26747

National Library of Canada | Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE | THÈSES CANADIENNES SUR MICROFICHE

NAME OF AUTHOR *NOM DE L'AUTEUR* Jean-François Dumais

TITLE OF THESIS *TITRE DE LA THÈSE* Dirac Spinor Kinematics

UNIVERSITY *UNIVERSITÉ* University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED / *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* M. Sc.

YEAR THIS DEGREE CONFERRED / *ANNÉE D'OBTENTION DE CE GRADE* 1975

NAME OF SUPERVISOR *NOM DU DIRECTEUR DE THÈSE* Professor Yasushi Takahashi

DATED / *DATÉ* August 13/75 SIGNED / *SIGNÉ* Jean-François Dumais

PERMANENT ADDRESS / *RÉSIDENCE FIXE* 11019 84th st, ste #201

Edmonton, Alberta.

THE UNIVERSITY OF ALBERTA

DIRAC SPINOR KINEMATICS

by

JEAN-FRANCOIS DUMAIS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

IN

THEORETICAL PHYSICS

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA

FALL , 1975

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled DIRAC SPINOR KINEMATICS submitted by Jean-Francois Dumais in partial fulfillment of the requirements for the degree of Master of Science in Theoretical Physics.

Y. Takahashi
Supervisor

M. Kruger

M. Razavy

S.K. Sehgal

Date July 7/75

ABSTRACT

The kinematic aspects of Dirac spinors are studied. All the properties of a totally arbitrary system of gamma matrices are derived without resorting to the theory of finite group representations or the theory of Clifford algebras. An algebraic method is then devised to find explicitely the similarity transformation arising in the fundamental theorem of gamma matrices. Next the Lie group of spinor transformations under the action of the orthochronous Lorentz group is studied in detail. The work ends with a thorough analysis of all the algebraic relations among the Dirac bilinears.

PREFACE

As described by the title the object of this simple work is the kinematics of Dirac spinors. Needless to say it was meant as a review. However it is hoped that the final product is not completely devoid of originality.

Chapter I recalls how one is led to the Dirac equation and its associated gamma matrices. The relativistic invariance of the equation, discussed in the second section, provides the physical motivation for the fundamental theorem of gamma matrices.

The first section of chapter II is a standard presentation of the properties of products of gamma matrices. The second section discusses the degree and reducibility of the representations of the fundamental relations $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}I$. The eventual originality of the discussion lies in a complete avoidance of the theory of finite group representations or the theory of Clifford algebras. The resulting treatment is self-contained and elementary; this might be of some pedagogical interest.

Chapter III deals with the fundamental theorem of gamma matrices. The usual proof is modified, leading to a shorter and perhaps more elegant one. Some elementary consequences of the fundamental theorem are then discussed; for instance it is shown that there exists no system of real gamma matrices.

While the main result of chapter III asserts the existence of a certain non-singular matrix S connecting two systems of gamma matrices, it says nothing about the explicit form of S. It is the aim of chapter IV to try to fill this gap.

Chapter V discusses in detail the transformation of spinors under orthochronous Lorentz transformations. The Lie group $S^{\uparrow}$ whose elements are those transformations is carefully studied. Several different ways of describing its elements are obtained. It is finally concluded that the subgroup $S_{+}^{\uparrow}$ corresponding to proper Lorentz transformations is isomorphic to SL(2,C).

Chapter VI deals with the tensors obtained by quadratic combinations of spinors. These include a scalar and a pseudo-scalar, a vector and a pseudo-vector and a twice contravariant antisymmetric tensor. These objects are not independent of each other. Covariant identities other than those given in (Pauli [1936]) are derived and used to provide a complete solution to the question of the algebraic dependence of the tensor components. This analysis is restricted to the case where ψ is an ordinary spinor and not a field-operator.

ACKNOWLEDGEMENTS

First I wish to thank my supervisor, Professor Y. Takahashi, for suggesting to me the topic of this thesis and for his availability for discussion.

To the National Research Council of Canada and to the Ministère de l'Education du Québec I am indebted for my financial support over the last two years.

Finally I thank Mrs. Lee Cech for her fast typing of the bulk of my thesis and Mrs. Joanne Koppens for completing the job.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

CHAPTER I

THE ORIGIN OF DIRAC γ MATRICES

(1) The Dirac equation

The Dirac γ matrices arise naturally when one seeks a relativistic equation for the wave function of an electron. We suppose that this wave function ψ is a function from spacetime into C^N for some N and we are looking for a differential equation describing its behavior. To have a close analogy with the Schroedinger equation we want it to be first order in time. In order to be invariant under Lorentz transformations it will have to be of first order in the space derivatives as well. The most general form of such a linear homogeneous equation with constant coefficients, expressing the time derivative of ψ in terms of the space derivatives of ψ and of ψ itself is clearly:

$$\left(\frac{\partial}{\partial x^o} + \alpha^k \frac{\partial}{\partial x^k} + \frac{imc}{\hbar}\beta\right)\psi = 0 \qquad (1.1)$$

where α^k and β are $N\times N$ complex matrices and k runs from 1 to 3. The x^k's are the space coordinates and $x^o = ct$. The constant in front of β takes care of the dimensions appropriately if m is a mass. To be consistent with the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$ we require that ψ satifies the Klein-Gordon equation as well. Simple multiplications give:

$$\left(\frac{\partial}{\partial x^{o}} - \alpha^{k}\frac{\partial}{\partial x^{k}} - \frac{imc}{\hbar}\beta\right)\left(\frac{\partial}{\partial x^{o}} + \alpha^{k}\frac{\partial}{\partial x^{k}} + \frac{imc}{\hbar}\beta\right)$$

$$= \frac{\partial^{2}}{\partial x^{o\,2}} - \frac{1}{2}(\alpha^{k}\alpha^{r} + \alpha^{r}\alpha^{k})\frac{\partial}{\partial x^{k}}\frac{\partial}{\partial x^{r}} - \frac{imc}{\hbar}(\alpha^{k}\beta + \beta\alpha^{k})\frac{\partial}{\partial x^{k}} + \frac{m^{2}c^{2}}{\hbar^{2}}\beta^{2}$$

Upon requiring that

$$\{\alpha^{k},\alpha^{r}\} = -2g^{kr}I,\ \{\alpha^{k},\beta\} = 0,\ \beta^{2} = I \qquad (1.2)$$

where $\{A,B\} \equiv AB + BA; (g^{\mu\nu}) \equiv \mathrm{diag}\ (1,-1,-1,-1)$, this differential operator reduces to the Klein-Gordon operator $\left(\frac{\partial^{2}}{\partial x^{o\,2}} - \nabla^{2} + \frac{m^{2}c^{2}}{\hbar^{2}}\right)$. Whence, when conditions (1.2) are imposed, any solution of (1.1) is a solution of the Klein-Gordon equation.

For the purpose of studying its relativistic invariance, eq. (1.1) is more conveniently written as:

$$(i\gamma^{\mu}\partial_{\mu} - \frac{mc}{\hbar})\psi = 0 \qquad (1.3a)$$

where $\gamma^{k} \equiv \beta\alpha^{k}$ and $\gamma^{o} \equiv \beta$ (1.3b)

The relations (1.2) are then equivalent to

$$\{\gamma^{\mu},\gamma^{\nu}\} = 2g^{\mu\nu}I,\ \mu,\nu = 0,1,2,3 \qquad (1.3c)$$

A system of 4 complex square matrices $\{\gamma^{\mu}\}$ satisfying this last equation will be called a system of γ matrices. Since their square is $\pm I$, γ matrices are non-singular. Another immediate consequence of (1.3c) is that the order N of γ matrices has to be even: $\det \gamma^1\gamma^2 = (-1)^N \det \gamma^2\gamma^1 = (-1)^N \det \gamma^1\gamma^2$, from which $(-1)^N = 1$. More will be said about this later.

(2) The relativistic invariance of the Dirac equation

From now on we will take units in which $c=\hbar=1$. Eq. (1.3a) then reads:

$$(-i\gamma^\mu\partial_\mu + m)\psi = 0 \qquad (1.4)$$

The "interaction" with an external electromagnetic field of 4-potential A^μ is achieved through the so-called minimal electromagnetic coupling in which $p_\mu \equiv i\partial_\mu$ is replaced by $p_\mu - eA_\mu = iD_\mu$, that is $D_\mu = \partial_\mu + ieA_\mu$, $e(<o)$ being the charge of the electron. Whence in the presence of an external electromagnetic field, eq. (1.4) becomes:

$$(-i\gamma^\mu D_\mu + m)\psi = 0 \qquad (1.4')$$

$$D_\mu \equiv \partial_\mu + ieA_\mu$$

We want to find a transformation law for ψ such that eq. (1.4') remains invariant under orthochronous Lorentz transformations. If Ω is such a Lorentz transformation it is assumed that the corresponding transformation for ψ is linear:

$$\psi'(x') = S(\Omega)\psi(x) \qquad \det S \neq 0 \tag{1.5}$$

Putting $y^\rho = \Omega^\rho{}_\mu x^\mu$ and rewriting (1.4') in terms of y^μ and ψ' yields:

$$[-i(S\Omega^\rho{}_\mu\gamma^\mu S^{-1})\frac{\partial}{\partial y^\rho} + (S\Omega^\rho{}_\mu\gamma^\mu S^{-1})eA_\rho + m]\psi' = 0$$

This will be identical in form with (1.4') if and only if

$$\hat{\gamma}^\rho \equiv \Omega^\rho{}_\mu\gamma^\mu = S^{-1}\gamma^\rho S \tag{1.6}$$

But it turns out that the $\hat{\gamma}^\rho$'s are also γ matrices:

$$\{\hat{\gamma}^\rho,\hat{\gamma}^\sigma\} = \Omega^\rho{}_\mu\Omega^\sigma{}_\nu\{\gamma^\mu,\gamma^\nu\} = 2\Omega^\rho{}_\mu\Omega^\sigma{}_\nu g^{\mu\nu} I = 2g^{\rho\sigma} I \quad .$$

Therefore the invariance of the theory will be guaranteed if we can show that any two sets of γ matrices are related by a similarity transformation as in eq. (1.6). In the coming sections the existence of such a similarity transformation will be proved in a way which, to our knowledge, is original to a certain extent.

The theory ought to be invariant under space-time translations as well. This is achieved by letting simply $\psi'(x') = \psi(x)$ under the translation $x' = x + a$.

Besides equation (1.4) we will sometimes refer to the so-called adjoint equation. The adjoint $\bar{\psi}$ of ψ is defined by

$$\bar{\psi} \quad \psi^{+}\gamma^{0} \tag{1.7}$$

Provided that the γ^{μ}'s are unitary equations (1.4) and (1.4') are easily seen to be equivalent to the following equations for $\bar{\psi}$:

$$\bar{\psi}(i\gamma^{\mu}\overleftarrow{\partial}_{\mu} + m) = 0 \tag{1.8}$$

$$\bar{\psi}(i\gamma^{\mu}\overleftarrow{D}^{*}_{\mu} + m) = 0, \quad D^{*}_{\mu} \equiv \partial_{\mu} - ieA_{\mu} \tag{1.8'}$$

From this equivalence it is clear that the invariance of equations (1.4), (1.4') implies that of equations (1.8), (1.8').

CHAPTER II

GENERAL PROPERTIES OF γ MATRICES

The aim of this chapter is to investigate the properties of a general system of γ matrices. The first part studies essentially the properties of their products. This is completely standard. The second part studies the degrees and reducibility of all possible representations of the relations $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}I$. Unlike the first part it might be original. This is because we have found a way of treating these questions without using the theory of representations of finite groups or the theory of Clifford algebras. Thus our treatment is self-contained and elementary.

(1) Products of γ^μ's

Let $\{\gamma^\mu\}$ be a system of arbitrary N×N γ matrices. Out of them we construct the following sixteen matrices which will play a great role in our considerations.

Table 1 List of the matrices γ^A

	I				
	γ^0		γ^1	γ^2	γ^3
$\gamma^1\gamma^0$	$\gamma^2\gamma^0$	$\gamma^3\gamma^0$	$\gamma^2\gamma^3$	$\gamma^3\gamma^1$	$\gamma^1\gamma^2$
	$\gamma^1\gamma^2\gamma^3$		$\gamma^0\gamma^2\gamma^3$	$\gamma^0\gamma^3\gamma^1$	$\gamma^0\gamma^1\gamma^2$
				$\gamma^0\gamma^1\gamma^2\gamma^3 \equiv \gamma^5$	

We will denote this set of 16 matrices by Γ and refer to its members by the symbol γ^A, $A=1,\ldots,16$. The inverse of γ^A will be denoted by γ_A. Indices on the γ^μ's will be raised and lowered with respect to $(g_{\mu\nu}) = \mathrm{diag}\,(1,-1,-1,-1)$. Notice that the γ^μ's are unitary if and only if $\gamma^{\mu +} = \gamma_\mu$. We also adopt the following notations:

$$\gamma^{[\mu\nu]} \equiv \begin{cases} \gamma^\mu \gamma^\nu \text{ if } \mu\neq\nu \\ 0 \quad \text{otherwise} \end{cases}$$

$$\gamma^{[\lambda\mu\nu]} \equiv \begin{cases} \gamma^\lambda\gamma^\mu\gamma^\nu \text{ if } \lambda,\mu,\nu \text{ are all different} \\ 0 \quad \text{otherwise} \end{cases}$$

$$\gamma^{[\mu\nu\lambda\rho]} \equiv \varepsilon_{\mu\nu\lambda\rho}\,\gamma^5$$

where $\varepsilon_{\mu\nu\lambda\rho}$ is zero if μ,ν,λ,ρ are not all different and is otherwise equal to the sign of the permutation $\binom{0\ 1\ 2\ 3}{\mu\ \nu\ \lambda\ \rho}$. Whence $\varepsilon_{0123}=1=-\varepsilon^{0123}$

$$\sigma^{\mu\nu} \equiv i\,\gamma^{[\mu\nu]}$$

All these quantities are completely antisymmetric with respect to their indices.

Our set Γ has remarkable properties which we now proceed to derive.

Proposition (2.1): The square of any member of Γ is I or -I.

Proof: This is an obvious consequence of the fact that the square of each γ^{μ} is $\pm I$ and they all anticommute.

In our table we have arranged the γ^{A}'s in such a way that all those with square +I are on the left, the others on the right.

Proposition (2.2): The product of two members of Γ is, up to a sign, again a member of Γ:

$$\gamma^{A}\gamma^{B} = \varepsilon_{AB}\,\gamma^{C(A,B)},\quad \varepsilon_{AB} = \pm 1 \tag{2.1}$$

Proof: This is again an obvious consequence of the anticommutation of the γ^{μ}'s and the fact that their square is $\pm I$.

Proposition (2.2'): The functions ε_{AB} and C(A,B) appearing in eq. (2.1) are the same for all systems of γ matrices.

Proof: This is trivally true by construction.

Equation (2.1) together with the apparently innocent proposition (2.2') will be the key to our proof of the fundamental theorem of γ matrices.

Proposition (2.3): All members of Γ other than I have vanishing trace.

Proof: We first prove it for the γ^{μ}'s. Let ν be given and choose $\mu \neq \nu$. From $\gamma^{\mu}\gamma^{\mu} + \gamma^{\mu}\gamma^{\mu} = \pm 2I$ we deduce

$$-\gamma^{\mu}(\gamma^{\nu}\gamma^{\mu}) + (\gamma^{\nu}\gamma^{\mu})\gamma^{\mu} = \pm 2\,\gamma^{\nu}$$

and whence $\pm 2\mathrm{Tr}\gamma^{\nu} = -\mathrm{Tr}(\gamma^{\mu}(\gamma^{\nu}\gamma^{\mu})) + \mathrm{Tr}[(\gamma^{\nu}\gamma^{\mu})\gamma^{\mu}] = 0$ because in general $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. This at the same time shows that $\mathrm{Tr}\gamma^{5} = 0$ because $\{\gamma^{\mu},\gamma^{5}\} = 0$.

Next if $\mu \neq \nu$ then $\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = 0$ because $\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = \mathrm{Tr}(-\gamma^{\nu}\gamma^{\mu}) = -\mathrm{Tr}(\gamma^{\nu}\gamma^{\mu}) = -\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu})$. Since $\gamma^{[\lambda\mu\nu]}$ is of the form $\pm\,\gamma^{5}\,\gamma^{\rho}$, the same argument shows that $\mathrm{Tr}\,\gamma^{[\lambda\mu\nu]} = 0$.

Proposition (2.4): If one fixes A in equation (2.1) and then lets B go from 1 to 16, C(A,B) goes over all the values in $\{1,\ldots,16\}$.

Proof: Since $\varepsilon_{AB} = \pm 1$, eq. (2.1) may be rewritten as

$$\gamma^{C(A,B)} = \varepsilon_{AB}\,\gamma^{A}\,\gamma^{B}$$

Whence $C(A,B) = C(A,B')$ implies $\varepsilon_{AB}\,\gamma^{A}\,\gamma^{B} = \varepsilon_{AB'}\,\gamma^{A}\,\gamma^{B'}$, from which $\varepsilon_{AB}\,\gamma^{B} = \varepsilon_{AB'}\,\gamma^{B'}$. But clearly this is possible only if $B = B'$; therefore if $B \neq B'$, then $C(A,B) \neq C(A,B')$ and the conclusion follows.

Proposition (2.5): The 16 γ^{A}'s are linearly independent.

Proof: Suppose we have a relation $\sum_{A=1}^{16} \alpha_A\,\gamma^{A} = 0$. Let us pick B in $\{1,\ldots,16\}$ and multiply by γ^{B}:

$$0 = \sum_{A=1}^{16} \alpha_A \gamma^{B}\gamma^{A} = \sum_{A=1}^{16} \alpha_A \varepsilon_{BA}\gamma^{C(B,A)}$$

From proposition (2.4), as A goes from 1 to 16 in this sum, $\gamma^{C(B,A)}$ goes over the whole set Γ. For $A=B$, $\gamma^{C(B,A)} = I$ and for all the other values, according to proposition (2.3), $\gamma^{C(B,A)}$ is a traceless matrix. Whence taking the trace of our equation yields $4\alpha_B \varepsilon_{BB} = 0$, or $\alpha_B = 0$. Since B was arbitrary, it follows that the γ^A's are independent.

One easily obtains the following product rules, some of which will be useful in the sequel.

$$\gamma^\mu \gamma^{[\rho\sigma]} = \varepsilon^{\mu\rho\sigma\theta} \gamma_\theta \gamma^5 + g^{\mu\rho} \gamma^\sigma - g^{\mu\sigma} \gamma^\rho \tag{2.2}$$

$$\gamma^\mu \gamma^{[\rho\sigma\theta]} = -\varepsilon^{\mu\rho\sigma\theta} \gamma^5 + g^{\mu\rho} \gamma^{[\sigma\theta]} - g^{\mu\sigma} \gamma^{[\rho\theta]} + g^{\mu\theta} \gamma^{[\rho\sigma]} \tag{2.3}$$

$$\gamma^\mu \gamma^5 = \frac{1}{6} \varepsilon^\mu{}_{\alpha\beta\delta} \gamma^{[\alpha\beta\delta]}$$

$$\gamma^{[\alpha\beta]} \gamma^{[\mu\nu]} = -\varepsilon^{\alpha\beta\mu\nu} \gamma^5 + 2g^{\mu[\alpha} \gamma^{\beta]} \gamma^\nu - 2g^{\nu[\alpha} \gamma^{\beta]} \gamma^\mu \tag{2.4}$$

$$\gamma^{[\rho\sigma]} \gamma_5 = \frac{1}{2} \varepsilon^{\rho\sigma}{}_{\alpha\beta} \gamma^{[\alpha\beta]} \equiv \tilde{\gamma}^{[\rho\sigma]} \tag{2.5}$$

$$\gamma^\mu \gamma^5 \gamma^{[\rho\sigma]} = -\varepsilon^{\mu\rho\sigma\theta} \gamma_\theta + g^{\mu\rho} \gamma_5 \gamma^\sigma - g^{\mu\sigma} \gamma_5 \gamma^\rho \tag{2.6}$$

From the relation (1.3c) it is possible to derive the commutators and anticommutators of all pairs of elements of Γ. We list here the results.

Table 2 Commutators and anticommutators of the γ^A's

$[\gamma^\mu,\gamma^\nu] = 2\,\gamma^{[\mu\nu]}$	$\{\gamma^\mu,\gamma^\nu\} = 2\,g^{\mu\nu}I$
$[\gamma^\mu,\gamma^5] = 2\gamma^\mu\gamma^5$	$\{\gamma^\mu,\gamma^5\} = 0$
$[\gamma^\lambda,\gamma^{[\mu\nu]}] = 2(g^{\lambda\mu}\gamma^\nu - g^{\lambda\nu}\gamma^\mu)$	$\{\gamma^\lambda,\gamma^{[\mu\nu]}\} = 2\,\gamma^{[\lambda\mu\nu]}$
$[\gamma^\mu,\gamma^{[\rho\sigma\tau]}] = 2\,\gamma^{[\mu\rho\sigma\tau]}$	$\{\gamma^\mu,\gamma^{[\rho\sigma\tau]}\} = 2(g^{\mu\rho}\gamma^{[\sigma\tau]} + g^{\tau\mu}\gamma^{[\rho\sigma]} + g^{\sigma\tau}\gamma^{[\mu\rho]})$
$[\gamma^5,\gamma^{[\mu\nu]}] = 0$	$\{\gamma^5,\gamma^{[\mu\nu]}\} = -\,\varepsilon^{\mu\nu\rho\sigma}\gamma_{[\rho\sigma]}$
$[\gamma^5,\gamma^{[\lambda\rho\sigma]}] = 2\varepsilon^{\lambda\rho\sigma\theta}\gamma_\theta$	$\{\gamma^5,\gamma^{[\lambda\rho\sigma]}\} = 0$
$[\gamma^{[\lambda\rho]},\gamma^{[\mu\nu]}] = 2(g^{\lambda\mu}\gamma^{[\rho\nu]} + g^{\nu\lambda}\gamma^{[\mu\rho]} + g^{\rho\nu}\gamma^{[\lambda\mu]} + g^{\mu\rho}\gamma^{[\nu\lambda]})$	$\{\gamma^{[\lambda\rho]},\gamma^{[\mu\nu]}\} = 2\gamma^{[\lambda\rho\mu\nu]} + 2(g^{\mu\rho}g^{\nu\lambda} - g^{\mu\lambda}g^{\nu\rho})I$
$[\gamma^{[\mu\nu]},\gamma^{[\lambda\rho\sigma]}] = 2\gamma^5(\varepsilon^{\lambda\rho\sigma\mu}\gamma^\nu - \varepsilon^{\lambda\rho\sigma\nu}\gamma^\mu)$	$\{\gamma^{[\mu\nu]},\gamma^{[\lambda\rho\sigma]}\} = -2\varepsilon^{\lambda\rho\sigma}{}_\theta\,\varepsilon^{\theta\mu\nu\beta}\gamma_\beta$ or equivalently $\{\gamma^{[\mu\nu]},\gamma^5\gamma^\lambda\} = 2\varepsilon^{\mu\nu\lambda\rho}\gamma_\rho$
$[\gamma^{[\lambda\rho\sigma]},\gamma^{[\alpha\beta\delta]}] = 2\varepsilon^{\lambda\rho\sigma\theta}\varepsilon^{\alpha\beta\delta\mu}\gamma_{[\theta\mu]}$ or equivalently $[\gamma^5\gamma^\mu,\gamma^5\gamma^\nu] = 2\gamma^{[\mu\nu]}$	$\{\gamma^{[\lambda\rho\sigma]},\gamma^{[\alpha\beta\delta]}\} = -2\varepsilon^{\lambda\rho\sigma}{}_\theta\,\varepsilon^{\theta\alpha\beta\delta}I$ or equivalently $\{\gamma^5\gamma^\mu,\gamma^5\gamma^\nu\} = 2g^{\mu\nu}I$

(2) The representations of the relations $\{\gamma^\mu,\gamma^\nu\}=2g^{\mu\nu}I$

We now come to the part of this chapter dealing with the order N of γ matrices and the irreducibility of such systems.

Theorem (2.6): The order N of γ matrices cannot be smaller than 4.

Proof: The dimension of the complex vector space of N×N complex matrices is N^2. Since by proposition (2.5) Γ is a set of 16 independent matrices, we see that N^2 has to be ≥ 16.

Theorem (2.7): The matrices Γ^μ defined by

$$\Gamma^o = \begin{pmatrix} \sigma_o & o \\ o & -\sigma_o \end{pmatrix} \qquad \Gamma^i = \begin{pmatrix} o & \sigma_i \\ -\sigma_i & o \end{pmatrix} \qquad i=1,2,3. \tag{2.7}$$

where $$\sigma_o=\begin{pmatrix} 1 & o \\ o & 1 \end{pmatrix}\ \sigma_1=\begin{pmatrix} o & 1 \\ 1 & o \end{pmatrix}\ \sigma_2=\begin{pmatrix} o & -i \\ i & o \end{pmatrix}\ \sigma_3=\begin{pmatrix} 1 & o \\ o & -1 \end{pmatrix} \tag{2.8}$$

form a system of unitary γ matrices.

Proof: From the relations $\sigma_i\sigma_j = i\epsilon_{ijk}\sigma_k + \delta_{ij}\sigma_o$ we obtain

$$\{\Gamma^i,\Gamma^j\} = -\begin{pmatrix} \{\sigma_i,\sigma_j\} & o \\ o & \{\sigma_i,\sigma_j\} \end{pmatrix} = 2g^{ij}I$$

$$\{\Gamma^i,\Gamma^o\} = \begin{pmatrix} o & \sigma_i \\ \sigma_i & o \end{pmatrix} - \begin{pmatrix} o & \sigma_i \\ \sigma_i & o \end{pmatrix} = 0$$

$$(\Gamma^o)^2 = I$$

The fact that each Γ^μ is unitary follows from

$$\Gamma^{o+} = \Gamma^o \qquad \Gamma^{i+} = -\Gamma^i \tag{2.9}$$

This system $\{\Gamma^{\mu}\}$ is referred to as the Dirac representation.

Our next theorem, in group theory language, would be said to reflect the irreducibility of any set of 4x4 γ matrices.

Theorem (2.8): Let $\{\gamma^{\mu}\}$ be a system of 4x4 γ matrices. Any matrix commuting with the four of them has to be a multiple of the identity.

Proof: Suppose M is such a matrix. It follows that it commutes with every γ^{A} in Γ. But proposition (2.5) says that, in the case under consideration, Γ is a basis. Therefore M commutes with everything. But it is easy to prove that any linear operator on a vector space commuting with all the others has to be a multiple of the identity. The conclusion applies to M.

Theorem (2.8) will be used repeatedly in the sequel. Together with the fundamental theorem of γ matrices, which is the subject of the next chapter, it lies at the basis of most of the constructions which we will make.

So far all we know about the possible order of γ matrices is that N is even and ≥ 4 (see the discussion following eq. (1.3c) and theorem (2.6)). It is in fact known that N has to be a multiple of 4. This can be proved by using the theory of representations of finite groups (the starting point of such an approach is to observe that the set $\bar{\Gamma} = \{\pm\gamma^{A};\gamma^{A}\epsilon\Gamma\}$ is a group) (Jansen and Boon [1967]). The proof we give here is (to our knowledge) original and owes absolutely nothing to group theory. We start with a little lemma which will also be used in the proof of another result.

Lemma: Let $\{\gamma^\mu\}$ be an arbitrary system of γ matrices. Then γ^o and $i\gamma^1\gamma^2$ can always be diagonalised simultaneously.

Proof: Let A be a matrix such that $A^2 = I$. Then A may be written as $A = I - 2P$ where P is a projector, namely $P = \frac{1}{2}(I-A)$, $(P^2=P)$. But a projector can always be diagonalised; whence so can A. Our two matrices commute and have square I. So the conclusion follows.

Theorem (2.9): The order N of γ matrices has to be a multiple of 4.

Proof: Let $\{\gamma^\mu\}$ be a system of γ matrices. By the previous lemma we may assume, by performing a similarity transformation on $\{\gamma^\mu\}$ if necessary, that γ^o and $i\gamma^1\gamma^2$ are diagonal. Since their square is I, their diagonal entries have to be ± 1. Moreover since by proposition (2.3) they are both traceless, the number of +1's has to be equal to the number of -1's. Again by performing a similarity transformation if necessary we may assume that $\gamma^o = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$. N is even so there is an integer n such that N=2n. Let us write $i\gamma^1\gamma^2 = \text{diag}(a_1,\ldots,a_n,b_1,\ldots,b_n)$. The matrix $\gamma^o i\gamma^1\gamma^2$ is also traceless and its trace is $a_1 + \ldots + a_n - (b_1 + \ldots + b_n)$. Now we have to distribute n +1's and n -1's among $a_1,\ldots, a_n,b_1,\ldots,b_n$.

Suppose that we put r - 1's $(0 \leq r \leq n)$ in $a_1, \ldots, a_n$. There are n-r left over to be distributed among $b_1, \ldots, b_n$ and therefore we have:

$$a_1 + \ldots + a_n = -r + (n-r) = n-2r$$

$$b_1 + \ldots + b_n = -(n-r) + r = -n + 2r$$

Whence $Tr\gamma^o i\gamma^1\gamma^2 = n-2r-(-n+2r) = 2n - 4r$. In order that this vanishes we must have $r = \frac{n}{2}$. Since r is an integer it follows that n is even; whence N = 2n is a multiple of 4.

<u>Corollary</u>: γ matrices of arbitrary order have determinant 1.

Proof: γ^o is similar to $\begin{bmatrix} I & \\ & -I \end{bmatrix}$ and N is a multiple of 4. Whence $\det \gamma^o = (-1)^{N/2} = 1$. A similar argument applies to the γ^k's.

The final result of this section, which we are about to present, is not the least in importance since it establishes in some sense the uniqueness of the relativistic equation for the electron wave function. As one might guess we are going to be concerned with the irreducibility of the representations of eq. (1.3c). It is well known that the only irreducible representations are provided by matrices of order 4. This result makes one think of group theory and of course, like theorem (2.9), it can be derived via the theory of representations of finite groups (Jansen and Boon [1967]). But as we did for theorem (2.9) we have found a simple way of proving it keeping away from group theory. Our proof is likely to have been thought of before but we have not met it anywhere.

Theorem (2.10): Any (unitary or not) representation of eq. (1.3c) can be reduced to one of degree 4. Whence the irreducible ones are those of degree 4.

Proof: Let $\{\gamma^\mu\}$ be such a representation. By the previous lemma we know that it is equivalent to one in which γ^o and $i\gamma^1\gamma^2$ are both diagonal. So we may assume without loss of generality that they are diagonal. In order that a subspace be invariant under $\{\gamma^\mu\}$ it is clear by proposition (2.2) that it is necessary and sufficient that it be invariant under the set Γ of table 1. Now pick a non vanishing vector u and define $u_A \equiv \gamma^A u$. Again by proposition (2.2) it is clear that the subspace V spanned by the u_A's is Γ-invariant. The special trick of the proof lies in an appropriate choice of u. From the proof of theorem (2.9) one easily checks that one can pick a non vanishing u such that $\gamma^o u = i\gamma^1\gamma^2 u = u$. We claim that the subspace V spanned by the corresponding u_A's is 4-dimensional. Indeed it is generated by $u, u_1 \equiv \gamma^1 u$, $u_3 \equiv \gamma^3 u$ and $u_{31} \equiv \gamma^3\gamma^1 u$. This is seen by letting Γ act on u: $Iu = u$, $\gamma^o u = u$, $\gamma^1 u = u_1$,

$\gamma^2 u = -\gamma^1\gamma^1\gamma^2 u = iu_1$, $\gamma^3 u = u_3$, $\gamma^1\gamma^2 u = -iu$,
$\gamma^3\gamma^1 u = u_{31}$, $\gamma^2\gamma^3 u = \gamma^3\gamma^1\gamma^1\gamma^2 u = -iu_{31}$, $\gamma^o\gamma^1 u = -u_1$,
$\gamma^o\gamma^2 u = -iu_1$, $\gamma^o\gamma^3 u = -u_3$, $\gamma^o\gamma^1\gamma^2 u = \gamma^1\gamma^2 u = -iu$,
$\gamma^o\gamma^3\gamma^1 u = u_{31}$, $\gamma^o\gamma^2\gamma^3 u = -iu_{31}$, $\gamma^1\gamma^2\gamma^3 u = \gamma^3\gamma^1\gamma^2 u = -iu_3$,
$\gamma^5 u = iu_3$.

By theorem (2.6) the vectors u, u_1, u_3, u_{31} are necessarily linearly independent. This completes the proof.

From now on, when we talk about γ matrices, unless otherwise stated, it will always be understood that these are 4×4 matrices. In physical applications the γ^μ's are always unitary. All the γ^A's are then unitary as well. But as Pauli did in his paper (Pauli [1936]) we will invoke this assumption only when needed: as has already been seen many results follow without it.

CHAPTER III

THE FUNDAMENTAL THEOREM OF γ MATRICES AND CONSEQUENCES

(1) The fundamental theorem

The so-called fundamental theorem of γ matrices is the basis upon which lies the relativistic invariance of the Dirac theory of electrons and positrons. This section is devoted to its proof and to the exposition of some of its consequences. The proof that we give is, to our knowledge, original and is, as will be seen, quite simple. It is based on a seldom used result of linear algebra. To preserve the continuity of exposition the proof of this result will be deferred to the end of this chapter. As will be seen later, we state it here in a restricted context which will be sufficient for our purpose.

Let us denote by $A_n(C)$ the algebra of $n \times n$ complex matrices. By an automorphism of $A_n(C)$ we mean a bijective linear map $h: A_n(C) \rightarrow A_n(C)$ which also preserves multiplication, that is for any μ, ν in C and M,N in $A_n(C)$ we have: $h(\mu M + \nu N) = \mu h(M) + \nu h(N)$, $h(MN) = h(M)h(N)$, $h(M)=0 \Rightarrow M=0$.

It is clear that, given a non-singular S in $A_n(C)$, the map $M \rightarrow SMS^{-1}$ is an automorphism of $A_n(C)$. The result of linear algebra we were referring to is the converse of this.

<u>Theorem</u> (3.1): If h is an automorphism of $A_n(C)$, then there exists a non-singular matrix S in $A_n(C)$ such that $h(M) = SMS^{-1}$ for all M.

Having stated this we may now give our proof of the fundamental theorem. As was mentioned in chapter II (section 1) the key of this proof is eq. (2.1) together with the trivial proposition (2.2').

<u>Theorem</u> (3.2): (The fundamental theorem of γ matrices): Let $\{\gamma^\mu\}$, $\{\hat{\gamma}^\mu\}$ be two systems of 4×4 γ matrices. Then there exists a non-singular matrix S such that $\hat{\gamma}^\mu = S^{-1}\gamma^\mu S$.

Proof: Let Γ and $\hat{\Gamma}$ be the two sets constructed from $\{\gamma^\mu\}$ and $\{\hat{\gamma}^\mu\}$ according to table 1. From proposition (2.5) we know that both Γ and $\hat{\Gamma}$ are basis of $A_4(C)$. Therefore we may define a linear map $h: A_4(C) \to A_4(C)$ by $h(\gamma^A) = \hat{\gamma}^A$ and this map is bijective. Moreover it preserves products. Indeed let $M = \sum_A \alpha_A \gamma^A$, $N = \sum_B \beta_B \gamma^B$. Then we have:

$$h(MN) = h(\sum_{AB} \alpha_A \beta_B \gamma^A \gamma^B) = \sum_{AB} \alpha_A \beta_B h(\epsilon_{AB} \gamma^{C(A,B)}) =$$

$$\sum_{AB} \alpha_A \beta_B \epsilon_{AB} \hat{\gamma}^{C(A,B)} = \sum_{AB} \alpha_A \beta_B \hat{\gamma}^A \hat{\gamma}^B = h(M)h(N)$$

where we have used propositions (2.2) and (2.2'). Hence h fullfills all the conditions of theorem (3.1) and it follows that there exits a non-singular S such that $\hat{\gamma}^A = h(\gamma^A) = S^{-1}\gamma^A S$ (q.e.d.).

Proposition (3.1): The matrix S of the fundamental theorem is unique up to a multiplicative factor.

Proof: Suppose that S and T satisfy

$$\hat{\gamma}^{\mu} = S^{-1} \gamma^{\mu} S = T^{-1} \gamma^{\mu} T$$

It follows that the commutator $[\gamma^{\mu}, ST^{-1}]$ vanishes. From theorem (2.8) we may therefore conclude that $ST^{-1} = cI$ or $T = \frac{1}{c} S$ (q.e.d.).

(2) Consequences of the fundamental theorem

The matrix S of the fundamental theorem is of course closely related to the systems of γ matrices from which it arises. This is illustrated in the following little result which we present here as a curiosity, since we shall not use it later.

Proposition (3.2): Let $\{\gamma^{\mu}\}$ and $\{\hat{\gamma}^{\mu}\}$ be two systems of γ matrices, with associated sets Γ and $\hat{\Gamma}$. The invertible matrix S such that $\hat{\gamma}^{\mu} = S^{-1} \gamma^{\mu} S$ has the same coordinates in both basis Γ and $\hat{\Gamma}$.

Proof: Let $S = \sum_A \alpha_A \gamma^A = \sum_A \beta_A \hat{\gamma}^A$. From $S\hat{\gamma}^A = \gamma^A S$ we get

$$\gamma^B S = \sum_A \alpha_A \gamma^B \gamma^A = S\hat{\gamma}^B = \sum_A \beta_A \hat{\gamma}^A \hat{\gamma}^B$$

Taking the trace on each side of $\sum_A \alpha_A \gamma^B \gamma^A = \sum_A \beta_A \hat{\gamma}^A \hat{\gamma}^B$ and using propositions (2.3) and (2.4) gives $\alpha_B = \beta_B$ (q.e.d.).

Another connection between $\{\gamma^\mu\}$ and S which will not be a mere curiosity for us is the following.

Proposition (3.3): Let $\{\gamma^\mu\}$, $\{\hat{\gamma}^\mu\}$ be two systems of γ matrices such that all the γ^μ's and $\hat{\gamma}^\mu$'s are unitary. Then the matrix S connecting the two systems can be chosen to be unitary. Such a choice is unique up to a phase factor.

Proof: We have from the unitarity of γ^μ, $\hat{\gamma}^\mu$ and $\hat{\gamma}^\mu = S^{-1}\gamma^\mu S$:

$$S^{-1}\gamma^o S = \hat{\gamma}^o = \hat{\gamma}^{o+} = S^+ \gamma^{o+} S^{-1+} = S^+ \gamma^o S^{-1+}$$

$$S^{-1}\gamma^i S = \hat{\gamma}^i = -\hat{\gamma}^{i+} = -S^+ \gamma^{i+} S^{-1+} = S^+ \gamma^i S^{-1+}$$

that is $S^{-1}\gamma^\mu S = S^+\gamma^\mu S^{-1+}$. From this we infer $[\gamma^\mu, SS^+] = 0$, and by theorem (2.8) conclude that $SS^+ = cI$. Now SS^+ is obviously self-adjoint and positive. Whence c is real and > 0. Taking $S' = \frac{1}{\sqrt{c}} S$ yields the required unitary matrix. The fact that the choice of a unitary S is unique up to a phase factor is obvious in view of proposition (3.1) and the unitarity condition.

The fundamental theorem allows us to draw other interesting general conclusions about γ matrices. For example, it says that γ^o has to be similar to the matrix Γ^o of eq. (2.7). Whence the characteristic and minimal polynomials of γ^o have to be $(t+1)^2(t-1)^2$ and $(t+1)(t-1)$ respectively, and its determinant is 1. To make similar remarks about γ^1 we observe that the set $\{\hat{\gamma}^\mu\}$ defined by $\hat{\gamma}^o = i\gamma^1$, $\hat{\gamma}^1 = i\gamma^o$, $\hat{\gamma}^2 = \gamma^2$, $\hat{\gamma}^3 = \gamma^3$

is also a system of γ matrices. Whence $i\gamma^1$ is similar to Γ^0, or γ^1 is similar to $-i\Gamma^0$. This shows that the characteristic and minimal polynomials of γ^1 (or any γ^i) are $(t+i)^2(t-i)^2$ and $(t+i)(t-i)$ respectively and its determinant is 1. These observations lead themselves immediately to the following algebraic characterization of matrices which can be γ matrices.

