time.

It seemed like pure fantasy until two researchers from the University of Maryland proved it through an experiment[7].

That really throws everything into question.

Because there are no provisions—or almost none—for cases involving complex spacetime coordinates...

We'll still present a first result: the generalization of the Poincaré group[8]

Remember: we defined the Lorentz group as the set of transformations that leave the pseudo-norm (1.91) invariant for every 4-vector (1.90) in real spacetime.

We generalize this by requiring the invariance of a pseudo-norm given by

$$r_C^\dagger \, g \, r_C \tag{1.328}$$

and this time it applies to a complex 4-vector

$$r_C = \begin{pmatrix} t_C \\ x_C \\ y_C \\ z_C \end{pmatrix} \; ; \; x_C^\mu \equiv x_R^\mu + i \, x_I^\mu \; , \mu = 0,\, 1,\, 2,\, 3. \tag{1.329}$$

In other words, this amounts to asking for the preservation of the real pseudo-norm

$$R^T \, G \, R \; ; \; G = diag(1, -1, -1, -1, 1, -1, -1, -1) \tag{1.330}$$

where

[7] Giovanelli, I.L. and Anlage, S.M. (2025). Physical Interpretation of Imaginary Time Delay, to appear in *Physical Review Letters*. DOI: http://doi.org/10.1103/nnk7-xy4v

[8] Debergh, N., D'Agostini, G., and Petit, J.-P. (2021) On the Poincaré Algebra in a Complex Space-Time Manifold. *Journal of Modern Physics*, 12, 218–228. DOI: 10.4236/jmp.2021.123017.

$$R = \begin{pmatrix} t_R \\ x_R \\ y_R \\ z_R \\ t_I \\ x_I \\ y_I \\ z_I \end{pmatrix} = \left\{ x^A \; ; \; A = 0,\, 1,\, 2, ...,\, 7 \right\} \tag{1.331}$$

By definition, the respective (real) Lie groups that ensure these respective conservations are $U(1,3,C)$ and $O(2,6)$. Surprisingly, since one might think that the two approaches are equivalent, they do not have the same number of parameters: $U(1,3,C)$ has 16 parameters, while $O(2,6)$ has 29 parameters. We will see why later.

We immediately impose the constraint that the determinants of the relevant matrices must be equal to 1. This allows us to move directly to algebra. This is, in fact, the procedure that was already followed for the Lorentz group. We therefore consider the groups $SU(1,3,C)$ and $SO(2,6)$ with 15 and 28 parameters, respectively.

Let's first consider the version with the eight real coordinates, since it is the most directly accessible based on what we know about the Lorentz group and its algebra.

1.10.1 The extended real Poincaré group

The (2+6)-D Poincaré group is defined as the set of matrices given by (compare to (1.152))

$$U_R(\Lambda_R, a_R) = \begin{pmatrix} \Lambda_R & a_R \\ 0 & 1 \end{pmatrix} \tag{1.332}$$

with

$$\Lambda_R^T \, G \, \Lambda_R = G \tag{1.333}$$

Its symmetric element is

$$U_R^{-1}(\Lambda_R, a_R) = \begin{pmatrix} \Lambda_R^{-1} & -\Lambda_R^{-1} a_R \\ 0 & 1 \end{pmatrix} \tag{1.334}$$

What types of transformations are covered by the request (1.333)?
Well, a lot!
Acting on the 8-vector given in (1.331), you find

- 12 transformations of boost-type (i.e. involving a -real or imaginary- time component and a -real or imaginary- space component):

$$\begin{cases} t'_a = \cosh(\theta^{0j}_{ab})\, t_a + \sinh(\theta^{0j}_{ab})\, x^j_b \\ x'^{,j}_b = \sinh(\theta^{0j}_{ab})\, t_a + \cosh(\theta^{0j}_{ab})\, x^j_b \end{cases} \tag{1.335}$$

with $j = 1, 2, 3$ and $a, b \in \{R, I\}$.

- 12 transformations of rotation-type (i.e. "rotations" between a -real or imaginary- space component and another -real or imaginary- space component):

$$\begin{cases} x'^{,k}_a = \cos(\theta^{jk}_{ab})\, x^k_a - \sin(\theta^{jk}_{ab})\, x^j_b \\ x'^{,j}_b = \sin(\theta^{jk}_{ab})\, x^k_a + \cos(\theta^{jk}_{ab})\, x^j_b \end{cases} \tag{1.336}$$

with $j = 1, 2, 3$, $j \neq k$ and $a, b \in \{R, I\}$.

- 4 rotations between real and imaginary parts of one of the component

$$\begin{cases} t'_R = \cos(\theta^{00})\, t_R + \sin(\theta^{00})\, t_I \\ t'_I = -\sin(\theta^{00})\, t_R + \cos(\theta^{00})\, t_I \end{cases} \tag{1.337}$$

$$\begin{cases} x'^{,j}_R = \cos(\theta^{jj})\, x^j_R - \sin(\theta^{jj})\, x^j_I \\ x'^{,j}_I = \sin(\theta^{jj})\, x^j_R + \cos(\theta^{jj})\, x^j_I \end{cases} \tag{1.338}$$

These latest transformations are particularly interesting. Indeed, if we set $\theta^{\mu\mu} = \pi$ for each of these transformations, we obtain the discrete symmetries linking the connected components of the Lorentz group. This means that these symmetries, which are discrete in the real case, actually become continuous if we extend spacetime to complex coordinates.

1.10.2 The extended real Poincaré algebra

The corresponding (2+6)-D algebra is associated with

$$Z = \begin{pmatrix} \omega & \gamma \\ 0 & 0 \end{pmatrix} \tag{1.339}$$

where

$$\omega^T G = -G\,\omega \tag{1.340}$$

Following Souriau[9], the adjoint representation is characterized by

$$Z\prime = U_R(\Lambda_R, a_R)\, Z\, U_R^{-1}(\Lambda_R, a_R) \tag{1.341}$$

Using (1.332) as well as (1.335), this leads to

$$\omega\prime = \Lambda_R\,\omega\,\Lambda_R^{-1} \tag{1.342}$$

and

$$\gamma\prime = -\Lambda_R\,\omega\,\Lambda_R^{-1}\,a_R + \Lambda_R\,\gamma \tag{1.343}$$

Then, we consider the torsor μ defined by

$$\mu(Z) \equiv \frac{1}{2}\,Tr(J\,\omega) + P^T\,G\,\gamma\ ;\ P\ \in \mathbb{R}^8,\, G\,J^T = -J\,G \tag{1.344}$$

and require its invariance

$$\mu(Z) = \mu\prime(Z\prime) \tag{1.345}$$

In other words, we have

$$\frac{1}{2}\,Tr(J\,\omega) + P^T\,G\,\gamma = \frac{1}{2}\,Tr(J\prime\,\Lambda_R\,\omega\,\Lambda_R^{-1}) + P\prime^T\,G\,(-\Lambda_R\,\omega\,\Lambda_R^{-1}\,a_R + \Lambda_R\,\gamma) \tag{1.346}$$

by use of Eqs. (1.338) and (1.339).

By equating the $\gamma-$terms, we obtain the transformation law for translation operators

$$P\prime = \Lambda_R\,P \tag{1.347}$$

and are left with

$$\frac{1}{2}\,Tr(J\,\omega) = \frac{1}{2}\,Tr(J\prime\,\Lambda_R\,\omega\,\Lambda_R^{-1}) - P^T\,G\,\omega\,\Lambda_R^{-1}\,a_R \tag{1.348}$$

Taking a look at the last term of Eq. (1.348), we can see that it is a number resulting of the product of a transposed 8-vector (P^T), a 8 by 8 matrix ($G\,\omega\,\Lambda_R^{-1}$) and a 8-vector ($a_R$). This means that we can rewrite it as

$$P^T\,G\,\omega\,\Lambda_R^{-1}\,a_R = Tr(G\,\omega\,\Lambda_R^{-1}\,a_R\,P^T)$$

Remembering also that the trace is invariant under a circular permutation, we obtain

[9]Souriau, J.M. (1997). *Structure of Dynamical Systems.* Progress in Mathematics. Springer (Boston). https://doi.org/10.1007/978-1-4612-0281-3

$$\frac{1}{2}\,Tr(\Lambda_R^{-1}\,J\prime\,\Lambda_R\,\omega) = \frac{1}{2}\,Tr(J\,\omega) + Tr(\Lambda_R^{-1}\,a_R\,P^T\,G\,\omega) \tag{1.349}$$

We then have

$$Tr(\Lambda_R^{-1}\,J\prime\,\Lambda_R\,\omega) = Tr(J\,\omega) + Tr(\Lambda_R^{-1}\,a_R\,P^T\,G\,\omega) - Tr(P\,a_R^T\,G\,\Lambda_R\,\omega) \tag{1.350}$$

We thus come to the conclusion that

$$J\prime = \Lambda_R\,J\,G\,\Lambda_R^T\,G + a_R\,P^T\,\Lambda_R^T\,G - \Lambda_R\,P\,a_R^T\,G \tag{1.351}$$

With Eqs. (1.331) and (1.332), we can write the transformations on the coordinates as

$$R\prime = \Lambda_R\,R + a_R \tag{1.352}$$

which can be reversed as

$$R = G\,\Lambda_R^T\,G\,R\prime - G\,\Lambda_R^T\,G\,a_R \tag{1.353}$$

This implies that the 8-gradient

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial t_R} \\ \frac{\partial}{\partial x_R} \\ \frac{\partial}{\partial y_R} \\ \frac{\partial}{\partial z_R} \\ \frac{\partial}{\partial t_I} \\ \frac{\partial}{\partial x_I} \\ \frac{\partial}{\partial y_I} \\ \frac{\partial}{\partial z_I} \end{pmatrix} = \left\{ \frac{\partial}{\partial x^A}\ ;\ A = 0,\,1,\,2,...,\,7 \right\} \tag{1.354}$$

transforms as

$$\nabla\prime = G\,\Lambda_R\,G\,\nabla \tag{1.355}$$

A simple comparison with Eq. (1.347) allows us to affirm that

$$P = k\,G\,\nabla \tag{1.356}$$

where k is a real constant. This constant can be chosen as $k = i\,\hbar$ (or $k = i$ if the Dirac constant is taken as being equal to 1, as it will be the case in the following) in accordance with physics[10]. That means that

[10] which, from a mathematical point of view, is less relevant since it leads us to consider complex extensions of algebras rather than the algebras themselves.

$$P_A = i\,\frac{\partial}{\partial x^A} \equiv i\,\partial_A\ ;\ P^A = i\,G^{AB}\,\frac{\partial}{\partial x^B} \equiv i\,\partial^A \tag{1.357}$$

Note that the simultaneous use of (1.347) and (1.352) leads to

$$J = (R\,P^T - P\,R^T)\,G \tag{1.358}$$

Indeed, we have successively

$$J\prime = R\prime\,P\prime^T\,G - P\prime\,R\prime^T\,G = (\Lambda_R\,R + a_R)\,(P^T\,\Lambda_R^T)\,G - (\Lambda_R\,P)\,(R^T\,\Lambda_R^T + a_R^T)\,G$$

$$\Leftrightarrow J\prime = \Lambda_R\,(R\,P^T - P\,R^T)\,\Lambda_R^T\,G + a_R\,P^T\,\Lambda_R^T\,G - \Lambda_R\,P\,a_R^T\,G$$

$$\Leftrightarrow J\prime = \Lambda_R\,J\,G\,\Lambda_R^T\,G + a_R\,P^T\,\Lambda_R^T\,G - \Lambda_R\,P\,a_R^T\,G$$

We recover (1.351) as expected.

If we expand the last constraint of (1.340), we have

$$J = \begin{pmatrix} 0 & J^{01} & J^{02} & J^{03} & J^{04} & J^{05} & J^{06} & J^{07} \\ J^{01} & 0 & J^{12} & J^{13} & J^{14} & J^{15} & J^{16} & J^{17} \\ J^{02} & -J^{12} & 0 & J^{23} & J^{24} & J^{25} & J^{26} & J^{27} \\ J^{03} & -J^{13} & -J^{23} & 0 & J^{34} & J^{35} & J^{36} & J^{37} \\ -J^{04} & J^{14} & J^{24} & j^{34} & 0 & J^{45} & J^{46} & J^{47} \\ J^{05} & -J^{15} & -J^{25} & -J^{35} & J^{45} & 0 & J^{56} & J^{57} \\ J^{06} & -J^{16} & -J^{26} & -J^{36} & J^{46} & -J^{56} & 0 & J^{67} \\ J^{07} & -J^{17} & -J^{27} & -J^{37} & J^{47} & -J^{57} & -J^{67} & 0 \end{pmatrix} \tag{1.359}$$

Comparing this expression to (1.358), it is clear that

$$J^{AB} = i\,(x^A\,\partial^B - x^B\,\partial^A) \tag{1.360}$$

Using

$$[x^A, \partial^B] = -G^{AB} \tag{1.361}$$

it is straightforward to find out the commutation relations of the (2+6)-D Poincaré algebra. They read (compare with Eqs. (1.173)-(1.175))

$$[P^A, P^B] = 0 \tag{1.362}$$

$$[J^{AB}, P^C] = -i\,(g^{AC}P^B - g^{BC}P^A) \tag{1.363}$$

$$[J^{AB}, J^{CD}] = i\,(g^{BC}J^{AD} + g^{AD}J^{BC} + g^{AC}J^{DB} + g^{DB}J^{CA}) \tag{1.364}$$

1.10.3 The complex Poincaré algebra

So far, we've been working with eight real coordinates. But the initial goal is still to reduce them to four complex coordinates!

However, to ensure consistency, we need to consider linear combinations of the J^{AB} operators highlighted earlier.

Indeed, let's consider, for example, the action of J^{41} on the eight real coordinates (the ones that undergo a change, at any rate!). Due to the fact that

$$J^{41} = -i\,t_I\,\frac{\partial}{\partial x_R} - i\,x_R\,\frac{\partial}{\partial t_I} \tag{1.365}$$

this generator only changes x_R and t_I, resulting in

$$J^{41}\,x_R = -i\,t_I\;;\; J^{41}\,t_I = -i\,x_R \tag{1.366}$$

To reconstruct the complex variable $x_C = x_R + ix_I$ and the complex time $t_C = t_R + it_I$, we must couple J^{41} to

$$J^{05} = -i\,t_R\,\frac{\partial}{\partial x_I} - i\,x_I\,\frac{\partial}{\partial t_R} \tag{1.367}$$

such that

$$J^{05}\,x_I = -i\,t_R\;;\; J^{05}\,t_R = -i\,x_I \tag{1.368}$$

This leaves us with two options:

- $J^{41} - J^{05}$: we then have

$$(J^{41} - J^{05})\,x_C = (J^{41} - J^{05})\,(x_R + ix_I) = -i\,t_I - t_R = -t_C \tag{1.369}$$

$$(J^{41} - J^{05})\,t_C = (J^{41} - J^{05})\,(t_R + it_I) = i\,x_I + x_R = x_C \tag{1.370}$$

- $J^{41} + J^{05}$: we then have

$$(J^{41} + J^{05})\,x_C = (J^{41} + J^{05})\,(x_R + ix_I) = -i\,t_I + t_R = t_C^* \tag{1.371}$$

$$(J^{41} + J^{05})\,t_C = (J^{41} + J^{05})\,(t_R + it_I) = -i\,x_I + x_R = x_C^* \tag{1.372}$$

We will clearly distinguish between the two options, since the first is a direct and fairly natural generalization of a standard Lorentzian boost. The second, on the other hand, is more unexpected, mixing the variables with their complex conjugates.

We therefore begin by defining

1. The complex equivalents of the Lorentz boosts:

$$M^{01} \equiv (J^{01} + J^{45}) \; ; \; M^{02} \equiv (J^{02} + J^{46}) \; ; \; M^{03} \equiv (J^{03} + J^{47}) \tag{1.373}$$

2. The complex equivalents of the Lorentz rotations:

$$M^{12} \equiv (J^{12} + J^{56}) \; ; \; M^{31} \equiv (J^{31} + J^{75}) \; ; \; M^{23} \equiv (J^{23} + J^{67}) \tag{1.374}$$

Let's take a closer look at these six generators.

We can write them in a more general way:

$$M^{\mu\nu} = i\, x_C^{\mu}\, \partial_C^{\nu} - i\, x_C^{\nu}\, \partial_C^{\mu} + i\, x_C^{\mu *}\, \partial_C^{*\nu} - i\, x_C^{\nu *}\, \partial_C^{*\mu} \tag{1.375}$$

having used (1.329) and its consequence

$$\partial_{\mu C} \equiv \frac{\partial}{\partial x_C^{\mu}} \; ; \; \partial_C^{\mu} \equiv g^{\mu\nu}\, \partial_{\nu C} \tag{1.376}$$

as well as

$$\partial_{\mu C}^{*} \equiv \frac{\partial}{\partial x_C^{\mu *}} \; ; \; \partial_C^{*\mu} \equiv g^{\mu\nu}\, \partial_{\nu C}^{*} \tag{1.377}$$

The commutation relations are then (compare with (1.134)):

$$\boxed{[M^{\mu\nu}, M^{\rho\sigma}] = i\, (g^{\rho\nu} M^{\mu\sigma} + g^{\sigma\nu} M^{\rho\mu} + g^{\rho\mu} M^{\sigma\nu} + g^{\sigma\mu} M^{\nu\rho})} \tag{1.378}$$

These are, of course, the commutation relations of the (usual) Lorentz algebra. The only difference is that the transformations involve complex coordinates rather than real ones.

What do these operators do?

We get

$$M^{0j}\, x_C^0 = -i\, x_C^j \; ; \; M^{0j}\, x_C^j = -i\, x_C^0 \tag{1.379}$$

$$M^{jk}\, x_C^j = i\, x_C^k \; ; \; M^{jk}\, x_C^k = -i\, x_C^j \tag{1.380}$$

One immediate implication of these results is that we will be able to represent the actions of these operators using 4-dimensional matrices by acting on the complex position 4-vector (1.329).

These matrices are then

$$M^{01} = \begin{pmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M^{02} = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M^{03} = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \tag{1.381}$$

$$M^{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, M^{31} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}, M^{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \tag{1.382}$$

But right now, we only have six operators... there are still twenty-two left!

Let's start with the following nine[11]:

$$M_C^{01} \equiv -J^{05} - J^{14}\,;\; M_C^{02} \equiv -J^{06} - J^{24}\,;\; M_C^{03} \equiv -J^{07} - J^{34} \tag{1.383}$$

$$M_C^{12} \equiv -J^{16} - J^{25}\,;\; M_C^{31} \equiv -J^{17} - J^{35}\,;\; M_C^{23} \equiv -J^{27} - J^{36} \tag{1.384}$$

$$M_C^{11} - M_C^{00} \equiv 2(J^{04} - J^{15})\,;\, M_C^{22} - M_C^{00} \equiv 2(J^{04} - J^{26})\,;\, M_C^{33} - M_C^{00} \equiv 2(J^{04} - J^{37}) \tag{1.385}$$

How do they affect the coordinates? Like this:

$$M_C^{0j}\, x_C^0 = x_C^j\,;\; M_C^{0j}\, x_C^j = -x_C^0 \tag{1.386}$$

$$M_C^{jk}\, x_C^j = -x_C^k\,;\; M_C^{jk}\, x_C^k = -x_C^j \tag{1.387}$$

$$(M_C^{jj} - M_C^{00})\, x_C^0 = -2x_C^0\,;\; (M_C^{jj} - M_C^{00})\, x_C^j = -2x_C^j \tag{1.388}$$

It is worth noting the significance of these operators, which map a coordinate to its opposite (up to a factor 2) and thus have the same effect as discrete symmetry operators within the framework of Lorentz algebra.

[11] The linear combinations might come as a surprise... they are, in fact, due to the decision to consider only $so(2,6)$ at the outset, rather than $o(2,6)$.

We can summarize these operators as follows:

$$M_C^{\mu\nu} = x_C^\mu \partial_C^\nu + x_C^\nu \partial_C^\mu - x_C^{\mu *} \partial_C^{*\nu} - x_C^{\nu *} \partial_C^{*\mu} \tag{1.389}$$

They combine with the previous six to form the algebra $su(1,3,C)$, characterized by the relations (1.378) and

$$[M^{\mu\nu}, M_C^{\rho\sigma}] = i\, g^{\rho\nu} M_C^{\sigma\mu} + i\, g^{\sigma\nu} M_C^{\mu\rho} - i\, g^{\rho\mu} M_C^{\nu\sigma} - i\, g^{\sigma\mu} M_C^{\rho\nu} \tag{1.390}$$

$$[M_C^{\mu\nu}, M_C^{\rho\sigma}] = i\, g^{\rho\nu} M^{\sigma\mu} - i\, g^{\sigma\nu} M^{\mu\rho} - i\, g^{\rho\mu} M^{\nu\sigma} + i\, g^{\sigma\mu} M^{\rho\nu} \tag{1.391}$$

Since the matrices of $su(1,3,C)$ are of dimension 4, it is not surprising that we also have

$$M_C^{01} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; M_C^{02} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; M_C^{03} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{1.392}$$

$$M_C^{12} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; M_C^{31} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \; M_C^{23} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{1.393}$$

$$M_C^{00} = 2 \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; M_C^{11} = 2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.394}$$

$$M_C^{22} = 2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \; M_C^{33} = 2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.395}$$

The set of sixteen generators of $u(1,3,C)$ yields the elements of the following group, after exponentiation:

$$B^{01} = \exp(i\varphi M^{01}) = \begin{pmatrix} \cosh\varphi & \sinh\varphi & 0 & 0 \\ \sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.396}$$

$$B^{02} = \exp(i\varphi M^{02}) = \begin{pmatrix} \cosh\varphi & 0 & \sinh\varphi & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\varphi & 0 & \cosh\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.397}$$

$$B^{03} = \exp(i\varphi M^{03}) = \begin{pmatrix} \cosh\varphi & 0 & 0 & \sinh\varphi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\varphi & 0 & 0 & \cosh\varphi \end{pmatrix} \tag{1.398}$$

$$R^{12} = \exp(i\theta M^{12}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & \sin\theta & 0 \\ 0 & -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.399}$$

$$R^{31} = \exp(i\theta M^{31}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & -\sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & \sin\theta & 0 & \cos\theta \end{pmatrix} \tag{1.400}$$

$$R^{23} = \exp(i\theta M^{23}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix} \tag{1.401}$$

$$B_C^{01} = \exp(i\varphi M_C^{01}) = \begin{pmatrix} \cosh\varphi & i\sinh\varphi & 0 & 0 \\ -i\sinh\varphi & \cosh\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.402}$$

$$B_C^{02} = \exp(i\varphi M_C^{02}) = \begin{pmatrix} \cosh\varphi & 0 & i\sinh\varphi & 0 \\ 0 & 1 & 0 & 0 \\ -i\sinh\varphi & 0 & \cosh\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.403}$$

$$B_C^{03} = \exp(i\varphi M_C^{03}) = \begin{pmatrix} \cosh\varphi & 0 & 0 & i\sinh\varphi \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -i\sinh\varphi & 0 & 0 & \cosh\varphi \end{pmatrix} \tag{1.404}$$

$$R_C^{12} = \exp(i\theta M_C^{12}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & i\sin\theta & 0 \\ 0 & i\sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.405}$$

$$R_C^{31} = \exp(i\theta M_C^{31}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & 0 & i\sin\theta \\ 0 & 0 & 1 & 0 \\ 0 & i\sin\theta & 0 & \cos\theta \end{pmatrix} \tag{1.406}$$

$$R_C^{23} = \exp(i\theta M_C^{23}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta & i\sin\theta \\ 0 & 0 & i\sin\theta & \cos\theta \end{pmatrix} \tag{1.407}$$

$$T = \exp(i\frac{\kappa}{2} M_C^{00}) = \begin{pmatrix} \exp(-i\kappa) & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.408}$$

$$P(1) = \exp(i\frac{\kappa}{2} M_C^{11}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \exp(i\kappa) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.409}$$

$$P(2) = \exp(i\frac{\kappa}{2} M_C^{22}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \exp(i\kappa) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{1.410}$$

$$P(3) = \exp(i\frac{\kappa}{2} M_C^{33}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \exp(i\kappa) \end{pmatrix} \tag{1.411}$$

Here again, we see that the transition to the complex plane has made it possible to transform the discrete P and T symmetry operators into continuous symmetries through a rotation of the coordinates in the complex plane. The discrete case is recovered by taking $\kappa = \pi$.

The remaining generators complete $su(1,3,C)$ to form $so(2,6)$.

What makes them so special?

Let's see.
Some of them are given by

$$N^{01} = -J^{01} + J^{45} \; ; \; N^{02} = -J^{02} + J^{46} \; ; \; N^{03} = -J^{03} + J^{47} \tag{1.412}$$

$$N^{12} = -J^{12} + J^{56} \; ; \; N^{31} = -J^{31} + J^{75} \; ; \; N^{23} = -J^{23} + J^{67} \tag{1.413}$$

They act on complex coordinates like this:

$$N^{0j}\, x^0_C = i\, x^{j*}_C \; ; \; N^{0j}\, x^j_C = i\, x^{0*}_C \tag{1.414}$$

$$N^{jk}\, x^j_C = -i\, x^{k*}_C \; ; \; N^{jk}\, x^k_C = i\, x^{j*}_C \tag{1.415}$$

It is precisely the complex conjugation in the right-hand side that makes these operators antiunitary and causes them to lie outside the algebra $su(1,3,C)$. That's why they cannot be associated with 4-dimensional matrices.

In general, they are written as

$$N^{\mu\nu} = -i\, x^{\mu*}_C \partial^\nu_C + i\, x^{\nu*}_C \partial^\mu_C - i\, x^\mu_C \partial^{*\nu}_C + i\, x^\nu_C \partial^{*\mu}_C \tag{1.416}$$

and satisfy the commutation relations

$$[M^{\mu\nu}, N^{\rho\sigma}] = -i\, g^{\rho\nu} N^{\sigma\mu} - i\, g^{\sigma\nu} N^{\mu\rho} - i\, g^{\rho\mu} N^{\nu\sigma} - i\, g^{\sigma\mu} N^{\rho\nu} \tag{1.417}$$

$$[M^{\mu\nu}_C, N^{\rho\sigma}] = i\, g^{\rho\nu} N^{\sigma\mu}_C + i\, g^{\sigma\nu} N^{\mu\rho}_C - i\, g^{\rho\mu} N^{\nu\sigma}_C - i\, g^{\sigma\mu} N^{\rho\nu}_C \tag{1.418}$$

$$[N^{\mu\nu}, N^{\rho\sigma}] = -i\, g^{\rho\nu} M^{\sigma\mu} - i\, g^{\sigma\nu} M^{\mu\rho} - i\, g^{\rho\mu} M^{\nu\sigma} - i\, g^{\sigma\mu} M^{\rho\nu} \tag{1.419}$$

We see new operators appearing in (1.418). They are defined by

$$N^{\mu\nu}_C = -x^{\mu*}_C \partial^\nu_C + x^{\nu*}_C \partial^\mu_C + x^\mu_C \partial^{*\nu}_C - x^\nu_C \partial^{*\mu}_C \tag{1.420}$$

and act as

$$N^{0j}_C\, x^0_C = x^{j*}_C \; ; \; N^{0j}_C\, x^j_C = x^{0*}_C \tag{1.421}$$

$$N_C^{jk}\, x_C^j = -x_C^{k*}\ ;\ N_C^{jk}\, x_C^k = x_C^{j*} \tag{1.422}$$

We are now ready for the final commutation relations, which, together with Eqs. (1.378), (1.390), (1.391), (1.417)-(1.419), define the extension of the Lorentz algebra to a complex spacetime. Here they are

$$[M^{\mu\nu}, N_C^{\rho\sigma}] = -i\, g^{\rho\nu} N_C^{\sigma\mu} - i\, g^{\sigma\nu} N_C^{\mu\rho} - i\, g^{\rho\mu} N_C^{\nu\sigma} - i\, g^{\sigma\mu} N_C^{\rho\nu} \tag{1.423}$$

$$[M_C^{\mu\nu}, N_C^{\rho\sigma}] = -i\, g^{\rho\nu} N^{\sigma\mu} - i\, g^{\sigma\nu} N^{\mu\rho} - i\, g^{\rho\mu} N^{\nu\sigma} - i\, g^{\sigma\mu} N^{\rho\nu} \tag{1.424}$$

$$[N^{\mu\nu}, N_C^{\rho\sigma}] = i\, g^{\rho\nu} M_C^{\sigma\mu} - i\, g^{\sigma\nu} M_C^{\mu\rho} - i\, g^{\mu\rho} M_C^{\nu\sigma} + i\, g^{\sigma\mu} M_C^{\rho\nu} \tag{1.425}$$

$$[N_C^{\mu\nu}, N_C^{\rho\sigma}] = -i\, g^{\rho\nu} M^{\sigma\mu} - i\, g^{\sigma\nu} M^{\mu\rho} - i\, g^{\rho\mu} M^{\nu\sigma} - i\, g^{\sigma\mu} M^{\rho\nu} \tag{1.426}$$

For the extension to Poincaré algebra, we must consider four additional operators associated with complex translations, in addition to the operators $M^{\mu\nu}$, $M_C^{\mu\nu}$, $N^{\mu\nu}$ and $N_C^{\mu\nu}$.

These operators are denoted by P^μ and are given by

$$P_C^\mu = i\, \partial_C^\mu = i\, g^{\mu\nu}\, \frac{\partial}{\partial x_C^\nu} \tag{1.427}$$

Of course, complex conjugates should also be considered:

$$P_C^{*\mu} = -i\, \partial_C^{*\mu} = -i\, g^{\mu\nu}\, \frac{\partial}{\partial x_C^{\nu *}} \tag{1.428}$$

The previous commutation relations are then supplemented by

$$[M^{\mu\nu}, P_C^\rho] = -i\, g^{\rho\mu} P_C^\nu + i\, g^{\rho\nu} P_C^\mu\ ;\ [M^{\mu\nu}, P_C^{*\rho}] = -i\, g^{\rho\mu} P^{*\nu} + i\, g^{\rho\nu} P^{*\mu} \tag{1.429}$$

$$[M_C^{\mu\nu}, P_C^\rho] = -g^{\mu\rho} P_C^\nu - g^{\rho\nu} P_C^\mu\ ;\ [M_C^{\mu\nu}, P_C^{*\rho}] = g^{\mu\rho} P_C^{\nu *} + g^{\rho\nu} P_C^{\mu *} \tag{1.430}$$

$$[N^{\mu\nu}, P_C^\rho] = -i\, g^{\rho\mu} P_C^{*\nu} + i\, g^{\rho\nu} P_C^{*\mu}\ ;\ [N^{\mu\nu}, P_C^{*\rho}] = -i\, g^{\rho\mu} P_C^\nu + i\, g^{\rho\nu} P_C^\mu \tag{1.431}$$

$$[N_C^{\mu\nu}, P_C^\rho] = g^{\rho\mu} P_C^{*\nu} - g^{\rho\nu} P_C^{*\mu}\ ;\ [N_C^{\mu\nu}, P_C^{*\rho}] = -g^{\rho\mu} P_C^\nu + g^{\rho\nu} P_C^\mu \tag{1.432}$$

$$[P^\mu, P^\nu] = [P^\mu, P^{*\nu}] = [P^{*\mu}, P^{*\nu}] = 0 \tag{1.433}$$

1.10.4 The Casimir operators

1.10.4.1 The first Casimir operator

The first Casimir equation is trivially

$$P^T G P \tag{1.434}$$

in its real version.

But it is, of course, the complex version that is particularly interesting:

$$C1 \equiv |P_C^0|^2 - |P_C^1|^2 - |P_C^2|^2 - |P_C^3|^2 \tag{1.435}$$

By analogy with the real spacetime case, it allows us to define the mass of a particle in this complex space we are dealing with:

$$\boxed{|P_C^0|^2 - |P_C^1|^2 - |P_C^2|^2 - |P_C^3|^2 = m^2} \tag{1.436}$$

Just as in the case of Poincaré algebra, the only constraint on the square of the mass is that it must be real... which allows for positive, negative, or imaginary masses.

The fundamental difference between what is proposed here and usual Poincaré algebra is that the energy (associated with P^0) is complex!

In the case of rest, where all momentum components except P_C^0 are set to zero, we thus have

$$E = m \exp(i\varphi) \tag{1.437}$$

The specific values $\varphi = 0$ or $\varphi = \pi$ correspond to the positive and negative energies found in any relativistic quantum theory (Dirac's theory -see Chapter 2- foremost among them), but the other values of φ clearly lead to something new.

1.10.4.2 Wigner's little groups

If we want to consider real, non-zero masses, we simply need to choose a complex energy such that

$$|P_C^0|^2 = m^2 \tag{1.438}$$

An example of this class is

$$k_C = \begin{pmatrix} m \exp(i\varphi) \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{1.439}$$

To be stabilized by an element of the group means to be annihilated by the corresponding element of the algebra. The matrices that satisfy this condition are:

$$M^{12}, M^{31}, M^{23}, M_C^{12}, M_C^{31}, M_C^{23}, M_C^{11} - M_C^{22}, M_C^{22} - M_C^{33} \tag{1.440}$$

Now, compare Eqs. (1.382), (1.393), (1.394) and (1.395) with the Gell-Mann matrices given in Eqs(1.247)-(1.249). It is clear that

$$M^{12} = \begin{pmatrix} 0 & 0 \\ 0 & 2F2 \end{pmatrix}, M^{31} = \begin{pmatrix} 0 & 0 \\ 0 & -2F5 \end{pmatrix}, M^{23} = \begin{pmatrix} 0 & 0 \\ 0 & 2F7 \end{pmatrix} \tag{1.441}$$

$$M_C^{12} = \begin{pmatrix} 0 & 0 \\ 0 & 2F1 \end{pmatrix}, M_C^{31} = \begin{pmatrix} 0 & 0 \\ 0 & 2F4 \end{pmatrix}, M_C^{23} = \begin{pmatrix} 0 & 0 \\ 0 & 2F6 \end{pmatrix} \tag{1.442}$$

$$M_C^{11} - M_C^{22} = \begin{pmatrix} 0 & 0 \\ 0 & 4F3 \end{pmatrix}, M_C^{22} - M_C^{33} = \begin{pmatrix} 0 & 0 \\ 0 & -2F3 + 2\sqrt{3}F8 \end{pmatrix} \tag{1.443}$$

In terms of the Cartan-Weyl basis, we have

$$H^1 = \frac{1}{4}(M_C^{11} - M_C^{22}),\ H^2 = \frac{\sqrt{3}}{12}(M_C^{11} + M_C^{22} - 2M_C^{33}) \tag{1.444}$$

$$E_{\alpha^1} = \frac{1}{2}(M^{12} - i\,M_C^{12}),\ E_{-\alpha^1} = \frac{1}{2}(M^{12} + i\,M_C^{12}) \tag{1.445}$$

$$E_{\alpha^2} = -\frac{i}{2}(M^{23} - i\,M_C^{23}),\ E_{-\alpha^2} = \frac{i}{2}(M^{23} + i\,M_C^{23}) \tag{1.446}$$

$$E_{\alpha^1+\alpha^2} = \frac{1}{2}(M^{31} + i\,M_C^{31}),\ E_{-\alpha^1-\alpha^2} = \frac{1}{2}(M^{31} - i\,M_C^{31}) \tag{1.447}$$

It is the straightforward to see that Eqs. (1.296)-(1.299) as well as Eqs. (1.301)-(1.304) are satisfied.

The root system and the matrix representation are unambiguous: the Wigner little group subtended by real non-zero masses is indeed $su(3, C)$ and it contains the subalgebra $so(3)$, generated by M^{12}, M^{31} and M^{23}, which emerges when only the real coordinates are considered.

The eigenstates of $H^1$and H^2 are

$$|\psi 1>= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, |\psi 2>= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, |\psi 3>= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, |\psi 4>= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad (1.448)$$

with respective pairs of eigenvalues

$$(0,0),\ (\frac{1}{2},\frac{\sqrt{3}}{6}),\ (-\frac{1}{2},\frac{\sqrt{3}}{6}),\ (0,-\frac{\sqrt{3}}{3}) \quad (1.449)$$

The weights diagram speaks for itself: the representation spanned by this little group is

$$1 \oplus 3 \quad (1.450)$$

i.e. $D(0,0)$, the trivial representation corresponding to weight $(0,0)$, and $D(1,0)$, the fundamental representation of dimension 3 associated with (1.310).

Let's move on to the next class, that of zero masses, for which we have, for example

$$|P_C^0|^2 - |P_C^3|^2 = 0 \quad (1.451)$$

A representative for this class is

$$k_C = \begin{pmatrix} \kappa\ \exp(i\varphi) \\ 0 \\ 0 \\ -\kappa\ \exp(i\varphi) \end{pmatrix} \quad (1.452)$$

The stabilizing algebra corresponds to

$$M^{12}, M_C^{12}, Q1 \equiv M^{02}+M^{23}, Q2 \equiv M^{01}-M^{31}, Q1_C \equiv M_C^{02}-M_C^{23}, Q2_C \equiv M_C^{01}-M_C^{31} \quad (1.453)$$

$$Y1 \equiv M_C^{11} - M_C^{22}, Y2 \equiv M_C^{00} + M_C^{33} - 2M_C^{03} \quad (1.454)$$

The non-zero commutation relations are

$$[M^{12}, Q1] = -i\,Q2\,,\ [M^{12,}Q2] = i\,Q1 \quad (1.455)$$

$$[M^{12}, Q1_C] = -i\,Q2_C\,,\ [M^{12,}Q2_C] = i\,Q1_C \quad (1.456)$$

$$[M_C^{12}, Q1] = i\,Q2_C\,,\ [M_C^{12}, Q2] = i\,Q1_C \tag{1.457}$$

$$[M_C^{12}, Q1_C] = -i\,Q2\,,\ [M_C^{12}, Q2_C] = -i\,Q1 \tag{1.458}$$

$$[M^{12}, M_C^{12}] = -i\,Y1\,,\ [Q1, Q1_C] = [Q2, Q2_C] = -i\,Y2 \tag{1.459}$$

$$[Y1, M^{12}] = -4i\,M_C^{12}\,,\ [Y1, M_C^{12}] = 4i\,M12M^{12} \tag{1.460}$$

$$[Y1, Q1] = -2i\,Q1_C\,,\ [Y1, Q1_C] = 2i\,Q1 \tag{1.461}$$

$$[Y1, Q2] = 2i\,Q2_C\,,\ [Y1, Q2_C] = -2i\,Q2 \tag{1.462}$$

Generator $Y2$ hasn't been forgotten: it commutes with all seven of the other generators!

This is a feature found, for example, in Heisenberg algebras. In fact, the algebra in question here is

$$so(3) \ominus h(4) \tag{1.463}$$

The subalgebra $so(3)$ is generated by

$$J_+ \equiv \frac{1}{2}(M^{12} + iM_C^{12})\,,\ J_- \equiv \frac{1}{2}(M^{12} - iM_C^{12})\,,\ J_z \equiv -\frac{1}{4}Y1 \tag{1.464}$$

whereas the part $h(4)$ is the Heisenberg algebra associated with

$$\{Q1, Q2, Q1_C, Q2_C, Y2\} \tag{1.465}$$

The notation $\ominus$ in (1.463) refers to a semidirect sum, implying that the generators of $so(3)$ do not commute with those of $h(4)$.

It is worth noting, moreover, that the result (1.463) seems more consistent in its complex form than the little group $E(2)$, which has raised many questions among researchers.

Finally, let's discuss the last class, that of imaginary masses, with

$$|P_C^3|^2 = -m^2 \equiv \xi^2 \tag{1.466}$$

A representative for this class is

$$k_C = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \xi\,\exp(i\varphi) \end{pmatrix} \tag{1.467}$$

The stabilizing algebra is then generated by

$$M^{01}, M^{02}, M^{12}, M_C^{01}, M_C^{02}, M_C^{12}, X1 \equiv M_C^{11} + M_C^{00}, X2 \equiv M_C^{22} + M_C^{00} \quad (1.468)$$

with commutation relations

$$[M^{01}, M^{02}] = -i\, M^{12}\,,\ [M^{12}, M^{01}] = i\, M^{02}\,,\ [M^{12}, M^{02}] = -i\, M^{01} \quad (1.469)$$

$$[M^{01}, M_C^{01}] = -i\, X1\,,\ [M^{01}, M_C^{02}] = -i\, M_C^{12}\,,\ [M^{01}, M_C^{12}] = -i\, M_C^{02} \quad (1.470)$$

$$[M^{02}, M_C^{01}] = -i\, M_C^{12}\,,\ [M^{02}, M_C^{02}] = -i\, X2\,,\ [M^{02}, M_C^{12}] = -i\, M_C^{01} \quad (1.471)$$

$$[M^{12}, M_C^{01}] = i\, M_C^{02}\,,\ [M^{12}, M_C^{02}] = -i\, M_C^{01}\,,\ [M^{12}, M_C^{12}] = -i\,(X1 - X2) \quad (1.472)$$

$$[M^{01}, X1] = -4i\, M_C^{01}\,,\ [M^{01}, X2] = -2i\, M_C^{01} \quad (1.473)$$

$$[M^{02}, X1] = -2i\, M_C^{02}\,,\ [M^{02}, X2] = -4i\, M_C^{02} \quad (1.474)$$

$$[M^{12}, X1] = 2i\, M_C^{12}\,,\ [M^{12}, X2] = -2i\, M_C^{12} \quad (1.475)$$

$$[M_C^{01}, X1] = 4i\, M^{01}\,,\ [M_C^{01}, X2] = 2i\, M^{01} \quad (1.476)$$

$$[M_C^{02}, X1] = 2i\, M^{02}\,,\ [M_C^{02}, X2] = 4i\, M^{02} \quad (1.477)$$

$$[M_C^{01}, M_C^{02}] = -i\, M^{12}\,,\ [M_C^{12}, M_C^{01}] = -i\, M^{02}\,,\ [M_C^{12}, M_C^{02}] = -i\, M^{01} \quad (1.478)$$

Once again, this closes to give rise to the algebra $su(1,2,C)$ which includes the algebra $so(1,2)$ (M^{01},M^{02},$M^{12}$and Eqs. (1.469)) of the real case as a substructure.

1.10.4.3 The second Casimir operator

We can reason in a similar way to what was done for the real case and propose the equivalent of the Pauli-Lubanski 4-vector in the form

$$W = c_{ABC}\, J^{AB}\, P^C \tag{1.479}$$

The constants c_{ABC} are determined by requiring W to commute with the momenta, as is the case for the standard Pauli-Lubanski 4-vector. We therefore require

$$[W, P^D] = -c_{ABC}(G^{AD}P^B - G^{BD}P^A)P^C = 2c^D_{BC}P^B P^C = 0 \tag{1.480}$$

This implies that c_{ABC} must be antisymmetric on B, C, in addition to being antisymmetric on A, B, by construction.

The forms of W are thus, in the generic form

$$W^{ABC} = J^{AB}P^C - J^{AC}P^B + J^{BC}P^A \tag{1.481}$$

That makes a total of 56 contributions! We can identify the four elements that appear in the real case:

$$W^{123} = W_0,\, W^{023} = -W_1,\, W^{013} = W_2,\, W^{012} = -W_3 \tag{1.482}$$

Operator (1.481) is indeed an antisymmetric tensor with respect to the indices A, B, and B, C. It is also symmetric with respect to the indices A and C.

