MATRIX THEORY AND APPLICATIONS TO RELATIVISTIC WAVE EQUATIONS

Thesis
Submitted to
The University of Madras
for
The Degree of Doctor of Philosophy
by

ACHYUTA RAU TEKUMALLA, M.Sc.,

'MATSCIENCE', THE INSTITUTE OF MATHEMATICAL SCIENCES,
MADRAS-600020, INDIA

JULY 1974

P R E F A C E

This thesis comprises the work done by the author during the period 1970-1974 under the supervision of Professor Alladi Ramakrishnan, Director, MATSCIENCE, The Institute of Mathematical Sciences, Madras.

The thesis consists of the work done by the author in the field of Matrix Theory and Applications to Relativistic Wave Equations. It is divided into Six Chapters; Chapter 1 being the Introduction, Chapter 2 on Bhabha equations for unique mass and spin, Chapter 3 on Relativistic wave equations for massless spin $\frac{1}{2}$ particle, Chapter 4 on Equivalent forms of the Dirac Hamiltonian, Chapter 5 on Unitary Foldy-Wouthuysen transformations for particles of arbitrary spin and Chapter 6 on a variety of applications of matrix methods to physical problems.

Ten papers which form the subject matter of this thesis have been published or are in the course of publication in established journals. Collaboration with some of my colleagues was necessitated by the nature and range of the problems dealt with and due acknowledgement has been made in the chapters.

The author is extremely grateful to Professor Alladi Ramakrishnan for his continued guidance and encouragement throughout the course of this work. He is indebted to MATSCIENCE for offering excellent facilities that enabled this work to be carried out. It is also a great pleasure to thank his colleagues Professor T.S.Santhanam and Dr.I.V.V.Raghavacharyulu and acknowledge the benefit of their valuable collaboration.

MATSCIENCE,
The Institute of Mathematical
Sciences, MADRAS-20.

JULY, 1974

A R Tekumalla

(A.R.TEKUMALLA)

TO ISHWAR, ROHINI and SHALINI

WHO DONT HAVE TO DESERVE ANYTHING FROM ME .

C O N T E N T S

......
......

CHAPTER. I

INTRODUCTION

Dirac's[1] relativistic wave equation for the electron is a fascinating example of the use of matrix theory in physical problems and is the starting point of the study of relativistic wave equations. The generalisations of the Dirac matrices and Clifford algebra and their various mathematical features and physical applications have been systematically studied by Alladi Ramakrishnan[2] and his collaborators in the last few years. Generalisations in a different direction involve attempts to find relativistic wave equations for particles of arbitrary spin.

The main object of this thesis is to study the matrix algebraic aspects of relativistic wave equations for arbitrary spin with special reference to spin-half. We also give, in the last chapter, some other interesting applications of matrix theory to physical problems.

In Chapter II of this work we deal with Bhabha equations for unique mass and spin[3,4]. Dirac himself, and later Fierz and Pauli, attempted generalisations of the Dirac equation for higher spin but their equations involved subsidiary conditions which are not compatible when one considers interactions with external fields. Bhabha[5] wrote linear equations of the form $(\beta_\mu p_\mu + \chi)\psi = 0$ without subsidiary conditions, β_μ being four matrices of suitable dimensions representing the spin properties of the particle but his equations had multi-spin and multi-mass solutions.

Later investigations by Harish-Chandra,[6] Umezawa and Visconti[6] and, more recently, by Capri[7] consist in finding algebraic conditions on the matrices involved in order that the equation describe a particle of unique mass and unique spin. An important consequence of this is that a heirarchy of inequivalent equations result for a particle of any given spin (in particular, for spin-half) by selecting out lower spins when higher spins are present. In this Chapter we offer arguments so that such equations are allowed to describe particles with lower spins. We give a justification for such alternatives by modifying the condition for unique spin to read $J^2\beta_0^{L-1} = s(s+1)\beta_0^{L-1}$ and the Umezawa-Visconti condition to read $\beta_0^{L+1} = \beta_0^{L-1}$, $2s \leqslant L \leqslant 2f$ where f is the maximum spin in the field function and s is the unique spin to be selected. We give a class of algebras of the type $\beta_\mu\{\beta_\nu, \beta_\lambda\} = 2g_{\nu\lambda}\beta_\mu$ leading to such equations for spin half and discuss their representations. We make explicit demonstrations to construct matrices which satisfy any desired minimal equation conditions and any one of our hierarchy of algebras. The procedure is analogous to the method of Alladi Ramakrishnan for finding the higher dimensional L-matrices or representations of Clifford algebra in higher dimensions. We also investigate the diagonalisability of such matrices and the existence of a hermitianising matrix η (such that $\eta\beta_\mu\eta^{-1} = \beta_\mu^\dagger$) which is essential if the equation is to be derivable from a Lagrangian etc. We show that for the

Capri equation, a hermitianising matrix does not exist and if a hermitianising matrix were to exist, the equation reduces to a trivial extension of the Dirac form by adding zero elements to the wave function and bloating its dimension.

We discuss the properties of the new spin-half equation in the presence of a minimally coupled electromagnetic field and show that the magnetic moment is the same as that of a Dirac particle. The non-existence of η in our opinion is not a very serious one as the equations of Fierz-Pauli type due to Fronsdal do not lead to a hermitianising matrix although the equations imply the conservation of several different charge current densities.

In Chapter III, we discuss relativistic wave equations for massless spin-half particles[8,9]. The Weyl equation is one such and is known to be C and P non-invariant but invariant under the combined CP operation. A second way of linearising the Klein-Gordon equation to describe a massless spin-half equation is to put the mass parameter equal to zero in the Dirac equation. Yet another method, and one inequivalent to the ~~above~~ above is to use singular idempotent combinations of the Dirac or Pauli matrices without explicitly putting the mass parameter as zero. This method was envisaged by Bhabha long ago. We show this method can be used to construct CP - noninvariant equations in which there has been some interest in the recent past. We give two component equations for a massless spin-half particle using a non-covariant factorisation of the Klein-Gordon equation as also the spherical factorisation (used by Biedenharn et al in the case of a massive particle). It is interesting to find that these equations can be CP and CPT non-invariant but never T - non-invariant which shows that CP and CPT violation can well be incorporated in a T-invariant relativistic equation for a massless particle. We then analyse the most general four component equation for a massless spin-half particle, starting with the five 4 x 4 matrices obeying the Clifford algebra $\{\alpha_\mu, \alpha_\nu\} = 2g_{\mu\nu}$ Some of these equations are CP and CPT non-invariant. The equation is of the form $(\underline{\alpha}\cdot\underline{p} + mA)\psi = i\frac{\partial\psi}{\partial t}$ where A is a matrix factorisable into two factors at least one of which is nilpotent. Also βA is factorisable into two factors at least one of which is idempotent in the sense $X^2 = aX$. The equations also yield a non-covariant factorisation of the Klein-Gordon equation and the equations given by Tokuoka, Santhanam and Fuschich[10] emerge as special cases of these equations. We study the invariance properties of these equations under the symmetry operations C,P and T and show that it is possible to write an equation with any prescribed invariance properties under these discrete transformations. We also give the transformations relating the Hamiltonian in these cases to the better known Dirac and Weyl Hamiltonians and point out that this similarity does not imply an identity of the invariance properties under the discrete symmetry transformations.

In Chapter IV, we discuss equivalent forms of the Dirac Hamiltonian. Several transformations yielding equivalent representations of the Dirac equation have been proposed, like the Foldy-Wouthuysen (FW) and the Cini-Touschek (CT) transformations and generalisations thereof[11]. We give here[12] a simple and elegant method of obtaining explicit forms of the FW transformation and its generalisations using the U-matrix method of Alladi Ramakrishnan[2]. In particular, we show that if H and H' are two forms of the Hamiltonian for a massive spin-half particle but with the same elementary divisors, they are related by a similarity transformation by S which can be simply written (except when $H' = -H$) as $S = (H + H')\chi$ where χ is non-singular and commutes with H' but is otherwise arbitrary. We then give a new generalisation of the FW transformation using the five 4×4 matrices obeying the Clifford algebra from which the FW and CT transformations and the transformations of Saavedra and De Vries emerge as particular cases. We also show how the U-matrix method mentioned above can be used the obtain the transformation which connects the Dirac Hamiltonian to the Hamiltonian discussed in the recent literature by Biedenharn[13] and others. We also give a new non-covariant factorisation of the Klein-Gordon equation for a massive spin-half particle distinct from the spherical factorisation advocated by Biedenharn and others.