Proposition (3.4): Let M be a complex 4×4 matrix. In order that M be the γ^0 of a $\{\gamma^\mu\}$ system it is necessary and sufficient that its characteristic and minimal polynomials be $(t+1)^2(t-1)^2$ and $(t+1)(t-1)$ respectively. In order that M be the γ^1 (or γ^2 or γ^3) of some $\{\gamma^\mu\}$ system it is necessary and sufficient that the characteristic and minimal polynomials of M be $(t+i)^2(t-i)^2$ and $(t+i)(t-i)$ respectively.

Before we close this chapter with the proof of theorem (3.1) we give a last application of the fundamental theorem. This theorem enabled us, in a rather curious way, to answer a question which arises naturally when dealing with γ matrices. When we look closely at the set $\{\Gamma^\mu\}$ of eq. (2.7) we notice that all these matrices, except Γ^2, are real. Whence it seems natural to ask whether one could find a system of γ matrices where all the matrices would be real. The answer is negative. As we will show below, the existence of such a system would imply the existence of a complex number c whose squared magnitude c^*c would be -1; this is of course absurd.

Proposition (3.5): There exists no system of γ matrices such that each γ^{μ} is real.

Proof: Consider again the particular system $\{\Gamma^{\mu}\}$ exhibited in eq. (2.7). If $\{\gamma^{\mu}\}$ is any system of γ matrices the fundamental theorem says that there exists a non-singular S such that $S^{-1}\gamma^{\mu}S = \Gamma^{\mu}$. (3.1)

Let $M = i\gamma^{[031]}$. Using $\Gamma^{2*} = -\Gamma^{2}$ and $(\Gamma^{\nu})^{*} = \Gamma^{\nu}$ for $\nu \neq 2$, one easily checks that the following equation holds true:

$$S^{-1*}(M\gamma^{\nu}M)^{*}S^{*} = \Gamma^{\nu}, \quad \nu=0,1,2,3$$

(Here * means complex conjugate.)

From this equation and equation (3.1) we obtain

$$S^{-1*}(M\gamma^{\nu}M)^{*}S^{*} = S^{-1}\gamma^{\nu}S$$

If we now suppose that all the γ^{ν}'s are real, this last equation may be rewritten as:

$$\gamma^{\nu} = \gamma^{\nu*} = M S^{*} S^{-1}\gamma^{\nu} S S^{*-1} M$$

from which we deduce that $[\gamma^{\nu}, MS^{*}S^{-1}] = 0$. This implies that there exists a number c such that $MS^{*}S^{-1} = cI$. This gives the equation $S = \frac{1}{c}M S^{*}$ and its complex conjugate $S^{*} = \frac{1}{c^{*}}(-M)S$. Substituting the second in the first yields $cc^{*} = -1$, a contradiction. Therefore the four γ^{μ}'s cannot all be real.

Though there is no system of real γ matrices, there are systems in which the γ^{μ}'s are purely imaginary. Their special interest is that they make the free Dirac equation (1.4) real; hence ψ is a solution if and only if its real and imaginary parts are separate solutions of the equation.

It is easy to give a fairly explicit description of all these systems. Let $\{\gamma^{\mu}\}$ be an arbitrary system of γ matrices. By the fundamental theorem there exists an invertible matrix T such that $\gamma^{\mu} = T^{-1}\Gamma^{\mu}T$, where $\{\Gamma^{\mu}\}$ is the particular system of equation (2.7). Due to the fact that Γ^2 is purely imaginary while Γ^o, Γ^1 and Γ^3 are real, the γ^{μ}'s will be purely imaginary if and only if $T^{-1*}\Gamma^2T^* = T^{-1}\Gamma^2T$ and $T^{-1*}\Gamma^{\nu}T^* = -T^{-1}\Gamma^{\nu}T$, for $\nu = 0, 1, 3$. This is equivalent to saying that T^*T^{-1} commutes with Γ^2 and anti-commutes with Γ^o, Γ^1 and Γ^3. This implies $T^*T^{-1} = \alpha\Gamma^2$ or $T^* = \alpha\Gamma^2T$. Clearly the number α has to be a phase factor. Upon writing $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, the general solution of $T^* = \alpha\Gamma^2T$ is easily seen to be

$$T = \begin{bmatrix} A & B \\ \alpha^*\sigma_2A^* & \alpha^*\sigma_2B^* \end{bmatrix}, \quad \text{A and B arbitrary, } |\alpha| = 1$$

This can be alternatively written as

$$T = \begin{bmatrix} u \\ v \\ -i\alpha^*v^* \\ i\alpha^*u \end{bmatrix} \quad \text{u and v arbitrary, } |\alpha| = 1$$

where u and v are row vectors. As long as the choice of A and B leads to an invertible matrix T, the matrices $T^{-1}\Gamma^{\mu}T$ will form a system of purely imaginary γ matrices. Conversely all the purely imaginary systems can be obtained that way.

The particular choice $A = \sigma_o + \sigma_2$, $B = \sigma_o - \sigma_2$ and $\alpha = -1$ gives the system

$$\gamma^o = \begin{pmatrix} \sigma_2 & \\ & -\sigma_2 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} -i\sigma_3 & \\ & -i\sigma_3 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} & -\sigma_2 \\ \sigma_2 & \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} i\sigma_1 & \\ & i\sigma_1 \end{pmatrix}$$

It is referred to as the Majorana representation.

(3) The proof of theorem (3.1)

We now come to the proof of theorem (3.1). As mentioned at the beginning of this chapter we stated it in a restricted context. It turns out that it is also true for $A_n(K)$, the algebra of n×n matrices over an arbitrary field K, for example the field of real numbers. We stated it for K=C because this was all we needed. Quite amusingly it wasn't our knowledge of this result which inspired our proof of the fundamental theorem but rather the study of the usual proofs of the fundamental theorem lead us to guess that such a result might be true. We were able to trace it in only one book (Herstein [1964]) where it is stated as a problem (problem 27, page 279). Therefore the proof we give here is ours. It is possible that a shorter proof could be given.

Theorem (4.1'): If h is an automorphism of $A_n(K)$, that is a bijective linear map preserving products of $A_n(K)$ onto itself, then there exists a non-singular matrix S in K such that $h(M) = SMS^{-1}$ for all M.

Proof: Throughout this proof we don't use the summation convention. We will denote by $|1\rangle,\ldots,|n\rangle$ the canonical basis of K^n, that is:

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad |2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{etc.,}$$

and by $\langle 1|,\ldots,\langle n|$ the dual basis: $\langle i|j\rangle = \delta_{ij}$. We denote by M_{ij} the matrix having null entries except at the intersection of the i^{th} row and the j^{th} column where it has a 1: $\langle k|M_{ij}|\ell\rangle = \delta_{ki}\,\delta_{j\ell}$. These matrices multiply according to $M_{ij}M_{k\ell} = \delta_{jk}M_{i\ell}$, and they form a basis of $A_n(K)$. Now since h is an automorphism the matrices $H_{ij} \equiv h(M_{ij})$ also form a basis and multiply according to

$$H_{ij}H_{k\ell} = \delta_{jk}H_{i\ell} \tag{3.2}$$

As often is the case it makes the argument simpler to think of H_{ij} as a linear operator instead of a matrix. Whence we consider H_{ij} as the linear operator on K^n whose matrix with respect to the canonical basis is H_{ij}. From eq. (3.2) it follows that $H_{11},\ldots,H_{nn}$ are n projectors such that $H_{ii}H_{jj} = 0$ if $i\neq j$. Moreover neither of them is 0 because h is an automorphism. It follows immediately from this that there exists a basis $|1\rangle',\ldots,|n\rangle'$ of K^n such that

$$H_{ii}|j\rangle' = \delta_{ij}\;|i\rangle' \tag{3.3}$$

From eq. (3.2) we also deduce that $H_{ij} = H_{ii}H_{ij}H_{jj}$. This together with (3.3) implies that

$$H_{ij}|k\rangle' = h_{ij}\,\delta_{jk}\;|i\rangle' \;,\quad h_{ij}\in K \;,\quad h_{ij}\neq 0 \tag{3.4}$$

Since H_{ii} is a projector, $h_{ii} = 1$. Moreover eqs. (3.2) and (3.4) clearly imply that $h_{ik}h_{kj} = h_{ij}$. Now let us define a new basis of K^n by

$|i>" = \frac{1}{h_{1i}} \; |i>'$. Then we have:

$$H_{ij}|j>" = \frac{h_{ij}}{h_{1j}} \; |i>' = \frac{h_{1i}h_{ij}}{h_{1j}} \; |i>" = |i>"$$

$$H_{ij} \; |k>" = 0 \quad \text{if} \quad k \neq j$$

Therefore the matrix of H_{ij} with respect to the basis $|1>",\ldots,|n>"$ is M_{ij}. So if S is the transition matrix from the basis $(|i>)$ to the basis $(|i>")$ $(|i>" = \sum_j S_{ji} \; |j>)$ one has $H_{ij} = S \; M_{ij} \; S^{-1}$, that is, $h(M_{ij}) = S \; M_{ij} S^{-1}$. Since h is linear and M_{ij} is a basis it follows that $h(M) = SMS^{-1}$ for an arbitrary M. (q.e.d.).

CHAPTER IV

AN ALGEBRAIC METHOD FOR FINDING THE MATRIX S OF THE FUNDAMENTAL THEOREM

(1) Finding S

The fundamental theorem of γ matrices asserts the existence of an invertible matrix S connecting two given sets of 4×4 γ matrices. But it does not tell us what S is. In later chapters we shall give the solution of this problem when the two sets $\{\gamma^{\mu}\}$, $\{\hat{\gamma}^{\mu}\}$ are related through a Lorentz transformation $\hat{\gamma}^{\mu} = \Omega^{\mu}{}_{\nu}\gamma^{\nu}$ by using Lie group techniques. This is what is usually done in the physical literature. In the present chapter we adopt a purely algebraic point of view and look at the general case: the γ^{μ}'s are not assumed [illegible]tary and the two sets $\{\hat{\gamma}^{\mu}\}$ and $\{\gamma^{\mu}\}$ are not [illegible] related through a Lorentz transformation. [illegible] people who investigate what happens if, instead [illegible] the field of real numbers to construct the M[illegible]ki space, one starts with a field having only a f[illegible] (enormous) number of elements. The aim of this is to [illegible]roduce a fundamental length in physics. In such a f[illegible]work it is clear that one couldn't use infinitesimal tr[illegible]sformations to obtain S in $\Omega^{\mu}{}_{\nu}\gamma^{\nu} = S^{-1}\gamma^{\mu}S$. The algebraic method which we set up in this section and the following one would provide a substitute.

The equations $\hat{\gamma}^{\mu} S = S\,\gamma^{\mu}$, where S is the unknown, give rise to a system of 64 linear equations with 16 unknowns. It is certainly not convenient to attack this directly. We have found a roundabout way which simplifies the task considerably. Consider again the particular set $\{\Gamma^{\mu}\}$ of eq. (2.7). By the fundamental theorem there exist S and T such that $\hat{\gamma}^{\mu} = T^{-1}\,\Gamma^{\mu}\,T$ and $\Gamma^{\mu} = S^{-1}\,\gamma^{\mu}\,S$. Clearly $\hat{\gamma}^{\mu} = (ST)^{-1}\,\gamma^{\mu}\,(ST)$. The point is now that, thanks to the simplicity of $\{\Gamma^{\mu}\}$, finding S and T is very simple as shown by the following result.

Proposition (4.1): Let $\{\gamma^{\mu}\}$ be a system of 4×4 γ matrices. Then the systems

$$\text{(a)}\quad \begin{matrix} (\gamma^{1} + i\gamma^{2})u = 0 \\ (I + \gamma^{o})u = 0 \end{matrix} \qquad\qquad \text{(b)}\quad \begin{matrix} w\,(\gamma^{1} - i\gamma^{2}) = 0 \\ w\,(I + \gamma^{o}) = 0 \end{matrix} \tag{4.1}$$

where u and w are a column and a row vector respectively, have non vanishing solutions which are determined up to multiplicative factors. The matrices S and T defined by

$$S = (\gamma^{3}u \quad \gamma^{1}u \quad u \quad \gamma^{3}\gamma^{1}u) \tag{4.2a}$$

$$T = \begin{bmatrix} -w\,\gamma^{3} \\ -w\,\gamma^{1} \\ w \\ w\,\gamma^{1}\gamma^{3} \end{bmatrix} \tag{4.2b}$$

satisfy $$\Gamma^{\mu} = S^{-1}\gamma^{\mu} S \qquad \gamma^{\mu} = T^{-1}\Gamma^{\mu} T . \tag{4.3}$$

Proof: The fundamental theorem guarantees the existence of the matrices S and T. So all we have to do is to show that they necessarily have the form stated above. We begin with S. We first decompose all our matrices into blocks of order 2:

$$\gamma^{\mu} \begin{bmatrix} A^{\mu} & B^{\mu} \\ C^{\mu} & D^{\mu} \end{bmatrix} \qquad S = \begin{bmatrix} X & Y \\ Z & U \end{bmatrix}.$$

The unknowns are now the four 2×2 matrices X,Y,Z,U. The equations which they must satisfy are derived from (4.3) and are as follows:

$$\begin{aligned} A^{i}X + B^{i}Z &= -Y\sigma_{i} & A^{o}X + B^{o}Z &= X \\ A^{i}Y + B^{i}U &= X\sigma_{i} & A^{o}Y + B^{o}U &= -Y \\ C^{i}X + D^{i}Z &= -U\sigma_{i} & C^{o}X + D^{o}Z &= Z \\ C^{i}Y + D^{i}U &= Z\sigma_{i} & C^{o}Y + D^{o}U &= -U \end{aligned} \tag{4.4}$$

The second and fourth equations on the left-hand side give

$$X = (A^{i}Y + B^{i}U)\sigma_{i} \quad \text{(no sum)} \tag{4.5}$$

$$Z = (C^{i}Y + D^{i}U)\sigma_{i} \quad \text{(no sum)} \tag{4.6}$$

This system of six equations can be rewritten as follows

$$X = (A^1Y + B^1U)\sigma_1 \qquad (4.7a)$$

$$Z = (C^1Y + D^1U)\sigma_1 \qquad (4.7b)$$

$$A^2Y + B^2U = (A^1Y + B^1U)\sigma_1\sigma_2 = (A^1Y + B^1U)i\sigma_3 \qquad (4.7c)$$

$$A^3Y + B^3U = (A^1Y + B^1U)\sigma_1\sigma_3 = -(A^1Y + B^1U)i\sigma_2 \qquad (4.7d)$$

$$C^2Y + D^2U = (C^1Y + D^1U)\sigma_1\sigma_2 = (C^1Y + D^1U)i\sigma_3 \qquad (4.7e)$$

$$C^3Y + D^3U = (C^1Y + D^1U)\sigma_1\sigma_3 = -(C^1Y + D^1U)i\sigma_2 \qquad (4.7f)$$

The first two give X and Z in terms of Y and U and the last four guarantee that the right-hand sides in (4.5) and (4.6) are independent of i. The problem is now reduced to finding Y and U.

From the identity

$$\gamma^\mu\gamma^\nu = \begin{bmatrix} A^\mu A^\nu + B^\mu C^\nu & A^\mu B^\nu + B^\mu D^\nu \\ C^\mu A^\nu + D^\mu C^\nu & C^\mu B^\nu + D^\mu D^\nu \end{bmatrix} \qquad (4.8)$$

we see that multiplying (4.7c) by A^2, (4.7e) by B^2 and adding, we obtain:

$$-Y = [(A^2A^1 + B^2C^1)Y + (A^2B^1 + B^2D^1)U]i\sigma_3$$

Similarly multiplying (4.7c) by C^2, (4.7e) by D^2 and adding, we obtain:

$$-U = [(C^2A^1 + D^2C^1)Y + (C^2B^1 + D^2D^1)U]i\sigma_3$$

These last two equations may be cast into one simpler equation:

$$i\,\gamma^1\,\gamma^2\,\binom{Y}{U} = \binom{Y\sigma_3}{U\sigma_3} \qquad (4.9)$$

Moreover one can check that the process can be followed backward so that (4.9) is really equivalent to (4.7c) and (4.7e).

In exactly the same way we obtain an equation equivalent to (4.7d) and (4.7f):

$$i\,\gamma^3\,\gamma^1\,\binom{Y}{U} = \binom{Y\sigma_2}{U\sigma_2} \qquad (4.10)$$

Now let's come back to the six left-out equations in (4.4). Substituting in them the expressions of X and Z given in (4.5) and (4.6) we obtain:

$$({A^i}^2 + B^iC^i + I)Y + (A^iB^i + B^iD^i)U = 0 \quad \text{(no sum)}$$

$$({D^i}^2 + C^iB^i + I)U + (D^iC^i + C^iA^i)Y = 0 \quad \text{(no sum)}$$

$$(A^oA^i + B^oC^i - A^i)Y + (A^oB^i + B^oD^i - B^i)U = 0$$

$$(C^oA^i + D^oC^i - C^i)Y + (C^oB^i + D^oD^i - D^i)U = 0$$

$$(A^o + I)Y + B^oU = 0$$

$$C^oY + (D^o + I)U = 0$$

From (4.8) we see that the factors multiplying Y and U in the first two equations are 0. Thus they are trivially satisfied and contain no information. The next two can be rewritten as

$$[\gamma^o\gamma^i - \gamma^i] \binom{Y}{U} = 0 \qquad (4.11')$$

and the last two as

$$[\gamma^o + I] \binom{Y}{U} = 0 \qquad (4.11)$$

But these equations are one and the same as (4.11') is $-\gamma^i$ times (4.11).

Let us summarize the results obtained so far. We found that Y and U are determined by (4.9), (4.10) and (4.11) and then X and Z follow from (4.5) and (4.6).

If we write $\binom{Y}{U} = (u\ v)$, where u and v are column vectors, the equations (4.9), (4.10) and (4.11) translate to:

$$i\ \gamma^1\gamma^2\ (u\ v) = (u\ -v)$$

$$i\ \gamma^3\gamma^1\ (u\ v) = (iv\ -iu)$$

$$(\gamma^o + I)(u\ v) = (0\ 0)$$

This system is equivalent to

$$(\gamma^1 + i\,\gamma^2)u = 0$$

$$(I + \gamma^o)u = 0$$

$$v = \gamma^3\gamma^1\,u$$

Similarly if we set $\binom{X}{Z} = (s\ t)$, where s, t are column vectors, we find that (4.5) and (4.6) can be written as:

$$(s\ t) = \gamma^1\,(v\ u)$$

Whence we may write finally:

$$S = \begin{pmatrix} X & Y \\ Z & U \end{pmatrix} = (s\ t\ u\ v)$$

$$= (\gamma^3 u,\ \gamma^1 u,\ u,\ \gamma^3\gamma^1 u)$$

where u is determined by

$$(\gamma^1 + i\,\gamma^2)u = 0$$

$$(I + \gamma^o)u = 0$$

As one might expect, once we know how to find S, it becomes a simple matter to find T (which has to be a multiple of S^{-1}).

The equations satisfied by T are

$$\Gamma^\mu\,T = T\,\gamma^\mu$$

Taking the adjoint and remembering that $\Gamma^{o^+} = \Gamma^o$ while $\Gamma^{i^+} = -\Gamma^i$, we obtain:

$$\gamma^{o^+} \cdot T^+ = T^+ \Gamma^o$$

$$-\gamma^{i^+} \Gamma^+ = T^+ \Gamma^i$$

But $\{\gamma^{o^+}, -\gamma^{i^+}\}$ is also a system of γ matrices. Hence the last two equations are just like the equations for S. Therefore using the solution just obtained for S we conclude:

$$T^+ = (-\gamma^{3^+} w^+, -\gamma^{1^+} w^+, w^+, \gamma^{3^+} \gamma^{1^+} w^+)$$

where w^+ is determined by:

$$(-\gamma^{1^+} - i\, \gamma^{2^+})w^+ = 0$$

$$(I + \gamma^{o^+})w^+ = 0$$

More conveniently we may now write:

$$T = \begin{bmatrix} -w\ \gamma^3 \\ -w\ \gamma^1 \\ w \\ w\ \gamma^1\gamma^3 \end{bmatrix}$$

where w is a row vector determined by

$$w(\gamma^1 - i\, \gamma^2) = 0$$

$$w(I + \gamma^o) = 0$$

We also asserted at the beginning that the equations $(\gamma^1 + i\,\gamma^2)u = 0$ and $(I + \gamma^o)u = 0$ determine u up to a multiplicative factor. To see that this is true one first checks that the statement is correct when γ^μ is simply Γ^μ. This is trivial and we don't do it here. Now an arbitrary system $\{\gamma^\mu\}$ is related to $\{\Gamma^\mu\}$ by a similarity transformation. Whence the equations $(\gamma^1 + i\,\gamma^2)u = (I + \gamma^o)u = 0$ can be viewed as the equations $(\Gamma^1 + i\,\Gamma^2)u = (I + \Gamma^o)u = 0$ formulated in another basis. Accordingly if the solutions of the second system form a one-dimensional subspace, so will the solutions of the first system. A similar comment applies to w. This completes the proof.

(2) Application to the equation $\Omega^\mu{}_\nu \gamma^\nu = S^{-1}\,\gamma^\mu\,S$

The customary way of solving the equation $\Omega^\mu{}_\nu \gamma^\nu = S^{-1}\,\gamma^\mu\,S$ for S uses infinitesimal transformations. One puts suitable constraints on S and shows that the corresponding solutions form a group, a "double-valued representation" of the orthochronous Lorentz group. The matrices S are then obtained by "exponentiating" the Lie algebra of this group. We shall discuss this method in detail later on and in particular use it to show that the group of the S matrices is isomorphic to a certain very concrete group. For the moment we want to show how the results of the last section can be used to find S explicitely.

This method definitely lacks the elegance of the one using infinitesimal transformations but it has the advantage of being purely algebraic.

Our aim is to find $S(\Omega)$ such that

$$\hat{\gamma}^{\mu} \equiv \Omega^{\mu}{}_{\nu}\gamma^{\nu} = S^{-1}\,\gamma^{\mu}\,S$$

The set $\{\gamma^{\mu}\}$ is fixed. By the fundamental theorem there exist M and $V(\Omega)$ such that:

$$\gamma^{\rho} = M^{-1}\,\Gamma^{\rho}\,M \qquad \Omega^{\mu}{}_{\rho}\Gamma^{\rho} = V^{-1}\,\Gamma^{\mu}\,V$$

From these equations we deduce:

$$\Omega^{\mu}{}_{\rho}\gamma^{\rho} = \Omega^{\mu}{}_{\rho}M^{-1}\Gamma^{\rho}M = M^{-1}V^{-1}\Gamma^{\mu}VM = M^{-1}V^{-1}M\gamma^{\mu}M^{-1}VM$$

Thus the problem of finding $S(\Omega)$ for an arbitrary Ω is reduced to that of finding $V(\Omega)$:

$$S(\Omega) = M^{-1}\,V(\Omega)\,M \qquad (4.12)$$

In the preceding section we have established that $V(\Omega)$ is given by:

$$V = \begin{bmatrix} -w\,\hat{\Gamma}^{3} \\ -w\,\hat{\Gamma}^{1} \\ w \\ w\,\hat{\Gamma}^{1}\hat{\Gamma}^{3} \end{bmatrix} , \qquad \hat{\Gamma}^{\mu} \equiv \Omega^{\mu}{}_{\nu}\Gamma^{\nu}$$

where w is a solution of

$$w(\hat{\Gamma}^1 - i\,\hat{\Gamma}^2) = 0 \qquad (4.13)$$

$$w(I + \hat{\Gamma}^o) = 0 \qquad (4.14)$$

So all we have to do is to find w. To this end we decompose it in the following way:

$$w = (u,v) = (u_1,\ u_2,\ v_1,\ v_2) \qquad (4.15)$$

Equation (4.13) then reads:

$$(u,v)\left[\begin{array}{c|c} c_o\sigma^o & c_k\sigma_k \\ \hline -c_k\sigma_k & -c_o\sigma^o \end{array}\right] = 0$$

where $$c_\mu \equiv \Omega^1{}_\mu - i\Omega^2{}_\mu \qquad (4.16)$$

Written explicitely this gives:

$$c_o u - v\, c_k\, \sigma_k = 0$$

$$u c_k \sigma_k - c_o v = 0$$

There are two cases to be considered: $c_o \neq 0$ and $c_o = 0$. When $c_o \neq 0$, the general solution is

$$(u,v) = (u,\ \frac{1}{c_o}\, u c_k\, \sigma_k)\ ,\quad u \text{ arbitrary} \qquad (4.17)$$

When $c_o = 0$, the general solution is $w = (u,v)$ where

$$u c_k\, \sigma_k = v c_k\, \sigma_k = 0 \qquad (4.18)$$

In either case the solutions form a two dimensional space. Next we come to equation (4.14). It reads:

$$(u,v)\begin{bmatrix}(1+\Omega^{o}{}_{o})\sigma^{o} & \Omega^{o}{}_{k}\sigma_{k}\\ -\Omega^{o}{}_{k}\sigma_{k} & (1-\Omega^{o}{}_{o})\sigma^{o}\end{bmatrix} = 0$$

or ,

$$u(1+\Omega^{o}{}_{o}) - v\,\Omega^{o}{}_{k}\sigma_{k} = 0$$

$$v\,\Omega^{o}{}_{k}\sigma_{k} + v(1-\Omega^{o}{}_{o}) = 0$$

For the sake of simplicity we assume that Ω lies in the orthochronous group: $\Omega^{o}{}_{o} \geq 1 > 0$. The general solution is then:

$$(u,v) = \left(\frac{1}{1+\Omega^{o}{}_{o}}\,v\,\Omega^{o}{}_{k}\sigma_{k},\ v\right) \quad , \quad v \text{ arbitrary} \qquad (4.19)$$

Now w must be a common solution to (4.13) and((4.14). For the case $c_o \neq 0$, eqs. (4.17) and (4.19) give:

$$\left(u,\ \frac{u}{c_o}\,\underset{\sim}{c}\cdot\underset{\sim}{\sigma}\right) = \left(\frac{1}{1+\Omega^{o}{}_{o}}\,v\,\underset{\sim}{\Omega}^{o}\cdot\underset{\sim}{\sigma},\ v\right)$$

where we have set $(\Omega^{o}{}_{1},\Omega^{o}{}_{2},\Omega^{o}{}_{3}) = \underset{\sim}{\Omega}^{o}$ and $(c_1,c_2,c_3) = \underset{\sim}{c}$.

This gives:

$$v = \frac{u}{c_o}(\underset{\sim}{c}\cdot\underset{\sim}{\sigma}) \qquad , \qquad u = \frac{1}{(1+\Omega^{o}{}_{o})c_o}\,u(\underset{\sim}{c}\cdot\underset{\sim}{\sigma})(\underset{\sim}{\Omega}^{o}\cdot\underset{\sim}{\sigma})$$

The second equation may be rewritten as:

$$u(\underset{\sim}{c}\times\underset{\sim}{\Omega}^{o})\cdot\underset{\sim}{\sigma} = -i\,c_o\,u$$

Upon setting $(\underset{\sim}{c} \times \underset{\sim}{\Omega}^o)_i = a_i$, this equation for u reads

$$(u_3 + ic_o)u_1 + (a_1 + ia_2)u_2 = 0$$

$$(a_1 - ia_2)u_1 + (ic_o - a_3)u_2 = 0$$

If $a_1 + ia_2$ and $a_3 + ic_o$ don't both vanish, the solution is $u = (a_1 + ia_2, -(a_3 + ic_o))$. If they both vanish, the solution is $u = (ic_o - a_3, ia_2 - a_1)$.

In the case where $c_o = 0$, eqs. (4.18) and (4.19) tell us that the solution (u,v) common to both systems must satisfy

$$(u,v) = \left(\frac{1}{1+\Omega^o{}_o} v\, \underset{\sim}{\Omega}^o \cdot \underset{\sim}{\sigma},\ v\right) \qquad v\, \underset{\sim}{c} \cdot \underset{\sim}{\sigma} = 0$$

The second equation, when written explicitely, reads:

$$c_3 v_1 + (c_1 + ic_2)v_2 = 0$$

$$(c_1 - ic_2)v_1 - c_3 v_2 = 0$$

If c_3 and $c_1 - ic_2$ don't both vanish, the solution is $v = (c_3, c_1 - ic_2)$. If they both vanish, the solution is $v = (1,0)$.

We now summarize the results. We have defined:

$$\hat{\Gamma}^\mu \equiv \Omega^\mu{}_\nu \Gamma^\nu = V^{-1}(\Omega)\Gamma^\mu V(\Omega), \quad \Omega^o{}_o > 0$$

$$c_\mu \equiv \Omega^1{}_\mu - i\Omega^2{}_\mu \quad \underset{\sim}{c} \equiv (c_1, c_2, c_3) \quad \underset{\sim}{\Omega}^o \equiv (\Omega^o{}_1, \Omega^o{}_2, \Omega^o{}_3)$$

$$\underset{\sim}{a} \equiv \underset{\sim}{c} \times \underset{\sim}{\Omega}^o$$

and found that

$$V(\Omega) = \begin{bmatrix} -w\ \hat{\Gamma}^3 \\ -w\ \hat{\Gamma}^1 \\ w \\ w\ \hat{\Gamma}^1\ \hat{\Gamma}^3 \end{bmatrix}$$

where w is to be chosen according to the following table.

Table 3 Vectors for constructing spinor transformations

$c_o \neq 0$		$c_o = 0$	
$w = (u, \frac{u}{c_o}\ \underset{\sim}{c} \cdot \underset{\sim}{\sigma})$		$w = (\frac{1}{1+\Omega^o_o}\ u\ \underset{\sim}{\Omega}^o \cdot \underset{\sim}{\sigma},\ u)$	
$\|a_1+ia_2\|+\|a_3+ic_o\| \neq 0$	$\|a_1+ia_2\|+\|a_3+ic_o\| = 0$	$\|c_3\|+\|c_1-ic_2\| \neq 0$	$\|c_3\|+\|c_1-ic_2\| = 0$
$u = (a_1+ia_2, -(a_3+ic_o))$	$u = (ic_o-a_3, ia_2-a_1)$	$u = (c_3, c_1-ic_2)$	$u = (1,0)$

As an illustration we obtain $V(\Omega)$ when Ω is a [illegible] boos[illegible]he x axis. If the frame K' moves with spee[illegible] β [illegible]ect to the frame K, the primed coordinates [illegible]re [illegible] the unprimed by $x' = \Omega x$ where:

$$\Omega = \left[\begin{array}{cc|c} \gamma & \gamma\beta & \\ -\gamma\beta & \gamma & 0 \\ \hline 0 & & I \end{array}\right], \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

In [illegible] case we have: $c_o = -\gamma\beta$, $\underset{\sim}{c} = (\gamma,-i,0)$, $\underset{\sim}{\Omega}^o = (-\gamma\beta,0,0)$. By [illegible]g table 3 we obtain $w = (0,2i\gamma\beta,-2i(\gamma+1),0)$ from whic[illegible]

$$V(\Omega) = -2i \begin{bmatrix} \gamma+1 & 0 & 0 & -\gamma\beta \\ 0 & \gamma+1 & -\gamma\beta & 0 \\ 0 & -\gamma\beta & \gamma+1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma+1 \end{bmatrix}$$

Since $V(\Omega)$ is determined up to a multiplicative factor we may drop the $-2i$. The resulting matrix is then:

$$V(\Omega) = (\gamma+1)I - \gamma\beta\, \Gamma^o\Gamma^1$$

It is customary to write $\gamma = ch\rho$ $\gamma\beta = sh\rho$. If we divide our $V(\Omega)$ by $2ch\rho/2$ we obtain the matrix $ch\rho/2I - sh\rho/2\Gamma^o\Gamma^1$. This expression agrees with the one obtained by the method of infinitesimal transformations as will be seen later.

CHAPTER V

THE GROUP OF SPINOR TRANSFORMATIONS

(1) Construction of the group

In this section we look at the set of all spinor transformations corresponding to all orthochronous Lorentz transformations. We study this set as a group. We give up the purely algebraic method devised in chapter IV and switch to the standard method using infinitesimal transformations, or the Lie algebras. The orthochronous Lorentz group will be denoted by $L^{\uparrow}$. It consists of all Lorentz transformations Ω with $\Omega^{o}{}_{o} > 0$. The proper subgroup of orthochronous transformations with determinant 1 will be denoted by $L^{\uparrow}_{+}$. We work throughout with a fixed system $\{\gamma^{\mu}\}$ which will be assumed to be unitary. We are interested in the solution $\Lambda(\Omega)$ to the equation

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \quad , \quad \Omega \in L^{\uparrow} \tag{5.1}$$

We already know that the solution Λ is determined up to a multiplicative factor. We want to remove as much as possible this arbitrariness. The first restriction that can be imposed is the following.

<u>Proposition</u> (5.1): The solution of eq. (5.1) may be chosen so as to satisfy $\Lambda^{+} = \gamma^{o}\Lambda^{-1}\gamma^{o}$.

Proof: Let us write $\hat{\gamma}^{\mu} = \Omega^{\mu}{}_{\nu}\gamma^{\nu}$ and let Λ be a solution of eq. (5.1). Then we have:

$$\hat{\gamma}^{\mu +} = \Omega^{\mu}{}_{\nu}\gamma^{\nu +} = \Omega^{\mu}{}_{\nu}\gamma^{o}\gamma^{\nu}\gamma^{o} = \gamma^{o}\hat{\gamma}^{\mu}\gamma^{o} = \gamma^{o}\Lambda^{-1}\gamma^{\mu}\Lambda\gamma^{o}$$

from which $\hat{\gamma}^{\mu} = \gamma^{o}\Lambda^{+}\gamma^{\mu +}\Lambda^{-1+}\gamma^{o} = (\gamma^{o}\Lambda^{+}\gamma^{o})\gamma^{\mu}(\gamma^{o}\Lambda^{+}\gamma^{o})^{-1} = \Lambda^{-1}\gamma^{\mu}\Lambda$

and $[\Lambda\gamma^{o}\Lambda^{+}\gamma^{o}, \gamma^{\mu}] = 0$. Therefore $\Lambda\gamma^{o}\Lambda^{+}\gamma^{o} = cI$ or $\Lambda^{+} = c\gamma^{o}\Lambda^{-1}\gamma^{o}$. This implies

$$\Lambda^{+}\Lambda = c\gamma^{o}\Lambda^{-1}\gamma^{o}\Lambda = c\gamma^{o}\hat{\gamma}^{o} = c\gamma^{o}(\Omega^{o}{}_{o}\gamma^{o} + \Omega^{o}{}_{k}\gamma^{k})$$

$$= c\Omega^{o}{}_{o}I + c\Omega^{o}{}_{k}\gamma^{o}\gamma^{k}$$

Taking the trace on both sides and using proposition (2.3) yields $T_{r}\Lambda^{+}\Lambda = 4c\Omega^{o}{}_{o}$. Since $\Lambda^{+}\Lambda$ is hermitian it follows that c is real. Since $\Lambda^{+}\Lambda$ is positive, it follows that c is positive ($\Omega^{o}{}_{o} > 0$). Now if we take $\Lambda' = \frac{1}{\sqrt{c}}\Lambda$, we obtain:

$$\Lambda'^{+} = \frac{1}{\sqrt{c}}\Lambda^{+} = \frac{1}{\sqrt{c}}c\gamma^{o}\Lambda^{-1}\gamma^{o} = \sqrt{c}\,\gamma^{o}\frac{(\Lambda')^{-1}}{\sqrt{c}}\gamma^{o} = \gamma^{o}(\Lambda')^{-1}\gamma^{o}.$$

The motivation for this restriction is to make the transformation of the adjoint $\bar{\psi} \equiv \psi^{+}\gamma^{o}$ simple:

$$\bar{\psi}' = \psi'^{+}\gamma^{o} = \psi^{+}\Lambda^{+}\gamma^{o} = \bar{\psi}\gamma^{o}\Lambda^{+}\gamma^{o} = \bar{\psi}\Lambda^{-1}$$

The usefulness of this will be appreciated in chapter VI. Before we go further we need to introduce a special matrix, called the B matrix, associated to any system of γ matrices. Let $\{\gamma^\mu\}$ be such an arbitrary system. The set $\{\gamma^{\mu *}\}$, where * means complex conjugate, is also a system of γ matrices. By the fundamental theorem there exists a matrix B such that $\gamma^{\mu *} = B^{-1}\gamma^\mu B$. We require that $|\det B| = 1$. This determines B up to a phase factor. If $\{\hat\gamma^\mu\}$ is another arbitrary system with corresponding $\hat B$ and $\hat\gamma^\mu = S^{-1}\gamma^\mu S$, then $\hat B$ is related to B by:

$$\hat B = e^{i\phi} S^{-1} B S^{*} \quad , \quad \phi \text{ real} \tag{5.2}$$

Indeed we have:

$$\hat\gamma^{\mu *} = S^{-1 *}\gamma^{\mu *} S^{*} = S^{-1 *} B^{-1}\gamma^\mu B S^{*}$$

$$= (S^{-1} B S^{*})^{-1} \hat\gamma^\mu S^{-1} B S^{*}$$

Moreover $|\det S^{-1} B S^{*}| = 1$. Since $\hat B$ is determined up to a phase, the conclusion follows.

The properties of the B matrix of use to us are contained in the following proposition.

<u>Proposition</u> (5.2): The B matrix of a unitary $\{\gamma^\mu\}$ system is unitary and antisymmetric.

Proof: Since γ^μ is unitary, so is $\gamma^{\mu *}$. By proposition (3.3) there exists a unitary matrix connecting γ^μ and $\gamma^{\mu *}$.

Clearly this can be taken as B and any phase multiple will also be unitary. We now have to show that it is antisymmetric. For the set $\{\Gamma^\mu\}$ of eq. (2.7) one easily checks that one can take $B' = \Gamma^o\Gamma^3\Gamma^1$. Notice that $B'^T = -B'$. By proposition (3.3) there exists a unitary U such that $\gamma^\mu = U^+\Gamma^\mu U$, and by eq. (5.2) one has

$$B = e^{i\phi}U^+B'U^* = e^{i\phi}(U^*)^TB'U^*$$

From this the antisymmetry of B' clearly implies that of B.

We are now prepared to put all the restrictions on Λ.

Theorem (5.1): The equations

$$\Omega^\mu{}_\nu\gamma^\nu = \Lambda^{-1}\gamma^\mu\Lambda \qquad (5.3a)$$

$$\Lambda^+ = \gamma^o\Lambda^{-1}\gamma^o \qquad (5.3b)$$

$$\Lambda^* = B^+\Lambda B \qquad (5.3c)$$

where Λ is the unknown and $\Omega\epsilon L^\dagger$ have exactly two solutions. One is -1 times the other.