We verify that it commutes with P^D:

$$[W^{ABC}, P^D] = [J^{AB}, P^D]P^C - [J^{AC}, P^D]P^B + [J^{BC}, P^D]P^A$$

$$\Leftrightarrow [W^{ABC}, P^D] = -G^{AD}[P^B, P^C] + G^{BD}[P^A, P^C] - G^{CD}[P^A, P^B] = 0 \tag{1.483}$$

How does it commute with J^{DE}? Here are the step-by-step calculations in detail:

$$\begin{aligned}[J^{AB}P^C - J^{AC}P^B + J^{BC}P^A, J^{DE}] &= [J^{AB}, J^{DE}]P^C - J^{AB}[J^{DE}, P^C] \\ &- [J^{AC}, J^{DE}]P^B + J^{AC}[J^{DE}, P^B] + [J^{BC}, J^{DE}]P^A - J^{BC}[J^{DE}, P^A]\end{aligned}$$

$$\begin{aligned}&= G^{BD}J^{AE}P^C + G^{BE}J^{DA}P^C + G^{AD}J^{EB}P^C + G^{AE}J^{BD}P^C + G^{CD}J^{AB}P^E - G^{CE}J^{AB}P^D \\ &- G^{CD}J^{AE}P^B - G^{CE}J^{DA}P^B - G^{AD}J^{EC}P^B - G^{AE}J^{CD}P^B - G^{BD}J^{AC}P^E + G^{BE}J^{AC}P^D \\ &+ G^{CD}J^{BE}P^A + G^{CE}J^{DB}P^A + G^{BD}J^{EC}P^A + G^{BE}J^{CD}P^A + G^{AD}J^{BC}P^E - G^{AE}J^{BC}P^D\end{aligned}$$

$$= G^{BD}(J^{AE}P^C - J^{AC}P^E + J^{EC}P^A) - G^{BE}(J^{AD}P^C - J^{AC}P^D + J^{DC}P^A)$$
$$+ G^{AD}(J^{EB}P^C - J^{EC}P^B + J^{BC}P^E) + G^{AE}(J^{BD}P^C - J^{CD}P^B - J^{BC}P^D)$$
$$+ G^{CD}(J^{BE}P^A + J^{AB}P^E - J^{AE}P^B) + G^{CE}(-J^{AB}P^D - J^{DA}P^B + J^{DB}P^A)$$

We thus have

$$[W^{ABC}, J^{DE}] = G^{AD}W^{BCE} - G^{AE}W^{BCD} - G^{BD}W^{ACE}$$
$$+ G^{BE}W^{ACD} + G^{CD}W^{ABE} - G^{CE}W^{ABD}$$

The second Casimir operator is then given by

$$C2 = W^{ABC}W_{ABC} \tag{1.484}$$

The information still needs to be completed by calculating the final commutation relations:

$$[W^{ABC}, W^{DEF}] = [J^{AB}P^C - J^{AC}P^B + J^{BC}P^A, J^{DE}P^F - J^{DF}P^E + J^{EF}P^D]$$

$$= [J^{AB}, J^{DE}]P^CP^F + J^{DE}[J^{AB}, P^F]P^C - J^{AB}[J^{DE}, P^C]P^F - [J^{AB}, J^{DF}]P^CP^F$$
$$- J^{DF}[J^{AB}, P^E]P^C + J^{AB}[J^{DF}, P^C]P^E + [J^{AB}, J^{EF}]P^CP^D + J^{EF}[J^{AB}, P^D]P^C$$
$$- J^{AB}[J^{EF}, P^C]P^D - [J^{AC}, J^{DE}]P^BP^F - J^{DE}[J^{AC}, P^F]P^B + J^{AC}[J^{DE}, P^B]P^F$$

$$+ [J^{AC}, J^{DF}]P^BP^E + J^{DF}[J^{AC}, P^E]P^B - J^{AC}[J^{DF}, P^B]P^E - [J^{AC}, J^{EF}]P^BP^D$$
$$- J^{EF}[J^{AC}, P^D]P^B + J^{AC}[J^{EF}, P^B]P^D + [J^{BC}, J^{DE}]P^AP^F + J^{DE}[J^{BC}, P^F]P^A$$
$$- J^{BC}[J^{DE}, P^A]P^F - [J^{BC}, J^{DF}]P^AP^E - J^{DF}[J^{BC}, P^E]P^A + J^{BC}[J^{DF}, P^A]P^E$$
$$+ [J^{BC}, J^{EF}]P^AP^D + J^{EF}[J^{BC}, P^D]P^A - J^{BC}[J^{EF}, P^A]P^D$$

We finally obtain

$$[W^{ABC}, W^{DEF}] = G^{AD}(W^{BCE}P^F - W^{BCF}P^E) + G^{AE}(W^{BCF}P^D - W^{BCD}P^F)$$
$$+ G^{AF}(W^{BCD}P^E - W^{BCE}P^D) + G^{BD}(W^{AEF}P^C - W^{CEF}P^A) + G^{BE}(W^{CDF}P^A - W^{ADF}P^C)$$
$$+ G^{BF}(W^{ACE}P^D - W^{ACD}P^E) + G^{CD}(W^{ABE}P^F - W^{ABF}P^E) + G^{CE}(W^{ABF}P^D - W^{ABD}P^F)$$
$$+ G^{CF}(W^{DAB}P^E - W^{EAB}P^D)$$

In the real case, the second Casimir operator is evaluated within a specific system: the rest frame. Here, the rest frame involves canceling out $P^1, P^2, P^3, P^5,\ P^6$ and P^7. Of the fifty-six W^{ABC} operators, twenty cancel out, and the second Casimir operator is written as follows

$$\begin{aligned}
C2 &= 4m^2(J_{12}^2 + J_{31}^2 + J_{23}^2) \\
&+ 4m^2(J_{15}^2 + J_{16}^2 + J_{17}^2 + J_{25}^2 + J_{26}^2 + J_{27}^2 + J_{35}^2 + J_{36}^2 + J_{37}^2 + J_{56}^2 + J_{57}^2 + J_{67}^2) \\
&- 4m^2\cos^2\varphi\,(J_{14}^2 + J_{24}^2 + J_{34}^2 + J_{45}^2 + J_{46}^2 + J_{47}^2) \\
&- 4m^2\sin^2\varphi\,(J_{01}^2 + J_{02}^2 + J_{03}^2 + J_{05}^2 + J_{06}^2 + J_{07}^2) \\
&-2m^2\sin 2\varphi\,(\{J^{01}, J^{14}\}+\{J^{02}, J^{24}\}+\{J^{03}, J^{34}\}+\{J^{05}, J^{45}\}+\{J^{06}, J^{46}\}+\{J^{07}, J^{47}\})
\end{aligned}$$

by using

$$P^0 = 2m\,\cos\varphi\,;\ P^4 = 2m\,\sin\varphi$$

We note that, in the last term of the second Casimir expression, all J^{AB} terms appear except one: J^{04}, which represents time rotation.

We also note the presence of the usual spin through J^{12}, J^{31} and J^{23}, which generates $so(3)$... but there are other possibilities! For example

$$J^{15}, J^{16}, J^{56}$$

This should come as no surprise: the traditional spin is associated with spatial rotations. Here, there are multiple combinations of 3D rotations between one real coordinate and two imaginary coordinates (nine in total), between two real coordinates and one imaginary coordinate (also nine), and a single rotation between the imaginary coordinates... That makes a total of twenty different spins!

The other parts of the Casimir operator (those involving φ, the angle of rotation that allows positive energies to be transferred to negative energies) concern the boosts or "rotations" between a time coordinate (real or imaginary) and two spatial coordinates (imaginary or real).

Chapter 2

The Dirac equation

The 17th century was the scene of a heated debate between the English physicist Isaac Newton and his Dutch counterpart, Christiaan Huygens. The former asserted that light is corpuscular in nature, while the latter maintained that it is a wave.

In 1909, Albert Einstein[1] brought everyone to a consensus by proposing a dual nature: light behaves both like a wave and a particle (known as a photon). This dual nature is an intrinsic characteristic: light is both at the same time, and it is only through certain experiments that one of these two aspects becomes apparent to the researcher conducting the experiment.

To illustrate this duality, here is an equation by the German physicist Max Planck[2]:

$$E = h\,\nu \tag{2.1}$$

The first term represents the particle aspect, the energy E. In the second term, we see the wave aspect represented by the frequency ν. To link these two aspects, we use what is known as Planck's constant: h.

The following version of Planck's equation is often preferred:

$$E = \hbar\,\omega \tag{2.2}$$

The new frequency ω is the angular frequency

$$\omega = 2\pi\,\nu \tag{2.3}$$

and so we see the Dirac constant appear:

$$\hbar = \frac{h}{2\pi} = 1,05\,10^{-34}\,J.s \tag{2.4}$$

[1] Einstein, A. (1909). Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift,* 10, 817-826

[2] Planck, M. (1901). Ueber das Gesetz der Energieverteilung im Normalspectrum. *Annalen der Physik*, 309 (4), 553-563

The French physicist and mathematician Louis de Broglie[3] takes it even further! In his 1924 dissertation, he argues that all matter has a wave-like nature. This is expressed by a second equation:

$$p = \frac{h}{\lambda} \Leftrightarrow p = \hbar\, k \tag{2.5}$$

(and, in three-dimensional space, $\overrightarrow{p} = \hbar\, \overrightarrow{k}$). This time, the particle world is characterized by the momentum p, while the wave world is represented by the wavelength λ or the angular wavenumber k. Once again, the two worlds are linked through Planck's constant or Dirac's constant.

These are extremely interesting equations, but they do not allow us to determine the properties of a quantum object.

It wasn't long before, a year later—in 1925—the Austrian physicist Erwin Schrödinger[4] developed the equation that bears his name, which would go on to become the foundation of non-relativistic quantum mechanics. Here's what it looks like:

$$i\hbar \frac{\partial \psi(t, \overrightarrow{r})}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) \psi(t, \overrightarrow{r}) + V(t, \overrightarrow{r})\, \psi(t, \overrightarrow{r}) \tag{2.6}$$

The real function $V(t, \overrightarrow{r})$ represents the potential energy, while the complex function $\psi(t, \overrightarrow{r})$ is the wavefunction. It is the latter that contains all the information, including the probability of finding the quantum object in a particular state, as given by the square of its modulus.

There is no strict proof of this equation as such, only a body of evidence that supports its form. Despite this, it is an incredibly effective equation that agrees remarkably well with experimental results!

How did Schrödinger come up with this equation?

It was a series of insights that helped him. We'll mention two of them.

The first is the time evolution of the quantum object. Starting from an initial time $t = 0$, we must go from a state $\psi(0, \overrightarrow{r})$ to a state $\psi(t, \overrightarrow{r})$. The conservation of probabilities—as we know from Wigner's theorem —requires that this be accomplished via the action of a unitary operator $U(t)$.

If we restrict ourselves to infinitesimal time translations, we therefore have

$$U(dt) = U(0) + dt\, \frac{dU}{dt}|_{t=0} = I + dt\, \frac{dU}{dt}|_{t=0}$$

So, we obtain

[3] de Broglie, L. (1924). *Recherches sur la théorie des Quanta.* PhD Thesis, Masson, Paris. https://doi.org/10.1051/anphys/192510030022

[4] For a summary in English, see Schrödinger, E. (1926). An Undulatory Theory of the Mechanics of Atoms and Molecules. Phys.Rev. 28, 1049-1070. https://doi.org/10.1103/PhysRev.28.1049

$$\psi(dt) = U(dt)\,\psi(0) = \psi(0) + dt\,\frac{dU}{dt}|_{t=0}\psi(0)$$

or

$$\frac{\psi(dt) - \psi(0)}{dt} = \frac{dU}{dt}|_{t=0}\psi(0)$$

Of course, the left-hand side is none other than the time derivative of the wavefunction when we consider the limit as $dt \to 0$. This ensures equality among operators

$$\frac{d}{dt} = \frac{dU}{dt}|_{t=0}$$

on the wavefunction.

The catch: this operator is not hermitian. In fact, when expressed this way, it leads to

$$U^\dagger U = I + (\frac{dU^\dagger}{dt}|_{t=0} + \frac{dU}{dt}|_{t=0})dt = I$$

The only way to restore a Hermitian operator rather than an anti-Hermitian one, as is currently the case, is to multiply the time derivative by the imaginary unit i. Thus, the time evolution of the wavefunction will be characterized by[5]

$$i\,\hbar\,\frac{\partial}{\partial t} \tag{2.7}$$

This explains the left-hand side of Schrödinger's equation.

The right-hand side is explained, in particular, by the second clue.

Keep in mind that wave-particle duality is the foundation of quantum mechanics. In particular, a specific wave solution is

$$\psi(t, \overrightarrow{r}) = A\,\exp(-i\,(\omega t - \overrightarrow{k}.\overrightarrow{r})) \tag{2.8}$$

All that remains is to transform it using Planck's equation and de Broglie's equation:

$$\psi(t, \overrightarrow{r}) = A\,\exp(-\frac{i}{\hbar}\,(Et - \overrightarrow{p}.\overrightarrow{r})) \tag{2.9}$$

Schrödinger then requires that this function be a solution to his equation.

What happens to the left-hand side? This:

$$i\hbar\,\frac{\partial\psi(t, \overrightarrow{r})}{\partial t} = E\,\psi(t, \overrightarrow{r}) \tag{2.10}$$

[5]The presence of the Dirac constant is necessary to ensure that the units are correct; without it, the first term of (2.6) would have the dimension of the inverse of time, while the second term would have the dimension of energy.

The right-hand side of (2.6) must therefore equal the total energy of the system. Now, the total energy is the sum of the kinetic energy and the potential energy. So $V(t, \overrightarrow{r})$ is indeed present!

As for the kinetic energy, we observe that

$$\triangle\psi(t, \overrightarrow{r}) = (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})\,\psi(t, \overrightarrow{r}) = -\frac{\overrightarrow{p}^2}{\hbar^2}\,\psi(t, \overrightarrow{r}) \tag{2.11}$$

We then need to multiply by $-\frac{\hbar^2}{2m}$ to obtain the equivalent of the classical kinetic energy $\frac{\overrightarrow{p}^2}{2m}$.

And that's how we can understand Schrödinger's equation in hindsight.

Schrödinger's equation remains, even today, a fascinating field of study... but it has two fatal flaws:

- It does not take spin into account (which, admittedly, was discovered later)
- It is not relativistic

On this last point, it comes as no surprise: Schrödinger's equation highlights an imbalance between time (first derivative) and spatial coordinates (second derivatives). Naturally, this didn't bode well for relativity!

It's the boosts that are going to throw a wrench in the works! After all, they bring time and space into alignment.

Remember that, for example, and limiting ourselves to a single spatial coordinate, a boost transformation is characterized by

$$\begin{cases} ct\prime = (\cosh\varphi)\; ct + (\sinh\varphi)\; x \\ x\prime = (\sinh\varphi)\; ct + (\cosh\varphi)\; x \end{cases}$$

In terms of derivatives, this means

$$\frac{\partial}{\partial x} = \frac{\partial x\prime}{\partial x}\frac{\partial}{\partial x\prime} + \frac{\partial t\prime}{\partial x}\frac{\partial}{\partial t\prime} = \cosh\varphi\,\frac{\partial}{\partial x\prime} + \frac{1}{c}\sinh\varphi\,\frac{\partial}{\partial t\prime}$$

$$\frac{\partial}{\partial t} = \frac{\partial x\prime}{\partial t}\frac{\partial}{\partial x\prime} + \frac{\partial t\prime}{\partial t}\frac{\partial}{\partial t\prime} = \cosh\varphi\,\frac{\partial}{\partial t\prime} + \frac{1}{c}\sinh\varphi\,\frac{\partial}{\partial x\prime}$$

So what happens to Schrödinger's equation?

Well, it's clear that it changes significantly, notably with cross-derivatives like $\frac{\partial^2}{\partial x\prime\partial t\prime}$!

If we want to study a relativistic particle—say, a relativistic electron—we need to quickly set aside Schrödinger's equation and turn to other theories...

Yes, but which ones?

Two equations will stand out from the rest.

2.1 The Klein-Gordon equation

Remember the first Casimir operator of the Poincaré group

$$C1 = P^{\mu}P_{\mu} \tag{2.12}$$

In the context of timelike 4-vectors, it is equal to m^2.

This is a purely relativistic result to which we must add the quantifications we just mentioned[6]:

$$P^0 = E \leftrightarrow i\frac{\partial}{\partial t}\ ;\ P_j = p_j \leftrightarrow -i\frac{\partial}{\partial x^j} \tag{2.13}$$

When you combine these two results (which is as simple as it is brilliant), you get the equation proposed by the Swedish physicist Oskar Klein[7] and his German counterpart Walter Gordon[8]:

$$(\frac{\partial^2}{\partial t^2} - (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}))\,\psi(t,\overrightarrow{r}) + m^2\,\psi(t,\overrightarrow{r}) = 0 \tag{2.14}$$

Using the definition of the d'Alembertian

$$\Box \equiv \frac{\partial^2}{\partial t^2} - (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}) \tag{2.15}$$

we then get

$$(\Box + m^2)\,\psi(t,\overrightarrow{r}) = 0 \tag{2.16}$$

Based on the relativistic conventions mentioned in Eqs. (130)-(131), the Klein-Gordon equation can be written simply as[9]

$$(\partial^{\mu}\partial_{\mu} + m^2)\,\psi(t,\overrightarrow{r}) = 0 \tag{2.17}$$

The Klein-Gordon equation begins to meet the requirements: it is invariant under Lorentz transformations, which is, of course, obvious given its formulation.

On the other hand:

- Still no sign of the spin
- A disaster in terms of probability

[6] From now on, to avoid unnecessary complexity, we set the Dirac constant and the speed of light to 1. In other words, we are working in the system of natural units.

[7] Oskar, K. (1926). Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik*, 37, 895-906. https://doi.org/10.1007/BF01397481

[8] Gordon, W. (1926). Der Comptoneffekt nach der Schrödingerschen Theorie. *Zeitschrift für Physik*, 40, 117-133. https://doi.org/10.1007/BF01390840

[9] Fun fact, we could also develop the "imaginary masses" version based on Wigner's third little group and propose $(\partial^{\mu}\partial_{\mu} - m^2)\,\psi(t,\overrightarrow{r}) = 0$

As for the first problem, it's no big deal: we'll reserve the Klein-Gordon equation for particles with zero spin. Well, we'll still have to come up with something else for the electron, but hang in there—we'll get there!

As for the second issue, it's more serious.

Indeed, let's consider the Klein-Gordon equation and multiply it on the left by ψ^*. Let's also take the complex conjugate of this equation and multiply the resulting equation on the left by ψ.

Let's subtract the two equations we obtained. We then have

$$\psi^* \frac{\partial^2 \psi}{\partial t^2} - \psi \frac{\partial^2 \psi^*}{\partial t^2} = \psi^* \triangle \psi - \psi \triangle \psi^* \tag{2.18}$$

This last equation can also be written as

$$\frac{\partial}{\partial t}(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}) = \overrightarrow{\nabla}(\psi^* \overrightarrow{\nabla} \psi - \psi \overrightarrow{\nabla} \psi^*) \tag{2.19}$$

It can be identified with a continuity equation

$$\frac{\partial \rho}{\partial t} + \overrightarrow{\nabla} . \overrightarrow{j} = 0 \tag{2.20}$$

like the one encountered in electromagnetism. The probability density ρ and the probability current density $\overrightarrow{j}$ are then defined by

$$\rho \equiv \frac{i}{2m}(\psi^* \frac{\partial \psi}{\partial t} - \psi \frac{\partial \psi^*}{\partial t}), \ \overrightarrow{j} \equiv -\frac{i}{2m}(\psi^* \overrightarrow{\nabla} \psi - \psi \overrightarrow{\nabla} \psi^*) \tag{2.21}$$

Now let's look at the "plane wave" solutions (2.9). It is then easy to see that the probability density is equal to

$$\rho = \frac{E}{m} |A|^2 \tag{2.22}$$

What's the problem, you might ask?

The problem is that the first Casimir operator in the Poincaré group implies Einstein's relation

$$E^2 = \overrightarrow{p}^2 + m^2 \tag{2.23}$$

which, in turn, leads to the presence of negative energies since

$$E = \pm \sqrt{\overrightarrow{p}^2 + m^2} \tag{2.24}$$

If we assume that mass m is positive (and that's what 99.99% of researchers do...), we end up with a negative probability! That's quite a problem!

Unless... we accept the existence of negative masses! In that case, everything falls into place!

We'll come back to this.

2.2 The Dirac equation

Attempts to reconcile relativity and quantum mechanics have thus far failed. One man would change all that: the British mathematician and physicist Paul Dirac.

In 1928, he wrote a landmark article [10] that provided the solution everyone had been waiting for.

We are providing an excerpt here, given the work's paramount importance.

Here's the main idea:

Quantum Theory of Electron. 613

§ 2. *The Hamiltonian for No Field.*

Our problem is to obtain a wave equation of the form (2) which shall be invariant under a Lorentz transformation and shall be equivalent to (1) in the limit of large quantum numbers. We shall consider first the case of no field, when equation (1) reduces to

$$(-p_0^2 + \mathbf{p}^2 + m^2c^2)\,\psi = 0 \tag{3}$$

if one puts

$$p_0 = \frac{\mathrm{W}}{c} = ih\frac{\partial}{c\,\partial t}.$$

The symmetry between p_0 and p_1, p_2, p_3 required by relativity shows that, since the Hamiltonian we want is linear in p_0, it must also be linear in p_1, p_2 and p_3. Our wave equation is therefore of the form

$$(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\,\psi = 0, \tag{4}$$

where for the present all that is known about the dynamical variables or operators α_1, α_2, α_3, β is that they are independent of p_0, p_1, p_2, p_3, *i.e.*, that they commute with t, x_1, x_2, x_3. Since we are considering the case of a particle moving in empty space, so that all points in space are equivalent, we should expect the Hamiltonian not to involve t, x_1, x_2, x_3. This means that α_1, α_2, α_3, β are independent of t, x_1, x_2, x_3, *i.e.*, that they commute with p_0, p_1, p_2, p_3. We are therefore obliged to have other dynamical variables besides the co-ordinates and momenta of the electron, in order that α_1, α_2, α_3, β may be functions of them. The wave function ψ must then involve more variables than merely x_1, x_2, x_3, t.

Equation (4) leads to

$$\begin{aligned} 0 &= (-p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)(p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta)\,\psi \\ &= [-p_0^2 + \Sigma\alpha_1^2 p_1^2 + \Sigma(\alpha_1\alpha_2 + \alpha_2\alpha_1)p_1p_2 + \beta^2 + \Sigma(\alpha_1\beta + \beta\alpha_1)p_1]\,\psi, \end{aligned} \tag{5}$$

Note that Dirac's starting point[11] is none other than Eq. (3), that is, Einstein's equation (2.23). He also considers quantification by taking $E = i\frac{\partial}{\partial t}$.

[10] Dirac, P.A.M. (1928). The Quantum Theory of the Electron. *Proceedings of the Royal Society in London*, Series A, 117, 610-624. https://doi.org/10.1098/rspa.1928.0023

[11] Its W notation is in fact the energy E

Finally, his proposal is as simple as it is brilliant: rather than harmonizing everything by considering second derivatives of the spacetime coordinates, as Klein and Gordon did, let's balance the whole thing by considering only first derivatives, for both time and space. The price to pay? The appearance of the mysterious $\alpha_1, \alpha_2, \alpha_3$ and β in (4), whose nature will be specified very soon.

Finally, note the two possible signs that appear in Eq. (5)...

But let us proceed as Dirac did, while adapting his reasoning to relativistic conventions.

What we must impose is an equation of the form

$$i\,\frac{\partial \psi}{\partial t} = \overrightarrow{\alpha}.\overrightarrow{p}\;\psi + m\beta\,\psi \tag{2.25}$$

In other words, we ask for

$$p_0\,\psi = -\alpha^j p_j\,\psi + m\beta\,\psi \tag{2.26}$$

We therefore must require

$$\begin{aligned} &-p_0^2 + p_1^2 + p_2^2 + p_3^2 \\ &\quad = (-p_0 + \alpha^1 p_1 + \alpha^2 p_2 + \alpha^3 p_3 - m\beta)\,(p_0 + \alpha^1 p_1 + \alpha^2 p_2 + \alpha^3 p_3 - m\beta) \end{aligned}$$

In the left-hand side, you recognize the Einstein relation and in the right-hand side, its factorization.

All you need to do is distribute (bearing in mind that the momenta commute, but not the α^j and β, which remain as mysterious as ever).

The following constraints are then revealed

$$(\alpha^1)^2 = (\alpha^2)^2 = (\alpha^3)^2 = I\;;\;\beta^2 = I \tag{2.27}$$

$$\alpha^j\alpha^k + \alpha^k\alpha^j \equiv \{\alpha^j, \alpha^k\} = 0,\; j \neq k \tag{2.28}$$

$$\{\alpha^j, \beta\} = 0,\; \forall j \tag{2.29}$$

In the last two relations, the anticommutator appears, defined, as we recall, by

$$\{A, B\} \equiv AB + BA \tag{2.30}$$

By setting

$$\beta = \alpha^0 \tag{2.31}$$

we can elegantly summarize these relations as

$$\{\alpha^\mu, \alpha^\nu\} = 2\delta^{\mu\nu} \tag{2.32}$$

It's immediately clear that these mysterious objects can't be numbers, because numbers aren't anti-commutative! Matrices, on the other hand, can be...

Dirac is therefore looking for four matrices whose squares are the identity matrix and which anticommute with one another.

These matrices are related to a well-known algebraic structure. Indeed, the Clifford algebras $Cl(p,q)$ are generated by $(p+q)$ matrices that anticommute with one another, such that p of them have a square equal to the identity matrix, while q of them have a square equal to minus the identity matrix.

Here, we are clearly dealing with $Cl(4)$.

How can we actually construct these matrices? Like this, for example

$$\alpha^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \alpha^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \alpha^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{2.33}$$

$$\alpha^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.34}$$

This is known as the Dirac formulation.

The Dirac equation thus takes the form

$$\begin{pmatrix} p_0 - m & 0 & p_3 & p_1 - ip_2 \\ 0 & p_0 - m & p_1 + ip_2 & -p_3 \\ p_3 & p_1 - ip_2 & p_0 + m & 0 \\ p_1 + ip_2 & -p_3 & 0 & p_0 + m \end{pmatrix} \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix} = 0 \tag{2.35}$$

These days, few people still use this formulation, unless they want to highlight the solutions to the Dirac equation (and we will be no exception).

Instead, they prefer to use a covariant formulation, which is completely equivalent. This latter approach amounts to factoring the Klein-Gordon equation as follows[12]:

[12] Note: There are two possible signs for mass!

$$(i\gamma^\nu \partial_\nu + m)(i\gamma^\mu \partial_\mu - m)\,\psi = 0 \tag{2.36}$$

The version that will be chosen by all Standard Model physicists is:

$$\boxed{(i\gamma^\mu \partial_\mu - m)\,\psi = 0} \tag{2.37}$$

The covariant formulation and the one introduced by Dirac are equivalent simply by multiplying one by γ^0.

Thus, we have

$$\gamma^0 = \alpha^0\ ,\ \gamma^1 = \alpha^0\alpha^1\ ,\ \gamma^2 = \alpha^0\alpha^2\ ,\ \gamma^3 = \alpha^0\alpha^3 \tag{2.38}$$

If the matrices γ continue to anticommute, their squares become equal to $-I$, except, of course, for the matrix γ^0. They thus generate the Clifford algebra $Cl(1,3)$.

There are several ways to construct these matrices, but we will use two of them:

- Dirac's realization:

$$\gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},\ \gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix},\ \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.39}$$

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{2.40}$$

- Chiral realization:

$$\gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},\ \gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix},\ \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.41}$$

$$\gamma^0 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.42}$$

Does Dirac's proposal resolve all the issues noted with the Klein-Gordon equation? Let's take a look!

2.2.1 The probabilistic interpretation associated with the Dirac equation

The Dirac equation is therefore

$$i\,\frac{\partial \psi}{\partial t} = \overrightarrow{\alpha}.\overrightarrow{p}\,\psi + m\alpha^0\,\psi \tag{2.43}$$

If we multiply this equation by $\psi^\dagger$ on the left, we get

$$i\,\psi^\dagger\frac{\partial \psi}{\partial t} = \psi^\dagger\overrightarrow{\alpha}.\overrightarrow{p}\,\psi + m\psi^\dagger\alpha^0\,\psi \tag{2.44}$$

If we consider the Hermitian conjugate equation of (2.43):

$$-i\,\frac{\partial \psi^\dagger}{\partial t} = -\psi^\dagger\overrightarrow{\alpha}.\overrightarrow{p} + m\,\psi^\dagger\alpha^0$$

and if we multiply it by ψ on the right, we get

$$-i\,\frac{\partial \psi^\dagger}{\partial t}\psi = -\psi^\dagger\overrightarrow{\alpha}.\overrightarrow{p}\psi + m\,\psi^\dagger\alpha^0\psi \tag{2.45}$$

Subtracting the two equations (2.44) and (2.45) from one another, we get

$$i\,\frac{\partial}{\partial t}(\psi^\dagger\psi) = -i\,\overrightarrow{\nabla}(\psi^\dagger\overrightarrow{\alpha}\psi) \tag{2.46}$$

This sets the probability density of occurrence and the probability current density to

$$\rho = \psi^\dagger\psi = |\psi^1|^2 + |\psi^2|^2 + |\psi^3|^2 + |\psi^4|^2 \tag{2.47}$$

$$\overrightarrow{j} = \psi^\dagger\overrightarrow{\alpha}\,\psi \tag{2.48}$$

We can therefore note with relief that the probability of presence has returned to positive values.

First problem solved!

Let's move on to the next one.

The task is to show that the Dirac equation accounts for spin. Since it is matrix-based, this looks promising! But these matrices must still be associated with a spin of $\frac{1}{2}$, since that is the value of the electron's spin.

To convince ourselves, we need to examine the relativistic invariance of the Dirac equation.

2.2.2 The relativistic invariance of the Dirac equation

To state the relativistic invariance of the Dirac equation is to show that

$$(\gamma^\mu p_\mu - m)\,\psi(x^\mu) = 0 \Leftrightarrow (\gamma^\mu p\prime_\mu - m)\,\psi\prime(x\prime^\mu) = 0 \tag{2.49}$$

The $x\prime^\mu$, $p\prime_\mu$ are determined by the Lorentz transformations, and we still need to determine what $\psi\prime(x\prime^\mu)$ is.

Let's first consider the rotation (1.87) as an initial example:

$$\begin{cases} x\prime^0 = x^0 \\ x\prime^1 = (\cos\theta)\,x^1 - (\sin\theta)\,x^2 \\ x\prime^2 = (\sin\theta)\,x^1 + (\cos\theta)\,x^2 \\ x\prime^3 = x^3 \end{cases} \tag{2.50}$$

These relations obviously have a direct impact on the derivatives, and we have

$$\begin{cases} p\prime_0 = p_0 \\ p\prime_1 = (\cos\theta)\,p_1 - (\sin\theta)\,p_2 \\ p\prime_2 = (\sin\theta)\,p_1 + (\cos\theta)\,p_2 \\ p\prime_3 = p_3 \end{cases} \tag{2.51}$$

While setting

$$\psi\prime = S_z\psi \tag{2.52}$$

the Dirac equation, after rotation, is then written as

$$(\gamma^0 p_0 + (\gamma^1\cos\theta + \gamma^2\sin\theta)p_1 + (-\gamma^1\sin\theta + \gamma^2\cos\theta)p_2 + \gamma^3 p_3 - m)\,S_z\psi = 0 \tag{2.53}$$

By applying S_z^{-1} from the left, we obtain the original Dirac equation if and only if

$$S_z^{-1}\gamma^0 S_z = \gamma^0\,,\; S_z^{-1}\gamma^3 S_z = \gamma^3 \tag{2.54}$$

$$S_z^{-1}(\gamma^1\cos\theta+\gamma^2\sin\theta)S_z=\gamma^1\ ,\ S_z^{-1}(-\gamma^1\sin\theta+\gamma^2\cos\theta)S_z=\gamma^2 \qquad (2.55)$$

Let's think a while within $Cl(1,3)$, since that is the algebra in question. It consists of 16 elements[13]:

$$I$$

$$\gamma^0,\gamma^1,\gamma^2,\gamma^3$$

$$\gamma^0\gamma^1,\gamma^0\gamma^2,\gamma^0\gamma^3,\gamma^1\gamma^2,\gamma^3\gamma^1,\gamma^2\gamma^3$$

$$\gamma^0\gamma^1\gamma^2,\gamma^0\gamma^3\gamma^1,\gamma^0\gamma^2\gamma^3,\gamma^1\gamma^2\gamma^3$$

$$\gamma^0\gamma^1\gamma^2\gamma^3$$

The first two constraints (2.54) indicate a commutation with γ^0 and γ^3. Only four elements satisfy these requirements:

$$I,\gamma^1\gamma^2,\gamma^0\gamma^3\gamma^1,\gamma^0\gamma^2\gamma^3$$

The matrix S_z is therefore, a priori, a linear combination of these elements: $S_z=c1+c2\,\gamma^1\gamma^2+c3\,\gamma^0\gamma^3\gamma^1+c4\,\gamma^0\gamma^2\gamma^3$.

It quickly becomes clear that $c3=c4=0$. The remaining parameters are fixed according to

$$\begin{cases} c1=c1\ \cos\theta+c2\ \sin\theta \\ c2=c1\ \sin\theta-c2\ \cos\theta \end{cases}$$

From there, it is easy to see, using a few well-known trigonometric formulas, that the matrix S_z satisfying these constraints is

$$S_z=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}\gamma^1\gamma^2 \qquad (2.56)$$

Similarly, the rotations (1.86) result in

$$S_x=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}\gamma^2\gamma^3\ ,\ S_y=\cos\frac{\theta}{2}+\sin\frac{\theta}{2}\gamma^3\gamma^1 \qquad (2.57)$$

Let's now consider the first boost of (1.88) as a second example:

[13] See Chapter 3 if it sounds mysterious for you.

$$\begin{cases} x\prime^0 = (\cosh\varphi)\, x^0 + (\sinh\varphi)\, x^1 \\ x\prime^1 = (\sinh\varphi)\, x^0 + (\cosh\varphi)\, x^1 \\ x\prime^2 = x^2 \\ x\prime^3 = x^3 \end{cases} \tag{2.58}$$

These relations also affect the momenta:

$$\begin{cases} p\prime_0 = (\cosh\varphi)\, p_0 - (\sinh\varphi)\, p_1 \\ p\prime_1 = -(\sinh\varphi)\, p_0 + (\cosh\varphi)\, p_1 \\ p\prime_2 = p_2 \\ p\prime_3 = p_3 \end{cases} \tag{2.59}$$

We set

$$\psi\prime = T_x \psi \tag{2.60}$$

After this boost transformation, the Dirac equation becomes

$$((\gamma^0\cosh\varphi - \gamma^1\sinh\varphi)p_0 + (-\gamma^0\sinh\varphi + \gamma^1\cosh\varphi)p_1 + \gamma^2 p_2 + \gamma^3 p_3 - m)\, T_x\psi = 0 \tag{2.61}$$

Invariance then implies the following constraints:

$$T_x^{-1}\gamma^2\, T_x = \gamma^2\ ,\ T_x^{-1}\gamma^3\, T_x = \gamma^3 \tag{2.62}$$

$$T_x^{-1}(\gamma^0\cosh\varphi - \gamma^1\sinh\varphi)T_x = \gamma^0\ ,\ T_x^{-1}(-\gamma^0\sinh\varphi + \gamma^1\cosh\varphi)T_x = \gamma^1 \tag{2.63}$$

We again fix a set of four matrices that satisfy the first two conditions, namely, commute with γ^2 and γ^3. These matrices are

$$I, \gamma^0\gamma^1, \gamma^0\gamma^2\gamma^3, \gamma^1\gamma^2\gamma^3$$

Starting from

$$c1 + c2\,\gamma^0\gamma^1 + c3\,\gamma^0\gamma^2\gamma^3 + c4\,\gamma^1\gamma^2\gamma^3$$

we can see that $c3$ and $c4$ must vanish while $c1$ and $c2$ are constrained through

$$\begin{cases} c1 = c1\ \cosh\varphi - c2\ \sinh\varphi \\ c2 = c1\ \sinh\varphi - c2\ \cosh\varphi \end{cases}$$

The well-known relations involving hyperbolic functions then yield to

$$T_x = \cosh\frac{\varphi}{2} + \sinh\frac{\varphi}{2}\gamma^0\gamma^1 \tag{2.64}$$

Similarly, the other boosts lead to

$$T_y = \cosh\frac{\varphi}{2} + \sinh\frac{\varphi}{2}\gamma^0\gamma^2\,,\; T_z = \cosh\frac{\varphi}{2} + \sinh\frac{\varphi}{2}\gamma^0\gamma^3 \tag{2.65}$$

These six transformations of the wavefunction have an algebraic counterpart that is obtained by considering infinitesimal angles. The associated matrices are then

$$J_x = \frac{1}{2}\gamma^2\gamma^3\,,\; J_y = \frac{1}{2}\gamma^3\gamma^1\,,\; J_z = \frac{1}{2}\gamma^1\gamma^2 \tag{2.66}$$

$$K_x = \frac{1}{2}\gamma^0\gamma^1\,,\; K_y = \frac{1}{2}\gamma^0\gamma^2\,,\; K_z = \frac{1}{2}\gamma^0\gamma^3 \tag{2.67}$$

Using the chiral representation, we can see that

$$\overrightarrow{J} = -\frac{i}{2}\begin{pmatrix} \overrightarrow{\sigma} & 0 \\ 0 & \overrightarrow{\sigma} \end{pmatrix}\,,\; \overrightarrow{K} = \frac{1}{2}\begin{pmatrix} -\overrightarrow{\sigma} & 0 \\ 0 & \overrightarrow{\sigma} \end{pmatrix} \tag{2.68}$$

In other words, it is the representation $D(\frac{1}{2},0)\oplus D(0,\frac{1}{2})$ of the Lorentz algebra—which we discussed earlier—that underlies the Dirac equation!

Combining left- and right-handed representations, it is described as enantiomorphic

The conclusions are clear:

- The Dirac equation is relativistic
- The Dirac equation describes two particles with spin $\mathbf{\frac{1}{2}}$

While the first of these conclusions was entirely expected, the second, on the other hand, came as a complete surprise! The Dirac equation thus made room not only for an electron but also for another particle with spin $\mathbf{\frac{1}{2}}$ ... Another particle, because the representations $D(\frac{1}{2},0)$ and $D(0,\frac{1}{2})$ are indeed different!

To learn more about this unexpected mystery guest, we need to turn to solving the Dirac equation.

But before that, let's discuss discrete symmetry operators.

2.2.3 How about discussing discrete symmetry operators?

So far, little has been said about them. Given their discrete nature, they don't really come into play in group theory, except for linking the connected components of a group.

But they will play an increasingly important role.

Let's list them.

2.2.3.1 The parity operator P

We have already mentioned it in the context of Lorentz transformations, of which it is a part.

Its effect is to change the signs of each of the three spatial coordinates and thus of their derivatives, while leaving time invariant. The action of this operator on the Dirac equation is thus

$$(\gamma^0 p_0 - \gamma^1 p_1 - \gamma^2 p_2 - \gamma^3 p_3 - m)\,\psi\prime = 0 \tag{2.69}$$

Let's recall Wigner's theorem: every symmetry operator in quantum mechanics must be either unitary or antiunitary.

Unitarity is a requirement for continuous transformations... But what about discrete transformations? We have a choice!

Well... physicists working within the Standard Model believe we don't really have much of a choice: P must be a unitary operator. Roughly speaking, their argument boils down to this: energy is associated with the operator $i\frac{\partial}{\partial t}$. Parity has no effect on $\frac{\partial}{\partial t}$. If we wish to consider only positive energies (and this is what Standard Model physicists want above all else), we must not change the sign of i, which means that P must necessarily be unitary.

But if we allow negative energies to exist, then P may be antiunitary[14]

Let's examine both cases.

If the operator P is considered to be unitary, the Dirac equation (2.69), after applying the parity operator, must be equivalent to

$$P_U(\gamma^0 p_0 + \gamma^1 p_1 + \gamma^2 p_2 + \gamma^3 p_3 - m)\,P_U^{-1} P_U \psi = 0 \tag{2.70}$$

The following identifications come directly from this:

$$\psi\prime = P_U \psi$$

and

$$P_U\gamma^0 P_U^{-1} = \gamma^0\,,\; P_U\gamma^1 P_U^{-1} = -\gamma^1\,,\; P_U\gamma^2 P_U^{-1} = -\gamma^2\,,\; P_U\gamma^3 P_U^{-1} = -\gamma^3 \tag{2.71}$$

It follows easily that P_U must commute with γ^0 and anticommute with γ^1, γ^2 and γ^3. Once again, as with any transformation, we search among all sixteen matrices of $Cl(1,3)$ that satisfy these constraints. We find

$$\boxed{P_U = \gamma^0} \tag{2.72}$$

[14] Debergh, N., Petit, J.-P. and D'Agostini, G. (2018). On evidence for negative energies and masses in the Dirac equation through a unitary time-reversal operator. *Journal of Physics Communications*, 2, 115012. https://doi.org/10.1088/2399-6528/aaedcc

If the operator P is considered antiunitary, however, it is written as a matrix followed by the complex conjugate $()^*$ and the equation (2.69) must be equivalent to

$$P_{AU}(-\gamma^{0^*}p_0 - \gamma^{1^*}p_1 - \gamma^{2^*}p_2 - \gamma^{3^*}p_3 - m)\,P_{AU}^{-1}P_{AU}\psi^* = 0 \tag{2.73}$$

since the p_μ are defined as the imaginary number i that multiplies the partial derivatives with respect to the spacetime coordinates.

Given the known realizations of the matrices γ^μ, this last equation reduces to

$$P_{AU}(-\gamma^0 p_0 - \gamma^1 p_1 + \gamma^2 p_2 - \gamma^3 p_3 - m)\,P_{AU}^{-1}P_{AU}\psi^* = 0 \tag{2.74}$$

So we have to impose

$$\psi\prime = P_{AU}\psi^*$$

and

$$P_{AU}\gamma^0 P_{AU}^{-1} = -\gamma^0\,,\; P_{AU}\gamma^1 P_{AU}^{-1} = \gamma^1\,,\; P_{AU}\gamma^2 P_{AU}^{-1} = -\gamma^2\,,\; P_{AU}\gamma^3 P_{AU}^{-1} = \gamma^3 \tag{2.75}$$

It follows that, from a matrix perspective, P_{AU} must commute with γ^1 and γ^3 and anticommute with γ^0 and γ^2. Consequently, we obtain

$$\boxed{P_{AU} = \gamma^0\gamma^2()^*} \tag{2.76}$$

Note this difference: $P_U^2 = I$, whereas $P_{AU}^2 = -I$.

2.2.3.2 The time-reversal operator T

Its effect is to change the signs of the single time coordinate and, consequently, of its derivative. The resulting Dirac equation is therefore

$$(-\gamma^0 p_0 + \gamma^1 p_1 + \gamma^2 p_2 + \gamma^3 p_3 - m)\,\psi\prime = 0 \tag{2.77}$$

What is the consensus among Standard Model physicists?

It once again hinges on the sign of the energy $i\frac{\partial}{\partial t}$. Here the time derivative changes sign! So, to avoid negative energies, the imaginary part i must also change sign! And this can only be achieved through a complex conjugation—and thus, an antiunitary operator. In contrast, negative energies are associated with a unitary T operator.

If the operator T is considered to be unitary, the equation (2.77) must be equivalent to

$$T_U(\gamma^0 p_0 + \gamma^1 p_1 + \gamma^2 p_2 + \gamma^3 p_3 - m)\,T_U^{-1}T_U\psi = 0 \tag{2.78}$$

The following identifications are a direct consequence of this:

$$\psi\prime = T_U\psi$$

and

$$T_U\gamma^0T_U^{-1} = -\gamma^0\,,\; T_U\gamma^1T_U^{-1} = \gamma^1\,,\; T_U\gamma^2T_U^{-1} = \gamma^2\,,\; T_U\gamma^3T_U^{-1} = \gamma^3 \tag{2.79}$$

It follows easily that T_U must anticommute with γ^0 and commute with γ^1, γ^2 and γ^3. Consequently, we have

$$\boxed{T_U = \gamma^1\gamma^2\gamma^3} \tag{2.80}$$

If the operator T is considered to be antiunitary (which, as we recall, is the consensus among Standard Model physicists), the equation (2.77) must be equivalent to

$$T_{AU}(-\gamma^{0^*}p_0 - \gamma^{1^*}p_1 - \gamma^{2^*}p_2 - \gamma^{3^*}p_3 - m)\,T_{AU}^{-1}T_{AU}\psi^* = 0 \tag{2.81}$$

or

$$T_{AU}(-\gamma^0p_0 - \gamma^1p_1 + \gamma^2p_2 - \gamma^3p_3 - m)\,T_{AU}^{-1}T_{AU}\psi^* = 0 \tag{2.82}$$

We thus ask for

$$\psi\prime = T_{AU}\psi^*$$

and

$$T_{AU}\gamma^0T_{AU}^{-1} = \gamma^0\,,\; T_{AU}\gamma^1T_{AU}^{-1} = -\gamma^1\,,\; T_{AU}\gamma^2T_{AU}^{-1} = \gamma^2\,,\; T_{AU}\gamma^3T_{AU}^{-1} = -\gamma^3 \tag{2.83}$$

It follows that the matrix part of T_{AU} must commute with γ^0 and γ^2 and anticommute with γ^1 and γ^3. Consequently, we obtain

$$\boxed{T_{AU} = \gamma^1\gamma^3()^*} \tag{2.84}$$

Note that, once again, $T_U^2 = I$, whereas $T_{AU}^2 = -I$.

2.2.3.3 The charge conjugation operator C

For this operator, which changes the sign of the charge, we need the Dirac equation with an electromagnetic field[15]:

[15]The field is represented there by the 4-vector potential A_μ

$$(\gamma^\mu(p_\mu - eA_\mu) - m)\,\psi = 0 \tag{2.85}$$

We take the complex conjugate of it and multiply on the left by $i\gamma^2$. We get

$$(i\gamma^2\gamma^{\mu^*}(-p_\mu - eA_\mu) - im\gamma^2)\,\psi^* = 0 \tag{2.86}$$

It is easy to verify using matrix realizations that

$$\gamma^2\gamma^{\mu^*} = -\gamma^\mu\gamma^2 \tag{2.87}$$

As a result, we have

$$(\gamma^\mu(p_\mu + eA_\mu) - m)\,(i\gamma^2\psi^*) = 0 \tag{2.88}$$

In other words, the operator that changes the sign of the charge is

$$\boxed{C = i\gamma^2()^*} \tag{2.89}$$

And since this only concerns the charge and doesn't affect the energy, everyone agrees: this operator is antiunitary.

There are several observations to be made at this stage.

The first is that the situation is like we've missed out one half the world... There is, in fact, one part of the world where the energy is positive (and everyone agrees on that) and another part where the energy is negative (and few people want that!).

These two perspectives represent different choices, but both are equally valid—according to Wigner's theorem!

Let's summarize the situation in a table:

	$E \geq 0$	$E < 0$
P	unitary	antiunitary
T	antiunitary	unitary

A second point to note is these two results

$$P_U T_{AU} = -P_{AU} T_U = \gamma^0\gamma^1\gamma^3()^* \tag{2.90}$$

$$CP_U T_{AU} = -CP_{AU} T_U = i\gamma^0\gamma^1\gamma^2\gamma^3 \equiv \gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.91}$$

This is an important point because it is not the P, C or T symmetries individually that play a crucial role in physics, but rather their product.