In Chapter V, we[14] give a unitary Foldy-Wouthuysen transformation and discuss related transformations for the Hamiltonian proposed by Weaver, Hammer and Good[15] for particles of arbitrary spin. The technique employed by Weaver, Hammer and Good is to postulate the rest frame wave function, and use a generalised FW transformation. One difficulty, as was realised earlier, is that except in the case of spin-half, the transformation is not unitary. Weaver[15] has constructed a unitary operator for spin one. Recently, there has been some misconception in the literature on the existence of a unitary FW transformation for spin greater than 1. We first establish here the existence of a unitary F W transformation for any spin and then explicitly calculate it. We first establish here the existance of a unitary F W transformation for any spin and then explicitly calculate it. We show that the unitary transformation can be written as $U = SX$ where S is the transformation used by Weaver, Hammer and Good and obtain conditions on X so that U is unitary. An additional condition on the hermiticity of X reduces U to the F W transformation in the conventional form for spin $\frac{1}{2}$ and to Weaver's transformation for spin 1. The explicit calculation of a hermitian X for arbitrary spin involves the evaluation of the function $f(z)$ where z is related to the spin matrices. We give methods for evaluating $f(z\theta)$ generally and also in particular when $f(z\theta)$ is an odd or even function in θ. We use these methods for explicitly writing the extreme relativistic form of the Hamiltonian and the unitary transformation leading to it.

In Chapter VI, we deal with several applications of matrix methods to physical problems. Section A is on quantum mechanics in finite dimensions[16]. We study the usual quantum mechanical commutation relation $[Q, P] = i$ when n, the dimension of the space on which these operators act is finite. Using the representation theory of Clifford Algebra developed by Alladi Ramakrishnan[2] and his collaborators, we explicitly compute the commutator $[Q, P]$ when n is finite. It turns out that the operator is strictly off-diagonal for finite n. This implies no "uncertainty" and no zero point energy, if these concepts have any meaning for finite n. Of course, we show that as $n \to \infty$ continuously, the operator reduces in the limiting case to the Dirac delta function. We elevate the commutator for finite n to what we call "Finite Quantum Mechanics". Section B is on symmetrical component networks for N-phase systems[17]. It has been shown by Parton in a matrix analysis of symmetrial component networks for N-phase systems that when N is prime, only one network is needed to segregate the components but when N contains factors, an additional network is required for each additional factor. We show that this result follows immediately from a simple property of the Sylvester matrix S $(S_{ij} = \omega^{ij}; i = 0, 1, \ldots, N-1)$ (ω = primitive N^{th} root of unity) that when N is composite, a replacement of ω by ω^i cannot be achieved by a permutation of the columns of the Sylvester matrix associated with ω. Further, the Sylvester matrix corresponding to ω^i may become singular in this case, for some values of i and in this case, a new network is required whose elements, however, are a subset of the first network. Section C is on Traces of products of Clifford elements[18]. Formulae exist for the evaluation of the traces of Pauli and Dirac matrices and products of their linear combinations. These are useful when one works with perturbation theory to higher orders. Here we study the trace properties of products of linear combinations of Clifford elements and generalised Clifford elements in n dimensions. The latter yields a possible generalisation of the concept of the pfaffian. In section D we give a new application of the rearrangement operation on a matrix defined by Alladi Ramakrishnan. The eigenvectors of the cyclic matrix C ($C_{ij} = \delta_{i+1,j}$ where $\delta_{rs} = 0$ except when $r = s \bmod n$ are given by the columns of the Sylvester matrix mentioned above. Thus the Sylvester matrix diagonalises any circulant (which is a linear combination of the powers of C. Here we show that the eigenvectors of the matrix D ($D_{ij} = \omega^j \delta_{i+1,j}$) which ω - commutes with C can be obtained from the Sylvester matrix by a repeated application of the rearrangement operation. A particularly interesting case arises when the matrix is even - dimensional in which case we have to perform a 'semi-rearrangement' or a rearrangement by a half-integral number of times which we define here. The matrix so obtained will diagonalise a complex circulant ie., a linear combination of powers of the D matrix. In section E, we give a theorem on the limiting properties of a stochastic matrix[19]. It is well known that for a stochastic matrix S ($S_{ij} \geq 0$ when $i \neq j$, $\sum_i S_{ij} = 0$) when

S is irreducible, $\lim_{t\to\infty} e^{St}\underline{\Pi} = \underline{\Pi}_0$ where $\underline{\Pi}$ is an arbitrary probability vector and $\underline{\Pi}_0$ is the eigenvector of S corresponding to eigenvalue 0 with sum normalised to 1. We have show that if the eigenvalues of S are all real, then after a sufficiently large t, the elements of $e^{St}\underline{\Pi}$ approach the elements of $\underline{\Pi}_0$ monotonically as $t\to\infty$. And when the eigenvalues of S are complex, then as $t\to\infty$, the elements of $e^{St}\underline{\Pi}$ approach the elements of $\underline{\Pi}_0$ as a damped oscillation about the final value. And finally the last section, section F, is on generalised Lucas Polynomials[20]. In connection with the problem of finding the M^{th} power of an N x N matrix, Barakat and Baumann introduced polynomials which they termed the generalised Lucas polynomials satisfying a difference equation and a set of initial conditions and suggested it is desirable to obtain them in closed form. We here show that these polynomials can be obtained directly from the symmetric functions which are of basic importance in combinatorial analysis. We give expressions in closed form for the generalised Lucas polynomials and also for their linear translates which occur in the above method. These expressions are useful in the evaluation of the matrix function f(z) which occurs in Chapter V earlier.

REFERENCES:

1. P.A.M.Dirac, Proc.Roy.Soc. (Lond.),A117, 610 (1928)
2. Alladi Ramakrishnan, "L-matrix Theory or The Grammar of Dirac Matrices", Tata-McGraw Hill, Bombay (1972)
3. A.R.Tekumalla and T.S.Santhanam, Progr.Theor.Phys., 50, 992 (1973).
   T.S.Santhanam and A.R.Tekumalla, communicated to Phys.Rev.
4. T.S.Santhanam and A.R.Tekumalla, Fort.der Phys., to appear Heft 8 or 9 (1974)
5. H.J.Bhabha, Rev. Mod. Phys., 17,200 (1945); 21, 451 (1949).
6. Harish-Chandra, Phys. Rev., 71, 793 (1947).
7. A.Z.Capri, Phys.Rev., 178, 2427 (1969); 187, 1811 (1969). Progr. Theor. Phys., 48, 1364 (1972).
8. T.S.Santhanam and A.R.Tekumalla, Lett. Nuov.Cim., 3,190 (1972).
9. A.R.Tekumalla and T.S.Santhanam, Lett. Nuov.Cim., 6, 99 (1973).
10. Z.Tokuoka, Progr.Theor.Phys., 37, 603 (1967).
   T.S.Santhanam and P.S.Chandrasekharan, Progr. Theor. Phys., 41, 264 (1969).
   V.I. Fushchich and A.C.Grishenko, Lett.Nuov.Cim., 4,927 (1970).
11. L.L. Foldy and S.A.Wouthuysen, Phys.Rev., 78, 29 (1950).
   I.Saavedra, Nucl.Phys. B2, 690 (1967).
   E.De Vries, Physica, 43, 45 (1969)
12. A.R.Tekumalla, communicated to Physica.
13. L.C.Biedenharn and M.Y.Han, Phys.Rev.Lett.,27,1167 (1971).
14. A.R.Tekumalla and T.S.Santhanam, to appear Lett.Nuov.Cim.,
15. D.L.Weaver,C.L.Hammer and R.H.Good,Jr. Phys.Rev., 135, B241 (1964).
   D.L.Weaver, Nuov.Cim., 53A, 665 (1968).
16. T.S.Santhanam and A.R.Tekumalla, communicated to Int.Jour. Theor.Phys.,
17. T.S.Santhanam and A.R.Tekumalla, Matrix and Tensor Quarterly, 21, 121 (1971).
18. A.R.Tekumalla and T.S.Santhanam, Matrix and Tensor Quarterly, 23, 99 (1973).
19. A.R.Tekumalla, Presented at Matscience Conference on Probability Theory and Stochastic Processes,Bangalore (1973).
20. I.V.V.Raghavacharyulu and A.R.Tekumalla, Jour.Math.Phys., 13, 321 (1972).