Proof: By proposition (5.1) we know that eqs. (5.3a), (5.3b) have a common solution Λ_o. We have:

$\Omega^\mu{}_\nu\gamma^{\nu*} = \Lambda^{-1*}\gamma^{\mu*}\Lambda^*$ which may be rewritten as

$\Omega^{\mu}{}_{\nu} B^{+}\gamma^{\nu}B = \Lambda^{-1*}B^{+}\gamma^{\mu}B\Lambda^{*}$, from which we deduce

$\Lambda^{-1}\gamma^{\mu}\Lambda = \Omega^{\mu}{}_{\nu}\gamma^{\nu} = B\Lambda^{-1*}B^{+}\gamma^{\mu}B\Lambda^{*}B^{+}$ or

$$[\gamma^{\mu}, B\Lambda^{*}B^{+}\Lambda^{-1}] = 0$$

By theorem (2.8) it follows as usual that $B\Lambda^{*}B^{+}\Lambda^{-1} = cI$, which implies $\frac{(\det \Lambda)^{*}}{\det \Lambda} = c^{4}$. Whence $|c| = 1$, or $c = e^{i\lambda}$. From $B\Lambda^{*}B^{+}\Lambda^{-1} = cI$, we get $\Lambda^{*} = e^{i\lambda}B^{+}\Lambda B$. So if we define $\Lambda' = e^{i\alpha}\Lambda$, we obtain:

$$\Lambda'^{*} = e^{-i\alpha}\Lambda^{*} = e^{i(\lambda-2\alpha)}B^{+}\Lambda' B$$

and the choice $\alpha=\lambda/2$ yields the solution to our three equations. Clearly if Λ is a solution, so is $-\Lambda$. Now suppose that $\tilde{\Lambda}$ is another solution to (5.3a,b,c). We know that $\tilde{\Lambda} = a\Lambda$ for some complex number a. Equation (5.3c) shows that a is real. Eq. (5.3b) then shows that its square is 1. Whence $\tilde{\Lambda} = \pm\Lambda$. This completes the proof.

For some purposes, especially when dealing with the Lie algebras, it is convenient to reexpress the three equations of theorem (5.2) in only two. This is achieved by the following result:

<u>Proposition</u> (5.3): Equations (5.3a,b,c) are equivalent to the two equations:

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \tag{5.4a}$$

$$\Lambda^{T}(B^{+}\gamma^{o})\Lambda = B^{+}\gamma^{o} \tag{5.4b}$$

Proof: We first show that (5.3) implies (5.4). Of course (5.3) implies (5.4a). Next we have, using (5.3b,c),

$$\Lambda^{T} = \gamma^{o*}\Lambda^{-1*}\gamma^{o*} = \gamma^{o*}B^{+}\Lambda^{-1}B\gamma^{o*} = B^{+}\gamma^{o}BB^{+}\Lambda^{-1}BB^{+}\gamma^{o}B$$
$$= B^{+}\gamma^{o}\Lambda^{-1}\gamma^{o}B$$

from which (5.4b) follows.

Now we have to show that (5.4) implies (5.3), that is, if $\bar{\Lambda}$ is a solution of (5.4a,b) then it satisfies (5.3b,c). Let Λ be a solution of (5.3). From what we've just seen, Λ is also a solution of (5.4). Therefore we may write:

$$\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o} = \Lambda^{-1} \quad , \quad \text{from which}$$

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda = \gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda$$

Similarly we have for $\bar{\Lambda}$:

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \bar{\Lambda}^{-1}\gamma^{\mu}\bar{\Lambda} = \gamma^{o}B\bar{\Lambda}^{T}B^{+}\gamma^{o}\gamma^{\mu}\bar{\Lambda}$$

From the fundamental theorem we know that $\bar{\Lambda} = c\Lambda$ and the last two equations for $\Omega^{\mu}{}_{\nu}\gamma^{\nu}$ give:

$$\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda = c^{2}\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda$$

or $c^2 = 1$. Whence $\bar{\Lambda} = \pm\Lambda$ and $\bar{\Lambda}$ is a solution of (5.3).

Equations (5.4a) and (5.4b) show that Λ can be interpreted as the transition matrix of a change of basis under which the matrices of the operators γ^μ become $\Omega^\mu{}_\nu\gamma^\nu$ and the matrix of the bilinear form defined by $B^+\gamma^o$ remains invariant. When we have gained some more information about our Λ matrices we will give a third completely different way of formulating (5.3a,b,c).

We now introduce some handy notation. The set of all Λ's solutions of (5.3a,b,c) when Ω goes over $L^\dagger$ will be denoted by $S^\dagger$. Given $\Lambda\varepsilon S^\dagger$, there is only one Ω in $L^\dagger$ such that Λ is a solution of (5.3a): this follows from the linear independence of the γ^μ's. The one Ω corresponding to Λ will be denoted by Ω_Λ and the map $\Lambda \to \Omega_\Lambda$ will be denoted by Π. Our first statement about the Λ's is the following.

<u>Theorem</u> (5.2): $S^\dagger$ is a six-dimensional Lie group locally isomorphic to $L^\dagger$. Π: $S^\dagger \to L^\dagger$: $\Omega \to \Omega_\Lambda$ is a homomorphism and a local isomorphism.

Proof: We first prove that $S^\dagger$ is a group. Let $\Lambda, \bar{\Lambda}$ be solutions of (5.4a) and (5.4b) with corresponding $\Omega, \bar{\Omega}$. Then we have:

$$\bar{\Omega}^\mu{}_\rho\Omega^\rho{}_\nu\gamma^\nu = \bar{\Omega}^\mu{}_\rho\Lambda^{-1}\gamma^\rho\Lambda = \Lambda^{-1}\bar{\Lambda}^{-1}\gamma^\mu\bar{\Lambda}\Lambda = (\bar{\Lambda}\Lambda)^{-1} \text{[illegible]}$$

Hence $\bar{\Lambda}\Lambda$ is a solution of (5.4a) corresponding to [illegible]

Moreover $\bar{\Lambda}\Lambda$ clearly satisfies (5.4b). Thus $S^{\dagger}$ is closed under multiplication. To show that Λ^{-1} belongs to $S^{\dagger}$ if Λ does we first observe that $(\Omega^{-1})^{\mu}{}_{\nu} = g^{\mu\rho}\Omega_{\rho}{}^{\sigma}g_{\sigma\nu}$. Whence, if Λ satisfies (5.4a), we have:

$$(\Omega^{-1})^{\mu}{}_{\nu}\gamma^{\nu} = g^{\mu\rho}\Omega_{\rho}{}^{\sigma}g_{\sigma\nu}\gamma^{\nu}$$

$$= g^{\mu\rho}\Omega_{\rho}{}^{\sigma}g_{\sigma\nu}\Omega^{\nu}{}_{\theta}\Lambda\gamma^{\theta}\Lambda^{-1}$$

$$= g^{\mu\rho}g_{\rho\theta}\Lambda\gamma^{\theta}\Lambda^{-1} = \Lambda\gamma^{\mu}\Lambda^{-1}$$

Whence Λ^{-1} is a solution of (5.4a) corresponding to $\Omega^{-1}\epsilon S^{\dagger}$. To show that $(\Lambda^{-1})^{T}B^{+}\gamma^{o}\Lambda^{-1} = B^{+}\gamma^{o}$ one simply multiplies eq. (5.4b) on the left by $(\Lambda^{-1})^{T}$ and on the right by Λ^{-1}. These considerations show at the same time that Π is an algebraic homomorphism. We now know that $S^{\dagger}$ is an algebraic subgroup of the Lie group GL(4,C) of 4×4 invertible complex matrices. To show that it is also a Lie subgroup we simply invoke the well-known Cartan theorem which says that an algebraic subgroup of a Lie group which is also a topologically closed subset is a Lie subgroup. It is easy to show that $S^{\dagger}$ is closed in GL(4,C). Let (Λ_n) be a sequence of elements of $S^{\dagger}$ converging to $L\epsilon GL(4,C)$. We want to show that L is in $S^{\dagger}$. It is clear by continuity that L satisfies (5.4b).

Therefore, all we have to do is to show that there exists Ω in $L^{\uparrow}$ such that (5.4a) holds with $\Lambda = L$. We have $(\Omega_{\Lambda_n})^{\mu}{}_{\nu} = \frac{1}{4} T_r(\Lambda_n^{-1} \gamma^{\mu} \Lambda_n \gamma_{\nu})$. Since $L^{\uparrow}$ is closed in GL(4,R), the group of 4×4 real invertible matrices, $\Omega^{\mu}{}_{\nu} \equiv \lim_{n\to\infty} (\Omega_{\Lambda_n})^{\mu}{}_{\nu}$ defines an orthochronous Lorentz transformation and by continuity one clearly has $\Omega^{\mu}{}_{\nu}\gamma^{\nu} = L^{-1}\gamma^{\mu}L$. Whence L belongs to $S^{\uparrow}$ and $S^{\uparrow}$ is a Lie group. We now want to show that $\Pi: \Lambda \to \Omega_{\Lambda}$ is a local isomorphism. We already know that it is a homomorphism. The fact that it is smooth follows clearly from $\Omega_{\Lambda}{}^{\mu}{}_{\nu} = \frac{1}{4} T_r(\Lambda^{-1}\gamma^{\mu}\Lambda\gamma_{\nu})$. So all that remains to be shown is that it is locally injective. Let $\Omega \in L^{\uparrow}$ and $\Lambda \in S^{\uparrow}$ such that $\Omega_{\Lambda} = \Omega$. There is only one other solution L to $\Omega_L = \Omega$ and it is $L = -\Lambda$. So all we have to do is to take a small neighborhood U_{Λ} of Λ such that $-U_{\Lambda} \equiv \{-\Lambda;\ \Lambda \in U_{\Lambda}\}$ and U_{Λ} don't intersect: Π restricted to U_{Λ} is clearly one to one.

(2) <u>Explicit form of the group elements</u>

We now proceed to obtain explicit expressions for the Λ's via infinitesimal transformations. Before we obtain the transformation laws of the spinors under the full orthochronous group $L^{\uparrow}$, we first investigate the subgroup of proper Lorentz transformations $L_{+}^{\uparrow}$. The corresponding Λ's obviously form a Lie-subgroup of $S^{\uparrow}$ which we will [illegible]

$S^\dagger_+$. It is clear that $S^\dagger_+$ is both open and closed in $S^\dagger$. Hence $S^\dagger$ is disconnected. We will show later that it has two components just as $L^\dagger$. We use the following standard basis of the Lie algebra of $L^\dagger_+$:

$$(I_{\alpha\beta})^\mu{}_\nu \equiv \delta^\mu{}_\alpha g_{\nu\beta} - \delta^\mu{}_\beta g_{\nu\alpha} \qquad (5.5)$$

One has $I_{\alpha\beta} = -I_{\beta\alpha}$ and a basis is obtained by taking $I_{01}, I_{02}, I_{03}, I_{12}, I_{13}, I_{23}$. The first three generate boosts; the last three generate rotations. The commutators are given by

$$[I_{\mu\nu}, I_{\kappa\lambda}] = -(g_{\mu\kappa}I_{\nu\lambda} + g_{\lambda\mu}I_{\kappa\nu} + g_{\nu\lambda}I_{\mu\kappa} + g_{\kappa\nu}I_{\lambda\mu}) \qquad (5.6)$$

To obtain the Lie algebra of $S^\dagger_+$ we simply use our local isomorphism $\Lambda \to \Omega_\Lambda$. If $\Omega(\tau)$ is a curve in $L^\dagger_+$ passing through I at $\tau = 0$, there is a unique curve $\Lambda(\tau)$ in $S^\dagger_+$ passing through I at $\tau = 0$ such that:

$$\Omega^\mu{}_\nu(\tau)\gamma^\nu = \Lambda^{-1}(\tau)\gamma^\mu\Lambda(\tau)$$

$$\Lambda^T(\tau)B^+\gamma^0\Lambda(\tau) = B^+\gamma^0$$

Taking the derivative of these equations at $\tau = 0$ gives:

$$[\dot\Lambda, \gamma^\mu] = -\dot\Omega^\mu{}_\nu\gamma^\nu$$

$$\dot\Lambda^T B^+\gamma^0 + B^+\gamma^0\dot\Lambda = 0$$

where we have set $\dot{\Lambda} \equiv \dot{\Lambda}(o)$, $\dot{\Omega}^{\mu}{}_{\nu} \equiv \dot{\Omega}^{\mu}{}_{\nu}(o)$. Now if the curve $\Omega(\tau)$ is chosen so that $\dot{\Omega} = I_{\alpha\beta}$ and we denote the corresponding $\dot{\Lambda}$ by $\dot{\Lambda}_{\alpha\beta}$, the equations for $\dot{\Lambda}$ read:

$$[\dot{\Lambda}_{\alpha\beta}, \gamma^{\mu}] = -(\delta^{\mu}{}_{\alpha} g_{\nu\beta} - \delta^{\mu}{}_{\beta} g_{\nu\alpha})\gamma^{\nu} = \delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta} \tag{5.7}$$

$$\dot{\Lambda}^{T}{}_{\alpha\beta} B^{+}\gamma^{o} + B^{+}\gamma^{o}\dot{\Lambda}_{\alpha\beta} = 0 \tag{5.8}$$

From table 2 we know that $[\gamma_{[\alpha\beta]}, \gamma^{\mu}] = 2(\delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta})$. Using this and eq. (5.7) we get $[\dot{\Lambda}_{\alpha\beta} - \frac{1}{2}\gamma_{[\alpha\beta]}, \gamma^{\mu}] = 0$, from which $\dot{\Lambda}_{\alpha\beta} = \frac{1}{2}\gamma_{[\alpha\beta]} + cI$. Inserting this in (5.8) yields:

$$0 = 2cB^{+}\gamma^{o} + \frac{1}{2}(\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} + B^{+}\gamma^{o}\gamma_{[\alpha\beta]}).$$ But

$$\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} = \gamma_{\beta}{}^{T}\gamma_{\alpha}{}^{T}B^{+}\gamma^{o} = \gamma^{\beta *}\gamma^{\alpha *}B^{+}\gamma^{o} = B^{+}\gamma^{\beta}\gamma^{\alpha}\gamma^{o}$$ and

$$\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} + B^{+}\gamma^{o}\gamma_{[\alpha\beta]} = B^{+}(\gamma^{\beta}\gamma^{\alpha}\gamma^{o} + \gamma^{o}\gamma_{[\alpha\beta]}) = 0$$

Therefore $c = 0$. Whence we may write:

$$\dot{\Lambda}_{\alpha\beta} = \frac{1}{2}\gamma_{[\alpha\beta]} \tag{5.9}$$

$$\Pi(\exp \frac{c^{\alpha\beta}}{2}\gamma_{[\alpha\beta]}) = \exp(c^{\alpha\beta}I_{\alpha\beta}) \tag{5.10}$$

Our first use of formula (5.10) is to get explicit forms. We first consider the transformation of spinors under rotations. When one rotates a frame through an angle ϕ around a unit vector $\underline{n}$, the coordinates x' in the new frame are related to the old coordinates $\underset{\sim}{x}$ by

$$x' = [\exp - \phi(\underset{\sim}{n}\cdot\underset{\sim}{A})]\,x \qquad \text{where}$$

$$A_1 \equiv I_{23} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_2 \equiv I_{31} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix},$$

$$A_3 \equiv I_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{5.11}$$

It is convenient at this stage to introduce the matrices Σ_i, defined by

$$\Sigma_1 = \sigma_{23}\ , \quad \Sigma_2 = \sigma_{31}\ , \quad \Sigma_3 = \sigma_{12} \tag{5.12}$$

where, as we recall, $\sigma_{\mu\nu} \equiv i\gamma_{[\mu\nu]}$. One easily checks that they satisfy the same relations as the Pauli matrices:

$$\Sigma_i\Sigma_j = i\epsilon_{ijk}\Sigma_k + \delta_{ij}I \tag{5.13}$$

from which one deduces:

$$(\underset{\sim}{a}\cdot\underset{\sim}{\Sigma})(\underset{\sim}{b}\cdot\underset{\sim}{\Sigma}) = (\underset{\sim}{a}\cdot\underset{\sim}{b})I + i(\underset{\sim}{a}\times\underset{\sim}{b})\cdot\underset{\sim}{\Sigma} \tag{5.14}$$

We now apply eq. (5.10) with $c^{\alpha\beta}I_{\alpha\beta} = -\phi\ \underset{\sim}{n}\cdot\underset{\sim}{A}$. The corresponding Λ on the left-hand side is $\exp(i\frac{\phi}{2}\ \underset{\sim}{n}\cdot\underset{\sim}{\Sigma})$. This is easily evaluated. Using (5.14) we see that $(\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p = I$ and $(\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p = \underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$ for p even and odd respectively. Hence:

$$\exp i\,\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma} = \sum_{p=0}^{\infty}\frac{(i\,\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p}{p!}$$

$$= \sum_{k=0}^{\infty}\frac{(i)^{2k}(\frac{\phi}{2})^{2k}I}{(2k)!} + \sum_{k=0}^{\infty}\frac{(i)^{2k+1}(\frac{\phi}{2})^{2k+1}}{(2k+1)!}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$$

$$= \cos\frac{\phi}{2}\,I + i\sin\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$$

Whence if we denote by $\Lambda_{\underset{\sim}{n}}(\phi)$ the transformation corresponding to a rotation of angle ϕ around the unit vector $\underset{\sim}{n}$ we obtain:

$$\Lambda_{\underset{\sim}{n}}(\phi) = \cos\frac{\phi}{2}\,I + i\sin\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma} \qquad (5.15)$$

One peculiar feature of this equation is that $\Lambda_{\underset{\sim}{n}}(2\pi) = -I$. This is characteristic of spin $\frac{1}{2}$ wave functions. The spin operator along direction $\underset{\sim}{n}$ is the infinitesimal generator of the unitary group of transformations of the internal variables under rotations around $\underset{\sim}{n}$. We have seen above that this group of transformations is given by $\exp(i\,\phi\,\frac{1}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})$. Whence the spin operator along direction $\underset{\sim}{n}$ is $\frac{1}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$. Using the eq. (5.13) we find indeed that the spin vector operator

$$\underset{\sim}{s} = \frac{1}{2}\,\underset{\sim}{\Sigma} \qquad (5.16)$$

satisfies the characteristic commutation relations of an angular momentum:

$$[S_i, S_j] = i\varepsilon_{ijk}S_k \qquad (5.17)$$

We have $S_3 = \frac{i}{2}\gamma_1\gamma_2 = \frac{i}{2}\gamma^1\gamma^2$. By the fundamental theorem $\frac{i}{2}\gamma^1\gamma^2$ is similar to $\frac{i}{2}\Gamma^1\Gamma^2$ (cf. eq. (2.7)). But $i\Gamma^1\Gamma^2 = \frac{1}{2}\begin{bmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{bmatrix}$. Whence the eigenvalues of S_3 are $\pm\frac{1}{2}$: $\underset{\sim}{S}$ is a spin $\frac{1}{2}$ operator. Of course this corresponds to the experimental spin of the electron. As far as physics is concerned the bizarre result $\Lambda_{\underset{\sim}{n}}(2\pi) = -I$ is of no consequence. All physical quantities arise through quadratic expressions which are insensitive to the interchange of ψ and $-\psi$. From a mathematical point of view the result shows that we cannot remove the sign arbitrariness of the elements of $S^{\uparrow}$ without giving up their property of forming a Lie group: $-I$ is obtained by exponentiating an element of the Lie algebra.

Having discussed rotations and spin we now turn to the transformation of spinors under boosts. Suppose that the frame K', coinciding with the frame K at t = 0, moves at speed $\beta\underset{\sim}{n}$ with respect to it. (Here $\underset{\sim}{n}$ is a unit vector and $\beta \equiv v/c = v$ with our choice of units.) Let $\rho = th^{-1}(\beta)$. Then the primed coordinates are related to the unprimed by $x' = \Omega_{\underset{\sim}{n}}(\rho)x$ where $\Omega_{\underset{\sim}{n}}(\rho) = \exp \rho\underset{\sim}{n}\cdot\underset{\sim}{I}_o$. Here we have set:

$$\underset{\sim}{I}_o \equiv (I_{01}, I_{02}, I_{03}) \quad \text{(cf. (5.[illegible]))} \qquad \text{([illegible])}$$

so that

$$\underset{\sim}{n}\cdot\underset{\sim}{I}_o = -\begin{bmatrix}0 & \underset{\sim}{n}^T\\ \underset{\sim}{n} & 0\end{bmatrix} \qquad \text{([illegible])}$$

Let $\underset{\sim}{\gamma} \equiv (\gamma^1, \gamma^2, \gamma^3)$. One easily checks the following product rule:

$$(\underset{\sim}{a}\cdot\underset{\sim}{\gamma})(\underset{\sim}{b}\cdot\underset{\sim}{\gamma}) = -[\underset{\sim}{a}\cdot\underset{\sim}{b}\ I + i\ \underset{\sim}{a}\times\underset{\sim}{b}\ \cdot\underset{\sim}{\Sigma}] \qquad (5.20)$$

We now apply eq. (5.10) with $c^{\alpha\beta} I_{\alpha\beta} = \rho \underset{\sim}{n}\cdot\underset{\sim}{I}_o$. The corresponding Λ on the left-hand side of (5.10) is $\exp(\frac{1}{2}\rho\ n^i\ \gamma_{[oi]})$. Using eq. (5.20) we obtain:

$$(n^i\ \gamma_{[oi]})^2 = -(\underset{\sim}{n}\cdot\underset{\sim}{\gamma})^2 = I .$$

Whence we have:

$$\exp(\frac{1}{2}\rho\ n^i\ \gamma_{[oi]}) = \sum_{p=0}^{\infty} \frac{(\frac{1}{2}\rho)^p}{p!} (n^i \gamma_{[oi]})^p$$

$$= \sum_{k=0}^{\infty} \frac{(\frac{\rho}{2})^{2k}}{(2k)!} I + \sum_{k=0}^{\infty} \frac{(\frac{\rho}{2})^{2k+1}}{(2k+1)!} (n^i\ \gamma_{[oi]})$$

$$= ch\frac{\rho}{2}\ I + sh\frac{\rho}{2}\ n^i\ \gamma_{[oi]}$$

Or, using the standard notation of eq. (1.3b):

$$\Lambda_{\underset{\sim}{n}}(\rho) = ch\frac{\rho}{2}\ I - sh\frac{\rho}{2}\ \underset{\sim}{n}\cdot\underset{\sim}{\alpha} \qquad (5.21)$$

One may check that this agrees with the expression obtained in chapter IV, section (2) with $\underset{\sim}{n} = (1,0,0)$.

An arbitrary proper Lorentz transformation Ω decomposes uniquely into the product of a boost and a rotation:

$\Lambda = \exp-(\phi \underset{\sim}{n}\cdot\underset{\sim}{A}) \exp(\rho \underset{\sim}{u}\cdot\underset{\sim}{I}_o)$, $\underset{\sim}{u}$: unit vector.

Therefore the transformation of spinors under the proper Lorentz group $L_+^\dagger$ is completely described by the formula:

$$\Pi[(\cos\frac{\phi}{2} I + i \sin\frac{\phi}{2} \underset{\sim}{n}\cdot\underset{\sim}{\Sigma})(ch\frac{\rho}{2} I - sh\frac{\rho}{2} \underset{\sim}{n}\cdot\underset{\sim}{\alpha})] \quad .$$

$$= \exp - \phi \underset{\sim}{n}\cdot\underset{\sim}{A} \exp \rho \underset{\sim}{u}\cdot\underset{\sim}{I}_o \qquad (5.22)$$

To give the transformation law of the spinors under an arbitrary element of the full orthochronous group $L^\dagger$ it is now sufficient to say how they transform under the space reflection $\Omega_s = \begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix}$, because any Ω in $L^\dagger$ which is not in $L_+^\dagger$ can be written as $\Omega = \Omega_s\Omega_+$, where Ω_+ is in $L_+^\dagger$. So let us find the Λ_s corresponding to Ω_s. Equation (5.4a) gives $\gamma^o = \Lambda_s^{-1}\gamma^o\Lambda_s$, $-\gamma^k = \Lambda_s^{-1}\gamma^k\Lambda_s$, from which it immediately follows that $\Lambda_s = c\gamma^o$. To determine c we use eq. (5.4b) which says: $c^2\gamma^{o*}B^+\gamma^o\gamma^o = B^+\gamma^o$ or $c^2B^+\gamma^o = B^+\gamma^o$. Whence $c = \pm 1$. Therefore:

$$\Lambda_s = \pm \gamma^o \qquad (5.23)$$

The group $S^\dagger$ that we have been considering arose through the transformations of the spinors under orthochronous Lorentz transformations. The reason we didn't consider the full Lorentz group is that eq. (5.4b) cannot be satisfied for time reversal $\Omega_t = \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix}$. For this Ω_t, eq. (5.4a) says that the corresponding Λ must anticommute with γ^o and commute with $\underset{\sim}{\gamma}$. The solution to this is $c\gamma^o\gamma^5$ and one easily checks that this cannot satisfy (5.4b).

(3) Charge conjugation and time reversal

Besides those induced by orthochronous Lorentz transformations, there are two other important types of spinor transformations: charge conjugation and time reversal.

Suppose that a given representation $\{\gamma^\mu\}$ has been chosen. Let B be the matrix of proposition (5.2): $\gamma^{\mu *} = B^{-1}\gamma^\mu B$. Then the charge-conjugate spinor ψ^c of the spinor ψ is defined by

$$\psi^c = \gamma^5 B\psi^*$$

where ψ^* is the complex conjugate of ψ. Let us denote by K the antiunitary operator of complex conjugation in that particular representation. Then $\psi^c = \gamma^5 BK\psi \equiv K_c\psi$. The operator BK commutes with γ^μ. Indeed we have:

$$\gamma^\mu BK\psi = \gamma^\mu B\psi^* = B\gamma^{\mu *}\psi^* = BK\gamma^\mu\psi$$

Suppose ψ satisfies the Dirac equation

$$[\gamma^\mu(i\partial_\mu - eA_\mu) - m]\psi = 0$$

Multiplying this on the left by $\gamma^5 BK$ and remembering that BK is an antilinear operator commuting with γ^μ we get

$$[\gamma^{\mu}(i\partial_{\mu} + eA_{\mu}) - m]\psi^{c} = 0$$

The only difference between the equations satisfied by ψ and ψ^c is the sign of the charge multiplying the vector potential. Whence ψ^c can be considered as the wave function of a particle of the same mass m but opposite charge $-e$ in the same electromagnetic field A_{μ}. This interpretation is consistent with the easily verified equation

$$K_c^{\,2} = I \ .$$

The other type of transformation is time reversal. This operation will be first defined for the electromagnetic field. Consider a classical electromagnetic field $\underset{\sim}{E}$, $\underset{\sim}{B}$. It satisfies Maxwell's equations:

$$\nabla\cdot\underset{\sim}{E} = 4\pi\rho \qquad\qquad \nabla\cdot\underset{\sim}{B} = 0$$

$$\nabla\times\underset{\sim}{B} = \frac{4\pi}{c}\underset{\sim}{J} + \frac{1}{c}\frac{\partial\underset{\sim}{E}}{\partial t} \qquad\qquad \nabla\times\underset{\sim}{E} = -\frac{1}{c}\frac{\partial\underset{\sim}{B}}{\partial t}$$

Let us define new vector fields $\underset{\sim}{\overline{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$ and $\underset{\sim}{\overline{B}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{B}(-t,\underset{\sim}{x})$. If $\underset{\sim}{E}$ and $\underset{\sim}{B}$ are really time dependent then $\underset{\sim}{\overline{E}}$ and $\underset{\sim}{\overline{B}}$ will not satisfy Maxwell's equations because $\nabla\times\underset{\sim}{\overline{E}} = \frac{1}{c}\frac{\partial\underset{\sim}{\overline{B}}}{\partial t} \neq -\frac{1}{c}\frac{\partial\underset{\sim}{\overline{B}}}{\partial t}$. However if we define instead $\underset{\sim}{\hat{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$ and $\underset{\sim}{\hat{B}}(t,\underset{\sim}{x}) = -\underset{\sim}{B}(-t,\underset{\sim}{x})$, then $\underset{\sim}{\hat{E}}$ and $\underset{\sim}{\hat{B}}$ satisfy Maxwell's equations with ρ and $\underset{\sim}{J}$ replaced by $\hat{\rho}(t,\underset{\sim}{x}) \equiv \rho(-t,\underset{\sim}{x})$ and $\underset{\sim}{\hat{J}}(t,\underset{\sim}{x}) \equiv -\underset{\sim}{J}(-t,\underset{\sim}{x})$ respectively.

In practice the new densities $\hat{\rho}$ and $\hat{\underset{\sim}{J}}$ could be obtained by reversing the motion of the charges acting as sources. It is physically clear that the fields resulting from this operation are $\hat{\underset{\sim}{E}}$ and $\hat{\underset{\sim}{B}}$. On the other hand, we cannot think of any physical operation resulting in the fields $\bar{\underset{\sim}{E}}, \bar{\underset{\sim}{B}}$ and this is consistent with Maxwell's equations which say that these fields do not exist.

One is thus led to define the operation of time reversal on an electromagnetic field as the replacements:

$$\underset{\sim}{E}(t,\underset{\sim}{x}) \rightarrow \hat{\underset{\sim}{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$$

$$\underset{\sim}{B}(t,\underset{\sim}{x}) \rightarrow \hat{\underset{\sim}{B}}(t,\underset{\sim}{x}) \equiv -\underset{\sim}{B}(-t,\underset{\sim}{x})$$

As an operation on the vector potential this amounts to:

$$A_\mu(t,\underset{\sim}{x}) \rightarrow \hat{A}_\mu(t,\underset{\sim}{x}) \equiv (A_o(-t,\underset{\sim}{x}), -\underset{\sim}{A}(-t,\underset{\sim}{x}))$$

Let $\psi'(t,\underset{\sim}{x}) \equiv \gamma^5\gamma^o\psi^c(-t,\underset{\sim}{x}) = \gamma^o BK\psi(-t,\underset{\sim}{x})$. By definition $\psi'(t,\underset{\sim}{x})$ is the spinor obtained from ψ by time reversal. It is easy to see that ψ' satisfies the equation

$$[\gamma^\mu(i\partial_\mu - e\hat{A}_\mu) - m]\psi' = 0$$

It thus describes the same particle as ψ but evolving in the electromagnetic field $\hat{A}_\mu$ obtained from A_μ by time reversal.

In the next section we come back to the group $S^\uparrow$ and take some time to obtain a global concrete picture of it.

(4) A concrete picture of the group $S^{\dagger}$

The first proposition describes the connected pieces of $S^{\dagger}$.

Proposition (5.4): The group $S_{+}^{\dagger}$ is connected. $S^{\dagger}$ has two components: $S_{+}^{\dagger}$ and $\gamma^{o}S_{+}^{\dagger} \equiv \{\gamma^{o}\Lambda;\ \Lambda\varepsilon S_{+}^{\dagger}\}$.

Proof: From eq. (5.15) we know that there is a continuous curve in $S_{+}^{\dagger}$ joining I to -I. Now let $\Omega \varepsilon L_{+}^{\dagger}$. There are two solutions, Λ and $-\Lambda$, to $\Pi(\Lambda) = \Omega$. From eq. (5.22) we know that at least one of them, say Λ, can be reached by a continuous curve lying in $S_{+}^{\dagger}$ and starting at I. But since there is a curve connecting I to -I in $S_{+}^{\dagger}$, it is clear that there is a curve lying in $S_{+}^{\dagger}$ connecting Λ to $-\Lambda$ (one simply takes the one connecting I to -I and multiplies it by Λ). Whence any element in $S_{+}^{\dagger}$ can be connected to I by a continuous curve lying in $S_{+}^{\dagger}$ and $S_{+}^{\dagger}$ is connected. We noticed after theorem (5.2) that $S^{\dagger}$ is disconnected. We have $S^{\dagger} = (S_{+}^{\dagger}) \cup (\gamma^{o}S_{+}^{\dagger})$ and both $S_{+}^{\dagger}$ and $\gamma^{o}S_{+}^{\dagger}$ are connected. Since they don't intersect it is clear that they are the two components of $S^{\dagger}$.

Proposition (5.5): Any Λ in $S^{\dagger}$ has determinant 1.

Proof: Eq. (5.4b) shows that $\det\Lambda = \underline{+}1$. Since by proposition (5.4) $S_{+}^{\dagger}$ is connected and $\det I = 1$ it follows that any Λ in $S_{+}^{\dagger}$ has determinant 1. If Λ is in $S^{\dagger}$, and not in $S_{+}^{\dagger}$, then $\Lambda = \gamma^{o}\Lambda_{+}$ where $\Lambda_{+} \varepsilon S_{+}^{\dagger}$.

whence $\det\Lambda = \det\gamma^{o}\,\det\Lambda_{+} = 1\times 1 = 1$, by the corollary of theorem (2.9).

After proposition (5.3) we said that we would give a third formulation of eqs. (5.3a,b,c). It is contained in the following proposition.

Proposition (5.6): The equations (5.3a,b,c) defining $S^{\uparrow}$ are equivalent to

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \quad , \quad \Omega \in L^{\uparrow}$$

$$\det\Lambda = 1$$

$$\mathrm{Tr}\Lambda^{*} = \mathrm{Tr}\Lambda$$

Proof: If Λ is a solution of (5.3a,b,c) it follows that it satisfies the first two conditions by proposition (5.5). The third condition that $\mathrm{Tr}\Lambda$ be real follows from (5.3c): $\mathrm{Tr}\Lambda^{*} = \mathrm{Tr}B^{+}\Lambda B = \mathrm{Tr}\Lambda BB^{+} = \mathrm{Tr}\Lambda$. Now let us see that the three conditions imply eqs. (5.3a, b,c). The first condition determines Λ up to an arbitrary complex multiplicative factor: if $\bar{\Lambda}$ is one of the two solutions to (5.3a,b,c) then $\Lambda = c\bar{\Lambda}$. Now the second condition above says that c is either ± 1 or $\pm i$, and the third condition eliminates the $\pm i$ possibility. Whence $\Lambda = \pm\bar{\Lambda}$. This completes the proof of the equivalence.

We now come to the main result of this section which provides an identification of $S_+^\uparrow$ with a surprisingly simple group.

Proposition (5.7): The group $S_+^\uparrow$ is isomorphic to the group SL(2,C) of complex 2×2 matrices with determinant 1.

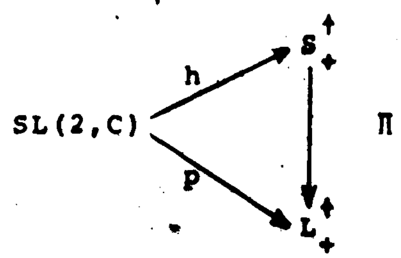
Before we prove this, it is necessary to make a few comments about SL(2,C) and its relation to $L_+^\uparrow$. SL(2,C) shares with $S_+^\uparrow$ the property that there exists a homomorphism P from it onto $L_+^\uparrow$ such that $P(M) = P(N)$ if and only if $M = \pm N$. This P is constructed as follows. One establishes a one to one linear correspondence between the 4-dimensional Minkowski space M and the real vector space H of 2×2 complex hermitian matrices by $x^\mu \to x^\mu \sigma_\mu$, where the σ_μ's are the Pauli matrices of eq. (2.8). Upon setting $\bar{x} \equiv x^\mu \sigma_\mu$, one easily checks that $\det\bar{x} = x^{o^2} - \underset{\sim}{x} \cdot \underset{\sim}{x}$. An element a of SL(2,c) induces a linear map $\tilde{a}: H \to H : h \to aha^+$. Now if $\tilde{a}(\bar{x}) = \bar{x}' = x'^\mu \sigma_\mu$ one has

$$(x'^o)^2 - \underset{\sim}{x}' \cdot \underset{\sim}{x}' = \det(\bar{x}') = \det(a\bar{x}a^+) = \det\bar{x} = x^{o^2} - \underset{\sim}{x} \cdot \underset{\sim}{x},$$

so that a determines a Lorentz transformation Ω_a. One then shows that $P: a \to \Omega_a$ is a homomorphism of SL(2,C) into $L_+^\uparrow$, that it is surjective and that $P(a) = P(b)$ if and only if $a = \pm b$. The details of this can be found in (Ruhl [1970]).

Suppose that we focus our attention on a small enough neighborhood V of I in $S_+^\uparrow$. To each Λ in V there will correspond a unique a_Λ in SL(2,C) close to the identity such that $\Pi(\Lambda) = P(a_\Lambda)$. If we take Λ' close to Λ there will be a unique $a_{\Lambda'}$ close to a_Λ with $\Pi(\Lambda') = P(a_{\Lambda'})$. By moving by little steps in $S_+^\uparrow$ we can extend the map $\Lambda \to a_\Lambda$ to the whole of $S_+^\uparrow$ and this map will be an isomorphism. All we have to do now is to make this intuitive argument rigorous. This is the only place in this work where we invoke mathematical notions which are not completely elementary. It is important to notice that since SL(2,C) is simply connected, the above discussion shows that it is also the universal covering group of $L_+^\uparrow$. Our theorem (5.2) together with the now established connectivity of $S_+^\uparrow$ (proposition (5.4)) shows that $S_+^\uparrow$ is a covering group for $L_+^\uparrow$. If we knew that $S_+^\uparrow$ was simply connected we could conclude at once that it is isomorphic to SL(2,C) by invoking the uniqueness (up to isomorphism) of the universal covering group. But we don't want to show directly that $S_+^\uparrow$ is simply connected. Instead we will use a result about topological groups which will enable us to prove the isomorphism without any calculation. By the same token we will have shown that $S_+^\uparrow$ is simply connected.

The theorem we use says the following. Let G be a topological group with universal covering group $(\tilde{G},p)$ ($p: \tilde{G} \to G$ covering homomorphism). Suppose that [illegible] is another covering group for G. Then there exists a unique continuous homomorphism $h:\tilde{G}\to\tilde{G}'$ such that $p'\circ h = p$. Moreover $(\tilde{G},h)$ is a covering group for $\tilde{G}'$. This theorem can be found in (Pichon [1973]). In our case we take $G = L_+^\uparrow$, $(\tilde{G},p) = (SL(2,C),P)$ and $(\tilde{G}',p') = (S_+^\uparrow,\Pi)$. The quoted theorem then says that there exists a unique continuous homomorphism $h:SL(2,C)\to S_+^\uparrow$ such that $\Pi\circ h = P$ as illustrated by the diagram:



Clearly, given any $a\in SL(2,C)$ h gives one of the two Λ's in $S_+^\uparrow$ such that $\Pi(\Lambda) = P(a)$, and this is accomplished in a continuous fashion. Our claim is that this is in fact an isomorphism. The theorem says that $(SL(2,C),h)$ is a covering group for $S_+^\uparrow$, whence h is surjective. So we only need to show that it is injective. The only two elements that h could map to $-I$ are $\pm I$. Since h is surjective and $h(I) = I$ we must have $h(-I) = -I$. But $-I$ is the only element other than I that h could have mapped to I. Whence we have that $h(a) = I$ implies $a = I$ and h is injective. Whence h is an isomorphism.

This section and the previous ones have provided a complete description of the groups of transformations $S_+^\uparrow$ and $S^\uparrow$. We now turn to the construction of tensors from spinors.

CHAPTER VI

TENSORS CONSTRUCTED FROM SPINORS

(1) Introduction

The aim of this chapter is to construct tensors from quadratic combinations of spinors and to study their relationships. These tensors are well known so we are not going to define anything new. However the section devoted to the study of their relationships might have some originality. In (Pauli [1936]) Pauli has shown how to derive some identities relating these tensors by using the Fierz identity. However the set of identities which he displayed is incomplete in the sense that it doesn't fully express the restrictions on the degrees of freedom in the tensor components. After having discussed the construction of the tensors we will provide a complete solution to the question of their algebraic dependence. In particular we shall give a set of almost independent covariant identities which tells exactly how the various tensors are related to each other. All the other identities are derivable from this particular set.

We must emphasize that our analysis is limited to the case where ψ is an ordinary spinor and not a field operator. We have not seriously investigated how much of the analysis carries through in this more general situation.

(2) Construction of the tensors

In order to obtain real quantities one has to take appropriate linear combinations of the products $\psi^{+}_{\mu}\psi_{\nu}$. The maximum number of linearly independent such combinations is clearly 16. It is in fact possible to construct 16 linearly independent quadratic forms which are all components of tensors or pseudotensors. We list them first and then proceed to show that they have the appropriate transformation laws.