Finally, we introduce the γ^5 matrix, which will play an important role in what follows.

2.2.4 Solutions to the Dirac equation

We take equation (2.43) and apply it to

$$\psi = \exp(-i(Et - p.\overrightarrow{r}))\,\varphi \tag{2.92}$$

We get

$$\begin{pmatrix} -E+m & 0 & p3 & p1-ip2 \\ 0 & -E+m & p1+ip2 & -p3 \\ p3 & p1-ip2 & -E-m & 0 \\ p1+ip2 & -p3 & 0 & -E-m \end{pmatrix} \begin{pmatrix} \varphi^1 \\ \varphi^2 \\ \varphi^3 \\ \varphi^4 \end{pmatrix} = 0 \tag{2.93}$$

In this last expression, the $pj's$ are no longer operators but numbers, thanks to the "plane wave" exponential. We are therefore dealing with a system of four equations with four unknowns:

$$\begin{cases} (E-m)\varphi^1 = p3\,\varphi^3 + (p1-ip2)\varphi^4 \\ (E-m)\varphi^2 = (p1+ip2)\varphi^3 - p3\,\varphi^4 \\ (E+m)\varphi^3 = p3\,\varphi^1 + (p1-ip2)\varphi^2 \\ (E+m)\varphi^4 = (p1+ip2)\varphi^1 - p3\,\varphi^2 \end{cases} \tag{2.94}$$

If the electron were at rest, only the first four terms would remain, subject to Einstein's constraint

$$E^2 = m^2 \tag{2.95}$$

If $E = m$, the solutions are referred to as positive-energy states and take the form

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{2.96}$$

If $E = -m$, then we have two negative-energy states, which are

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} ; \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.97}$$

The term has been coined: negative-energy states!

They are inherent in any relativistic theory, since we rely on Einstein's equation, which specifies only the square of the energy! But keep in mind that the vast majority of Standard Model physicists do everything—absolutely everything—to avoid these states.

In the meantime, here they are. Of the four solutions, two are associated with positive energies and two are linked to negative energies.

If we solve the system in general, we get[16]

$$\varphi^3 = \frac{p3}{E+m} ; \varphi^4 = \frac{p1 + ip2}{E+m} \tag{2.98}$$

if we set $\varphi^1 = 1$ and $\varphi^2 = 0$ and

$$\varphi^3 = \frac{p1 - ip2}{E+m} ; \varphi^4 = -\frac{p3}{E+m} \tag{2.99}$$

if $\varphi^1 = 0$ and $\varphi^2 = 1$.

These are the two states of positive energy.

The two states of negative energy are determined by successively setting $\varphi^3 = 1, \varphi^4 = 0$ and $\varphi^3 = 0, \varphi^4 = 1$, yielding respectively

$$\varphi^1 = \frac{p3}{E-m} ; \varphi^2 = \frac{p1 + ip2}{E-m} \tag{2.100}$$

$$\varphi^1 = \frac{p1 - ip2}{E-m} ; \varphi^2 = -\frac{p3}{E-m} \tag{2.101}$$

[16]The system does not specify all four functions, but only two of them... which leads to different formulations depending on the choices made, based on how they will be used later on. You'll find one possible option here, but there will be others in following Chapters.

The full range of solutions is therefore

$$\psi_{+,1} = \sqrt{\frac{E+m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} 1 \\ 0 \\ \frac{p3}{E+m} \\ \frac{p1+ip2}{E+m} \end{pmatrix} \tag{2.102}$$

$$\psi_{+,2} = \sqrt{\frac{E+m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} 0 \\ 1 \\ \frac{p1-ip2}{E+m} \\ -\frac{p3}{E+m} \end{pmatrix} \tag{2.103}$$

$$\psi_{-,1} = \sqrt{\frac{E-m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} \frac{p3}{E-m} \\ \frac{p1+ip2}{E-m} \\ 1 \\ 0 \end{pmatrix} \tag{2.104}$$

$$\psi_{-,2} = \sqrt{\frac{E-m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} \frac{p1-ip2}{E-m} \\ -\frac{p3}{E-m} \\ 0 \\ 1 \end{pmatrix} \tag{2.105}$$

The normalization factor $\sqrt{\frac{E\pm m}{2E}}$ is explained by the square norm of the solutions. For example, if it were absent, we would have

$$\psi^{\dagger}_{+,1}\psi_{+,1} = 1 + \frac{\overrightarrow{p}^2}{(E+m)^2} = \frac{2E}{(E+m)}$$

However, the desired value must be equal to 1.

2.2.5 So, should we accept these negative energies or not?

That's the big knot in the Dirac equation. And it won't win him any friends...

Let's quote Niels Bohr's charming remarks:

"Dirac's theory is good for nothing except as bait to capture elephants in Africa. If they saw it written on a sign, the elephants—which are intelligent animals—would be mesmerized. We could then easily capture them and send them to a zoo."

or Werner Heisenberg:

"The saddest chapter in modern physics is and will remain Dirac's theory... I consider it scholarly rubbish that no one can take seriously."

Well, it just goes to show that we know how to communicate in a considerate way in the world of scientific research...

What is it that bothers these two giants of quantum physics so much?

Negative energies and their interpretation via the Dirac sea.

The latter can be summarized by the following diagram:

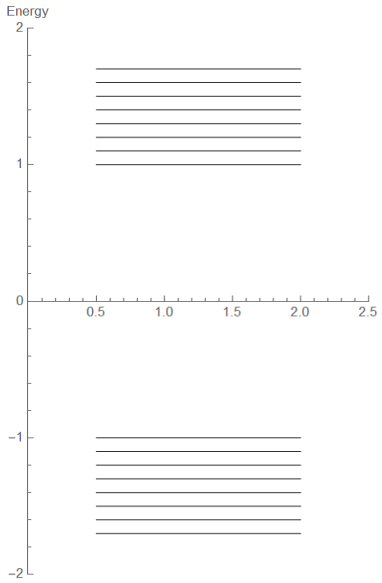


Figure 2.1: Dirac's sea

Since everything starts from Einstein's relation (2.23), we have a double series of energies. Those corresponding to rest in Eqs. (2.95), then at the first non-zero momentum $p1 = ||\overrightarrow{p1}||$:

$$E = \pm\sqrt{p1^2 + m^2} \tag{2.106}$$

and so on. The greater the momentum, the greater the positive energies and the smaller the negative energies (though they are large in absolute terms).

Dirac's interpretation is that the lowest-energy negative states are occupied first (by the fastest electrons).

Why, then, do some occupy the higher energy levels? Because they are forced to do so by the Pauli exclusion principle.

The Austrian physicist Wolfgang Pauli demonstrated the behavior of fermions—particles with half-integer spin, which include electrons. They are antisocial in nature: no more than two electrons per energy level, and even then, these two electrons must be such that one has a spin up while the other has a spin down.

So, as soon as the lowest energy level is occupied by its two electrons, the next two electrons occupy the next higher energy level, and so on.

The negative energy levels are thus all occupied—with the possible exception of one or two "holes"—and the subsequent electrons occupy the positive energy levels.

Dirac claimed that we would never observe electrons with negative energy, much like the underwater creatures that live in the deep sea. That doesn't mean they don't exist...

But he was so harshly criticized that he proposed another explanation in 1931[17]. Since electrons with negative energy cannot exist, we'll send them into the positive energy zone!

Look at Figure (2.1): to make this big leap from $E = -m$ to $E = m$, the "politically correct" electron must have an energy equal to $2m$. Its charge, of course, remains unchanged.

Let's think about this for a moment... an energy of $2m$, a charge differential of zero... this could also correspond to two particles:

1. an electron with spin $1/2$, mass m and charge $(-e)$
2. a particle with spin $1/2$, mass m, and charge e

This mysterious particle therefore has all the characteristics of an electron, except for its charge, which has the same absolute value but a positive sign.

Dirac gave this particle a name: the antielectron.

The electron is described by the left representation $D(\frac{1}{2}, 0)$ of the Lorentz algebra; the antielectron is described by the right representation $D(0, \frac{1}{2})$ of the Lorentz algebra. These are different representations, so the particles are different as well.

The Dirac equation, which is characterized by these two representations, thus implies the existence of this mysterious antielectron.

But theoretical particles are proposed all the time without their existence necessarily being confirmed experimentally...

That won't be the case here.

A year later, in 1932, the American physicist Carl Anderson proved the existence of this antielectron. He achieved this by studying the particles present in cosmic radiation using a cloud chamber. Indeed, when a charged particle passes through this cloud chamber, its trajectory is revealed by the vapor condensing into fine droplets. The charge of this particle can be determined by observing

[17] Dirac, P.A.M. (1931). Quantised singularities in the electromagnetic field. *Proceedings of the Royal Society of London*, SeriesA133, 60-72. https://doi.org/10.1098/rspa.1931.0130

how its trajectory is deflected when the chamber is subjected to a magnetic field.

And what happens? Anderson observes two different curvatures[18]:

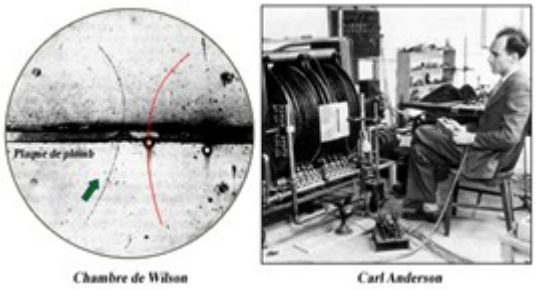


Figure 2.2: Anderson's experiment

An electron would have left a mark along the arc on the right. The fact that the curvature of the observed arc—on the left—is different indicates that we are dealing with a particle other than the electron.

Experimental confirmation of a particle very similar to the electron, then! Rather than calling it an antielectron, it was decided to name it the "positron". This marked the birth of antimatter!

And that's when you want to say, "Not so fast!"

Yes... because the radius of curvature of the observed arc is given by

$$R = \frac{mv}{qB} \tag{2.107}$$

In this equation, m is the mass of the particle, v is its velocity, q is its charge, and B is the strength of the magnetic field to which it is subjected in the experiment.

The field and the velocity must necessarily be identical regardless of the nature of the particle. This means that the sign of the radius of curvature can depend only on $\frac{m}{q}$.

Just think about it... it's 1932, and the conclusion is obvious: it's the charges that have opposite signs! But are we still convinced of that today?

The question is certainly worth asking, and that is precisely what Belgian researcher Daniel Dubois and his colleague, French researcher Gilles Nibart, do in an article with an unequivocal title[19].

Did Anderson's experiment demonstrate the existence of a particle with opposite charge or negative mass?

[18] Image source: https://laradioactivite.com/articles/le-phenomene/positon

[19] Dubois, D.M. and Nibart, G.N. (2022). What if the positron never existed? What if the antielectron is a negative mass electron? *Journal of Physics*: Conference Series 2197, 012006. DOI 10.1088/1742-6596/2197/1/012006

2.2.6 The Feynman-Stueckelberg approach

Attentive readers will have noticed that, while the "mysterious guest" in Dirac's equation was named the antielectron by its creator, it was the positron that emerged from Anderson's experiment.

Antielectron and positron... are they the same thing?

Well, not really!

We are in an era when negative energies are causing a stir among physicists—with the exception of Dirac.

What should we do if they appear anyway?

Remember, energy is

$$E = i\,\frac{\partial}{\partial t} \tag{2.108}$$

If we want to change the sign of the energy, we have only two options: play with the imaginary i, or imagine a time arrow that runs in reverse!

Changing the sign of the imaginary part i is what Dirac did in 1931. Indeed, his charge conjugation operator is antiunitary, and thus possesses complex conjugation.

The idea of reversing the arrow of time was first proposed by the Swiss physicist Ernst Stueckelberg[20] and later taken up by the American physicist Richard Feynman[21].

Feynman actually came up with a brilliant idea based on this: the diagrams that bear his name.

We'll explain using this example:

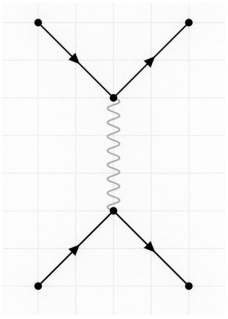


Figure 2.3: Feynman diagram

[20] Stueckelberg, E.C.G. (1942). La mécanique du point matériel en théorie de relativité et en théorie des quanta. *Helv. Phys. Acta*, 15, 23-37. https://doi.org/10.5169/seals-111289

[21] Feynman, R.P. (1949). The theory of positrons. Phys.Rev., 76, 749-759. https://doi.org/10.1103/PhysRev.76.749

There are conventions for reading Feynman diagrams: here, the arrows represent two electrons, and the wavy line represents a photon.

This diagram is interpreted as follows: the horizontal direction represents time; the vertical direction represents position; on the left side, two electrons move closer to each other over time. When the distance between them reaches a certain value, since they have identical charges, they will repel each other. This action is due to electromagnetic interaction and, more specifically, to its agent, the photon.

But what if we rotate this diagram 90 degrees? We get this:

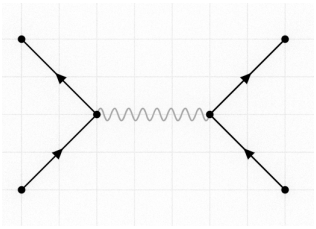


Figure 2.4: Time reversal in the Feynman diagram

Some arrows still represent electrons for which time flows from the past to the future. But the other arrows indicate that time flows from the future to the past! Feynman tells us that these will be associated with antimatter, specifically the positron.

So, this diagram, read from left to right, tells us a story.

Episode One: an electron e^- encounters a positron e^+ and they annihilate to produce energy in the form of photons γ:

$$e^- + e^+ \rightarrow \gamma + \gamma \tag{2.109}$$

Episode Two: this energy will spontaneously create matter and antimatter:

$$\gamma + \gamma \rightarrow e^- + e^+ \tag{2.110}$$

Does this perspective make a difference in solving the Dirac equation?

Yes. It implies that we must now set

$$\psi(t, \overrightarrow{r}) = N \exp(i(Et - \overrightarrow{p}.\overrightarrow{r}))\, v = N \exp(i(Et - i\overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} v1 \\ v2 \\ v3 \\ v4 \end{pmatrix} \tag{2.111}$$

reducing the Dirac equation to

$$\begin{cases} (E+m)v1 = p3\,v3 + (p1 - ip2)v4 \\ (E+m)v2 = -p3\,v4 + (p1 + ip2)v3 \\ (E-m)v3 = p3\,v1 + (p1 - ip2)v2 \\ (E-m)v4 = -p3\,v2 + (p1 + ip2)v1 \end{cases} \tag{2.112}$$

The four solutions obtained are then

$$v_{+,1} = \begin{pmatrix} \frac{p3}{E+m} \\ \frac{p1+ip2}{E+m} \\ 1 \\ 0 \end{pmatrix}, v_{+,2} = \begin{pmatrix} \frac{p1-ip2}{E+m} \\ -\frac{p3}{E+m} \\ 0 \\ 1 \end{pmatrix} \tag{2.113}$$

$$v_{-,1} = \begin{pmatrix} 1 \\ 0 \\ \frac{p3}{E-m} \\ \frac{p1+ip2}{E-m} \end{pmatrix}, v_{-,2} = \begin{pmatrix} 0 \\ 1 \\ \frac{p1-ip2}{E-m} \\ -\frac{p3}{E-m} \end{pmatrix} \tag{2.114}$$

up to a normalization factor. As before, it is obtained by setting

$$v_{+,1}^\dagger v_{+,1} = v_{+,2}^\dagger v_{+,2} = v_{-,1}^\dagger v_{-,1} = v_{-,2}^\dagger v_{-,2} = 1$$

which results in

$$N = \sqrt{\frac{E+m}{2E}} \tag{2.115}$$

for positive energy states and

$$N = \sqrt{\frac{E-m}{2E}} \tag{2.116}$$

for negative energy states.

That's right... with this approach, we can't escape negative energy either!

But let's summarize the different approaches.

2.2.7 We should probably wrap things up, right?

So we can summarize the previous developments this way:

ELECTRON	
$E > 0$; spin up	$E > 0$; spin down
$\sqrt{\frac{E+m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} 1 \\ 0 \\ \frac{p3}{E+m} \\ \frac{p1+ip2}{E+m} \end{pmatrix}$	$\sqrt{\frac{E+m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} 0 \\ 1 \\ \frac{p1-ip2}{E+m} \\ -\frac{p3}{E+m} \end{pmatrix}$
POSITRON	
$E > 0$; spin up	$E > 0$; spin down
$\sqrt{\frac{E+m}{2E}} \exp(i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} \frac{p1-ip2}{E+m} \\ -\frac{p3}{E+m} \\ 0 \\ 1 \end{pmatrix}$	$\sqrt{\frac{E+m}{2E}} \exp(i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} \frac{p3}{E+m} \\ \frac{p1+ip2}{E+m} \\ 1 \\ 0 \end{pmatrix}$

ANTIELECTRON	
$E < 0$; spin up	$E < 0$; spin down
$\sqrt{\frac{E-m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} \frac{p3}{E-m} \\ \frac{p1+ip2}{E-m} \\ 1 \\ 0 \end{pmatrix}$	$\sqrt{\frac{E-m}{2E}} \exp(-i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} \frac{p1-ip2}{E-m} \\ -\frac{p3}{E-m} \\ 0 \\ 1 \end{pmatrix}$
ANTIPOSITRON	
$E < 0$; spin up	$E < 0$; spin down
$\sqrt{\frac{E-m}{2E}} \exp(i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} 0 \\ 1 \\ \frac{p1-ip2}{E-m} \\ -\frac{p3}{E-m} \end{pmatrix}$	$\sqrt{\frac{E-m}{2E}} \exp(i(Et - \overrightarrow{p}.\overrightarrow{x})) \begin{pmatrix} 1 \\ 0 \\ \frac{p3}{E-m} \\ \frac{p1+ip2}{E-m} \end{pmatrix}$

The electron and the antielectron are the two particles identified by Dirac as early as 1928. The former has positive energy, while the latter is characterized by negative energy.

Stueckelberg and Feynman retained the electron, but abandoned the antielectron in favor of the positron—which has positive energy—by performing a time reversal.

The electron (positive matter) is the common element in both approaches. Another common element is the exclusion of the antipositron (negative matter).

It is the sign of antimatter that distinguishes the two approaches: negative in Dirac's model, it becomes positive in the Stueckelberg-Feynman model.

In both approaches, the spin is specified by applying the rotation matrix about the z-axis to the bispinors, which are assumed to be at rest. For example

$$\frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

If the eigenvalue is $1/2$ (as in this case), the state is spin up. If the eigenvalue is -1/2, the state is spin down.

But how do we transition from one state to another? By applying which operator?

That is what we will now examine.

The charge conjugation operator is

$$C = i\gamma^2()^* = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}()^* \tag{2.117}$$

When applied to the electron's spin-up state, it gives

$$\sqrt{\frac{E+m}{2E}}\exp(i(Et - \overrightarrow{p}.\overrightarrow{r}))\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ \frac{p3}{E+m} \\ \frac{p1-ip2}{E+m} \end{pmatrix}$$

$$= \sqrt{\frac{E+m}{2E}} \exp(i(Et - \overrightarrow{p}.\overrightarrow{r})) \begin{pmatrix} \frac{p1-ip2}{E+m} \\ -\frac{p3}{E+m} \\ 0 \\ 1 \end{pmatrix}$$

We thus recover the spin-up state of the positron. It is easy to see that the same holds true for the spin-down states.

The states of the positron are therefore obtained by applying the charge conjugation operator C to the states of the electron, and vice versa.

The states of the antiproton can also be obtained by applying the charge conjugation operator C to the states of the antielectron, and vice versa.

The operator C therefore changes the sign of the charge but not that of the energy or the spin.

How can an electron in the spin-up state transition to one of the two states of the antielectron? Well, it can't... unless we change the sign of the mass! Generally speaking, this antielectron would only make sense if it had both negative energy and negative mass.

If we change the sign of the mass—recalling that this was an option included in Dirac's original approach in 1928—then it is the operator represented by the matrix

$$\gamma^5 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.118}$$

that mediates the transition between electron states and antielectron states.

The γ^5 operator therefore changes the sign of the energy but not that of the charge or the spin. This change must be accompanied by a change in the sign of the mass.

The explanation for the change in the sign of the mass is very simple:

$$\gamma^5(\gamma^\mu(p_\mu - eA_\mu) - m)\,\psi = -(\gamma^\mu(p_\mu - eA_\mu) + m)\,\gamma^5\psi \tag{2.119}$$

since γ^5 anticommutes with each of the γ^μ.

Since spin plays a less significant role in classifying states based on the sign of their energy[22], we conclude that there are four types of states (positive-mass matter, negative-mass matter, positive-mass antimatter and negative-mass

[22]Clearly, energy and mass have the same sign. These wavefunctions are therefore also classified according to the sign of the mass.

antimatter) that can be represented as follows using the chiral realization of the γ^μ matrices:

Matter $E \geq 0$
$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$
Matter $E < 0$
$CP_UT_{AU}\psi = -CP_{AU}T_U\psi = \gamma^5\psi = \begin{pmatrix} -\psi^1 \\ -\psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$
Antimatter $E \geq 0$
$C\psi = i\gamma^2\psi^* = \begin{pmatrix} \psi^{4^*} \\ -\psi^{3^*} \\ -\psi^{2^*} \\ \psi^{1^*} \end{pmatrix}$
Antimatter $E < 0$
$P_UT_{AU}\psi = -P_{AU}T_U\psi = i\gamma^5\gamma^2\psi^* = \begin{pmatrix} -\psi^{4^*} \\ \psi^{3^*} \\ -\psi^{2^*} \\ \psi^{1^*} \end{pmatrix}$

Table 2.2: List of matter-antimatter wavefunctions

2.2.8 Chirality

We have just seen that it is the matrix γ^5 that links the electron and its positive-energy antiparticle to their negative-energy (and negative-mass) counterparts.

This matrix is well known in Dirac theory: it is the matrix associated with chirality.

Chirality? What does that mean?

Whenever you need to tell your left from your right, you're using chirality.

Take a look at your two hands. They have the same parts: five fingers, the same relative sizes, and so on. And yet, they're different! On your right hand, the thumb is on the left, and on your left hand, the thumb is on the right. Even if you press your palms together to swap the positions of the thumbs and other fingers, your hands are still different. You had to rotate your right hand counterclockwise and your left hand clockwise to bring them together.

Based on this observation, no matter what you do, your left hand is always different from your right hand... we say that your hands are chiral. In other words, it is necessary to distinguish a right side from a left side.

However, when you bring your two hands together, the whole is no longer chiral: mirror symmetry is preserved for both hands as a whole. We say that the system consisting of your two hands is enantiomorphic.

For reasons that aren't entirely clear, nature tends to favor chirality. This is especially true when it comes to small objects like molecules or particles.

It's easy to tell whether a macroscopic object is chiral or not. It's harder to say the same for a microscopic object.

It was the French chemist and physicist Louis Pasteur who introduced the concept of chirality at the molecular level.

Two diagrams will suffice to explain the concept.

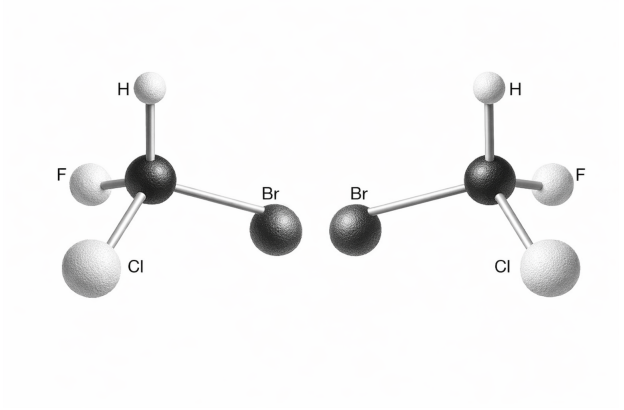


Figure 2.5: Example of a chiral molecule : CHBrClF or Bromochlorofluoromethane

On the left is the original molecule. On the right is its mirror image. If we rotate the mirror image so that the fluorine returns to its original position, we see that the positions of the chlorine and bromine have switched. The molecule

is therefore chiral.

Is the orientation really that important?

Oh yes! The left-handed form of a molecule may have therapeutic effects, while its right-handed form can be harmful. For example, L-DOPA has proven to be an effective treatment for reducing tremors, particularly in cases of Parkinson's disease. On the other hand, the D-DOPA form not only fails to provide any relief to patients but can even cause harmful side effects such as granulocytopenia, which weakens the immune system.

Another example: the infamous thalidomide. The right-handed form has a sedative effect and alleviates nausea, while the left-handed form is teratogenic.

However, unlike chemical laboratories, which can produce both right-handed and left-handed versions, it appears that nature favors certain versions over others. Sugars, for example, are exclusively dextrorotatory (i.e., right-handed) molecules. This is indicated by the right-handed polarization of light. DNA is also found exclusively in its right-handed form. Conversely, all amino acids are levorotatory (i.e., left-handed).

Why does nature favor one version over the other? Why this version rather than the other? Various hypotheses have been put forward, including the extraterrestrial origin of the imbalance. There appears to be a chiral force that favors either the right-handed or left-handed form, depending on the molecule being studied. This chiral force is thought to originate from circularly polarized light[23]. And, in fact, such infrared light was reportedly detected in the Orion Nebula in 1998.

But we have to admit that researchers don't have any truly convincing explanation at this point...

Well, all of that has to do with molecules. But what about elementary particles?

In the absence of other information, researchers use a highly mathematical definition based on group theory. The chirality of an electron or any particle is defined through a representation of the Lorentz algebra. Remember that there are two representations associated with the Dirac equation:

$$D(\frac{1}{2},0)\ ;\ D(0,\frac{1}{2})$$

The first is the left-handed version and applies only to the electron. The second is the right-handed version and applies only to the positron.

From the perspective of chirality, matter would therefore be left-handed and antimatter right-handed. The Dirac equation, which contains both versions, is enantiomorphic.

But how do we switch from one version to the other? Using the γ^5 matrix!

The two parallel worlds we mentioned in the previous Sections—the world of positive energies and masses (that is, the world of electrons and positrons)

[23]The magnitude of the electric field remains constant, while its direction changes as it rotates.

and the world of negative energies and masses (in other words, the world of antielectrons and antipositrons)—are thus connected by the chirality matrix.

Our world, for reasons unknown, has favored left-handed matter and right-handed antimatter, but by reversing the chirality, its mirror world—composed of right-handed antimatter and left-handed matter—could, in principle, become visible to us.

Chapter 3

The geometric algebra point of view

We mentioned Clifford algebra $Cl(1,3)$ in the previous chapter, and it is obvious that we cannot speak of this algebraic structure without referring to geometric algebras. The reason is simple: it was the English mathematician William Clifford who invented the concept of geometric algebra in 1878. He unfortunately passed away a year later and this idea did not have time to receive the publicity it deserved. This viewpoint is thus relatively unknown in the scientific world despite the remarkable work done in particular by David Hestenes[1] to rehabilitate it.

Geometric algebras are extremely powerful, yet also disconcerting at first glance. For example, you will be asked to forget one of the things you have learned since your earliest childhood, namely that you don't add apples and pears...

But let's look at this in detail.

3.1 The foundations

We introduced the concept of algebra in chapter 1, by discussing Lie algebras. Algebras are vector spaces equipped with a specific "multiplication" law and, in the case of Lie algebras, it is the "commutator" law. Similarly, a geometric algebra is a vector space equipped with a law called the geometric product[2]:

$$ab \equiv a.b + a \wedge b \tag{3.1}$$

[1] For a review, see, for instance, Hestenes, D. (2003). Spacetime Physics with Geometric Algebra. *American Journal Physics*, 71, 691-704. https://doi.org/10.1119/1.1571836

[2] The elements of this vector space are denoted by a lowercase letter, without an arrow.

The geometric law combines two vectors (also called 1-vectors) a and b to give rise to

- a scalar (or 0-vector) through $a.b$
- a bivector (or 2-vector) through $a \wedge b$

Everyone knows what a scalar is. The term bivector is undoubtedly more mysterious.

The idea is quite simple. We know vectors can be seen as segments of lines to which we give an orientation. Similarly, bivectors are oriented segments of planes. The shape of these parts of a plan is irrelevant as long as they have the same area and orientation:

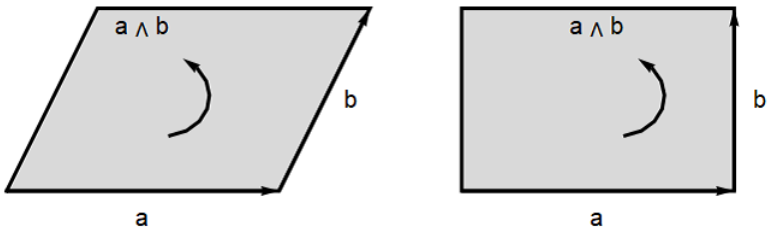


Figure 3.1: Two representations of the same bivector $a \wedge b$

Later, we will also define trivectors -or 3-vectors- as oriented segments of 3D spaces and so on.

The definition (3.1) implies two products:

- an inner product or dot product denoted by ".". It is the usual scalar product when only vectors are involved. One of its characteristics is that it is symmetric:

$$a.b = b.a \tag{3.2}$$

- an outer product or wedge product denoted by " $\wedge$ ". In contrast to the first product, it is antisymmetric:

$$a \wedge b = -b \wedge a \tag{3.3}$$

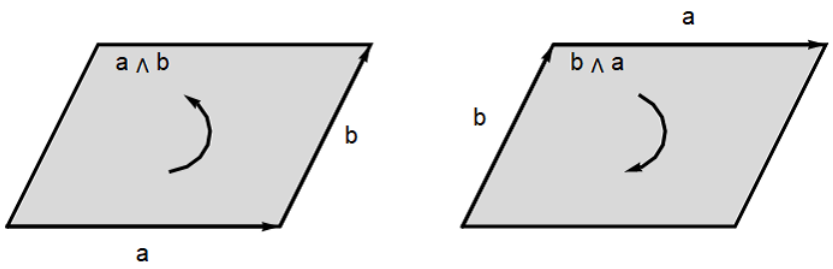


Figure 3.2: Antisymmetry of the wedge product

When the two vectors a and b are parallel, the outer product vanishes due to antisymmetry and only the symmetric contribution of the inner product remains. In this case,

$$ab = ba \tag{3.4}$$

A consequence of this is[3]

$$aa = a.a = \alpha,\ \alpha \in \mathbb{R} \tag{3.5}$$

This is known as the contraction axiom. It follows that

$$a^{-1} = \frac{a}{\alpha} \tag{3.6}$$

When the two vectors a and b are orthogonal, the inner product becomes zero, leaving only the antisymmetric contribution of the outer product. In this case, we have

$$ab = -ba \tag{3.7}$$

The degree of commutativity in the geometric product therefore measures, in a way, the direction of a with respect to b. If they are collinear, we have commutativity and if they are orthogonal, anticommutativity emerges:

$$a\ //\ b \Leftrightarrow a \wedge b = 0 \Leftrightarrow ab = a.b \Leftrightarrow ab = ba$$

$$a\ \perp\ b \Leftrightarrow a.b = 0 \Leftrightarrow ab = a \wedge b \Leftrightarrow ab = -ba$$

This explains the two terms in (3.1). They result from the decomposition of the vector b into two components, one parallel to a and the other perpendicular to a:

$$ab = a(b_{//} + b_{\perp}) = a.b_{//} + a \wedge b_{\perp}$$

[3]The reason why α can be negative is related to the metric of the original vector space.

In other words, the geometric product contracts parallel directions and expand perpendicular directions:

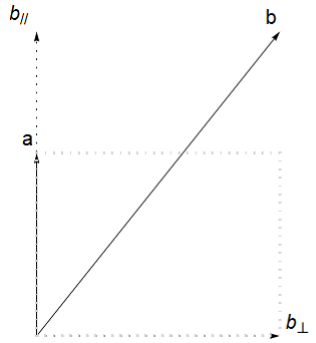


Figure 3.3: Visualization of the geometric product

Generally speaking, we have

$$a.b = \frac{1}{2}(ab + ba) \; ; \; a \wedge b = \frac{1}{2}(ab - ba) \tag{3.8}$$

From two 1-vectors, the inner product creates a 0-vector: it lowers the grade of the vectors. On the contrary, starting with two 1-vectors, the outer product creates a 2-vector: it raises the grade of the vectors. This product therefore allows the creation of 3-vectors, 4-vectors, etc. (depending on the dimension of the initial vector space).

More specifically, if the initial vector space has dimension n and an orthonormal basis denoted by e_1, e_2, ..., e_n, the geometric product allows to create 2^n elements:

- 1 0-vector: 1
- n 1-vectors: $e_j, \; j = 1, 2, ..., n$
- $\frac{n(n-1)}{2}$ 2-vectors: $e_j e_k, \; j, k = 1, 2, ..., n$
- ...
- 1 n-vector: $e_1 e_2 e_3 ... e_n$

The unique n-vector is called the pseudoscalar unit. It is the one who determines the orientation of the space.

A multivector is a real linear combination of these 2^n elements:

$$c^0 + c^j \, e_j + c^{jk} \, e_j e_k + ... + c^{n+1} \, e_1 e_2 e_3 ... e_n$$

The set of these multivectors, equipped with addition, multiplication by a scalar and geometric product, is defined as the geometric algebra itself or, equivalently, the Clifford algebra $Cl(p,q)$ if p and q basis vectors have a square equal to +1 and -1, respectively.

Let's illustrate these considerations on $\mathbb{R}^2$.
There are two basis vectors, orthonormal:

$$e_1\,, e_2 \tag{3.9}$$

The inner product gives rise, through e_1^2 and e_2^2, to the 0-vector 1 while the outer product is such that

$$e_1 \wedge e_2 = e_1 e_2 \tag{3.10}$$

The corresponding geometric algebra is thus of dimension 4. No need to say that it coincides with the Clifford algebra $Cl(2)$.
We can summarize this in a table:

$\mathbb{R}^2$	$Cl(2)$	Grade
	1	0
e_1	e_1	1
e_2	e_2	1
	$e_1 e_2$	2

A multivector of $Cl(2)$ is:

$$v = c^0 + c^1\, e_1 + c^2\, e_2 + c^3\, e_1 e_2, \;\; c^0, c^1, c^2, c^3 \in \mathbb{R} \tag{3.11}$$

It is also written as

$$v =< A >_0 + < A >_1 + < A >_2$$

referring to the grade of the elements of the algebra $A \equiv Cl(2)$.
The even part of this multivector is

$$v_0 = c^0 + c^3\, e_1 e_2 =< A >_0 + < A >_2$$

while the odd part is just a vector:

$$v_1 = c^1\, e_1 + c^2\, e_2 =< A >_1$$

Two major observations at this stage:

- The square of the pseudoscalar unit is equal to -1:

$$(e_1e_2)^2 = e_1e_2e_1e_2 = -e_1^2e_2^2 = -1 \tag{3.12}$$

It is then customary to say that this unit plays the role of the imaginary i. In other words, the even part of $Cl(2)$ is isomorphic to the set of complex numbers.

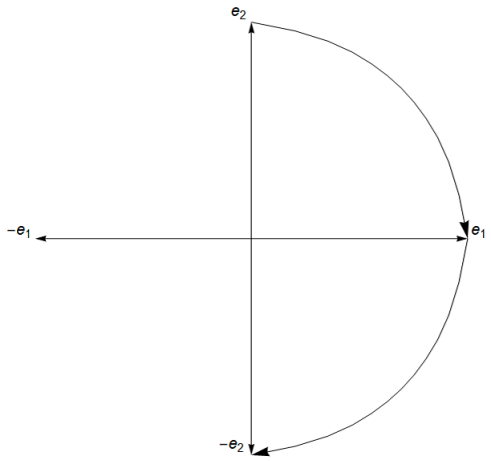
- The pseudoscalar unit rotates e_1 to $-e_2$:

$$e_1e_2e_1 = -e_1^2e_2 = -e_2 \tag{3.13}$$

It therefore rotates e_1 by 90° clockwise. It also rotates e_2 to e_1:

$$e_1e_2e_2 = e_1e_2^2 = e_1 \tag{3.14}$$

which is also a 90° clockwise rotation.



It is true for the basis vectors but also for any vector of $\mathbb{R}^2$. The position occupied by the pseudoscalar unit reveals the direction of rotation: clockwise if the pseudoscalar unit multiplies the vector on the left and counterclockwise if it is applied on the right

$$e_1e_2\,v_1 = c^2e_1 - c^1e_2\ ;\ v_1\,e_1e_2 = -c^2e_1 + c^1e_2 \tag{3.15}$$

This is yet another illustration of the mantra of geometric algebra: the geometric product contracts parallel directions and expand perpendicular directions.

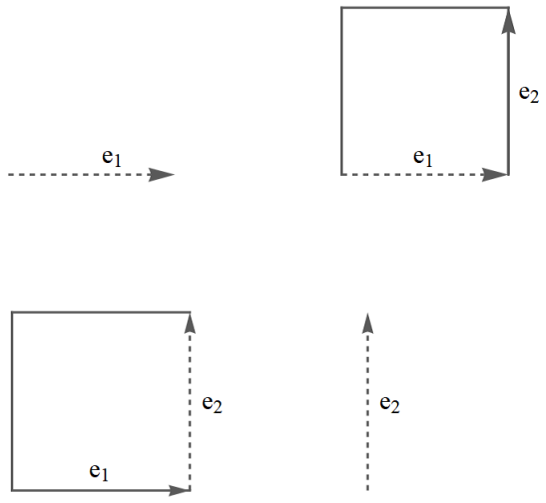


Figure 3.4: Contraction of parallel directions-expansion of perpendicular directions

3.2 Reflections and rotations

Besides what we have just mentioned, there is another crucial thing to understand in geometric algebra: some multivectors represent geometric transformations rather than geometric objects.

Thus, Clifford algebras are naturally connected with transformations such as reflections and rotations.

Let's stay in $\mathbb{R}^2$ for a moment. What follows can be transposed to any vector space, but $\mathbb{R}^2$ does indeed allow for optimal visualization.

Consider a 1-vector a and reflect it with respect to the 1-vector u :

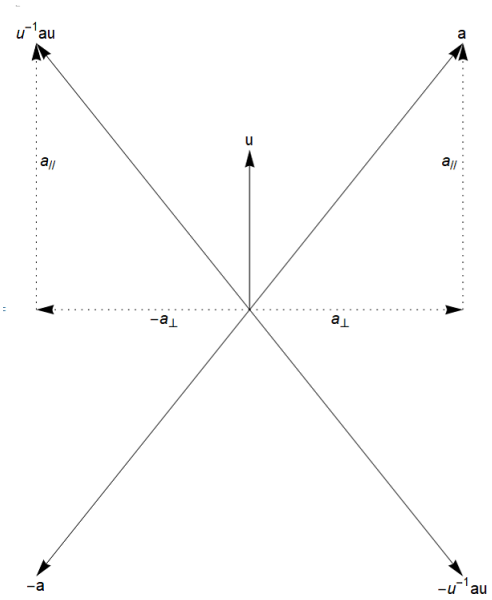


Figure 3.5: Reflection of a

This reflection is given by

$$u^{-1}a\,u$$

Indeed, we have

$$u^{-1}a\,u = u^{-1}(a_{//} + a_\perp)\,u$$

Remember now (3.4) and (3.7) and you will be convinced that

$$a_{//}\,u = u\,a_{//}\ ;\ a_\perp\,u = -u\,a_\perp$$

leading to

$$u^{-1}a\,u = a_{//} - a_\perp$$

which is the desired result.

Furthermore, it is clear that the reflection of a across a line perpendicular to u is marked through

$$-u^{-1}a\,u$$

As a numerical example, we consider the vector

$$a = e_1 + 2\,e_2$$

Let's reflect it along

$$u = \frac{1}{5}(3\,e_1 + 4\,e_2) = u^{-1}$$

Note that u has been normalized. This is something we will do systematically going forward. If that weren't the case, this normalization factor would be carried over to the calculation of the inverse.

We have

$$u^{-1} a\, u = \frac{1}{25}(3\,e_1 + 4\,e_2)\,(e_1 + 2\,e_2)\,(3\,e_1 + 4\,e_2) = \frac{1}{25}(41\,e_1 + 38\,e_2)$$

A simple graph proves that this vector is indeed the object sought.

It is well known that two successive reflections are equivalent to one rotation and that the angle of rotation is exactly twice the angle between the two intersecting lines. This is denoted, in algebraic geometry, by the appearance of what is called a rotor

$$R \equiv uv$$

in the sense that

$$R^{-1} a\, R = v^{-1}(u^{-1} a\, u)\, v$$

A rotation is therefore considered equivalent to the geometric product of two (normalized) 1-vectors:

$$R = uv = u.v + u \wedge v$$

Moreover we know that

$$u.v = \cos\frac{\theta}{2} \tag{3.16}$$

if $\frac{\theta}{2}$ is the angle between u and v. On the other side, the wedge product is related to the area of the parallelogram built on u and v:

$$u \wedge v = \sin\frac{\theta}{2}\, e_1 e_2 \tag{3.17}$$

These two results lead to

$$R = uv = (\cos\frac{\theta}{2} + \sin\frac{\theta}{2}\, e_1 e_2) \tag{3.18}$$

Now, $e_1 e_2$ behaves like i and it is clear that

$$R = uv = \exp(e_1 e_2\, \frac{\theta}{2}) \tag{3.19}$$

In general, the rotor thus specifies the angle of rotation, as well as the plane (u, v) in which the rotation takes place and its direction (clockwise if $u \wedge v$ and counterclockwise if $v \wedge u$).

We note that

$$R^{-1} = vu = \exp(-e_1e_2\,\frac{\theta}{2})$$

So if we want to subject the vector a to a rotation of amplitude θ in the direction u to v, we simply need to consider

$$a\prime \equiv R^{-1}a\,R = \exp(-e_1e_2\,\frac{\theta}{2})\,a\,\exp(e_1e_2\,\frac{\theta}{2}) \tag{3.20}$$

which gives the desired result.

One might be surprised by this formula, frequently called "the sandwich formula" for obvious reasons. One might indeed think that it would be simpler to apply the rotor only once, doubling the amplitude of the angle and defining

$$a\prime \equiv a\,R(2\theta) = a\,\exp(e_1e_2\,\theta)$$

And that would be perfectly the case if we were to limit ourselves permanently to $\mathbb{R}^2$! But you only need to go up to $\mathbb{R}^3$ to be convinced that this option is no longer valid.

Let's consider an example to prove that this formula leads to a dead end. Take the vector

$$a = 2e_1 + 3e_2 + 4e_3$$

Let's rotate it 90° in the plane (e_1, e_2) (which is equivalent to saying around e_3). Equation (3.20) gives the desired result:

$$a\prime = \exp(-e_1e_2\,\frac{\pi}{4})\,a\,\exp(e_1e_2\,\frac{\pi}{4})$$

$$\Leftrightarrow a\prime = (\cos(\frac{\pi}{4}) - e_1e_2\sin(\frac{\pi}{4}))\,a\,(\cos(\frac{\pi}{4}) + e_1e_2\sin(\frac{\pi}{4}))$$

$$\Leftrightarrow a\prime = \frac{1}{2}(1 - e_1e_2)\,(2e_1 + 3e_2 + 4e_3)\,(1 + e_1e_2)$$

$$\Leftrightarrow a\prime = \frac{1}{2}(-e_1 + 5e_2 + 4e_3 - 4e_1e_2e_3)\,(1 + e_1e_2)$$

$$a\prime = -3e_1 + 2e_2 + 4e_3$$

However, the simplified proposal leads to

$$a\prime = a\,\exp(e_1e_2\,\frac{\pi}{2}) = (2e_1 + 3e_2 + 4e_3)\,e_1e_2 = -3e_1 + 2e_2 + 4e_1e_2e_3$$

The presence of the third term means that $a\prime$ is no longer a vector and we must therefore necessarily resort to "the sandwich formula" (3.20).

3.3 The space algebra

The initial vector space is $\mathbb{R}^3$ and the geometric algebra is $Cl(3)$ -also called the space algebra in this case-:

$\mathbb{R}^3$	$Cl(3)$	Grade
	1	0
e_1	e_1	1
e_2	e_2	1
e_3	e_3	1
	e_1e_2	2
	e_3e_1	2
	e_2e_3	2
	$e_1e_2e_3$	3

The space algebra is made of 8 elements : 1 scalar, 3 vectors, 3 bivectors (also called pseudovectors) and 1 trivector (or pseudoscalar). If you feel you recognize this series of numbers, you are not seeing things because it is indeed a line of Pascal's triangle, a characteristic that will be present in every geometric algebra.

A multivector is here

$$v = c^0 + c^1e_1 + c^2e_2 + c^3e_3 + c^{12}e_1e_2 + c^{31}e_3e_1 + c^{23}e_2e_3 + c^4e_1e_2e_3$$

with c^0, c^1, ..., $c^4 \in \mathbb{R}$ or

$$v =< A >_0 + < A >_1 + < A >_2 + < A >_3$$

Let's return for a moment to the term "pseudo". What do we get when we apply the pseudoscalar unit to the scalar or vectors? This:

$$e_1e_2e_3\,1 = e_1e_2e_3 \tag{3.21}$$

$$e_1e_2e_3\,e_1 = e_2e_3\ ;\ e_1e_2e_3\,e_2 = e_3e_1\ ;\ e_1e_2e_3\,e_3 = e_1e_2 \tag{3.22}$$

So, respectively the pseudoscalar unit and the pseudovectors. We will say that scalars and pseudoscalars as well as vectors and pseudovectors avec Hodge duals.