CHAPTER. II

LINEAR EQUATIONS OF BHABHA TYPE FOR ARBITRARY SPIN[*+]

1. Introduction:

Dirac's epoch making paper of 1948[1] was the starting point of the study of linear relativistic wave equations. The generalisation of the Dirac matrices and Clifford algebra and their various physical applications have been systematically studied by Alladi Ramakrishnan[2] and his collaborators. Generalisations in a different direction involve attempts to find relativistic wave equations for particles of arbitrary spin. A great many such attempts have been made. Dirac himself[3] and later Fierz, and Fierz and Pauli[4] attempted such equations. Soon it was realised that the subsidiary conditions involved are not compatible when one considers interactions with external fields, for instance, the electromagnetic field. This can be rectified only at great cost to the mathematical elegance. Fierz-Pauli equation for a spin 3/2 particle has been written in vector covariant indices (instead of dotted and undotted indices) by Rarita and Schwinger[5]. It is however, desirable to write an equation without auxiliary conditions and also with simple Lorentz transformation properties. Bhabha[6] wrote linear equations of this type without subsidiary conditions but it turned out that his equations, besides having

---

* A.R.Tekumalla and T.S.Santhanam, Progr.Theor.Phys., 50, 982(1973).
+ T.S.Santhanam and A.R.Tekumalla, Fort.Schritte der Phys., (to appear 1974)
T.S.Santhanam and A.R.Tekumalla, communicated to Phys.Rev.,

several spin components, fail to satisfy the Klein-Gordon equation resulting in multi-mass solutions. Later investigations by Harish-Chandra[8,9] and Bhabha[10] in this direction consisted in finding algebraic conditions on the matrices in order that the wave equation describe a particle of unique mass.

In the sixties, equations of Schroedinger type were more popular in the description of particles of higher spin. Attempts by Weinberg[11], by Weaver, Hammer and Good[12] and by Pursey[13] are notable in this direction. In such theories, the Hamiltonian is a polynomial in the derivative operator. It has, however, been shown by Wightman[14] that Weinberg equations lead to unstable representations of the Poincaré group. Physically this meant that the spectrum of the Hamiltonian also contains non-real eigenvalues. Of course, this causes absolutely no trouble in Weinberg's construction of a free field theory where one excludes them from the Fourier expansion of the field by definition.

So a revival of linear relativistic wave equations came up. Techniques have been developed (Capri[15]) to pick up unique spin. The idea is to find algebraic conditions on the matrices entering the linear equation (besides the Harish-Chandra condition to ensure unique mass) to guarantee unique spin. A reversal of the systematics to pick up lower spins results in a class of equations for a spin-half particle, inequivalent to the Dirac equation. Of course, one sacrifices the diagonalisability of β_0 to weaker block form (Capri[16]). But subsequent investigations by Santhanam et al[17] have shown that the matrices involved in Capri's version satisfy a lower degree minimal equation than required. However, the consistency of such equations in the general frame-work has been proved.

But while the new approach justifies the existence of such an alternative equation, it makes it obvious that a hermitianising matrix in this case does not exist[18,19]. If on the other hand, one insists on the existence of a hermitianising matrix (if the equation is to be derivable from a Lagrangian) then the above equation reduces to a trivial extension of the Dirac form by adding zero elements to the wave function and bloating its dimension. However, we do not see any compelling reason to insist on the existence of the hermitianising matrix. In fact, the equations of Fierz-Pauli type due to Fronsdal[20] do not lead to a hermitianising matrix although the equations imply the conservation of several charge current densities.

In section 2 we outline the general method of Bhabha for solving the commutator equations to be satisfied by the matrices. In section 3, we outline the work of Harish-Chandra to get unique mass conditions. In section 4, we discuss the conditions of Umezawa and Visconti[21] and in section 5, the methods of Bakri[22], Hurley[23] and the method of Capri[15] so that the wave function describe a particle of unique mass and unique spin. In section 6, we discuss the reversal problem, of a class of linear equations for a particle of spin half. In section 7, we give the representations of the new algebras involved.

In section 8, representations of the hierarchy of algebras are discussed. We give here a method for constructing matrices which obey these higher algebras starting from the Dirac matrices. In section 9, we study the solutions of the new equations and in section 10, we study the characteristics of the new equations in the presence of a minimally coupled electromagnetic field[18,24,25]. Finally, in section 11, we give our conclusions.

2. The Method of Bhabha:

Bhabha[6,7] investigated relativistic wave equations for describing the behaviour of elementary particles of any integral or half-integral spin on the assumption that these equations must always be written in the absence of interaction in the form

$$(p_\mu \beta^\mu + \chi)\psi = 0 \; ; \; \mu = 0, 1, 2, 3. \tag{2.1}$$

where the p_μ are the differential operators $-i\frac{\partial}{\partial x_\mu}$ and β_μ are four numerical matrices describing the spin properties of the particle. χ is an arbitrary parameter related to the mass. He derived algebraic conditions on the matrices β_μ on the assumption that all properties of the particle are derivable from eq.(2.1) without the use of any further subsidiary conditions. He postulates that the predictions of the theory with regard to the result of any possible observation shall be in accordance with the requirements of the principle of special relativity. This is the requirement of Lorentz invariance.

Any general transformation t of the Lorentz group is one whose coefficients t_{μ}^{ν} are all real and which leaves the metric form unaltered, that is

$$g_{\mu\nu} = g_{\lambda\sigma} t_{\mu}^{\lambda} t_{\nu}^{\sigma}$$

where the repeated indices are summed over and the metric tensor is chosen to have the form

$$g_{00} = -g_{11} = -g_{22} = -g_{33} = 1$$

$$g_{\mu\nu} = 0 \quad \text{for } \mu \neq \nu$$

The effect of a Lorentz transformation on eq.(2.1) is to transform the p_{μ} and β^{μ} to p'_{μ} and β'^{μ} as

$$p'_{\mu} = p_{\nu} (t^{-1})^{\nu}_{\mu}, \quad \beta'^{\mu} = t^{\mu}_{\nu} \beta^{\nu}$$

where t^{-1} is the inverse of t. If the β^{μ} are square matrices of dimension d, then the requirement of invariance under the Lorentz group means that there exists a nonsingular matrix D(t) of dimensions $d \times d$ which brings the transformed matrices α'^{μ} back to their original form by a similarity transformation

$$\beta^{\mu} = D(t) \beta'^{\mu} D(t)^{-1}$$

The β^{μ} are therefore quantities which transform according to

$$\beta'^{\mu} = t^{\mu}_{\nu}\,(D(t)\,\beta^{\nu}\,D(t)^{-1})$$

and the transformation on the wave function is given by

$$\psi(x) \rightarrow \psi'(x') = D^{-1}(t)\,\psi(t^{-1}x)$$

Written in terms of the generators of the homogeneous Lorentz group, the above equation for β'^{μ} reads

$$\begin{aligned}[\beta^{\mu}, I^{\nu\lambda}] &\equiv \beta^{\mu} I^{\nu\lambda} - I^{\nu\lambda}\beta^{\mu} \\ &= g^{\mu\nu}\beta^{\lambda} - g^{\mu\lambda}\beta^{\nu} \qquad (2.2)\end{aligned}$$

where the six matrices $I^{\nu\lambda}$ are the infinitesimal transformations of the representation D . $I^{\nu\lambda}$ are the nucleus of the representation D and are antisymmetric in ν and λ . They satisfy the commutation relations,

$$[I^{\mu\nu}, I^{\lambda\sigma}] = -g^{\mu\lambda} I^{\nu\sigma} + g^{\mu\sigma} I^{\nu\lambda} + g^{\nu\lambda} I^{\mu\sigma} - g^{\nu\sigma} I^{\mu\lambda}$$

Conversely, any set of six matrices satisfying this relation can be used to build up a representation of the homogenous Lorentz group. We need concern ourselves first only with an irreducible set of matrices β^{μ} and D satisfying eq.(2.2), since every other representation can be made up of a direct sum of these. The representation D by itself is not irreducible and therefore its nucleus consisting of the six matrices $I^{\mu\nu}$ is not either. But in this problem, the six $I^{\mu\nu}$ have also to satisfy eqs.(2.2) and the collection of the ten matrices $I^{\mu\nu}$ and β^{μ} is irreducible.

Bhabha investigated the most general solution for β_{μ} of the commutator eq.(2.2). In fact, χ can also be an operator the only condition on that being that it should commute with $I_{\mu\nu}$. This property has been particularly used by Tokuoka[26], Sen-Gupta[27] and Santhanam and Chandrasekaran[28] for constructing equation for neutrino and by Brulin and Hjalmars[29] for spin-2 theories (see also Harish-Chandra[8]).