$S \equiv \bar{\psi}\psi$	: scalar
$P \equiv \bar{\psi}\gamma^{5}\psi$	: pseudo-scalar
$V^{\mu} \equiv \bar{\psi}\gamma^{\mu}\psi$	: future pointing timelike or null vector
$P^{\mu} \equiv i\bar{\psi}\gamma^{\mu}\gamma^{5}\psi$	: spacelike or null pseudo-vector
$S^{\mu\nu} \equiv i\bar{\psi}\gamma^{[\mu\nu]}\psi$	: antisymmetric tensor

One easily checks that all these quantities are real. Their linear independence follows from that of the γ^{A}'s. From the condition (5.3b), $\bar{\psi} \equiv \psi^{+}\gamma^{o}$ transforms under the action of $L^{\uparrow}$ according to $\bar{\psi}' = \bar{\psi}\Lambda^{-1}$. From the proof of proposition (5.1), the condition (5.3b) can only be satisfied when dealing with orthochronous Lorentz transformations. Accordingly our quantities behave as claimed only under this type of transformations.

It is trivial to check that S and V^{μ} are a scalar and a vector. Now let us see that $S^{\mu\nu}$ is an antisymmetric tensor. The antisymmetry follows at once from that of $\gamma^{[\mu\nu]}$. We have $S'^{\mu\nu} = i\bar{\psi}'\gamma^{[\mu\nu]}\psi' = i\bar{\psi}\Lambda^{-1}\gamma^{[\mu\nu]}\Lambda\psi$. Clearly if $\mu \neq \nu$ $S'^{\mu\nu} = 0$. Suppose now that $\mu \neq \nu$. Then:

$$\begin{aligned} S'^{\mu\nu} &= i\bar{\psi}\Lambda^{-1}\gamma^{\mu}\Lambda\Lambda^{-1}\gamma^{\nu}\Lambda\psi \\ &= i\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{\rho}\gamma^{\sigma}\psi \\ &= i\sum_{\rho\neq\sigma}\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{[\rho\sigma]}\psi + i\sum_{\rho=\sigma}\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{\rho}\gamma^{\sigma}\psi \\ &= \Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}T^{\rho\sigma} + i\bar{\psi}(-\Omega^{\mu}{}_{i}\Omega^{\nu}{}_{i} + \Omega^{\mu}{}_{o}\Omega^{\nu}{}_{o})\psi \end{aligned}$$

The second term vanishes because we assumed $\mu\neq\nu$ and Ω is a Lorentz transformation. To treat the other two cases we need to use permutations. If $\gamma'^{\mu} \equiv \Omega^{\mu}{}_{\nu}\gamma^{\nu}$, then

$$\gamma'^{5} = \gamma'^{0}\gamma'^{1}\gamma'^{2}\gamma'^{3} = \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\gamma'^{\sigma_0}\gamma'^{\sigma_1}\gamma'^{\sigma_2}\gamma'^{\sigma_3}$$

where S_4 is the set of permutations of (0123).

$$= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\Omega^{\sigma_0}{}_{\rho_0}\Omega^{\sigma_1}{}_{\rho_1}\Omega^{\sigma_2}{}_{\rho_2}\Omega^{\sigma_3}{}_{\rho_3}\gamma^{\rho_0}\gamma^{\rho_1}\gamma^{\rho_2}\gamma^{\rho_3}$$

Clearly the only terms contributing to this sum are those for which all the ρ_{ν}'s are different. So we may write

$$\begin{aligned} \gamma'^{5} &= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\sum_{\tau\in S_4}(-1)^{\tau}\,\Omega^{\sigma_0}{}_{\tau_0}\cdots\Omega^{\sigma_3}{}_{\tau_3}\gamma^{5} \\ &= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}(-1)^{\sigma}\det(\Omega^{\mu}{}_{\nu})\gamma^{5} = \det(\Omega^{\mu}{}_{\nu})\gamma^{5} \end{aligned}$$

From this we obtain:

$$P'^{\mu} = \bar{\psi}\Lambda^{-1}\gamma^{\mu}\Lambda\Lambda^{-1}\gamma^{5}\Lambda\psi = \Omega^{\mu}{}_{\nu}\bar{\psi}\gamma^{\nu}\hat{\gamma}^{5}\psi = \det(\Omega)\Omega^{\mu}{}_{\nu}P^{\nu} \quad .$$

Similarly $P' = \det(\Omega)P$.

The physical interpretation of the vector V^{μ} is the probability current of the electron and it is denoted by j^{μ}. $j^{o} = \psi^{+}\psi$ is indeed positive and can thus represent a probability density. The conservation of total probability is guaranteed, under appropriate boundary conditions at infinity, because j^{μ} satisfies the continuity equation. Indeed upon multiplying $(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0$ on the left by $\bar{\psi}$ and the adjoint equation $\bar{\psi}(i\gamma^{\mu}\overleftarrow{\partial}_{\mu} + m) = 0$ on the right by ψ and adding one obtains $\partial_{\mu}j^{\mu} = 0$. This also works when an external electromagnetic field is present.

The tensor $-2S^{\mu\nu}$ is interpreted as a spin density (Messiah [1964]). The other tensors are used to couple the Dirac field with other fields.

(3) <u>Covariant identities</u>

Let us now come to the identities connecting the tensors. We <u>fix</u> the space-time point x and look at how the components of the tensors are related to each other at that point. They are 16 real-valued functions of $\psi(x)$. (We emphasize again that in this study ψ is an ordinary spinor and not a field operator.) Since $\psi(x)$ contains 8 independent real parameters, the components of the tensors can be considered as 16 real functions of 8 real variables. It is therefore clear that at most 8 of them can be independent. It turns out that only seven are independent and the nine others are determined by the first seven. The basic tool for studying these functions in a way which is independent of a particular choice of unitary γ^{μ}'s is the Fierz identity. It is a very remarkable identity satisfied by any set of 4×4 γ matrices. Let $\{\gamma^{\mu}\}$ be such a set. By proposition (2.5) we know that the 16 γ^{A}'s span the entire vector space of complex 4×4 matrices. If M is an arbitrary 4×4 matrix there exist coefficients α_A such that $M = \sum_A \alpha_A \gamma^A$. Multiplying this by $\gamma_B \equiv (\gamma^B)^{-1}$ and taking the trace yields $\alpha_B = \frac{1}{4} T_r(M\gamma_B)$. Whence we have:

$$M_{\alpha\beta} = \frac{1}{4} \sum_A \gamma^A{}_{\alpha\beta} (\gamma_A)_{\nu\mu} M_{\mu\nu}$$

But, since M is arbitrary, this implies

$$\frac{1}{4} \sum_A \gamma^A{}_{\alpha\beta} (\gamma_A)_{\nu\mu} = \delta_{\alpha\mu} \delta_{\beta\nu} \quad (6.1)$$

This is the Fierz identity. It is in fact a tensor product identity and is seen most clearly when written as such. The tensor product of two linear operators M and N, $M\otimes N$, is defined by $M\otimes N(u\otimes v) = Mu\otimes Nv$ for arbitrary vectors u,v. Given $M\otimes N$ we define $(M\otimes N)^{\pi}$ by $(M\otimes N)^{\pi}(u\otimes v) = Nv\otimes Mu$, so that

$(M \otimes N)^{\pi} = (M \otimes N)(I \otimes I)^{\pi}$. It is then a trivial matter to check that the Fierz identity may be rewritten as:

$$\frac{1}{4} \sum_A \gamma^A \otimes \gamma_A = (I \otimes I)^{\pi} \qquad (6.2)$$

The advantage of this notation is that many equations become clearer because of the elimination of the indices. There is another algebraic tool which we shall need. If $\{\gamma^\mu\}$ is a system of γ matrices, so is $\{\gamma^{\mu^T}\}$, the set of transposes. By the fundamental theorem, there exists an invertible matrix T such that $\gamma^{\mu^T} = T\gamma^\mu T^{-1}$. This matrix T is the other tool which we will use. It is anti-symmetric. Indeed by taking the transpose of the equation defining T one obtains:

$$\gamma^\mu = T^{-1^T} \gamma^{\mu^T} T^T = T^{-1^T} T^{-1} \gamma^\mu T T^T$$

from which $[\gamma^\mu, T^T T^{-1}] = 0$. By theorem (2.8) it follows that $T^T T^{-1} = cI$ or $T^T = cT$. From this we obtain $T = cT^T = c^2 T$, whence $c = \pm 1$. Thus T is either symmetric or antisymmetric. Suppose that it was symmetric. Then one can easily check that the ten matrices $T\gamma^{[\mu\nu]}$ and $T\gamma^{[\lambda\mu\nu]}$ would be antisymmetric. But, since there are at most six linearly independent 4×4 antisymmetric matrices, this is a contradiction; therefore, T is antisymmetric. This elegant argument is taken from (Pauli [1936]).

We are now prepared to study the identities satisfied by the tensors. The first step is to write the Fierz identity in five different ways:

$$4(I\otimes I)^{\pi} = I\otimes I+\gamma^{\mu}\otimes\gamma_{\mu}-\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}+\gamma^{\mu}\gamma^{5}\otimes\gamma_{\mu}\gamma^{5}+\gamma^{5}\otimes\gamma_{5} \qquad (6.3)$$

$$4(\gamma^{5}\otimes\gamma_{5})^{\pi} = I\otimes I-\gamma^{\mu}\otimes\gamma_{\mu}-\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}-\gamma^{\mu}\gamma^{5}\otimes\gamma_{\mu}\gamma^{5}+\gamma^{5}\otimes\gamma_{5} \qquad (6.4)$$

$$4(\gamma^{5}\otimes I)^{\pi} = \gamma^{5}\otimes I+I\otimes\gamma^{5}+\gamma^{\mu}\otimes\gamma_{\mu}\gamma^{5}-\gamma_{\mu}\gamma^{5}\otimes\gamma^{\mu}+\frac{1}{2}\tilde{\gamma}^{[\mu\nu]}\otimes\gamma_{[\mu\nu]} \qquad (6.5)$$

$$4(I\otimes\gamma^{5})^{\pi} = \gamma^{5}\otimes I+I\otimes\gamma^{5}-\gamma^{\mu}\otimes\gamma_{\mu}\gamma^{5}+\gamma_{\mu}\gamma^{5}\otimes\gamma^{\mu}+\frac{1}{2}\tilde{\gamma}^{[\mu\nu]}\otimes\gamma_{[\mu\nu]} \qquad (6.6)$$

$$4(T^{-1}\otimes T)^{\pi} = I\otimes I+\gamma^{\mu}\otimes\gamma_{\mu}^{T}+\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}^{T}-\gamma^{\mu}\gamma^{5}\otimes(\gamma_{\mu}\gamma^{5})^{T}+\gamma^{5}\otimes\gamma_{5}^{T} \qquad (6.7)$$

The matrix $\tilde{\gamma}_{[\mu\nu]}$ is defined by $\tilde{\gamma}_{[\mu\nu]} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}\gamma^{[\alpha\beta]}$. This corresponds to the definition of the dual of an antisymmetric tensor $T_{\mu\nu}$ as $\tilde{T}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}T^{\alpha\beta}$. From eq. (2.5) we have: $\tilde{\gamma}_{[\mu\nu]} = \gamma_{5}\gamma_{[\mu\nu]}$, where we recall that $\gamma_5 \equiv (\gamma^5)^{-1} = -\gamma^5$. Eq. (6.3) is just (6.2) written out explicitely. Eq. (6.4) is obtained by multiplying (6.3) on the left by $\gamma^{5}\otimes\gamma_{5}$. Eq. (6.5) is obtained by multiplying (6.3) on the left by $\gamma^{5}\otimes I$. Eq. (6.6) is obtained by reversing all tensor products in (6.5); this can be achieved by multiplying on right and left by $(I\otimes I)^{\pi}$. Finally eq. (6.7) is obtained by multiplying (6.3) on the left by $I\otimes T$ and on the right by $I\otimes T^{-1}$.

For a while it will be convenient to replace P^{μ} and $S^{\mu\nu}$ by the pure imaginary quantities $\tilde{V}^{\mu} \equiv -iP^{\mu} = \overline{\psi}\gamma^{\mu}\gamma^{5}\psi$, and $T^{\mu\nu} \equiv -iS^{\mu\nu} = \overline{\psi}\gamma^{[\mu\nu]}\psi$.

By multiplying (6.3) on the left by $\overline{\psi}\otimes\overline{\psi}$ and on the right by $\psi\otimes\psi$ we obtain:

$$4S^2 = S^2 + V^{\mu}V_{\mu} - \frac{1}{2}T^{\mu\nu}T_{\mu\nu} + \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

By doing the same thing with (6.4) we get:

$$-4P^2 = S^2 - V^{\mu}V_{\mu} - \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

By adding and substracting these two equations we derive the equivalent system:

$$T^{\mu\nu}T_{\mu\nu} = 2(P^2-S^2) \qquad (6.8)$$

$$V^{\mu}V_{\mu} + \tilde{V}^{\mu}\tilde{V}_{\mu} = 2(P^2+S^2) \qquad (6.9)$$

Performing the same operations on (6.5) yields:

$$\tilde{T}^{\mu\nu}T_{\mu\nu} = 4PS \qquad (6.10)$$

Next if we multiply (6.7) on the left by $\overline{\psi}\otimes\psi^{T}$ and on the right by $\psi\otimes\overline{\psi}^{T}$ we obtain:

$$4(\overline{\psi}T^{-1}\overline{\psi}^{T})(\psi^{T}T\psi) = S^2 + V^{\mu}V_{\mu} + \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

where we have used the fact that an expression like $\psi^{T}\gamma_{\mu}^{T}\overline{\psi}^{T}$, being a number, is equal to its transpose $\overline{\psi}\gamma_{\mu}\psi$. Now since

T is antisymmetric we have $\psi^T T\psi = (\psi^T T\psi)^T = -\psi^T T\psi = 0$. Whence our last equation reduces to

$$S^2 + V^\mu V_\mu + \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^\mu\tilde{V}_\mu - P^2 = 0$$

When combined with (6.8) and (6.9), this gives

$$V^\mu V_\mu = \tilde{V}^\mu\tilde{V}_\mu = P^2+S^2 \qquad (6.11)$$

Upon substracting (6.5) from (6.6) and then multiplying on the left by $I\otimes T$ and on the right by $I\otimes T^{-1}$ we get

$$2[T^{-1}\otimes T\gamma^5 - \gamma^5 T^{-1}\otimes T]^\pi = \gamma_\mu\otimes(\gamma_\mu\gamma^5)^T + \gamma_\mu\gamma^5\otimes\gamma^{\mu^T}$$

Then by treating this as we treated eq. (6.7) we obtain:

$$V^\mu\tilde{V}_\mu = 0 \qquad (6.12)$$

Eqs. (6.8), (6.10), (6.11) and (6.12) are the identities which Pauli displayed. We now proceed to derive others.

By taking the difference between (6.3) and (6.4) we obtain

$$2[I\otimes I - \gamma^5\otimes\gamma_5]^\pi = \gamma^\rho\otimes\gamma_\rho + \gamma^\rho\gamma^5\otimes\gamma_\rho\gamma^5 \qquad (6.13)$$

Multiplying this on the left by $I\otimes\gamma_\mu$ yields

$$2[I\otimes\gamma_\mu + \gamma^5\otimes\gamma_\mu\gamma^5]^\pi = \gamma^\rho\otimes\gamma_\mu\gamma_\rho + \gamma^\rho\gamma^5\otimes\gamma_\mu\gamma_\rho\gamma^5$$

$$= \gamma_\mu\otimes I + \gamma^\rho\otimes\gamma_{[\mu\rho]} + \gamma_\mu\gamma^5\otimes\gamma^5 + \gamma^\rho\gamma^5 \ldots$$

By applying $\bar{\psi}\otimes\bar{\psi}$ to the left and $\psi\otimes\psi$ to the right we get:

$$2[SV_\mu + P\tilde{V}_\mu] = SV_\mu + P\tilde{V}_\mu + T_{\mu\rho}V^\rho - \tilde{T}_{\mu\rho}\tilde{V}^\rho \qquad (6.14)$$

Next by taking the sum of (6.3) and (6.4) we get:

$$2[I\otimes I - \gamma^5\otimes\gamma^5]^\pi = I\otimes I - \gamma^5\otimes\gamma^5 - \frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}$$

Multiplying this on the left by $I\otimes\gamma_\mu$ yields:

$$2[I\otimes\gamma_\mu - \gamma^5\otimes\gamma_\mu\gamma^5]^\pi = I\otimes\gamma_\mu - \gamma^5\otimes\gamma_\mu\gamma^5 - \frac{1}{2}\gamma^{[\alpha\beta]}\otimes\gamma_\mu\gamma_{[\alpha\beta]}$$

Now, by using the product rule (2.2), this may be rewritten as

$$2[I\otimes\gamma_\mu - \gamma^5\otimes\gamma_\mu\gamma^5]^\pi = I\otimes\gamma_\mu - \gamma^5\otimes\gamma_\mu\gamma^5 - \tilde{\gamma}_{[\mu\nu]}\otimes\gamma^\nu\gamma^5 - g_{\mu\alpha}\gamma^{[\alpha\beta]}\otimes\gamma_\beta$$

which, when taken between $\bar{\psi}\otimes\bar{\psi}$ and $\psi\otimes\psi$, gives

$$2[SV_\mu - P\tilde{V}_\mu] = SV_\mu - P\tilde{V}_\mu - \tilde{T}_{\mu\nu}\tilde{V}^\nu - T_{\mu\beta}V^\beta$$

If we combine this with eq. (6.14) we obtain the equivalent system:

$$T_{\mu\rho}V^\rho = P\tilde{V}_\mu \qquad (6.15)$$

$$\tilde{T}_{\mu\rho}\tilde{V}^\rho = -SV_\mu \qquad (6.16)$$

We consider again eq. (6.13) and multiply it on the right by $\gamma^\mu\otimes\gamma_\rho$; this yields:

$$2[\gamma_\rho\otimes\gamma^\mu - \gamma_\rho\gamma^5\otimes\gamma^\mu\gamma_5]^\pi = \gamma^\nu\gamma^\mu\otimes\gamma_\nu\gamma_\rho + \gamma^\nu\gamma^\mu\gamma^5\otimes\gamma_\nu\gamma_\rho\gamma^5$$

$$= \delta^\mu{}_\rho[I\otimes I+\gamma^5\otimes\gamma^5] + \gamma^{[\nu\mu]}\otimes\gamma_{[\nu\rho]} + \gamma^{[\nu\mu]}\gamma^5\otimes\gamma_{[\nu\rho]}\gamma^5$$

$$+ I\otimes\gamma^{[\mu}\gamma_{\rho]} + \gamma_{[\rho}\gamma^{\mu]}\otimes I + \gamma^5\otimes\gamma^{[\mu}\gamma_{\rho]}\gamma^5 + \gamma_{[\rho}\gamma^{\mu]}\gamma^5\otimes\gamma^5$$

Operating on this with $\bar\psi\otimes\bar\psi$ and $\psi\otimes\psi$, in the now familiar way, we obtain:

$$T^{\mu\nu}T_\nu{}^\rho + \tilde T^{\mu\nu}\tilde T_\nu{}^\rho = g^{\mu\rho}(S^2+P^2) - 2(V^\mu V^\rho + \tilde V^\mu\tilde V^\rho) \qquad (6.17)$$

As long as $T_{\mu\nu}$ is antisymmetric, the following identities hold true:

$$T^{\mu\nu}T_{\nu\rho} - \tilde T^{\mu\nu}\tilde T_{\nu\rho} = -\frac{1}{2}(T^{\alpha\beta}T_{\alpha\beta})\delta^\mu{}_\rho \qquad (6.18)$$

$$T^{\mu\nu}\tilde T_{\nu\rho} = -\frac{1}{4}(T^{\alpha\beta}\tilde T_{\alpha\beta})\delta^\mu{}_\rho \qquad (6.19)$$

Combining (6.18) with (6.8) and (6.17) yields

$$T^{\mu\nu}T_\nu{}^\rho = -(V^\mu V^\rho + \tilde V^\mu\tilde V^\rho) + g^{\mu\rho}S^2 \qquad (6.20)$$

$$\tilde T^{\mu\nu}\tilde T_\nu{}^\rho = -(V^\mu V^\rho + \tilde V^\mu\tilde V^\rho) + g^{\mu\rho}P^2 \qquad (6.21)$$

Fortunately we have now nearly exhausted the set of all possible invariant quadratic identities! There remain only two. Multiplying (6.20) by V^ρ and using (6.11), (6.12) and (6.15) we obtain:

$$PT^{\mu\nu}\tilde V_\nu = -(S^2+P^2)V^\mu + S^2V^\mu = -P^2V^\mu$$

Since this is true irrespective of ψ we may divide by P to get

$$T^{\mu\nu}\tilde{V}_{\nu} = -PV^{\mu} \tag{6.22}$$

Next we use this together with (6.19) and (6.10) to get:

$$\tilde{T}^{\mu\nu}V_{\nu} = -\frac{1}{P}\,\tilde{T}^{\mu\nu}T_{\nu\rho}\tilde{V}^{\rho} = -\frac{1}{P}\left(-\frac{1}{4}\,\tilde{T}^{\alpha\beta}T_{\alpha\beta}\right)\tilde{V}^{\mu} = S\tilde{V}^{\mu},$$

that is:

$$\tilde{T}^{\mu\nu}V_{\nu} = S\tilde{V}^{\mu} \tag{6.23}$$

We now reexpress all the identities obtained in terms of our original real-valued functions P^{μ} and $S^{\mu\nu}$.

$$V^{\mu}P_{\mu} = 0 \tag{6.24}$$

$$V^{\mu}V_{\mu} = P^2+S^2 \tag{6.25}$$

$$P^{\mu}P_{\mu} = -(P^2+S^2) \tag{6.26}$$

$$S^{\mu\nu}S_{\mu\nu} = 2(S^2-P^2) \tag{6.27}$$

$$\tilde{S}^{\mu\nu}S_{\mu\nu} = -4PS \tag{6.28}$$

$$S^{\mu\nu}S_{\nu\rho} = (V^{\mu}V^{\rho} - P^{\mu}P^{\rho}) - g^{\mu\rho}S^2 \tag{6.29}$$

$$\tilde{S}^{\mu\nu}\tilde{S}_{\nu\rho} = (V^{\mu}V^{\rho} - P^{\mu}P^{\rho}) - g^{\mu\rho}P^2 \tag{6.30}$$

$$S^{\mu\rho}\tilde{S}_{\rho\sigma} = PS\,\delta^{\mu}_{\sigma} \tag{6.31}$$

$$S_{\mu\rho}V^{\rho} = PP_{\mu} \tag{6.32}$$

$$\tilde{S}_{\mu\rho}V^{\rho} = SP_{\mu} \tag{6.33}$$

$$S_{\mu\rho}P^{\rho} = PV_{\mu} \tag{6.34}$$

$$\tilde{S}_{\mu\rho}P^{\rho} = SV_{\mu} \tag{6.35}$$

(4) The information contained in the identities

Clearly the left hand sides of the last set of equations comprise all possible quadratic invariant combinations. But these equations are so numerous that their content is far from clear. They form a highly redundant system: for instance, (6.24) is an obvious consequence of any one of the last four equations. In fact only nine equations are independent. This leaves seven independent functions: one possible choice is P, V^k, P^k. We shall now show how one can pick a system of almost independent identities implying all the others and then proceed to demonstrate that P, V^k, P^k are independent. First we fix the notation. We define:

$$\underset{\sim}{V} = \begin{bmatrix} V^1 \\ V^2 \\ V^3 \end{bmatrix} \qquad \underset{\sim}{P} = \begin{bmatrix} P^1 \\ P^2 \\ P^3 \end{bmatrix} \tag{6.36}$$

$$\underset{\sim}{S} = \begin{bmatrix} S_{23} \\ S_{31} \\ S_{12} \end{bmatrix} \qquad \underset{\sim}{K} = \begin{bmatrix} S_{10} \\ S_{20} \\ S_{30} \end{bmatrix} \tag{6.37}$$

so that

$$(S_{\mu\nu}) = \begin{bmatrix} 0 & -\underset{\sim}{K}^T \\ \underset{\sim}{K} & \omega(\underset{\sim}{S}) \end{bmatrix} \qquad (\tilde{S}_{\mu\nu}) = \begin{bmatrix} 0 & \underset{\sim}{S}^T \\ -\underset{\sim}{S} & \omega(\underset{\sim}{K}) \end{bmatrix} \tag{6.38}$$

where

$$\omega(\underset{\sim}{U}) \equiv \begin{bmatrix} 0 & U^3 & -U^2 \\ -U^3 & 0 & U^1 \\ U^2 & -U^1 & 0 \end{bmatrix}$$

$\mathbf{S}$ is twice the spin density vector. One easily establishes that

$$\mathbf{K}^2 - \mathbf{S}^2 = -\frac{1}{2} S^{\alpha\beta} S_{\alpha\beta} \qquad (6.39)$$

$$\mathbf{K} \cdot \mathbf{S} = \frac{1}{4} S^{\alpha\beta} \tilde{S}_{\alpha\beta} \qquad (6.40)$$

These are the analogues of the familiar invariants of the electromagnetic field.

Our claim is that the following set of identities is complete:

$$S_{\mu\rho} V^{\rho} = p P_{\mu} \qquad (6.41)$$

$$\tilde{S}_{\mu\rho} V^{\rho} = s P_{\mu} \qquad (6.42)$$

$$V^{\mu} V_{\mu} = s^2 + p^2 = -P^{\mu} P_{\mu} \qquad (6.43)$$

First (6.41) clearly implies (6.24). Now in our new notation, eqs. (6.41) and (6.42) read

$$p \begin{bmatrix} P^0 \\ -\mathbf{P} \end{bmatrix} = \begin{bmatrix} -\mathbf{K} \cdot \mathbf{V} \\ V^0 \mathbf{K} + \mathbf{V} \times \mathbf{S} \end{bmatrix} \qquad s \begin{bmatrix} P^0 \\ -\mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{S} \cdot \mathbf{V} \\ -V^0 \mathbf{S} + \mathbf{V} \times \mathbf{K} \end{bmatrix} \qquad (6.44)$$

The second equation gives

$$\mathbf{S} = \frac{1}{V^0} (\mathbf{V} \times \mathbf{K} + s\mathbf{P})$$

from which

$$\mathbf{V} \times \mathbf{S} = \frac{1}{V^0} \left((\mathbf{V} \cdot \mathbf{K}) \mathbf{V} - \mathbf{V}^2 \mathbf{K} + s \mathbf{V} \times \mathbf{P} \right)$$

Inserting this into the first part of (6.44) we obtain

$$V^{o2}\mathbf{K} - PP^{o}\mathbf{V} - \mathbf{V}^{2}\mathbf{K} + S\mathbf{V}\times\mathbf{P} = -V^{o}P\mathbf{P}$$

that is:
$$\mathbf{K} = \frac{1}{V^{\mu}V_{\mu}}\left[P(P^{o}\mathbf{V} - V^{o}\mathbf{P}) - S\mathbf{V}\times\mathbf{P}\right]$$

From this we get:

$$\mathbf{V}\times\mathbf{K} = \frac{1}{V^{\mu}V_{\mu}}\left[-PV^{o}\mathbf{V}\times\mathbf{P} - S(\mathbf{V}\cdot\mathbf{P})\mathbf{V} + S\mathbf{V}^{2}\mathbf{P}\right]$$

Inserting this into the second part of (6.44) yields

$$\mathbf{S} = \frac{1}{V^{\mu}V_{\mu}}\left[-P\mathbf{V}\times\mathbf{P} - \frac{\mathbf{S}}{V^{o}}(\mathbf{V}\cdot\mathbf{P})\mathbf{V} + \frac{\mathbf{S}}{V^{o}}\mathbf{V}^{2}\mathbf{P} + \frac{V^{\mu}V_{\mu}}{V^{o}}S\mathbf{P}\right]$$

Equation (6.24), which may be written

$$V^{o}P^{o} = \mathbf{V}\cdot\mathbf{P} \quad , \tag{6.45}$$

is a consequence of (6.41). This can be used to cancel the V^{o} in the denominator of the second term on the right hand side of the next to last equation. Thus we obtain the following expressions for $\mathbf{K}$ and $\mathbf{S}$:

$$\mathbf{K} = \frac{1}{V^{\mu}V_{\mu}}\left[P(P^{o}\mathbf{V} - V^{o}\mathbf{P}) - S\mathbf{V}\times\mathbf{P}\right] \tag{6.46}$$

$$\mathbf{S} = \frac{1}{V^{\mu}V_{\mu}}\left[S(V^{o}\mathbf{P} - P^{o}\mathbf{V}) - P\mathbf{V}\times\mathbf{P}\right] \tag{6.47}$$

We haven't yet used the $SP^{o} = \mathbf{S}\cdot\mathbf{V}$ part of eq. (6.44). In fact this is now a consequence of (6.45) and (6.47). This is the extent to which the identities (6.41,42,43) are not completely independent.

Next we obtain

$$SP_{\mu}P^{\mu} = \tilde{S}_{\mu\rho}V^{\rho}P^{\mu} \qquad \text{from (6.42)}$$

$$= \frac{1}{P}\,\tilde{S}_{\mu\rho}S^{\mu\lambda}V_{\lambda}V^{\rho} \qquad \text{from (6.41)}$$

$$= \frac{1}{4P}\,(\tilde{S}_{\alpha\beta}S^{\alpha\beta})V_{\rho}V^{\rho} \qquad \text{from (6.19)}$$

$$= -\frac{1}{4P}\,(\tilde{S}_{\alpha\beta}S^{\alpha\beta})P_{\mu}P^{\mu} \qquad \text{from (6.43)}$$

Whence $\tilde{S}_{\alpha\beta}S^{\alpha\beta} = -4PS$; this is (6.28).

The explicit calculations will not be given but one can deduce, using only (6.45,46,47), the relation

$$S^{\mu\nu}S_{\nu\rho} = \frac{S^2+P^2}{V^{\mu}V_{\mu}}\,(V^{\mu}V_{\rho} - P^{\mu}P_{\rho}) - \delta^{\mu}{}_{\rho}S^2$$

When coupled with (6.43) this equation gives (6.29). It is also seen to imply (6.27) by letting $\rho=\mu$ and using $P^{\mu}P_{\mu} = -V^{\mu}V_{\mu}$. Combining (6.29), (6.27) and the identity (6.18), which is valid for arbitrary antisymmetric tensors, one is led to (6.30). Equation (6.31) follows from the general identity (6.19) and equation (6.28). Finally equations (6.34) and (6.35) are a consequence of (6.41), (6.42) together with the general identity (6.19) and eq. (6.28).

We have shown that the equations (6.41,42,43) imply all the others in (6.24) to (6.35).

Now suppose that $\underset{\sim}{V}, \underset{\sim}{P}$ and P are given. The equations $V^{\mu}V_{\mu} = - P^{\mu}P_{\mu}$ and $V^{\mu}P_{\mu} = 0$ read:

$$V^{o^2} + P^{o^2} = \underset{\sim}{V}^2 + \underset{\sim}{P}^2 \quad .$$

$$V^o P^o = \underset{\sim}{P} \cdot \underset{\sim}{V}$$

By themselves these equations are sufficient to determine V^o and P^o up to sign. But we know a priori that $V^o = \psi^+\psi \geq 0$. If V^o is 0 then $\psi=0$ and everything vanishes. If $V^o \neq 0$ then $V^o > 0$ and the two above equations take care of the signs as well as of the magnitudes of V^o and P^o. Now that V^{μ} is determined, the equation $V^{\mu}V_{\mu} = S^2+P^2$ gives S up to sign, P being supposedly given. Then, unless both S and P are zero, equations (6.46) and (6.47) determine $S_{\mu\nu}$ uniquely. If both S and P vanish, V^{μ} and P^{μ} are two orthogonal null* vectors; this implies that they are linearly dependent. In this case the system becomes degenerate since it fails to determine $S_{\mu\nu}$ uniquely. Indeed one may check by looking at (6.44) that it only demands that $\underset{\sim}{K}$ and $\underset{\sim}{S}$ be two vectors of equal norm orthogonal to each other and orthogonal to $\underset{\sim}{V}$. Whence there is one degree of freedom left: their position in the plane orthogonal to $\underset{\sim}{V}$. Due to the fact that V^{μ} and P^{μ} are linearly dependent, one can easily check that the whole set of equations (6.24) to (6.35) says nothing else about $S_{\mu\nu}$ than what is already implied by (6.46) and (6.47).

* A 4-vector A^{μ} is said to be null if $A^{\mu}A_{\mu} = 0$.

Thus we have shown that, except in the degenerate case where V^{μ} is a null vector, the system (6.41,42,43) determines the other functions once $\underset{\sim}{V}$, $\underset{\sim}{P}$ and P are given. Moreover this system is complete in the sense that it implies all the other equations in (6.24) to (6.35). To show that the system is really algebraically complete we have to demonstrate that there are no non trivial relations between $\underset{\sim}{V}$, $\underset{\sim}{P}$ and P, that is these are independent functions. One way to do this would be to use a particular set of γ matrices, for instance that of eq. (2.7), to write down explicitely these functions. But we have found an alternative method which is less cumbersome and doesn't rely on a special choice of the system $\{\gamma^{\mu}\}$. It amounts to showing that their differentials can be made linearly independent with an appropriate choice of ψ.

The seven functions P, $\underset{\sim}{V}$, $\underset{\sim}{P}$ are of the form $\psi \to \bar{\psi}M^{A}\psi$ where the M^{A}'s are γ^{5}, γ^{k} and $i\gamma^{k}\gamma^{5}$ respectively. Let us denote them by f^{A}. To emphasize their dependence upon eight real variables we write ψ as $\Phi+i\phi$ where both Φ and ϕ are real. Then we have:

$$f^{A}(\Phi,\phi) = (\bar{\Phi}-i\bar{\phi})M^{A}(\Phi+i\phi)$$

$$= \bar{\Phi}M^{A}\Phi + \bar{\phi}M^{A}\phi + i(\bar{\Phi}M^{A}\phi - \bar{\phi}M^{A}\Phi).$$

Now the differential of f^{A} at the point (Φ,ϕ) is defined by

$$f^A(\Phi+h,\phi+k) - f^A(\Phi,\phi) = df^A_{(\Phi,\phi)}(h,k) + \theta(h,k),$$

df^A being linear and $\theta(h,k)$ satisfying $\lim_{(h,k)\to 0} \frac{\theta(h,k)}{\|(h,k)\|} = 0.$

One can easily check that.

$$df^A_{(\Phi,\phi)}(h,k) = (\bar{\Phi}-i\bar{\phi})M^A h+(\bar{\phi}+i\bar{\Phi})M^A k+\bar{h}M^A(\Phi+i\phi)+\bar{k}M^A(\phi-i\Phi)$$

Suppose that these differentials were linearly dependent. That means that there would be real numbers λ_A, not all vanishing, such that

$$\lambda_A df^A_{(\Phi,\phi)}(h,k) = 0 \quad \text{independently of } h \text{ and } k.$$

This would imply the two equations

$$(\bar{\Phi}-i\bar{\phi})Mh + \bar{h}M(\Phi+i\phi) = 0 \quad \text{for all } h \qquad (6.48)$$

$$(\bar{\phi}+i\bar{\Phi})Mk + \bar{k}M(\phi-i\Phi) = 0 \quad \text{for all } k \qquad (6.49)$$

where we have set $\lambda_A M^A \equiv M$. Upon replacing the second term in (6.48) by its transpose, to which it is equal, we obtain

$$(\bar{\Phi}-i\bar{\phi})Mh + (\Phi^T+i\phi^T)M^T\gamma^{o^T}h = 0$$

This being supposedly true for arbitrary h we deduce

$$(\bar{\Phi}-i\bar{\phi})M + (\Phi^T+i\phi^T)M^T\gamma^{o^T} = 0 \qquad (6.50)$$

Doing the same thing with (6.49) and multiplying the resulting equation by -i yields:

$$(\bar{\Phi}-i\bar{\phi})M - (\phi^T+i\phi^T)M^T\gamma^{o^T} = 0 \qquad (6.51)$$

Substracting (6.51) from (6.50) gives $(\phi^T+i\phi^T)M^T\gamma^{o^T} = 0$; this is equivalent to $\gamma^o M\psi = 0$ or $M\psi = 0$. Whence the problem is reduced to finding a ψ such that the seven vectors $M^A\psi$ are linearly independent over the reals. Suppose that we have a relation

$$[\lambda_o\gamma^5 + \lambda_k\gamma^k + i\mu_k\gamma^k\gamma^5]\psi = 0 \qquad (6.52)$$

where all λ's and μ's are real. Let ψ be an eigenvector of γ^o: $\gamma^o\psi = \rho\psi$. ρ is not 0 since γ^o is invertible. We also know that such an eigenvector exists since γ^o can be diagonalized.

Multiplying (6.52) by γ^o and then dividing by ρ we obtain:

$$[-\lambda_o\gamma^5 - \lambda_k\gamma^k + i\mu_k\gamma^k\gamma^5]\psi = 0$$

Together with (6.52) this implies:

$$\mu_k\gamma^k\psi = 0$$

$$[\lambda_o\gamma^5 + \lambda_k\gamma^k]\psi = 0$$

Multiplying the last equation by $\psi^+\gamma^5$ we get:

$$-\lambda_o\psi^+\psi + \lambda_k\psi^+\gamma^5\gamma^k\psi$$

National Library of Canada | Bibliothèque nationale du Canada

CANADIAN THESES ON MICROFICHE | ***THÈSES CANADIENNES SUR MICROFICHE***

NAME OF AUTHOR *NOM DE L'AUTEUR* Jean-François Dumais

TITLE OF THESIS *TITRE DE LA THÈSE* Dirac Spinor Kinematics

UNIVERSITY *UNIVERSITÉ* University of Alberta

DEGREE FOR WHICH THESIS WAS PRESENTED / *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* M. Sc.

YEAR THIS DEGREE CONFERRED / *ANNÉE D'OBTENTION DE CE GRADE* 1975

NAME OF SUPERVISOR / *NOM DU DIRECTEUR DE THÈSE* Professor Yasushi Takahashi

DATED / *DATÉ* august 13/75 SIGNED / *SIGNÉ* Jean-François Dumais

PERMANENT ADDRESS / *RÉSIDENCE FIXE* 11019 84th st, ste #201

Edmonton, Alberta.

NL-91 (3-74)

THE UNIVERSITY OF ALBERTA

DIRAC SPINOR KINEMATICS

by

© JEAN-FRANCOIS DUMAIS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
OF MASTER OF SCIENCE

IN

THEORETICAL PHYSICS

DEPARTMENT OF PHYSICS

EDMONTON, ALBERTA

FALL, 1975

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled DIRAC SPINOR KINEMATICS submitted by Jean-Francois Dumais in partial fulfillment of the requirements for the degree of Master of Science in Theoretical Physics.

Y. Takahashi

Supervisor

M. Razavy

S.K. Sehgal

Date July 7/75

ABSTRACT

The kinematic aspects of Dirac spinors are studied. All the properties of a totally arbitrary system of gamma matrices are derived without resorting to the theory of finite group representations or the theory of Clifford algebras. An algebraic method is then devised to find explicitely the similarity transformation arising in the fundamental theorem of gamma matrices. Next the Lie group of spinor transformations under the action of the orthochronous Lorentz group is studied in detail. The work ends with a thorough analysis of all the algebraic relations among the Dirac bilinears.

PREFACE

As described by the title the object of this simple work is the kinematics of Dirac spinors. Needless to say it was meant as a review. However it is hoped that the final product is not completely devoid of originality.

Chapter I recalls how one is led to the Dirac equation and its associated gamma matrices. The relativistic invariance of the equation, discussed in the second section, provides the physical motivation for the fundamental theorem of gamma matrices.

The first section of chapter II is a standard presentation of the properties of products of gamma matrices. The second section discusses the degree and reducibility of the representations of the fundamental relations $\{\gamma^{\mu},\gamma^{\nu}\} = 2g^{\mu\nu}I$. The eventual originality of the discussion lies in a complete avoidance of the theory of finite group representations or the theory of Clifford algebras. The resulting treatment is self-contained and elementary; this might be of some pedagogical interest.