In a more general way, the Hodge dual of a k-vector X is defined as

$$*X \equiv (-1)^{\frac{k(k-1)}{2}} IX$$

where I is the pseudoscalar unit. It is such that

$$X * X = I$$

A particularly clear way to illustrate this duality is through the use of the matrix realization of the space algebra. In this case, we are using Pauli matrices:

$$1 \to I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{3.23}$$

$$e_1 \to \sigma 1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, e_2 \to \sigma 2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, e_3 \to \sigma 3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.24}$$

$$e_2e_3 \to i\sigma 1 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, e_3e_1 \to i\sigma 2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, e_1e_2 \to i\sigma 3 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \tag{3.25}$$

$$e_1e_2e_3 \to iI_2 = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} \tag{3.26}$$

Therefore, as in the case of $Cl(2)$, we note that considering the dual of an element amounts to multiplying it by i.

We note that the role of the imaginary i can be attributed to four elements of $Cl(3)$, namely e_2e_3, e_3e_1, e_1e_2 besides the pseudoscalar unit $e_1e_2e_3$. This leads to identifying the even part $\{1, e_2e_3, e_3e_1, e_1e_2\}$ with the set of quaternions[4].

The position vector of the space algebra is defined by

$$x = x^j e_j = \begin{pmatrix} x^3 & x^1 - ix^2 \\ x^1 + ix^2 & -x^3 \end{pmatrix} \tag{3.27}$$

The fact that it undergoes a rotation of amplitude θ in the plane (e_1, e_2), for instance is noted through

[4]Quaternions were introduced by the Irish mathematician William Hamilton in 1843. They are generated by i, j and k such that $i^2 = j^2 = k^2 = ijk = -1$.

$$R^{-1}x\,R = \exp(-e_1e_2\,\frac{\theta}{2})\,x\,\exp(e_1e_2\,\frac{\theta}{2}) \tag{3.28}$$

This gives rise to

$$\exp(-e_1e_2\,\frac{\theta}{2})\,e_1\,\exp(e_1e_2\,\frac{\theta}{2}) = \cos(\theta)\,e_1 + \sin(\theta)\,e_2,$$

$$\exp(-e_1e_2\,\frac{\theta}{2})\,e_2\,\exp(e_1e_2\,\frac{\theta}{2}) = -\sin(\theta)\,e_1 + \cos(\theta)\,e_2$$

that is to say

$$x\prime = (x^1\cos(\theta) - x^2\sin(\theta))e_1 + (x^1\sin(\theta) + x^2\cos(\theta))e_2 + x^3e_3 \tag{3.29}$$

Besides, we have

$$R\,x\,R^{-1} = \exp(e_1e_2\,\frac{\theta}{2})\,x\,\exp(-e_1e_2\,\frac{\theta}{2}) \tag{3.30}$$

$$\Leftrightarrow x\prime = (x^1\cos(\theta) + x^2\sin(\theta))e_1 + (-x^1\sin(\theta) + x^2\cos(\theta))e_2 + x^3e_3 \tag{3.31}$$

Let's conclude this Section by saying a few words about the cross product. It is very close to the wedge product, and yet these two concepts are not equivalent. Indeed, we know that

$$a \times b = (a^2b^3 - a^3b^2)e_1 + (a^3b^1 - a^1b^3)e_2 + (a^1b^2 - a^2b^1)e_3 \tag{3.32}$$

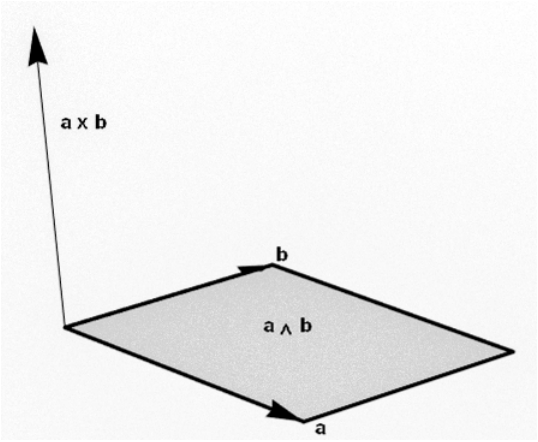
if

$$a = a^1e_1 + a^2e_2 + a^3e_3\ ;\ b = b^1e_1 + b^2e_2 + b^3e_3$$

On the other hand, we have

$$a \wedge b = (a^2b^3 - a^3b^2)e_2e_3 + (a^3b^1 - a^1b^3)e_3e_1 + (a^1b^2 - a^2b^1)e_1e_2 \tag{3.33}$$

Objects (3.32) and (3.33) therefore have exactly the same norm but their nature is different: axial vector for one and bivector for the other. This result states that the cross product is the Hodge dual of the wedge product (only in $\mathbb{R}^3$, of course). The duality operation interchanges a plane with a vector perpendicular to that plane:



A rather remarkable consequence concerns mirror symmetry. Take the example of a loop carrying an electric current. This generates a magnetic field, usually defined by a cross product. Its symmetrical counterpart is a magnetic field directed in the opposite direction (hence the name axial vector...). However, if we consider the wedge product instead of the cross product, we will not have two different vectors but a single bivector with a different orientation on either side of the mirror.

3.4 The 4D-space algebra

Here we go up to $\mathbb{R}^4$ and $Cl(4)$, respectively. The elements concerned are given in the table below:

$\mathbb{R}^4$	$Cl(4)$	Grade
	1	0
e_1, e_2, e_3, e_4	e_1, e_2, e_3, e_4	1
	e_1e_2, e_1e_3, e_1e_4, e_2e_3, e_2e_4, e_3e_4	2
	$e_1e_2e_3$, $e_1e_2e_4$, $e_1e_3e_4$, $e_2e_3e_4$	3
	$e_1e_2e_3e_4$	4

One possible matrix realization is

$$1 \to \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{3.34}$$

$$e_j \to \begin{pmatrix} 0 & -i\sigma j \\ i\sigma j & 0 \end{pmatrix}, \; e_4 \to \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, \; j = 1, 2, 3 \tag{3.35}$$

$$e_j e_4 \to \begin{pmatrix} -i\sigma j & 0 \\ 0 & i\sigma j \end{pmatrix}, \; e_j e_k \to i\varepsilon_{jkl} \begin{pmatrix} \sigma l & 0 \\ 0 & \sigma l \end{pmatrix}, \; j, k, l = 1, 2, 3 \tag{3.36}$$

$$e_j e_k e_4 \to i\varepsilon_{jkl} \begin{pmatrix} 0 & \sigma l \\ \sigma l & 0 \end{pmatrix}, \; e_1 e_2 e_3 \to \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}, \; j, k, l = 1, 2, 3 \tag{3.37}$$

$$e_1 e_2 e_3 e_4 \to \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \tag{3.38}$$

One might think that this is simply an immediate generalization of previous cases. It is by looking at rotations that we understand that this is hardly the case. This will be particularly evident if we look at the rotation matrices for $\mathbb{R}^2$, $\mathbb{R}^3$ and $\mathbb{R}^4$, respectively:

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

In $\mathbb{R}^2$, there is only one plane, let's say (e_1, e_2). Nothing is fixed except the point which is the center of rotation.

In $\mathbb{R}^3$, there can be three rotation planes: (e_1, e_2), (e_1, e_3) and (e_2, e_3). Here, we have chosen (e_1, e_2). The e_3 axis is therefore fixed... hence the presence of the "1" at the intersection of the third row and the third column. We could, in fact, speak of rotation around e_3 rather than rotation in (e_1, e_2). But geometric algebra is clear: we must refer to the plane of rotation.

In $\mathbb{R}^4$, an entire plane is fixed. Marked here by the vectors e_3 and e_4, it is orthogonal to the plane in which the rotation takes place. Since an entire plane is freed up, another rotation can be installed there:

$$R(\theta,\varphi) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\varphi & -\sin\varphi \\ 0 & 0 & \sin\varphi & \cos\varphi \end{pmatrix} \tag{3.39}$$

We are thus witnessing a simultaneous rotation in two orthogonal planes. And this is where it becomes remarkable. We know that it is impossible to transform a left hand drawn on a sheet of paper into a right hand on the same sheet without going through a reversal, that is to say a movement that seeks the third dimension. Similarly, we observe that it is impossible to change the chirality of a three-dimensional object without going through a higher dimension. So, unlike what happens in 3D, these rotations $R(\theta,\varphi)$ can induce reflections. For instance, taking $\theta = 0$ and $\varphi = \pi$, we have

$$R(0,\pi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.40}$$

Thus, from our 3D perspective, we observe the $z \to -z$ reflection. Of course, it is also accompanied by a reflection at the level of the fourth dimension, but this escapes us.

We can also observe total reflection when considering isoclinic rotations.

If the two angles θ and φ are equal, the rotation $R(\theta,\varphi)$ is called left isoclinic:

$$R_L(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{pmatrix} = \cos\theta + \sin\theta\, e_1e_3 \tag{3.41}$$

If the two angles are opposite, the rotation is right isoclinic:

$$R_R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & \sin\theta \\ 0 & 0 & -\sin\theta & \cos\theta \end{pmatrix} = \cos\theta + \sin\theta\, e_2e_4 \tag{3.42}$$

An isoclinic rotation of 180° is thus equivalent to

$$R_L(\pi) = R_R(\pi) = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{3.43}$$

and each of the coordinates goes to its opposite.

Another notable aspect of the transition to 4D is the existence of non-simple bivectors. A bivector is simple if its square is real. Any of the $Cl(3)$ bivectors is simple because

$$(c^{12}e_1e_2 + c^{31}e_3e_1 + c^{23}e_2e_3)^2 = -(c^{12})^2 - (c^{31})^2 - (c^{23})^2$$

For $Cl(4)$, this is no longer the case. For instance, we have

$$(e_1e_2 + e_3e_4)^2 = -2 + 2\,e_1e_2e_3e_4$$

This is where the pseudoscalar unit will take on a new dimension. Indeed, the operators

$$P_\pm \equiv \frac{1}{2}\left(1 \pm e_1e_2e_3e_4\right) \tag{3.44}$$

are projection operators in the sense that

$$P_\pm^2 = P_\pm,\ P_+P_- = P_-P_+ = 0 \tag{3.45}$$

Acting on the $Cl(4)$ bivectors, they give rise to

$$P_+e_1e_2 = -P_+e_3e_4 = \frac{1}{2}\left(e_1e_2 - e_3e_4\right) \tag{3.46}$$

$$P_+e_3e_1 = -P_+e_2e_4 = \frac{1}{2}\left(e_3e_1 - e_2e_4\right) \tag{3.47}$$

$$P_+e_2e_3 = -P_+e_1e_4 = \frac{1}{2}\left(e_2e_3 - e_1e_4\right) \tag{3.48}$$

$$P_-e_1e_2 = P_-e_3e_4 = \frac{1}{2}\left(e_1e_2 + e_3e_4\right) \tag{3.49}$$

$$P_-e_3e_1 = P_-e_2e_4 = \frac{1}{2}\left(e_3e_1 + e_2e_4\right) \tag{3.50}$$

$$P_-e_2e_3 = P_-e_1e_4 = \frac{1}{2}\left(e_2e_3 + e_1e_4\right) \tag{3.51}$$

Consequently, the projectors have the effect of combining the bivectors in such a way that they generate two $so(3)$ in direct sum. In other words, the simple bivectors, when multiplied by the projectors $P_{\pm}$, are related to isoclinic rotations.

3.5 The spacetime algebra

As its name suggests, this geometric algebra is based on spacetime and its four elementary vectors, one -e_0-with a square equal to 1 and the other three -e_1, e_2, e_3- with a square equal to -1.

Clearly, this algebra is none other than Clifford algebra $Cl(1,3)$ and it's very close to $Cl(4)$. Like this last one, it is generated by 16 elements:

- 1 scalar (grade 0): 1
- 4 vectors (grade 1): e_0, e_1, e_2, e_3
- 6 bivectors (grade 2): e_0e_1, e_0e_2, e_0e_3, e_1e_2, e_3e_1, e_2e_3
- 4 pseudovectors (grade 3): $e_0e_1e_2$, $e_0e_3e_1$, $e_0e_2e_3$, $e_1e_2e_3$
- 1 pseudoscalar (grade 4): $e_0e_1e_2e_3$

Some of these elements have a square equal to 1:

$$1,\ e_0,\ e_0e_1,\ e_0e_2,\ e_0e_3,\ e_1e_2e_3$$

They are said to be timelike. Others have a square equal to -1:

$$e_1,\ e_2,\ e_3,\ e_1e_2,\ e_3e_1,\ e_2e_3,\ e_0e_1e_2,\ e_0e_3e_1,\ e_0e_2e_3,\ e_0e_1e_2e_3$$

They are spacelike elements.

A linear combination can give rise to lightlike objects. For instance, we have

$$(e_0 + e_1)^2 = 0$$

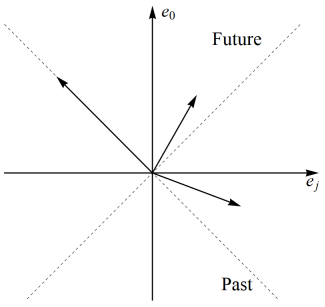


Figure 3.6: Timelike, spacelike and lightlike 1-vectors

A matrix form is available[5]:

$$1 \to \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \tag{3.52}$$

$$e_0 \to \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}, e_j \to \begin{pmatrix} 0 & \sigma j \\ -\sigma j & 0 \end{pmatrix}, \ j = 1, 2, 3 \tag{3.53}$$

$$e_0 e_j \to \begin{pmatrix} -\sigma j & 0 \\ 0 & \sigma j \end{pmatrix}, e_j e_k \to -i\varepsilon_{jkl} \begin{pmatrix} \sigma l & 0 \\ 0 & \sigma l \end{pmatrix}, \ j, k, l = 1, 2, 3 \tag{3.54}$$

$$e_0 e_j e_k \to -i\varepsilon_{jkl} \begin{pmatrix} 0 & \sigma l \\ \sigma l & 0 \end{pmatrix}, e_1 e_2 e_3 \to \begin{pmatrix} 0 & -iI \\ iI & 0 \end{pmatrix}, \ j, k, l = 1, 2, 3 \tag{3.55}$$

$$e_0 e_1 e_2 e_3 \to \begin{pmatrix} iI & 0 \\ 0 & -iI \end{pmatrix} \tag{3.56}$$

If we compare with the realization of $Cl(4)$, we find that the basis vectors e_j have been multiplied here by $-i$.

Of these 16 elements, 8 form a subalgebra. These elements are the even part of the spacetime algebra:

$$1, e_0 e_j, e_j e_k, e_0 e_1 e_2 e_3, \ j, k = 1, 2, 3 \tag{3.57}$$

It is obvious to see because these are the only "block"-diagonal matrices. We also have

$$(e_0 e_j)^2 = 1 \, , \, e_0 e_j e_0 e_k = -e_j e_k$$

so that we can identify this subalgebra with the space algebra discussed earlier:

$$1 \leftrightarrow 1, \sigma j \leftrightarrow -e_0 e_j, \sigma j \sigma k \leftrightarrow -e_j e_k, \sigma 1 \sigma 2 \sigma 3 \leftrightarrow e_0 e_1 e_2 e_3 \tag{3.58}$$

It is this observation that leads to what is called spacetime split. A spacetime vector $X = x^\mu e_\mu$ can be split according to

[5] It is noted here that the pseudoscalar unit is no longer totally assimilable to the imaginary i, in the sense that it does not commute with each of the sixteen elements of the spacetime algebra.

$$Xe_0 = x^0 + x^j\, e_j e_0 = t + x^j\, \sigma j$$

in the observer's frame e_0.

The most general element of this even part will be written as

$$M_0 =< A >_0 + < A >_2 + < A >_4$$

i.e.

$$M_0 = c_1 + c_2\, e_0e_1 + c_3\, e_0e_2 + c_4\, e_0e_3 + c_5\, e_1e_2 + c_6\, e_3e_1 + c_7\, e_2e_3 + c_8\, e_0e_1e_2e_3 \tag{3.59}$$

or

$$M_0 = \begin{pmatrix} a & b & 0 & 0 \\ c & d & 0 & 0 \\ 0 & 0 & d^* & -c^* \\ 0 & 0 & -b^* & a^* \end{pmatrix} \tag{3.60}$$

with

$$\begin{cases} a = c_1 - c_4 - ic_5 + ic_8 \\ b = -c_2 + ic_3 - c_6 - ic_7 \\ c = -c_2 - ic_3 + c_6 - ic_7 \\ d = c_1 + c_4 + ic_5 + ic_8 \end{cases} \tag{3.61}$$

This naturally leads to a first conjugation of $N_0 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$:

$$\widehat{N_0} \equiv \begin{pmatrix} d^* & -c^* \\ -b^* & a^* \end{pmatrix} \tag{3.62}$$

What is the physical meaning of this operation ?

To find out, let's take advantage of the fact that the even part of the space-time algebra coincides with the space algebra. Going from N_0 to $\widehat{N_0}$ or vice-versa leads to the following changes

$$c_2 \leftrightarrow -c_2,\ c_3 \leftrightarrow -c_3,\ c_4 \leftrightarrow -c_4,\ c_8 \leftrightarrow -c_8 \tag{3.63}$$

that is to say (see (3.58)-(3.59))

$$\sigma j \leftrightarrow -\sigma j$$

In other words, this conjugation amounts to inverting all the vectors of $\mathbb{R}^3$. It is therefore the equivalent of the parity operator, from the point of view of geometric algebra.

This is for the even part of the spacetime algebra. In what concerns the odd part, we have

$$M_1 =< A >_1 + < A >_3$$

that is to say

$$M_1 = d_1\, e_0 + d_2\, e_1 + d_3\, e_2 + d_4\, e_3 + d_5\, e_0e_1e_2 + d_6\, e_0e_3e_1 + d_7\, e_0e_2e_3 + d_8\, e_1e_2e_3 \tag{3.64}$$

It is equivalent to

$$M_1 = \begin{pmatrix} 0 & N_1 \\ \widehat{N_1} & 0 \end{pmatrix} \tag{3.65}$$

In general, a multivector in spacetime algebra is therefore written

$$M = M_0 + M_1 = \begin{pmatrix} N_0 & N_1 \\ \widehat{N_1} & \widehat{N_0} \end{pmatrix} \tag{3.66}$$

Another remarkable point is that the even part of the spacetime algebra can also be written as the sum of a vector of the space algebra and its dual. Indeed, Eq. (3.59) can also be written as

$$M_0 = c_1 + c_2\, e_0e_1 + c_3\, e_0e_2 + c_4\, e_0e_3 + e_0e_1e_2e_3(c_8 + c_7\, e_0e_1 + c_6\, e_0e_2 + c_5\, e_0e_3) \tag{3.67}$$

or

$$M_0 = x + y\, e_0e_1e_2e_3 \tag{3.68}$$

This leads us to define four types of sets in spacetime algebra:

1. The even part (E): $M_0 = x + y\, e_0e_1e_2e_3$
2. The odd part (O): $M_0\, e_0 = x\, e_0 - y\, e_1e_2e_3$
3. The dual even part (DE): $M_0\, e_0e_1e_2e_3 = -y + x\, e_0e_1e_2e_3$
4. The dual odd part (DO): $M_0\, e_1e_2e_3 = y\, e_0 + x\, e_1e_2e_3$

3.6 Some tedious but necessary calculations

Before continuing, it is useful to present some results, all related to conjugations.

3.6.1 The conjugation $\widehat{M}$

We have already mentioned a first type of conjugation: that which consists of reversing the direction of all the basis vectors of the space algebra. This one is called the grade involution and is noted using the symbol $\widehat{M}$.

This means that for the even multivector (3.59), we have

$$N_0 = (c_1+ic_8)I-(c_2+ic_7)\sigma 1-(c_3+ic_6)\sigma 2-(c_4+ic_5)\sigma 3 = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (3.69)$$

and

$$\widehat{N_0} = (c_1-ic_8)I+(c_2-ic_7)\sigma 1+(c_3-ic_6)\sigma 2+(c_4-ic_5)\sigma 3 = \begin{pmatrix} d^* & -c^* \\ -b^* & a^* \end{pmatrix} \quad (3.70)$$

When applied to the even part M_0 of the spacetime algebra, it is clear that this conjugation has no effect:

$$\widehat{M_0} = \begin{pmatrix} N_0 & 0 \\ 0 & \widehat{N_0} \end{pmatrix} =< A >_0 + < A >_2 + < A >_4 \quad (3.71)$$

However, it changes the sign of the odd part of this algebra:

$$\widehat{M_1} = -\begin{pmatrix} 0 & N_1 \\ \widehat{N_1} & 0 \end{pmatrix} = - < A >_1 - < A >_3 \quad (3.72)$$

Moreover, it is easy to convince oneself that

$$\widehat{M_A M_B} = \widehat{M_A}\widehat{M_B} \quad (3.73)$$

for any matrices M_A and M_B of dimension 2.

3.6.2 The conjugation $\widetilde{M}$

The second conjugation or reverse conjugation $\widetilde{M}$ corresponds to permuting the order of the vectors within the wedge product in the sense that

$e_j e_k \to e_k e_j = -e_j e_k$; $e_j e_k e_l \to e_l e_k e_j = -e_j e_k e_l$; $e_0 e_1 e_2 e_3 \to e_3 e_2 e_1 e_0 = e_0 e_1 e_2 e_3$

In other words, we have (see Eqs. (3.59) and (3.64))

$$\widetilde{M_0} =< A >_0 - < A >_2 + < A >_4 \tag{3.74}$$

$$\widetilde{M_1} =< A >_1 - < A >_3 \tag{3.75}$$

This has various implications, including

$$\widetilde{M_0} = \begin{pmatrix} \overline{N_0} & 0 \\ 0 & \widetilde{N_0} \end{pmatrix} \tag{3.76}$$

with

$$\overline{N_0} \equiv \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \widetilde{N_0} \equiv \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} \tag{3.77}$$

Similarly, we have

$$\widetilde{M_1} = \begin{pmatrix} 0 & \widetilde{N_1} \\ \overline{N_1} & 0 \end{pmatrix} \tag{3.78}$$

A final remark is that

$$\widetilde{M_A M_B} = \widetilde{M_B}\widetilde{M_A} \tag{3.79}$$

3.6.3 The conjugation $\overline{M}$

This third conjugation, referred to as the Clifford conjugation, arises naturally, as we have seen above. It results, in fact, from the successive applications, whatever the order is, of the two preceding conjugations.

So we have

$$\overline{M_0} =< A >_0 - < A >_2 + < A >_4 \tag{3.80}$$

$$\overline{M_1} = - < A >_1 + < A >_3 \tag{3.81}$$

and

$$\overline{M_0} = \begin{pmatrix} \overline{N_0} & 0 \\ 0 & \widetilde{N_0} \end{pmatrix}, \overline{M_1} = - \begin{pmatrix} 0 & \widetilde{N_1} \\ \overline{N_1} & 0 \end{pmatrix} \tag{3.82}$$

Of course, we also get

$$\overline{M_A M_B} = \overline{M_B}\,\overline{M_A} \tag{3.83}$$

3.7 Relativistic invariance through the spacetime algebra

As in the space algebra approach, a position vector is defined as a linear combination of the basis elements:

$$X \equiv x^\mu e_\mu = \begin{pmatrix} 0 & 0 & x^0+x^3 & x^1-ix^2 \\ 0 & 0 & x^1+ix^2 & x^0-x^3 \\ x^0-x^3 & -x^1+ix^2 & 0 & 0 \\ -x^1-ix^2 & x^0+x^3 & 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} 0 & x \\ \widehat{x} & 0 \end{pmatrix} \tag{3.84}$$

This is therefore an element of the odd part of the spacetime algebra.

It contains the position vector of the space algebra:

$$x = x^0 + \overrightarrow{x}.\overrightarrow{\sigma} = \begin{pmatrix} x^0+x^3 & x^1-ix^2 \\ x^1+ix^2 & x^0-x^3 \end{pmatrix} \tag{3.85}$$

as well as its conjugate counterpart, which clearly corresponds to an action of parity

$$\widehat{x} = x^0 - \overrightarrow{x}.\overrightarrow{\sigma} = \begin{pmatrix} x^0-x^3 & -x^1+ix^2 \\ -x^1-ix^2 & x^0+x^3 \end{pmatrix} \tag{3.86}$$

Note that the spacetime metric is then nothing other than the determinant of x or $\hat{x}$, indifferently:

$$det(x) = det(\widehat{x}) = x^\mu x_\mu$$

Before going any further, it is useful to note the following observations.

First, we can see that

$$e_0 = \gamma^0,\ e_j = \gamma^j\ (j = 1, 2, 3) \tag{3.87}$$

Since the metric is different from that of the space algebra, this leads us to write that (compare with (3.28)-(3.30)):

$$X\prime = \exp(-e_1e_2\,\frac{\theta}{2})\,X\,\exp(e_1e_2\,\frac{\theta}{2}) \tag{3.88}$$

$$\Leftrightarrow X\prime = x^0e_0+(x^1\cos(\theta)+x^2\sin(\theta))e_1+(-x^1\sin(\theta)+x^2\cos(\theta))e_2+x^3e_3 \tag{3.89}$$

and

$$X\prime = \exp(e_1e_2\,\frac{\theta}{2})\,X\,\exp(-e_1e_2\,\frac{\theta}{2}) \tag{3.90}$$

$$\Leftrightarrow X\prime = x^0e_0+(x^1\cos(\theta)-x^2\sin(\theta))e_1+(-x\sin(\theta)+x^2\cos(\theta))e_2+x^3e_3 \tag{3.91}$$

The metric therefore implied a change in the direction of rotation. Here is an illustration on $X = e_2$:

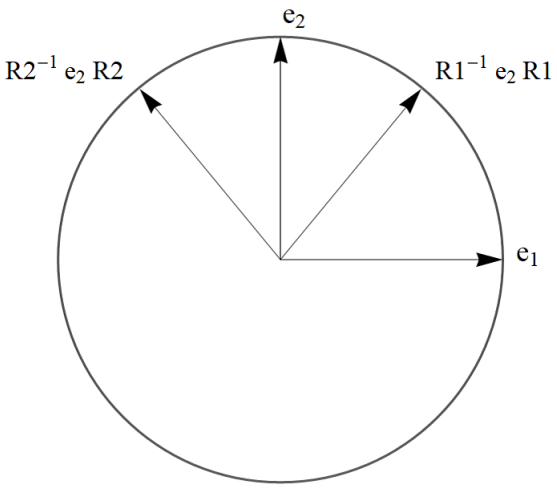


Figure 3.7: Actions of the rotors $R1 = \exp(e_1e_2\,\frac{\theta}{2})$ and $R2 = \exp(-e_1e_2\,\frac{\theta}{2})$ on e_2

Let's now consider hyperbolic rotations or boosts.
As we have seen in Chapter 1, boosts are characterized by (1.81) or

$$\begin{cases} t\prime = \frac{1}{\sqrt{1-v^2}}\,(t - v\,x) \\ x\prime = \frac{1}{\sqrt{1-v^2}}\,(x - v\,t) \end{cases}$$

when (1+1) dimensions are considered and $c = 1$.

Take now the timelike vector

$$u = 2\, e_0$$

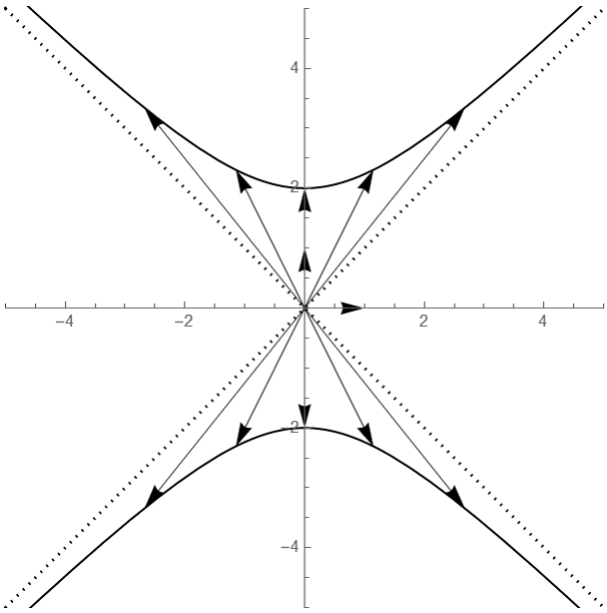
A boost transforms it into

$$u\prime = \frac{2}{\sqrt{1-v^2}}\, e_0 - \frac{2v}{\sqrt{1-v^2}}\, e_1$$

So, if $v = \pm 0.5$ or $v = \pm 0.8$, for example, we respectively have

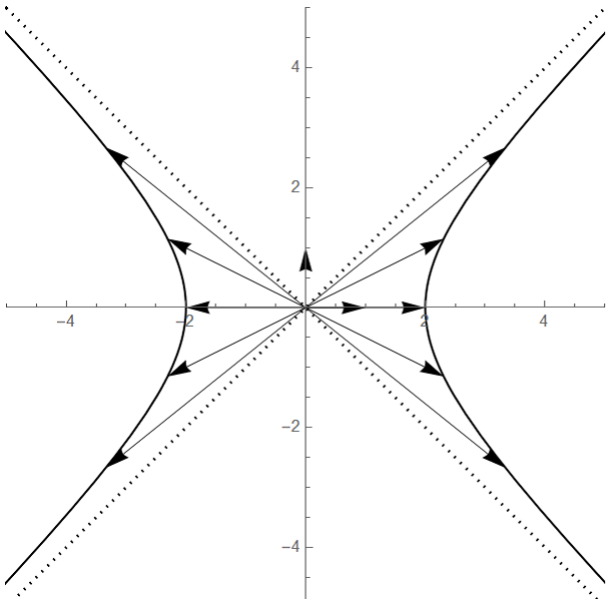
$$u\prime = 2.31\, e_0 \mp 1.15\, e_1,\ u\prime = 3.33\, e_0 \mp 2.67\, e_1$$

These images are represented in the graph below, and, of course, we can confirm that a boost does not change the nature of the initial vector in any way: starting from a timelike vector, directed towards the future, we obtain a set of timelike vectors, also directed towards the future. Their extremities form the upper half of the equilateral hyperbola with equation $y^2 - x^2 = 4$:



From a timelike vector directed towards the past, we get a set of vectors with the same features.

It is also true if we start with a spacelike vector:



From the perspective of geometric algebra, a boost is considered a rotation in the plane bounded by the timelike vector e_0 and a spacelike vector e_j.

So, we respectively have

$$X\prime = \exp(-e_0 e_1 \frac{\varphi}{2})\, X\, \exp(e_0 e_1 \frac{\varphi}{2}) \tag{3.92}$$

$$\Leftrightarrow X\prime = (x^0 \cosh(\varphi) + x^1 \sinh(\varphi))e_0 + (x^0 \sinh(\varphi) + x^1 \cosh(\varphi))e_1 + x^2 e_2 + x^3 e_3 \tag{3.93}$$

and

$$X\prime = \exp(e_0 e_1 \frac{\varphi}{2})\, X\, \exp(-e_0 e_1 \frac{\varphi}{2}) \tag{3.94}$$

$$\Leftrightarrow X\prime = (x^0 \cosh(\varphi) - x^1 \sinh(\varphi))e_0 + (-x^0 \sinh(\varphi) + x^1 \cosh(\varphi))e_1 + x^2 e_2 + x^3 e_3 \tag{3.95}$$

Once again, an illustration will be welcome:

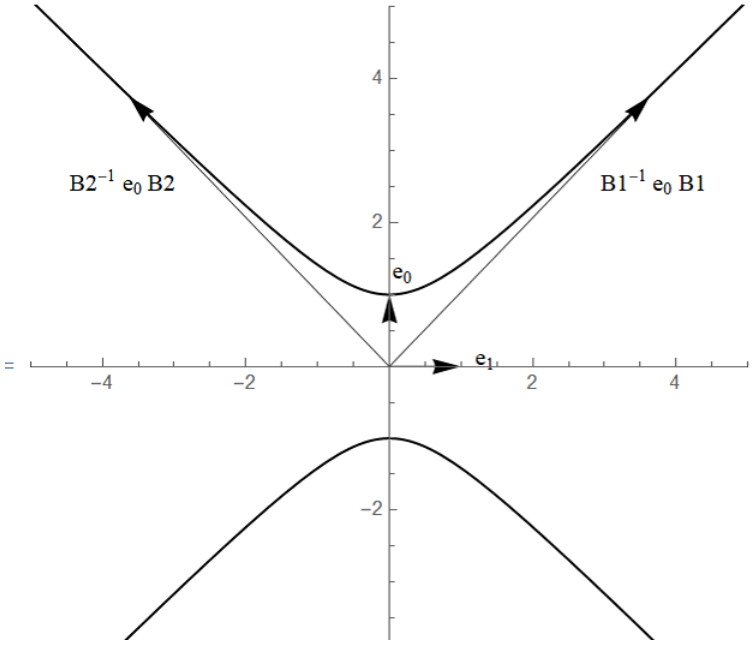


Figure 3.8: Actions of the rotors $B1 = \exp(e_0 e_1 \frac{\varphi}{2})$ and $B2 = \exp(-e_0 e_1 \frac{\varphi}{2})$ on e_0

We draw inspiration from the rotors in (3.88)-(3.90) to propose

$$R_1(\theta) = \exp(-e_2 e_3 \frac{\theta}{2}),\ R_2(\theta) = \exp(-e_3 e_1 \frac{\theta}{2}),\ R_3(\theta) = \exp(-e_1 e_2 \frac{\theta}{2}) \quad (3.96)$$

Indeed, with the realization (3.54), it is clear that

$$R_j(\theta) = \begin{pmatrix} \cos(\frac{\theta}{2}) + i\sigma j \sin(\frac{\theta}{2}) & 0 \\ 0 & \cos(\frac{\theta}{2}) + i\sigma j \sin(\frac{\theta}{2}) \end{pmatrix},\ j = 1,\ 2,\ 3 \quad (3.97)$$

It has the form (3.66) with $N_1 = 0$ and

$$N_0 = \cos(\frac{\theta}{2}) + i\sigma j \sin(\frac{\theta}{2}) \quad (3.98)$$

So, for example, when $j = 1$, we are led to

$$N_0 = \begin{pmatrix} \cos(\frac{\theta}{2}) & i \sin(\frac{\theta}{2}) \\ i \sin(\frac{\theta}{2}) & \cos(\frac{\theta}{2}) \end{pmatrix} \quad (3.99)$$

and we find the rotation in the plane (e_2, e_3) marked by the matrix $R_1(\theta)$ given in (1.86).

We can now turn our attention to the boosts that are associated with

$$B_1(\varphi) = \exp(e_0 e_1 \frac{\varphi}{2}),\ B_2(\varphi) = \exp(e_0 e_2 \frac{\varphi}{2}),\ B_3(\varphi) = \exp(e_0 e_3 \frac{\varphi}{2}), \quad (3.100)$$

Using the realization (3.54), we obtain

$$B_j(\varphi) = \begin{pmatrix} \cosh(\frac{\varphi}{2}) - \sigma j \sinh(\frac{\varphi}{2}) & 0 \\ 0 & \cosh(\frac{\varphi}{2}) + \sigma j \sinh(\frac{\varphi}{2}) \end{pmatrix},\ j = 1,\ 2,\ 3 \quad (3.101)$$

In other words, it is once again an even element of the spacetime algebra. We have

$$N_{0,j} = \cosh(\frac{\varphi}{2}) - \sigma j \sinh(\frac{\varphi}{2}) \quad (3.102)$$

and, for $j = 1$ for example, we recover the hyperbolic rotation in the plane (e_0, e_1) marked by the matrix $B_1(\varphi)$ given in (1.88).

By a similar line of reasoning for $R_2(\theta)$, $R_3(\theta)$, $B_2(\varphi)$ and $B_3(\varphi)$,we find the six Lorentz transformations, as expected. The six matrices $N_{0,j}$ are 2x2 matrices with complex elements and whose determinant is equal to 1. In other words, these are the matrices of the group $SL(2, C)$. This is not surprising since the group $SL(2, C)$ is the universal covering of the Lorentz group.

What's more surprising is that there's a seventh transformation! It is related to the pseudoscalar unit since it is defined by

$$R(\lambda) = \exp(-e_0 e_1 e_2 e_3 \frac{\lambda}{2}) \quad (3.103)$$

It is easy to see that

$$X\prime = R^{-1}(\lambda)\, X\, R(\lambda) = \cos(\lambda)\, X + \sin(\lambda)\, IX \quad (3.104)$$

Clearly, we have

$$R(\lambda) = \begin{pmatrix} N_0 & 0 \\ 0 & \widehat{N_0} \end{pmatrix},\ N_0 = \begin{pmatrix} \exp(-\frac{i}{2}\lambda) & 0 \\ 0 & \exp(-\frac{i}{2}\lambda) \end{pmatrix} \quad (3.105)$$

Since it is not defined through bivectors and, therefore, not related to SL(2,C), N_0 has a determinant different from 1, in general.

Eq. (3.104) means that

$$x\prime = \exp(i\lambda)\, x \quad (3.106)$$

In other words, this -unexpected- seventh transformation is nothing but a complex rotation of the position.

All the Lorentz transformations can be summarized by

$$X\prime = \begin{pmatrix} \overline{N_0} & 0 \\ 0 & \widetilde{N_0} \end{pmatrix} X \begin{pmatrix} N_0 & 0 \\ 0 & \widehat{N_0} \end{pmatrix} \tag{3.107}$$

i.e.

$$x\prime = \overline{N_0}\, x\, \widehat{N_0} \tag{3.108}$$

The seventh one corresponds to

$$X\prime = \begin{pmatrix} \widetilde{N_0} & 0 \\ 0 & \overline{N_0} \end{pmatrix} X \begin{pmatrix} N_0 & 0 \\ 0 & \widehat{N_0} \end{pmatrix} \tag{3.109}$$

or

$$x\prime = \widetilde{N_0}\, x\, \widehat{N_0} \tag{3.110}$$

We can, equivalently, analyze transformations based on the odd part of the spacetime algebra[6]. This time, we have

$$x\prime = \overline{N_1}\, \widehat{x}\, \widehat{N_1} \tag{3.111}$$

or

$$x\prime = \widetilde{N_1}\, \widehat{x}\, \widehat{N_1} \tag{3.112}$$

instead of (3108) and (3.110). Taking (see (3.98) and the following)

$$N_{1,j} = \cos(\frac{\theta}{2}) + i\sigma j\, \sin(\frac{\theta}{2}),\ j = 1, 2, 3 \tag{3.113}$$

$$N_{1,j} = \cosh(\frac{\varphi}{2}) - \sigma j\, \sinh(\frac{\varphi}{2}),\ j = 1, 2, 3 \tag{3.114}$$

and

$$N_1 = \exp(-\frac{i}{2}\lambda)I \tag{3.115}$$

we see that (3.113)-(3.115) lead us to the same transformations as those mentioned above, up to the fact that

$$x^1 \to -x^1,\ x^2 \to -x^2,\ x^3 \to -x^3$$

[6]Debergh, N. and Petit, J.P. (2023) On Spacetime Algebra and Its Relation with Negative Mass. Reports on Mathematical Physics, 91, 165-181. DOI: 10.1016/S0034-4877(23)00023-X

This is obviously due to the fact that x was replaced by $\widehat{x}$. This means that going from an even transformation to an odd transformation actually involves the parity operator, in a natural way. In other words, and based on the table 1.1, we are dealing here with improper orthochronous transformations.

Is there therefore a way to highlight retrochronous transformations? And how?

To do this, one only needs to consider Hodge's duals. But not only that. In fact, we are going to consider the complex version of the spacetime algebra.

Indeed, an even transformation when modified by the pseudoscalar unit (3.56)

$$M_0 \to -ie_0e_1e_2e_3M_0$$

is such that

$$x\prime = -\overline{N_0}\, x\, \widehat{N_0} \tag{3.116}$$

It is equivalent to ask for

$$x^0 \to -x^0,\, x^1 \to -x^1,\, x^2 \to -x^2,\, x^3 \to -x^3 \tag{3.117}$$

So, it is the PT transformation that is related to this case. Referring once again to the table 1.1, we can therefore affirm that it is the proper retrochronous transformations that are underlying here.

What is missing then are the improper retrochronous transformations which are obtained from the only time-reversal operator T. In the framework of the spacetime algebra, they are found by considering the dual odd transformations.

It's already time for a recap.

We have seen that the spacetime agebra highlights the six usual Lorentz transformations but also a seventh transformation whose nature is a complex rotation on coordinates. Moreover, improper versions via odd transformations, as well as retrochronous versions through dual transformations, appear here in a very natural way.

Let's analyze this more precisely by observing a representative of some of the transformations:

Proper/orthochronous
$R_1 = \begin{pmatrix} \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} & 0 \\ 0 & \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} \end{pmatrix}$
$B_1 = \begin{pmatrix} \cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} & 0 \\ 0 & \cosh\frac{\varphi}{2} + \sigma 1 \sinh\frac{\varphi}{2} \end{pmatrix}$
$R = \begin{pmatrix} \exp(-i\frac{\lambda}{2}) & 0 \\ 0 & \exp(i\frac{\lambda}{2}) \end{pmatrix}$
Proper/retrochronous
$R_1 = \begin{pmatrix} \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} & 0 \\ 0 & -\cos\frac{\theta}{2} - i\sigma 1 \sin\frac{\theta}{2} \end{pmatrix}$
$B_1 = \begin{pmatrix} \cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} & 0 \\ 0 & -\cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} \end{pmatrix}$
$R = \begin{pmatrix} \exp(-i\frac{\lambda}{2}) & 0 \\ 0 & -\exp(i\frac{\lambda}{2}) \end{pmatrix}$
Improper/orthochronous
$R_1 = \begin{pmatrix} 0 & \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} & 0 \end{pmatrix}$
$B_1 = \begin{pmatrix} 0 & \cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} \\ \cosh\frac{\varphi}{2} + \sigma 1 \sinh\frac{\varphi}{2} & 0 \end{pmatrix}$
$R = \begin{pmatrix} 0 & \exp(-i\frac{\lambda}{2}) \\ \exp(i\frac{\lambda}{2}) & 0 \end{pmatrix}$
Improper/retrochronous
$R_1 = \begin{pmatrix} 0 & \cos\frac{\theta}{2} + i\sigma 1 \sin\frac{\theta}{2} \\ -\cos\frac{\theta}{2} - i\sigma 1 \sin\frac{\theta}{2} & 0 \end{pmatrix}$
$B_1 = \begin{pmatrix} 0 & \cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} \\ -\cosh\frac{\varphi}{2} - \sigma 1 \sinh\frac{\varphi}{2} & 0 \end{pmatrix}$
$R = \begin{pmatrix} 0 & \exp(-i\frac{\lambda}{2}) \\ -\exp(i\frac{\lambda}{2}) & 0 \end{pmatrix}$

The first observation is that each of these matrices has a determinant equal to 1. We are therefore dealing with the Lie group $SL(4,C)$. It clearly generalizes the group $SL(2,C)$ which underlies the usual Lorentz transformations.

The second observation is that the proper transformations clearly involve two simultaneous rotations. To obtain the retrochronous case from the orthochronous one, we must perform

$$\theta \to \theta + \pi,\ \varphi \to \varphi + i\pi,\ \lambda \to \lambda + \pi$$

in the plane (e_2, e_3) while the transformations in (e_0, e_1) (for the specific cases we have considered) are invariant.

A third observation is that geometric algebra allows us to consider, through the odd part of the spacetime algebra, the discrete symmetries P and T. The fact of considering the odd part ensures in fact that we keep a determinant equal to 1 whereas this is impossible in the traditional approach.

Two tables will be particularly useful for summarizing the results at this stage.