Define the spinor

$$A^{l\dot{m}} = \beta^{\mu}\sigma_{\mu}^{l\dot{m}}, \quad \beta^{\mu} = \frac{1}{2}\,\sigma_{\dot{m}l}^{\mu}A^{l\dot{m}} \tag{2.3}$$

where the σ's are the Pauli matrices and the rows and columns are labelled by upper undotted and dotted indexes respectively. The Greek indices take values 0,1,2,3 and the Latin indices 1,2,3. σ_0 is the unit matrix. Similarly, define the two symmetric spinors,

$$4K^{m}_{n} = -I_{\mu\nu}\,\sigma^{\mu m\dot{l}}\,\sigma^{\nu}_{\dot{l}n} \tag{2.4}$$

$$4L^{\dot{l}}_{\dot{m}} = I_{\mu\nu}\,\sigma^{\mu}_{\dot{m}n}\,\sigma^{\nu n\dot{l}} \tag{2.4a}$$

Then eqn.(2.2) takes the form

$$2[A^{k\dot{l}}, K^{sr}] = \epsilon^{ks}A^{r\dot{l}} + \epsilon^{kr}A^{s\dot{l}} \tag{2.5}$$

$$2[A^{k\dot{l}}, L^{\dot{m}\dot{n}}] = \epsilon^{\dot{l}\dot{m}}A^{k\dot{n}} + \epsilon^{\dot{l}\dot{n}}A^{k\dot{m}} \tag{2.6}$$

where ϵ_{mn} and ϵ^{mn} are defined as

$$\epsilon_{11} = \epsilon_{22} = \epsilon^{11} = \epsilon^{22} = 0$$

$$\epsilon_{12} = -\epsilon_{21} = \epsilon^{12} = -\epsilon^{21} = 1$$

The matrices $I_{\rho\sigma}$ in eqn. (2.2) must be a reducible representation of the nucleus of the proper Lorentz group and they can be written in the form

$$I_{\rho\sigma} = \begin{bmatrix} I_{\rho\sigma}(k_1, l_1) & & & \\ & I_{\rho\sigma}(k_2, l_2) & & \\ & & \cdot & \\ & & & \cdot \\ & & & & \cdot \end{bmatrix} \tag{2.7}$$

where all the empty rectangles are filled with zeros. K and L take a form similar to that of $I_{\rho\sigma}$. Corresponding to this reduction eqns.(2.5) and (2.6) become

$$\langle k_\sigma l_\sigma | A^{k\dot{l}} | k_\tau l_\tau \rangle K^{sr}(k_\tau) - K^{sr}(k_\sigma) \langle k_\sigma l_\sigma | A^{k\dot{l}} | k_\tau l_\tau \rangle$$

$$= \frac{1}{2} \epsilon^{ks} \langle k_\sigma l_\sigma | A^{r\dot{l}} | k_\tau l_\tau \rangle - \frac{1}{2} \epsilon^{kr} \langle k_\sigma l_\sigma | A^{s\dot{l}} | k_\tau l_\tau \rangle \tag{2.8}$$

$$\langle k_\sigma l_\sigma | A^{k\dot{l}} | k_\tau l_\tau \rangle L^{\dot{m}\dot{n}}(l_\tau) - L^{\dot{m}\dot{n}}(l_\sigma) \langle k_\sigma l_\sigma | A^{k\dot{l}} | k_\tau l_\tau \rangle$$

$$= \frac{1}{2} \epsilon^{\dot{l}\dot{m}} \langle k_\sigma l_\sigma | A^{k\dot{n}} | k_\tau l_\tau \rangle + \frac{1}{2} \epsilon^{\dot{l}\dot{n}} \langle k_\sigma l_\sigma | A^{k\dot{m}} | k_\tau l_\tau \rangle \tag{2.9}$$

from which it is clear that

$$\langle k_\sigma l_\sigma | A^{k\dot{l}} | k_\tau l_\tau \rangle = \langle k_\sigma | A^{k} | k_\tau \rangle \langle l_\sigma | A^{\dot{l}} | l_\tau \rangle \tag{2.10}$$

the expansion on the right being really the direct product of two rectangular matrices of dimensions $(2k_\sigma+1)(2k_\tau+1)$ **and** $(2l_\sigma+1)(2l_\tau+1)$ **respectively. The equations (2.8) and (2.9) reduce to**

$$\langle k_\sigma | A^{k} | k_\tau \rangle K^{sr}(k_\tau) - K^{sr}(k_\sigma) \langle k_\sigma | A^{k} | k_\tau \rangle$$

$$= \frac{1}{2} \epsilon^{ks} \langle k_\sigma | A^{r} | k_\tau \rangle + \frac{1}{2} \epsilon^{kr} \langle k_\sigma | A^{s} | k_\tau \rangle \tag{2.11}$$

$$\langle l_\sigma | A^{\dot{l}} | l_\tau \rangle L^{\dot{m}\dot{n}}(l_\tau) - L^{\dot{m}\dot{n}}(l_\sigma) \langle l_\sigma | A^{\dot{l}} | l_\tau \rangle$$

$$= \frac{1}{2} \epsilon^{\dot{l}\dot{m}} \langle l_\sigma | A^{\dot{n}} | l_\tau \rangle + \frac{1}{2} \epsilon^{\dot{l}\dot{n}} \langle l_\sigma | A^{\dot{m}} | l_\tau \rangle \tag{2.12}$$

Dirac, in connection with the direct products of representations $D(\frac{1}{2},\frac{1}{2})$ and $D(k,l)$ of the Lorentz group introduced the two matrices $u^r(k)$ and $v^r(k)$ of dimensions $(2k+1)\times 2k$ and $2k\times(2k+1)$ respectively[30]. These matrices satisfy

$$-u_m(k+\tfrac{1}{2})v^m(k+\tfrac{1}{2}) = v_m(k)u^m(k) = 2k+1 \quad (2.13a)$$

$$v_m(k)v^m(k+\tfrac{1}{2}) = u_m(k+\tfrac{1}{2})u^m(k) = 0 \quad (2.13b)$$

$$-v^m(k+\tfrac{1}{2})u_n(k+\tfrac{1}{2}) = K^m_n(k) + (k+1)\delta^m_n \quad (2.14a)$$

$$u^m(k)v_n(k) = K^m_n(k) - k\delta^m_n \quad (2.14b)$$

After a little algebraic manipulation we can show that

$$u^r(k)K^{mn}(k-\tfrac{1}{2}) - K^{mn}(k)u^r(k) = \frac{1}{2}\epsilon^{rm}u^n(k) + \frac{1}{2}\epsilon^{rn}u^m(k) \quad (2.15)$$

$$v^r(k)K^{mn}(k) - K^{mn}(k-\tfrac{1}{2})u^r(k) = \frac{1}{2}\epsilon^{rm}v^n(k) + \frac{1}{2}\epsilon^{rn}v^m(k) \quad (2.16)$$

It is the similarity of forms between the eqns.(2.11), (2.12) and (2.15), (2.16) that led Bhabha to the conclusion that the non-vanishing elements of the matrix A and hence β are given by

$$\langle k,l \mid A^{k\dot{l}} \mid k+\tfrac{1}{2}, l-\tfrac{1}{2}\rangle = c\,v^{k}(k+\tfrac{1}{2})\,u^{\dot{l}}(l) \tag{2.17a}$$

$$\langle k+\tfrac{1}{2}, l-\tfrac{1}{2} \mid A^{k\dot{l}} \mid k,l\rangle = d\,u^{k}(k+\tfrac{1}{2})\,v^{\dot{l}}(l) \tag{2.17b}$$

$$\langle k,l \mid A^{k\dot{l}} \mid k+\tfrac{1}{2}, l+\tfrac{1}{2}\rangle = c\,v^{k}(k+\tfrac{1}{2})\,v^{\dot{l}}(l+\tfrac{1}{2}) \tag{2.18a}$$

$$\langle k+\tfrac{1}{2}, l+\tfrac{1}{2} \mid A^{k\dot{l}} \mid k,l\rangle = d\,u^{k}(k+\tfrac{1}{2})\,u^{\dot{l}}(l+\tfrac{1}{2}) \tag{2.18b}$$

where, by properly choosing a similarity transformation leaving $I_{\rho\sigma}$ invariant one can reduce the constants to obey $|c| = |d|$ when $cd \neq 0$. The matrices we shall be considering in section 6 and the ones considered Capri fall under the special case when the product may be zero. This procedure enables one to completely determine the form of $\beta\mu$ apart from the above arbitrary constants.