Chapter III deals with the fundamental theorem of gamma matrices. The usual proof is modified, leading to a shorter and perhaps more elegant one. Some elementary consequences of the fundamental theorem are then discussed; for instance it is shown that there exists no system of real gamma matrices.

While the main result of chapter III asserts the existence of a certain non-singular matrix S connecting two systems of gamma matrices, it says nothing about the explicit form of S. It is the aim of chapter IV to try to fill this gap.

Chapter V discusses in detail the transformation of spinors under orthochronous Lorentz transformations. The Lie group $S^{\uparrow}$ whose elements are those transformations is carefully studied. Several different ways of describing its elements are obtained. It is finally concluded that the subgroup $S_{+}^{\uparrow}$ corresponding to proper Lorentz transformations is isomorphic to SL(2,C).

Chapter VI deals with the tensors obtained by quadratic combinations of spinors. These include a scalar and a pseudo-scalar, a vector and a pseudo-vector and a twice contravariant antisymmetric tensor. These objects are not independent of each other. Covariant identities other than those given in (Pauli [1936]) are derived and used to provide a complete solution to the question of the algebraic dependence of the tensor components. This analysis is restricted to the case where ψ is an ordinary spinor and not a field-operator.

ACKNOWLEDGEMENTS

First I wish to thank my supervisor, Professor Y. Takahashi, for suggesting to me the topic of this thesis and for his availability for discussion.

To the National Research Council of Canada and to the Ministère de l'Education du Québec I am indebted for my financial support over the last two years.

Finally I thank Mrs. Lee Cech for her fast typing of the bulk of my thesis and Mrs. Joanne Koppens for completing the job.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

THE ORIGIN OF DIRAC γ MATRICES

(1) The Dirac equation

The Dirac γ matrices arise naturally when one seeks a relativistic equation for the wave function of an electron. We suppose that this wave function ψ is a function from spacetime into C^N for some N and we are looking for a differential equation describing its behavior. To have a close analogy with the Schroedinger equation we want it to be first order in time. In order to be invariant under Lorentz transformations it will have to be of first order in the space derivatives as well. The most general form of such a linear homogeneous equation with constant coefficients, expressing the time derivative of ψ in terms of the space derivatives of ψ and of ψ itself is clearly:

$$\left(\frac{\partial}{\partial x^o} + \alpha^k \frac{\partial}{\partial x^k} + \frac{imc}{\hbar}\beta\right)\psi = 0 \qquad (1.1)$$

where α^k and β are N×N complex matrices and k runs from 1 to 3. The x^k's are the space coordinates and $x^o = ct$. The constant in front of β takes care of the dimensions appropriately if m is a mass. To be consistent with the relativistic energy-momentum relation $E^2 = p^2c^2 + m^2c^4$ we require that ψ satifies the Klein-Gordon equation as well. Simple multiplications give:

$$\left(\frac{\partial}{\partial x^o} - \alpha^k \frac{\partial}{\partial x^k} - \frac{imc}{\hbar}\beta\right)\left(\frac{\partial}{\partial x^o} + \alpha^k \frac{\partial}{\partial x^k} + \frac{imc}{\hbar}\beta\right)$$

$$= \frac{\partial^2}{\partial x^{o2}} - \frac{1}{2}(\alpha^k\alpha^r + \alpha^r\alpha^k)\frac{\partial}{\partial x^k}\frac{\partial}{\partial x^r} - \frac{imc}{\hbar}(\alpha^k\beta + \beta\alpha^k)\frac{\partial}{\partial x^k} + \frac{m^2c^2}{\hbar^2}\beta^2$$

Upon requiring that

$$\{\alpha^k,\alpha^r\} = -2g^{kr}I,\ \{\alpha^k,\beta\} = 0,\ \beta^2 = I \qquad (1.2)$$

where $\{A,B\} \equiv AB + BA; (g^{\mu\nu}) \equiv \text{diag}\ (1,-1,-1,-1)$, this differential operator reduces to the Klein-Gordon operator $(\frac{\partial^2}{\partial x^{o2}} - \nabla^2 + \frac{m^2c^2}{\hbar^2})$. Whence, when conditions (1.2) are imposed, any solution of (1.1) is a solution of the Klein-Gordon equation.

For the purpose of studying its relativistic invariance, eq. (1.1) is more conveniently written as:

$$(i\gamma^\mu\partial_\mu - \frac{mc}{\hbar})\psi = 0 \qquad (1.3a)$$

where $$\gamma^k \equiv \beta\alpha^k \text{ and } \gamma^o \equiv \beta \qquad (1.3b)$$

The relations (1.2) are then equivalent to

$$\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}I,\ \mu,\nu = 0,1,2,3 \qquad (1.3c)$$

A system of 4 complex square matrices $\{\gamma^\mu\}$ satisfying this last equation will be called a system of γ matrices. Since their square is $\pm I$, γ matrices are non-singular. Another immediate consequence of (1.3c) is that the order N of γ matrices has to be even: $\det \gamma^1\gamma^2 = (-1)^N \det \gamma^2\gamma^1 = (-1)^N \det \gamma^1\gamma^2$, from which $(-1)^N = 1$. More will be said about this later.

(2) The relativistic invariance of the Dirac equation

From now on we will take units in which $c=\hbar=1$. Eq. (1.3a) then reads:

$$(-i\gamma^\mu\partial_\mu + m)\psi = 0 \qquad (1.4)$$

The "interaction" with an external electromagnetic field of 4-potential A^μ is achieved through the so-called minimal electromagnetic coupling in which $p_\mu \equiv i\partial_\mu$ is replaced by $p_\mu - eA_\mu = iD_\mu$, that is $D_\mu = \partial_\mu + ieA_\mu$, $e(<o)$ being the charge of the electron. Whence in the presence of an external electromagnetic field, eq. (1.4) becomes:

$$(-i\gamma^\mu D_\mu + m)\psi = 0 \qquad (1.4')$$

$$D_\mu \equiv \partial_\mu + ieA_\mu$$

We want to find a transformation law for ψ such that eq. (1.4') remains invariant under orthochronous Lorentz transformations. If Ω is such a Lorentz transformation it is assumed that the corresponding transformation for ψ is linear:

$$\psi'(x') = S(\Omega)\psi(x) \qquad \det S \neq 0 \tag{1.5}$$

Putting $y^\rho = \Omega^\rho_{\ \mu} x^\mu$ and rewriting (1.4') in terms of y^μ and ψ' yields:

$$[-i(S\Omega^\rho_{\ \mu}\gamma^\mu S^{-1})\frac{\partial}{\partial y^\rho} + (S\Omega^\rho_{\ \mu}\gamma^\mu S^{-1})eA_\rho + m]\psi' = 0$$

This will be identical in form with (1.4') if and only if

$$\hat{\gamma}^\rho \equiv \Omega^\rho_{\ \mu}\gamma^\mu = S^{-1}\gamma^\rho S \tag{1.6}$$

But it turns out that the $\hat{\gamma}^\rho$'s are also γ matrices:

$$\{\hat{\gamma}^\rho, \hat{\gamma}^\sigma\} = \Omega^\rho_{\ \mu}\Omega^\sigma_{\ \nu}\{\gamma^\mu, \gamma^\nu\} = 2\Omega^\rho_{\ \mu}\Omega^\sigma_{\ \nu} g^{\mu\nu} I = 2g^{\rho\sigma} I .$$

Therefore the invariance of the theory will be guaranteed if we can show that any two sets of γ matrices are related by a similarity transformation as in eq. (1.6). In the coming sections the existence of such a similarity transformation will be proved in a way which, to our knowledge, is original to a certain extent.

The theory ought to be invariant under space-time translations as well. This is achieved by letting simply $\psi'(x') = \psi(x)$ under the translation $x' = x + a$.

Besides equation (1.4) we will sometimes refer to the so-called adjoint equation. The adjoint $\bar{\psi}$ of ψ is defined by

$$\bar{\psi} \quad \psi^{+} \gamma^{0} \tag{1.7}$$

Provided that the γ^{μ}'s are unitary equations (1.4) and (1.4') are easily seen to be equivalent to the following equations for $\bar{\psi}$:

$$\bar{\psi}(i\gamma^{\mu + *}\partial_{\mu} + m) = 0 \tag{1.8}$$

$$\bar{\psi}(i\gamma^{\mu + *}D^{*}_{\mu} + m) = 0, \quad D^{*}_{\mu} \equiv \partial_{\mu} - ieA_{\mu} \tag{1.8'}$$

From this equivalence it is clear that the invariance of equations (1.4), (1.4') implies that of equations (1.8), (1.8').

CHAPTER II

GENERAL PROPERTIES OF γ MATRICES

The aim of this chapter is to investigate the properties of a general system of γ matrices. The first part studies essentially the properties of their products. This is completely standard. The second part studies the degrees and reducibility of all possible representations of the relations $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}I$. Unlike the first part it might be original. This is because we have found a way of treating these questions without using the theory of representations of finite groups or the theory of Clifford algebras. Thus our treatment is self-contained and elementary.

(1) Products of γ^μ's

Let $\{\gamma^\mu\}$ be a system of arbitrary N×N γ matrices. Out of them we construct the following sixteen matrices which will play a great role in our considerations.

Table 1 List of the matrices γ^A

	I				
	γ^0		γ^1	γ^2	γ^3
$\gamma^1\gamma^0$	$\gamma^2\gamma^0$	$\gamma^3\gamma^0$	$\gamma^2\gamma^3$	$\gamma^3\gamma^1$	$\gamma^1\gamma^2$
	$\gamma^1\gamma^2\gamma^3$		$\gamma^0\gamma^2\gamma^3$	$\gamma^0\gamma^3\gamma^1$	$\gamma^0\gamma^1\gamma^2$
				$\gamma^0\gamma^1\gamma^2\gamma^3\equiv\gamma^5$	

We will denote this set of 16 matrices by Γ and refer to its members by the symbol γ^A, $A=1,\ldots,16$. The inverse of γ^A will be denoted by γ_A. Indices on the γ^μ's will be raised and lowered with respect to $(g_{\mu\nu}) = \text{diag}\,(1,-1,-1,-1)$. Notice that the γ^μ's are unitary if and only if $\gamma^{\mu^+} = \gamma_\mu$. We also adopt the following notations:

$$\gamma^{[\mu\nu]} \equiv \begin{cases} \gamma^\mu\, \gamma^\nu & \text{if } \mu \neq \nu \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma^{[\lambda\mu\nu]} \equiv \begin{cases} \gamma^\lambda \gamma^\mu \gamma^\nu & \text{if } \lambda,\mu,\nu \text{ are all different} \\ 0 & \text{otherwise} \end{cases}$$

$$\gamma^{[\mu\nu\lambda\rho]} \equiv \epsilon_{\mu\nu\lambda\rho}\, \gamma^5$$

where $\epsilon_{\mu\nu\lambda\rho}$ is zero if μ,ν,λ,ρ are not all different and is otherwise equal to the sign of the permutation $\begin{pmatrix} 0 & 1 & 2 & 3 \\ \mu & \nu & \lambda & \rho \end{pmatrix}$. Whence $\epsilon_{0123}=1=-\epsilon^{0123}$

$$\sigma^{\mu\nu} \equiv i\, \gamma^{[\mu\nu]}$$

All these quantities are completely antisymmetric with respect to their indices.

Our set Γ has remarkable properties which we now proceed to derive.

Proposition (2.1): The square of any member of Γ is I or -I.

Proof: This is an obvious consequence of the fact that the square of each γ^{μ} is $\pm$I and they all anticommute.

In our table we have arranged the γ^{A}'s in such a way that all those with square +I are on the left, the others on the right.

Proposition (2.2): The product of two members of Γ is, up to a sign, again a member of Γ:

$$\gamma^{A}\gamma^{B} = \varepsilon_{AB}\, \gamma^{C(A,B)}, \quad \varepsilon_{AB} = \pm 1 \tag{2.1}$$

Proof: This is again an obvious consequence of the anticommutation of the γ^{μ}'s and the fact that their square is $\pm$I.

Proposition (2.2'): The functions ε_{AB} and C(A,B) appearing in eq. (2.1) are the same for all systems of γ matrices.

Proof: This is trivally true by construction.

Equation (2.1) together with the apparently innocent proposition (2.2') will be the key to our proof of the fundamental theorem of γ matrices.

Proposition (2.3): All members of Γ other than I have vanishing trace.

Proof: We first prove it for the γ^{μ}'s. Let ν be given and choose $\mu \neq \nu$. From $\gamma^{\mu}\gamma^{\mu} + \gamma^{\mu}\gamma^{\mu} = \pm 2I$ we deduce

$$-\gamma^{\mu}(\gamma^{\nu}\gamma^{\mu}) + (\gamma^{\nu}\gamma^{\mu})\gamma^{\mu} = \pm 2\, \gamma^{\nu}$$

and whence $\pm 2\,\mathrm{Tr}\,\gamma^{\nu} = -\mathrm{Tr}(\gamma^{\mu}(\gamma^{\nu}\gamma^{\mu})) + \mathrm{Tr}[(\gamma^{\nu}\gamma^{\mu})\gamma^{\mu}] = 0$ because in general $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$. This at the same time shows that $\mathrm{Tr}\gamma^{5} = 0$ because $\{\gamma^{\mu},\gamma^{5}\} = 0$.

Next if $\mu \neq \nu$ then $\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = 0$ because $\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu}) = \mathrm{Tr}(-\gamma^{\nu}\gamma^{\mu}) = -\mathrm{Tr}(\gamma^{\nu}\gamma^{\mu}) = -\mathrm{Tr}(\gamma^{\mu}\gamma^{\nu})$. Since $\gamma^{[\lambda\mu\nu]}$ is of the form $\pm\,\gamma^{5}\,\gamma^{\rho}$, the same argument shows that $\mathrm{Tr}\,\gamma^{[\lambda\mu\nu]} = 0$.

<u>Proposition</u> (2.4): If one fixes A in equation (2.1) and then lets B go from 1 to 16, C(A,B) goes over all the values in $\{1,\ldots,16\}$.

Proof: Since $\varepsilon_{AB} = \pm 1$, eq. (2.1) may be rewritten as

$$\gamma^{C(A,B)} = \varepsilon_{AB}\,\gamma^{A}\,\gamma^{B}$$

Whence $C(A,B) = C(A,B')$ implies $\varepsilon_{AB}\,\gamma^{A}\,\gamma^{B} = \varepsilon_{AB'}\,\gamma^{A}\,\gamma^{B'}$, from which $\varepsilon_{AB}\,\gamma^{B} = \varepsilon_{AB'}\,\gamma^{B'}$. But clearly this is possible only if $B = B'$; therefore if $B \neq B'$, then $C(A,B) \neq C(A,B')$ and the conclusion follows.

<u>Proposition</u> (2.5): The 16 γ^{A}'s are linearly independent.

Proof: Suppose we have a relation $\sum_{A=1}^{16} \alpha_{A}\,\gamma^{A} = 0$. Let us pick B in $\{1,\ldots,16\}$ and multiply by γ^{B}:

$$0 = \sum_{A=1}^{16} \alpha_{A}\gamma^{B}\gamma^{A} = \sum_{A=1}^{16} \alpha_{A}\varepsilon_{BA}\gamma^{C(B,A)}$$

From proposition (2.4), as A goes from 1 to 16 in this sum, $\gamma^{C(B,A)}$ goes over the whole set Γ. For $A=B$, $\gamma^{C(B,A)} = I$ and for all the other values, according to proposition (2.3), $\gamma^{C(B,A)}$ is a traceless matrix. Whence taking the trace of our equation yields $4\alpha_B \varepsilon_{BB} = 0$, or $\alpha_B = 0$. Since B was arbitrary, it follows that the γ^A's are independent.

One easily obtains the following product rules, some of which will be useful in the sequel.

$$\gamma^\mu \gamma^{[\rho\sigma]} = \varepsilon^{\mu\rho\sigma\theta}\gamma_\theta \gamma^5 + g^{\mu\rho}\gamma^\sigma - g^{\mu\sigma}\gamma^\rho \qquad (2.2)$$

$$\gamma^\mu \gamma^{[\rho\sigma\theta]} = -\varepsilon^{\mu\rho\sigma\theta}\gamma^5 + g^{\mu\rho}\gamma^{[\sigma\theta]} - g^{\mu\sigma}\gamma^{[\rho\theta]} + g^{\mu\theta}\gamma^{[\rho\sigma]} \qquad (2.3)$$

$$\gamma^\mu \gamma^5 = \frac{1}{6}\varepsilon^\mu{}_{\alpha\beta\delta}\,\gamma^{[\alpha\beta\delta]}$$

$$\gamma^{[\alpha\beta]}\gamma^{[\mu\nu]} = -\varepsilon^{\alpha\beta\mu\nu}\gamma^5 + 2g^{\mu[\alpha}\gamma^{\beta]}\gamma^\nu - 2g^{\nu[\alpha}\gamma^{\beta]}\gamma^\mu \qquad (2.4)$$

$$\gamma^{[\rho\sigma]}\gamma_5 = \frac{1}{2}\varepsilon^{\rho\sigma}{}_{\alpha\beta}\,\gamma^{[\alpha\beta]} \equiv \tilde{\gamma}^{[\rho\sigma]} \qquad (2.5)$$

$$\gamma^\mu \gamma^5 \gamma^{[\rho\sigma]} = -\varepsilon^{\mu\rho\sigma\theta}\gamma_\theta + g^{\mu\rho}\gamma_5\gamma^\sigma - g^{\mu\sigma}\gamma_5\gamma^\rho \qquad (2.6)$$

From the relation (1.3c) it is possible to derive the commutators and anticommutators of all pairs of elements of Γ. We list here the results.

Table 2 Commutators and anticommutators of the γ^A's

$[\gamma^\mu,\gamma^\nu] = 2\,\gamma^{[\mu\nu]}$	$\{\gamma^\mu,\gamma^\nu\} = 2\,g^{\mu\nu}I$
$[\gamma^\mu,\gamma^5] = 2\gamma^\mu\gamma^5$	$\{\gamma^\mu,\gamma^5\} = 0$
$[\gamma^\lambda,\gamma^{[\mu\nu]}] = 2(g^{\lambda\mu}\gamma^\nu - g^{\lambda\nu}\gamma^\mu)$	$\{\gamma^\lambda,\gamma^{[\mu\nu]}\} = 2\,\gamma^{[\lambda\mu\nu]}$
$[\gamma^\mu,\gamma^{[\rho\sigma\tau]}] = 2\,\gamma^{[\mu\rho\sigma\tau]}$	$\{\gamma^\mu,\gamma^{[\rho\sigma\tau]}\} = 2(g^{\mu\rho}\gamma^{[\sigma\tau]} + g^{\tau\mu}\gamma^{[\rho\sigma]} + g^{\sigma\tau}\gamma^{[\mu\rho]})$
$[\gamma^5,\gamma^{[\mu\nu]}] = 0$	$\{\gamma^5,\gamma^{[\mu\nu]}\} = -\,\epsilon^{\mu\nu\rho\sigma}\gamma_{[\rho\sigma]}$
$[\gamma^5,\gamma^{[\lambda\rho\sigma]}] = 2\epsilon^{\lambda\rho\sigma\theta}\gamma_\theta$	$\{\gamma^5,\gamma^{[\lambda\rho\sigma]}\} = 0$
$[\gamma^{[\lambda\rho]},\gamma^{[\mu\nu]}] = 2(g^{\lambda\mu}\gamma^{[\rho\nu]} + g^{\nu\lambda}\gamma^{[\mu\rho]} + g^{\rho\nu}\gamma^{[\lambda\mu]} + g^{\mu\rho}\gamma^{[\nu\lambda]})$	$\{\gamma^{[\lambda\rho]},\gamma^{[\mu\nu]}\} = 2\gamma^{[\lambda\rho\mu\nu]} + 2(g^{\mu\rho}g^{\nu\lambda} - g^{\mu\lambda}g^{\nu\rho})I$
$[\gamma^{[\mu\nu]},\gamma^{[\lambda\rho\sigma]}] = 2\gamma^5(\epsilon^{\lambda\rho\sigma\mu}\gamma^\nu - \epsilon^{\lambda\rho\sigma\nu}\gamma^\mu)$	$\{\gamma^{[\mu\nu]},\gamma^{[\lambda\rho\sigma]}\} = -2\epsilon^{\lambda\rho\sigma}{}_\theta\,\epsilon^{\theta\mu\nu\beta}\gamma_\beta$ or equivalently $\{\gamma^{[\mu\nu]},\gamma^5\gamma^\lambda\} = 2\epsilon^{\mu\nu\lambda\rho}\gamma_\rho$
$[\gamma^{[\lambda\rho\sigma]},\gamma^{[\alpha\beta\delta]}] = 2\epsilon^{\lambda\rho\sigma\theta}\epsilon^{\alpha\beta\delta\mu}\gamma_{[\theta\mu]}$ or equivalently $[\gamma^5\gamma^\mu,\gamma^5\gamma^\nu] = 2\gamma^{[\mu\nu]}$	$\{\gamma^{[\lambda\rho\sigma]},\gamma^{[\alpha\beta\delta]}\} = -2\epsilon^{\lambda\rho\sigma}{}_\theta\,\epsilon^{\theta\alpha\beta\delta}I$ or equivalently $\{\gamma^5\gamma^\mu,\gamma^5\gamma^\nu\} = 2g^{\mu\nu}I$

(2) The representations of the relations $\{\gamma^\mu,\gamma^\nu\}=2g^{\mu\nu}I$

We now come to the part of this chapter dealing with the order N of γ matrices and the irreducibility of such systems.

Theorem (2.6): The order N of γ matrices cannot be smaller than 4.

Proof: The dimension of the complex vector space of N×N complex matrices is N^2. Since by proposition (2.5) Γ is a set of 16 independent matrices, we see that N^2 has to be ≥ 16.

Theorem (2.7): The matrices Γ^μ defined by

$$\Gamma^o = \begin{pmatrix} \sigma_o & o \\ o & -\sigma_o \end{pmatrix} \qquad \Gamma^i = \begin{pmatrix} o & \sigma_i \\ -\sigma_i & o \end{pmatrix} \qquad i=1,2,3. \tag{2.7}$$

where $\sigma_o=\begin{pmatrix} 1 & o \\ o & 1 \end{pmatrix}$ $\sigma_1=\begin{pmatrix} o & 1 \\ 1 & o \end{pmatrix}$ $\sigma_2=\begin{pmatrix} o & -i \\ i & o \end{pmatrix}$ $\sigma_3=\begin{pmatrix} 1 & o \\ o & -1 \end{pmatrix}$ (2.8)

form a system of unitary γ matrices.

Proof: From the relations $\sigma_i\sigma_j = i\varepsilon_{ijk}\sigma_k + \delta_{ij}\sigma_o$ we obtain

$$\{\Gamma^i,\Gamma^j\} = -\begin{pmatrix} \{\sigma_i,\sigma_j\} & o \\ o & \{\sigma_i,\sigma_j\} \end{pmatrix} = 2g^{ij}I$$

$$\{\Gamma^i,\Gamma^o\} = \begin{pmatrix} o & \sigma_i \\ \sigma_i & o \end{pmatrix} - \begin{pmatrix} o & \sigma_i \\ \sigma_i & o \end{pmatrix} = 0$$

$$(\Gamma^o)^2 = I$$

The fact that each Γ^μ is unitary follows from

$$\Gamma^{o+} = \Gamma^o \qquad \Gamma^{i+} = -\Gamma^i \tag{2.9}$$

This system $\{\Gamma^{\mu}\}$ is referred to as the Dirac representation.

Our next theorem, in group theory language, would be said to reflect the irreducibility of any set of 4x4 γ matrices.

Theorem (2.8): Let $\{\gamma^{\mu}\}$ be a system of 4x4 γ matrices. Any matrix commuting with the four of them has to be a multiple of the identity.

Proof: Suppose M is such a matrix. It follows that it commutes with every γ^{A} in Γ. But proposition (2.5) says that, in the case under consideration, Γ is a basis. Therefore M commutes with everything. But it is easy to prove that any linear operator on a vector space commuting with all the others has to be a multiple of the identity. The conclusion applies to M.

Theorem (2.8) will be used repeatedly in the sequel. Together with the fundamental theorem of γ matrices, which is the subject of the next chapter, it lies at the basis of most of the constructions which we will make.

So far all we know about the possible order of γ matrices is that N is even and ≥ 4 (see the discussion following eq. (1.3c) and theorem (2.6)). It is in fact known that N has to be a multiple of 4. This can be proved by using the theory of representations of finite groups (the starting point of such an approach is to observe that the set $\bar{\Gamma} = \{\pm\gamma^{A};\gamma^{A}\epsilon\Gamma\}$ is a group) (Jansen and Boon [1967]). The proof we give here is (to our knowledge) original and owes absolutely nothing to group theory. We start with a little lemma which will also be used in the proof of another result.

Lemma: Let $\{\gamma^\mu\}$ be an arbitrary system of γ matrices. Then γ^o and $i\gamma^1\gamma^2$ can always be diagonalised simultaneously.

Proof: Let A be a matrix such that $A^2 = I$. Then A may be written as $A = I - 2P$ where P is a projector, namely $P = \frac{1}{2}(I-A)$, $(P^2=P)$. But a projector can always be diagonalised; whence so can A. Our two matrices commute and have square I. So the conclusion follows.

Theorem (2.9): The order N of γ matrices has to be a multiple of 4.

Proof: Let $\{\gamma^\mu\}$ be a system of γ matrices. By the previous lemma we may assume, by performing a similarity transformation on $\{\gamma^\mu\}$ if necessary, that γ^o and $i\gamma^1\gamma^2$ are diagonal. Since their square is I, their diagonal entries have to be ± 1. Moreover since by proposition (2.3) they are both traceless, the number of +1's has to be equal to the number of -1's. Again by performing a similarity transformation if necessary we may assume that $\gamma^o = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix}$. N is even so there is an integer n such that N=2n. Let us write $i\gamma^1\gamma^2 = \text{diag}(a_1,\dots,a_n,b_1,\dots,b_n)$. The matrix $\gamma^o i\gamma^1\gamma^2$ is also traceless and its trace is $a_1 + \dots + a_n - (b_1 + \dots + b_n)$. Now we have to distribute n +1's and n -1's among $a_1,\dots,a_n,b_1,\dots,b_n$.

Suppose that we put r - 1's ($0 \leq r \leq n$) in $a_1, \ldots, a_n$. There are n-r left over to be distributed among $b_1, \ldots, b_n$ and therefore we have:

$$a_1 + \ldots + a_n = -r + (n-r) = n-2r$$

$$b_1 + \ldots + b_n = -(n-r) + r = -n + 2r$$

Whence $Tr\gamma^o i\gamma^1\gamma^2 = n-2r-(-n+2r) = 2n - 4r$. In order that this vanishes we must have $r = \frac{n}{2}$. Since r is an integer it follows that n is even; whence $N = 2n$ is a multiple of 4.

<u>Corollary</u>: γ matrices of arbitrary order have determinant 1.

Proof: γ^o is similar to $\begin{bmatrix} I & \\ & -I \end{bmatrix}$ and N is a multiple of 4. Whence $\det \gamma^o = (-1)^{N/2} = 1$. A similar argument applies to the γ^k's.

The final result of this section, which we are about to present, is not the least in importance since it establishes in some sense the uniqueness of the relativistic equation for the electron wave function. As one might guess we are going to be concerned with the irreducibility of the representations of eq. (1.3c). It is well known that the only irreducible representations are provided by matrices of order 4. This result makes one think of group theory and of course, like theorem (2.9), it can be derived via the theory of representations of finite groups (Jansen and Boon [1967]). But as we did for theorem (2.9) we have found a simple way of proving it keeping away from group theory. Our proof is likely to have been thought of before but we have not met it anywhere.

Theorem (2.10): Any (unitary or not) representation of eq. (1.3c) can be reduced to one of degree 4. Whence the irreducible ones are those of degree 4.

Proof: Let $\{\gamma^\mu\}$ be such a representation. By the previous lemma we know that it is equivalent to one in which γ^o and $i\gamma^1\gamma^2$ are both diagonal. So we may assume without loss of generality that they are diagonal. In order that a subspace be invariant under $\{\gamma^\mu\}$ it is clear by proposition (2.2) that it is necessary and sufficient that it be invariant under the set Γ of table 1. Now pick a non vanishing vector u and define $u_A \equiv \gamma^A u$. Again by proposition (2.2) it is clear that the subspace V spanned by the u_A's is Γ-invariant. The special trick of the proof lies in an appropriate choice of u. From the proof of theorem (2.9) one easily checks that one can pick a non vanishing u such that $\gamma^o u = i\gamma^1\gamma^2 u = u$. We claim that the subspace V spanned by the corresponding u_A's is 4-dimensional. Indeed it is generated by $u, u_1 \equiv \gamma^1 u$, $u_3 \equiv \gamma^3 u$ and $u_{31} \equiv \gamma^3\gamma^1 u$. This is seen by letting Γ act on u: $Iu = u$, $\gamma^o u = u$, $\gamma^1 u = u_1$,

$$\gamma^2 u = -\gamma^1\gamma^1\gamma^2 u = iu_1, \quad \gamma^3 u = u_3, \quad \gamma^1\gamma^2 u = -iu,$$

$$\gamma^3\gamma^1 u = u_{31}, \quad \gamma^2\gamma^3 u = \gamma^3\gamma^1\gamma^1\gamma^2 u = -iu_{31}, \quad \gamma^0\gamma^1 u = -u_1,$$

$$\gamma^0\gamma^2 u = -iu_1, \quad \gamma^0\gamma^3 u = -u_3, \quad \gamma^0\gamma^1\gamma^2 u = \gamma^1\gamma^2 u = -iu,$$

$$\gamma^0\gamma^3\gamma^1 u = u_{31}, \quad \gamma^0\gamma^2\gamma^3 u = -iu_{31}, \quad \gamma^1\gamma^2\gamma^3 u = \gamma^3\gamma^1\gamma^2 u = -iu_3,$$

$$\gamma^5 u = iu_3 .$$

By theorem (2.6) the vectors u, u_1, u_3, u_{31} are necessarily linearly independent. This completes the proof.

From now on, when we talk about γ matrices, unless otherwise stated, it will always be understood that these are 4×4 matrices. In physical applications the γ^μ's are always unitary. All the γ^A's are then unitary as well. But as Pauli did in his paper (Pauli [1936]) we will invoke this assumption only when needed: as has already been seen many results follow without it.

CHAPTER III

THE FUNDAMENTAL THEOREM OF γ MATRICES AND CONSEQUENCES

(1) The fundamental theorem

The so-called fundamental theorem of γ matrices is the basis upon which lies the relativistic invariance of the Dirac theory of electrons and positrons. This section is devoted to its proof and to the exposition of some of its consequences. The proof that we give is, to our knowledge, original and is, as will be seen, quite simple. It is based on a seldom used result of linear algebra. To preserve the continuity of exposition the proof of this result will be deferred to the end of this chapter. As will be seen later, we state it here in a restricted context which will be sufficient for our purpose.

Let us denote by $A_n(C)$ the algebra of $n \times n$ complex matrices. By an automorphism of $A_n(C)$ we mean a bijective linear map $h: A_n(C) \to A_n(C)$ which also preserves multiplication, that is for any μ, ν in C and M, N in $A_n(C)$ we have:

$h(\mu M + \nu N) = \mu h(M) + \nu h(N)$, $h(MN) = h(M)h(N)$, $h(M)=0 \Rightarrow M=0$.

It is clear that, given a non-singular S in $A_n(C)$, the map $M \to SMS^{-1}$ is an automorphism of $A_n(C)$. The result of linear algebra we were referring to is the converse of this.

Theorem (3.1): If h is an automorphism of $A_n(C)$, then there exists a non-singular matrix S in $A_n(C)$ such that $h(M) = SMS^{-1}$ for all M.

Having stated this we may now give our proof of the fundamental theorem. As was mentioned in chapter II (section 1) the key of this proof is eq. (2.1) together with the trivial proposition (2.2').

Theorem (3.2): (The fundamental theorem of γ matrices): Let $\{\gamma^\mu\}$, $\{\hat{\gamma}^\mu\}$ be two systems of 4×4 γ matrices. Then there exists a non-singular matrix S such that $\hat{\gamma}^\mu = S^{-1}\gamma^\mu S$.

Proof: Let Γ and $\hat{\Gamma}$ be the two sets constructed from $\{\gamma^\mu\}$ and $\{\hat{\gamma}^\mu\}$ according to table 1. From proposition (2.5) we know that both Γ and $\hat{\Gamma}$ are basis of $A_4(C)$. Therefore we may define a linear map $h: A_4(C) \to A_4(C)$ by $h(\gamma^A) = \hat{\gamma}^A$ and this map is bijective. Moreover it preserves products. Indeed let $M = \sum_A \alpha_A \gamma^A$, $N = \sum_B \beta_B \gamma^B$. Then we have:

$$h(MN) = h(\sum_{AB} \alpha_A \beta_B \gamma^A \gamma^B) = \sum_{AB} \alpha_A \beta_B h(\epsilon_{AB} \gamma^{C(A,B)}) =$$

$$\sum_{AB} \alpha_A \beta_B \epsilon_{AB} \hat{\gamma}^{C(A,B)} = \sum_{AB} \alpha_A \beta_B \hat{\gamma}^A \hat{\gamma}^B = h(M)h(N)$$

where we have used propositions (2.2) and (2.2').
Hence h fullfills all the conditions of theorem (3.1) and it follows that there exits a non-singular S such that $\hat{\gamma}^A = h(\gamma^A) = S^{-1}\gamma^A S$ (q.e.d.).

Proposition (3.1): The matrix S of the fundamental theorem is unique up to a multiplicative factor.

Proof: Suppose that S and T satisfy

$$\hat{\gamma}^{\mu} = S^{-1}\gamma^{\mu}S = T^{-1}\gamma^{\mu}T$$

It follows that the commutator $[\gamma^{\mu}, ST^{-1}]$ vanishes. From theorem (2.8) we may therefore conclude that $ST^{-1} = cI$ or $T = \frac{1}{c}S$ (q.e.d.).

(2) Consequences of the fundamental theorem

The matrix S of the fundamental theorem is of course closely related to the systems of γ matrices from which it arises. This is illustrated in the following little result which we present here as a curiosity, since we shall not use it later.

Proposition (3.2): Let $\{\gamma^{\mu}\}$ and $\{\hat{\gamma}^{\mu}\}$ be two systems of γ matrices, with associated sets Γ and $\hat{\Gamma}$. The invertible matrix S such that $\hat{\gamma}^{\mu} = S^{-1}\gamma^{\mu}S$ has the same coordinates in both basis Γ and $\hat{\Gamma}$.

Proof: Let $S = \sum_A \alpha_A \gamma^A = \sum_A \beta_A \hat{\gamma}^A$. From $S\hat{\gamma}^A = \gamma^A S$ we get

$$\gamma^B S = \sum_A \alpha_A \gamma^B \gamma^A = S\hat{\gamma}^B = \sum_A \beta_A \hat{\gamma}^A \hat{\gamma}^B$$

Taking the trace on each side of $\sum_A \alpha_A \gamma^B \gamma^A = \sum_A \beta_A \hat{\gamma}^A \hat{\gamma}^B$ and using propositions (2.3) and (2.4) gives

$\alpha_B = \beta_B$ (q.e.d.).

Another connection between $\{\gamma^\mu\}$ and S which will not be a mere curiosity for us is the following.

Proposition (3.3): Let $\{\gamma^\mu\}$, $\{\hat{\gamma}^\mu\}$ be two systems of γ matrices such that all the γ^μ's and $\hat{\gamma}^\mu$'s are unitary. Then the matrix S connecting the two systems can be chosen to be unitary. Such a choice is unique up to a phase factor.

Proof: We have from the unitarity of γ^μ, $\hat{\gamma}^\mu$ and $\hat{\gamma}^\mu = S^{-1}\gamma^\mu S$:

$$S^{-1}\gamma^o S = \hat{\gamma}^o = \hat{\gamma}^{o+} = S^+\gamma^{o+} S^{-1+} = S^+\gamma^o S^{-1+}$$

$$S^{-1}\gamma^i S = \hat{\gamma}^i = -\hat{\gamma}^{i+} = -S^+\gamma^{i+} S^{-1+} = S^+\gamma^i S^{-1+}$$

that is $S^{-1}\gamma^\mu S = S^+\gamma^\mu S^{-1+}$. From this we infer $[\gamma^\mu, SS^+] = 0$, and by theorem (2.8) conclude that $SS^+ = cI$. Now SS^+ is obviously self-adjoint and positive. Whence c is real and > 0. Taking $S' = \frac{1}{\sqrt{c}} S$ yields the required unitary matrix. The fact that the choice of a unitary S is unique up to a phase factor is obvious in view of proposition (3.1) and the unitarity condition.

The fundamental theorem allows us to draw other interesting general conclusions about γ matrices. For example, it says that γ^o has to be similar to the matrix Γ^o of eq. (2.7). Whence the characteristic and minimal polynomials of γ^o have to be $(t+1)^2(t-1)^2$ and $(t+1)(t-1)$ respectively, and its determinant is 1. To make similar remarks about γ^i we observe that the set $\{\hat{\gamma}^\mu\}$ defined by $\hat{\gamma}^o = i\gamma^1$, $\hat{\gamma}^1 = i\gamma^o$, $\hat{\gamma}^2 = \gamma^2$, $\hat{\gamma}^3 = \gamma^3$

is also a system of γ matrices. Whence $i\gamma^1$ is similar to Γ^o, or γ^1 is similar to $-i\Gamma^o$. This shows that the characteristic and minimal polynomials of γ^1 (or any γ^i) are $(t+i)^2(t-i)^2$ and $(t+i)(t-i)$ respectively and its determinant is 1. These observations lead themselves immediately to the following algebraic characterization of matrices which can be γ matrices.

Proposition (3.4): Let M be a complex 4×4 matrix. In order that M be the γ^o of a $\{\gamma^\mu\}$ system it is necessary and sufficient that its characteristic and minimal polynomials be $(t+1)^2(t-1)^2$ and $(t+1)(t-1)$ respectively. In order that M be the γ^1 (or γ^2 or γ^3) of some $\{\gamma^\mu\}$ system it is necessary and sufficient that the characteristic and minimal polynomials of M be $(t+i)^2(t-i)^2$ and $(t+i)(t-i)$ respectively.

Before we close this chapter with the proof of theorem (3.1) we give a last application of the fundamental theorem. This theorem enabled us, in a rather curious way, to answer a question which arises naturally when dealing with γ matrices. When we look closely at the set $\{\Gamma^\mu\}$ of eq. (2.7) we notice that all these matrices, except Γ^2, are real. Whence it seems natural to ask whether one could find a system of γ matrices where all the matrices would be real. The answer is negative. As we will show below, the existence of such a system would imply the existence of a complex number c whose squared magnitude c^*c would be -1; this is of course absurd.

Proposition (3.5): There exists no system of γ matrices such that each γ^{μ} is real.

Proof: Consider again the particular system $\{\Gamma^{\mu}\}$ exhibited in eq. (2.7). If $\{\gamma^{\mu}\}$ is any system of γ matrices the fundamental theorem says that there exists a non-singular S such that $S^{-1}\,\gamma^{\mu}\,S = \Gamma^{\mu}$. (3.1)

Let $M = i\,\gamma^{[031]}$. Using $\Gamma^{2*} = -\Gamma^{2}$ and $(\Gamma^{\nu})^{*} = \Gamma^{\nu}$ for $\nu \neq 2$, one easily checks that the following equation holds true:

$$S^{-1*}\,(M\,\gamma^{\nu}\,M)^{*}\,S^{*} = \Gamma^{\nu}, \qquad \nu=0,1,2,3$$

(Here * means complex conjugate.)