Proper/orthochronous		Improper/orthochronous
Even		Odd
$x\prime = \overline{N_0}\, x\, \widehat{N_0}$	Lorentz	$x\prime = \overline{N_1}\, \widehat{x}\, \widehat{N_1}$
$x\prime = \widetilde{N_0}\, x\, \widehat{N_0}$		$x\prime = \widetilde{N_1}\, \widehat{x}\, \widehat{N_1}$
Proper/retrochronous		Improper/retrochronous
Dual even		Dual odd
$x\prime = -\overline{N_0}\, x\, \widehat{N_0}$	Lorentz	$x\prime = -\overline{N_1}\, \widehat{x}\, \widehat{N_1}$
$x\prime = -\widetilde{N_0}\, x\, \widehat{N_0}$		$x\prime = -\widetilde{N_1}\, \widehat{x}\, \widehat{N_1}$

	Proper/orthochronous	Improper/orthochronous
R_1	$\cos(\frac{\theta}{2}) - \sin(\frac{\theta}{2})\, e_2e_3$	$\cos(\frac{\theta}{2})\, e_0 - \sin(\frac{\theta}{2})\, e_0e_2e_3$
R_2	$\cos(\frac{\theta}{2}) - \sin(\frac{\theta}{2})\, e_3e_1$	$\cos(\frac{\theta}{2})\, e_0 - \sin(\frac{\theta}{2})\, e_0e_3e_1$
R_3	$\cos(\frac{\theta}{2}) - \sin(\frac{\theta}{2})\, e_1e_2$	$\cos(\frac{\theta}{2})\, e_0 - \sin(\frac{\theta}{2})\, e_0e_1e_2$
B_1	$\cosh(\frac{\varphi}{2}) + \sinh(\frac{\varphi}{2})\, e_0e_1$	$\cosh(\frac{\varphi}{2})\, e_0 - \sinh(\frac{\varphi}{2})\, e_1$
B_2	$\cosh(\frac{\varphi}{2}) + \sinh(\frac{\varphi}{2})\, e_0e_2$	$\cosh(\frac{\varphi}{2})\, e_0 - \sinh(\frac{\varphi}{2})\, e_2$
B_3	$\cosh(\frac{\varphi}{2}) + \sinh(\frac{\varphi}{2})\, e_0e_3$	$\cosh(\frac{\varphi}{2})\, e_0 - \sinh(\frac{\varphi}{2})\, e_3$
R	$\cos(\frac{\lambda}{2}) - \sin(\frac{\lambda}{2})\, e_0e_1e_2e_3$	$\cos(\frac{\lambda}{2})\, e_0 + \sin(\frac{\lambda}{2})\, e_1e_2e_3$
	Proper/retrochronous	Improper/retrochronous
R_1	$-i\cos(\frac{\theta}{2})\, e_0e_1e_2e_3 - i\sin(\frac{\theta}{2})\, e_0e_1$	$i\cos(\frac{\theta}{2})\, e_1e_2e_3 + i\sin(\frac{\theta}{2})\, e_1$
R_2	$-i\cos(\frac{\theta}{2})\, e_0e_1e_2e_3 - i\sin(\frac{\theta}{2})\, e_0e_2$	$i\cos(\frac{\theta}{2})\, e_1e_2e_3 + i\sin(\frac{\theta}{2})\, e_2$
R_3	$-i\cos(\frac{\theta}{2})\, e_0e_1e_2e_3 - i\sin(\frac{\theta}{2})\, e_0e_3$	$i\cos(\frac{\theta}{2})\, e_1e_2e_3 + i\sin(\frac{\theta}{2})\, e_3$
B_1	$-i\cosh(\frac{\varphi}{2})\, e_0e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_2e_3$	$i\cosh(\frac{\varphi}{2})\, e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_0e_2e_3$
B_2	$-i\cosh(\frac{\varphi}{2})\, e_0e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_3e_1$	$i\cosh(\frac{\varphi}{2})\, e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_0e_3e_1$
B_3	$-i\cosh(\frac{\varphi}{2})\, e_0e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_1e_2$	$i\cosh(\frac{\varphi}{2})\, e_1e_2e_3 - i\sinh(\frac{\varphi}{2})\, e_0e_1e_2$
R	$-i\cos(\frac{\lambda}{2})\, e_0e_1e_2e_3 - i\sin(\frac{\lambda}{2})$	$i\cos(\frac{\lambda}{2})\, e_1e_2e_3 - i\sin(\frac{\lambda}{2})\, e_0$

It is remarkable that the spacetime algebra contains within itself the set of transformations of the complete Lorentz group.

Furthermore, if we denote M as the generic matrix of the seven proper orthochronous transformations, we note that

- the improper orthochronous transformations are obtained through $M\, e_0$, i.e. under the action, on the right, of the unitary parity operator P;
- the proper retrochronous transformations are obtained through $-iM\, e_0e_1e_2e_3$, i.e. under the action, on the right, of the chirality operator, also responsible for the change of sign in the mass;
- the improper retrochronous transformations are obtained through $iM\, e_1e_2e_3$, i.e. under the action, on the right, of the unitary time-reversal operator T;

One final observation to conclude this Section.

The elements of the dual, associated with retrochronous transformations, require consideration of a complex extension of the spacetime algebra.

This is a development quite similar to what was proposed in (3.23)-(3.26) in the sense that we leave $Cl(1,3)$ to consider $Cl(2,3)$ through the new vector basis

$$e_0\prime = e_0,\ e_1\prime = e_1,\ e_2\prime = e_2,\ e_3\prime = e_3,\ e_4\prime = ie_0e_1e_2e_3 \tag{3.118}$$

The pseudovectors now read

$$e_0\prime e_1\prime e_2\prime e_3\prime = e_0e_1e_2e_3,\ e_0\prime e_1\prime e_2\prime e_4\prime = -ie_3 \tag{3.119}$$

and

$$e_0\prime e_1\prime e_3\prime e_4\prime = ie_2,\ e_0\prime e_2\prime e_3\prime e_4\prime == -ie_1,\ e_1\prime e_2\prime e_3\prime e_4\prime = -ie_0 \tag{3.120}$$

The algebra spanned by the corresponding bivectors is obviously $so(2,3)$. So there are four additional transformations, besides the Lorentz transformations: a circular rotation and three hyperbolic rotations.

Note that these additional transformations have nothing to do with those mentioned in the Section 1.10 devoted to the complex extension of the Poincaré group. They encompass here a fifth dimension.

For the transformations discussed in the first Chapter, we must go up to $Cl(2,6)$, whose 28 bivectors generate $so(2,6)$. We then recover the set of transformations of Section 1.10. For example,

$$\cos(\frac{\theta}{2}) - \sin(\frac{\theta}{2})\, e_0e_4$$

leads to (1.337), where e_0 is associated with real time while e_4 is related to imaginary time.

3.8 The Dirac-Hestenes equation

In geometric algebra, it is common practice to replace the column matrix describing the wavefunction with a square matrix. More specifically, the wavefunction is a multivector in the even part of the spacetime algebra, being of the form (3.60):

$$\psi_0 = \begin{pmatrix} \psi^1 & -\psi^{4*} & 0 & 0 \\ \psi^2 & \psi^{3*} & 0 & 0 \\ 0 & 0 & \psi^3 & -\psi^{2*} \\ 0 & 0 & \psi^4 & \psi^{1*} \end{pmatrix} \tag{3.121}$$

The next step is to transform the Dirac equation (2.85) concerning a bispinor into a new version that would act on the matrix (3.121).

To that end, we note that

$$\gamma^0\psi = \begin{pmatrix} \psi^3 \\ \psi^4 \\ \psi^1 \\ \psi^2 \end{pmatrix} \Leftrightarrow \gamma^0\psi_0\gamma^0 = \begin{pmatrix} \psi^3 & -\psi^{2*} & 0 & 0 \\ \psi^4 & \psi^{1*} & 0 & 0 \\ 0 & 0 & \psi^1 & -\psi^{4*} \\ 0 & 0 & \psi^2 & \psi^{3*} \end{pmatrix} \tag{3.122}$$

$$\gamma^1\psi = \begin{pmatrix} \psi^4 \\ \psi^3 \\ -\psi^2 \\ -\psi^1 \end{pmatrix} \Leftrightarrow \gamma^1\psi_0\gamma^0 = \begin{pmatrix} \psi^4 & \psi^{1*} & 0 & 0 \\ \psi^3 & -\psi^{2*} & 0 & 0 \\ 0 & 0 & -\psi^2 & -\psi^{3*} \\ 0 & 0 & -\psi^1 & \psi^{4*} \end{pmatrix} \tag{3.123}$$

$$\gamma^2\psi = \begin{pmatrix} -i\psi^4 \\ i\psi^3 \\ i\psi^2 \\ -i\psi^1 \end{pmatrix} \Leftrightarrow \gamma^2\psi_0\gamma^0 = \begin{pmatrix} -i\psi^4 & -i\psi^{1*} & 0 & 0 \\ i\psi^3 & -i\psi^{2*} & 0 & 0 \\ 0 & 0 & i\psi^2 & i\psi^{3*} \\ 0 & 0 & -i\psi^1 & i\psi^{4*} \end{pmatrix} \tag{3.124}$$

$$\gamma^3\psi = \begin{pmatrix} \psi^3 \\ -\psi^4 \\ -\psi^1 \\ \psi^2 \end{pmatrix} \Leftrightarrow \gamma^3\psi_0\gamma^0 = \begin{pmatrix} \psi^3 & -\psi^{2*} & 0 & 0 \\ -\psi^4 & -\psi^{1*} & 0 & 0 \\ 0 & 0 & -\psi^1 & \psi^{4*} \\ 0 & 0 & \psi^2 & \psi^{3*} \end{pmatrix} \tag{3.125}$$

$$i\psi = \begin{pmatrix} i\psi^1 \\ i\psi^2 \\ i\psi^3 \\ i\psi^4 \end{pmatrix} \Leftrightarrow \psi_0\gamma^2\gamma^1 = \begin{pmatrix} -i\psi^1 & -i\psi^{4*} & 0 & 0 \\ -i\psi^2 & i\psi^{3*} & 0 & 0 \\ 0 & 0 & -i\psi^3 & -i\psi^{2*} \\ 0 & 0 & -i\psi^4 & i\psi^{1*} \end{pmatrix} \tag{3.126}$$

If we multiply Eq. (2.85) by $-i$, we obtain, in direct correspondence via (3.122)-(3.126):

$$\gamma^\mu \partial_\mu \psi_0 \gamma^0 + e\gamma^\mu A_\mu \psi_0 \gamma^0 \gamma^2 \gamma^1 + m\,\psi_0 \gamma^2 \gamma^1 = 0$$

Then, we simply multiply on the right by $\gamma^2\gamma^1\gamma^0$ to obtain the final version of the Dirac equation in geometric algebra:

$$\boxed{\gamma^\mu(\partial_\mu \psi_0 \gamma^2 \gamma^1 - e\,A_\mu \psi_0) = m\,\psi_0 \gamma^0} \tag{3.127}$$

This equation is known as the Dirac-Hestenes equation or real Dirac equation.

We can also do the reverse process, namely starting from this equation and finding the Dirac equation in its traditional form. To do this, it suffices to consider u such that

$$\gamma^0 u = u,\ \gamma^2\gamma^1 u = i\,u,\ \psi_0\, u = \psi \Rightarrow u = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{3.128}$$

Then multiplying (3.127) by u on the right, we find the Dirac equation.

3.9 The real 2x2 Dirac equation

One remarkable thing is that since the wavefunction is expressed solely using the even part of the spacetime algebra, and this part is isomorphic to the space algebra, it is possible to reformulate the Dirac equation in terms of 2-dimensional matrices. To this end, we note that

$$\psi_0 = \begin{pmatrix} \varphi & 0 \\ 0 & \widehat{\varphi} \end{pmatrix} \tag{3.129}$$

if

$$\varphi \equiv \begin{pmatrix} \psi^1 & -\psi^{4*} \\ \psi^2 & \psi^{3*} \end{pmatrix} \tag{3.130}$$

The next step is to decompose the Dirac-Hestenes equation into blocks along its antidiagonal to hightlight the corresponding equation. Indeed, when multiplying (3.127) by $\gamma^2\gamma^1$ on the right, we have

$$\left(\begin{pmatrix} 0 & \nabla \\ \widehat{\nabla} & 0 \end{pmatrix} - e \begin{pmatrix} 0 & A \\ \widehat{A} & 0 \end{pmatrix} \right) \begin{pmatrix} \varphi & 0 \\ 0 & \widehat{\varphi} \end{pmatrix} = m \begin{pmatrix} \varphi & 0 \\ 0 & \widehat{\varphi} \end{pmatrix} \begin{pmatrix} \sigma 2 \sigma 1 & 0 \\ 0 & \sigma 2 \sigma 1 \end{pmatrix}$$

Here, we have defined[7]

$$\nabla \equiv \partial_0 - \overrightarrow{\sigma}.\overrightarrow{\nabla} \, , \, A \equiv A_0 - \overrightarrow{\sigma}.\overrightarrow{A} \tag{3.131}$$

and consequently

$$\widehat{\nabla} = \partial_0 + \overrightarrow{\sigma}.\overrightarrow{\nabla} \, , \, \widehat{A} \equiv A_0 + \overrightarrow{\sigma}.\overrightarrow{A} \tag{3.132}$$

So, Eq. (3.127) now reduces to

$$\boxed{\nabla \widehat{\varphi} = e \, A \, \widehat{\varphi} \, \sigma 2 \sigma 1 + m \, \varphi \, \sigma 2 \sigma 1} \tag{3.133}$$

and its conjugate

$$\boxed{\widehat{\nabla} \varphi = e \, \widehat{A} \, \varphi \, \sigma 2 \sigma 1 + m \, \widehat{\varphi} \, \sigma 2 \sigma 1} \tag{3.134}$$

However, we can also construct them starting from the Dirac equation:

$$(\gamma^{\mu}(\partial_{\mu} + ie \, A_{\mu}) + im) \begin{pmatrix} \xi \\ \eta \end{pmatrix} = 0 \tag{3.135}$$

which is equivalent to

$$\begin{cases} (\nabla + ie \, A) \, \eta + im \, \xi = 0 \\ (\widehat{\nabla} + ie \, \widehat{A}) \, \xi + im \, \eta = 0 \end{cases} \tag{3.136}$$

if the chiral realization is used.

A side remark: we had noted that the negative energy and mass solutions were obtained by applying the matrix γ^5 associated with chirality to the positive energy and mass solutions. In other words, via the changes

$$\xi \to \xi \, , \, \eta \to -\eta$$

This is particularly clear here since the equations (3.136) are invariant under these changes accompanied by $m \to -m$.

Now, if we take the complex conjugate of these equations (3.136) and multiply them on the left by $-i\sigma 2$, we get

[7] In the 4x4 formulation, ∇ is defined through $\nabla \equiv \gamma^{\mu} \partial_{\mu}$.

$$\begin{cases} (\widehat{\nabla} - ie\,\widehat{A})\,(-i\sigma 2\,\eta^*) - im\,(-i\sigma 2\,\xi^*) = 0 \\ (\nabla - ie\,A)\,(-i\sigma 2\,\xi^*) - im\,(-i\sigma 2\,\eta^*) = 0 \end{cases} \tag{3.137}$$

after noticing that

$$-i\sigma 2\,\nabla^* = \widehat{\nabla}\,(-i\sigma 2)\,,\; -i\sigma 2\,\widehat{\nabla}^* = \nabla\,(-i\sigma 2) \tag{3.138}$$

Taking care of both results (3.136)-(3.137), we have

$$\nabla \begin{pmatrix} \eta & -i\sigma 2\,\xi^* \end{pmatrix} + ie\,A \begin{pmatrix} \eta & i\sigma 2\,\xi^* \end{pmatrix} + im \begin{pmatrix} \xi & i\sigma 2\,\eta^* \end{pmatrix} \tag{3.139}$$

We conclude by noting that

$$\varphi = \begin{pmatrix} \xi & -i\sigma 2\,\eta^* \end{pmatrix} \tag{3.140}$$

so that Eq.(3.139) coincides with Eq. (3.133).

This equation (3.133) is called the real 2x2 Dirac equation.
There are a few observations to note at this stage.

- If we take

$$\varphi\prime = c\,\varphi \Rightarrow \widehat{\varphi}\prime = c^*\widehat{\varphi}$$

where c is a complex constant, we have

$$\nabla\widehat{\varphi}\prime = e\,A\,\widehat{\varphi}\prime\,\sigma 2\sigma 1 + m\,\frac{c^*}{c}\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.141}$$

An imaginary constant c thus allows us to highlight negative masses. This demonstrates once again the power of the pseudoscalar unit.

- If we ask for

$$\varphi\prime = c\,\varphi\,\sigma 3 \Rightarrow \widehat{\varphi}\prime = -c^*\widehat{\varphi}\,\sigma 3$$

we obtain

$$\nabla\widehat{\varphi}\prime = e\,A\,\widehat{\varphi}\prime\,\sigma 2\sigma 1 - m\,\frac{c^*}{c}\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.142}$$

This time, it is a real constant c that allows us to consider negative masses.

We therefore have two cases in which the mass becomes negative: $\varphi\prime = i\,\varphi$ and $\varphi\prime = \varphi\,\sigma 3$. At the level of the wavefunction ψ, both cases lead respectively to

$$\psi\prime = \gamma^5\,\psi,\; \psi\prime = -i\,\gamma^5\,\psi \tag{3.143}$$

These are two clearly equivalent results in the traditional approach to the Dirac equation, since it is only in geometric algebra that the imaginary i plays a distinctive role due to the simultaneous presence of φ and $\widehat{\varphi}$ within the same equation.

Let's move on to the next two results.

- If we have

$$\varphi\prime = c\,\varphi\,\sigma 1 \Rightarrow \widehat{\varphi}\prime = -c^*\widehat{\varphi}\,\sigma 1$$

the real 2x2 Dirac equation is now

$$\nabla\widehat{\varphi}\prime = -e\,A\,\widehat{\varphi}\prime\,\sigma 2\sigma 1 + m\,\frac{c^*}{c}\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.144}$$

So, once again, a negative mass is conceivable provided that we take c to be imaginary. But the novelty is that we are clearly dealing with a charge conjugation.

This is a result that we will also obtain in the following case.

- If we consider

$$\varphi\prime = c\,\varphi\,\sigma 2 \Rightarrow \widehat{\varphi}\prime = -c^*\widehat{\varphi}\,\sigma 2$$

we still have

$$\nabla\widehat{\varphi}\prime = -e\,A\,\widehat{\varphi}\prime\,\sigma 2\sigma 1 + m\,\frac{c^*}{c}\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.145}$$

Again, the two cases coincide at the level of the traditional approach, up to a multiplication by i, and give rise to

$$\psi\prime = \begin{pmatrix} 0 & 0 & 0 & -c \\ 0 & 0 & c & 0 \\ 0 & c^* & 0 & 0 \\ -c^* & 0 & 0 & 0 \end{pmatrix} \psi^*$$

This is an expected result since we recognize the γ^2 matrix followed by a complex conjugation or the same with the γ^5 matrix in addition if we change the sign of the mass with an imaginary c.

3.10 Symmetries of the real 2x2 Dirac equation

The Dirac equation is this set of four equations:

$$\begin{cases} (\partial_0 + \partial_3)\psi^1 + (\partial_1 - i\partial_2)\psi^2 + im\psi^3 = 0 \\ (\partial_0 + \partial_3)\psi^{4*} - (\partial_1 - i\partial_2)\psi^{3*} - im\psi^{2*} = 0 \\ (\partial_0 - \partial_3)\psi^2 + (\partial_1 + i\partial_2)\psi^1 + im\psi^4 = 0 \\ (\partial_0 - \partial_3)\psi^{3*} - (\partial_1 + i\partial_2)\psi^{4*} - im\psi^{1*} = 0 \end{cases} \tag{3.146}$$

as well as their complex conjugates.

It is already clear at this stage that the following transformations

$$\psi^1 \leftrightarrow -\psi^{4*},\ \psi^2 \leftrightarrow \psi^{3*} \tag{3.147}$$

leave these equations invariant. Take a moment to look at Table 2.2 and you will be convinced that these transformations link matter with positive mass and antimatter with positive mass.

At the level of $\varphi =$(3.130), this leads to

$$\varphi \leftrightarrow \varphi\,\sigma 1 \tag{3.148}$$

as already noticed.

Having made this initial observation, let us now consider the real 2x2 free Dirac equation

$$\nabla\widehat{\varphi} - m\,\varphi\,\sigma 2\sigma 1 = 0 \tag{3.149}$$

as well as its transformation

$$\nabla\prime\widehat{\varphi\prime} - m\prime\,\varphi\prime\,\sigma 2\sigma 1 = 0 \tag{3.150}$$

Knowing the transformations affecting the coordinates, and therefore $\nabla\prime$, we still need to find the values of $m\prime$ and

$$\varphi\prime \equiv \begin{pmatrix} \varphi^1 & \varphi^2 \\ \varphi^3 & \varphi^4 \end{pmatrix} \tag{3.151}$$

3.10.1 The even (proper) symmetries

In general, taking into account (3.69) and (3.108), we have

$$\begin{cases} x^0\prime + x^3\prime = (|a|^2 + |b|^2)x^0 + (a^*b + ab^*)x^1 + i(a^*b - ab^*)x^2 + (|a|^2 - |b|^2)x^3 \\ x^0\prime - x^3\prime = (|c|^2 + |d|^2)x^0 + (c^*d + cd^*)x^1 + i(c^*d - cd^*)x^2 + (|c|^2 - |d|^2)x^3 \\ x^1\prime - ix^2\prime = (ac^* + bd^*)x^0 + (ad^* + bc^*)x^1 + i(-ad^* + bc^*)x^2 + (ac^* - bd^*)x^3 \\ x^1\prime + ix^2\prime = (a^*c + b^*d)x^0 + (a^*d + b^*c)x^1 + i(a^*d - b^*c)x^2 + (a^*c - b^*d)x^3 \end{cases}$$

This implies

$$\begin{cases} \partial_0 - \partial_3 = (|b|^2 + |d|^2)\partial_0\prime + (b^*d + bd^*)\partial_1\prime + i(-b^*d + bd^*)\partial_2\prime + (|b|^2 - |d|^2)\partial_3\prime \\ \partial_0 + \partial_3 = (|a|^2 + |c|^2)\partial_0\prime + (a^*c + ac^*)\partial_1\prime + i(-a^*c + ac^*)\partial_2\prime + (|a|^2 - |c|^2)\partial_3\prime \\ \partial_1 + i\partial_2 = (ab^* + cd^*)\partial_0\prime + (b^*c + ad^*)\partial_1\prime + i(-b^*c + ad^*)\partial_2\prime + (ab^* - cd^*)\partial_3\prime \\ \partial_1 - i\partial_2 = (a^*b + c^*d)\partial_0\prime + (bc^* + a^*d)\partial_1\prime + i(bc^* - a^*d)\partial_2\prime + (a^*b - c^*d)\partial_3\prime \end{cases}$$

In other words, we have

$$\nabla' = N_0 \, \nabla \, \widetilde{N_0} \tag{3.152}$$

Let us now take the rotation $R_1(\theta)$ characterized by

$$a = d = \cos(\frac{\theta}{2}),\ b = c = -i\ \sin(\frac{\theta}{2}) \tag{3.153}$$

as an example. Eq. (3.150) reads

$$\begin{cases} (\partial_0 - \sin(\theta)\partial_2 - \cos(\theta)\partial_3)\varphi^{4*} + (\partial_1 - i\cos(\theta)\partial_2 + i\sin(\theta)\partial_3)\varphi^{2*} + im\prime\,\varphi^1 = 0 \\ (\partial_0 - \sin(\theta)\partial_2 - \cos(\theta)\partial_3)\varphi^{3*} + (\partial_1 - i\cos(\theta)\partial_2 + i\sin(\theta)\partial_3)\varphi^{1*} + im\prime\,\varphi^2 = 0 \\ (\partial_0 + \sin(\theta)\partial_2 + \cos(\theta)\partial_3)\varphi^{2*} + (\partial_1 + i\cos(\theta)\partial_2 - i\sin(\theta)\partial_3)\varphi^{4*} - im\prime\,\varphi^3 = 0 \\ (\partial_0 + \sin(\theta)\partial_2 + \cos(\theta)\partial_3)\varphi^{1*} + (\partial_1 + i\cos(\theta)\partial_2 - i\sin(\theta)\partial_3)\varphi^{3*} - im\prime\,\varphi^4 = 0 \end{cases}$$

All of this must be related to (3.149), which can only be done iff

$$\begin{cases} \varphi^1 = A\,(i\cos(\frac{\theta}{2})\,\psi^1 + \sin(\frac{\theta}{2})\,\psi^2) + B\,(\sin(\frac{\theta}{2})\,\psi^{3*} - i\cos(\frac{\theta}{2})\,\psi^{4*}) \\ \varphi^2 = \exp(i\,\chi)(-B^*\,(i\cos(\frac{\theta}{2})\,\psi^1 + \sin(\frac{\theta}{2})\,\psi^2) - A^*\,(\sin(\frac{\theta}{2})\,\psi^{3*} - i\cos(\frac{\theta}{2})\,\psi^{4*})) \\ \varphi^3 = A\,(\sin(\frac{\theta}{2})\,\psi^1 + i\cos(\frac{\theta}{2})\,\psi^2) + B\,(i\cos(\frac{\theta}{2})\,\psi^{3*} - \sin(\frac{\theta}{2})\,\psi^{4*}) \\ \varphi^4 = \exp(i\,\chi)(-B^*\,(\sin(\frac{\theta}{2})\,\psi^1 + i\cos(\frac{\theta}{2})\,\psi^2) - A^*\,(i\cos(\frac{\theta}{2})\,\psi^{3*} - \sin(\frac{\theta}{2})\,\psi^{4*})) \end{cases} \tag{3.154}$$

In these equations, A and B are complex constants and χ is defined by

$$m\prime = m\,\exp(-i\,\chi) \tag{3.155}$$

This result stems from the fact that the equations naturally impose

$$|m\prime|^2 = |m|^2$$

Eqs. (3.154) mean that

$$\varphi\prime = N_0\,\varphi \begin{pmatrix} iA & -iB^*\exp(i\chi) \\ iB & -iA^*\exp(i\chi) \end{pmatrix} \tag{3.156}$$

and, as we will see, this will be the case for each of the Lorentz transformations.

Let us dwell for a moment on the matrix

$$N \equiv \begin{pmatrix} iA & -iB^*\exp(i\chi) \\ iB & -iA^*\exp(i\chi) \end{pmatrix} \tag{3.157}$$

It can be successively equal to the eight elements of the space algebra, namely

1. I if $A = -i$, $B = 0$, $\chi = 0$
2. $\sigma 1$ if $A = 0$, $B = -i$, $\chi = 0$
3. $\sigma 2$ if $A = 0$, $B = 1$, $\chi = 0$
4. $\sigma 3$ if $A = -i$, $B = 0$, $\chi = \pi$
5. $i\,I$ if $A = 1$, $B = 0$, $\chi = \pi$
6. $i\,\sigma 1$ if $A = 0$, $B = 1$, $\chi = \pi$
7. $i\,\sigma 2$ if $A = 0$, $B = i$, $\chi = \pi$
8. $i\,\sigma 3$ if $A = 1$, $B = 0$, $\chi = 0$

We now have all the pieces in hand to analyze the complete set of Lorentz transformations. Eq. (3.150) can be written successively

$$N_0\,\nabla\,\widetilde{N_0}\,\widehat{N_0}\,\widehat{\varphi}\,\widehat{N} - m\prime\,N_0\,\varphi\,N\,\sigma 2\sigma 1 = 0 \tag{3.158}$$

$$N_0\,(\nabla\,\widetilde{N_0}\,\widehat{N_0}\,\widehat{\varphi}\,\widehat{N} - m\,\exp(-i\chi)\,\varphi\,N\,\sigma 2\sigma 1) = 0 \tag{3.159}$$

$$N_0 \nabla (\widetilde{N_0}\, \widehat{N_0}\, \widehat{\varphi}\, \widehat{N} + \exp(-i\chi)\, \widehat{\varphi}\, \sigma 2\sigma 1\, N\, \sigma 2\sigma 1) = 0 \tag{3.160}$$

Through these differents steps, we have made use of (3.150), (3.152), (3.155) and (3.156).

It is easy to convince ourselves that

$$\sigma 2\sigma 1\, N\, \sigma 2\sigma 1 = -\exp(i\chi)\, \widehat{N} \tag{3.161}$$

We thus obtain

$$N_0 \nabla (\widetilde{N_0}\, \widehat{N_0}\, \widehat{\varphi N} - \widehat{\varphi N}) = N_0 \nabla (det(N_0)^* - 1)\widehat{\varphi N} = 0 \tag{3.162}$$

by use of

$$\widetilde{N_0}\, \widehat{N_0} = (ad - bc)^* I = det(N_0)^* I \tag{3.163}$$

Now, each of the six Lorentz transformations is such that $det(N_0) = 1$.Eq. (3.162) is therefore satisfied and we can conclude that Eqs. (3.155) and (3.156) characterize the even Lorentz symmetries of the real 2x2 Dirac equation.

The additional symmetry still needs to be considered. Indeed, in this case, we have

$$x\prime = \exp(i\lambda)\, x \tag{3.164}$$

and we get

$$\nabla\prime = \exp(-i\lambda)\, \nabla \tag{3.165}$$

This implies that, while (3.155) is still true, while (3.156) now reads

$$\varphi\prime = N_0\, \varphi \begin{pmatrix} iA & iB^* \exp(i\chi) \\ iB & iA^* \exp(i\chi) \end{pmatrix} = N_0\, \varphi\, N\, \sigma 3 \tag{3.166}$$

with $N_0 = \exp(-\frac{i}{2}\lambda)\, I$. In other words, Eq. (3.156) is still valid up to the transformation $\chi \to \chi + \pi$.

3.10.2 The odd (improper) symmetries

Now, let us turn to improper transformations which are obtained from proper transformations by performing

$$x \to \widehat{x}$$

And as we have emphasized, this is the action of the parity operator. So, we have

$$\left\{\begin{array}{l} x^{0\prime}+x^{3\prime}=(|a|^2+|b|^2)x^0-(a^*b+ab^*)x^1-i(a^*b-ab^*)x^2-(|a|^2-|b|^2)x^3 \\ x^{0\prime}-x^{3\prime}=(|c|^2+|d|^2)x^0-(c^*d+cd^*)x^1-i(c^*d-cd^*)x^2-(|c|^2-|d|^2)x^3 \\ x^{1\prime}-ix^{2\prime}=(ac^*+bd^*)x^0-(ad^*+bc^*)x^1-i(-ad^*+bc^*)x^2-(ac^*-bd^*)x^3 \\ x^{1\prime}+ix^{2\prime}=(a^*c+b^*d)x^0-(a^*d+b^*c)x^1-i(a^*d-b^*c)x^2-(a^*c-b^*d)x^3 \end{array}\right.$$

leading to

$$\left\{\begin{array}{l} \partial_0-\partial_3=(|a|^2+|c|^2)\partial_{0\prime}+(a^*c+ac^*)\partial_{1\prime}+i(-a^*c+ac^*)\partial_{2\prime}+(|a|^2-|c|^2)\partial_{3\prime} \\ \partial_0+\partial_3=(|b|^2+|d|^2)\partial_{0\prime}+(b^*d+bd^*)\partial_{1\prime}+i(-b^*d+bd^*)\partial_{2\prime}+(|b|^2-|d|^2)\partial_{3\prime} \\ \partial_1+i\partial_2=-(ab^*+cd^*)\partial_{0\prime}-(b^*c+ad^*)\partial_{1\prime}-i(-b^*c+ad^*)\partial_{2\prime}-(ab^*-cd^*)\partial_{3\prime} \\ \partial_1-i\partial_2=-(a^*b+c^*d)\partial_{0\prime}-(bc^*+a^*d)\partial_{1\prime}-i(bc^*-a^*d)\partial_{2\prime}-(a^*b-c^*d)\partial_{3\prime} \end{array}\right.$$

Therefore, Eq. (3.152) becomes

$$\nabla'=N_1\,\widehat{\nabla}\,\widetilde{N_1} \tag{3.167}$$

Let us focus once again on the rotation $R_1(\theta)$ characterized by (3.153). The key is once again to connect Eqs. (3.146) to

$$\left\{\begin{array}{l} (\partial_0+\sin(\theta)\partial_2+\cos(\theta)\partial_3)\varphi^{4*}+(-\partial_1+i\cos(\theta)\partial_2-i\sin(\theta)\partial_3)\varphi^{2*}+im\prime\,\varphi^1=0 \\ (\partial_0+\sin(\theta)\partial_2+\cos(\theta)\partial_3)\varphi^{3*}+(-\partial_1+i\cos(\theta)\partial_2-i\sin(\theta)\partial_3)\varphi^{1*}+im\prime\,\varphi^2=0 \\ (\partial_0-\sin(\theta)\partial_2-\cos(\theta)\partial_3)\varphi^{2*}+(-\partial_1-i\cos(\theta)\partial_2+i\sin(\theta)\partial_3)\varphi^{4*}-im\prime\,\varphi^3=0 \\ (\partial_0-\sin(\theta)\partial_2-\cos(\theta)\partial_3)\varphi^{1*}+(-\partial_1-i\cos(\theta)\partial_2+i\sin(\theta)\partial_3)\varphi^{3*}-im\prime\,\varphi^4=0 \end{array}\right.$$

As it can be verified, the solutions are

$$\left\{\begin{array}{c} \varphi^1=B\,(\sin(\frac{\theta}{2})\,\psi^{1*}-i\cos(\frac{\theta}{2})\,\psi^{2*})+iA\,(\cos(\frac{\theta}{2})\,\psi^3-i\sin(\frac{\theta}{2})\,\psi^4) \\ \varphi^2=\exp(i\,\chi)(iA^*\,(i\sin(\frac{\theta}{2})\,\psi^{1*}+\cos(\frac{\theta}{2})\,\psi^{2*})-B^*\,(i\cos(\frac{\theta}{2})\,\psi^3+\sin(\frac{\theta}{2})\,\psi^4)) \\ \varphi^3=iB\,(\cos(\frac{\theta}{2})\,\psi^{1*}+i\sin(\frac{\theta}{2})\,\psi^{2*})+A\,(\sin(\frac{\theta}{2})\,\psi^3+i\cos(\frac{\theta}{2})\,\psi^4) \\ \varphi^4=\exp(i\,\chi)(-iA^*\,(\cos(\frac{\theta}{2})\,\psi^{1*}+i\sin(\frac{\theta}{2})\,\psi^{2*})-B^*\,(\sin(\frac{\theta}{2})\,\psi^3+i\cos(\frac{\theta}{2})\,\psi^4)) \end{array}\right. \tag{3.168}$$

The conclusion is inescapable: Eq. (3.155) is still valid, but Eq. (3.156) is now written as

$$\boxed{\varphi\prime = N_1\,\widehat{\varphi}\,N} \tag{3.169}$$

This result can be extended to each of the other Lorentz transformation. Indeed, from (3.150) we have

$$N_1\,\widehat{\nabla}\,\widetilde{N_1}\,\widehat{N_1}\,\varphi\,\widehat{N} - m\prime\,N_1\,\widehat{\varphi}\,N\,\sigma 2\sigma 1 = 0 \tag{3.170}$$

$$\Leftrightarrow N_1\,(\widehat{\nabla}\,\widetilde{N_1}\,\widehat{N_1}\,\varphi\,\widehat{N} - m\,\exp(-i\chi)\,\widehat{\varphi}\,N\,\sigma 2\sigma 1) = 0 \tag{3.171}$$

$$N_1\,\widehat{\nabla}\,(\widetilde{N_1}\,\widehat{N_1} - 1)\varphi\,\widehat{N} = N_1\,\widehat{\nabla}\,(det(N_1)^* - 1)\varphi\,\widehat{N} = 0 \tag{3.172}$$

And this is obviously true because $det(N_1) = 1$ for Lorentz transformations.

The case of the symmetry R is slightly different, as it was for the even operators. We have here

$$\varphi\prime = N_1\,\widehat{\varphi}\begin{pmatrix} iA & iB^*\exp(i\chi) \\ iB & iA^*\exp(i\chi) \end{pmatrix} = N_1\,\widehat{\varphi}\,N\,\sigma 3 \tag{3.173}$$

3.10.3 The discrete symmetries P and T

We have seen the essential role played by the N matrix.

Before analyzing the parity and time-reversal operators, let us examine its impact on the real 2x2 Dirac equation with interaction (3.133). The question is: are there other solutions to this equation, and what are they? In other words, we are looking for $e\prime$, $m\prime$ satisfying

$$\nabla\widehat{\varphi\prime} = e\prime\,A\,\widehat{\varphi\prime}\,\sigma 2\sigma 1 + m\prime\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.174}$$

with

$$\varphi\prime = \varphi\,N \tag{3.175}$$

We have

$$\nabla\,(\widehat{\varphi}\,\widehat{N}) = e\prime\,A\,\widehat{\varphi}\,\widehat{N}\,\sigma 2\sigma 1 + m\prime\,\varphi\,N\,\sigma 2\sigma 1$$

or, by using (3.133)

$$e\,A\,\widehat{\varphi}\,\sigma 2\sigma 1\,\widehat{N} + m\,\varphi\,\sigma 2\sigma 1\,\widehat{N} = e\prime\,A\,\widehat{\varphi}\,\widehat{N}\,\sigma 2\sigma 1 + m\prime\,\varphi\,N\,\sigma 2\sigma 1 \tag{3.176}$$

The system

$$\begin{cases} e\,\sigma 2\sigma 1\,\widehat{N} = e\prime\,\widehat{N}\,\sigma 2\sigma 1 \\ m\,\sigma 2\sigma 1\,\widehat{N} = m\prime\,N\,\sigma 2\sigma 1 \end{cases} \tag{3.177}$$

follows from this.

The first equation leads to

- $e\prime = e$ for $B = 0$
- $e\prime = -e$ for $A = 0$

The second equation confirms a rotation of the mass in the complex plane in the sense that

$$m\prime = m\,\exp(-i\chi) \tag{3.178}$$

So, the real 2x2 Dirac equation involves four types of "matter", as expected:

Positive matter	$B = 0\ ;\ \chi = 0$	$\psi\prime = iA\,\psi$
Positive antimatter	$A = 0\ ;\ \chi = 0$	$\psi\prime = -B\,\gamma^2\psi^*$
Negative matter	$B = 0\ ;\ \chi = \pi$	$\psi\prime = iA\,\gamma^5\psi$
Negative antimatter	$A = 0\ ;\ \chi = \pi$	$\psi\prime = -B\,\gamma^5\gamma^2\psi^*$

Therefore, up to insignificant complex numbers, we do indeed find the results in Table 2.2

3.10.3.1 The parity operator

The action of the parity operator on derivatives implies that

$$\nabla\prime = \widehat{\nabla} \tag{3.179}$$

Eq. (3.174) now reads

$$\widehat{\nabla}\widehat{\varphi}\prime = e\prime\,A\,\widehat{\varphi}\prime\,\sigma 2\sigma 1 + m\prime\,\varphi\prime\,\sigma 2\sigma 1 \tag{3.180}$$

Using the same approach as previously developed, it is easy to convince oneself that

$$\varphi\prime = \widehat{\varphi}\,\widehat{N}\ ;\ \widehat{\varphi}\prime = \varphi\,N \tag{3.181}$$

and

$$\begin{cases} e\,\sigma 2\sigma 1\,N = e\prime\,N\,\sigma 2\sigma 1 \\ m\,\sigma 2\sigma 1\,N = m\prime\,\widehat{N}\,\sigma 2\sigma 1 \end{cases} \tag{3.182}$$

Therefore, the conclusions are obvious:

- $e\prime = e$ for $B = 0$
- $e\prime = -e$ for $A = 0$
- $m\prime = m\,\exp(i\chi)$

We can summarize these results in a table:

Positive matter	$B = 0\ ;\ \chi = 0$	$\psi\prime = iA\,\gamma^0\psi$
Positive antimatter	$A = 0\ ;\ \chi = 0$	$\psi\prime = -B\,\gamma^0\gamma^2\psi^*$
Negative matter	$B = 0\ ;\ \chi = \pi$	$\psi\prime = -A\,\gamma^1\gamma^2\gamma^3\psi$
Negative antimatter	$A = 0\ ;\ \chi = \pi$	$\psi\prime = iB\,\gamma^3\gamma^1\psi^*$

In other words, we have

$$\varphi\prime = \widehat{\varphi} \tag{3.183}$$

and

$$\psi\prime = \gamma^0\psi \tag{3.184}$$

when $A = -i$, $B = 0$, $\chi = 0$. So, $N = I$ corresponds to the parity operator in its unitary form (2.72). This confirms what we have stated on several occasions: conjugation $\widehat{\varphi}$ is nothing other than the action of parity. Moreover, this leads to the fact that mass is unchanged.

The antiunitary version is given in (2.76) and gives rise to

$$\psi\prime = \gamma^0\gamma^2\psi^* \tag{3.185}$$

It corresponds to $N = -\sigma 2$ so that

$$\varphi\prime = \widehat{\varphi}\,\sigma 2 \tag{3.186}$$

Here again, the mass retains its positive sign.

This table shows several things. One of them is that the antiunitary operator in the traditional version uses antimatter with positive mass.

Note also that, as usual, antimatter is represented by a complex conjugation and a multiplication by γ^2 in the traditional version while changing the sign of the mass is equivalent to multiplying the wavefunctions by γ^5. Negative matter and negative antimatter are also the Hodge duals of positive matter and positive antimatter.

3.10.3.2 The time-reversal operator

The action of the time reversal operator is $\partial_0 \rightarrow -\partial_0$ which translates, in geometric algebra, to

$$\nabla\prime = -\widehat{\nabla} \tag{3.187}$$

In other words, the equation to consider is here

$$\widehat{\nabla}\widehat{\varphi}\prime = -e\prime\, A\, \widehat{\varphi}\prime\, \sigma 2\sigma 1 - m\prime\, \varphi\prime\, \sigma 2\sigma 1 \tag{3.188}$$

The solutions of this equation are

$$\varphi\prime = \widehat{\varphi}\, \widehat{N} \tag{3.189}$$

and

$$\begin{cases} e\, \sigma 2\sigma 1\, N = -e\prime\, N\, \sigma 2\sigma 1 \\ m\, \sigma 2\sigma 1\, N = -m\prime\, \widehat{N}\, \sigma 2\sigma 1 \end{cases} \tag{3.190}$$

We conclude that

- $e\prime = e$ for $A = 0$
- $e\prime = -e$ for $B = 0$
- $m\prime = -m\, \exp(i\chi)$

The table is now

Positive matter	$A = 0\ ;\ \chi = \pi$	$\psi\prime = iB\, \gamma^3\gamma^1\psi^*$
Positive antimatter	$B = 0\ ;\ \chi = \pi$	$\psi\prime = -A\, \gamma^1\gamma^2\gamma^3\psi$
Negative matter	$A = 0\ ;\ \chi = 0$	$\psi\prime = -B\, \gamma^0\gamma^2\psi^*$
Negative antimatter	$B = 0\ ;\ \chi = 0$	$\psi\prime = iA\, \gamma^0\psi$

If we take $A = -1$, $B = 0$, $\chi = \pi$, we obtain $N = -I$ and the unitary version given by (2.80).

The choice $A = 0$, $B = i$, $\chi = \pi$ leads to $N = i\sigma 2$ and the traditional anti-unitary version associated with (2.84). Both versions are associated with positive masses.

The observation is clear: the time-reversal operator is none other than the dual of parity for the wavefunction relevant to algebraic geometry.

This implies that

- the parity operator for positive matter coincides with the time-reversal operator associated with negative antimatter
- the parity operator for negative antimatter coincides with the time-reversal operator associated with positive matter
- the parity operator for negative matter coincides with the time-reversal operator associated with positive antimatter
- the parity operator for positive antimatter coincides with the time-reversal operator associated with negative matter

The symmetry of our world is therefore more than remarkable!

A word on (3.155) to end this Section.

We could also have mentioned an eighth transformation, besides the Lorentz transformations and the R-symmetry. This results from the fact that the matrix N_0 is complex, and therefore its determinant is also complex, in general. We note it as

$$det(N_0) = n_0 \, \exp(-i\lambda)$$

Let us think about it. Eq. (3.152) can also be written as

$$\nabla = \overline{N_0} \, \nabla\prime \, \widehat{N_0} \tag{3.191}$$

as it can be easily verified. The real 2x2 Dirac equation (3.149) is then equivalent to

$$\overline{N_0} \, \nabla\prime \, \widehat{N_0} \, \widehat{\varphi} - m \, \varphi \, \sigma 2 \sigma 1 = 0 \tag{3.192}$$

The transformed wavefunction is

$$\varphi\prime \equiv N_0 \, \varphi \, N \Rightarrow \widehat{\varphi\prime} = \widehat{N_0} \, \widehat{\varphi} \, \widehat{N} \tag{3.193}$$

This implies that

$$\overline{N_0} \, (\nabla\prime \widehat{\varphi\prime} \widehat{N}^{-1}) - m \, \varphi \, \sigma 2 \sigma 1 = 0 \tag{3.194}$$

Asking for invariance is requiring that

$$\nabla\prime \widehat{\varphi\prime} - m\prime \, \varphi\prime \sigma 2 \sigma 1 = 0 \tag{3.195}$$

We thus have

$$\overline{N_0} \, (m\prime \, N_0 \, \varphi \, N \, \sigma 2 \sigma 1 \, \widehat{N}^{-1} \, \sigma 2 \sigma 1) + m \, \varphi = 0 \tag{3.196}$$

Moreover, it is easy to prove that

$$N \, \sigma 2 \sigma 1 \, \widehat{N}^{-1} \, \sigma 2 \sigma 1 = - \exp(i\chi) \, I \tag{3.197}$$

leading us to

$$m\prime \exp(i\chi)\,\overline{N_0}\,N_0 = m \tag{3.198}$$

Keep in mind (3.77) and you will be convinced that

$$(ad - bc)\,m\prime = \exp(-i\chi)\,m \tag{3.199}$$

The factor $(ad - bc)$ is simply the determinant of N_0. In other words, we have

$$m\prime = \frac{1}{n_0}\exp(-i(\chi - \lambda))\,m \tag{3.200}$$

Of course, for each of the Lorentz transformations, we have $n_0 = 1$ and $\lambda = 0$, so we recover the only result (3.155).

We could have considered a transformation such as $n_0 \neq 1$ and $\lambda = 0$, for instance. However, this transformation does not preserve the pseudonorm of spacetime:

$$x^\mu\prime x_\mu\prime = det(x\prime) = det(N_0 x \widetilde{N_0}) = n_0\,\exp(i\lambda)\,x^\mu x_\mu\,n_0\,\exp(-i\lambda) = n_0^2\,x^\mu x_\mu$$

n_0 is therefore a parameter of dilation (or contraction if $n_0 < 1$) of spacetime distances. It is such that

$$\boxed{m\prime = \frac{1}{n_0}\exp(-i\,\chi)\,m} \tag{3.201}$$

The conclusion is inescapable: where distances become smaller due to the contraction factor n_0, masses and therefore energies become greater.