Wild[31] has obtained the form of the matrices $\beta\mu$ in the J^2 diagonal form where J^2 is the invariant of the rotation subgroup. Wild's form is very useful in the discussion of spins. The non-vanishing components of β_0 in the Wild basis are given by

$$\langle (k,l)_\sigma | \beta_0 | (k-\tfrac{1}{2}, l+\tfrac{1}{2})_\tau \rangle \rightarrow c_{\sigma\tau} [(k+j-l)(j+l-1+k)]^{\frac{1}{2}} \delta_{jj'} \otimes I ,$$

$$\langle (k,l)_\sigma | \beta_0 | (k-\tfrac{1}{2}, l-\tfrac{1}{2})_\tau \rangle \rightarrow c_{\sigma\tau} (-1)^{k+l+j} [(k+l-j)(k+l+j+1)]^{\frac{1}{2}} \times \delta_{jj'} \otimes I ,$$

$$\langle (k, k+\tfrac{1}{2})_\sigma | \beta_0 | (k+\tfrac{1}{2}, k)_\tau \rangle \rightarrow c_{\sigma\tau} (-1)^{\{j\}+1} (j+\tfrac{1}{2}) \delta_{jj'} \otimes I ,$$

(2.19)

where I represents a $(2j+1)(2j+1)$ dimensional unit matrix,

$|k-l| \leqslant j' \leqslant k+l$ **and** $\{j\}$ **= integral part of j.**

All other components are obtained from

$$\langle kl | \beta_0 | k'l' \rangle = (-1)^{2k+2} \langle k'l' | \beta_0 | kl \rangle ,$$

$$\langle kl | \beta_0 | k'l' \rangle = -\langle lk | \beta_0 | l'k' \rangle .$$

(2.20)

Further, if r,t,s,u denote two pairs of inequivalent irreducible representations of the proper Lorentz group such that under parity $r \to t$, $s \to u$ **then** $c_{rs} = -c_{tu}$

The generators $I_{\rho\sigma}$ **obey the commutator equation**

$$[I_{\rho\sigma}, I_{\tau\upsilon}] = -g_{\rho\tau} I_{\sigma\upsilon} + g_{\rho\upsilon} I_{\sigma\tau} + g_{\sigma\tau} I_{\rho\upsilon} - g_{\sigma\upsilon} I_{\rho\tau} .$$

(2.21)

Bhabha tried to close the algebra satisfied by the $I_{\rho\sigma}$ and β_μ namely eqs. (2.2) and (2.21) by assuming that

$$I_{\rho\sigma} = \lambda [\beta_\rho, \beta_\sigma] \tag{2.22}$$

which we know is true for the Dirac and Kemmer equations. λ is a numerical constant. Eq. (2.22) does not hold for any of the alternative forms of the equations for spin greater than one given by Dirac, Fierz and Pauli. But eq. (2.22) is of course consistent with eqs. (2.2) and (2.21) for any spin as can be verified by direct substitution of eq. (2.22) in eq. (2.2) to obtain eq.(2.21). Now $I_{\rho\sigma}$ and β_μ form the Lorentz group in five dimensions. For, by introducing the additional index 4 and defining

$$I^{\mu 4} \equiv \beta^\mu ; \quad g^{44} = -1, \quad g^{m4} = 0, \; m \neq 4,$$

eqs. (2.2), (2.21) and (2.22) are all combined into one equation namely eq. (2.2) where the indices now run from 0 to 4. The ten matrices I^{LM} with $L, M = 0, 1, \ldots, 4$ now satisfy the same commutation rules as the ten infinitesimal transformations of the Lorentz group in five dimensions. The problem of finding the irreducible representations of the Lorentz group can be related to that of finding the irreducible representations of the orthogonal group in five dimensions, the solution of which is well known. In this case the wave function transforms as a representation $R_5(\lambda_1, \lambda_2)$ where λ_1, λ_2 are both integers or zero, or both half-integers with $\lambda_1 \geq \lambda_2 \geq 0$, and this describes a particle of maximum spin λ_1, irrespective of λ_2. For particles with a maximum spin f, the number of different equations possible is $f+\frac{1}{2}$ if f is half integer and f + 1 if f is an integer and these equations are inequivalent and describe particles with different physical properties, the only equation with unique spin s being obtained from $R_5(s,s)$ Further the assumption of eq. (2.22) implies that a particle of spin greater than one must appear with several values for the rest mass with rational multiples of the lowest value, as each component of the wave function by itself does not satisfy a second order wave equation but one of higher order consisting of products of the usual second order wave operators as will be shown in the next section. Thus a particle of spin 3/2 must appear with two possible rest masses, one three times the other while a particle of spin 2 has two rest masses one twice the other, the lower values of the rest mass being more stable in each case. The condition for unique mass due to Harish-Chandra will be considered in the next section.

It has also been shown by Madhava Rao, et al[32,33] that in the case of spin 3/2 the algebra of β's is simply the direct product of the Dirac algebra and another algebra (ξ - algebra) of rank 42 so that the original algebra is of rank 672. Madhava Rao et al[33] have also shown that the ξ- algebra has three irreducible representations of dimensions 1,4 and 5 and have explicitly written the irreducible representations.

3. The Harish-Chandra Condition:

In 1947 Harish-Chandra showed that by imposing certain conditions on the minimal equation obeyed by the matrices $\beta\mu$ which are solutions of the commutator eq.(2.2), the wave function can be made to obey a Klein-Gordon equation with unique mass.

Let the minimal equation for β_0 be given by

$$\beta_0^m + a_1\beta_0^{m-1} + \cdots + a_m = 0 \qquad (3.1)$$

This can be written in the form

$$\beta_0^m + g_{00}\, a_2\beta_0^{m-2} + (g_{00})^2 a_4\beta_0^{m-4} + \cdots$$
$$+ (a_1\beta_0^{m-1} + g_{00}\, a_3\beta_0^{m-3} + \cdots) = 0$$

β and $t\beta$ are equivalent, t being a Lorentz transformation. Therefore $\beta_0' = t_0^{\mu}\beta_\mu$ has the same minimal equation. Put $t_0^{\mu} = y^{\mu}$. Then

$$y^{\mu}\beta_\mu = t_0^{\mu} t_0^{\nu} g_{\mu\nu} = g_{00} = 1 \qquad (3.2)$$

and $\beta_0' = y^{\mu}\beta_\mu$. Then

$$\{(y^{\nu}\beta_\nu)^m + a_2(y^{\mu_1}y^{\mu_2}g_{\mu_1\mu_2})(y^{\nu}\beta_\nu)^{m-2}$$
$$+ a_4(y^{\mu_1}y^{\mu_2}g_{\mu_1\mu_2})^2(y^{\nu}\beta_\nu)^{m-4} + \cdots\}$$
$$+ \{a_1(y^{\nu}\beta_\nu)^{m-1} + a_3 y^{\mu_1}y^{\mu_2}g_{\mu_1\mu_2}(y^{\nu}\beta_\nu)^{m-3} + \cdots\} = 0$$

Now the y's are arbitrary except for eq.(3.2). Hence put

$$y^\nu = \frac{z^\nu}{(g_{\lambda\sigma} z^\lambda z^\sigma)^{\frac{1}{2}}} \equiv \frac{z^\nu}{Z},$$

where the z's are independent indeterminates. Now on multiplying by Z^m we find

$$\{(z^\nu \beta_\nu)^m + a_2 z^{\mu_1} z^{\mu_2} g_{\mu_1\mu_2} (z^\nu \beta_\nu)^{m-2} + \cdots\}$$

$$+ Z\{a_1 (z^\nu \beta_\nu)^{m-1} + \cdots\} = 0 \qquad (3.3)$$

This must be an identity in z. Therefore from the irrationality of Z in z, we get

$$a_1 (z^\nu \beta_\nu)^{m-1} + \cdots = 0$$

Now put $z^0 = 1,\ z^\ell = 0,\ \ell \neq 0$. Then

$$a_1 \beta_0^{m-1} + \cdots = 0$$

If $a_1, a_3, \cdots$ are not all zero, this is an equation satisfied by β_0 of degree lower than the minimal equation which is clearly impossible since, by definition, the minimal equation is the equation of lowest degree satisfied by β_0. Hence, $a_1 = a_3 = \cdots = 0$ and only the expression in the first brackets of eq. (3.3) survive. Equating coefficients of z in this equation, we get

$$\sum^{(p)}(\beta_{\mu_1}\cdots\beta_{\mu_m}+a_2 g_{\mu_1\mu_2}\beta_{\mu_3}\cdots\beta_{\mu_m}$$

$$+a_4 g_{\mu_1\mu_2}g_{\mu_3\mu_4}\beta_{\mu_5}\cdots\beta_{\mu_m}+\cdots)=0 \qquad (3.4)$$

where $\sum^{(p)}$ denotes a summation over all permutations of $\mu_1, \mu_2, \ldots, \mu_m$. From eq.(3.4), it follows

$$\{(p^\mu\beta_\mu)^m+a_2p^2(p^\mu\beta_\mu)^{m-2}+a_4p^4(p^\mu\beta_\mu)^{m-4}+\cdots\}\psi=0 \qquad (3.5a)$$

where $p^\mu p^\mu = p^2$. On using the wave equation (2.1), this reduces to

$$(\chi^m+a_2\chi^{m-2}p^2+a_4\chi^{m-4}p^4+\cdots)\psi=0 \qquad (3.5b)$$

Eq. (3.5a) and (3.5b) contain only even or odd powers of $p^\mu\beta_\mu$ and χ respectively and may be written in factorised form as