From this equation and equation (3.1) we obtain

$$S^{-1*}\,(M\,\gamma^{\nu}\,M)^{*}\,S^{*} = S^{-1}\,\gamma^{\nu}\,S$$

If we now suppose that all the γ^{ν}'s are real, this last equation may be rewritten as:

$$\gamma^{\nu} = \gamma^{\nu *} = M\,S^{*}\,S^{-1}\,\gamma^{\nu}\,S\,S^{*-1}\,M$$

from which we deduce that $[\gamma^{\nu}, MS^{*}\,S^{-1}] = 0$. This implies that there exists a number c such that $MS^{*}\,S^{-1} = cI$. This gives the equation $S = \frac{1}{c}M\,S^{*}$ and its complex conjugate $S^{*} = \frac{1}{c^{*}}\,(-M)S$. Substituting the second in the first yields $cc^{*} = -1$, a contradiction. Therefore the four γ^{μ}'s cannot all be real.

Though there is no system of real γ matrices, there are systems in which the γ^{μ}'s are purely imaginary. Their special interest is that they make the free Dirac equation (1.4) real; hence ψ is a solution if and only if its real and imaginary parts are separate solutions of the equation.

It is easy to give a fairly explicit description of all these systems. Let $\{\gamma^{\mu}\}$ be an arbitrary system of γ matrices. By the fundamental theorem there exists an invertible matrix T such that $\gamma^{\mu} = T^{-1}\Gamma^{\mu}T$, where $\{\Gamma^{\mu}\}$ is the particular system of equation (2.7). Due to the fact that Γ^2 is purely imaginary while Γ^o, Γ^1 and Γ^3 are real, the γ^{μ}'s will be purely imaginary if and only if $T^{-1*}\Gamma^2 T^* = T^{-1}\Gamma^2 T$ and $T^{-1*}\Gamma^{\nu}T^* = -T^{-1}\Gamma^{\nu}T$, for $\nu = 0, 1, 3$. This is equivalent to saying that T^*T^{-1} commutes with Γ^2 and anti-commutes with Γ^o, Γ^1 and Γ^3. This implies $T^*T^{-1} = \alpha\Gamma^2$ or $T^* = \alpha\Gamma^2 T$. Clearly the number α has to be a phase factor. Upon writing $T = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$, the general solution of $T^* = \alpha\Gamma^2 T$ is easily seen to be

$$T = \begin{bmatrix} A & B \\ \alpha^*\sigma_2 A^* & \alpha^*\sigma_2 B^* \end{bmatrix}, \quad \text{A and B arbitrary, } |\alpha| = 1$$

This can be alternatively written as

$$T = \begin{bmatrix} u \\ v \\ -i\alpha^* v^* \\ i\alpha^* u \end{bmatrix} \quad \text{u and v arbitrary, } |\alpha| = 1$$

where u and v are row vectors. As long as the choice of A and B leads to an invertible matrix T, the matrices $T^{-1}\Gamma^{\mu}T$ will form a system of purely imaginary γ matrices. Conversely all the purely imaginary systems can be obtained that way.

The particular choice $A = \sigma_o + \sigma_2$, $B = \sigma_o - \sigma_2$ and $\alpha = -1$ gives the system

$$\gamma^o = \begin{pmatrix} \sigma_2 & \\ & -\sigma_2 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} -i\sigma_3 & \\ & -i\sigma_3 \end{pmatrix},$$

$$\gamma^2 = \begin{pmatrix} & -\sigma_2 \\ \sigma_2 & \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} i\sigma_1 & \\ & i\sigma_1 \end{pmatrix}$$

It is referred to as the Majorana representation.

(3) The proof of theorem (3.1)

We now come to the proof of theorem (3.1). As mentioned at the beginning of this chapter we stated it in a restricted context. It turns out that it is also true for $A_n(K)$, the algebra of n×n matrices over an arbitrary field K, for example the field of real numbers. We stated it for K=C because this was all we needed. Quite amusingly it wasn't our knowledge of this result which inspired our proof of the fundamental theorem but rather the study of the usual proofs of the fundamental theorem lead us to guess that such a result might be true. We were able to trace it in only one book (Herstein [1964]) where it is stated as a problem (problem 27, page 279). Therefore the proof we give here is ours. It is possible that a shorter proof could be given.

Theorem (4.1'): If h is an automorphism of $A_n(K)$, that is a bijective linear map preserving products of $A_n(K)$ onto itself, then there exists a non-singular matrix S in K such that $h(M) = SMS^{-1}$ for all M.

Proof: Throughout this proof we don't use the summation convention. We will denote by $|1\rangle,\ldots,|n\rangle$ the canonical basis of K^n, that is:

$$|1\rangle = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad |2\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \quad \text{etc.},$$

and by $\langle 1|,\ldots,\langle n|$ the dual basis: $\langle i|j\rangle = \delta_{ij}$. We denote by M_{ij} the matrix having null entries except at the intersection of the i^{th} row and the j^{th} column where it has a 1: $\langle k|M_{ij}|\ell\rangle = \delta_{ki}\,\delta_{j\ell}$. These matrices multiply according to $M_{ij}M_{k\ell} = \delta_{jk}M_{i\ell}$, and they form a basis of $A_n(K)$. Now since h is an automorphism the matrices $H_{ij} \equiv h(M_{ij})$ also form a basis and multiply according to

$$H_{ij}H_{k\ell} = \delta_{jk}H_{i\ell} \tag{3.2}$$

As often is the case it makes the argument simpler to think of H_{ij} as a linear operator instead of a matrix. Whence we consider H_{ij} as the linear operator on K^n whose matrix with respect to the canonical basis is H_{ij}. From eq. (3.2) it follows that $H_{11},\ldots,H_{nn}$ are n projectors such that $H_{ii}H_{jj} = 0$ if $i\neq j$. Moreover neither of them is 0 because h is an automorphism. It follows immediately from this that there exists a basis $|1\rangle',\ldots,|n\rangle'$ of K^n such that

$$H_{ii}|j\rangle' = \delta_{ij}\;|i\rangle' \tag{3.3}$$

From eq. (3.2) we also deduce that $H_{ij} = H_{ii}H_{ij}H_{jj}$. This together with (3.3) implies that

$$H_{ij}|k\rangle' = h_{ij}\,\delta_{jk}\;|i\rangle' \;,\quad h_{ij}\in K \;,\quad h_{ij}\neq 0 \tag{3.4}$$

Since H_{ii} is a projector, $h_{ii} = 1$. Moreover eqs. (3.2) and (3.4) clearly imply that $h_{ik}h_{kj} = h_{ij}$. Now let us define a new basis of K^n by

$|i>'' = \frac{1}{h_{1i}} \; |i>'$. Then we have:

$$H_{ij}|j>'' = \frac{h_{ij}}{h_{1j}} \; |i>' = \frac{h_{1i}h_{ij}}{h_{1j}} \; |i>'' = |i>''$$

$$H_{ij} \; |k>'' = 0 \quad \text{if} \quad k \neq j$$

Therefore the matrix of H_{ij} with respect to the basis $|1>'',\ldots,|n>''$ is M_{ij}. So if S is the transition matrix from the basis $(|i>)$ to the basis $(|i>'')$ $(|i>'' = \sum_j S_{ji} \; |j>)$ one has $H_{ij} = S \; M_{ij} \; S^{-1}$, that is, $h(M_{ij}) = S \; M_{ij} S^{-1}$. Since h is linear and M_{ij} is a basis it follows that $h(M) = SMS^{-1}$ for an arbitrary M. (q.e.d.).

CHAPTER IV

AN ALGEBRAIC METHOD FOR FINDING THE MATRIX S OF THE FUNDAMENTAL THEOREM

(1) Finding S

The fundamental theorem of γ matrices asserts the existence of an invertible matrix S connecting two given sets of 4×4 γ matrices. But it does not tell us what S is. In later chapters we shall give the solution of this problem when the two sets $\{\gamma^\mu\}$, $\{\hat{\gamma}^\mu\}$ are related through a Lorentz transformation $\hat{\gamma}^\mu = \Omega^\mu{}_\nu \gamma^\nu$ by using Lie group techniques. This is what is usually done in the physical literature. In the present chapter we adopt a purely algebraic point of view and look at the general case: the γ^μ's are not assumed [illegible]tary and the two sets $\{\hat{\gamma}^\mu\}$ and $\{\gamma^\mu\}$ are not [illegible] related through a Lorentz transformation. [illegible] people who investigate what happens if, instead [illegible] the field of real numbers to construct the M[illegible]ki space, one starts with a field having only a f[illegible] (enormous) number of elements. The aim of this is to[illegible]troduce a fundamental length in physics. In such a f[illegible]work it is clear that one couldn't use infinitesimal tr[illegible]sformations to obtain S in $\Omega^\mu{}_\nu \gamma^\nu = S^{-1} \gamma^\mu S$. The algebraic method which we set up in this section and the following one would provide a substitute.

The equations $\hat{\gamma}^{\mu}S = S\,\gamma^{\mu}$, where S is the unknown, give rise to a system of 64 linear equations with 16 unknowns. It is certainly not convenient to attack this directly. We have found a roundabout way which simplifies the task considerably. Consider again the particular set $\{\Gamma^{\mu}\}$ of eq. (2.7). By the fundamental theorem there exist S and T such that $\hat{\gamma}^{\mu} = T^{-1}\,\Gamma^{\mu}\,T$ and $\Gamma^{\mu} = S^{-1}\,\gamma^{\mu}\,S$. Clearly $\hat{\gamma}^{\mu} = (ST)^{-1}\,\gamma^{\mu}\,(ST)$. The point is now that, thanks to the simplicity of $\{\Gamma^{\mu}\}$, finding S and T is very simple as shown by the following result.

Proposition (4.1): Let $\{\gamma^{\mu}\}$ be a system of 4×4 γ matrices. Then the systems

$$\text{(a)}\quad \begin{matrix} (\gamma^1 + i\gamma^2)u = 0 \\ (I + \gamma^o)u = 0 \end{matrix} \qquad\qquad \text{(b)}\quad \begin{matrix} w\,(\gamma^1 - i\gamma^2) = 0 \\ w\,(I + \gamma^o) = 0 \end{matrix} \tag{4.1}$$

where u and w are a column and a row vector respectively, have non vanishing solutions which are determined up to multiplicative factors. The matrices S and T defined by

$$S = (\gamma^3 u \quad \gamma^1 u \quad u \quad \gamma^3\gamma^1 u) \tag{4.2a}$$

$$T = \begin{bmatrix} -w\,\gamma^3 \\ -w\,\gamma^1 \\ w \\ w\,\gamma^1\gamma^3 \end{bmatrix} \tag{4.2b}$$

satisfy $\Gamma^{\mu} = S^{-1} \gamma^{\mu} S \qquad \gamma^{\mu} = T^{-1} \Gamma^{\mu} T$ (4.3)

Proof: The fundamental theorem guarantees the existence of the matrices S and T. So all we have to do is to show that they necessarily have the form stated above. We begin with S. We first decompose all our matrices into blocks of order 2:

$$\gamma^{\mu} \quad \begin{bmatrix} A^{\mu} & B^{\mu} \\ C^{\mu} & D^{\mu} \end{bmatrix} \qquad S = \begin{bmatrix} X & Y \\ Z & U \end{bmatrix}.$$

The unknowns are now the four 2×2 matrices X,Y,Z,U. The equations which they must satisfy are derived from (4.3) and are as follows:

$$\begin{aligned} A^{i}X + B^{i}Z &= -Y\sigma_{i} & A^{o}X + B^{o}Z &= X \\ A^{i}Y + B^{i}U &= X\sigma_{i} & A^{o}Y + B^{o}U &= -Y \\ C^{i}X + D^{i}Z &= -U\sigma_{i} & C^{o}X + D^{o}Z &= Z \\ C^{i}Y + D^{i}U &= Z\sigma_{i} & C^{o}Y + D^{o}U &= -U \end{aligned} \tag{4.4}$$

The second and fourth equations on the left-hand side give

$$X = (A^{i}Y + B^{i}U)\sigma_{i} \quad \text{(no sum)} \tag{4.5}$$

$$Z = (C^{i}Y + D^{i}U)\sigma_{i} \quad \text{(no sum)} \tag{4.6}$$

This system of six equations can be rewritten as follows

$$X = (A^1 Y + B^1 U)\sigma_1 \tag{4.7a}$$

$$Z = (C^1 Y + D^1 U)\sigma_1 \tag{4.7b}$$

$$A^2 Y + B^2 U = (A^1 Y + B^1 U)\sigma_1\sigma_2 = (A^1 Y + B^1 U)i\sigma_3 \tag{4.7c}$$

$$A^3 Y + B^3 U = (A^1 Y + B^1 U)\sigma_1\sigma_3 = -(A^1 Y + B^1 U)i\sigma_2 \tag{4.7d}$$

$$C^2 Y + D^2 U = (C^1 Y + D^1 U)\sigma_1\sigma_2 = (C^1 Y + D^1 U)i\sigma_3 \tag{4.7e}$$

$$C^3 Y + D^3 U = (C^1 Y + D^1 U)\sigma_1\sigma_3 = -(C^1 Y + D^1 U)i\sigma_2 \tag{4.7f}$$

The first two give X and Z in terms of Y and U and the last four guarantee that the right-hand sides in (4.5) and (4.6) are independent of i. The problem is now reduced to finding Y and U.

From the identity

$$\gamma^\mu\gamma^\nu = \begin{bmatrix} A^\mu A^\nu + B^\mu C^\nu & A^\mu B^\nu + B^\mu D^\nu \\ C^\mu A^\nu + D^\mu C^\nu & C^\mu B^\nu + D^\mu D^\nu \end{bmatrix} \tag{4.8}$$

we see that multiplying (4.7c) by A^2, (4.7e) by B^2 and adding, we obtain:

$$-Y = [(A^2A^1 + B^2C^1)Y + (A^2B^1 + B^2D^1)U]i\sigma_3$$

Similarly multiplying (4.7c) by C^2, (4.7e) by D^2 and adding, we obtain:

$$-U = [(C^2A^1 + D^2C^1)Y + (C^2B^1 + D^2D^1)U]i\sigma_3$$

These last two equations may be cast into one simpler equation:

$$i\,\gamma^1\gamma^2\begin{pmatrix}Y\\U\end{pmatrix} = \begin{pmatrix}Y\sigma_3\\U\sigma_3\end{pmatrix} \qquad (4.9)$$

Moreover one can check that the process can be followed backward so that (4.9) is really equivalent to (4.7c) and (4.7e).

In exactly the same way we obtain an equation equivalent to (4.7d) and (4.7f):

$$i\,\gamma^3\gamma^1\begin{pmatrix}Y\\U\end{pmatrix} = \begin{pmatrix}Y\sigma_2\\U\sigma_2\end{pmatrix} \qquad (4.10)$$

Now let's come back to the six left-out equations in (4.4). Substituting in them the expressions of X and Z given in (4.5) and (4.6) we obtain:

$$({A^i}^2 + B^iC^i + I)Y + (A^iB^i + B^iD^i)U = 0 \quad \text{(no sum)}$$

$$({D^i}^2 + C^iB^i + I)U + (D^iC^i + C^iA^i)Y = 0 \quad \text{(no sum)}$$

$$(A^oA^i + B^oC^i - A^i)Y + (A^oB^i + B^oD^i - B^i)U = 0$$

$$(C^oA^i + D^oC^i - C^i)Y + (C^oB^i + D^oD^i - D^i)U = 0$$

$$(A^o + I)Y + B^oU = 0$$

$$C^oY + (D^o + I)U = 0$$

From (4.8) we see that the factors multiplying Y and U in the first two equations are 0. Thus they are trivially satisfied and contain no information. The next two can be rewritten as

$$[\gamma^{o}\gamma^{i} - \gamma^{i}] \binom{Y}{U} = 0 \qquad (4.11')$$

and the last two as

$$[\gamma^{o} + I] \binom{Y}{U} = 0 \qquad (4.11)$$

But these equations are one and the same as (4.11') is $-\gamma^{i}$ times (4.11).

Let us summarize the results obtained so far. We found that Y and U are determined by (4.9), (4.10) and (4.11) and then X and Z follow from (4.5) and (4.6).

If we write $\binom{Y}{U} = (u\ v)$, where u and v are column vectors, the equations (4.9), (4.10) and (4.11) translate to:

$$i\ \gamma^{1}\gamma^{2}\ (u\ v) = (u\ -v)$$

$$i\ \gamma^{3}\gamma^{1}\ (u\ v) = (iv\ -iu)$$

$$(\gamma^{o} + I)(u\ v) = (0\ 0)$$

This system is equivalent to

$$(\gamma^1 + i\,\gamma^2)u = 0$$

$$(I + \gamma^o)u = 0$$

$$v = \gamma^3\gamma^1\, u$$

Similarly if we set $\binom{X}{Z} = (s\ t)$, where s, t are column vectors, we find that (4.5) and (4.6) can be written as:

$$(s\ t) = \gamma^1\ (v\ u)$$

Whence we may write finally:

$$S = \begin{pmatrix} X & Y \\ Z & U \end{pmatrix} = (s\ t\ u\ v)$$

$$= (\gamma^3 u,\ \gamma^1 u,\ u,\ \gamma^3\gamma^1 u)$$

where u is determined by

$$(\gamma^1 + i\,\gamma^2)u = 0$$

$$(I + \gamma^o)u = 0$$

As one might expect, once we know how to find S, it becomes a simple matter to find T (which has to be a multiple of S^{-1}).

The equations satisfied by T are

$$\Gamma^\mu\ T = T\ \gamma^\mu$$

Taking the adjoint and remembering that $\Gamma^{o^+} = \Gamma^o$ while $\Gamma^{i^+} = -\Gamma^i$, we obtain:

$$\gamma^{o^+} \cdot T^+ = T^+ \Gamma^o$$

$$-\gamma^{i^+} \Gamma^+ = T^+ \Gamma^i$$

But $\{\gamma^{o^+}, -\gamma^{i^+}\}$ is also a system of γ matrices. Hence the last two equations are just like the equations for S. Therefore using the solution just obtained for S we conclude:

$$T^+ = (-\gamma^{3^+} w^+, -\gamma^{1^+} w^+, w^+, \gamma^{3^+} \gamma^{1^+} w^+)$$

where w^+ is determined by:

$$(-\gamma^{1^+} - i\, \gamma^{2^+})w^+ = 0$$

$$(I + \gamma^{o^+})w^+ = 0$$

More conveniently we may now write:

$$T = \begin{bmatrix} -w\ \gamma^3 \\ -w\ \gamma^1 \\ w \\ w\ \gamma^1\gamma^3 \end{bmatrix}$$

where w is a row vector determined by

$$w(\gamma^1 - i\, \gamma^2) = 0$$

$$w(I + \gamma^o) = 0$$

We also asserted at the beginning that the equations $(\gamma^1 + i\,\gamma^2)u = 0$ and $(I + \gamma^o)u = 0$ determine u up to a multiplicative factor. To see that this is true one first checks that the statement is correct when γ^μ is simply Γ^μ. This is trivial and we don't do it here. Now an arbitrary system $\{\gamma^\mu\}$ is related to $\{\Gamma^\mu\}$ by a similarity transformation. Whence the equations $(\gamma^1 + i\,\gamma^2)u = (I + \gamma^o)u = 0$ can be viewed as the equations $(\Gamma^1 + i\,\Gamma^2)u = (I + \Gamma^o)u = 0$ formulated in another basis. Accordingly if the solutions of the second system form a one-dimensional subspace, so will the solutions of the first system. A similar comment applies to w. This completes the proof.

(2) Application to the equation $\Omega^\mu{}_\nu\gamma^\nu = S^{-1}\,\gamma^\mu\,S$.

The customary way of solving the equation $\Omega^\mu{}_\nu\gamma^\nu = S^{-1}\,\gamma^\mu\,S$ for S uses infinitesimal transformations. One puts suitable constraints on S and shows that the corresponding solutions form a group, a "double-valued representation" of the orthochronous Lorentz group. The matrices S are then obtained by "exponentiating" the Lie algebra of this group. We shall discuss this method in detail later on and in particular use it to show that the group of the S matrices is isomorphic to a certain very concrete group. For the moment we want to show how the results of the last section can be used to find S explicitely.

This method definitely lacks the elegance of the one using infinitesimal transformations but it has the advantage of being purely algebraic.

Our aim is to find $S(\Omega)$ such that

$$\hat{\gamma}^{\mu} \equiv \Omega^{\mu}{}_{\nu}\gamma^{\nu} = S^{-1}\,\gamma^{\mu}\,S \quad .$$

The set $\{\gamma^{\mu}\}$ is fixed. By the fundamental theorem there exist M and $V(\Omega)$ such that:

$$\gamma^{\rho} = M^{-1}\,\Gamma^{\rho}\,M \qquad \Omega^{\mu}{}_{\rho}\Gamma^{\rho} = V^{-1}\,\Gamma^{\mu}\,V$$

From these equations we deduce:

$$\Omega^{\mu}{}_{\rho}\gamma^{\rho} = \Omega^{\mu}{}_{\rho}M^{-1}\Gamma^{\rho}M = M^{-1}V^{-1}\Gamma^{\mu}VM = M^{-1}V^{-1}M\gamma^{\mu}M^{-1}VM \quad .$$

Thus the problem of finding $S(\Omega)$ for an arbitrary Ω is reduced to that of finding $V(\Omega)$:

$$S(\Omega) = M^{-1}\,V(\Omega)\,M \qquad (4.12)$$

In the preceding section we have established that $V(\Omega)$ is given by:

$$V = \begin{bmatrix} -w\;\hat{\Gamma}^{3} \\ -w\;\hat{\Gamma}^{1} \\ w \\ w\;\hat{\Gamma}^{1}\hat{\Gamma}^{3} \end{bmatrix} \quad , \qquad \hat{\Gamma}^{\mu} \equiv \Omega^{\mu}{}_{\nu}\Gamma^{\nu}$$

where w is a solution of

$$w(\hat{\Gamma}^1 - i\,\hat{\Gamma}^2) = 0 \qquad (4.13)$$

$$w(I + \hat{\Gamma}^o) = 0 \qquad (4.14)$$

So all we have to do is to find w. To this end we decompose it in the following way:

$$w = (u,v) = (u_1, u_2, v_1, v_2) \qquad (4.15)$$

Equation (4.13) then reads:

$$(u,v)\left[\begin{array}{c|c} c_o\sigma^o & c_k\sigma_k \\ \hline -c_k\sigma_k & -c_o\sigma^o \end{array}\right] = 0$$

where $c_\mu \equiv \Omega^1{}_\mu - i\Omega^2{}_\mu$ (4.16)

Written explicitely this gives:

$$c_o u - v\, c_k\, \sigma_k = 0$$

$$uc_k\sigma_k - c_o v = 0$$

There are two cases to be considered: $c_o \neq 0$ and $c_o = 0$. When $c_o \neq 0$, the general solution is

$$(u,v) = (u, \frac{1}{c_o} uc_k\, \sigma_k) \quad , \quad u \text{ arbitrary} \qquad (4.17)$$

When $c_o = 0$, the general solution is $w = (u,v)$ where

$$uc_k\, \sigma_k = vc_k\, \sigma_k = 0 \qquad (4.18)$$

In either case the solutions form a two dimensional space. Next we come to equation (4.14). It reads:

$$(u,v)\begin{bmatrix}(1+\Omega^{o}{}_{o})\sigma^{o} & \Omega^{o}{}_{k}\sigma_{k} \\ -\Omega^{o}{}_{k}\sigma_{k} & (1-\Omega^{o}{}_{o})\sigma^{o}\end{bmatrix} = 0$$

or ,

$$u(1+\Omega^{o}{}_{o}) - v\,\Omega^{o}{}_{k}\sigma_{k} = 0$$

$$v\,\Omega^{o}{}_{k}\sigma_{k} + v(1-\Omega^{o}{}_{o}) = 0$$

For the sake of simplicity we assume that Ω lies in the orthochronous group: $\Omega^{o}{}_{o} \geq 1 > 0$. The general solution is then:

$$(u,v) = \left(\frac{1}{1+\Omega^{o}{}_{o}}\,v\,\Omega^{o}{}_{k}\sigma_{k},\ v\right) \quad , \quad v \text{ arbitrary} \tag{4.19}$$

Now w must be a common solution to (4.13) and (4.14). For the case $c_o \neq 0$, eqs. (4.17) and (4.19) give:

$$\left(u,\ \frac{u}{c_o}\,\underset{\sim}{c}\cdot\underset{\sim}{\sigma}\right) = \left(\frac{1}{1+\Omega^{o}{}_{o}}\,v\,\underset{\sim}{\Omega}^{o}\cdot\underset{\sim}{\sigma},\ v\right)$$

where we have set $(\Omega^{o}{}_{1},\Omega^{o}{}_{2},\Omega^{o}{}_{3}) = \underset{\sim}{\Omega}^{o}$ and $(c_1,c_2,c_3) = \underset{\sim}{c}$. This gives:

$$v = \frac{u}{c_o}\,(\underset{\sim}{c}\cdot\underset{\sim}{\sigma}) \qquad u = \frac{1}{(1+\Omega^{o}{}_{o})c_o}\,u(\underset{\sim}{c}\cdot\underset{\sim}{\sigma})(\underset{\sim}{\Omega}^{o}\cdot\underset{\sim}{\sigma})$$

The second equation may be rewritten as:

$$u(\underset{\sim}{c}\times\underset{\sim}{\Omega}^{o})\cdot\underset{\sim}{\sigma} = -i\,c_o\,u$$

Upon setting $(\underset{\sim}{c} \times \underset{\sim}{\Omega}^o)_i = a_i$, this equation for u reads

$$(u_3 + ic_o)u_1 + (a_1 + ia_2)u_2 = 0$$

$$(a_1 - ia_2)u_1 + (ic_o - a_3)u_2 = 0$$

If $a_1 + ia_2$ and $a_3 + ic_o$ don't both vanish, the solution is $u = (a_1 + ia_2, -(a_3 + ic_o))$. If they both vanish, the solution is $u = (ic_o - a_3, ia_2 - a_1)$.

In the case where $c_o = 0$, eqs. (4.18) and (4.19) tell us that the solution (u,v) common to both systems must satisfy

$$(u,v) = \left(\frac{1}{1+\Omega^o_o} v \, \underset{\sim}{\Omega}^o \cdot \underset{\sim}{\sigma}, \; v\right) \qquad v \, \underset{\sim}{c} \cdot \underset{\sim}{\sigma} = 0$$

The second equation, when written explicitely, reads:

$$c_3 v_1 + (c_1 + ic_2)v_2 = 0$$

$$(c_1 - ic_2)v_1 - c_3 v_2 = 0$$

If c_3 and $c_1 - ic_2$ don't both vanish, the solution is $v = (c_3, c_1 - ic_2)$. If they both vanish, the solution is $v = (1,0)$.

We now summarize the results. We have defined:

$$\hat{\Gamma}^\mu \equiv \Omega^\mu{}_\nu \Gamma^\nu = V^{-1}(\Omega)\Gamma^\mu V(\Omega), \; \Omega^o{}_o > 0$$

$$c_\mu \equiv \Omega^1{}_\mu - i\Omega^2{}_\mu \quad \underset{\sim}{c} \equiv (c_1, c_2, c_3) \quad \underset{\sim}{\Omega}^o \equiv (\Omega^o{}_1, \Omega^o{}_2, \Omega^o{}_3)$$

$$\underset{\sim}{a} \equiv \underset{\sim}{c} \times \underset{\sim}{\Omega}^o$$

and found that

$$V(\Omega) = \begin{bmatrix} -w\,\hat{\Gamma}^3 \\ -w\,\hat{\Gamma}^1 \\ w \\ w\,\hat{\Gamma}^1\,\hat{\Gamma}^3 \end{bmatrix}$$

where w is to be chosen according to the following table.

Table 3 Vectors for constructing spinor transformations

$c_o \neq 0$		$c_o = 0$	
$w = (u, \frac{u}{c_o}\,\underset{\sim}{c}\cdot\underset{\sim}{\sigma})$		$w = (\frac{1}{1+\Omega^o_o}\,u\,\underset{\sim}{\Omega}^o\cdot\underset{\sim}{\sigma},\ u)$	
$\lvert a_1+ia_2\rvert+\lvert a_3+ic_o\rvert \neq 0$	$\lvert a_1+ia_2\rvert+\lvert a_3+ic_o\rvert = 0$	$\lvert c_3\rvert+\lvert c_1-ic_2\rvert \neq 0$	$\lvert c_3\rvert+\lvert c_1-ic_2\rvert = 0$
$u = (a_1+ia_2, -(a_3+ic_o))$	$u = (ic_o-a_3, ia_2-a_1)$	$u = (c_3, c_1-ic_2)$	$u = (1,0)$

As an illustration we obtain $V(\Omega)$ when Ω is a [illegible] boos[illegible]the x axis. If the frame K' moves with spee[illegible] β [illegible]ect to the frame K, the primed coordinates are [illegible] the unprimed by $x' = \Omega x$ where:

$$\Omega = \left[\begin{array}{cc|c} \gamma & \gamma\beta & \\ -\gamma\beta & \gamma & 0 \\ \hline 0 & & I \end{array}\right] . \qquad \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

In [illegible] case we have: $c_o = -\gamma\beta$, $\underset{\sim}{c} = (\gamma, -i, 0)$, $\underset{\sim}{\Omega}^o = (-\gamma\beta, 0, 0)$. By [illegible]g table 3 we obtain $w = (0, 2i\gamma\beta, -2i(\gamma+1), 0)$ from whic[illegible]

$$V(\Omega) = -2i \begin{bmatrix} \gamma+1 & 0 & 0 & -\gamma\beta \\ 0 & \gamma+1 & -\gamma\beta & 0 \\ 0 & -\gamma\beta & \gamma+1 & 0 \\ -\gamma\beta & 0 & 0 & \gamma+1 \end{bmatrix}$$

Since $V(\Omega)$ is determined up to a multiplicative factor we may drop the $-2i$. The resulting matrix is then:

$$V(\Omega) = (\gamma+1)I - \gamma\beta\, \Gamma^o \Gamma^1$$

It is customary to write $\gamma = ch\rho$ $\gamma\beta = sh\rho$. If we divide our $V(\Omega)$ by $2ch\rho/2$ we obtain the matrix $ch\rho/2 I - sh\rho/2 \Gamma^o\Gamma^1$. This expression agrees with the one obtained by the method of infinitesimal transformations as will be seen later.

CHAPTER V

THE GROUP OF SPINOR TRANSFORMATIONS

(1) Construction of the group

In this section we look at the set of all spinor transformations corresponding to all orthochronous Lorentz transformations. We study this set as a group. We give up the purely algebraic method devised in chapter IV and switch to the standard method using infinitesimal transformations, or the Lie algebras. The orthochronous Lorentz group will be denoted by $L^{\uparrow}$. It consists of all Lorentz transformations Ω with $\Omega^{o}{}_{o} > 0$. The proper subgroup of orthochronous transformations with determinant 1 will be denoted by $L_{+}^{\uparrow}$. We work throughout with a fixed system $\{\gamma^{\mu}\}$ which will be assumed to be unitary. We are interested in the solution $\Lambda(\Omega)$ to the equation

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \quad , \quad \Omega \in L^{\uparrow} \tag{5.1}$$

We already know that the solution Λ is determined up to a multiplicative factor. We want to remove as much as possible this arbitrariness. The first restriction that can be imposed is the following.

Proposition (5.1): The solution of eq. (5.1) may be chosen so as to satisfy $\Lambda^{+} = \gamma^{o}\Lambda^{-1}\gamma^{o}$.

Proof: Let us write $\hat{\gamma}^{\mu} = \Omega^{\mu}{}_{\nu}\gamma^{\nu}$ and let Λ be a solution of eq. (5.1). Then we have:

$$\hat{\gamma}^{\mu +} = \Omega^{\mu}{}_{\nu}\gamma^{\nu +} = \Omega^{\mu}{}_{\nu}\gamma^{0}\gamma^{\nu}\gamma^{0} = \gamma^{0}\hat{\gamma}^{\mu}\gamma^{0} = \gamma^{0}\Lambda^{-1}\gamma^{\mu}\Lambda\gamma^{0}$$

from which $\hat{\gamma}^{\mu} = \gamma^{0}\Lambda^{+}\gamma^{\mu +}\Lambda^{-1 +}\gamma^{0} = (\gamma^{0}\Lambda^{+}\gamma^{0})\gamma^{\mu}(\gamma^{0}\Lambda^{+}\gamma^{0})^{-1} = \Lambda^{-1}\gamma^{\mu}\Lambda$

and $[\Lambda\gamma^{0}\Lambda^{+}\gamma^{0}, \gamma^{\mu}] = 0$. Therefore $\Lambda\gamma^{0}\Lambda^{+}\gamma^{0} = cI$ or

$\Lambda^{+} = c\gamma^{0}\Lambda^{-1}\gamma^{0}$. This implies

$$\Lambda^{+}\Lambda = c\gamma^{0}\Lambda^{-1}\gamma^{0}\Lambda = c\gamma^{0}\hat{\gamma}^{0} = c\gamma^{0}(\Omega^{0}{}_{0}\gamma^{0} + \Omega^{0}{}_{k}\gamma^{k})$$

$$= c\Omega^{0}{}_{0}I + c\Omega^{0}{}_{k}\gamma^{0}\gamma^{k}$$

Taking the trace on both sides and using proposition (2.3) yields $T_r\Lambda^{+}\Lambda = 4c\Omega^{0}{}_{0}$. Since $\Lambda^{+}\Lambda$ is hermitian it follows that c is real. Since $\Lambda^{+}\Lambda$ is positive, it follows that c is positive ($\Omega^{0}{}_{0} > 0$). Now if we take $\Lambda' = \frac{1}{\sqrt{c}}\Lambda$, we obtain:

$$\Lambda'^{+} = \frac{1}{\sqrt{c}}\Lambda^{+} = \frac{1}{\sqrt{c}}c\gamma^{0}\Lambda^{-1}\gamma^{0} = \sqrt{c}\,\gamma^{0}\frac{(\Lambda')^{-1}}{\sqrt{c}}\gamma^{0} = \gamma^{0}(\Lambda')^{-1}\gamma^{0}.$$

The motivation for this restriction is to make the transformation of the adjoint $\bar{\psi} \equiv \psi^{+}\gamma^{0}$ simple:

$$\bar{\psi}' = \psi'^{+}\gamma^{0} = \psi^{+}\Lambda^{+}\gamma^{0} = \bar{\psi}\gamma^{0}\Lambda^{+}\gamma^{0} = \bar{\psi}\Lambda^{-1}.$$

The usefulness of this will be appreciated in chapter VI. Before we go further we need to introduce a special matrix, called the B matrix, associated to any system of γ matrices. Let $\{\gamma^\mu\}$ be such an arbitrary system. The set $\{\gamma^{\mu *}\}$, where * means complex conjugate, is also a system of γ matrices. By the fundamental theorem there exists a matrix B such that $\gamma^{\mu *} = B^{-1}\gamma^\mu B$. We require that $|\det B| = 1$. This determines B up to a phase factor. If $\{\hat{\gamma}^\mu\}$ is another arbitrary system with corresponding $\hat{B}$ and $\hat{\gamma}^\mu = S^{-1}\gamma^\mu S$, then $\hat{B}$ is related to B by:

$$\hat{B} = e^{i\phi}S^{-1}BS^* \quad , \quad \phi \text{ real} \tag{5.2}$$

Indeed we have:

$$\hat{\gamma}^{\mu *} = S^{-1*}\gamma^{\mu *}S^* = S^{-1*}B^{-1}\gamma^\mu BS^*$$

$$= (S^{-1}BS^*)^{-1}\hat{\gamma}^\mu S^{-1}BS^*$$

Moreover $|\det S^{-1}BS^*| = 1$. Since $\hat{B}$ is determined up to a phase, the conclusion follows.

The properties of the B matrix of use to us are contained in the following proposition.

Proposition (5.2): The B matrix of a unitary $\{\gamma^\mu\}$ system is unitary and antisymmetric.

Proof: Since γ^μ is unitary, so is $\gamma^{\mu *}$. By proposition (3.3) there exists a unitary matrix connecting γ^μ and $\gamma^{\mu *}$.

Clearly this can be taken as B and any phase multiple will also be unitary. We now have to show that it is antisymmetric. For the set $\{\Gamma^\mu\}$ of eq. (2.7) one easily checks that one can take $B' = \Gamma^0\Gamma^3\Gamma^1$. Notice that $B'^T = -B'$. By proposition (3.3) there exists a unitary U such that $\gamma^\mu = U^+\Gamma^\mu U$, and by eq. (5.2) one has

$$B = e^{i\phi}U^+B'U^* = e^{i\phi}(U^*)^T B'U^*$$

From this the antisymmetry of B' clearly implies that of B.

We are now prepared to put all the restrictions on Λ.

<u>Theorem</u> (5.1): The equations

$$\Omega^\mu{}_\nu\gamma^\nu = \Lambda^{-1}\gamma^\mu\Lambda \qquad (5.3a)$$

$$\Lambda^+ = \gamma^0\Lambda^{-1}\gamma^0 \qquad (5.3b)$$

$$\Lambda^* = B^+\Lambda B \qquad (5.3c)$$

where Λ is the unknown and $\Omega\varepsilon L^\uparrow$ have exactly two solutions. One is -1 times the other.

Proof: By proposition (5.1) we know that eqs. (5.3a), (5.3b) have a common solution Λ. We have:

$\Omega^\mu{}_\nu\gamma^{\nu *} = \Lambda^{-1*}\gamma^{\mu *}\Lambda^*$ which may be rewritten as

$\Omega^{\mu}{}_{\nu} B^{+} \gamma^{\nu} B = \Lambda^{-1*} B^{+} \gamma^{\mu} B \Lambda^{*}$, from which we deduce

$\Lambda^{-1} \gamma^{\mu} \Lambda = \Omega^{\mu}{}_{\nu} \gamma^{\nu} = B \Lambda^{-1*} B^{+} \gamma^{\mu} B \Lambda^{*} B^{+}$ or

$$[\gamma^{\mu}, B \Lambda^{*} B^{+} \Lambda^{-1}] = 0$$

By theorem (2.8) it follows as usual that $B \Lambda^{*} B^{+} \Lambda^{-1} = cI$, which implies $\frac{(\det \Lambda)^{*}}{\det \Lambda} = c^{4}$. Whence $|c| = 1$, or $c = e^{i\lambda}$. From $B \Lambda^{*} B^{+} \Lambda^{-1} = cI$, we get $\Lambda^{*} = e^{i\lambda} B^{+} \Lambda B$. So if we define $\Lambda' = e^{i\alpha} \Lambda$, we obtain:

$$\Lambda'^{*} = e^{-i\alpha} \Lambda^{*} = e^{i(\lambda - 2\alpha)} B^{+} \Lambda' B$$

and the choice $\alpha = \lambda/2$ yields the solution to our three equations. Clearly if Λ is a solution, so is $-\Lambda$. Now suppose that $\tilde{\Lambda}$ is another solution to (5.3a,b,c). We know that $\tilde{\Lambda} = a\Lambda$ for some complex number a. Equation (5.3c) shows that a is real. Eq. (5.3b) then shows that its square is 1. Whence $\tilde{\Lambda} = \pm\Lambda$. This completes the proof.

For some purposes, especially when dealing with the Lie algebras, it is convenient to reexpress the three equations of theorem (5.2) in only two. This is achieved by the following result:

<u>Proposition</u> (5.3): Equations (5.3a,b,c) are equivalent to the two equations:

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \qquad (5.4a)$$

$$\Lambda^{T}(B^{+}\gamma^{o})\Lambda = B^{+}\gamma^{o} \qquad (5.4b)$$

Proof: We first show that (5.3) implies (5.4). Of course (5.3) implies (5.4a). Next we have, using (5.3b,c),

$$\Lambda^{T} = \gamma^{o*}\Lambda^{-1*}\gamma^{o*} = \gamma^{o*}B^{+}\Lambda^{-1}B\gamma^{o*} = B^{+}\gamma^{o}BB^{+}\Lambda^{-1}BB^{+}\gamma^{o}B$$
$$= B^{+}\gamma^{o}\Lambda^{-1}\gamma^{o}B$$

from which (5.4b) follows.