The time has come to conclude this Section.
What have we learned or confirmed?

- Dirac's equation highlights, in a completely natural way, negative masses (or energies).
- In the context of geometric algebra, these negative masses are obtained from the operator associated with chirality, as in the traditional approach. But they also appear thanks to the time-reversal operator, whether it is unitary or antiunitary.
- A rotation of the mass in the complex plane arises naturally through the invariance of the real 2x2 Dirac equation.
- Geometric algebra also allows us to put in evidence an additional symmetry to the Lorentz transformations and this symmetry corresponds to a rotation, in the complex plane, of the position as well as of the mass.
- Finally, there is a contraction factor which, while reducing the distances of spacetime, increases the masses and therefore the energies that are relative to them.

Chapter 4

Perspectives in quantum field theory

Group theory is clear: negative masses are an integral part of the symmetry group of any relativistic theory. Geometric algebra confirms this view through a rotation of mass in the complex plane, which occurs naturally.

If negative energies regain their place, if negative masses must be taken into account, that means we have to rewrite the entire quantum field theory!

But let's talk about it.

And since it uses annihilation and creation operators, let's see what they consist of using the simple example of the harmonic oscillator.

4.1 The harmonic oscillator

The one-dimensional harmonic oscillator is the system whose Hamiltonian is given by

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega^2 x^2 \tag{4.1}$$

In quantum theory, momentum p is defined by $-i\frac{d}{dx}$, such that

$$[p, x] = -i \tag{4.2}$$

By defining the operators a and $a^\dagger$ as

$$a \equiv \sqrt{\frac{m\omega}{2}}x + \frac{i}{\sqrt{2m\omega}}p\,;\; a^\dagger \equiv \sqrt{\frac{m\omega}{2}}x - \frac{i}{\sqrt{2m\omega}}p$$

and inverting them as

$$p = -i\sqrt{\frac{m\omega}{2}}(a - a^\dagger)\,;\; x = \frac{1}{\sqrt{2m\omega}}(a + a^\dagger) \tag{4.3}$$

this Hamiltonian is written as

$$H = \frac{\omega}{2}\{a, a^\dagger\} \tag{4.4}$$

The above definitions also imply

$$[a, a^\dagger] = 1 \tag{4.5}$$

and, using this result, the Hamiltonian is

$$H = \omega(a^\dagger a + \frac{1}{2}) \tag{4.6}$$

It's easy to convince oneself that

$$[H, a] = -\omega a\,;\, [H, a^\dagger] = \omega a^\dagger \tag{4.7}$$

The operators a and $a^\dagger$ are then regarded as scaling operators:

- a decreases the energy value of ω
- $a^\dagger$ increases the energy value of ω

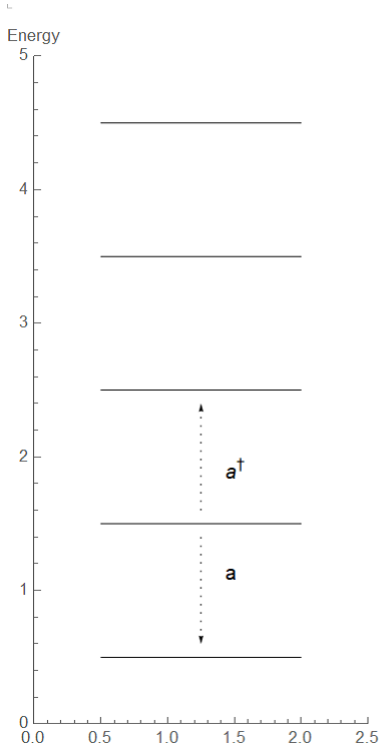


Figure 4.1: Harmonic oscillator energies

For these reasons, a is called the annihilation operator, while $a^\dagger$ is called the creation operator.

It is therefore customary to define the ground state $|0>$ by

$$a|0>=0 \tag{4.8}$$

from which we immediately conclude that the lowest energy of the harmonic oscillator is $\frac{\omega}{2}$.

The other states $|n>$ and other energies $(n+\frac{1}{2})\,\omega$ are obtained by applying the operator $a^\dagger$, n times to $|0>$.

4.2 The Klein-Gordon scalar field

The basic idea is simple and relates to what is known as second quantization: replacing wavefunctions with fields.

A field is defined at every point $(t, \overrightarrow{r})$ in spacetime. If the result is a number, the field is called scalar and denoted by $\phi(t, \overrightarrow{r})$.

This field is more consistent with physical reality because it involves multiple particles rather than a single one, as quantum mechanics has done up to now. It is also more difficult to grasp because there are an infinite number of degrees of freedom (at least one per point...).

It all starts with a Lagrangian density, which, in the case of the Klein-Gordon equation, is given by

$$L=\frac{1}{2}\partial_\mu\partial^\mu\phi-\frac{1}{2}m^2\phi^2=\frac{1}{2}\dot{\phi}^2-\frac{1}{2}(\overrightarrow{\nabla}\phi)^2-\frac{1}{2}m^2\phi^2 \tag{4.9}$$

The scalar field ϕ is real and depends on the four spacetime coordinates $(t, \overrightarrow{r})$, which are summarized by the notation r.

How can we verify that this Lagrangian density is valid? By using the Euler-Lagrange equation

$$\frac{\partial L}{\partial\phi}-\partial_\mu\left(\frac{\partial L}{\partial(\partial_\mu\phi)}\right)=0 \tag{4.10}$$

Indeed, we have successively

$$\frac{\partial L}{\partial\phi}=-m^2\phi \tag{4.11}$$

$$\frac{\partial L}{\partial(\partial_\mu\phi)}=\frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}(\partial_\mu\partial^\mu\phi)=\frac{1}{2}\frac{\partial}{\partial(\partial_\mu\phi)}(g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi)=\partial^\mu\phi \tag{4.12}$$

The Euler-Lagrange equation thus becomes (up to a sign, which is insignificant)

$$\left(\partial_\mu\partial^\mu+m^2\right)\phi(r)=0 \tag{4.13}$$

This is indeed the Klein-Gordon equation.

For every field ϕ, there corresponds a conjugate moment denoted by Π_ϕ and defined by

$$\Pi_\phi \equiv \frac{\partial L}{\partial \dot{\phi}} \tag{4.14}$$

For the Klein-Gordon field, we have

$$\Pi_\phi = \dot{\phi} \tag{4.15}$$

This allows us to introduce a Hamiltonian density given by

$$H \equiv \Pi_\phi \dot{\phi} - L \tag{4.16}$$

In the Klein-Gordon context, it follows that

$$H = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\overrightarrow{\nabla}\phi)^2 + \frac{1}{2}m^2\phi^2 \tag{4.17}$$

We are now ready to write the Klein-Gordon field in detail using annihilation and creation operators.

Keep in mind that this field must be such that

$$[\phi(t,\overrightarrow{r1}),\phi(t,\overrightarrow{r2})] = 0\,;\; [\Pi_\phi(t,\overrightarrow{r1}),\Pi_\phi(t,\overrightarrow{r2})] = 0 \tag{4.18}$$

$$[\phi(t,\overrightarrow{r1}),\Pi_\phi(t,\overrightarrow{r2})] = i\,\delta^{(3)}(\overrightarrow{r1}-\overrightarrow{r2}) \tag{4.19}$$

where $\delta^{(3)}$ is the Dirac distribution (which means here that only the case $\overrightarrow{r1} = \overrightarrow{r2}$ yields a non-zero result). These claims are, of course, based on the entirely analogous results of quantum mechanics:

$$[x,x] = 0\,;\; [p,p] = 0\,;\; [x,p] = i \tag{4.20}$$

The first step in deriving the Klein-Gordon field is to use a Fourier transform and its inverse:

$$\widetilde{\phi}(t,\overrightarrow{p}) = \int d\overrightarrow{r}\; \exp(-i\overrightarrow{p}.\overrightarrow{r})\,\phi(t,\overrightarrow{r}) \tag{4.21}$$

$$\phi(t,\overrightarrow{r}) = \frac{1}{(2\pi)^3}\int d\overrightarrow{p}\; \exp(i\overrightarrow{p}.\overrightarrow{r})\,\widetilde{\phi}(t,\overrightarrow{p}) \tag{4.22}$$

We then recall that $\phi(t,\overrightarrow{r})$ is a solution to the Klein-Gordon equation

$$(\partial_t^2 - \overrightarrow{\nabla}^2 + m^2)\,\phi(t,\overrightarrow{r}) = 0 \tag{4.23}$$

Of course, this means that

$$(\partial_t^2 + \overrightarrow{p}^2 + m^2)\,\widetilde{\phi}(t,\overrightarrow{p}) = 0 \tag{4.24}$$

By defining

$$\omega_{\overrightarrow{p}} \equiv \sqrt{\overrightarrow{p}^2 + m^2} \; (= E_{\overrightarrow{p}}) \tag{4.25}$$

the previous equation can be interpreted as that of a harmonic oscillator vibrating at the frequency $\omega_{\overrightarrow{p}}$. The solution is therefore obvious:

$$\widetilde{\phi}(t, \overrightarrow{p}) = \phi 1(\overrightarrow{p}) \exp(-i\,\omega_{\overrightarrow{p}} t) + \phi 2(\overrightarrow{p}) \exp(i\,\omega_{\overrightarrow{p}} t) \tag{4.26}$$

At first glance, there are two "constants" of integration, but we must not forget that $\phi(t, \overrightarrow{r})$ must be real; in other words, $\phi(t, \overrightarrow{r}) = \phi^\dagger(t, \overrightarrow{r})$, which implies, via the Fourier transform, that

$$(\widetilde{\phi}(t, \overrightarrow{p}))^\dagger = \widetilde{\phi}(t, -\overrightarrow{p}) \tag{4.27}$$

When applied to the highlighted solution, this leads to

$$\phi 2(\overrightarrow{p}) = \phi 1^\dagger(-\overrightarrow{p}) \tag{4.28}$$

So we have

$$\widetilde{\phi}(t, \overrightarrow{p}) = \phi 1(\overrightarrow{p}) \exp(-i\,\omega_{\overrightarrow{p}} t) + \phi 1^\dagger(-\overrightarrow{p}) \exp(i\,\omega_{\overrightarrow{p}} t) \tag{4.29}$$

and

$$\Pi_{\widetilde{\phi}}(t, \overrightarrow{p}) = -i\,\omega_{\overrightarrow{p}} \phi 1(\overrightarrow{p}) \exp(-i\,\omega_{\overrightarrow{p}} t) + i\,\omega_{\overrightarrow{p}} \phi 1^\dagger(-\overrightarrow{p}) \exp(i\,\omega_{\overrightarrow{p}} t) \tag{4.30}$$

We then factor out $\phi 1(\overrightarrow{p})$ as follows:

$$\phi 1(\overrightarrow{p}) = \frac{1}{2}(\widetilde{\phi}(t, \overrightarrow{p}) + \frac{i}{\omega_{\overrightarrow{p}}} \Pi_{\widetilde{\phi}}(t, \overrightarrow{p})) \exp(i\,\omega_{\overrightarrow{p}} t) \tag{4.31}$$

Returning to the original fields, we get

$$\phi 1(\overrightarrow{p}) = \frac{1}{2} \int d\overrightarrow{r} \; \exp(-i(\overrightarrow{p}.\overrightarrow{r} - \omega_{\overrightarrow{p}} t)) \, (\phi(t, \overrightarrow{r}) + \frac{i}{\omega_{\overrightarrow{p}}} \Pi_\phi(t, \overrightarrow{r})) \tag{4.32}$$

Let's calculate the commutation relations for this field and its Hermitian conjugate. We have

$$[\phi 1(\overrightarrow{p1}), \phi 1(\overrightarrow{p2})] = \frac{1}{4} \int d\overrightarrow{r1} d\overrightarrow{r2} \; \exp(-i(\overrightarrow{p1}.\overrightarrow{r1} - \omega_{\overrightarrow{p1}} t)) \exp(-i(\overrightarrow{p2}.\overrightarrow{r2} - \omega_{\overrightarrow{p2}} t))$$
$$[\phi(t, \overrightarrow{r1}) + \frac{i}{\omega_{\overrightarrow{p1}}} \Pi_\phi(t, \overrightarrow{r1}), \phi(t, \overrightarrow{r2}) + \frac{i}{\omega_{\overrightarrow{p2}}} \Pi_\phi(t, \overrightarrow{r2})]$$

Using the results (4.18)-(4.19), we obtain

$$[\phi 1(\overrightarrow{p1}),\phi 1(\overrightarrow{p2})] = \frac{1}{4}\int d\overrightarrow{r1}d\overrightarrow{r2}\,\exp(-i(\overrightarrow{p1}.\overrightarrow{r1}-\omega_{\overrightarrow{p1}}t))\exp(-i(\overrightarrow{p2}.\overrightarrow{r2}-\omega_{\overrightarrow{p2}}t))$$
$$(\frac{i}{\omega_{\overrightarrow{p2}}}-\frac{i}{\omega_{\overrightarrow{p1}}})(i\delta^{(3)}(\overrightarrow{r1}-\overrightarrow{r2}))$$

and

$$[\phi 1(\overrightarrow{p1}),\phi 1(\overrightarrow{p2})] = \frac{1}{4}\int d\overrightarrow{r1}\,\exp(-i(\overrightarrow{p1}.\overrightarrow{r1}-\omega_{\overrightarrow{p1}}t))\exp(-i(\overrightarrow{p2}.\overrightarrow{r1}-\omega_{\overrightarrow{p2}}t))(\frac{1}{\omega_{\overrightarrow{p1}}}-\frac{1}{\omega_{\overrightarrow{p2}}})$$

The distribution

$$\int d\overrightarrow{r1}\,\exp(-i(\overrightarrow{p1}+\overrightarrow{p2}).\overrightarrow{r1}) = (2\pi)^3\delta^{(3)}(\overrightarrow{p1}+\overrightarrow{p2}) \tag{4.33}$$

implies $\overrightarrow{p1} = -\overrightarrow{p2}$ and we have

$$[\phi 1(\overrightarrow{p1}),\phi 1(\overrightarrow{p2})] = \frac{(2\pi)^3}{4}\,\exp(2i\,\omega_{\overrightarrow{p1}}t)\,(\frac{1}{\omega_{\overrightarrow{p1}}}-\frac{1}{\omega_{\overrightarrow{p1}}}) = 0 \tag{4.34}$$

since, by definition, $\omega_{\overrightarrow{p1}} = \omega_{-\overrightarrow{p1}}$.

We apply the same reasoning to the next commutator. We have, in order:

$$[\phi 1(\overrightarrow{p1}),\phi 1^{\dagger}(\overrightarrow{p2})] = \frac{1}{4}\int d\overrightarrow{r1}d\overrightarrow{r2}\,\exp(-i(\overrightarrow{p1}.\overrightarrow{r1}-\omega_{\overrightarrow{p1}}t))\exp(i(\overrightarrow{p2}.\overrightarrow{r2}-\omega_{\overrightarrow{p2}}t))$$
$$[\phi(t,\overrightarrow{r1})+\frac{i}{\omega_{\overrightarrow{p1}}}\Pi_{\phi}(t,\overrightarrow{r1}),\phi(t,\overrightarrow{r2})-\frac{i}{\omega_{\overrightarrow{p2}}}\Pi_{\phi}(t,\overrightarrow{r2})]$$

and

$$[\phi 1(\overrightarrow{p1}),\phi 1^{\dagger}(\overrightarrow{p2})] = \frac{1}{4}\int d\overrightarrow{r1}\,\exp(-i(\overrightarrow{p1}.\overrightarrow{r1}-\omega_{\overrightarrow{p1}}t))\exp(i(\overrightarrow{p2}.\overrightarrow{r1}-\omega_{\overrightarrow{p2}}t))(\frac{1}{\omega_{\overrightarrow{p1}}}+\frac{1}{\omega_{\overrightarrow{p2}}})$$

In total, we get

$$[\phi 1(\overrightarrow{p1}),\phi 1^{\dagger}(\overrightarrow{p2})] = \frac{(2\pi)^3}{4}\,\frac{2}{\omega_{\overrightarrow{p1}}}\,\delta^{(3)}(\overrightarrow{p1}-\overrightarrow{p2}) \tag{4.35}$$

These results lead us to define the annihilation and creation operators as

$$a(\overrightarrow{p}) \equiv \sqrt{2\omega_{\overrightarrow{p}}}\,\phi 1(\overrightarrow{p})\,;\, a^{\dagger}(\overrightarrow{p}) \equiv \sqrt{2\omega_{\overrightarrow{p}}}\,\phi 1^{\dagger}(\overrightarrow{p}) \tag{4.36}$$

We thus have

$$[a(\overrightarrow{p1}), a(\overrightarrow{p2})] = 0 = [a^\dagger(\overrightarrow{p1}), a^\dagger(\overrightarrow{p2})] \tag{4.37}$$

$$[a(\overrightarrow{p1}), a^\dagger(\overrightarrow{p2})] = (2\pi)^3\, \delta^{(3)}(\overrightarrow{p1} - \overrightarrow{p2}) \tag{4.38}$$

As they come from the Klein-Gordon field, related to the spin 0 case, these operators are said bosonic.

Putting all the information together, the Klein-Gordon field can ultimately be written as

$$\phi(t, \overrightarrow{r}) = \frac{1}{(2\pi)^3} \int d\overrightarrow{p}\, \frac{1}{\sqrt{2\omega_{\overrightarrow{p}}}} \left(a(\overrightarrow{p})\, \exp(i(\overrightarrow{p}.\overrightarrow{r} - \omega_{\overrightarrow{p}} t)) + a^\dagger(\overrightarrow{p})\, \exp(-i(\overrightarrow{p}.\overrightarrow{r} - \omega_{\overrightarrow{p}} t))\right)$$

Note the change $\overrightarrow{p} \to -\overrightarrow{p}$ in the second term. Such a change is perfectly valid since we are integrating over all values of $\overrightarrow{p}$.

We also have

$$\Pi_\phi(t, \overrightarrow{r}) = -\frac{i}{(2\pi)^3} \int d\overrightarrow{r}\, \sqrt{\frac{\omega_{\overrightarrow{p}}}{2}} \left(a(\overrightarrow{p})\, \exp(i(\overrightarrow{p}.\overrightarrow{r} - \omega_{\overrightarrow{p}} t)) - a^\dagger(\overrightarrow{p})\, \exp(-i(\overrightarrow{p}.\overrightarrow{r} - \omega_{\overrightarrow{p}} t))\right)$$

So, everything seems to be going well at first glance, but in fact, there's a major problem. And this problem lies in the Hamiltonian.

To highlight it, and to simplify the calculations somewhat, we use the following versions of the fields (assuming $t = 0$):

$$\phi(\overrightarrow{r}) = \frac{1}{(2\pi)^3} \int d\overrightarrow{p}\, \frac{1}{\sqrt{2\omega_{\overrightarrow{p}}}}\, \exp(i\overrightarrow{p}.\overrightarrow{r})\, (a(\overrightarrow{p}) + a^\dagger(-\overrightarrow{p})) \tag{4.39}$$

$$\Pi_\phi(\overrightarrow{r}) = -\frac{i}{(2\pi)^3} \int d\overrightarrow{p}\, \sqrt{\frac{\omega_{\overrightarrow{p}}}{2}}\, \exp(i\overrightarrow{p}.\overrightarrow{r})\, (a(\overrightarrow{p}) - a^\dagger(-\overrightarrow{p})) \tag{4.40}$$

The Hamiltonian density is

$$H = \frac{1}{2}(\Pi_\phi^2 + (\overrightarrow{\nabla}\phi)^2 + m^2\phi^2) \tag{4.41}$$

Let's calculate each of the contributions.

The first one gives

$$\frac{1}{2}\Pi_\phi^2 = -\frac{1}{4(2\pi)^6}\int d\overrightarrow{p1}d\overrightarrow{p2}\,\sqrt{\omega_{\overrightarrow{p1}}\omega_{\overrightarrow{p2}}}\,\exp(i(\overrightarrow{p1}+\overrightarrow{p2}).\overrightarrow{x}) \\ (a(\overrightarrow{p1})-a^\dagger(-\overrightarrow{p1}))(a(\overrightarrow{p2})-a^\dagger(-\overrightarrow{p2}))$$

The second is equal to

$$\frac{1}{2}(\overrightarrow{\nabla}\phi)^2 = -\frac{1}{4(2\pi)^6}\int d\overrightarrow{p1}d\overrightarrow{p2}\,\frac{\overrightarrow{p1}\,\overrightarrow{p2}}{\sqrt{\omega_{\overrightarrow{p1}}\omega_{\overrightarrow{p2}}}}\,\exp(i(\overrightarrow{p1}+\overrightarrow{p2}).\overrightarrow{x}) \\ (a(\overrightarrow{p1})+a^\dagger(-\overrightarrow{p1}))(a(\overrightarrow{p2})+a^\dagger(-\overrightarrow{p2}))$$

Finally, the third one is simply

$$\frac{1}{2}m^2\phi^2 = \frac{m^2}{4(2\pi)^6}\int d\overrightarrow{p1}d\overrightarrow{p2}\,\frac{1}{\sqrt{\omega_{\overrightarrow{p1}}\omega_{\overrightarrow{p2}}}}\,\exp(i(\overrightarrow{p1}+\overrightarrow{p2}).\overrightarrow{x}) \\ (a(\overrightarrow{p1})+a^\dagger(-\overrightarrow{p1}))(a(\overrightarrow{p2})+a^\dagger(-\overrightarrow{p2}))$$

You're probably getting used to this by now: the exponential function $\exp(i(\overrightarrow{p1}+\overrightarrow{p2}).\overrightarrow{x})$ yields a meaningful result only if $\overrightarrow{p2} = -\overrightarrow{p1}$ ($\equiv -\overrightarrow{p}$).

We then have

$$H = \frac{1}{4(2\pi)^3}\int d\overrightarrow{p}\,(-\sqrt{\omega_{\overrightarrow{p}}\omega_{\overrightarrow{-p}}}\,(a(\overrightarrow{p})-a^\dagger(-\overrightarrow{p}))(a(\overrightarrow{p})-a^\dagger(-\overrightarrow{p})) \\ + \frac{\overrightarrow{p}^2+m^2}{\sqrt{\omega_{\overrightarrow{p}}\omega_{\overrightarrow{-p}}}}\,(a(\overrightarrow{p})+a^\dagger(-\overrightarrow{p}))(a(\overrightarrow{p})+a^\dagger(-\overrightarrow{p})))$$

Here, too, we're starting to get the hang of it:

$$\omega_{\overrightarrow{p}} = \omega_{-\overrightarrow{p}}\ ;\ \overrightarrow{p}^2+m^2 = \omega_{\overrightarrow{p}}^2 \tag{4.42}$$

This leads to the following simplifications:

$$H = \frac{1}{4(2\pi)^3}\int d\overrightarrow{p}\,\omega_{\overrightarrow{p}}\,(-(a(\overrightarrow{p})-a^\dagger(-\overrightarrow{p}))(a(\overrightarrow{p})-a^\dagger(-\overrightarrow{p})) \\ + (a(\overrightarrow{p})+a^\dagger(-\overrightarrow{p}))(a(\overrightarrow{p})+a^\dagger(-\overrightarrow{p})))$$

And, after applying the distributive property and the standard substitution $-\overrightarrow{p} \to \overrightarrow{p}$, we have

$$H = \frac{1}{2(2\pi)^3} \int d\overrightarrow{p}\, \omega_{\overrightarrow{p}}\, \{a(\overrightarrow{p}), a^\dagger(\overrightarrow{p})\} \tag{4.43}$$

This is the "quantum field theory" version of the Klein-Gordon Hamiltonian.

You have to admit, it looks a lot like the output of a simple harmonic oscillator (4.4). The problem—and it's no small one!—is that while version (4.6) wasn't particularly problematic for this simple oscillator, it becomes a disaster in the context of quantum field theory, since we have

$$H = \frac{1}{(2\pi)^3} \int d\overrightarrow{p}\, \omega_{\overrightarrow{p}}\, (a^\dagger(\overrightarrow{p})a(\overrightarrow{p}) + \frac{1}{2}(2\pi)^3\delta^{(3)}(0)) \tag{4.44}$$

If the vacuum state is defined by

$$a(\overrightarrow{p})\,|0>=0\,,\ \forall\overrightarrow{p} \tag{4.45}$$

its energy is then given by

$$E_0 = \frac{1}{2} \int d\overrightarrow{p}\, \omega_{\overrightarrow{p}}\, \delta^{(3)}(0) \tag{4.46}$$

And here, it's no longer just a matter of adding a single $\frac{\omega}{2}$, but of taking into account an infinite number of $\frac{\omega_{\overrightarrow{p}}}{2}$! To be more precise, it's actually a double infinity...

The first follows from the definition of the Dirac distribution:

$$(2\pi)^3\delta^{(3)}(\overrightarrow{p}) = \int d\overrightarrow{x}\, \exp(i\overrightarrow{p}.\overrightarrow{x}) \tag{4.47}$$

If we let $\overrightarrow{p}$ tend to 0, we obtain

$$(2\pi)^3\delta^{(3)}(0) = \int d\overrightarrow{x} \tag{4.48}$$

And what do we have on the right-hand side? An infinite volume.

This is what's known as infrared divergence. Physicists get rid of it by reducing it to a finite volume V. So, they assume

$$(2\pi)^3\delta^{(3)}(0) = V \tag{4.49}$$

They then reason that if this divergence arises, it is because we are calculating the total energy rather than the energy density, which they define as

$$\varepsilon_0 \equiv \frac{E_0}{V} = \frac{\frac{1}{2}\int d\overrightarrow{p}\, \omega_{\overrightarrow{p}}\, \delta^{(3)}(0)}{(2\pi)^3\delta^{(3)}(0)} \tag{4.50}$$

As a result, they obtain

$$\varepsilon_0 = \frac{1}{2(2\pi)^3} \int d\overrightarrow{p}\, \omega_{\overrightarrow{p}} = \frac{1}{2(2\pi)^3} \int d\overrightarrow{p}\, \sqrt{\overrightarrow{p}^{\,2} + m^2} = \frac{4\pi}{2(2\pi)^3} \int dp\, p^2\, \sqrt{p^2 + m^2}$$

The last equality follows from a transformation to spherical coordinates using the Jacobian $p^2 \sin\theta$.

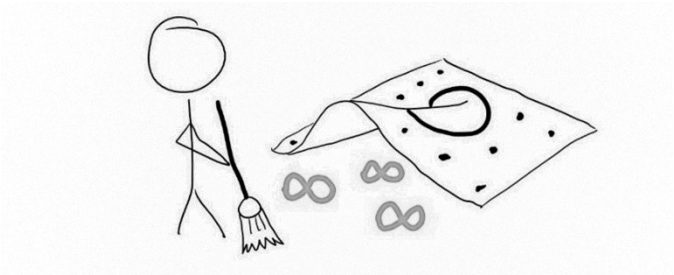
All that remains is to calculate a simple integral... and it diverges wildly! This is the second divergence, which is now called "ultraviolet".

So playing around with a finite volume wasn't enough: there remains a major divergence.

One might say that it manifests at high energies (since the problem stems from the infinite limit of the integral) and that we must therefore simply avoid those energies... while remaining aware that a new theory will be needed to resolve all these issues.

That's still a lot to swallow, especially since those same people were so picky about negative energy!

One of the best images[1] to sum up the situation is this one:



But let's move on to the Dirac equation and see what the situation is in quantum field theory.

4.3 The Dirac field

Before discussing the Lagrangian or the field, let's return to the Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\,\psi(r) = 0 \tag{4.51}$$

The solutions to this equation will play a crucial role in what follows.

We will therefore consider the "standard" solutions found in the physics literature: this requires using the chiral realization and making appropriate choices for the two functions, within the spinors, which can be considered free.

Is this the simplest option? No. But it is the one that almost everyone uses.

4.3.1 Properties of solutions to the Dirac equation

We define, as usual, the positive energy solutions ($p^0 > 0$) as

$$\psi(r) = u(p)\,\exp(-ipr) \tag{4.52}$$

[1] Source: https://www.youtube.com/watch?v=0OQ7BhlfAJY

We thus have

$$(\gamma^\mu p_\mu - m)\, u(p) = 0$$

With the chiral realization, this is equivalent to

$$\begin{pmatrix} -m & p\sigma \\ p\overline{\sigma} & -m \end{pmatrix} \begin{pmatrix} u1(p) \\ u2(p) \end{pmatrix} = 0 \tag{4.53}$$

In this formulation, we have defined

$$p\sigma \equiv p_\mu \sigma^\mu = p_0 - \overrightarrow{p}.\overrightarrow{\sigma}\,; p\overline{\sigma} \equiv p_0 + \overrightarrow{p}.\overrightarrow{\sigma} \tag{4.54}$$

Note that in the geometric algebra language, it would have simply written as

$$p\sigma = p,\ p\overline{\sigma} = \widehat{p}$$

The Dirac equation then boils down to having

$$u1(p) = \frac{p\sigma}{m}\, u2(p) \tag{4.55}$$

since the other constraint

$$u2(p) = \frac{p\overline{\sigma}}{m}\, u1(p) \tag{4.56}$$

is automatically satisfied with

$$u2(p) = \frac{1}{m} p\overline{\sigma}\, u1(p) = \frac{1}{m^2} p\overline{\sigma}\, p\sigma\, u2(p) = \frac{1}{m^2}(p_0^2 - \overrightarrow{p}^2)\, u2(p) = u2(p)$$

How can we reveal symmetry in these solutions? In the case of rest where $\overrightarrow{p} = 0$ and $p_0 = m$, it is simple since we can take $u1(\overrightarrow{p} = 0) = u2(\overrightarrow{p} = 0)$.

The normalization factor can be fixed to simplify subsequent calculations. Therefore, we choose

$$u^{(s)}(\overrightarrow{p} = 0) = \sqrt{m} \begin{pmatrix} \xi^{(s)} \\ \xi^{(s)} \end{pmatrix} \tag{4.57}$$

by distinguishing spin up ($(s) = (+)$) from spin down ($(s) = (-)$) by two different choices for $\xi^{(s)}$, for example:

$$\xi^{(+)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\ ;\ \xi^{(-)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{4.58}$$

Simply apply a boost to exit rest mode:

$$u^{(s)}(\overrightarrow{p}) = \exp(-i\eta\,\widehat{p}.\overrightarrow{K})\,u^{(s)}(\overrightarrow{p}=0) \tag{4.59}$$

Here, we did not make three angles (φ^1, φ^2, φ^3) appear via $\exp(-i\,\overrightarrow{\varphi}.\overrightarrow{K})$ but η and $\widehat{p}$ defined by

$$\cosh\eta \equiv \frac{E}{m}\,;\,\sinh\eta \equiv \frac{||\overrightarrow{p}||}{m}\,;\,\widehat{p} \equiv \frac{\overrightarrow{p}}{||\overrightarrow{p}||} \tag{4.60}$$

The definition of η is, of course, related to $E^2 = \overrightarrow{p}^{\,2} + m^2$.

In the chiral realization, we have successively

$$\exp(-i\eta\,\widehat{p}.\overrightarrow{K}) = I + \frac{\eta}{2}\begin{pmatrix} -\widehat{p}.\overrightarrow{\sigma} & 0 \\ 0 & \widehat{p}.\overrightarrow{\sigma} \end{pmatrix} + \frac{\eta^2}{8}I + \frac{\eta^3}{48}\begin{pmatrix} -\widehat{p}.\overrightarrow{\sigma} & 0 \\ 0 & \widehat{p}.\overrightarrow{\sigma} \end{pmatrix} + \cdots$$

$$= \begin{pmatrix} \cosh\frac{\eta}{2} - \sinh\frac{\eta}{2}\,\widehat{p}.\overrightarrow{\sigma} & 0 \\ 0 & \cosh\frac{\eta}{2} + \sinh\frac{\eta}{2}\,\widehat{p}.\overrightarrow{\sigma} \end{pmatrix} = \begin{pmatrix} K^{(+)} & 0 \\ 0 & K^{(-)} \end{pmatrix}$$

where

$$K^{(+)} \equiv \frac{1}{2}\exp(\frac{\eta}{2})\,(1 - \widehat{p}.\overrightarrow{\sigma}) + \frac{1}{2}\exp(-\frac{\eta}{2})\,(1 + \widehat{p}.\overrightarrow{\sigma}) \tag{4.61}$$

$$K^{(-)} \equiv \frac{1}{2}\exp(\frac{\eta}{2})\,(1 + \widehat{p}.\overrightarrow{\sigma}) + \frac{1}{2}\exp(-\frac{\eta}{2})\,(1 - \widehat{p}.\overrightarrow{\sigma}) \tag{4.62}$$

Given the definition of η, it is clear that

$$\exp(\frac{\eta}{2}) = \sqrt{\frac{E + ||\overrightarrow{p}||}{m}}\,;\,\exp(-\frac{\eta}{2}) = \sqrt{\frac{E - ||\overrightarrow{p}||}{m}} \tag{4.63}$$

Therefore, we deduce that

$$\exp(-i\eta\,\widehat{p}.\overrightarrow{K}) = \frac{1}{\sqrt{m}}\begin{pmatrix} \sqrt{E - ||\overrightarrow{p}||} & 0 \\ 0 & \sqrt{E + ||\overrightarrow{p}||} \end{pmatrix} \tag{4.64}$$

if $\widehat{p}.\overrightarrow{\sigma} = 1$ and

$$\exp(-i\eta\,\widehat{p}.\overrightarrow{K}) = \frac{1}{\sqrt{m}}\begin{pmatrix} \sqrt{E + ||\overrightarrow{p}||} & 0 \\ 0 & \sqrt{E - ||\overrightarrow{p}||} \end{pmatrix} \tag{4.65}$$

if $\widehat{p}.\overrightarrow{\sigma} = -1$. Which we can summarize as

$$\exp(-i\eta\,\widehat{p}.\overrightarrow{K}) = \frac{1}{\sqrt{m}}\begin{pmatrix} \sqrt{E-\overrightarrow{p}.\overrightarrow{\sigma}} & 0 \\ 0 & \sqrt{E+\overrightarrow{p}.\overrightarrow{\sigma}} \end{pmatrix} = \frac{1}{\sqrt{m}}\begin{pmatrix} \sqrt{p\sigma} & 0 \\ 0 & \sqrt{p\overline{\sigma}} \end{pmatrix} \tag{4.66}$$

We thus have two positive energy solutions:

$$u^{(s)}(p) = \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s)} \\ \sqrt{p\overline{\sigma}}\,\xi^{(s)} \end{pmatrix} \tag{4.67}$$

And what about negative energy solutions?
We can define

$$\psi(r) = v(p)\,\exp(-ipr) = v(-p)\,\exp(ipr) \tag{4.68}$$

which we will more readily write

$$\psi(r) = v(p)\,\exp(ipr) \tag{4.69}$$

since, in any case, $v(p)$ is to be determined.

The change of sign in the exponential function induces a change of sign in the mass (which confirms, once again, that negative energy and negative mass go hand in hand):

$$(\gamma^\mu p_\mu + m)\,v(p) = 0 \tag{4.70}$$

Everything then unfolds in the same way as it was done for positive energies, and we obtain

$$v^{(s)}(p) = \begin{pmatrix} \sqrt{p\sigma}\,\eta^{(s)} \\ -\sqrt{p\overline{\sigma}}\,\eta^{(s)} \end{pmatrix} \tag{4.71}$$

the spinor $\eta^{(s)}$ being free, with the only constraint of being of norm equal to 1 (and orthogonal to each other if the s differ).

We will need a whole series of properties relating to these four solutions.
Let's highlight them.

First, this one:

$$\sum_s u^{(s)}(p)\overline{u}^{(s)}(p) = \gamma^\mu p_\mu + m \tag{4.72}$$

with

$$\overline{u}^{(s)}(p) \equiv u^{\dagger(s)}(p)\,\gamma^0 \tag{4.73}$$

It is easily demonstrated since

$$\sum_s u^{(s)}(p)\overline{u}^{(s)}(p) = \sum_s \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s)} \\ \sqrt{p\overline{\sigma}}\,\xi^{(s)} \end{pmatrix} \begin{pmatrix} \sqrt{p\overline{\sigma}}\,\xi^{(s)\dagger} & \sqrt{p\sigma}\,\xi^{(s)\dagger} \end{pmatrix} \tag{4.74}$$

It is clear that

$$\sum_s \xi^{(s)}\xi^{(s)\dagger} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \end{pmatrix} = I \tag{4.75}$$

So this results in

$$\sum_s u^{(s)}(p)\overline{u}^{(s)}(p) = \begin{pmatrix} \sqrt{(p\sigma)(p\overline{\sigma})} & p\sigma \\ p\overline{\sigma} & \sqrt{(p\sigma)(p\overline{\sigma})} \end{pmatrix} \tag{4.76}$$

We also remember that

$$(p\sigma)(p\overline{\sigma}) = (E - \overrightarrow{p}.\overrightarrow{\sigma})(E + \overrightarrow{p}.\overrightarrow{\sigma}) = E^2 - \overrightarrow{p}^2 = m^2 \tag{4.77}$$

The conclusion is inescapable:

$$\sum_s u^{(s)}(p)\overline{u}^{(s)}(p) = \begin{pmatrix} m & E - \overrightarrow{p}.\overrightarrow{\sigma} \\ E + \overrightarrow{p}.\overrightarrow{\sigma} & m \end{pmatrix} = \gamma^\mu p_\mu + m \tag{4.78}$$

Next, these two properties:

$$\overline{u}^{(s\prime)}(p)u^{(s)}(p) = 2m\,\delta^{s\prime s}\,;\; u^{(s\prime)\dagger}(p)u^{(s)}(p) = 2p_0\,\delta^{s\prime s} \tag{4.79}$$

We also demonstrate them. We have

$$\overline{u}^{(s\prime)}(p)u^{(s)}(p) = \begin{pmatrix} \sqrt{p\overline{\sigma}}\,\xi^{(s\prime)\dagger} & \sqrt{p\sigma}\,\xi^{(s\prime)\dagger} \end{pmatrix} \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s)} \\ \sqrt{p\overline{\sigma}}\,\xi^{(s)} \end{pmatrix} \tag{4.80}$$

and

$$\xi^{(s\prime)\dagger}\xi^{(s)} = \delta^{s\prime s} \tag{4.81}$$

We thus obtain

$$\overline{u}^{(s\prime)}(p)u^{(s)}(p) = \delta^{s\prime s}(\sqrt{p\overline{\sigma}}\,\sqrt{p\sigma} + \sqrt{p\sigma}\,\sqrt{p\overline{\sigma}}) = 2m\,\delta^{s\prime s} \tag{4.82}$$

In the same way, we have

$$u^{(s\prime)\dagger}(p)u^{(s)}(p) = \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s\prime)\dagger} & \sqrt{p\overline{\sigma}}\,\xi^{(s\prime)\dagger} \end{pmatrix} \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s)} \\ \sqrt{p\overline{\sigma}}\,\xi^{(s)} \end{pmatrix} = 2p_0\delta^{s\prime s} \quad (4.83)$$

We also have the corresponding properties for negative energy states:

$$\sum_s v^{(s)}(p)\overline{v}^{(s)}(p) = \gamma^\mu p_\mu - m \quad (4.84)$$

$$\overline{v}^{(s\prime)}(p)v^{(s)}(p) = -2m\,\delta^{s\prime s} \quad (4.85)$$

$$v^{(s\prime)\dagger}(p)v^{(s)}(p) = 2p_0\,\delta^{s\prime s} \quad (4.86)$$

And, finally, mixed relations:

$$\overline{u}^{(s\prime)}(p)v^{(s)}(p) = \overline{v}^{(s\prime)}(p)u^{(s)}(p) = 0 \quad (4.87)$$

$$u^{(s\prime)\dagger}(\overrightarrow{p})v^{(s)}(-\overrightarrow{p}) = v^{(s\prime)\dagger}(\overrightarrow{p})u^{(s)}(-\overrightarrow{p}) = 0 \quad (4.88)$$

We demonstrate one for each of the results.

$$\overline{u}^{(s\prime)}(p)v^{(s)}(p) = \begin{pmatrix} \sqrt{p\overline{\sigma}}\,\xi^{(s\prime)\dagger} & \sqrt{p\sigma}\,\xi^{(s\prime)\dagger} \end{pmatrix} \begin{pmatrix} \sqrt{p\sigma}\,\eta^{(s)} \\ -\sqrt{p\overline{\sigma}}\,\eta^{(s)} \end{pmatrix} \quad (4.89)$$

It is clear that the only way to obtain a potentially non-zero result is to consider $s\prime = s$. We then have

$$\sqrt{p\overline{\sigma}}\,\sqrt{p\sigma} - \sqrt{p\sigma}\,\sqrt{p\overline{\sigma}} = 0$$

The second (where the change of sign on $\overrightarrow{p}$ justified the change in notation):

$$u^{(s\prime)\dagger}(\overrightarrow{p})v^{(s)}(-\overrightarrow{p}) = \begin{pmatrix} \sqrt{p\sigma}\,\xi^{(s\prime)\dagger} & \sqrt{p\overline{\sigma}}\,\xi^{(s\prime)\dagger} \end{pmatrix} \begin{pmatrix} \sqrt{p\overline{\sigma}}\,\eta^{(s)} \\ -\sqrt{p\sigma}\,\eta^{(s)} \end{pmatrix} = 0 \quad (4.90)$$

We can finally move on to quantifying the solutions!

4.3.2 The Dirac field

The Dirac Lagrangian is defined by

$$L \equiv \overline{\psi}\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi \tag{4.91}$$

Since it involves two fields, ψ and $\overline{\psi}$, there are two resulting Euler-Lagrange equations:

$$\frac{\partial L}{\partial \psi} - \partial_{\mu}\left(\frac{\partial L}{\partial(\partial_{\mu}\psi)}\right) = 0 \tag{4.92}$$

and

$$\frac{\partial L}{\partial \overline{\psi}} - \partial_{\mu}\left(\frac{\partial L}{\partial(\partial_{\mu}\overline{\psi})}\right) = 0 \tag{4.93}$$

The second one gives the result directly:

$$\left(i\gamma^{\mu}\partial_{\mu} - m\right)\psi = 0$$

that is to say the Dirac equation.

The first one implies

$$\frac{\partial L}{\partial \psi} = -\overline{\psi}m$$

and

$$\frac{\partial L}{\partial(\partial_{\mu}\psi)} = i\overline{\psi}\gamma^{\mu} \;\rightarrow\; \partial_{\mu}\left(\frac{\partial L}{\partial(\partial_{\mu}\psi)}\right) = i\partial_{\mu}\overline{\psi}\gamma^{\mu}$$

It is customary to write this last result as

$$i\overline{\psi}\overleftarrow{\partial_{\mu}}\gamma^{\mu}$$

This yields the second equation,

$$\overline{\psi}\left(i\overleftarrow{\partial_{\mu}}\gamma^{\mu} + m\right) = 0 \tag{4.94}$$

which is, essentially, the Hermitian conjugate of the first.

We also have

$$\Pi = \frac{\partial L}{\partial(\partial_0\psi)} = i\overline{\psi}\gamma^0 = i\psi^{\dagger} \tag{4.95}$$

Since $u^{(s)}(p)\,\exp(-ipx)$ and $v^{(s)}(p)\,\exp(ipx)$ are solutions to the Dirac equation, the spinors

$$u^{(s)}(\overrightarrow{p})\,\exp(i\overrightarrow{p}.\overrightarrow{x})\,;\; v^{(s)}(\overrightarrow{p})\,\exp(-i\overrightarrow{p}.\overrightarrow{x}) \tag{4.96}$$

are eigenfunctions of its Hamiltonian, with respective eigenvalues $E_{\overrightarrow{p}}$ and $-E_{\overrightarrow{p}}$.

But here's the thing... negative energies aren't welcome, as you know!

Furthermore, physicists keep in mind the form (4.39) of the Klein-Gordon field. It therefore seems to them that the best option is to assume that the Dirac field is linear with respect to the annihilation and creation operators.

But annihilate or create what?

Here is what Dirac wrote on page 293 of his book[2]:

"The annihilation of a negative-energy electron is to be understood as the creation of a hole in the sea of negative-energy electrons, or the creation of a positron."

In other words, the Dirac field will be linear, and its linearity will apply not only to the annihilation operator $b_s(\overrightarrow{p})$ of a positive-energy electron but also to the creation operator $c_s^\dagger(\overrightarrow{p})$ of a positron, which is also positive-energy.

The Dirac field is thus defined by

$$\psi(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (b_s(\overrightarrow{p})\,u^{(s)}(\overrightarrow{p})\exp(-ipx)+c_s^\dagger(\overrightarrow{p})\,v^{(s)}(\overrightarrow{p})\exp(ipx)) \tag{4.97}$$

Its Hermitian conjugate is then

$$\psi^\dagger(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (b_s^\dagger(\overrightarrow{p})\,u^{(s)\dagger}(\overrightarrow{p})\exp(ipx)+c_s(\overrightarrow{p})\,v^{(s)\dagger}(\overrightarrow{p})\exp(-ipx)) \tag{4.98}$$

and we also have

$$\overline{\psi}(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (b_s^\dagger(\overrightarrow{p})\,\overline{u}^{(s)}(\overrightarrow{p})\exp(ipx)+c_s(\overrightarrow{p})\,\overline{v}^{(s)}(\overrightarrow{p})\exp(-ipx)) \tag{4.99}$$

These two fields are also linear with respect to the annihilation and creation operators.