$$(p^\mu\beta_\mu-c_1^2p^2)(p^\mu\beta_\mu-c_2^2p^2)\cdots=0 \qquad (3.6a)$$

and

$$(\chi^2-c_1^2p^2)(\chi^2-c_2^2p^2)\cdots\cdots=0 \qquad (3.6b)$$

respectively with a further multiplication by $p^{\mu}\beta_{\mu}$ and χ respectively if the spin is integral. $c_1, c_2, \ldots$ and the non zero eigenvalues of β_0. Eq. (3.5b), or (3.6b), is the differential equation (with numerical coefficients) of the lowest order which the most general solution ψ of the wave eq. (2.1) would satisfy. It is clear from eq. (3.6a) or (3.6b), how the multiple mass arises in Bhabha's theory. If we now demand that eq. (3.5b) reduce to the usual second order Klein Gordon equation, we get $a_4 = a_6 = \cdots = 0$. Therefore from eq. (3.4)

$$\sum{}^{(p)} \beta_{\mu_1} \cdots \beta_{\mu_m} = \sum{}^{(p)} g_{\mu_1\mu_2} \beta_{\mu_3} \cdots \beta_{\mu_m} \tag{3.7}$$

where $\sum^{(p)}$ is the sum over all permutations of the indices. This eq. (3.7) is completely equivalent to

$$(z^{\mu}\beta_{\mu})^{m} = z^{2}(z^{\mu}\beta_{\mu})^{m-2} \tag{3.8}$$

(where the z are arbitrary) which is the natural generalisation of the Dirac matrices for which

$$(z^{\mu}\beta_{\mu})^{2} = z^{2}$$

and of the Duffin Kemmer matrices for which

$$(z^{\mu}\beta_{\mu})^{2} = z^{2}(z^{\mu}\beta_{\mu})$$

When we take $z_0 = 1$ and $z_j = 0$, we get the equation in terms of β_0, namely

$$\beta_0^m = \beta_0^{m-2} \qquad m \geqslant 2 \tag{3.9}$$

or

$$\beta_0^{n+1} = \beta_0^{n-1} \qquad n \geqslant 1$$

which is the Harish-Chandra condition in the well known form. One can also look upon it as the condition that β_0 has only one pair of non-zero eigenvalues which, however, can occur several times.

Since eq.(3.4) is already in a proper tensor form, one might at first sight think that the invariance of an irreducible representation under the proper Lorentz group is guaranteed. This, however, is in general, not the case, since for $n \geqslant 2$ eq. (3.7) does not generate a finite algebra. Some other stronger tensor condition compatible with eq. (3.7) is required to make the algebra finite. Thus for example, for n = 2 we have the Duffin Kemmer commutation relations

$$\beta_\mu \beta_\nu \beta_\lambda + \beta_\lambda \beta_\nu \beta_\mu = g_{\mu\nu} \beta_\lambda + g_{\lambda\nu} \beta_\mu$$

In the Dirac and the Duffin Kemmer cases, one can choose a representation consistent with eq. (3.7) in which the β_μ are Hermitian or anti-hermitian. But this is obviously impossible for n > 2 for if β_0 is hermitian, it is diagonalisable. Further, since β_0 satisfies the minimal equation $\beta_0^{n-1}(\beta_0^2 - 1) = 0$ all its eigenvalues satisfy the equation too and hence its eigenvalues are all 0 or $\pm$ 1 . But a diagonal matrix (and hence a diagonalisable matrix) with eigenvalues 0, ± 1 obviously satisfies $\beta_0^3 = \beta_0$. Hence for hermitian β_0, $n \leq 2$.

4. Umezawa-Visconti Condition:

We shall now discuss the proof due to Umezawa and Visconti[21] to show that n = 2f where f is the maximum spin contained in the field function. They assume the existence of the Klein-Gordon divisor $d(\partial)$ such that

$$d(\partial)\,\Lambda(\partial) = (\Box - \chi^2)\, I \tag{4.1}$$

where

$$\Lambda(\partial) = -(\beta_\mu \partial_\mu + \chi) \tag{4.2}$$

They also assume that $d(\partial)$ is a polynomial in the derivative operators

$$\begin{aligned} d(\partial) &= \alpha_0 + \alpha_\mu \partial_\mu + \alpha_{\mu\nu}\partial_\mu \partial_\nu + \cdots \\ &= \sum_{l=0}^{L} \alpha_{\mu_1 \cdots \mu_l}\, \partial_{\mu_1} \partial_{\mu_2} \cdots \partial_{\mu_l} \end{aligned} \tag{4.3}$$

where the coefficients $\alpha_{\mu_1 \cdots \mu_l}$ are obviously symmetric in the indices. They show that for a consistent quantisation of the field, the components of the wave function ψ obey the commutation relations $[\psi_\alpha(x), \bar{\psi}_\beta(x')] = i d_{\alpha\beta}(\partial)\Delta(x-x')$ Thus $d(\partial)$ transforms like the tensor product $\psi \otimes \psi$ of the field function. Since ψ contains spins $(f, f-1, \cdots)$ where f is the maximum spin, by simple Clebsch-Gordon theorem it follows that $d(\partial)$ contains spins (2f,2f-1,......). Eq.(4.3) can be rewritten as

$$d(\partial) = \sum_{l=0}^{2f} \alpha_{\mu_1 \cdots \mu_l} \partial_{\mu_1} \partial_{\mu_2} \cdots \partial_{\mu_l} + \sum_{l'=2f+1}^{L} \alpha_{\mu_1 \cdots \mu_{l'}} \partial_{\mu_1} \partial_{\mu_2} \cdots \partial_{\mu_{l'}}$$

The terms for which $l > 2f$ can be regrouped as

$$\alpha'_{\mu_1 \cdots \mu_l} \partial_{\mu_1} \cdots \partial_{\mu_{2f}} (\Box)^{(l-2f)/2}$$

$$= \alpha_{\mu_1} \cdots \alpha_{\mu_l} \partial_{\mu_1} \cdots \partial_{\mu_l} ;$$

$$l > 2f$$

since 2f is the maximum spin in $d(\partial)$ and the rest of the terms can only contribute to a power of $\Box$. Obviously $\alpha_{\mu_1 \cdots \mu_l}$ must be zero for $l-2f$ odd.

From eq. (4.1) it follows that

$$\alpha_0 \chi = \chi^2 I$$

$$\alpha_0 \beta_\mu + \chi \alpha_\nu = 0$$

$$(\alpha_\mu \beta_\nu + \alpha_\nu \beta_\mu) + 2\chi\, \alpha_{\mu\nu} = -2 g_{\mu\nu}$$

$$\sum^{(p)} \beta_{\mu_1} \alpha_{\mu_1 \ldots \mu_\ell} - \chi\, \alpha_{\mu_1 \ldots \mu_\ell} = 0$$

$$\text{for } \ell > 2 \tag{4.4}$$

where $\sum^{(p)}$ denotes the summation over terms given by taking all possible permutations over the suffixes. Using these equations it follows that

$$\alpha_{\mu_1 \ldots \mu_\ell} = 0; \quad \ell \geq 2f \tag{4.5}$$

even when $\ell - 2f$ is even. Hence the polynomial $d(\partial)$ should terminate at $L \leq 2f$. If in addition we require the field function to contain the maximum spin f,

$\alpha_{\mu_1 \ldots \mu_{2f}} \neq 0$ then $L = 2f$. This is the proof of Umezawa and Visconti.

5. <u>Equations for Unique Spin and Unique Mass:</u>

As we are here mainly concerned with linear relativistic equations of Bhabha type, we do not discuss other types of equations where the problem of unique spin and mass has also been attacked. In such cases, subsidiary conditions of the nature of differential equations are imposed to select out unique spin and unique mass[13,14]. Therefore the general problem of describing a particle with unique mass and unique spin satisfying a linear relativistic equation of Bhabha's type exists. Recently, Bakri[22] has derived field equations of the form

$$\mathcal{P}_\mu \equiv (I_{\mu\nu} p_\nu + \chi \beta_\mu)\psi = p_\mu \psi \tag{5.1}$$

which describe particles of definite mass and definite spin. Here β_μ are the Bhabha matrices. These matrices were derived from the finite dimensional representation of the inhomogeneous de Sitter group $SO(4,1)$ with the invariant $p_A p_A = 0$ where p_A is the momentum-energy-mass five vector with components p_K = cartesian components of momentum $p_4 = ip_0$ p_0 = energy, and $p_5 = \chi$. S_{AB}, the generators of the homogeneous de Sitter group are given by

$$S_{\mu\nu} \equiv I_{\mu\nu} = \lambda_1 [\beta_\mu, \beta_\nu], \tag{5.2}$$

$$S_{\mu 5} = \beta_\mu , \tag{5.3}$$

and their finite dimensional irreducible representations are those of SO(5), $R_5(\lambda_1, \lambda_2)$ characterised by two non-negative integers (for bosons) or half-integers (for fermions), such that $\lambda_1 > \lambda_2$. Eq.(5.1) leads directly to the Bhabha equation