Now we have to show that (5.4) implies (5.3), that is, if $\bar{\Lambda}$ is a solution of (5.4a,b) then it satisfies (5.3b,c). Let Λ be a solution of (5.3). From what we've just seen, Λ is also a solution of (5.4). Therefore we may write:

$$\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o} = \Lambda^{-1} \quad , \quad \text{from which}$$

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda = \gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda$$

Similarly we have for $\bar{\Lambda}$:

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \bar{\Lambda}^{-1}\gamma^{\mu}\bar{\Lambda} = \gamma^{o}B\bar{\Lambda}^{T}B^{+}\gamma^{o}\gamma^{\mu}\bar{\Lambda}$$

From the fundamental theorem we know that $\bar{\Lambda} = c\Lambda$ and the last two equations for $\Omega^{\mu}{}_{\nu}\gamma^{\nu}$ give:

$$\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda = c^{2}\gamma^{o}B\Lambda^{T}B^{+}\gamma^{o}\gamma^{\mu}\Lambda$$

or $c^2 = 1$. Whence $\bar{\Lambda} = \pm\Lambda$ and $\bar{\Lambda}$ is a solution of (5.3).

Equations (5.4a) and (5.4b) show that Λ can be interpreted as the transition matrix of a change of basis under which the matrices of the operators γ^μ become $\Omega^\mu{}_\nu\gamma^\nu$ and the matrix of the bilinear form defined by $B^+\gamma^0$ remains invariant. When we have gained some more information about our Λ matrices we will give a third completely different way of formulating (5.3a,b,c).

We now introduce some handy notation. The set of all Λ's solutions of (5.3a,b,c) when Ω goes over $L^\dagger$ will be denoted by $S^\dagger$. Given $\Lambda \epsilon S^\dagger$, there is only one Ω in $L^\dagger$ such that Λ is a solution of (5.3a): this follows from the linear independence of the γ^μ's. The one Ω corresponding to Λ will be denoted by Ω_Λ and the map $\Lambda \to \Omega_\Lambda$ will be denoted by Π. Our first statement about the Λ's is the following.

<u>Theorem</u> (5.2): $S^\dagger$ is a six-dimensional Lie group locally isomorphic to $L^\dagger$. Π: $S^\dagger \to L^\dagger$: $\Omega \to \Omega_\Lambda$ is a homomorphism and a local isomorphism.

Proof: We first prove that $S^\dagger$ is a group. Let $\Lambda, \bar{\Lambda}$ be solutions of (5.4a) and (5.4b) with corresponding $\Omega, \bar{\Omega}$. Then we have:

$$\bar{\Omega}^\mu{}_\rho\Omega^\rho{}_\nu\gamma^\nu = \bar{\Omega}^\mu{}_\rho\Lambda^{-1}\gamma^\rho\Lambda = \Lambda^{-1}\bar{\Lambda}^{-1}\gamma^\mu\bar{\Lambda}\Lambda = (\bar{\Lambda}\Lambda)^{-1}\gamma^\mu(\bar{\Lambda}\Lambda)$$

Hence $\bar{\Lambda}\Lambda$ is a solution of (5.4a) corresponding to [illegible]

Moreover $\Lambda\Lambda'$ clearly satisfies (5.4b). Thus $S^\dagger$ is closed under multiplication. To show that Λ^{-1} belongs to $S^\dagger$ if Λ does we first observe that $(\Omega^{-1})^\mu{}_\nu = g^{\mu\rho}\Omega_\rho{}^\sigma g_{\sigma\nu}$. Whence, if Λ satisfies (5.4a), we have:

$$\begin{aligned}(\Omega^{-1})^\mu{}_\nu\gamma^\nu &= g^{\mu\rho}\Omega_\rho{}^\sigma g_{\sigma\nu}\gamma^\nu \\ &= g^{\mu\rho}\Omega_\rho{}^\sigma g_{\sigma\nu}\Omega^\nu{}_\theta\Lambda\gamma^\theta\Lambda^{-1} \\ &= g^{\mu\rho}g_{\rho\theta}\Lambda\gamma^\theta\Lambda^{-1} = \Lambda\gamma^\mu\Lambda^{-1}\end{aligned}$$

Whence Λ^{-1} is a solution of (5.4a) corresponding to $\Omega^{-1}\epsilon S^\dagger$. To show that $(\Lambda^{-1})^T B^+\gamma^o\Lambda^{-1} = B^+\gamma^o$ one simply multiplies eq. (5.4b) on the left by $(\Lambda^{-1})^T$ and on the right by Λ^{-1}. These considerations show at the same time that Π is an algebraic homomorphism. We now know that $S^\dagger$ is an algebraic subgroup of the Lie group GL(4,C) of 4×4 invertible complex matrices. To show that it is also a Lie subgroup we simply invoke the well-known Cartan theorem which says that an algebraic subgroup of a Lie group which is also a topologically closed subset is a Lie subgroup. It is easy to show that $S^\dagger$ is closed in GL(4,C). Let (Λ_n) be a sequence of elements of $S^\dagger$ converging to $L\epsilon GL(4,C)$. We want to show that L is in $S^\dagger$. It is clear by continuity that L satisfies (5.4b).

Therefore, all we have to do is to show that there exists Ω in $L^{\uparrow}$ such that (5.4a) holds with $\Lambda = L$. We have $(\Omega_{\Lambda_n})^{\mu}{}_{\nu} = \frac{1}{4} T_r(\Lambda_n^{-1} \gamma^{\mu}\Lambda_n\gamma_{\nu})$. Since $L^{\uparrow}$ is closed in GL(4,R), the group of 4×4 real invertible matrices, $\Omega^{\mu}{}_{\nu} \equiv \lim_{n\to\infty} (\Omega_{\Lambda_n})^{\mu}{}_{\nu}$ defines an orthochronous Lorentz transformation and by continuity one clearly has $\Omega^{\mu}{}_{\nu}\gamma^{\nu} = L^{-1}\gamma^{\mu}L$. Whence L belongs to $S^{\uparrow}$ and $S^{\uparrow}$ is a Lie group. We now want to show that $\Pi: \Lambda \to \Omega_{\Lambda}$ is a local isomorphism. We already know that it is a homomorphism. The fact that it is smooth follows clearly from $\Omega_{\Lambda}{}^{\mu}{}_{\nu} = \frac{1}{4} T_r(\Lambda^{-1}\gamma^{\mu}\Lambda\gamma_{\nu})$. So all that remains to be shown is that it is locally injective. Let $\Omega \in L^{\uparrow}$ and $\Lambda \in S^{\uparrow}$ such that $\Omega_{\Lambda} = \Omega$. There is only one other solution L to $\Omega_L = \Omega$ and it is $L = -\Lambda$. So all we have to do is to take a small neighborhood U_{Λ} of Λ such that $-U_{\Lambda} \equiv \{-\Lambda;\ \Lambda\epsilon U_{\Lambda}\}$ and U_{Λ} don't intersect: Π restricted to U_{Λ} is clearly one to one.

(2) <u>Explicit form of the group elements</u>

We now proceed to obtain explicit expressions for the Λ's via infinitesimal transformations. Before we obtain the transformation laws of the spinors under the full orthochronous group $L^{\uparrow}$, we first investigate the subgroup of proper Lorentz transformations $L_+^{\uparrow}$. The corresponding Λ's obviously form a Lie-subgroup of $S^{\uparrow}$ which we will denote by

$S^{\dagger}_{+}$. It is clear that $S^{\dagger}_{+}$ is both open and closed in $S^{\dagger}$. Hence $S^{\dagger}$ is disconnected. We will show later that it has two components just as $L^{\dagger}$. We use the following standard basis of the Lie algebra of $L^{\dagger}_{+}$:

$$(I_{\alpha\beta})^{\mu}{}_{\nu} \equiv \delta^{\mu}{}_{\alpha} g_{\nu\beta} - \delta^{\mu}{}_{\beta} g_{\nu\alpha} \tag{5.5}$$

One has $I_{\alpha\beta} = -I_{\beta\alpha}$ and a basis is obtained by taking $I_{01}, I_{02}, I_{03}, I_{12}, I_{13}, I_{23}$. The first three generate boosts; the last three generate rotations. The commutators are given by

$$[I_{\mu\nu}, I_{\kappa\lambda}] = -(g_{\mu\kappa} I_{\nu\lambda} + g_{\lambda\mu} I_{\kappa\nu} + g_{\nu\lambda} I_{\mu\kappa} + g_{\kappa\nu} I_{\lambda\mu}) \tag{5.6}$$

To obtain the Lie algebra of $S^{\dagger}_{+}$ we simply use our local isomorphism $\Lambda \to \Omega_{\Lambda}$. If $\Omega(\tau)$ is a curve in $L^{\dagger}_{+}$ passing through I at $\tau = 0$, there is a unique curve $\Lambda(\tau)$ in $S^{\dagger}_{+}$ passing through I at $\tau = 0$ such that:

$$\Omega^{\mu}{}_{\nu}(\tau)\gamma^{\nu} = \Lambda^{-1}(\tau)\gamma^{\mu}\Lambda(\tau)$$

$$\Lambda^{T}(\tau) B^{+}\gamma^{0}\Lambda(\tau) = B^{+}\gamma^{0}$$

Taking the derivative of these equations at $\tau = 0$ gives:

$$[\dot{\Lambda}, \gamma^{\mu}] = -\dot{\Omega}^{\mu}{}_{\nu}\gamma^{\nu}$$

$$\dot{\Lambda}^{T} B^{+}\gamma^{0} + B^{+}\gamma^{0}\dot{\Lambda} = 0$$

where we have set $\dot{\Lambda} \equiv \dot{\Lambda}(o)$, $\dot{\Omega}^{\mu}{}_{\nu} \equiv \dot{\Omega}^{\mu}{}_{\nu}(o)$. Now if the curve $\Omega(\tau)$ is chosen so that $\dot{\Omega} = I_{\alpha\beta}$ and we denote the corresponding $\dot{\Lambda}$ by $\dot{\Lambda}_{\alpha\beta}$, the equations for $\dot{\Lambda}$ read:

$$[\dot{\Lambda}_{\alpha\beta}, \gamma^{\mu}] = -(\delta^{\mu}{}_{\alpha} g_{\nu\beta} - \delta^{\mu}{}_{\beta} g_{\nu\alpha})\gamma^{\nu} = \delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta} \tag{5.7}$$

$$\dot{\Lambda}^{T}{}_{\alpha\beta} B^{+}\gamma^{o} + B^{+}\gamma^{o}\dot{\Lambda}_{\alpha\beta} = 0 \tag{5.8}$$

From table 2 we know that $[\gamma_{[\alpha\beta]}, \gamma^{\mu}] = 2(\delta^{\mu}{}_{\beta}\gamma_{\alpha} - \delta^{\mu}{}_{\alpha}\gamma_{\beta})$. Using this and eq. (5.7) we get $[\dot{\Lambda}_{\alpha\beta} - \frac{1}{2}\gamma_{[\alpha\beta]}, \gamma^{\mu}] = 0$, from which $\dot{\Lambda}_{\alpha\beta} = \frac{1}{2}\gamma_{[\alpha\beta]} + cI$. Inserting this in (5.8) yields:

$$0 = 2cB^{+}\gamma^{o} + \frac{1}{2}(\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} + B^{+}\gamma^{o}\gamma_{[\alpha\beta]}).$$ But

$$\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} = \gamma_{\beta}{}^{T}\gamma_{\alpha}{}^{T}B^{+}\gamma^{o} = \gamma^{\beta *}\gamma^{\alpha *}B^{+}\gamma^{o} = B^{+}\gamma^{\beta}\gamma^{\alpha}\gamma^{o}$$ and

$$\gamma_{[\alpha\beta]}{}^{T}B^{+}\gamma^{o} + B^{+}\gamma^{o}\gamma_{[\alpha\beta]} = B^{+}(\gamma^{\beta}\gamma^{\alpha}\gamma^{o} + \gamma^{o}\gamma_{[\alpha\beta]}) = 0$$

Therefore $c = 0$. Whence we may write:

$$\dot{\Lambda}_{\alpha\beta} = \frac{1}{2}\gamma_{[\alpha\beta]} \tag{5.9}$$

$$\Pi(\exp \frac{c^{\alpha\beta}}{2}\gamma_{[\alpha\beta]}) = \exp(c^{\alpha\beta}I_{\alpha\beta}) \tag{5.10}$$

Our first use of formula (5.10) is to get explicit forms. We first consider the transformation of spinors under rotations. When one rotates a frame through an angle ϕ around a unit vector $\underset{\sim}{a}$, the coordinates x' in the new frame are related to the old coordinates $\underset{\sim}{x}$ by

$$x' = [\exp - \phi(\underset{\sim}{n}\cdot\underset{\sim}{A})]\,x \qquad \text{where}$$

$$A_1 \equiv I_{23} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_2 \equiv I_{31} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix},$$

$$A_3 \equiv I_{12} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (5.11)$$

It is convenient at this stage to introduce the matrices Σ_i, defined by

$$\Sigma_1 = \sigma_{23} \quad , \quad \Sigma_2 = \sigma_{31} \quad , \quad \Sigma_3 = \sigma_{12} \qquad (5.12)$$

where, as we recall, $\sigma_{\mu\nu} \equiv i\gamma_{[\mu\nu]}$. One easily checks that they satisfy the same relations as the Pauli matrices:

$$\Sigma_i \Sigma_j = i\epsilon_{ijk}\Sigma_k + \delta_{ij} I \qquad (5.13)$$

from which one deduces:

$$(\underset{\sim}{a}\cdot\underset{\sim}{\Sigma})(\underset{\sim}{b}\cdot\underset{\sim}{\Sigma}) = (\underset{\sim}{a}\cdot\underset{\sim}{b})I + i(\underset{\sim}{a}\times\underset{\sim}{b})\cdot\underset{\sim}{\Sigma} \qquad (5.14)$$

We now apply eq. (5.10) with $c^{\alpha\beta} I_{\alpha\beta} = -\phi\, \underset{\sim}{n}\cdot\underset{\sim}{A}$. The corresponding Λ on the left-hand side is $\exp(i\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})$. This is easily evaluated. Using (5.14) we see that $(\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p = I$ and $(\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p = \underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$ for p even and odd respectively. Hence:

$$\exp i\,\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma} = \sum_{p=0}^{\infty}\frac{(i\,\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})^p}{p!}$$

$$= \sum_{k=0}^{\infty}\frac{(i)^{2k}(\frac{\phi}{2})^{2k} I}{(2k)!} + \sum_{k=0}^{\infty}\frac{(i)^{2k+1}(\frac{\phi}{2})^{2k+1}}{(2k+1)!}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$$

$$= \cos\frac{\phi}{2}\,I + i\sin\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$$

Whence if we denote by $\Lambda_{\underset{\sim}{n}}(\phi)$ the transformation corresponding to a rotation of angle ϕ around the unit vector $\underset{\sim}{n}$ we obtain:

$$\Lambda_{\underset{\sim}{n}}(\phi) = \cos\frac{\phi}{2}\,I + i\sin\frac{\phi}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma} \tag{5.15}$$

One peculiar feature of this equation is that $\Lambda_{\underset{\sim}{n}}(2\pi) = -I$. This is characteristic of spin $\frac{1}{2}$ wave functions. The spin operator along direction $\underset{\sim}{n}$ is the infinitesimal generator of the unitary group of transformations of the internal variables under rotations around $\underset{\sim}{n}$. We have seen above that this group of transformations is given by $\exp\,(i\,\phi\,\frac{1}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma})$. Whence the spin operator along direction $\underset{\sim}{n}$ is $\frac{1}{2}\,\underset{\sim}{n}\cdot\underset{\sim}{\Sigma}$. Using the eq. (5.13) we find indeed that the spin vector operator

$$\underset{\sim}{S} = \frac{1}{2}\,\underset{\sim}{\Sigma} \tag{5.16}$$

satisfies the characteristic commutation relations of an angular momentum:

$$[S_i, S_j] = i\varepsilon_{ijk}S_k \tag{5.17}$$

We have $S_3 = \frac{i}{2}\gamma_1\gamma_2 = \frac{i}{2}\gamma^1\gamma^2$. By the fundamental theorem $\frac{i}{2}\gamma^1\gamma^2$ is similar to $\frac{i}{2}\Gamma^1\Gamma^2$ (cf. eq. (2.7)). But $i\Gamma^1\Gamma^2 = \frac{1}{2}\begin{bmatrix}\sigma_3 & 0\\ 0 & \sigma_3\end{bmatrix}$. Whence the eigenvalues of S_3 are $\pm\frac{1}{2}$: $\underset{\sim}{S}$ is a spin $\frac{1}{2}$ operator. Of course this corresponds to the experimental spin of the electron. As far as physics is concerned the bizarre result $\Lambda_{\underset{\sim}{n}}(2\pi) = -I$ is of no consequence. All physical quantities arise through quadratic expressions which are insensitive to the interchange of ψ and $-\psi$. From a mathematical point of view the result shows that we cannot remove the sign arbitrariness of the elements of $S^{\uparrow}$ without giving up their property of forming a Lie group: $-I$ is obtained by exponentiating an element of the Lie algebra.

Having discussed rotations and spin we now turn to the transformation of spinors under boosts. Suppose that the frame K', coinciding with the frame K at $t = 0$, moves at speed $\beta\underset{\sim}{n}$ with respect to it. (Here $\underset{\sim}{n}$ is a unit vector and $\beta \equiv v/c = v$ with our choice of units.) Let $\rho = \text{th}^{-1}(\beta)$. Then the primed coordinates are related to the unprimed by $x' = \Omega_{\underset{\sim}{n}}(\rho)x$ where $\Omega_{\underset{\sim}{n}}(\rho) = \exp \rho\underset{\sim}{n}\cdot\underset{\sim}{I}_0$. Here we have set:

$$\underset{\sim}{I}_0 \equiv (I_{01}, I_{02}, I_{03}) \quad \text{(cf. (5.[illegible]))} \qquad \text{([illegible])}$$

so that

$$\underset{\sim}{n}\cdot\underset{\sim}{I}_0 = -\begin{bmatrix}0 & \underset{\sim}{n}^T\\ \underset{\sim}{n} & 0\end{bmatrix} \qquad \text{([illegible])}$$

Let $\underset{\sim}{\gamma} \equiv (\gamma^1,\gamma^2,\gamma^3)$. One easily checks the following product rule:

$$(\underset{\sim}{a}\cdot\underset{\sim}{\gamma})(\underset{\sim}{b}\cdot\underset{\sim}{\gamma}) = -[\underset{\sim}{a}\cdot\underset{\sim}{b}\ I + i\ \underset{\sim}{a}\times\underset{\sim}{b}\ \cdot\underset{\sim}{\Sigma}] \qquad (5.20)$$

We now apply eq. (5.10) with $c^{\alpha\beta}I_{\alpha\beta} = \rho\underset{\sim}{n}\cdot\underset{\sim}{I}_o$. The correspon-ding Λ on the left-hand side of (5.10) is $\exp(\frac{1}{2}\rho\ n^i\ \gamma_{[oi]})$. Using eq. (5.20) we obtain:

$$(n^i\ \gamma_{[oi]})^2 = -(\underset{\sim}{n}\cdot\underset{\sim}{\gamma})^2 = I\ .$$

Whence we have:

$$\exp(\tfrac{1}{2}\rho\ n^i\ \gamma_{[oi]}) = \sum_{p=0}^{\infty} \frac{(\frac{1}{2}\rho)^p}{p!}(n^i\gamma_{[oi]})^p$$

$$= \sum_{k=0}^{\infty}\frac{(\frac{\rho}{2})^{2k}}{(2k)!} I + \sum_{k=0}^{\infty}\frac{(\frac{\rho}{2})^{2k+1}}{(2k+1)!}(n^i\ \gamma_{[oi]})$$

$$= \mathrm{ch}\frac{\rho}{2}\ I + \mathrm{sh}\frac{\rho}{2}\ n^i\ \gamma_{[oi]}$$

Or, using the standard notation of eq. (1.3b):

$$\Lambda_{\underset{\sim}{n}}(\rho) = \mathrm{ch}\frac{\rho}{2}\ I - \mathrm{sh}\frac{\rho}{2}\ \underset{\sim}{n}\cdot\underset{\sim}{\alpha} \qquad (5.21)$$

One may check that this agrees with the expression obtained in chapter IV, section (2) with $\underset{\sim}{n} = (1,0,0)$.

An arbitrary proper Lorentz transformation Ω decomposes uniquely into the product of a boost and a rotation:

$$\Lambda = \exp-(\phi \underset{\sim}{n}\cdot\underset{\sim}{A}) \exp(\rho \underset{\sim}{u}\cdot\underset{\sim}{I}_o) \quad , \quad \underset{\sim}{u}: \text{unit vector.}$$

Therefore the transformation of spinors under the proper Lorentz group $L_+^\uparrow$ is completely described by the formula:

$$\Pi[(\cos\tfrac{\phi}{2} I + i \sin\tfrac{\phi}{2} \underset{\sim}{n}\cdot\underset{\sim}{\Sigma})(ch\tfrac{\rho}{2} I - sh\tfrac{\rho}{2} \underset{\sim}{n}\cdot\underset{\sim}{\alpha})] \quad .$$

$$= \exp - \phi \underset{\sim}{n}\cdot\underset{\sim}{A} \exp \rho \underset{\sim}{u}\cdot\underset{\sim}{I}_o \tag{5.22}$$

To give the transformation law of the spinors under an arbitrary element of the full orthochronous group $L^\uparrow$ it is now sufficient to say how they transform under the space reflection $\Omega_s = \begin{bmatrix} 1 & 0 \\ 0 & -I \end{bmatrix}$, because any Ω in $L^\uparrow$ which is not in $L_+^\uparrow$ can be written as $\Omega = \Omega_s\Omega_+$, where Ω_+ is in $L_+^\uparrow$. So let us find the Λ_s corresponding to Ω_s. Equation (5.4a) gives $\gamma^o = \Lambda_s^{-1}\gamma^o\Lambda_s$, $-\gamma^k = \Lambda_s^{-1}\gamma^k\Lambda_s$, from which it immediately follows that $\Lambda_s = c\gamma^o$. To determine c we use eq. (5.4b) which says: $c^2\gamma^{o*}B^+\gamma^o\gamma^o = B^+\gamma^o$ or $c^2B^+\gamma^o = B^+\gamma^o$. Whence $c = \pm 1$. Therefore:

$$\Lambda_s = \pm \gamma^o \tag{5.23}$$

The group $S^\uparrow$ that we have been considering arose through the transformations of the spinors under orthochronous Lorentz transformations. The reason we didn't consider the full Lorentz group is that eq. (5.4b) cannot be satisfied for time reversal $\Omega_t = \begin{bmatrix} -1 & 0 \\ 0 & I \end{bmatrix}$. For this Ω_t, eq. (5.4a) says that the corresponding Λ must anticommute with γ^o and commute with $\underset{\sim}{\gamma}$. The solution to this is $c\gamma^o\gamma^5$ and one easily checks that this cannot satisfy (5.4b).

(3) Charge conjugation and time reversal

Besides those induced by orthochronous Lorentz transformations, there are two other important types of spinor transformations: charge conjugation and time reversal.

Suppose that a given representation $\{\gamma^\mu\}$ has been chosen. Let B be the matrix of proposition (5.2): $\gamma^{\mu *} = B^{-1}\gamma^\mu B$. Then the charge-conjugate spinor ψ^c of the spinor ψ is defined by

$$\psi^c = \gamma^5 B\psi^*$$

where ψ^* is the complex conjugate of ψ. Let us denote by K the antiunitary operator of complex conjugation in that particular representation. Then $\psi^c = \gamma^5 BK\psi \equiv K_c\psi$. The operator BK commutes with γ^μ. Indeed we have:

$$\gamma^\mu BK\psi = \gamma^\mu B\psi^* = B\gamma^{\mu *}\psi^* = BK\gamma^\mu\psi$$

Suppose ψ satisfies the Dirac equation

$$[\gamma^\mu(i\partial_\mu - eA_\mu) - m]\psi = 0$$

Multiplying this on the left by $\gamma^5 BK$ and remembering that BK is an antilinear operator commuting with γ^μ we get

$$[\gamma^{\mu}(i\partial_{\mu} + eA_{\mu}) - m]\psi^{c} = 0$$

The only difference between the equations satisfied by ψ and ψ^c is the sign of the charge multiplying the vector potential. Whence ψ^c can be considered as the wave function of a particle of the same mass m but opposite charge $-e$ in the same electromagnetic field A_μ. This interpretation is consistent with the easily verified equation

$$K_c^{\,2} = I \quad .$$

The other type of transformation is time reversal. This operation will be first defined for the electromagnetic field. Consider a classical electromagnetic field $\underset{\sim}{E}$, $\underset{\sim}{B}$. It satisfies Maxwell's equations:

$$\nabla\cdot\underset{\sim}{E} = 4\pi\rho \qquad\qquad \nabla\cdot\underset{\sim}{B} = 0$$

$$\nabla\times\underset{\sim}{B} = \frac{4\pi}{c}\underset{\sim}{J} + \frac{1}{c}\frac{\partial\underset{\sim}{E}}{\partial t} \qquad\qquad \nabla\times\underset{\sim}{E} = -\frac{1}{c}\frac{\partial\underset{\sim}{B}}{\partial t}$$

Let us define new vector fields $\overline{\underset{\sim}{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$ and $\overline{\underset{\sim}{B}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{B}(-t,\underset{\sim}{x})$. If $\underset{\sim}{E}$ and $\underset{\sim}{B}$ are really time dependent then $\overline{\underset{\sim}{E}}$ and $\overline{\underset{\sim}{B}}$ will not satisfy Maxwell's equations because $\nabla\times\overline{\underset{\sim}{E}} = \frac{1}{c}\frac{\partial\overline{\underset{\sim}{B}}}{\partial t} \neq -\frac{1}{c}\frac{\partial\overline{\underset{\sim}{B}}}{\partial t}$. However if we define instead $\hat{\underset{\sim}{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$ and $\hat{\underset{\sim}{B}}(t,\underset{\sim}{x}) = -\underset{\sim}{B}(-t,\underset{\sim}{x})$, then $\hat{\underset{\sim}{E}}$ and $\hat{\underset{\sim}{B}}$ satisfy Maxwell's equations with ρ and $\underset{\sim}{J}$ replaced by $\hat{\rho}(t,\underset{\sim}{x}) \equiv \rho(-t,\underset{\sim}{x})$ and $\hat{\underset{\sim}{J}}(t,\underset{\sim}{x}) \equiv -\underset{\sim}{J}(-t,\underset{\sim}{x})$ respectively.

In practice the new densities $\hat{\rho}$ and $\hat{\underset{\sim}{J}}$ could be obtained by reversing the motion of the charges acting as sources. It is physically clear that the fields resulting from this operation are $\hat{\underset{\sim}{E}}$ and $\hat{\underset{\sim}{B}}$. On the other hand, we cannot think of any physical operation resulting in the fields $\bar{\underset{\sim}{E}}, \bar{\underset{\sim}{B}}$ and this is consistent with Maxwell's equations which say that these fields do not exist.

One is thus led to define the operation of time reversal on an electromagnetic field as the replacements:

$$\underset{\sim}{E}(t,\underset{\sim}{x}) \rightarrow \hat{\underset{\sim}{E}}(t,\underset{\sim}{x}) \equiv \underset{\sim}{E}(-t,\underset{\sim}{x})$$

$$\underset{\sim}{B}(t,\underset{\sim}{x}) \rightarrow \hat{\underset{\sim}{B}}(t,\underset{\sim}{x}) \equiv -\underset{\sim}{B}(-t,\underset{\sim}{x})$$

As an operation on the vector potential this amounts to:

$$A_\mu(t,\underset{\sim}{x}) \rightarrow \hat{A}_\mu(t,\underset{\sim}{x}) \equiv (A_o(-t,\underset{\sim}{x}), -\underset{\sim}{A}(-t,\underset{\sim}{x}))$$

Let $\quad \psi'(t,\underset{\sim}{x}) \equiv \gamma^5\gamma^o\psi^c(-t,\underset{\sim}{x}) = \gamma^o BK\psi(-t,\underset{\sim}{x})$. By definition $\psi'(t,\underset{\sim}{x})$ is the spinor obtained from ψ by time reversal. It is easy to see that ψ' satisfies the equation

$$[\gamma^\mu(i\partial_\mu - e\hat{A}_\mu) - m]\psi' = 0$$

It thus describes the same particle as ψ but evolving in the electromagnetic field $\hat{A}_\mu$ obtained from A_μ by time reversal.

In the next section we come back to the group S^+ and take some time to obtain a global concrete picture of it.

(4) <u>A concrete picture of the group $S^{\dagger}$</u>

The first proposition describes the connected pieces of $S^{\dagger}$.

<u>Proposition</u> (5.4): The group $S_{+}^{\dagger}$ is connected. $S^{\dagger}$ has two components: $S_{+}^{\dagger}$ and $\gamma^{o}S_{+}^{\dagger} \equiv \{\gamma^{o}\Lambda;\ \Lambda\epsilon S_{+}^{\dagger}\}$.

Proof: From eq. (5.15) we know that there is a continuous curve in $S_{+}^{\dagger}$ joining I to -I. Now let $\Omega \in L_{+}^{\dagger}$. There are two solutions, Λ and $-\Lambda$, to $\Pi(\Lambda) = \Omega$. From eq. (5.22) we know that at least one of them, say Λ, can be reached by a continuous curve lying in $S_{+}^{\dagger}$ and starting at I. But since there is a curve connecting I to -I in $S_{+}^{\dagger}$, it is clear that there is a curve lying in $S_{+}^{\dagger}$ connecting Λ to $-\Lambda$ (one simply takes the one connecting I to -I and multiplies it by Λ). Whence any element in $S_{+}^{\dagger}$ can be connected to I by a continuous curve lying in $S_{+}^{\dagger}$ and $S_{+}^{\dagger}$ is connected. We noticed after theorem (5.2) that $S^{\dagger}$ is disconnected. We have $S^{\dagger} = (S_{+}^{\dagger}) \cup (\gamma^{o}S_{+}^{\dagger})$ and both $S_{+}^{\dagger}$ and $\gamma^{o}S_{+}^{\dagger}$ are connected. Since they don't intersect it is clear that they are the two components of $S^{\dagger}$.

<u>Proposition</u> (5.5): Any Λ in $S^{\dagger}$ has determinant 1.

Proof: Eq. (5.4b) shows that $\det\Lambda = \pm 1$. Since by proposition (5.4) $S_{+}^{\dagger}$ is connected and $\det I = 1$ it follows that any Λ in $S_{+}^{\dagger}$ has determinant 1. If Λ is in $S^{\dagger}$, and not in $S_{+}^{\dagger}$, then $\Lambda = \gamma^{o}\Lambda_{+}$ where $\Lambda_{+} \in S_{+}^{\dagger}$.

whence $\det\Lambda = \det\gamma^{0}\,\det\Lambda_{+} = 1\times 1 = 1$, by the corollary of theorem (2.9).

After proposition (5.3) we said that we would give a third formulation of eqs. (5.3a,b,c). It is contained in the following proposition.

<u>Proposition</u> (5.6): The equations (5.3a,b,c) defining $S^{\dagger}$ are equivalent to

$$\Omega^{\mu}{}_{\nu}\gamma^{\nu} = \Lambda^{-1}\gamma^{\mu}\Lambda \quad , \quad \Omega \in L^{\dagger}$$

$$\det\Lambda = 1$$

$$\mathrm{Tr}\Lambda^{*} = \mathrm{Tr}\Lambda$$

Proof: If Λ is a solution of (5.3a,b,c) it follows that it satisfies the first two conditions by proposition (5.5). The third condition that $\mathrm{Tr}\Lambda$ be real follows from (5.3c): $\mathrm{Tr}\Lambda^{*} = \mathrm{Tr}B^{+}\Lambda B = \mathrm{Tr}\Lambda BB^{+} = \mathrm{Tr}\Lambda$. Now let us see that the three conditions imply eqs. (5.3a, b,c). The first condition determines Λ up to an arbitrary complex multiplicative factor: if $\bar{\Lambda}$ is one of the two solutions to (5.3a,b,c) then $\Lambda = c\bar{\Lambda}$. Now the second condition above says that c is either ± 1 or $\pm i$, and the third condition eliminates the $\pm i$ possibility. Whence $\Lambda = \pm\bar{\Lambda}$. This completes the proof of the equivalence.

We now come to the main result of this section which provides an identification of $S_+^\uparrow$ with a surprisingly simple group.

<u>Proposition</u> (5.7): The group $S_+^\uparrow$ is isomorphic to the group SL(2,C) of complex 2×2 matrices with determinant 1.

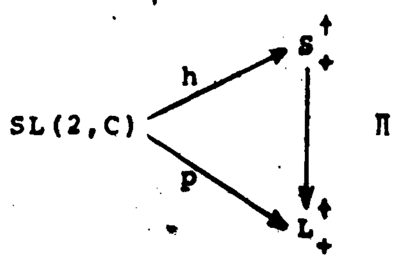
Before we prove this, it is necessary to make a few comments about SL(2,C) and its relation to $L_+^\uparrow$. SL(2,C) shares with $S_+^\uparrow$ the property that there exists a homomorphism P from it onto $L_+^\uparrow$ such that $P(M) = P(N)$ if and only if $M = \pm N$. This P is constructed as follows. One establishes a one to one linear correspondence between the 4-dimensional Minkowski space M and the real vector space H of 2×2 complex hermitian matrices by $x^\mu \to x^\mu \sigma_\mu$, where the σ_μ's are the Pauli matrices of eq. (2.8). Upon setting $\bar{x} \equiv x^\mu \sigma_\mu$, one easily checks that $\det\bar{x} = x^{o^2} - \underset{\sim}{x}\cdot\underset{\sim}{x}$. An element a of SL(2,C) induces a linear map $\tilde{a}: H \to H : h \to aha^+$. Now if $\tilde{a}(\bar{x}) = \bar{x}' = x'^\mu \sigma_\mu$ one has

$$(x'^o)^2 - \underset{\sim}{x}'\cdot\underset{\sim}{x}' = \det(\bar{x}') = \det(a\bar{x}a^+) = \det\bar{x} = x^{o^2} - \underset{\sim}{x}\cdot\underset{\sim}{x},$$

so that a determines a Lorentz transformation Ω_a. One then shows that $P: a \to \Omega_a$ is a homomorphism of SL(2,C) into $L_+^\uparrow$, that it is surjective and that $P(a) = P(b)$ if and only if $a = \pm b$. The details of this can be found in (Ruhl [1970]).

Suppose that we focus our attention on a small enough neighborhood V of I in $S_+^\uparrow$. To each A in V there will correspond a unique a_Λ in SL(2,C) close to the identity such that $\Pi(\Lambda) = P(a_\Lambda)$. If we take Λ' close to Λ there will be a unique $a_{\Lambda'}$ close to a_Λ with $\Pi(\Lambda') = P(a_{\Lambda'})$. By moving by little steps in $S_+^\dagger$ we can extend the map $\Lambda \rightarrow a_\Lambda$ to the whole of $S_+^\dagger$ and this map will be an isomorphism. All we have to do now is to make this intuitive argument rigorous. This is the only place in this work where we invoke mathematical notions which are not completely elementary. It is important to notice that since SL(2,C) is simply connected, the above discussion shows that it is also the universal covering group of $L_+^\dagger$. Our theorem (5.2) together with the now established connectivity of $S_+^\dagger$ (proposition (5.4)) shows that $S_+^\dagger$ is a covering group for $L_+^\dagger$. If we knew that $S_+^\dagger$ was simply connected we could conclude at once that it is isomorphic to SL(2,C) by invoking the uniqueness (up to isomorphism) of the universal covering group. But we don't want to show directly that $S_+^\dagger$ is simply connected. Instead we will use a result about topological groups which will enable us to prove the isomorphism without any calculation. By the same token we will have shown that $S_+^\dagger$ is simply connected.

The theorem we use says the following. Let G be a topological group with universal covering group $(\tilde{G},p)$ ($p: \tilde{G} \rightarrow G$ covering homomorphism). Suppose that [illegible] is another covering group for G. Then there exists a unique continuous homomorphism $h:\tilde{G}\to\tilde{G}'$ such that $p'\circ h = p$. Moreover $(\tilde{G},h)$ is a covering group for $\tilde{G}'$. This theorem can be found in (Pichon [1973]). In our case we take $G = L_+^{\uparrow}$, $(\tilde{G},p) = (SL(2,C),P)$ and $(\tilde{G}',p') = (S_+^{\uparrow},\Pi)$. The quoted theorem then says that there exists a unique continuous homomorphism $h:SL(2,C)\to S_+^{\uparrow}$ such that $\Pi\circ h = P$ as illustrated by the diagram:



Clearly, given any $a\in SL(2,C)$ h gives one of the two Λ's in $S_+^{\uparrow}$ such that $\Pi(\Lambda) = P(a)$, and this is accomplished in a continuous fashion. Our claim is that this is in fact an isomorphism. The theorem says that $(SL(2,C),h)$ is a covering group for $S_+^{\uparrow}$, whence h is surjective. So we only need to show that it is injective. The only two elements that h could map to $-I$ are $\pm I$. Since h is surjective and $h(I) = I$ we must have $h(-I) = -I$. But $-I$ is the only element other than I that h could have mapped to I. Whence we have that $h(a) = I$ implies $a = I$ and h is injective. Whence h is an isomorphism.

This section and the previous ones have provided a complete description of the groups of transformations $S_+^{\uparrow}$ and $S^{\uparrow}$. We now turn to the construction of tensors from spinors.

CHAPTER VI

TENSORS CONSTRUCTED FROM SPINORS

(1) Introduction

The aim of this chapter is to construct tensors from quadratic combinations of spinors and to study their relationships. These tensors are well known so we are not going to define anything new. However the section devoted to the study of their relationships might have some originality. In (Pauli [1936]) Pauli has shown how to derive some identities relating these tensors by using the Fierz identity. However the set of identities which he displayed is incomplete in the sense that it doesn't fully express the restrictions on the degrees of freedom in the tensor components. After having discussed the construction of the tensors we will provide a complete solution to the question of their algebraic dependence. In particular we shall give a set of almost independent covariant identities which tells exactly how the various tensors are related to each other. All the other identities are derivable from this particular set.

We must emphasize that our analysis is limited to the case where ψ is an ordinary spinor and not a field operator. We have not seriously investigated how much of the analysis carries through in this more general situation.

(2) Construction of the tensors

In order to obtain real quantities one has to take appropriate linear combinations of the products $\psi^{+}_{\mu}\psi_{\nu}$. The maximum number of linearly independent such combinations is clearly 16. It is in fact possible to construct 16 linearly independent quadratic forms which are all components of tensors or pseudotensors. We list them first and then proceed to show that they have the appropriate transformation laws.

$S \equiv \bar{\psi}\psi$: scalar

$P \equiv \bar{\psi}\gamma^{5}\psi$: pseudo-scalar

$V^{\mu} \equiv \bar{\psi}\gamma^{\mu}\psi$: future pointing timelike or null vector

$P^{\mu} \equiv i\bar{\psi}\gamma^{\mu}\gamma^{5}\psi$: spacelike or null pseudo-vector

$S^{\mu\nu} \equiv i\bar{\psi}\gamma^{[\mu\nu]}\psi$: antisymmetric tensor

One easily checks that all these quantities are real. Their linear independence follows from that of the γ^{A}'s. From the condition (5.3b), $\bar{\psi} \equiv \psi^{+}\gamma^{0}$ transforms under the action of $L^{\uparrow}$ according to $\bar{\psi}' = \bar{\psi}\Lambda^{-1}$. From the proof of proposition (5.1), the condition (5.3b) can only be satisfied when dealing with orthochronous Lorentz transformations. Accordingly our quantities behave as claimed only under this type of transformations.