Note (and this is obvious, if only because of the factor $\sqrt{2E_{\overrightarrow{p}}}$) that all energies are considered positive. This is a postulate on which all subsequent developments are based.

Let's go back to the fields. These are fermionic fields... so it is the anti-commutation relations (rather than the commutation relations) that will come into play here.

We will therefore set

[2] Dirac, P.A.M. (1958). *The Principles of Quantum Mechanics*, 4th Edition, Oxford University Press

$$\{b_r(\overrightarrow{p1}), b_s^\dagger(\overrightarrow{p2})\} = \{c_r(\overrightarrow{p1}), c_s^\dagger(\overrightarrow{p2})\} = (2\pi)^3\delta^{(3)}(\overrightarrow{p1} - \overrightarrow{p2})\delta_{rs} \tag{4.100}$$

(the other relations are always zero).

We then calculate the anticommutation relations between the fields at the same time, which amounts to considering

$$\psi(\overrightarrow{x}) = \psi(t=0, \overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s (b_s(\overrightarrow{p})\, u^{(s)}(\overrightarrow{p}) \exp(i\overrightarrow{p}.\overrightarrow{x}) + c_s^\dagger(\overrightarrow{p})\, v^{(s)}(\overrightarrow{p}) \exp(-i\overrightarrow{p}.\overrightarrow{x}))$$

which can also be written more simply as

$$\psi(\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s (b_s(\overrightarrow{p})\, u^{(s)}(\overrightarrow{p}) + c_s^\dagger(-\overrightarrow{p})\, v^{(s)}(-\overrightarrow{p})) \exp(i\overrightarrow{p}.\overrightarrow{x}) \tag{4.101}$$

For example, we have

$$\{\psi(\overrightarrow{x}), \psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^6}\int d^3p1 d^3p2 \frac{1}{\sqrt{2E_{\overrightarrow{p1}}E_{\overrightarrow{p2}}}} \sum_s (\{b_s(\overrightarrow{p1}), b_s^\dagger(\overrightarrow{p2})\} u^{(s)}(\overrightarrow{p1})\overline{u}^{(s)}(\overrightarrow{p2}) + \{c_s(-\overrightarrow{p1}), c_s^\dagger(-\overrightarrow{p2})\} v^{(s)}(-\overrightarrow{p1})\overline{v}^{(s)}(-\overrightarrow{p2}))\, \gamma^0 \exp(i(\overrightarrow{p1}.\overrightarrow{x} - \overrightarrow{p2}.\overrightarrow{y}))$$

We have already used all the zero anticommutation relations in this result. By exploiting those that are not zero, we get

$$\{\psi(\overrightarrow{x}), \psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p1 \frac{1}{2E_{\overrightarrow{p1}}} \sum_s (u^{(s)}(\overrightarrow{p1})\overline{u}^{(s)}(\overrightarrow{p1}) + v^{(s)}(-\overrightarrow{p1})\overline{v}^{(s)}(-\overrightarrow{p1}))\, \gamma^0 \exp(i(\overrightarrow{p1}.\overrightarrow{x} - \overrightarrow{p1}.\overrightarrow{y}))$$

Now is the time to remember (4.72) as well as

$$\sum_s v^{(s)}(p)\overline{v}^{(s)}(p) = \gamma^\mu p_\mu - m \Leftrightarrow \sum_s v^{(s)}(-\overrightarrow{p})\overline{v}^{(s)}(-\overrightarrow{p}) = \gamma^0 p_0 - \gamma^j p_j - m \tag{4.102}$$

We then have

$$\{\psi(\overrightarrow{x}), \psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p1 \frac{1}{2E_{\overrightarrow{p1}}} \exp(i\overrightarrow{p1}.(\overrightarrow{x} - \overrightarrow{y})) ((\gamma^0 p_0 + \gamma^j p_j + m)\gamma^0 + (\gamma^0 p_0 - \gamma^j p_j - m)\gamma^0)$$

This leads to

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p1\,\frac{1}{2E_{\overrightarrow{p1}}}\,\exp(i\overrightarrow{p1}.(\overrightarrow{x}-\overrightarrow{y}))2p_0$$
$$= \frac{1}{(2\pi)^3}\int d^3p1\,\exp(i\overrightarrow{p1}.(\overrightarrow{x}-\overrightarrow{y})) = \delta^{(3)}(\overrightarrow{x}-\overrightarrow{y})$$

i.e. the expected anticommutation relation between the Dirac fields.

Let us now calculate the Hamiltonian.

The Hamiltonian density is

$$H = \Pi\dot{\psi} - L = i\psi^\dagger\partial_0\psi - \psi^\dagger\gamma^0(i\gamma^\mu\partial_\mu - m)\psi = -i\psi^\dagger\gamma^0\gamma^j\partial_j\psi + m\psi^\dagger\gamma^0\psi$$

In other words, we have

$$H = \overline{\psi}\,(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\,\psi \tag{4.103}$$

First, we calculate

$$(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\,\psi = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\,\sum_s (b_s(\overrightarrow{p})(\overrightarrow{\gamma}.\overrightarrow{p} + m)u^{(s)}(\overrightarrow{p})$$
$$+ c_s^\dagger(-\overrightarrow{p})(\overrightarrow{\gamma}.\overrightarrow{p} + m)v^{(s)}(-\overrightarrow{p}))\exp(i\overrightarrow{p}.\overrightarrow{x})$$

Next, we need to call on

$$(\gamma^\mu p_\mu - m)u(p) = 0 \Leftrightarrow (\overrightarrow{\gamma}.\overrightarrow{p} + m)u^{(s)}(\overrightarrow{p}) = \gamma^0 p_0\,u^{(s)}(\overrightarrow{p}) \tag{4.104}$$

$$(\gamma^\mu p_\mu + m)v(p) = 0 \Leftrightarrow (\overrightarrow{\gamma}.\overrightarrow{p} + m)v^{(s)}(-\overrightarrow{p}) = -\gamma^0 p_0\,v^{(s)}(-\overrightarrow{p}) \tag{4.105}$$

This gives us

$$(-i\overrightarrow{\gamma}.\overrightarrow{\nabla}+m)\,\psi = \frac{1}{(2\pi)^3}\int d^3p\,\sqrt{\frac{E_{\overrightarrow{p}}}{2}}\gamma^0\sum_s (b_s(\overrightarrow{p})u^{(s)}(\overrightarrow{p}) - c_s^\dagger(-\overrightarrow{p})v^{(s)}(-\overrightarrow{p}))\exp(i\overrightarrow{p}.\overrightarrow{x})$$

Of course, here—as is generally the case throughout the calculations—we consider only positive energies, taking into account that $p_0 = E_{\overrightarrow{p}}$.

We are now ready to calculate the Hamiltonian:

$$H = \int d^3x\,\psi^\dagger\gamma^0(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\psi \tag{4.106}$$

We have

$$H = \frac{1}{(2\pi)^6}\int d^3x\, d^3p1\, d^3p2\, \frac{1}{2}\sqrt{\frac{E_{\overrightarrow{p1}}}{E_{\overrightarrow{p2}}}} \sum_r (b_r^\dagger(\overrightarrow{p2})u^{(r)\dagger}(\overrightarrow{p2}) + c_r(-\overrightarrow{p2})v^{(r)\dagger}(-\overrightarrow{p2}))\exp(-i\overrightarrow{p2}.\overrightarrow{y}) \sum_s (b_s(\overrightarrow{p1})u^{(s)}(\overrightarrow{p1}) - c_r^\dagger(-\overrightarrow{p1})v^{(s)}(-\overrightarrow{p1}))\exp(i\overrightarrow{p1}.\overrightarrow{x})$$

We bring the properties

$$u^{(r)\dagger}(\overrightarrow{p2})v^{(s)}(-\overrightarrow{p1}) = v^{(r)\dagger}(-\overrightarrow{p2})u^{(s)}(\overrightarrow{p1}) = 0 \tag{4.107}$$

into play as well as the integration over spatial coordinates, which yields a distribution $\delta^{(3)}(\overrightarrow{p1} - \overrightarrow{p2})$, in other words, an equality $\overrightarrow{p1} = \overrightarrow{p2} = \overrightarrow{p}$.

We obtain

$$H = \frac{1}{(2\pi)^3}\int d^3p\, \frac{1}{2}\sum_{r,s}(b_r^\dagger(\overrightarrow{p})b_s(\overrightarrow{p})u^{(r)\dagger}(\overrightarrow{p})u^{(s)}(\overrightarrow{p}) - c_r(\overrightarrow{p})c_s^\dagger(\overrightarrow{p})v^{(r)\dagger}(\overrightarrow{p})v^{(s)}(\overrightarrow{p}))$$

by making the substitution $\overrightarrow{p} \to -\overrightarrow{p}$ in the second term.

This expression can be further simplified by using

$$u^{(r)\dagger}(\overrightarrow{p})u^{(s)}(\overrightarrow{p}) = 2p_0\delta^{rs} = v^{(r)\dagger}(\overrightarrow{p})v^{(s)}(\overrightarrow{p}) \tag{4.108}$$

We finally have

$$\boxed{H = \frac{1}{(2\pi)^3}\int d^3p\, E_{\overrightarrow{p}} \sum_s (b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p}) - c_s(\overrightarrow{p})c_s^\dagger(\overrightarrow{p}))} \tag{4.109}$$

If we take into account the anticommutation relations involving the operators $c_s(\overrightarrow{p})$ and $c_s^\dagger(\overrightarrow{p})$, this Hamiltonian can also be written as

$$\boxed{H = \frac{1}{(2\pi)^3}\int d^3p\, E_{\overrightarrow{p}} \sum_s (b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p}) + c_s^\dagger(\overrightarrow{p})c_s(\overrightarrow{p})) + E_0} \tag{4.110}$$

It is in the vacuum energy that we find, once again, a divergence through $\delta^{(3)}(0)$:

$$E_0 = -2\int d^3p\, E_{\overrightarrow{p}}\, \delta^{(3)}(0) \tag{4.111}$$

If you're observant, you've probably noticed that this fermionic divergence has the opposite sign of the bosonic divergence of the Klein-Gordon Hamiltonian... the vacuum energy is negative here!

This is what gave rise to the idea of supersymmetry! Researchers were inspired to propose that if a fermionic field were coupled to a bosonic field, the divergences would cancel each other out and there would no longer be any problem...

Unfortunately, this implied the existence of new particles, the superpartners. Each fermion would have had to correspond to a new boson particle, and vice versa. Unfortunately... because these new particles have never been detected, despite intensive and costly research.

4.3.3 Discrete symmetry operators

We will study discrete symmetry operators within the framework of this quantum field theory.

The approach will always be the same:

- Examine the commutation or anticommutation relations of the discrete symmetry in question with the ten generators of the Poincaré algebra
- Deduce its action on spin-$\frac{1}{2}$ states
- Deduce its action on creation and annihilation operators
- Apply these results to the Dirac field
- Analyze the impact of these changes on spinors, which are solutions to the Dirac equation
- Conclude

4.3.3.1 The parity operator P

Let us recall that its action is characterized by

$$t \to t\,;\ \overrightarrow{x} \to -\overrightarrow{x}$$

which, with respect to the ten generators of the Poincaré algebra

$$J^{\mu\nu} = i(x^\mu \partial^\nu - x^\nu \partial^\mu),\ P^\mu = i\partial^\mu$$

clearly leads to

$$\{P_U, \overrightarrow{P}\} = [P_U, P_0] = [P_U, \overrightarrow{J}] = \{P_U, \overrightarrow{K}\} = 0 \tag{4.112}$$

if it is considered unitary (and recall that this is necessary if we wish to avoid negative energies).

We then introduce the action of P on spin-$\frac{1}{2}$ states.

In the unitary case, the parity operator commutes with rotations, which implies that spin remains unchanged. We therefore set

$$P_U|\overrightarrow{p} = 0, s = \frac{1}{2} >= \eta\,|\overrightarrow{p} = 0, s = \frac{1}{2} > \tag{4.113}$$

where the factor η is a simple phase factor[3]

For the spin-down state, we have, in order:

$$P_U|\overrightarrow{p}=0,s=-\frac{1}{2}>=P_U(J_x-iJ_y)|\overrightarrow{p}=0,s=\frac{1}{2}>=(J_x-iJ_y)P_U|\overrightarrow{p}=0,s=\frac{1}{2}> \tag{4.114}$$

The last equality follows from the fact that the parity operator commutes with $\overrightarrow{J}$ and has no effect on i.

We then have

$$P_U|\overrightarrow{p}=0,s=-\frac{1}{2}>=\eta\,|\overrightarrow{p}=0,s=-\frac{1}{2}> \tag{4.115}$$

We'll now exit the rest state by applying a boost:

$$P_U|\overrightarrow{p},s_\pm>=P_UB(\overrightarrow{p})|\overrightarrow{p}=0,s_\pm> \tag{4.116}$$

But what is inside $B(\overrightarrow{p})$? Essentially, an exponential function with the boosts $\overrightarrow{K}$ multiplied by i as its argument. P_U does not affect the imaginary part but results in a "-" sign due to the presence of the $\overrightarrow{K}$ with which it anticommutes.

We thus have

$$\begin{aligned}P_U|\overrightarrow{p},s_\pm>&=P_UB(\overrightarrow{p})|\overrightarrow{p}=0,s_\pm>=P_UB(\overrightarrow{p})P_U^{-1}P_U|\overrightarrow{p}=0,s_\pm>\\&=\eta\,P_UB(\overrightarrow{p})P_U^{-1}|\overrightarrow{p}=0,s_\pm>=\eta\,B(-\overrightarrow{p})|\overrightarrow{p}=0,s_\pm>\end{aligned}$$

In the end, we get

$$P_U|\overrightarrow{p},s_\pm>=\eta\,|-\overrightarrow{p},s_\pm> \tag{4.117}$$

Antimatter leads to the same result, but with a phase factor that is, a priori, different from η.

We are now at the stage of examining the impact of discrete symmetries on the creation and annihilation operators.

To do this, we consider the vacuum state, which remains unchanged under the action of discrete symmetry operators.

So, for example, we will have

$$P_U^{-1}b_s^\dagger(\overrightarrow{p})P_U|0>=P_U^{-1}b_s^\dagger(\overrightarrow{p})|0>=\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}P_U^{-1}|\overrightarrow{p},s>=\frac{\eta}{\sqrt{2E_{\overrightarrow{p}}}}|-\overrightarrow{p},s>=\eta b_s^\dagger(-\overrightarrow{p})$$

We will thus note the actions of P_U on the creation and annihilation operators for matter and antimatter:

[3] A quick note: For convenience and to avoid cumbersome notation, we denote all phase factors as η. It goes without saying that these factors are not all equal, and we will distinguish between them whenever necessary.

$$P_U^{-1} b_s(\overrightarrow{p}) P_U = \eta_b^* \, b_s(-\overrightarrow{p}) \, ; \, P_U^{-1} c_s(\overrightarrow{p}) P_U = \eta_c^* \, c_s(-\overrightarrow{p}) \tag{4.118}$$

$$P_U^{-1} b_s^\dagger(\overrightarrow{p}) P_U = \eta_b \, b_s^\dagger(-\overrightarrow{p}) \, ; \, P_U^{-1} c_s^\dagger(\overrightarrow{p}) P_U = \eta_c \, c_s^\dagger(-\overrightarrow{p}) \tag{4.119}$$

We apply the results related to P_U to the Dirac field:

$$P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\int d^3p \, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s (\eta_b^* \, b_s(-\overrightarrow{p}) \, u^{(s)}(\overrightarrow{p}) \exp(-ipx) + \eta_c \, c_s^\dagger(-\overrightarrow{p}) \, v^{(s)}(\overrightarrow{p}) \exp(ipx))$$

A little trick: change $\overrightarrow{p}$ to $-\overrightarrow{p}$ and introduce $\widehat{x} \equiv (t, -\overrightarrow{x})$.
We then have

$$px = p^0 x^0 - \overrightarrow{p} \, . \, \overrightarrow{x} \rightarrow p\widehat{x}$$

under parity action and

$$P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\int d^3p \, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s (\eta_b^* \, b_s(\overrightarrow{p}) \, u^{(s)}(-\overrightarrow{p}) \exp(-ip\widehat{x}) + \eta_c \, c_s^\dagger(\overrightarrow{p}) \, v^{(s)}(-\overrightarrow{p}) \exp(ip\widehat{x}))$$

After all, it's easy to convince oneself that

$$u^{(s)}(-\overrightarrow{p}) = \gamma^0 u^{(s)}(\overrightarrow{p}) \, ; \, v^{(s)}(-\overrightarrow{p}) = -\gamma^0 v^{(s)}(\overrightarrow{p}) \tag{4.120}$$

to get

$$P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\gamma^0 \int d^3p \, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s (\eta_b^* \, b_s(\overrightarrow{p}) \, u^{(s)}(\overrightarrow{p}) \exp(-ip\widehat{x}) - \eta_c \, c_s^\dagger(\overrightarrow{p}) \, v^{(s)}(\overrightarrow{p}) \exp(ip\widehat{x}))$$

By setting η_c to be a real number equal to $-\eta_b$ (which indicates that the electron and positron will have opposite intrinsic parities), we finally obtain

$$P_U^{-1} \, \psi(t, \overrightarrow{x}) \, P_U = \eta_b \, \gamma^0 \psi(t, -\overrightarrow{x}) \tag{4.121}$$

This result will come as no surprise to anyone, given what we have discussed about the parity operator in quantum mechanics.

4.3.3.2 The time-reversal operator

Its action is characterized by

$$t \to -t\,;\ \overrightarrow{x} \to \overrightarrow{x}$$

We know that it is antiunitary if we want to avoid negative energies.

This means that it is impossible to represent it by a single matrix... this matrix would necessarily have to be followed by a complex conjugation.

As before, we are primarily interested in the relationships between the discrete symmetry operator and the operators of the Poincaré algebra:

$$\{T_{AU}, \overrightarrow{P}\} = [T_{AU}, P_0] = \{T_{AU}, \overrightarrow{J}\} = [T_{AU}, \overrightarrow{K}] = 0 \tag{4.122}$$

So we will have

$$T_{AU}|\overrightarrow{p} = 0, s = \frac{1}{2} >= \eta\,|\overrightarrow{p} = 0, s = -\frac{1}{2} > \tag{4.123}$$

And, regarding the spin-down state,

$$\begin{aligned} T_{AU}|\overrightarrow{p} = 0, s = -\frac{1}{2} >= T_{AU}(J_x - iJ_y)|\overrightarrow{p} = 0, s = \frac{1}{2} >= -(J_x + iJ_y)T_{AU}|\overrightarrow{p} = 0, s = \frac{1}{2} > \\ = -\eta\,(J_x + iJ_y)|\overrightarrow{p} = 0, s = -\frac{1}{2} >= -\eta\,|\overrightarrow{p} = 0, s = \frac{1}{2} > \end{aligned}$$

These relations can be summarized as follows:

$$T_{AU}|\overrightarrow{p} = 0, s_\pm >= \pm\eta\,|\overrightarrow{p} = 0, s_\mp > \tag{4.124}$$

Finally, if we apply a boost to these states:

$$\begin{aligned} T_{AU}|\overrightarrow{p}, s_\pm >= T_{AU}B(\overrightarrow{p})|\overrightarrow{p} = 0, s_\pm >= T_{AU}B(\overrightarrow{p})T_{AU}^{-1}T_{AU}|\overrightarrow{p} = 0, s_\pm > \\ = \pm\eta\,T_{AU}B(\overrightarrow{p})T_{AU}^{-1}|\overrightarrow{p} = 0, s_\mp >= \pm\eta\,B(-\overrightarrow{p})|\overrightarrow{p} = 0, s_\mp > \end{aligned}$$

and, in the end,

$$T_{AU}|\overrightarrow{p}, s_\pm >= \pm\eta\,|-\overrightarrow{p}, s_\mp > \tag{4.125}$$

What we'll note

$$T_{AU}|\overrightarrow{p}, s >= i\eta\,(\sigma 2)_{ss'}|-\overrightarrow{p}, s' > \tag{4.126}$$

We will thus note the actions of T_{AU} on the creation and annihilation operators for matter and antimatter:

$$T_{AU}^{-1} b_s(\overrightarrow{p}) T_{AU} = -i\eta_b^*(\sigma 2)_{ss'} b_{s'}(-\overrightarrow{p})\,;\; T_{AU}^{-1} c_s(\overrightarrow{p}) T_{AU} = -i\eta_c^*(\sigma 2)_{ss'} c_{s'}(-\overrightarrow{p}) \tag{4.127}$$

$$T_{AU}^{-1} b_s^\dagger(\overrightarrow{p}) T_{AU} = -i\eta_b(\sigma 2)_{ss'} b_{s'}^\dagger(-\overrightarrow{p})\,;\; T_{AU}^{-1} c_s^\dagger(\overrightarrow{p}) T_{AU} = -i\eta_c(\sigma 2)_{ss'} c_{s'}^\dagger(-\overrightarrow{p}) \tag{4.128}$$

We apply these results to the Dirac field

$$T_{AU}^{-1}\psi(x)T_{AU} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (-i\eta_b^*(\sigma 2)_{ss'} b_{s'}(-\overrightarrow{p})\,u^{(s)*}(\overrightarrow{p})\exp(ipx) - i\eta_c(\sigma 2)_{ss'} c_{s'}^\dagger(-\overrightarrow{p})\,v^{(s)*}(\overrightarrow{p})\exp(-ipx))$$

Once again, we change $\overrightarrow{p}$ to $-\overrightarrow{p}$ and introduce now $\widetilde{x} \equiv (-t, \overrightarrow{x})$.

We then have

$$px = p^0x^0 - \overrightarrow{p}.\overrightarrow{x} \to -p\widetilde{x} \tag{4.129}$$

and

$$T_{AU}^{-1}\psi(x)T_{AU} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (-i\eta_b^*(\sigma 2)_{ss'} b_{s'}(\overrightarrow{p})\,u^{(s)*}(-\overrightarrow{p})\exp(-ip\widetilde{x}) - i\eta_c(\sigma 2)_{ss'} c_{s'}^\dagger(\overrightarrow{p})\,v^{(s)*}(-\overrightarrow{p})\exp(ip\widetilde{x}))$$

We need to try to connect $-i(\sigma 2)_{ss'} b_{s'}(\overrightarrow{p})\,u^{(s)*}(-\overrightarrow{p})$ to $b_s(\overrightarrow{p})u^{(s)}(\overrightarrow{p})$.

But given the form of the spinors, it's anything but easy!

In fact, they involve operators such as $\sqrt{p\sigma}$. However, taking the complex conjugate of these operators is not done by considering the complex conjugate of the radicand...

Instead, we must return to the alternative formulation of these operators:

$$\sqrt{p\sigma} = \frac{1}{2}(\sqrt{E + ||\overrightarrow{p}||}(1 - \widehat{p}.\overrightarrow{\sigma}) + \sqrt{E - ||\overrightarrow{p}||}(1 + \widehat{p}.\overrightarrow{\sigma})) \tag{4.130}$$

$$\sqrt{p\overline{\sigma}} = \frac{1}{2}(\sqrt{E - ||\overrightarrow{p}||}(1 - \widehat{p}.\overrightarrow{\sigma}) + \sqrt{E + ||\overrightarrow{p}||}(1 + \widehat{p}.\overrightarrow{\sigma})) \tag{4.131}$$

The complex conjugation then applies only to $\sigma 2$, and to account for this single sign change, we write

$$(\sqrt{p\sigma})^* = \frac{1}{2}\sigma 2(\sqrt{E+||\overrightarrow{p}||}(1+\widehat{p}.\overrightarrow{\sigma}) + \sqrt{E-||\overrightarrow{p}||}(1-\widehat{p}.\overrightarrow{\sigma}))\sigma 2 = \sigma 2\sqrt{p\overline{\sigma}}\sigma 2 \tag{4.132}$$

$$(\sqrt{p\overline{\sigma}})^* = \frac{1}{2}\sigma 2(\sqrt{E-||\overrightarrow{p}||}(1+\widehat{p}.\overrightarrow{\sigma}) + \sqrt{E+||\overrightarrow{p}|\&}(1-\widehat{p}.\overrightarrow{\sigma}))\sigma 2 = \sigma 2\sqrt{p\sigma}\sigma 2 \tag{4.133}$$

We then have

$$\sum_s -i(\sigma 2)_{ss'} b_{s'}(\overrightarrow{p}) u^{(s)*}(-\overrightarrow{p}) = -b_-(\overrightarrow{p}) \begin{pmatrix} \sigma 2\sqrt{p\sigma}\sigma 2\xi^{(+)} \\ \sigma 2\sqrt{p\overline{\sigma}}\sigma 2\xi^{(+)} \end{pmatrix} + b_+(\overrightarrow{p}) \begin{pmatrix} \sigma 2\sqrt{p\sigma}\sigma 2\xi^{(-)} \\ \sigma 2\sqrt{p\overline{\sigma}}\sigma 2\xi^{(-)} \end{pmatrix}$$

It is clear that

$$\sigma 2\,\xi^{(+)} = i\xi^{(-)}\,;\, \sigma 2\,\xi^{(-)} = -i\xi^{(+)} \tag{4.134}$$

from which it comes

$$\sum_s -i(\sigma 2)_{ss'} b_{s'}(\overrightarrow{p}) u^{(s)*}(-\overrightarrow{p}) = -i \begin{pmatrix} \sigma 2 & 0 \\ 0 & \sigma 2 \end{pmatrix} \sum_s b_s(\overrightarrow{p}) u^{(s)}(\overrightarrow{p}) = -\gamma^1\gamma^3 \sum_s b_s(\overrightarrow{p}) u^{(s)}(\overrightarrow{p})$$

Similarly, one can convince oneself that

$$\sum_s -i(\sigma 2)_{ss'} c^\dagger_{s'}(\overrightarrow{p}) v^{(s)*}(-\overrightarrow{p}) = -\gamma^1\gamma^3 \sum_s c^\dagger_s(\overrightarrow{p}) v^{(s)}(\overrightarrow{p})$$

We thus obtain

$$T^{-1}_{AU}\psi(x)T_{AU} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (\eta^*_b(-\gamma^1\gamma^3)b_s(\overrightarrow{p})\,u^{(s)}(\overrightarrow{p})\exp(-ip\widetilde{x}) + \eta_c(-\gamma^1\gamma^3)c^\dagger_s(\overrightarrow{p})v^{(s)}(\overrightarrow{p})\exp(ip\widetilde{x}))$$

If we take η_c to be a real number equal to η_b, we finally have

$$\boxed{T^{-1}_{AU}\,\psi(t,\overrightarrow{x})\,T_{AU} = -\eta_b\,\gamma^1\gamma^3\,\psi(-t,\overrightarrow{x})} \tag{4.135}$$

This is perfectly consistent with what had been demonstrated in the field of quantum mechanics.

4.3.3.3 The charge conjugation operator

This operator is an internal symmetry operator and has no effect on the coordinates. All it does is convert matter into antimatter and vice versa.

As a result,

$$C|\overrightarrow{p},s>_b=\eta_b|\overrightarrow{p},s>_c\,;\,C|\overrightarrow{p},s>_c=\eta_c|\overrightarrow{p},s>_b \tag{4.136}$$

with implications for the creation and annihilation operators:

$$C^{-1}b_s(\overrightarrow{p})C=\eta_c c_s(\overrightarrow{p})\,;\,C^{-1}c_s(\overrightarrow{p})C=\eta_b b_s(\overrightarrow{p}) \tag{4.137}$$

$$C^{-1}b_s^\dagger(\overrightarrow{p})C=\eta_c^* c_s^\dagger(\overrightarrow{p})\,;\,C^{-1}c_s^\dagger(\overrightarrow{p})C=\eta_b^* b_s^\dagger(\overrightarrow{p}) \tag{4.138}$$

What are the implications for the Dirac field?

$$C^{-1}\psi(x)C=\frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s(\eta_c c_s(\overrightarrow{p})\,u^{(s)}(\overrightarrow{p})\exp(-ipx) + \eta_b^* b_s^\dagger(\overrightarrow{p})\,v^{(s)}(\overrightarrow{p})\exp(ipx))$$

Given the expression we've obtained, we should be able to relate it to

$$\psi^\dagger(x)=\frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s(b_s^\dagger(\overrightarrow{p})u^{(s)*}(\overrightarrow{p})\exp(ipx) + c_s(\overrightarrow{p})v^{(s)*}(\overrightarrow{p})\exp(-ipx))$$

However, to do this, we must find a relation between $u^{(s)*}(\overrightarrow{p})$ and $v^{(s)}(\overrightarrow{p})$, as well as between $v^{(s)*}(\overrightarrow{p})$ and $u^{(s)}(\overrightarrow{p})$.

We have

$$u^{(s)*}(\overrightarrow{p})=\begin{pmatrix}\sigma 2\sqrt{p\overline{\sigma}}\sigma 2\xi^{(s)}\\ \sigma 2\sqrt{p\sigma}\sigma 2\xi^{(s)}\end{pmatrix}=-i\begin{pmatrix}0 & \sigma 2\\ -\sigma 2 & 0\end{pmatrix}\begin{pmatrix}\sqrt{p\sigma}\eta^{(s)}\\ -\sqrt{p\overline{\sigma}}\eta^{(s)}\end{pmatrix}=-i\gamma^2 v^{(s)}(\overrightarrow{p})$$

by choosing

$$\eta^{(s)}=i\sigma 2\,\xi^{(s)} \tag{4.139}$$

The complex conjugate of the last result then leads to

$$v^{(s)*}(\overrightarrow{p})=-i\gamma^2 u^{(s)}(\overrightarrow{p}) \tag{4.140}$$

As a result,

$$C^{-1}\psi(x)C = \frac{1}{(2\pi)^3}(-i\gamma^2)\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s(\eta_c c_s(\overrightarrow{p})\,v^{(s)*}(\overrightarrow{p})\exp(-ipx) + \eta_b^* b_s^\dagger(\overrightarrow{p})\,u^{(s)*}(\overrightarrow{p})\exp(ipx))$$

and

$$C^{-1}\psi(x)C = \frac{1}{(2\pi)^3}(-i\gamma^2)\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\,(\sum_s(\eta_c^* c_s^\dagger(\overrightarrow{p})\,v^{(s)}(\overrightarrow{p})\exp(ipx) + \eta_b b_s(\overrightarrow{p})\,u^{(s)}(\overrightarrow{p})\exp(-ipx)))^\dagger$$

By choosing real and equal phases, this expression becomes

$$C^{-1}\,\psi(x)\,C = -i\eta_b\,(\psi^\dagger(x)\gamma^2)^T \tag{4.141}$$

4.4 What if we took it a step further?

In order to generalize what was presented earlier, we must first introduce the solutions to the Dirac equation for both positive and negative masses.

4.4.1 Solutions to the Dirac equation: Would you like to know a little more?

In Hamiltonian form, the Dirac equation is written as

$$i\,\frac{\partial\psi}{\partial t} = (-i\,\overrightarrow{\alpha}.\overrightarrow{\nabla} \pm m\beta)\,\psi \tag{4.142}$$

Here, you'll find both versions, and the equation with the "-" sign in front of m is the one for negative masses (in other words, m is considered strictly positive).

If we set

$$\psi = \exp(-i\varepsilon Et + i\,\overrightarrow{p}.\overrightarrow{x})\,\varphi \tag{4.143}$$

this equation becomes

$$\varepsilon E\,\varphi = (\overrightarrow{\alpha}.\overrightarrow{p} \pm m\beta)\,\varphi \tag{4.144}$$

where $\varepsilon = +1$ refers to positive energies and $\varepsilon = -1$ to negative energies.

With the Dirac realization, these equations are written as

$$\varphi 1 = \frac{1}{\mp m + \varepsilon E}\,\overrightarrow{\sigma}.\overrightarrow{p}\,\varphi 2\,;\;\varphi 2 = \frac{1}{\pm m + \varepsilon E}\,\overrightarrow{\sigma}.\overrightarrow{p}\,\varphi 1 \tag{4.145}$$

This gives us the full set of possibilities.

Positive energy and positive mass
$\exp(-iEt + i\overrightarrow{p}.\overrightarrow{x}) = \exp(-ipx)$
$u^{(s)}(\overrightarrow{p}, m) = N \begin{pmatrix} u1 \\ u2 \\ \frac{(p3u1+(p1-ip2)u2)}{E+m} \\ \frac{(p1+ip2)u1-p3u2)}{E+m} \end{pmatrix}; v^{(s)}(\overrightarrow{p}, m) = N \begin{pmatrix} \frac{(p3u3+(p1-ip2)u4)}{E-m} \\ \frac{((p1+ip2)u3-p3u4)}{E-m} \\ u3 \\ u4 \end{pmatrix}$
Positive energy and negative mass
$\exp(-iEt + i\overrightarrow{p}.\overrightarrow{x}) = \exp(-ipx)$
$u^{(s)}(\overrightarrow{p}, -m) = N \begin{pmatrix} u1 \\ u2 \\ \frac{(p3u1+(p1-ip2)u2)}{E-m} \\ \frac{(p1+ip2)u1-p3u2)}{E-m} \end{pmatrix}; v^{(s)}(\overrightarrow{p}, -m) = N \begin{pmatrix} \frac{(p3u3+(p1-ip2)u4)}{E+m} \\ \frac{((p1+ip2)u3-p3u4)}{E+m} \\ u3 \\ u4 \end{pmatrix}$
Negative energy and positive mass
$\exp(iEt - i\overrightarrow{p}.\overrightarrow{x}) = \exp(ipx)$
$u^{(s)}(-\overrightarrow{p}, -m) = N \begin{pmatrix} u1 \\ u2 \\ -\frac{(p3u1+(p1-ip2)u2)}{E-m} \\ -\frac{(p1+ip2)u1-p3u2)}{E-m} \end{pmatrix}; v^{(s)}(-\overrightarrow{p}, -m) = N \begin{pmatrix} -\frac{(p3u3+(p1-ip2)u4)}{E+m} \\ -\frac{((p1+ip2)u3-p3u4)}{E+m} \\ u3 \\ u4 \end{pmatrix}$
Negative energy and negative mass
$\exp(iEt - i\overrightarrow{p}.\overrightarrow{x}) = \exp(ipx)$
$u^{(s)}(-\overrightarrow{p}, m) = N \begin{pmatrix} u1 \\ u2 \\ -\frac{(p3u1+(p1-ip2)u2)}{E+m} \\ -\frac{(p1+ip2)u1-p3u2)}{E+m} \end{pmatrix}, v^{(s)}(-\overrightarrow{p}, m) = N \begin{pmatrix} -\frac{(p3u3+(p1-ip2)u4)}{E-m} \\ -\frac{((p1+ip2)u3-p3u4)}{E-m} \\ u3 \\ u4 \end{pmatrix}$

One might naively think that there are sixteen solutions in total... Nothing could be further from the truth.

In fact, it's not hard to convince oneself that

$$v^{(s)}(\pm\overrightarrow{p},m)=\frac{p}{E-m}u^{(s)}(\pm\overrightarrow{p},m)\,;\,u^{(s)}(\pm\overrightarrow{p},-m)=\frac{p}{E-m}v^{(s)}(\pm\overrightarrow{p},-m) \tag{4.146}$$

Therefore, we need to consider only eight of the sixteen solutions. We choose

$$u^{(s)}(\overrightarrow{p},m)\,,\,v^{(s)}(\overrightarrow{p},-m)\,,\,v^{(s)}(-\overrightarrow{p},-m)\,,\,u^{(s)}(-\overrightarrow{p},m) \tag{4.147}$$

to be consistent with the choice made in the "standard" literature, which considers only $u^{(s)}(\overrightarrow{p},m)\,,\,v^{(s)}(\overrightarrow{p},-m)$.

We conclude this SubSection by setting a normalization factor and highlighting a few properties.

The final solutions to consider are:

$$u^{(s)}(\overrightarrow{p},m)=\sqrt{E+m}\begin{pmatrix}\xi^{(s)}\\ \frac{\overrightarrow{\sigma}.\overrightarrow{p}}{E+m}\,\xi^{(s)}\end{pmatrix} \tag{4.148}$$

$$u^{(s)}(-\overrightarrow{p},m)=\sqrt{E+m}\begin{pmatrix}\xi^{(s)}\\ -\frac{\overrightarrow{\sigma}.\overrightarrow{p}}{E+m}\,\xi^{(s)}\end{pmatrix} \tag{4.149}$$

$$v^{(s)}(\overrightarrow{p},-m)=\sqrt{E+m}\begin{pmatrix}\frac{\overrightarrow{\sigma}.\overrightarrow{p}}{E+m}\,\eta^{(s)}\\ \eta^{(s)}\end{pmatrix} \tag{4.150}$$

$$v^{(s)}(-\overrightarrow{p},-m)=\sqrt{E+m}\begin{pmatrix}-\frac{\overrightarrow{\sigma}.\overrightarrow{p}}{E+m}\,\eta^{(s)}\\ \eta^{(s)}\end{pmatrix} \tag{4.151}$$

They satisfy the following properties:

$$u^{(s)\dagger}(\pm\overrightarrow{p},m)u^{(r)}(\pm\overrightarrow{p},m)=2E\,\delta^{rs} \tag{4.152}$$

$$\overline{u}^{(s)}(\pm\overrightarrow{p},m)u^{(r)}(\pm\overrightarrow{p},m)=2m\,\delta^{rs} \tag{4.153}$$

$$\sum u^{(s)}(\pm\overrightarrow{p},m)\overline{u}^{(s)}(\pm\overrightarrow{p},m)=E\gamma^0\mp\overrightarrow{\gamma}.\overrightarrow{p}+m \tag{4.154}$$

$$v^{(s)\dagger}(\pm\overrightarrow{p},-m)v^{(r)}(\pm\overrightarrow{p},-m)=2E\,\delta^{rs} \tag{4.155}$$

$$\overline{v}^{(s)}(\pm\overrightarrow{p},-m)v^{(r)}(\pm\overrightarrow{p},-m)=-2m\,\delta^{rs} \tag{4.156}$$

$$\sum v^{(s)}(\pm\overrightarrow{p},-m)\overline{v}^{(s)}(\pm\overrightarrow{p},-m) = E\gamma^0 \mp \overrightarrow{\gamma}.\overrightarrow{p} - m \tag{4.157}$$

In these relations, as usual, the notation † denotes the Hermitian conjugate, and

$$\overline{u}^{(s)}(\pm\overrightarrow{p},m) \equiv u^{(s)\dagger}(\pm\overrightarrow{p},m)\gamma^0\,;\; \overline{v}^{(s)}(\pm\overrightarrow{p},m) \equiv v^{(s)\dagger}(\pm\overrightarrow{p},m)\gamma^0 \tag{4.158}$$

4.4.2 Different Dirac fields

Let's take a moment to revisit the Dirac field, using the conventions from the previous Section:

$$\psi(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (b_s(\overrightarrow{p})u^{(s)}(\overrightarrow{p},m)\exp(-ipx) + c_s^\dagger(\overrightarrow{p})v^{(s)}(\overrightarrow{p},-m)\exp(ipx))$$

In the first term, we recognize the destruction of a fermion with positive energy and mass—in other words, the annihilation of an electron.

But what about the second term?

It represents the creation of a fermion with positive energy and negative mass, but the wave argument has changed sign! It is this change that makes it a positron! This is antimatter according to the Feynman-Stueckelberg interpretation.

To avoid any ambiguity, it must be written using equations.

The spinor $v^{(s)}(\overrightarrow{p},-m)$ is solution of

$$(\gamma^\mu p_\mu + m)\,v^{(s)}(\overrightarrow{p},-m) = 0 \tag{4.159}$$

but the wavefunction $v^{(s)}(\overrightarrow{p},-m)\exp(ipx)$ is solution of

$$i\frac{\partial\psi}{\partial t} = (-i\overrightarrow{\alpha}.\overrightarrow{\nabla} + m\beta)\,\psi \tag{4.160}$$

The Dirac field is therefore what we might call a "positive energy and positive mass" field. When $\exp(-ipx) \to \exp(ipx)$, matter becomes antimatter.

This change in the wavefunction is crucial.

Consider the standard Dirac equation

$$(\gamma^\mu p_\mu - m)\,u^{(s)}(\overrightarrow{p},m) = 0 \tag{4.161}$$

Depending on the choice of plane wave, the mass will change sign:

$$i\frac{\partial}{\partial t}(u^{(s)}(\overrightarrow{p},m)\exp(-ipx)) = (-i\overrightarrow{\alpha}.\overrightarrow{\nabla} + m\beta)\,u^{(s)}(\overrightarrow{p},m)\exp(-ipx) \tag{4.162}$$

$$i\frac{\partial}{\partial t}(u^{(s)}(\overrightarrow{p},m)\exp(ipx)) = (-i\overrightarrow{\alpha}.\overrightarrow{\nabla} - m\beta)\,u^{(s)}(\overrightarrow{p},m)\exp(ipx) \qquad (4.163)$$

This leads us to create a summary table—which will be very useful later on—in which we summarize all eight solutions, four associated with a positive mass and four associated with a negative mass:

Positive mass	Negative mass
$u^{(s)}(\overrightarrow{p},m)\exp(-ipx)$	$u^{(s)}(\overrightarrow{p},m)\exp(ipx)$
$v^{(s)}(\overrightarrow{p},-m)\exp(ipx)$	$v^{(s)}(\overrightarrow{p},-m)\exp(-ipx)$
$v^{(s)}(-\overrightarrow{p},m)\exp(ipx)$	$v^{(s)}(-\overrightarrow{p},m)\exp(-ipx)$
$u^{(s)}(-\overrightarrow{p},-m)\exp(-ipx)$	$u^{(s)}(-\overrightarrow{p},-m)\exp(ipx)$

One field that differs from Dirac's—and that one might think of—is the "positive mass" field:

$$\begin{aligned}\psi(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s((C1\,a_s(\overrightarrow{p})+C2\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(-ipx)\\ +(C3\,b_s(\overrightarrow{p})+C4\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(ipx)\\ +(C5\,c_s(\overrightarrow{p})+C6\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(ipx)\\ +(C7\,d_s(\overrightarrow{p})+C8\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(-ipx))\end{aligned}$$

The operators $a_s(\overrightarrow{p})$ and $a_s^\dagger(\overrightarrow{p})$ annihilate and create an electron, respectively; the operators $b_s(\overrightarrow{p})$ and $b_s^\dagger(\overrightarrow{p})$ annihilate and create a positron, respectively; you get the idea...

The only nonzero anticommutation relations are

$$\{a_s(\overrightarrow{p1}),a_r^\dagger(\overrightarrow{p2})\} = \{b_s(\overrightarrow{p1}),b_r^\dagger(\overrightarrow{p2})\} = (2\pi)^3\delta^{(3)}(\overrightarrow{p1}-\overrightarrow{p2})\delta_{rs} \qquad (4.164)$$

$$\{c_s(\overrightarrow{p1}),c_r^\dagger(\overrightarrow{p2})\} = \{d_s(\overrightarrow{p1}),d_r^\dagger(\overrightarrow{p2})\} = (2\pi)^3\delta^{(3)}(\overrightarrow{p1}-\overrightarrow{p2})\delta_{rs} \qquad (4.165)$$

The coefficients $C_j, j = 1,\ldots,8$, are complex numbers that will be determined later. Clearly, the Dirac field corresponds to $C_1 = C_4 = 1; C_2 = C_3 = C_5 = C_6 = C_7 = C_8 = 0$.

The first thing to note is that

$$\psi(t=0,\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s((C1\,a_s(\overrightarrow{p})+C2\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)$$
$$+(C3\,b_s(-\overrightarrow{p})+C4\,b_s^\dagger(-\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},-m)+(C5\,c_s(-\overrightarrow{p})+C6\,c_s^\dagger(-\overrightarrow{p}))v^{(s)}(\overrightarrow{p},m)$$
$$+(C7\,d_s(\overrightarrow{p})+C8\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m))\exp(i\overrightarrow{p}.\overrightarrow{x})$$

This means that, when time is not taken into account, the "positive mass–negative energy" solutions coincide with the "positive mass–positive energy" solutions.