$$(\beta_\mu p_\mu + \chi)\psi = 0, \tag{5.4}$$

to the mass relation

$$p_A p_A \psi \equiv (p_\mu p_\mu + \chi^2)\psi = 0 \tag{5.5}$$

and to the equation

$$W_\mu W_\mu \psi = \chi^2 \lambda_2(\lambda_2+1)\psi, \tag{5.6}$$

where

$$W_\mu = \frac{\lambda_1}{2} \epsilon_{\mu\nu\alpha\beta} p_\nu I_{\alpha\beta} \tag{5.7}$$

is the Pauli Lubanski pseudovector. ψ decomposes under the Poincare group into several components $\psi(S_1, S_2)$. $\psi(S_1, S_2)$ transforms according to the irreducible representation $D(S_1, S_2)$ of the homogeneous Lorentz group. SO(4,1) decomposes under the homogeneous Lorentz group into all inequivalent $D(S_1, S_2)$

with

$$\delta \leq |S_1 - S_2| \leq \lambda_2 \leq S_1 + S_2 \leq \lambda_1 \tag{5.8}$$

where $\delta = 0$ for bosons and $\delta = \frac{1}{2}$ for fermions. Eqns. (5.5) and (5.6) are satisfied by each Poincare component separately. Eqn. (5.6) defines the spin in the rest system. Thus the theory describes a particle with a definite mass and definite spin λ_2. The important point in such an approach is that the Lorentz $I_{\mu\nu}$ and the Bhabha matrices β_μ are related by the above eqn. (5.2) while we know very well that apart from the Dirac and the Kemmer equations, the other known equations for higher spins do not respect this relation if they are to be described completely by Bhabha eqn. (5.4). Bakri overcomes this by taking only solutions of eqn. (5.1) and not (5.4).

Another class of relativistic wave equations has been derived by Hurley[23]. He begins with Galilei-covariant wave equations for massive particles with any integer or half-integer spin and generalises them to obtain Lorentz invariant equations. He imposes a minimality condition on the number of components possessed by the relativistic wave function and shows that the index transformation properties of the wave function may be either those of the $D = (S,0) \oplus (S-\frac{1}{2},\frac{1}{2})$ representation or $D = (0,S) \oplus (\frac{1}{2},S-\frac{1}{2})$ and with the incorporation of reflection symmetry

$$D = (s,0) \oplus (s-\tfrac{1}{2}, \tfrac{1}{2}) \oplus (\tfrac{1}{2}, s-\tfrac{1}{2}) \oplus (0,s)$$

Thus the wave function possesses 4 (2s + 1) independent components, has no subsidiary conditions and describes a particle with unique mass $m \neq 0$ and a unique spin. His equation is of the form

$$(p_\mu \beta^\mu + m)\psi = 0$$

where β_μ are the matrices

$$\beta_0 = \begin{array}{c} \left[\begin{array}{cccccc} 0 & 1 & 0 & & & \\ 1 & 0 & 0 & & 0 & \\ 0 & 0 & 0 & & & \\ & & & 0 & 1 & 0 \\ & 0 & & 1 & 0 & 0 \\ & & & 0 & 0 & 0 \end{array}\right] \end{array} \begin{array}{l} (2S+1) \\ (2S+1) \\ (2S-1) \\ (2S+1) \\ (2S+1) \\ (2S-1) \end{array}$$

$$\beta_i = \left[\begin{array}{cccccc} 0 & -S_i & K_i^\dagger & & & \\ S_i & 0 & 0 & & 0 & \\ K_i & 0 & 0 & & & \\ & & & 0 & S_i & -K_i^\dagger \\ & 0 & & -S_i & 0 & 0 \\ & & & K_i & 0 & 0 \end{array}\right]$$

The term in paranthesis are the number of rows in the corresponding sub-matrix, $\underline{S}$ are the (2s + 1) x (2s + 1) dimensional spin matrices and $\underline{K}$ are (2s-1) x (2s+1) dimensional rectangular matrices defined upto a phase by

$$K_i S_j - \Sigma_j K_i = i \epsilon_{ijk} K_k$$

and

$$S_i S_j + K_i^{+} K_j = i s \epsilon_{ijk} S_k + s^2 \delta_{ij}$$

where Σ_i are the 2s - 1 dimensional spin matrices , ϵ_{ijk} is the Levi-Civita symbol and δ_{ij} is the Kronecker symbol.

The matrices of Hurley are easily understood in terms of Bhabha's analysis, though Hurley arrives at them without recourse to Bhabha's work. In section 2, we found that Bhabha has given the most general solution to the commutator eq. (2.2) when ψ transforms like the representation of the Lorentz group

$$D = \sum \oplus n_i D(k_i, l_i)$$

and he showed that the only non-vanishing elements of β_μ are $\langle k, l | \beta^\mu | k \pm \frac{1}{2}, l \pm \frac{1}{2} \rangle$. Therefore the simplest non-trivial construction for β_μ satisfying eq. (2.2) is obtained by the choice of D as

$$D = (k, l) \oplus (k \pm \tfrac{1}{2}, l \pm \tfrac{1}{2})$$

To obtain spin s, the simplest choice is

$$D = (s,0) \oplus (s-\tfrac{1}{2}, \tfrac{1}{2})$$

Hurley's choice is thus the simplest choice of D covered in Bhabha's general analysis although he arrived at it without recourse to it. In fact Hurley's β_0 is identical to what one would obtain using Bhabha's prescription and transforming to the J^2 diagonal basis. It is a matter of simple verification that Hurley's matrices obey the algebra

$$\sum^{(p)} (\beta_\mu \beta_\nu \beta_\lambda - g_{\mu\nu} \beta_\lambda) = 0$$

and hence the minimal equation $\beta_0^3 = \beta_0$ or, in general,

$$(p_\mu \beta^\mu)^3 = p^2 (p_\mu \beta^\mu)$$

Here $\sum^{(p)}$ is a sum over all permutations of the indices. This minimal equation is a characteristic of Duffin - Kemmer matrices for a spin 1 particle and the algebra has the Duffin-Kemmer algebra as a subalgebra. Of course, since β_μ satisfies a Cubic minimal equation, it is diagonalisable and can be hermitian or anti-hermitian as indeed Hurley's matrices are. The matrix also satisfies the Harish-Chandra condition with n = 1 and hence the equation describes a particle of unique mass. However, when $s > 1$, the matrices do not satisfy the

Umezawa-Visconti condition for a consistent quantisation of the field since the maximum spin in the field function is s while the minimal equation is always cubic[23]. In fact the Klein-Gordon divisor defined (as in section 4) as

$$d(\partial)\Lambda(\partial) = (\Box - m^2) I,$$

$$d(\partial) = \sum_{l=0}^{L} \alpha_{\mu_1} \cdots \alpha_{\mu_l} \partial_{\mu_1} \cdots \partial_{\mu_l}$$

$$\Lambda(\partial) = -(\beta_\mu \partial_\mu + m)$$

and which can be rewritten as

$$d(\partial) = m + i\underline{\beta}\cdot\underline{\partial} + \frac{1}{m}\left[(i\underline{\beta}\cdot\underline{\partial})^2 - \Box\right]^2 + \frac{1}{m^2}\left[(i\underline{\beta}\cdot\underline{\partial})^3 - \Box(i\underline{\beta}\cdot\underline{\partial})\right] + \cdots + \frac{1}{m^n}\left[(i\underline{\beta}\cdot\underline{\partial})^{n+1} - \Box(i\underline{\beta}\cdot\underline{\partial})^{n-1}\right]$$

cannot contain terms beyond n = 2 because of the minimal equation and is given for the whole class of equations, for any spin, by

$$d(\partial) = m + (i\underline{\beta}\cdot\underline{\partial}) + \frac{1}{m}\left[(i\underline{\beta}\cdot\underline{\partial})^2 - \Box\right]$$

Umezawa and Visconti have shown that for a consistent quantisation of the field, the ψ's obey the commutation relation

$$[\psi_\alpha(x), \bar{\psi}_\beta(x')] = i d_{\alpha\beta}(\partial)\Delta(x-x')$$

so that d should transform like $\psi \otimes \psi$ giving n = 2f. Thus for $s>1$, the equation of Hurley violates the Umezawa-Visconti relation for a consistent quantisation of the field.

We have seen in Sec. 2, that Bhabha has given the most general solution of Eqn. (2.2) involving two parameters which may or may not satisfy eqn. (2.22). Using this form, an attempt has been made by Capri[15] to describe a particle of unique spin and unique mass by a linear relativistic equation of Bhabha's type. Instead of having an equation yielding multi-mass solutions, what is done to select the maximum spin is to introduce lower spins to start with and eventually make the corresponding submatrices in β_0^{2s-1} in the block diagonal form to be nilpotent. Here we shall describe such an attempt by Capri for a particle of spin 3/2 .