It is trivial to check that S and V^{μ} are a scalar and a vector. Now let us see that $S^{\mu\nu}$ is an antisymmetric tensor. The antisymmetry follows at once from that of $\gamma^{[\mu\nu]}$. We have $S'^{\mu\nu} = i\bar{\psi}'\gamma^{[\mu\nu]}\psi' = i\bar{\psi}\Lambda^{-1}\gamma^{[\mu\nu]}\Lambda\psi$. Clearly if $\mu \neq \nu$ $S'^{\mu\nu} = 0$. Suppose now that $\mu \neq \nu$. Then:

$$
\begin{aligned}
S'^{\mu\nu} &= i\bar{\psi}\Lambda^{-1}\gamma^{\mu}\Lambda\Lambda^{-1}\gamma^{\nu}\Lambda\psi \\
&= i\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{\rho}\gamma^{\sigma}\psi \\
&= i\sum_{\rho\neq\sigma}\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{[\rho\sigma]}\psi + i\sum_{\rho=\sigma}\Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}\bar{\psi}\gamma^{\rho}\gamma^{\sigma}\psi \\
&= \Omega^{\mu}{}_{\rho}\Omega^{\nu}{}_{\sigma}T^{\rho\sigma} + i\bar{\psi}(-\Omega^{\mu}{}_{i}\Omega^{\nu}{}_{i} + \Omega^{\mu}{}_{o}\Omega^{\nu}{}_{o})\psi
\end{aligned}
$$

The second term vanishes because we assumed $\mu \neq \nu$ and Ω is a Lorentz transformation. To treat the other two cases we need to use permutations. If $\gamma'^{\mu} \equiv \Omega^{\mu}{}_{\nu}\gamma^{\nu}$, then

$$
\gamma'^{5} = \gamma'^{0}\gamma'^{1}\gamma'^{2}\gamma'^{3} = \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\gamma'^{\sigma_o}\gamma'^{\sigma_1}\gamma'^{\sigma_2}\gamma'^{\sigma_3}
$$

where S_4 is the set of permutations of (0123).

$$
= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\Omega^{\sigma_o}{}_{\rho_o}\Omega^{\sigma_1}{}_{\rho_1}\Omega^{\sigma_2}{}_{\rho_2}\Omega^{\sigma_3}{}_{\rho_3}\gamma^{\rho_o}\gamma^{\rho_1}\gamma^{\rho_2}\gamma^{\rho_3}
$$

Clearly the only terms contributing to this sum are those for which all the ρ_{ν}'s are different. So we may write

$$
\begin{aligned}
\gamma'^{5} &= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}\sum_{\tau\in S_4}(-1)^{\tau}\,\Omega^{\sigma_o}{}_{\tau_o}\ldots\Omega^{\sigma_3}{}_{\tau_3}\gamma^{5} \\
&= \frac{1}{4!}\sum_{\sigma\in S_4}(-1)^{\sigma}(-1)^{\sigma}\det(\Omega^{\mu}{}_{\nu})\gamma^{5} = \det(\Omega^{\mu}{}_{\nu})\gamma^{5}
\end{aligned}
$$

From this we obtain:

$$P'^{\mu} = \overline{\psi}\Lambda^{-1}\gamma^{\mu}\Lambda\Lambda^{-1}\gamma^{5}\Lambda\psi = \Omega^{\mu}{}_{\nu}\overline{\psi}\gamma^{\nu}\hat{\gamma}^{5}\psi = \det(\Omega)\Omega^{\mu}{}_{\nu}P^{\nu} \quad .$$

Similarly $P' = \det(\Omega)P$.

The physical interpretation of the vector V^{μ} is the probability current of the electron and it is denoted by j^{μ}. $j^{0} = \psi^{+}\psi$ is indeed positive and can thus represent a probability density. The conservation of total probability is guaranteed, under appropriate boundary conditions at infinity, because j^{μ} satisfies the continuity equation. Indeed upon multiplying $(i\gamma^{\mu}\partial_{\mu} - m)\psi = 0$ on the left by $\overline{\psi}$ and the adjoint equation $\overline{\psi}(i\gamma^{\mu}\overleftarrow{\partial}_{\mu} + m) = 0$ on the right by ψ and adding one obtains $\partial_{\mu}j^{\mu} = 0$. This also works when an external electromagnetic field is present.

The tensor $-2S^{\mu\nu}$ is interpreted as a spin density (Messiah [1964]). The other tensors are used to couple the Dirac field with other fields.

(3) <u>Covariant identities</u>

Let us now come to the identities connecting the tensors. We <u>fix</u> the space-time point x and look at how the components of the tensors are related to each other at that point. They are 16 real-valued functions of $\psi(x)$. (We emphasize again that in this study ψ is an ordinary spinor and not a field operator.) Since $\psi(x)$ contains 8 independent real parameters, the components of the tensors can be considered as 16 real functions of 8 real variables. It is therefore clear that at most 8 of them can be independent. It turns out that only seven are independent and the nine others are determined by the first seven. The basic tool for studying these functions in a way which is independent of a particular choice of unitary γ^{μ}'s is the Fierz identity. It is a very remarkable identity satisfied by any set of 4×4 γ matrices. Let $\{\gamma^{\mu}\}$ be such a set. By proposition (2.5) we know that the 16 γ^{A}'s span the entire vector space of complex 4×4 matrices. If M is an arbitrary 4×4 matrix there exist coefficients α_A such that $M = \sum_A \alpha_A \gamma^A$. Multiplying this by $\gamma_B \equiv (\gamma^B)^{-1}$ and taking the trace yields $\alpha_B = \frac{1}{4} T_r(M\gamma_B)$. Whence we have:

$$M_{\alpha\beta} = \frac{1}{4} \sum_A \gamma^A{}_{\alpha\beta} (\gamma_A)_{\nu\mu} M_{\mu\nu}$$

But, since M is arbitrary, this implies

$$\frac{1}{4} \sum_A \gamma^A{}_{\alpha\beta} (\gamma_A)_{\nu\mu} = \delta_{\alpha\mu}\delta_{\beta\nu} \quad (6.1)$$

This is the Fierz identity. It is in fact a tensor product identity and is seen most clearly when written as such. The tensor product of two linear operators M and N, $M\otimes N$, is defined by $M\otimes N(u\otimes v) = Mu\otimes Nv$ for arbitrary vectors u,v. Given $M\otimes N$ we define $(M\otimes N)^{\pi}$ by $(M\otimes N)^{\pi}(u\otimes v) = Mv\otimes Nu$, so that

$(M \bullet N)^{\pi} = (M \bullet N)(I \bullet I)^{\pi}$. It is then a trivial matter to check that the Fierz identity may be rewritten as:

$$\frac{1}{4} \sum_A \gamma^A \bullet \gamma_A = (I \bullet I)^{\pi} \tag{6.2}$$

The advantage of this notation is that many equations become clearer because of the elimination of the indices. There is another algebraic tool which we shall need. If $\{\gamma^{\mu}\}$ is a system of γ matrices, so is $\{\gamma^{\mu^T}\}$, the set of transposes. By the fundamental theorem, there exists an invertible matrix T such that $\gamma^{\mu^T} = T\gamma^{\mu}T^{-1}$. This matrix T is the other tool which we will use. It is anti-symmetric. Indeed by taking the transpose of the equation defining T one obtains:

$$\gamma^{\mu} = T^{-1^T}\gamma^{\mu^T}T^T = T^{-1}\gamma^{\mu^T}T$$

from which $[\gamma^{\mu^T}, T^T T^{-1}] = 0$. By theorem (2.8) it follows that $T^T T^{-1} = cI$ or $T^T = cT$. From this we obtain $T = cT^T = c^2 T$, whence $c = \pm 1$. Thus T is either symmetric or antisymmetric. Suppose that it was symmetric. Then one can easily check that the ten matrices $T\gamma^{[\mu\nu]}$ and $T\gamma^{[\lambda\mu\nu]}$ would be antisymmetric. But, since there are at most six linearly independent 4×4 antisymmetric matrices, this is a contradiction; therefore, T is antisymmetric. This elegant argument is taken from (Pauli [1936]).

We are now prepared to study the identities satisfied by the tensors. The first step is to write the Fierz identity in five different ways:

$$4(I\otimes I)^{\pi} = I\otimes I+\gamma^{\mu}\otimes\gamma_{\mu}-\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}+\gamma^{\mu}\gamma^{5}\otimes\gamma_{\mu}\gamma^{5}+\gamma^{5}\otimes\gamma_{5} \qquad (6.3)$$

$$4(\gamma^{5}\otimes\gamma_{5})^{\pi} = I\otimes I-\gamma^{\mu}\otimes\gamma_{\mu}-\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma_{[\mu\nu]}-\gamma^{\mu}\gamma^{5}\otimes\gamma_{\mu}\gamma^{5}+\gamma^{5}\otimes\gamma_{5} \qquad (6.4)$$

$$4(\gamma^{5}\otimes I)^{\pi} = \gamma^{5}\otimes I+I\otimes\gamma^{5}+\gamma^{\mu}\otimes\gamma_{\mu}\gamma^{5}-\gamma_{\mu}\gamma^{5}\otimes\gamma^{\mu}+\frac{1}{2}\tilde{\gamma}^{[\mu\nu]}\otimes\gamma_{[\mu\nu]} \qquad (6.5)$$

$$4(I\otimes\gamma^{5})^{\pi} = \gamma^{5}\otimes I+I\otimes\gamma^{5}-\gamma^{\mu}\otimes\gamma_{\mu}\gamma^{5}+\gamma_{\mu}\gamma^{5}\otimes\gamma^{\mu}+\frac{1}{2}\tilde{\gamma}^{[\mu\nu]}\otimes\gamma_{[\mu\nu]} \qquad (6.6)$$

$$4(T^{-1}\otimes T)^{\pi} = I\otimes I+\gamma^{\mu}\otimes\gamma_{\mu}^{T}+\frac{1}{2}\gamma^{[\mu\nu]}\otimes\gamma^{T}_{[\mu\nu]}-\gamma^{\mu}\gamma^{5}\otimes(\gamma_{\mu}\gamma^{5})^{T}+\gamma^{5}\otimes\gamma_{5}^{T} \qquad (6.7)$$

The matrix $\tilde{\gamma}_{[\mu\nu]}$ is defined by $\tilde{\gamma}_{[\mu\nu]} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}\gamma^{[\alpha\beta]}$. This corresponds to the definition of the dual of an antisymmetric tensor $T_{\mu\nu}$ as $\tilde{T}_{\mu\nu} = \frac{1}{2}\epsilon_{\mu\nu\alpha\beta}T^{\alpha\beta}$. From eq. (2.5) we have: $\tilde{\gamma}_{[\mu\nu]} = \gamma_{5}\gamma_{[\mu\nu]}$, where we recall that $\gamma_{5} \equiv (\gamma^{5})^{-1} = -\gamma^{5}$. Eq. (6.3) is just (6.2) written out explicitely. Eq. (6.4) is obtained by multiplying (6.3) on the left by $\gamma^{5}\otimes\gamma_{5}$. Eq. (6.5) is obtained by multiplying (6.3) on the left by $\gamma^{5}\otimes I$. Eq. (6.6) is obtained by reversing all tensor products in (6.5); this can be achieved by multiplying on right and left by $(I\otimes I)^{\pi}$. Finally eq. (6.7) is obtained by multiplying (6.3) on the left by $I\otimes T$ and on the right by $I\otimes T^{-1}$.

For a while it will be convenient to replace P^{μ} and $S^{\mu\nu}$ by the pure imaginary quantities $\tilde{V}^{\mu} \equiv -iP^{\mu} = \overline{\psi}\gamma^{\mu}\gamma^{5}\psi$, and $T^{\mu\nu} \equiv -iS^{\mu\nu} = \overline{\psi}\gamma^{[\mu\nu]}\psi$.

By multiplying (6.3) on the left by $\overline{\psi}\bullet\overline{\psi}$ and on the right by $\psi\bullet\psi$ we obtain:

$$4S^2 = S^2 + V^{\mu}V_{\mu} - \frac{1}{2}T^{\mu\nu}T_{\mu\nu} + \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

By doing the same thing with (6.4) we get:

$$-4P^2 = S^2 - V^{\mu}V_{\mu} - \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

By adding and substracting these two equations we derive the equivalent system:

$$T^{\mu\nu}T_{\mu\nu} = 2(P^2-S^2) \qquad (6.8)$$

$$V^{\mu}V_{\mu} + \tilde{V}^{\mu}\tilde{V}_{\mu} = 2(P^2+S^2) \qquad (6.9)$$

Performing the same operations on (6.5) yields:

$$\tilde{T}^{\mu\nu}T_{\mu\nu} = 4PS \qquad (6.10)$$

Next if we multiply (6.7) on the left by $\overline{\psi}\bullet\psi^{T}$ and on the right by $\psi\bullet\overline{\psi}^{T}$ we obtain:

$$4(\overline{\psi}T^{-1}\overline{\psi}^{T})(\psi^{T}T\psi) = S^2 + V^{\mu}V_{\mu} + \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^{\mu}\tilde{V}_{\mu} - P^2$$

where we have used the fact that an expression like $\psi^{T}\gamma_{\mu}\overline{\psi}^{T}$, being a number, is equal to its transpose $\overline{\psi}\gamma_{\mu}\psi$. Now since

T is antisymmetric we have $\psi^T T\psi = (\psi^T T\psi)^T = -\psi^T T\psi = 0$. Whence our last equation reduces to

$$S^2 + V^\mu V_\mu + \frac{1}{2}T^{\mu\nu}T_{\mu\nu} - \tilde{V}^\mu\tilde{V}_\mu - P^2 = 0$$

When combined with (6.8) and (6.9), this gives

$$V^\mu V_\mu = \tilde{V}^\mu\tilde{V}_\mu = P^2+S^2 \qquad (6.11)$$

Upon substracting (6.5) from (6.6) and then multiplying on the left by $I\otimes T$ and on the right by $I\otimes T^{-1}$ we get

$$2[T^{-1}\otimes T\gamma^5 - \gamma^5 T^{-1}\otimes T]^\pi = \gamma_\mu\otimes(\gamma_\mu\gamma^5)^T + \gamma_\mu\gamma^5\otimes\gamma^{\mu^T}$$

Then by treating this as we treated eq. (6.7) we obtain:

$$V^\mu\tilde{V}_\mu = 0 \qquad (6.12)$$

Eqs. (6.8), (6.10), (6.11) and (6.12) are the identities which Pauli displayed. We now proceed to derive others.

By taking the difference between (6.3) and (6.4) we obtain

$$2[I\otimes I - \gamma^5\otimes\gamma_5]^\pi = \gamma^\rho\otimes\gamma_\rho + \gamma^\rho\gamma^5\otimes\gamma_\rho\gamma^5 \qquad (6.13)$$

Multiplying this on the left by $I\otimes\gamma_\mu$ yields

$$2[I\otimes\gamma_\mu + \gamma^5\otimes\gamma_\mu\gamma^5]^\pi = \gamma^\rho\otimes\gamma_\mu\gamma_\rho + \gamma^\rho\gamma^5\otimes\gamma_\mu\gamma_\rho\gamma^5$$

$$= \gamma_\mu\otimes I + \gamma^\rho\otimes\gamma_{[\mu\rho]} + \gamma_\mu\gamma^5\otimes\gamma^5 + \gamma^\rho\gamma^5 \ldots$$

By applying $\bar{\psi}\bullet\bar{\psi}$ to the left and $\psi\bullet\psi$ to the right we get:

$$2[SV_\mu + P\tilde{V}_\mu] = SV_\mu + P\tilde{V}_\mu + T_{\mu\rho}V^\rho - \tilde{T}_{\mu\rho}\tilde{V}^\rho \qquad (6.14)$$

Next by taking the sum of (6.3) and (6.4) we get:

$$2[I\bullet I - \gamma^5\bullet\gamma^5]^\pi = I\bullet I - \gamma^5\bullet\gamma^5 - \frac{1}{2}\gamma^{[\mu\nu]}\bullet\gamma_{[\mu\nu]}$$

Multiplying this on the left by $I\bullet\gamma_\mu$ yields:

$$2[I\bullet\gamma_\mu - \gamma^5\bullet\gamma_\mu\gamma^5]^\pi = I\bullet\gamma_\mu - \gamma^5\bullet\gamma_\mu\gamma^5 - \frac{1}{2}\gamma^{[\alpha\beta]}\bullet\gamma_\mu\gamma_{[\alpha\beta]}$$

Now, by using the product rule (2.2), this may be rewritten as

$$2[I\bullet\gamma_\mu - \gamma^5\bullet\gamma_\mu\gamma^5]^\pi = I\bullet\gamma_\mu - \gamma^5\bullet\gamma_\mu\gamma^5 - \tilde{\gamma}_{[\mu\nu]}\bullet\gamma^\nu\gamma^5 - g_{\mu\alpha}\gamma^{[\alpha\beta]}\bullet\gamma_\beta$$

which, when taken between $\bar{\psi}\bullet\bar{\psi}$ and $\psi\bullet\psi$, gives

$$2[SV_\mu - P\tilde{V}_\mu] = SV_\mu - P\tilde{V}_\mu - \tilde{T}_{\mu\nu}\tilde{V}^\nu - T_{\mu\beta}V^\beta$$

If we combine this with eq. (6.14) we obtain the equivalent system:

$$T_{\mu\rho}V^\rho = P\tilde{V}_\mu \qquad (6.15)$$

$$\tilde{T}_{\mu\rho}\tilde{V}^\rho = -SV_\mu \qquad (6.16)$$

We consider again eq. (6.13) and multiply it on the right by $\gamma^\mu\bullet\gamma_\rho$; this yields:

$$2[\gamma_\rho \otimes \gamma^\mu - \gamma_\rho\gamma^5 \otimes \gamma^\mu\gamma_5]^\pi = \gamma^\nu\gamma^\mu \otimes \gamma_\nu\gamma_\rho + \gamma^\nu\gamma^\mu\gamma^5 \otimes \gamma_\nu\gamma_\rho\gamma^5$$

$$= \delta^\mu{}_\rho[I\otimes I + \gamma^5 \otimes \gamma^5] + \gamma^{[\nu\mu]} \otimes \gamma_{[\nu\rho]} + \gamma^{[\nu\mu]}\gamma^5 \otimes \gamma_{[\nu\rho]}\gamma^5$$

$$+ I \otimes \gamma^{[\mu}\gamma_{\rho]} + \gamma_{[\rho}\gamma^{\mu]} \otimes I + \gamma^5 \otimes \gamma^{[\mu}\gamma_{\rho]}\gamma^5 + \gamma_{[\rho}\gamma^{\mu]}\gamma^5 \otimes \gamma^5$$

Operating on this with $\bar{\psi}\otimes\bar{\psi}$ and $\psi\otimes\psi$, in the now familiar way, we obtain:

$$T^{\mu\nu}T_\nu{}^\rho + \tilde{T}^{\mu\nu}\tilde{T}_\nu{}^\rho = g^{\mu\rho}(S^2+P^2) - 2(V^\mu V^\rho + \tilde{V}^\mu\tilde{V}^\rho) \tag{6.17}$$

As long as $T_{\mu\nu}$ is antisymmetric, the following identities hold true:

$$T^{\mu\nu}T_{\nu\rho} - \tilde{T}^{\mu\nu}\tilde{T}_{\nu\rho} = -\frac{1}{2}(T^{\alpha\beta}T_{\alpha\beta})\delta^\mu{}_\rho \tag{6.18}$$

$$T^{\mu\nu}\tilde{T}_{\nu\rho} = -\frac{1}{4}(T^{\alpha\beta}\tilde{T}_{\alpha\beta})\delta^\mu{}_\rho \tag{6.19}$$

Combining (6.18) with (6.8) and (6.17) yields

$$T^{\mu\nu}T_\nu{}^\rho = -(V^\mu V^\rho + \tilde{V}^\mu\tilde{V}^\rho) + g^{\mu\rho}S^2 \tag{6.20}$$

$$\tilde{T}^{\mu\nu}\tilde{T}_\nu{}^\rho = -(V^\mu V^\rho + \tilde{V}^\mu\tilde{V}^\rho) + g^{\mu\rho}P^2 \tag{6.21}$$

Fortunately we have now nearly exhausted the set of all possible invariant quadratic identities! There remain only two. Multiplying (6.20) by V^ρ and using (6.11), (6.12) and (6.15) we obtain:

$$PT^{\mu\nu}\tilde{V}_\nu = -(S^2+P^2)V^\mu + S^2V^\mu = -P^2V^\mu$$

Since this is true irrespective of ψ we may divide by P to get

$$T^{\mu\nu}\tilde{V}_\nu = -PV^\mu \qquad (6.22)$$

Next we use this together with (6.19) and (6.10) to get:

$$\tilde{T}^{\mu\nu}V_\nu = -\frac{1}{P}\tilde{T}^{\mu\nu}T_{\nu\rho}\tilde{V}^\rho = -\frac{1}{P}\left(-\frac{1}{4}\tilde{T}^{\alpha\beta}T_{\alpha\beta}\right)\tilde{V}^\mu = S\tilde{V}^\mu,$$

that is:

$$\tilde{T}^{\mu\nu}V_\nu = S\tilde{V}^\mu \qquad (6.23)$$

We now reexpress all the identities obtained in terms of our original real-valued functions P^μ and $S^{\mu\nu}$.

$$V^\mu P_\mu = 0 \qquad (6.24)$$

$$V^\mu V_\mu = P^2+S^2 \qquad (6.25)$$

$$P^\mu P_\mu = -(P^2+S^2) \qquad (6.26)$$

$$S^{\mu\nu}S_{\mu\nu} = 2(S^2-P^2) \qquad (6.27)$$

$$\tilde{S}^{\mu\nu}S_{\mu\nu} = -4PS \qquad (6.28)$$

$$S^{\mu\nu}S_{\nu\rho} = (V^\mu V^\rho - P^\mu P^\rho) - g^{\mu\rho}S^2 \qquad (6.29)$$

$$\tilde{S}^{\mu\nu}\tilde{S}_{\nu\rho} = (V^\mu V^\rho - P^\mu P^\rho) - g^{\mu\rho}P^2 \qquad (6.30)$$

$$\tilde{S}^{\mu\rho}S_{\rho\sigma} = PS\,\delta^\mu_\sigma \qquad (6.31)$$

$$S_{\mu\rho}V^\rho = PP_\mu \qquad (6.32)$$

$$\tilde{S}_{\mu\rho}V^\rho = SP_\mu \qquad (6.33)$$

$$S_{\mu\rho}P^\rho = PV_\mu \qquad (6.34)$$

$$\tilde{S}_{\mu\rho}P^\rho = SV_\mu \qquad (6.35)$$

(4) The information contained in the identities

Clearly the left hand sides of the last set of equations comprise all possible quadratic invariant combinations. But these equations are so numerous that their content is far from clear. They form a highly redundant system: for instance, (6.24) is an obvious consequence of any one of the last four equations. In fact only nine equations are independent. This leaves seven independent functions: one possible choice is P, V^k, P^k. We shall now show how one can pick a system of almost independent identities implying all the others and then proceed to demonstrate that P, V^k, P^k are independent. First we fix the notation. We define:

$$\underset{\sim}{V} = \begin{bmatrix} V^1 \\ V^2 \\ V^3 \end{bmatrix} \qquad \underset{\sim}{P} = \begin{bmatrix} P^1 \\ P^2 \\ P^3 \end{bmatrix} \tag{6.36}$$

$$\underset{\sim}{S} = \begin{bmatrix} S_{23} \\ S_{31} \\ S_{12} \end{bmatrix} \qquad \underset{\sim}{K} = \begin{bmatrix} S_{10} \\ S_{20} \\ S_{30} \end{bmatrix} \tag{6.37}$$

so that

$$(S_{\mu\nu}) = \begin{bmatrix} 0 & -\underset{\sim}{K}^T \\ \underset{\sim}{K} & \omega(\underset{\sim}{S}) \end{bmatrix} \qquad (\tilde{S}_{\mu\nu}) = \begin{bmatrix} 0 & \underset{\sim}{S}^T \\ -\underset{\sim}{S} & \omega(\underset{\sim}{K}) \end{bmatrix} \tag{6.38}$$

where

$$\omega(\underset{\sim}{U}) \equiv \begin{bmatrix} 0 & U^3 & -U^2 \\ -U^3 & 0 & U^1 \\ U^2 & -U^1 & 0 \end{bmatrix}$$

$\mathbf{S}$ is twice the spin density vector. One easily establishes that

$$\mathbf{K}^2 - \mathbf{S}^2 = -\frac{1}{2} S^{\alpha\beta} S_{\alpha\beta} \qquad (6.39)$$

$$\mathbf{K} \cdot \mathbf{S} = \frac{1}{4} S^{\alpha\beta} \tilde{S}_{\alpha\beta} \qquad (6.40)$$

These are the analogues of the familiar invariants of the electromagnetic field.

Our claim is that the following set of identities is complete:

$$S_{\mu\rho} V^{\rho} = P P_{\mu} \qquad (6.41)$$

$$\tilde{S}_{\mu\rho} V^{\rho} = S P_{\mu} \qquad (6.42)$$

$$V^{\mu} V_{\mu} = S^2 + P^2 = -P^{\mu} P_{\mu} \qquad (6.43)$$

First (6.41) clearly implies (6.24). Now in our new notation, eqs. (6.41) and (6.42) read

$$P \begin{bmatrix} P^0 \\ -\mathbf{P} \end{bmatrix} = \begin{bmatrix} -\mathbf{K}\cdot\mathbf{V} \\ V^0 \mathbf{K} + \mathbf{V}\times\mathbf{S} \end{bmatrix} \qquad S \begin{bmatrix} P^0 \\ -\mathbf{P} \end{bmatrix} = \begin{bmatrix} \mathbf{S}\cdot\mathbf{V} \\ -V^0 \mathbf{S} + \mathbf{V}\times\mathbf{K} \end{bmatrix} \qquad (6.44)$$

The second equation gives

$$\mathbf{S} = \frac{1}{V^0} (\mathbf{V}\times\mathbf{K} + S\mathbf{P})$$

from which
$$\mathbf{V}\times\mathbf{S} = \frac{1}{V^0} \left((\mathbf{V}\cdot\mathbf{K})\,\mathbf{V} - \mathbf{V}^2 \mathbf{K} + S\,\mathbf{V}\times\mathbf{P} \right)$$

Inserting this into the first part of (6.44) we obtain

$$V^{o^2}\underset{\sim}{K} - PP^o\underset{\sim}{V} - \underset{\sim}{V}^2\underset{\sim}{K} + S\underset{\sim}{V}\times\underset{\sim}{P} = - V^o P\underset{\sim}{P}$$

that is: $$\underset{\sim}{K} = \frac{1}{V^\mu V_\mu}\,[P(P^o\underset{\sim}{V} - V^o\underset{\sim}{P}) - S\underset{\sim}{V}\times\underset{\sim}{P}]$$

From this we get:

$$\underset{\sim}{V}\times\underset{\sim}{K} = \frac{1}{V^\mu V_\mu}\,[-PV^o\underset{\sim}{V}\times\underset{\sim}{P} - S(\underset{\sim}{V}\cdot\underset{\sim}{P})\underset{\sim}{V} + S\underset{\sim}{V}^2\underset{\sim}{P}]$$

Inserting this into the second part of (6.44) yields

$$\underset{\sim}{S} = \frac{1}{V^\mu V_\mu}\,[-P\underset{\sim}{V}\times\underset{\sim}{P} - \frac{S}{V^o}(\underset{\sim}{V}\cdot\underset{\sim}{P})\underset{\sim}{V} + \frac{S}{V^o}\,\underset{\sim}{V}^2\underset{\sim}{P} + \frac{V^\mu V_\mu}{V^o}\,S\underset{\sim}{P}]$$

Equation (6.24), which may be written

$$V^oP^o = \underset{\sim}{V}\cdot\underset{\sim}{P} \quad , \qquad (6.45)$$

is a consequence of (6.41). This can be used to cancel the V^o in the denominator of the second term on the right hand side of the next to last equation. Thus we obtain the following expressions for $\underset{\sim}{K}$ and $\underset{\sim}{S}$:

$$\underset{\sim}{K} = \frac{1}{V^\mu V_\mu}\,[P(P^o\underset{\sim}{V} - V^o\underset{\sim}{P}) - S\underset{\sim}{V}\times\underset{\sim}{P}] \qquad (6.46)$$

$$\underset{\sim}{S} = \frac{1}{V^\mu V_\mu}\,[S(V^o\underset{\sim}{P} - P^o\underset{\sim}{V}) - P\underset{\sim}{V}\times\underset{\sim}{P}] \qquad (6.47)$$

We haven't yet used the $SP^o = \underset{\sim}{S}\cdot\underset{\sim}{V}$ part of eq. (6.44). In fact this is now a consequence of (6.45) and (6.47). This is the extent to which the identities (6.41,42,43) are not completely independent.

Next we obtain

$$SP_{\mu}P^{\mu} = \tilde{S}_{\mu\rho}V^{\rho}P^{\mu} \qquad \text{from (6.42)}$$

$$= \frac{1}{P}\,\tilde{S}_{\mu\rho}S^{\mu\lambda}V_{\lambda}V^{\rho} \qquad \text{from (6.41)}$$

$$= \frac{1}{4P}\,(\tilde{S}_{\alpha\beta}S^{\alpha\beta})V_{\rho}V^{\rho} \qquad \text{from (6.19)}$$

$$= -\frac{1}{4P}\,(\tilde{S}_{\alpha\beta}S^{\alpha\beta})P_{\mu}P^{\mu} \qquad \text{from (6.43)}$$

Whence $\tilde{S}_{\alpha\beta}S^{\alpha\beta} = -4PS$; this is (6.28).

The explicit calculations will not be given but one can deduce, using only (6.45,46,47), the relation

$$S^{\mu\nu}S_{\nu\rho} \overset{o}{=} \frac{S^2+P^2}{V^{\mu}V_{\mu}}\,(V^{\mu}V_{\rho} - P^{\mu}P_{\rho}) - \delta^{\mu}{}_{\rho}S^2 \quad .$$

When coupled with (6.43) this equation gives (6.29). It is also seen to imply (6.27) by letting $\rho=\mu$ and using $P^{\mu}P_{\mu} = -V^{\mu}V_{\mu}$. Combining (6.29), (6.27) and the identity (6.18), which is valid for arbitrary antisymmetric tensors, one is led to (6.30). Equation (6.31) follows from the general identity (6.19) and equation (6.28). Finally equations (6.34) and (6.35) are a consequence of (6.41), (6.42) together with the general identity (6.19) and eq. (6.28).

We have shown that the equations (6.41,42,43) imply all the others in (6.24) to (6.35).

Now suppose that $\underset{\sim}{V}$, $\underset{\sim}{P}$ and P are given. The equations $V^\mu V_\mu = -P^\mu P_\mu$ and $V^\mu P_\mu = 0$ read:

$$V^{o\,2} + P^{o\,2} = \underset{\sim}{V}^2 + \underset{\sim}{P}^2 \quad .$$

$$V^o P^o = \underset{\sim}{P} \cdot \underset{\sim}{V}$$

By themselves these equations are sufficient to determine V^o and P^o up to sign. But we know a priori that $V^o = \psi^+\psi \geq 0$. If V^o is 0 then $\psi=0$ and everything vanishes. If $V^o \neq 0$ then $V^o > 0$ and the two above equations take care of the signs as well as of the magnitudes of V^o and P^o. Now that V^μ is determined, the equation $V^\mu V_\mu = S^2+P^2$ gives S up to sign, P being supposedly given. Then, unless both S and P are zero, equations (6.46) and (6.47) determine $S_{\mu\nu}$ uniquely. If both S and P vanish, V^μ and P^μ are two orthogonal null* vectors; this implies that they are linearly dependent. In this case the system becomes degenerate since it fails to determine $S_{\mu\nu}$ uniquely. Indeed one may check by looking at (6.44) that it only demands that $\underset{\sim}{K}$ and $\underset{\sim}{S}$ be two vectors of equal norm orthogonal to each other and orthogonal to $\underset{\sim}{V}$. Whence there is one degree of freedom left: their position in the plane orthogonal to $\underset{\sim}{V}$. Due to the fact that V^μ and P^μ are linearly dependent, one can easily check that the whole set of equations (6.24) to (6.35) says nothing else about $S_{\mu\nu}$ than what is already implied by (6.46) and (6.47).

* A 4-vector A^μ is said to be null if $A^\mu A_\mu = 0$.

Thus we have shown that, except in the degenerate case where V^{μ} is a null vector, the system (6.41,42,43) determines the other functions once $\underset{\sim}{V}$, $\underset{\sim}{P}$ and P are given. Moreover this system is complete in the sense that it implies all the other equations in (6.24) to (6.35). To show that the system is really algebraically complete we have to demonstrate that there are no non trivial relations between $\underset{\sim}{V}$, $\underset{\sim}{P}$ and P, that is these are independent functions. One way to do this would be to use a particular set of γ matrices, for instance that of eq. (2.7), to write down explicitely these functions. But we have found an alternative method which is less cumbersome and doesn't rely on a special choice of the system $\{\gamma^{\mu}\}$. It amounts to showing that their differentials can be made linearly independent with an appropriate choice of ψ.

The seven functions P, $\underset{\sim}{V}$, $\underset{\sim}{P}$ are of the form $\psi \to \bar{\psi}M^{A}\psi$ where the M^{A}'s are γ^{5}, γ^{k} and $i\gamma^{k}\gamma^{5}$ respectively. Let us denote them by f^{A}. To emphasize their dependence upon eight real variables we write ψ as $\Phi+i\phi$ where both Φ and ϕ are real. Then we have:

$$f^{A}(\Phi,\phi) = (\bar{\Phi}-i\bar{\phi})M^{A}(\Phi+i\phi)$$

$$= \bar{\Phi}M^{A}\Phi + \bar{\phi}M^{A}\phi + i(\bar{\Phi}M^{A}\phi - \bar{\phi}M^{A}\Phi).$$

Now the differential of f^{A} at the point (Φ,ϕ) is defined by

$$f^A(\Phi+h,\phi+k) - f^A(\Phi,\phi) = df^A_{(\Phi,\phi)}(h,k) + \theta(h,k),$$

df^A being linear and $\theta(h,k)$ satisfying $\lim\limits_{(h,k)\to 0} \frac{\theta(h,k)}{\|(h,k)\|} = 0.$

One can easily check that.

$$df^A_{(\Phi,\phi)}(h,k) = (\bar{\Phi}-i\bar{\phi})M^A h+(\bar{\phi}+i\bar{\Phi})M^A k+\bar{h}M^A(\Phi+i\phi)+\bar{k}M^A(\phi-i\Phi)$$

Suppose that these differentials were linearly dependent. That means that there would be real numbers λ_A, not all vanishing, such that

$$\lambda_A df^A_{(\Phi,\phi)}(h,k) = 0 \quad \text{independently of h and k.}$$

This would imply the two equations

$$(\bar{\Phi}-i\bar{\phi})Mh + \bar{h}M(\Phi+i\phi) = 0 \quad \text{for all h} \tag{6.48}$$

$$(\bar{\phi}+i\bar{\Phi})Mk + \bar{k}M(\phi-i\Phi) = 0 \quad \text{for all k} \tag{6.49}$$

where we have set $\lambda_A M^A \equiv M$. Upon replacing the second term in (6.48) by its transpose, to which it is equal, we obtain

$$(\bar{\Phi}-i\bar{\phi})Mh + (\Phi^T+i\phi^T)M^T\gamma^{o\,T} h = 0$$

This being supposedly true for arbitrary h we deduce

$$(\bar{\Phi}-i\bar{\phi})M + (\Phi^T+i\phi^T)M^T\gamma^{o\,T} = 0 \tag{6.50}$$

Doing the same thing with (6.49) and multiplying the resulting equation by -i yields:

$$(\bar{\Phi}-i\bar{\phi})M - (\Phi^T+i\phi^T)M^T\gamma^{o^T} = 0 \qquad (6.51)$$

Substracting (6.51) from (6.50) gives $(\Phi^T+i\phi^T)M^T\gamma^{o^T} = 0$; this is equivalent to $\gamma^o M\psi = 0$ or $M\psi = 0$. Whence the problem is reduced to finding a ψ such that the seven vectors $M^A\psi$ are linearly independent over the reals. Suppose that we have a relation

$$[\lambda_o\gamma^5 + \lambda_k\gamma^k + i\mu_k\gamma^k\gamma^5]\psi = 0 \qquad (6.52)$$

where all λ's and μ's are real. Let ψ be an eigenvector of γ^o: $\gamma^o\psi = \rho\psi$. ρ is not 0 since γ^o is invertible. We also know that such an eigenvector exists since γ^o can be diagonalized.

Multiplying (6.52) by γ^o and then dividing by ρ we obtain:

$$[-\lambda_o\gamma^5 - \lambda_k\gamma^k + i\mu_k\gamma^k\gamma^5]\psi = 0$$

Together with (6.52) this implies:

$$\mu_k\gamma^k\psi = 0$$

$$[\lambda_o\gamma^5 + \lambda_k\gamma^k]\psi = 0$$

Multiplying the last equation by $\psi^+\gamma^5$ we get:

$$-\lambda_o\psi^+\psi + \lambda_k\psi^+\gamma^5\gamma^k\psi \ldots$$

[illegible] first term [illegible] real whereas [illegible] be purely imaginary. [illegible] and we are left with

[illegible]

[illegible]

[illegible] the first term is real while the second is a pure imaginary. Therefore [illegible] and we are left with

[illegible]

[illegible] same with [illegible] we obtain [illegible] = [illegible] and it then follows that [illegible] = [illegible] also. The same argument could clearly be applied to the equation [illegible] = 0. Whence the M^A's are linearly independent over the reals if we choose [illegible] to be an [illegible] of [illegible]. This shows that the functions F, V^k, P^k are independent in the neighborhood of this point [illegible]. Whence it is impossible to express some of them in terms of the others.

Let us summarize what we did. We have first obtained all the possible invariant quadratic identities among the tensors defined at the beginning. Next we have shown that all the information contained in this set of identities is already present in the following subset:

The way to [illegible] the algebraic content of this system of identities is as follows: once the values of the [illegible] independent functions I, V^k, P^k are given, the [illegible] allow one to compute the values of the nine [illegible] functions S, V^o, P^o, $C_{\mu\nu}$ ([illegible]). Though the system contains ten equations, one of them, the time component of the second part in ([illegible].44), can be deduced from the others. There appropriately remain nine independent equations to reduce to seven the number of independent functions in the total of sixteen.

Apart from those which were derived in (Pauli [illegible]), the identities ([illegible].24) to ([illegible].35) don't seem to be very well-known. While trying to see if they were new or not, the only place where we could trace them was in a short paper (Penney [19[illegible]4]). They are stated there without proof and in a slightly different notation; no analysis of their exact algebraic content is provided.

Due to its restriction to the case of ordinary spinors, the interest of the above analysis is quite limited. Results of greater physical value would be

REFERENCES

[illegible]erner, H. [19[illegible]]: "Representations of Groups", North-Holland Publishing Company, Amsterdam.

[illegible], T.N. [1964]: "[illegible] in Algebra", Xerox [illegible] Publishing, Waltham, Massachussetts.

[illegible] and Boon, M. [19[illegible]]: "The Theory of Finite Groups. Applications in Physics", North-Holland Publishing Company, Amsterdam.

Messiah, A. [1964]: "Mécanique Quantique", Dunod, Paris.

Pauli, W. [1936]: "Contributions Mathématiques à la Théorie des Matrices de Dirac", Ann. Inst. Henri Poincaré 6, [illegible].

[illegible]enney, F. [1964]: "Algebra of Dirac Bilinears", J. Math. Phys. 5, 1657.

[illegible]chon, J. [1973]: "Groupes de Lie", Hermann, Paris.

[illegible]hl, W. [1970]: "The Lorentz Group and Harmonic Analysis", W.A. Benjamin, Inc., New-York.

Takahashi, Y. [1969]: "An Introduction to Field Quantization", Pergamon Press, Toronto.