We can therefore simplify things by considering only

$$\psi(t=0,\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p1\,\frac{1}{\sqrt{2E_{\overrightarrow{p1}}}}\sum_s((C1\,a_s(\overrightarrow{p1})+C2\,a_s^\dagger(\overrightarrow{p1}))u^{(s)}(\overrightarrow{p1},m)$$
$$+(C3\,b_s(-\overrightarrow{p1})+C4\,b_s^\dagger(-\overrightarrow{p1}))v^{(s)}(-\overrightarrow{p1},-m))\exp(i\overrightarrow{p1}.\overrightarrow{x})$$

and

$$\psi^\dagger(t=0,\overrightarrow{y}) = \frac{1}{(2\pi)^3}\int d^3p2\,\frac{1}{\sqrt{2E_{\overrightarrow{p2}}}}\sum_r((C1^*\,a_r^\dagger(\overrightarrow{p2})+C2^*\,a_r(\overrightarrow{p2}))u^{(r)\dagger}(\overrightarrow{p2},m)$$
$$+(C3^*\,b_r^\dagger(-\overrightarrow{p2})+C4^*\,b_r(-\overrightarrow{p2}))v^{(r)\dagger}(-\overrightarrow{p2},-m))\exp(-i\overrightarrow{p2}.\overrightarrow{y})$$

The anticommutation relations involving the annihilation and creation operators then imply

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{2E_{\overrightarrow{p}}}\sum_s((|C1|^2+|C2|^2)u^{(s)}(\overrightarrow{p},m)\overline{u}^{(s)}(\overrightarrow{p},m)$$
$$+(|C3|^2+|C4|^2)v^{(s)}(-\overrightarrow{p},-m)\overline{v}^{(s)}(-\overrightarrow{p},-m))\gamma^0\exp(i\overrightarrow{p}.(\overrightarrow{x}-\overrightarrow{y}))$$

And, using the properties of spinors, we obtain

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{2E_{\overrightarrow{p}}}\sum_s((|C1|^2+|C2|^2)(E_{\overrightarrow{p}}\gamma^0-\overrightarrow{\gamma}.\overrightarrow{p}+m)$$
$$+(|C3|^2+|C4|^2)(E_{\overrightarrow{p}}\gamma^0+\overrightarrow{\gamma}.\overrightarrow{p}-m))\gamma^0\exp(i\overrightarrow{p}.(\overrightarrow{x}-\overrightarrow{y}))$$

The expected relation

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \delta^{(3)}(\overrightarrow{x}-\overrightarrow{y}) \tag{4.166}$$

follows from this if we impose

$$|C1|^2 + |C2|^2 = |C3|^2 + |C4|^2 = 1 \tag{4.167}$$

Next, we need to calculate the Hamiltonian

$$H = \int d^3x\, \psi^\dagger(\overrightarrow{x})\gamma^0(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\psi(\overrightarrow{x}) \tag{4.168}$$

The first step is to note that

$$\begin{aligned}(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\psi(\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p\, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \\ \sum_s((C1\, a_s(\overrightarrow{p}) + C2\, a_s^\dagger(\overrightarrow{p}))(\overrightarrow{\gamma}.\overrightarrow{p} + m)u^{(s)}(\overrightarrow{p}, m) \\ + (C3\, b_s(-\overrightarrow{p}) + C4\, b_s^\dagger(-\overrightarrow{p}))(\overrightarrow{\gamma}.\overrightarrow{p} + m)v^{(s)}(-\overrightarrow{p}, -m))\exp(i\overrightarrow{p}.\overrightarrow{x})\end{aligned}$$

And, recalling the equations governing the spinors $u^{(s)}(\overrightarrow{p}, m)$ and $v^{(s)}(-\overrightarrow{p}, -m)$, we obtain

$$\begin{aligned}(-i\overrightarrow{\gamma}.\overrightarrow{\nabla} + m)\psi(\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p \sqrt{\frac{E_{\overrightarrow{p}}}{2}}\,\gamma^0 \\ \sum_s((C1\, a_s(\overrightarrow{p}) + C2\, a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p}, m) \\ - (C3\, b_s(-\overrightarrow{p}) + C4\, b_s^\dagger(-\overrightarrow{p}))v^{(s)}(-\overrightarrow{p}, -m))\exp(i\overrightarrow{p}.\overrightarrow{x})\end{aligned}$$

The final Hamiltonian is then given by

$$\begin{aligned}H = \frac{1}{(2\pi)^3}\int d^3p \sum_s((C1^*\, a_s^\dagger(\overrightarrow{p}) + C2^*\, a_s(\overrightarrow{p}))(C1\, a_s(\overrightarrow{p}) + C2\, a_s^\dagger(\overrightarrow{p})) \\ - (C3^*\, b_s^\dagger(\overrightarrow{p}) + C4^*\, b_s(\overrightarrow{p}))(C3\, b_s(\overrightarrow{p}) + C4\, b_s^\dagger(\overrightarrow{p}))\end{aligned}$$

There are two points to discuss.

1. $C2 = C3 = 0\,;\ |C1|^2 = |C4|^2 = 1$

This corresponds to the annihilation of an electron and the creation of a positron, which is the standard theory.

The Hamiltonian is

$$H = \frac{1}{(2\pi)^3}\int d^3p \sum_s(a_s^\dagger(\overrightarrow{p})a_s(\overrightarrow{p}) + b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p})) - E_0 \tag{4.169}$$

with

$$E_0 = 2\int d^3p\, E_{\overrightarrow{p}}\,\delta^{(3)}(0) \tag{4.170}$$

$2.C1 = C4 = 0\,;\ |C2|^2 = |C3|^2 = 1$

This corresponds to the annihilation of a positron and the creation of an electron.

The Hamiltonian is then

$$H = \frac{1}{(2\pi)^3}\int d^3p \sum_s(-a_s^\dagger(\overrightarrow{p})a_s(\overrightarrow{p}) - b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p})) + E_0 \tag{4.171}$$

This means that the vacuum energy is still associated with a divergence, but of the opposite sign compared to the previous case.

Of course, this shouldn't particularly excite us, since it's just a "switch" between destruction and creation.

But let's keep this result in mind as we move on to the next case, which is that of negative masses.

If we return to the summary table, we see that it's natural to propose defining the field of negative masses by[4]

$$\begin{aligned}\psi(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s((&D1\,a_s(\overrightarrow{p}) + D2\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(ipx)\\ &+ (D3\,b_s(\overrightarrow{p}) + D4\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(-ipx)\\ &+ (D5\,c_s(\overrightarrow{p}) + D6\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(-ipx)\\ &+ (D7\,d_s(\overrightarrow{p}) + D8\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(ipx))\end{aligned}$$

As noted earlier, we see that, when time is not taken into account, the "negative mass–negative energy" and "negative mass–positive energy" solutions are inseparable.

We therefore have

$$\begin{aligned}\psi(t=0,\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p1\,\frac{1}{\sqrt{2E_{\overrightarrow{p1}}}}\sum_s((&D1\,a_s(-\overrightarrow{p1}) + D2\,a_s^\dagger(-\overrightarrow{p1}))u^{(s)}(-\overrightarrow{p1},m)\\ &+ (D3\,b_s(\overrightarrow{p1}) + D4\,b_s^\dagger(\overrightarrow{p1}))v^{(s)}(\overrightarrow{p1},-m))\exp(i\overrightarrow{p1}.\overrightarrow{x})\end{aligned}$$

[4]The annihilation and creation operators have been designated by the same letters, just to avoid going through the entire alphabet... but it is clear that these are different operators from those used for positive masses.

and

$$\psi^\dagger(t=0,\overrightarrow{y}) = \frac{1}{(2\pi)^3}\int d^3p2\,\frac{1}{\sqrt{2E_{\overrightarrow{p2}}}}\sum_r((D1^*\,a_r^\dagger(-\overrightarrow{p2})+D2^*\,a_r(-\overrightarrow{p2}))u^{(r)\dagger}(-\overrightarrow{p2},m) + (D3^*\,b_r^\dagger(\overrightarrow{p2})+D4^*\,b_r(\overrightarrow{p2}))v^{(r)\dagger}(\overrightarrow{p2},-m))\exp(-i\overrightarrow{p2}.\overrightarrow{y})$$

We then successively have

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{2E_{\overrightarrow{p}}}\sum_s((|D1|^2+|D2|^2)u^{(s)}(-\overrightarrow{p},m)\overline{u}^{(s)}(-\overrightarrow{p},m) + (|D3|^2+|D4|^2)v^{(s)}(\overrightarrow{p},-m)\overline{v}^{(s)}(\overrightarrow{p},-m))\gamma^0\exp(i\overrightarrow{p}.(\overrightarrow{x}-\overrightarrow{y}))$$

and

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{2E_{\overrightarrow{p}}}\sum_s((|D1|^2+|D2|^2)(E_{\overrightarrow{p}}\gamma^0+\overrightarrow{\gamma}.\overrightarrow{p}+m) + (|D3|^2+|D4|^2)(E_{\overrightarrow{p}}\gamma^0-\overrightarrow{\gamma}.\overrightarrow{p}-m))\gamma^0\exp(i\overrightarrow{p}.(\overrightarrow{x}-\overrightarrow{y}))$$

We do indeed find

$$\{\psi(\overrightarrow{x}),\psi^\dagger(\overrightarrow{y})\} = \delta^{(3)}(\overrightarrow{x}-\overrightarrow{y}) \tag{4.172}$$

provided that the coefficients are set according to

$$|D1|^2+|D2|^2 = |D3|^2+|D4|^2 = 1 \tag{4.173}$$

Now it's the Hamiltonian's turn! Here we have

$$H = \overline{\psi}(-i\overrightarrow{\gamma}.\overrightarrow{\nabla}-m)\psi \tag{4.174}$$

You're starting to get the hang of these calculations...

It comes

$$(-i\overrightarrow{\gamma}.\overrightarrow{\nabla}-m)\psi(\overrightarrow{x}) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \sum_s((D1\,a_s(-\overrightarrow{p})+D2\,a_s^\dagger(-\overrightarrow{p}))(\overrightarrow{\gamma}.\overrightarrow{p}-m)u^{(s)}(-\overrightarrow{p},m) + (D3\,b_s(\overrightarrow{p})+D4\,b_s^\dagger(\overrightarrow{p}))(\overrightarrow{\gamma}.\overrightarrow{p}-m)v^{(s)}(\overrightarrow{p},-m))\exp(i\overrightarrow{p}.\overrightarrow{x})$$

from which it follows that

$$(-i\overrightarrow{\gamma}.\overrightarrow{\nabla}-m)\psi(\overrightarrow{x})=\frac{1}{(2\pi)^3}\int d^3p\sqrt{\frac{E_{\overrightarrow{p}}}{2}}\,\gamma^0$$
$$\sum_s(-(D1\,a_s(-\overrightarrow{p})+D2\,a_s^\dagger(-\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},m)$$
$$+(D3\,b_s(\overrightarrow{p})+D4\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m))\exp(i\overrightarrow{p}.\overrightarrow{x})$$

So, the Hamiltonian is ultimately

$$H=\frac{1}{(2\pi)^3}\int d^3p\sum_s(-(D1^*\,a_s^\dagger(\overrightarrow{p})+D2^*\,a_s(\overrightarrow{p}))(D1\,a_s(\overrightarrow{p})+D2\,a_s^\dagger(\overrightarrow{p}))$$
$$+(D3^*\,b_s^\dagger(\overrightarrow{p})+D4^*\,b_s(\overrightarrow{p}))(D3\,b_s(\overrightarrow{p})+D4\,b_s^\dagger(\overrightarrow{p}))$$

Once again, we have two cases to discuss.

1.$D2=D3=0\,;\ |D1|^2=|D4|^2=1$

This corresponds to the annihilation of an "electron" with negative mass and the creation of a "positron" that also has negative mass.

The Hamiltonian is

$$H=\frac{1}{(2\pi)^3}\int d^3p\sum_s(-a_s^\dagger(\overrightarrow{p})a_s(\overrightarrow{p})-b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p}))+E_0 \tag{4.175}$$

2.$D1=D4=0\,;\ |D2|^2=|D3|^2=1$

This corresponds to the annihilation of a "positron" and the creation of an "electron".

The Hamiltonian is then

$$H=\frac{1}{(2\pi)^3}\int d^3p\sum_s(a_s^\dagger(\overrightarrow{p})a_s(\overrightarrow{p})+b_s^\dagger(\overrightarrow{p})b_s(\overrightarrow{p}))-E_0 \tag{4.176}$$

This is where things get interesting right away! In fact, it is possible to offset the negative divergence observed when an electron is destroyed and a positron is created by the positive divergence calculated when their negatively charged counterparts are annihilated and created...

This result is attributed to Canadian researcher Benoît Guay[5].

And that is exactly why supersymmetric theories were developed!

However, despite long, intensive, and costly research, no experiment has yet detected even a single supersymmetric particle...

Perhaps if we looked into the realm of negative masses, things might change...

[5]Guay, B. (2022). Experimental test to verify whether quantum states of conjugated mass and energy exist. *Mod. Phys. Lett. A*37, 2250013 https://doi.org/10.1142/S0217732322500134

4.4.3 The Effect of Discrete Symmetry Operators on Solutions

Let's examine the effects of the various parity and time-reversal operators on the solutions to the Dirac equation.

4.4.3.1 The unitary parity operator P_U

As a reminder,

$$P_U^{-1} a_s(\overrightarrow{p}) P_U = \eta_a^* a_s(-\overrightarrow{p})\,;\; P_U^{-1} a_s^\dagger(\overrightarrow{p}) P_U = \eta_a a_s^\dagger(-\overrightarrow{p}) \tag{4.177}$$

And, of course, similar relations hold true for the other operators of annihilation and creation.

Its effect on the positive-mass field is therefore as follows:

$$\begin{aligned} P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (C1\,\eta_a^*\,a_s(-\overrightarrow{p}) + C2\,\eta_a\,a_s^\dagger(-\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(-ipx) \\ + (C3\,\eta_b^*\,b_s(-\overrightarrow{p}) + C4\,\eta_b\,b_s^\dagger(-\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(ipx) \\ + (C5\,\eta_c^*\,c_s(-\overrightarrow{p}) + C6\,\eta_c\,c_s^\dagger(-\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(ipx) \\ + (C7\,\eta_d^*\,d_s(-\overrightarrow{p}) + C8\,\eta_d\,d_s^\dagger(-\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(-ipx)) \end{aligned}$$

By reintroducing $\widehat{x} \equiv (t, -\overrightarrow{x})$, we can write

$$\begin{aligned} P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (C1\,\eta_a^*\,a_s(\overrightarrow{p}) + C2\,\eta_a\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},m)\exp(-ip\widehat{x}) \\ + (C3\,\eta_b^*\,b_s(\overrightarrow{p}) + C4\,\eta_b\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},-m)\exp(ip\widehat{x}) \\ + (C5\,\eta_c^*\,c_s(\overrightarrow{p}) + C6\,\eta_c\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},m)\exp(ip\widehat{x}) \\ + (C7\,\eta_d^*\,d_s(\overrightarrow{p}) + C8\,\eta_d\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},-m)\exp(-ip\widehat{x})) \end{aligned}$$

It's time to go back to the spinors and see that

$$u^{(s)}(\mp\overrightarrow{p},m) = \gamma^0 u^{(s)}(\pm\overrightarrow{p},m)\,;\; v^{(s)}(\mp\overrightarrow{p},m) = -\gamma^0 v^{(s)}(\pm\overrightarrow{p},m) \tag{4.178}$$

to get

$$P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\gamma^0\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (C1\,\eta_a^*\,a_s(\overrightarrow{p})+C2\,\eta_a\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(-ip\widehat{x})$$
$$-(C3\,\eta_b^*\,b_s(\overrightarrow{p})+C4\,\eta_b\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(ip\widehat{x})$$
$$-(C5\,\eta_c^*\,c_s(\overrightarrow{p})+C6\,\eta_c\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(ip\widehat{x})$$
$$+(C7\,\eta_d^*\,d_s(\overrightarrow{p})+C8\,\eta_d\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(-ip\widehat{x}))$$

Finally, we have

$$P_U^{-1}\psi(t,\overrightarrow{x})P_U = \eta_a\,\gamma^0\,\psi(t,-\overrightarrow{x}) \tag{4.179}$$

provided that all parities η are real and satisfy $\eta_a = -\eta_b = -\eta_c = \eta_d$.

The same reasoning applies to the negative mass field:

$$P_U^{-1}\psi(x)P_U = \frac{1}{(2\pi)^3}\gamma^0\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (D1\,\eta_a^*\,a_s(\overrightarrow{p})+D2\,\eta_a\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(ip\widehat{x})$$
$$-(D3\,\eta_b^*\,b_s(\overrightarrow{p})+D4\,\eta_b\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(-ip\widehat{x})$$
$$-(D5\,\eta_c^*\,c_s(\overrightarrow{p})+D6\,\eta_c\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(-ip\widehat{x})$$
$$+(D7\,\eta_d^*\,d_s(\overrightarrow{p})+D8\,\eta_d\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(ip\widehat{x}))$$

and

$$P_U^{-1}\psi(t,\overrightarrow{x})P_U = \eta_a\,\gamma^0\,\psi(t,-\overrightarrow{x}) \tag{4.180}$$

4.4.3.2 The antiunitary time-reversal operator T_{AU}

We would also like to highlight his work:

$$T_{AU}^{-1}a_s(\overrightarrow{p})T_{AU} = -i\eta_a^*(\sigma 2)_{ss'}a_{s'}(-\overrightarrow{p})\,;\; T_{AU}^{-1}a_s^\dagger(\overrightarrow{p})T_{AU} = -i\eta_a(\sigma 2)_{ss'}a_s^\dagger(-\overrightarrow{p}) \tag{4.181}$$

It should also be noted that $px = p^0x^0 - \overrightarrow{p}\cdot\overrightarrow{x} \to -p\widetilde{x}$, where $\widetilde{x} \equiv (-t,\overrightarrow{x})$. Furthermore, since the operator T_{AU} is antiunitary, it changes the sign of the imaginary unit i. Overall, the exponentials do not change sign.

As for the field of positive masses, we have

$$T_{AU}^{-1}\psi(x)T_{AU} = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}$$
$$\sum_s(-iC1\,\eta_a^*\,(\sigma 2)_{ss'}a_{s'}(\overrightarrow{p}) - iC2\,\eta_a\,(\sigma 2)_{ss'}a_{s'}^\dagger(\overrightarrow{p}))u^{(s)*}(-\overrightarrow{p},m)\exp(-ip\widetilde{x})$$
$$+(-iC3\,\eta_b^*\,(\sigma 2)_{ss'}b_{s'}(\overrightarrow{p}) - iC4\,\eta_b\,(\sigma 2)_{ss'}b_{s'}^\dagger(\overrightarrow{p}))v^{(s)*}(-\overrightarrow{p},-m)\exp(ip\widetilde{x})$$
$$+(-iC5\,\eta_c^*\,(\sigma 2)_{ss'}c_{s'}(\overrightarrow{p}) - iC6\,\eta_c\,(\sigma 2)_{ss'}c_{s'}^\dagger(\overrightarrow{p}))v^{(s)*}(\overrightarrow{p},m)\exp(ip\widetilde{x})$$
$$+(-iC7\,\eta_d^*\,(\sigma 2)_{ss'}d_{s'}(\overrightarrow{p}) - iC8\,\eta_d\,(\sigma 2)_{ss'}d_{s'}^\dagger(\overrightarrow{p}))u^{(s)*}(\overrightarrow{p},-m)\exp(-ip\widetilde{x})$$

The goal, then, is to find a matrix M such that

$$Mu^{(-)}(\overrightarrow{p},m) = -u^{(+)*}(-\overrightarrow{p},m)\,;\; Mu^{(+)}(\overrightarrow{p},m) = u^{(-)*}(-\overrightarrow{p},m) \tag{4.182}$$

$$Mv^{(-)}(\overrightarrow{p},-m) = -v^{(+)*}(-\overrightarrow{p},-m)\,;\; Mv^{(+)}(\overrightarrow{p},-m) = v^{(-)*}(-\overrightarrow{p},-m) \tag{4.183}$$

$$Mv^{(-)}(-\overrightarrow{p},m) = -v^{(+)*}(\overrightarrow{p},m)\,;\; Mv^{(+)}(-\overrightarrow{p},m) = v^{(-)*}(\overrightarrow{p},m) \tag{4.184}$$

$$Mu^{(-)}(-\overrightarrow{p},-m) = -u^{(+)*}(\overrightarrow{p},-m)\,;\; Mu^{(+)}(-\overrightarrow{p},-m) = u^{(-)*}(\overrightarrow{p},-m) \tag{4.185}$$

The matrix $M = -\gamma^1\gamma^3$ clearly satisfies the requirements.

Thus, we have

$$T_{AU}^{-1}\psi(t,\overrightarrow{x})T_{AU} = -\eta_a\,\gamma^1\gamma^3\,\psi(-t,\overrightarrow{x}) \tag{4.186}$$

provided that the η's are real and $\eta_a = \eta_b = \eta_c = \eta_d$.

In the field of negative masses, the same relations must be imposed, and consequently, the same matrix satisfies the requirements. Thus, we again have the same relation.

4.4.3.3 The antiunitary parity operator P_{AU}

Here is an operator that, to the best of our knowledge, has never been discussed in the literature... in quantum field theory, at least.

It is such that

$$[P_{AU},\overrightarrow{P}] = \{P_{AU},P^0\} = \{P_{AU},\overrightarrow{J}\} = [P_{AU},\overrightarrow{K}] = 0 \tag{4.187}$$

The parity operator anticommutes with rotations, which implies a reverse spin. We therefore set

$$P_{AU}|\overrightarrow{p}=0,s=\frac{1}{2}>=\eta\,|\overrightarrow{p}=0,s=-\frac{1}{2}> \tag{4.188}$$

For the spin-down state, we have, in order

$$P_{AU}|\overrightarrow{p}=0,s=-\frac{1}{2}>=P_{AU}(J_x-iJ_y)|\overrightarrow{p}=0,s=\frac{1}{2}>=-(J_x+iJ_y)P_{AU}|\overrightarrow{p}=0,s=\frac{1}{2}>$$

The last equality follows from the fact that the parity operator anticommutes with $\overrightarrow{J}$ and changes the sign of i.

We then have

$$P_{AU}|\overrightarrow{p}=0,s=-\frac{1}{2}>=-\eta\,|\overrightarrow{p}=0,s=\frac{1}{2}> \tag{4.189}$$

We come out of the rest with a boost:

$$P_{AU}|\overrightarrow{p},s_{\pm}>=P_{AU}B(\overrightarrow{p})|\overrightarrow{p}=0,s_{\pm}>$$

Here, P_{AU} acts on the imaginary part but not on the $\overrightarrow{K}$ with which it commutes.

Thus, we have

$$\begin{aligned}P_{AU}|\overrightarrow{p},s_{\pm}>&=P_{AU}B(\overrightarrow{p})|\overrightarrow{p}=0,s_{\pm}>=P_{AU}B(\overrightarrow{p})P_{AU}^{-1}P_{AU}|\overrightarrow{p}=0,s_{\pm}>\\&=\pm\eta P_{AU}B(\overrightarrow{p})P_{AU}^{-1}|\overrightarrow{p}=0,s_{\mp}>=\pm\eta B(-\overrightarrow{p})|\overrightarrow{p}=0,s_{\mp}>\end{aligned}$$

In the end, we get

$$P_{AU}|\overrightarrow{p},s_{\pm}>=\pm\eta|-\overrightarrow{p},s_{\mp}> \tag{4.190}$$

The change in the phase's sign has certain consequences. The first is that $P_{AU}^2=-I$ since

$$P_{AU}^2|\overrightarrow{p},s_{+}>=\eta^*P_{AU}|-\overrightarrow{p},s_{-}>=-|\eta|^2|\overrightarrow{p},s_{+}>=-|\overrightarrow{p},s_{+}> \tag{4.191}$$

$$P_{AU}^2|\overrightarrow{p},s_{-}>=-\eta^*P_{AU}|-\overrightarrow{p},s_{+}>=-|\eta|^2|\overrightarrow{p},s_{-}>=-|\overrightarrow{p},s_{-}> \tag{4.192}$$

and the second is the consequence that follows from the first:

$$P_{AU}^{-1}=-P_{AU} \tag{4.193}$$

Note that the equations can also be written in this form:

$$P_{AU}|\overrightarrow{p},s>=i\eta\,(\sigma 2)_{ss'}|-\overrightarrow{p},s'> \tag{4.194}$$

For now, there is no difference from $T_{AU}\ldots$ but we must take the following result into account.

According to the results of quantum mechanics, P_{AU} is fixed according to

$$P_{AU} = \gamma^0\gamma^2()^* \tag{4.195}$$

That means we're going to have to assess

$$P_{AU}u^{(s)}(\pm\overrightarrow{p}, m)\ ;\ P_{AU}v^{(s)}(\pm\overrightarrow{p}, -m)$$

We have

$$P_{AU}u^{(s)}(\overrightarrow{p}, m) \sim \begin{pmatrix} 0 & \sigma 2 \\ \sigma 2 & 0 \end{pmatrix} \begin{pmatrix} \xi^{(s)} \\ \frac{\overrightarrow{\sigma}.\overrightarrow{p}}{E+m}\xi^{(s)} \end{pmatrix} = \begin{pmatrix} -\frac{\overrightarrow{\sigma}^*.\overrightarrow{p}}{E+m}\sigma 2\xi^{(s)} \\ \sigma 2\xi^{(s)} \end{pmatrix} \tag{4.196}$$

For $s = +$, this leads to

$$P_{AU}u^{(+)}(\overrightarrow{p}, m) = iv^{(-)*}(-\overrightarrow{p}, -m) \tag{4.197}$$

and for $s = -$

$$P_{AU}u^{(-)}(\overrightarrow{p}, m) = -iv^{(+)*}(-\overrightarrow{p}, -m) \tag{4.198}$$

We will write

$$P_{AU}u^{(s)}(\overrightarrow{p}, m) = -(\sigma 2)_{ss'}v^{(s')*}(-\overrightarrow{p}, -m) \tag{4.199}$$

Similarly, it is shown that

$$P_{AU}u^{(s)}(-\overrightarrow{p}, m) = -(\sigma 2)_{ss'}v^{(s')*}(\overrightarrow{p}, -m) \tag{4.200}$$

$$P_{AU}v^{(s)}(\overrightarrow{p}, -m) = -(\sigma 2)_{ss'}u^{(s')*}(-\overrightarrow{p}, m) \tag{4.201}$$

$$P_{AU}v^{(s)}(-\overrightarrow{p}, -m) = -(\sigma 2)_{ss'}u^{(s')*}(\overrightarrow{p}, m) \tag{4.202}$$

We will need to take this into account for the P_{AU} actions on the creation and annihilation operators.

So we have

$$P_{AU}^{-1} a_s(\overrightarrow{p}) P_{AU} = \eta_a^* (\sigma 2)_{ss'} b_{s'}(-\overrightarrow{p}) \; ; \; P_{AU}^{-1} a_s^\dagger(\overrightarrow{p}) P_{AU} = -\eta_a (\sigma 2)_{ss'} b_{s'}^\dagger(-\overrightarrow{p}) \tag{4.203}$$

$$P_{AU}^{-1} b_s(\overrightarrow{p}) P_{AU} = \eta_b^* (\sigma 2)_{ss'} a_{s'}(-\overrightarrow{p}) \; ; \; P_{AU}^{-1} b_s^\dagger(\overrightarrow{p}) P_{AU} = -\eta_b (\sigma 2)_{ss'} a_{s'}^\dagger(-\overrightarrow{p}) \tag{4.204}$$

$$P_{AU}^{-1} c_s(\overrightarrow{p}) P_{AU} = \eta_c^* (\sigma 2)_{ss'} d_{s'}(-\overrightarrow{p}) \; ; \; P_{AU}^{-1} c_s^\dagger(\overrightarrow{p}) P_{AU} = -\eta_c (\sigma 2)_{ss'} d_{s'}^\dagger(-\overrightarrow{p}) \tag{4.205}$$

$$P_{AU}^{-1} d_s(\overrightarrow{p}) P_{AU} = \eta_d^* (\sigma 2)_{ss'} c_{s'}(-\overrightarrow{p}) \; ; \; P_{AU}^{-1} d_s^\dagger(\overrightarrow{p}) P_{AU} = -\eta_d (\sigma 2)_{ss'} c_{s'}^\dagger(-\overrightarrow{p}) \tag{4.206}$$

It is interesting to note that the effect of P_{AU} is, in a sense, the combined action of T_{AU} (for the change $\overrightarrow{p} \to -\overrightarrow{p}$ and $s_\pm \to s_\mp$) and the charge conjugation operator (in the sense that a$\leftrightarrow b, a^\dagger \leftrightarrow b^\dagger$). *So here is another way to obtain antimatter.*

Let's apply these results to the positive mass field.

$$\begin{aligned} P_{AU}^{-1} \psi(x) P_{AU} = \frac{1}{(2\pi)^3} \int d^3p \, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \\ \sum_s (C1 \, \eta_a^* \, (\sigma 2)_{ss'} b_{s'}(\overrightarrow{p}) - C2 \, \eta_a \, (\sigma 2)_{ss'} \, b_{s'}^\dagger(\overrightarrow{p})) u^{(s)*}(-\overrightarrow{p}, m) \exp(ip\widehat{x}) \\ + (C3 \, \eta_b^* \, (\sigma 2)_{ss'} \, a_{s'}(\overrightarrow{p}) - C4 \, \eta_b \, (\sigma 2)_{ss'} \, a_{s'}^\dagger(\overrightarrow{p})) v^{(s)*}(-\overrightarrow{p}, -m) \exp(-ip\widehat{x}) \\ + (C5 \, \eta_c^* \, (\sigma 2)_{ss'} \, d_{s'}(\overrightarrow{p}) - C6 \, \eta_c \, (\sigma 2)_{ss'} \, d_{s'}^\dagger(\overrightarrow{p})) v^{(s)*}(\overrightarrow{p}, m) \exp(-ip\widehat{x}) \\ + (C7 \, \eta_d^* \, (\sigma 2)_{ss'} \, c_{s'}(\overrightarrow{p}) - C8 \, \eta_d \, (\sigma 2)_{ss'} \, c_{s'}^\dagger(\overrightarrow{p})) u^{(s)*}(\overrightarrow{p}, -m) \exp(ip\widehat{x})) \end{aligned}$$

Given the results we have just shown, this field is equivalent to

$$\begin{aligned} P_{AU}^{-1} \psi(x) P_{AU} = \frac{1}{(2\pi)^3} \gamma^0 \gamma^2 \int d^3p \, \frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \\ \sum_s (-C1 \, \eta_a^* \, b_s(\overrightarrow{p}) + C2 \, \eta_a \, b_s^\dagger(\overrightarrow{p})) v^{(s)}(\overrightarrow{p}, -m) \exp(ip\widehat{x}) \\ + (-C3 \, \eta_b^* \, a_s(\overrightarrow{p}) + C4 \, \eta_b \, a_s^\dagger(\overrightarrow{p})) u^{(s)}(\overrightarrow{p}, m) \exp(-ip\widehat{x}) \\ + (-C7 \, \eta_d^* \, c_s(\overrightarrow{p}) + C8 \, \eta_d \, c_s^\dagger(\overrightarrow{p})) v^{(s)}(-\overrightarrow{p}, m) \exp(ip\widehat{x})) \\ + (-C5 \, \eta_c^* \, d_s(\overrightarrow{p}) + C6 \, \eta_c \, d_s^\dagger(\overrightarrow{p})) u^{(s)}(-\overrightarrow{p}, -m) \exp(-ip\widehat{x}) \end{aligned}$$

To recover the entire field, we must set

$$-C3\eta_b^* = C\,C1,\; C\,C1\eta_a^* = -C^2\,C3,\; C4\eta_b = C\,C2,\; C2\eta_a = C\,C4 \qquad (4.207)$$

$$-C7\eta_d^* = C\,C5, -C5\eta_c^* = C\,C7,\; C8\eta_d = C\,C6,\; C6\eta_c = C\,C8 \qquad (4.208)$$

However, these constraints are incompatible with what we noted regarding the Hamiltonian.

For example, the first case is $C2 = C3 = 0, |C1|^2 = |C4|^2 = 1$, and this clearly contradicts what is stated above...

The result will, in fact, be that starting from one field, we will find another field. This is a result that Frédéric Henry-Couannier[6] had already obtained concerning the unitary time-reversal operator.

Thus, starting from the Dirac field

$$\psi(x) = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (C1a_s(\overrightarrow{p})u^{(s)}(\overrightarrow{p},m)\exp(-ipx) + C4b_s^\dagger(\overrightarrow{p})v^{(s)}(\overrightarrow{p},-m)\exp(ipx))$$

we find

$$P_{AU}^{-1}\psi(x)P_{AU} = \frac{1}{(2\pi)^3}\gamma^0\gamma^2\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}\sum_s (-C1\eta_a^* b_s(\overrightarrow{p})v^{(s)}(\overrightarrow{p},-m)\exp(ip\widehat{x}) + C4\eta_b a_s^\dagger(\overrightarrow{p})u^{(s)}(\overrightarrow{p},m)\exp(-ip\widehat{x}))$$

i.e. the creation of an electron that has undergone the transformation $\overrightarrow{x} \to -\overrightarrow{x}$ and the annihilation of a positron, using the same geometric operation.

The situation is identical for the field of negative masses, since we have:

$$\begin{aligned}P_{AU}^{-1}\psi(x)P_{AU} = {} & \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}} \\ & \sum_s (D1\,\eta_a^*\,(\sigma 2)_{ss'}\,b_{s'}(\overrightarrow{p}) - D2\,\eta_a\,(\sigma 2)_{ss'}\,b_{s'}^\dagger(\overrightarrow{p}))u^{(s)*}(-\overrightarrow{p},m)\exp(-ip\widehat{x}) \\ & + (D3\,\eta_b^*\,(\sigma 2)_{ss'}\,a_{s'}(\overrightarrow{p}) - D4\,\eta_b\,(\sigma 2)_{ss'}\,a_{s'}^\dagger(\overrightarrow{p}))v^{(s)*}(-\overrightarrow{p},-m)\exp(ip\widehat{x}) \\ & + (D5\,\eta_c^*\,(\sigma 2)_{ss'}\,d_{s'}(\overrightarrow{p}) - D6\,\eta_c\,(\sigma 2)_{ss'}\,d_{s'}^\dagger(\overrightarrow{p}))v^{(s)*}(\overrightarrow{p},m)\exp(ip\widehat{x}) \\ & + (D7\,\eta_d^*\,(\sigma 2)_{ss'}\,c_{s'}(\overrightarrow{p}) - D8\,\eta_d\,(\sigma 2)_{ss'}\,c_{s'}^\dagger(\overrightarrow{p}))u^{(s)*}(\overrightarrow{p},-m)\exp(-ip\widehat{x}))\end{aligned}$$

It follows that

[6] Henry-Couannier, F. (2005). Negative energies and time reversal in Quantum Field Theory, Preprint hal-00001476v6

$$P_{AU}^{-1}\psi(x)P_{AU} = \frac{1}{(2\pi)^3}\gamma^0\gamma^2\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}$$
$$\sum_s(-D1\,\eta_a^*\,b_s(\overrightarrow{p}) + D2\,\eta_a\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(-ip\widehat{x})$$
$$+(-D3\,\eta_b^*\,a_s(\overrightarrow{p}) + D4\,\eta_b\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(ip\widehat{x})$$
$$+(-D5\,\eta_c^*\,d_s(\overrightarrow{p}) + D6\,\eta_c\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(ip\widehat{x})$$
$$+(-D7\,\eta_d^*\,c_s(\overrightarrow{p}) + D8\,\eta_d\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(-ip\widehat{x}))$$

Once again, to recover the entire field would require setting

$$C\,D1 = -D3\eta_b^*,\ C\,D2 = D4\eta_b,\ C\,D3 = -D1\eta_a^*,\ C\,D4 = D2\eta_a \tag{4.209}$$

$$C\,D5 = -D7\eta_d^*,\ C\,D6 = D8\eta_d,\ C\,D7 = -D5\eta_c^*,\ C\,D8 = D6\eta_c \tag{4.210}$$

This is incompatible with the constraints associated with the Hamiltonian.

So, once again, we find another field arising from a certain field. Thus, the annihilation of a "negative-mass electron" and the creation of a "negative-mass positron" will become the annihilation of a "negative-mass positron" and the creation of a "negative-mass electron" both of which have undergone the mirror symmetry $\overrightarrow{x} \to -\overrightarrow{x}$.

4.4.3.4 The unitary time-reversal operator T_U

It is such that

$$[T_U, \overrightarrow{P}] = \{T_U, P^0\} = [T_U, \overrightarrow{J}] = \{T_U, \overrightarrow{K}\} = 0 \tag{4.211}$$

So, we will have

$$T_U|\overrightarrow{p} = 0, s = \frac{1}{2} >= \eta\,|\overrightarrow{p} = 0, s = \frac{1}{2} > \tag{4.212}$$

and

$$T_U|\overrightarrow{p} = 0, s = -\frac{1}{2} >= T_U(J_x - iJ_y)|\overrightarrow{p} = 0, s = \frac{1}{2} >= (J_x - iJ_y)T_U|\overrightarrow{p} = 0, s = \frac{1}{2} >$$

to finally obtain

$$T_U|\overrightarrow{p} = 0, s = -\frac{1}{2} >= \eta\,|\overrightarrow{p} = 0, s = -\frac{1}{2} > \tag{4.213}$$

With the boost, it comes

$$T_U|\overrightarrow{p}, s_\pm >= T_U B(\overrightarrow{p})|\overrightarrow{p} = 0, s_\pm >= T_U B(\overrightarrow{p}) T_U^{-1} T_U|\overrightarrow{p} = 0, s_\pm >$$
$$= \eta T_U B(\overrightarrow{p}) T_U^{-1}|\overrightarrow{p} = 0, s_\pm >= \eta B(-\overrightarrow{p})|\overrightarrow{p} = 0, s_\pm >$$

In the end, we get

$$T_U|\overrightarrow{p}, s_\pm >= \eta|-\overrightarrow{p}, s_\pm > \tag{4.214}$$

There seems to be no difference from what P_U does... And yet, there is!

In fact, using Dirac's formulation, we have

$$T_U = \gamma^1\gamma^2\gamma^3 = \begin{pmatrix} 0 & -iI \\ iI & 0 \end{pmatrix} \tag{4.215}$$

with the result that

$$T_U u^{(s)}(\pm\overrightarrow{p}, m) = iv^{(s)}(\mp\overrightarrow{p}, -m)\,;\, T_U v^{(s)}(\pm\overrightarrow{p}, -m) = -iu^{(s)}(\mp\overrightarrow{p}, m) \tag{4.216}$$

We therefore require

$$T_U^{-1} a_s(\overrightarrow{p}) T_U = \eta_a^* b_s(-\overrightarrow{p})\,;\, T_U^{-1} a_s^\dagger(\overrightarrow{p}) T_U = \eta_a b_s^\dagger(-\overrightarrow{p}) \tag{4.217}$$

$$T_U^{-1} b_s(\overrightarrow{p}) T_U = \eta_b^* a_s(-\overrightarrow{p})\,;\, T_U^{-1} b_s^\dagger(\overrightarrow{p}) T_U = \eta_b a_s^\dagger(-\overrightarrow{p}) \tag{4.218}$$

$$T_U^{-1} c_s(\overrightarrow{p}) T_U = \eta_c^* d_s(-\overrightarrow{p})\,;\, T_U^{-1} c_s^\dagger(\overrightarrow{p}) T_U = \eta_c d_s^\dagger(-\overrightarrow{p}) \tag{4.219}$$

$$T_U^{-1} d_s(\overrightarrow{p}) T_U = \eta_d^* c_s(-\overrightarrow{p})\,;\, T_U^{-1} d_s^\dagger(\overrightarrow{p}) T_U = \eta_d c_s^\dagger(-\overrightarrow{p}) \tag{4.220}$$

In the positive mass field, this results in

$$T_U^{-1}\psi(x) T_U = \frac{1}{(2\pi)^3}\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}$$
$$\sum_s (C1\,\eta_a^*\, b_s(\overrightarrow{p}) + C2\,\eta_a\, b_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p}, m)\exp(ip\widetilde{x})$$
$$+ (C3\,\eta_b^*\, a_s(\overrightarrow{p}) + C4\,\eta_b\, a_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p}, -m)\exp(-ip\widetilde{x})$$
$$+ (C5\,\eta_c^*\, d_s(\overrightarrow{p}) + C6\,\eta_c\, d_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p}, m)\exp(-ip\widetilde{x})$$
$$+ (C7\,\eta_d^*\, c_s(\overrightarrow{p}) + C8\,\eta_d\, c_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p}, -m)\exp(ip\widetilde{x}))$$

In this expression, we note that $\exp(-ipx)$ has been replaced by $\exp(ip\widetilde{x})$.

We can also write

$$T_U^{-1}\psi(x)T_U = \frac{1}{(2\pi)^3}\gamma^1\gamma^2\gamma^3\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}$$
$$\sum_s (i(C1\,\eta_a^*\,b_s(\overrightarrow{p}) + C2\,\eta_a\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(-p,-m)\exp(ip\widetilde{x})$$
$$- i(C3\,\eta_b^*\,a_s(\overrightarrow{p}) + C4\,\eta_b\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(-ip\widetilde{x})$$
$$- i(C5\,\eta_c^*\,d_s(\overrightarrow{p}) + C6\,\eta_c\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(-ip\widetilde{x})$$
$$+ i(C7\,\eta_d^*\,c_s(\overrightarrow{p}) + C8\,\eta_d\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(ip\widetilde{x}))$$

Let's be clear: if we were to force a relation like

$$T_U^{-1}\psi(x)T_U = D\,\gamma^1\gamma^2\gamma^3\,\psi(\widetilde{x}) \tag{4.221}$$

we should require

$$D\,C1 = -iC3\eta_b^*\,;\; D\,C3 = iC1\eta_a^*\,;\; D\,C2 = -iC4\eta_b\,;\; D\,C4 = iC2\eta_a \tag{4.222}$$

$$D\,C5 = iC7\eta_d^*\,;\; D\,C7 = -iC5\eta_c^*\,;\; D\,C6 = iC8\eta_d\,;\; D\,C8 = -iC6\eta_c \tag{4.223}$$

However, such relations are incompatible with the constraints imposed by the Hamiltonian...

Once again, these are different fields that one would obtain if the annihilation and creation operators had been interchanged via the transformation $t \to -t$.

Let's finally move on to the field of negative masses. We have

$$T_U^{-1}\psi(x)T_U = \frac{1}{(2\pi)^3}\gamma^1\gamma^2\gamma^3\int d^3p\,\frac{1}{\sqrt{2E_{\overrightarrow{p}}}}$$
$$\sum_s (i(D1\,\eta_a^*\,b_s(\overrightarrow{p}) + D2\,\eta_a\,b_s^\dagger(\overrightarrow{p}))v^{(s)}(\overrightarrow{p},-m)\exp(-ip\widetilde{x})$$
$$- i(D3\,\eta_b^*\,a_s(\overrightarrow{p}) + D4\,\eta_b\,a_s^\dagger(\overrightarrow{p}))u^{(s)}(\overrightarrow{p},m)\exp(ip\widetilde{x})$$
$$- i(D5\,\eta_c^*\,d_s(\overrightarrow{p}) + D6\,\eta_c\,d_s^\dagger(\overrightarrow{p}))u^{(s)}(-\overrightarrow{p},-m)\exp(ip\widetilde{x})$$
$$+ i(D7\,\eta_d^*\,c_s(\overrightarrow{p}) + D8\,\eta_d\,c_s^\dagger(\overrightarrow{p}))v^{(s)}(-\overrightarrow{p},m)\exp(-ip\widetilde{x}))$$

This would confirm that

$$T_U^{-1}\psi(x)T_U = C\,\gamma^1\gamma^2\gamma^3\,\psi(\widetilde{x}) \tag{4.224}$$

if

$$C\,D1 = -iD3\eta_b^*\,;\, C\,D3 = iD1\eta_a^*\,;\, C\,D2 = -iD4\eta_b\,;\, C\,D4 = iD2\eta_a \quad (4.225)$$

$$C\,D5 = iD7\eta_d^*\,;\, C\,D7 = -iD5\eta_c^*\,;\, C\,D6 = iD8\eta_d\,;\, C\,D8 = -iD6\eta_c \quad (4.226)$$

But we must set these relations aside and consider only parts of fields in order to discover other parts...

In any case, these "iconoclastic" versions of P_{AU} and T_U open up new horizons...

Should we see this as the creation of a different kind of antimatter?

Only time will tell...

Conclusion

The central question addressed here is that of the physical relevance of negative masses within the framework of relativistic quantum mechanics and quantum field theory. The analysis suggests that, from the perspective of fundamental mathematical structures, there is no a priori reason to rule them out. Indeed, negative masses appear naturally in group theory, and more specifically in the unitary irreducible representations of the Poincaré group, which forms the basis of spacetime symmetries. Their presence is therefore not artificial, but results directly from the algebraic properties of the formalism.

Furthermore, it has been shown that the introduction of such masses could play a regulatory role in certain theoretical contexts. Similar to the mechanisms proposed in supersymmetry, where contributions with different spins help to cancel out divergences, particles associated with negative masses could contribute to a natural annihilation of these divergences arising in quantum field theory. This analogy suggests that studying them could offer a conceptual alternative or complement to the approaches currently being explored.

These factors argue in favor of a thorough examination of these solutions, similar to the scrutiny that has been given to supersymmetric theories. A first step—one of considerable scope—would be to reformulate quantum field theory by explicitly incorporating these additional degrees of freedom. Such an undertaking would require revisiting the very foundations of the theory, both from a formal standpoint and in terms of its physical implications.

However, even if such an extension were to prove coherent and fruitful, it would constitute only an initial step. Indeed, previous work, notably that of Terletsky[7], has highlighted remarkable properties concerning composite systems of particles with different masses. These results show that the overall nature of a system depends on the relative combination of positive and negative masses, leading to configurations whose effective mass can be positive, zero, negative, or even imaginary. In particular, the introduction of negative masses inevitably implies the possibility of imaginary masses in certain configurations.

This observation is of major conceptual importance. It suggests that the acceptance of negative masses cannot be considered in isolation from a broader expansion of the theoretical framework, one that includes objects whose prop-

[7] Terletsky, J.P. (1962). Masses propres positives, négatives et imaginaires. *Journal de Physique et le Radium*, 23, 910-920. 10.1051/jphysrad:019620023011091000

erties go beyond those traditionally accepted in particle physics. In this sense, the systematic exclusion of these solutions could limit our understanding of fundamental structures.

In conclusion, the study of negative masses is not merely a matter of theoretical curiosity, but represents a serious avenue for exploring new physical structures. It calls for a reconsideration of certain implicit assumptions of the Standard Model and paves the way for potentially fruitful developments, both mathematically and physically. The challenge now is to determine whether these extensions can be coherently integrated into an experimentally testable framework, and whether they can shed new light on the current limits of theoretical physics.