In this approach, the aim is to construct a set of four matrices which satisfy the following conditions:

a) Lorentz invariance: This is ensured by following Bhabha's prescription for writing the non-vanishing elements of $\beta\mu$ as outlined in section 2.

b) Unique rest mass: To ensure that the particle satisfies a Klein-Gordon equation for a unique rest mass m, β_μ must satisfy Harish-Chandra's condition, i.e. it must have eqn.(3.9) as the minimal equation. It immediately follows that, for $n>1$ in the Jordan form, β_μ becomes, in block form

$$\beta_\mu = \left[\begin{array}{c|c} N & 0 \\ \hline 0 & -1 \end{array}\right] \tag{5.9}$$

where N is a matrix nilpotent of order n-1. Therefore, for $n>2$, β_μ is non-diagonisable and hence cannot be hermitian. (This can also be seen from the fact that the singular matrices β have eigenvalues ± 1 and 0 and so if they are hermitian and hence diagonalisable they will always obey a minimal equation of degree ≤ 3 and Harish-Chandra condition will not be satisfied).

c) Hermitianisability: In order to derive eqn. (2.1) from a Lagrangian and to be able to define a current and energy momentum tensor, we require the existence of a hermitianising matrix such that

$$\eta^{-1}\beta_\mu^{+}\eta = \beta_\mu \tag{5.10}$$

d) Unique spin: In general, Bhabha's equation yields solutions which contain not only the maximum spin but in fact all lower spins. In order to pick up the solutions corresponding to a unique spin s, we demand, after Capri, the additional property

$$J^2 E = s(s+1)E \tag{5.11}$$

where E is a projection operator which projects on to the subspace spanned by the eigenvectors of β corresponding to the non zero eigenvalues. Explicitly, E is given by (in J^2 diag. basis)

$$E = \beta_0^{2s-1} \tag{5.12}$$

To take a concrete example, we consider spin 3/2. We start with the reducible representation of the Lorentz group

$$D = (1, \tfrac{1}{2}) \oplus (0, \tfrac{1}{2}) \oplus (\tfrac{1}{2}, 0) \oplus (\tfrac{1}{2}, 1) \tag{5.13}$$

To satisfy condition (a), we use Bhabha's relations (2.17) and (2.18) to write the non-vanishing elements of $A^{\ell m}$

$$A^{\ell m} = \begin{bmatrix} \multicolumn{2}{c}{\mathbf{0}} & c_1 u^\ell(1) u^m(\tfrac{1}{2}) & c_2 u^\ell(1) v^m(1) \\ & & c_3 v^\ell(\tfrac{1}{2}) u^m(\tfrac{1}{2}) & c_4 v^\ell(\tfrac{1}{2}) v^m(1) \\ -c_4 v^\ell(1) v^m(\tfrac{1}{2}) & -c_3 u^\ell(\tfrac{1}{2}) v^m(\tfrac{1}{2}) & \multicolumn{2}{c}{\mathbf{0}} \\ -c_2 v^\ell(1) u^m(1) & -c_1 u^\ell(\tfrac{1}{2}) u_m(1) & & \end{bmatrix} \tag{5.14}$$

The products in the above matrix are direct products. To enable us to impose condition (d), we go over to the Wild basis which yields for β_0 the form

$$\beta_0 = \left[\begin{array}{cccccc} & & & 0 & 2c_2 & 0 \\ & O & & \sqrt{3}c_1 & 0 & -c_2 \\ & & & c_3 & 0 & \sqrt{3}c_4 \\ 0 & \sqrt{3}c_4 & c_3 & & & \\ 2c_2 & 0 & 0 & & O & \\ 0 & -c_2 & \sqrt{3}c_1 & & & \end{array}\right]$$

(5.15)

whence

$$\beta_0^2 = \left[\begin{array}{cccccc} 4c_2^2 & 0 & 0 & & & \\ 0 & c_2^2+3c_1c_4 & \sqrt{3}c_1(c_3-c_2) & & O & \\ 0 & \sqrt{3}c_4(c_3-c_2) & c_3^2+3c_1c_4 & & & \\ & & & c_3^2+3c_1c_4 & 0 & \sqrt{3}c_1(c_3-c_2) \\ & O & & 0 & 4c_2^2 & 0 \\ & & & \sqrt{3}c_4(c_3-c_2) & 0 & c_2^2+3c_1c_4 \end{array}\right]$$

(5.16)

Further J^2 is given by

$$J^2 = \text{diagonal}\left\{\tfrac{3}{2}\left(\tfrac{3}{2}+1\right)I_4,\ \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)I_2,\ \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)I_2,\ \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)I_2,\ \tfrac{3}{2}\left(\tfrac{3}{2}+1\right)I_4,\ \tfrac{1}{2}\left(\tfrac{1}{2}+1\right)I_2\right\}, \tag{5.17}$$

where I_ℓ is the unit matrix of dimension ℓ. Now, for ψ to represent a solution corresponding to spin 3/2 only we demand that in the Wild basis

$$J^2\beta_0^2 = \tfrac{3}{2}\left(\tfrac{3}{2}+1\right)\beta_0^2 \tag{5.18}$$

Hence, from Eqn.(5.16), it follows that

$$4c_2^2 = 1$$

and the matrix

$$\begin{bmatrix} c_2^2 + 3c_1c_4 & \sqrt{3}c_1(c_3-c_2) \\ \sqrt{3}c_4(c_3-c_2) & c_3^2+3c_1c_4 \end{bmatrix} = 0 \tag{5.19}$$

This yields the solution

$$c_2 = \tfrac{1}{2}, \quad c_3 = \tfrac{1}{2}, \quad c_1 = \frac{i}{2\sqrt{3}}, \quad c_4 = \frac{i}{2\sqrt{3}}. \tag{5.20}$$

We note here that β_0^2 is a good projection operator, it is idempotent and hermitian. Eqn.(5.20) yields Fierz-Pauli equation as written by Gupta[34]. We notice that $GC_4 \neq 0$. The hermitianising matrix can be easily computed as

$$\eta = \begin{bmatrix} O & \begin{matrix} 0 & P_{\frac{1}{2},1} \\ -1 & 0 \end{matrix} \\ \begin{matrix} 0 & -1 \\ P_{1,\frac{1}{2}} & 0 \end{matrix} & O \end{bmatrix} \tag{5.21}$$

where P_{ij} is the permutation matrix.

6. <u>A Class of Equations for Spin 1/2 Particles</u>:

A natural question arises whether one cannot add higher spins and later make the blocks in β_0^{2f-1} (f = maximum spin) corresponding to these higher spins nilpotent to yield different equations for spin-half and whether these equations will be inequivalent. The procedure here is just the reverse of what we did in the earlier section to get an equation for a unique higher spin. An example for a spin 1/2 particle has actually been constructed by Capri[16]. Taking the specific example of the representation chosen in the preceding section for the field function

$$D = \left(1, \tfrac{1}{2}\right) \oplus \left(0, \tfrac{1}{2}\right) \oplus \left(\tfrac{1}{2}, 0\right) \oplus \left(\tfrac{1}{2}, 1\right) \tag{6.1}$$

we impose the following conditions on the matrix β_0 , namely, $\beta_0^4 = \beta_0^2$ as minimal equation to ensure that the particle has a unique mass and $J^2\beta_0^2 = \frac{1}{2}(\frac{1}{2}+1)\beta_0^2$ to make ψ represent a particle with spin 1/2. The latter condition is incorporated by demanding that the diagonal blocks in β_0^2 corresponding to spin 3/2 be nilpotent and that of spin 1/2 blocks idempotent and then finally fixing the constants that are otherwise arbitrary by demanding that β_0 satisfy the minimal equation $\beta_0^4 = \beta_0^2$. Capri has initiated such an attempt. In fact, in the Wild form of β_0^2 written out explicitly in the preceding section and by imposing the condition $J^2\beta_0^2 = \frac{1}{2}(\frac{1}{2}+1)\beta_0^2$ instead of eqn.(5.18) one gets the following solutions on the constants c_1, c_3, c_4 when $c_2=0$

1. $c_4=0,\ c_3^2=1,$ c_1 arbitrary

2. $c_1=0,\ c_3^2=1,$ c_2 arbitrary

3. $c_3=0,\ c_1c_4=\frac{1}{3},$

4. $c_3=0,\ c_4=0$ c_1 arbitrary (6.2)

Capri's matrices are obtained corresponding to the first set of values in eqn. (6.2). Substituting these in eqn.(5.14), β_μ are finally given by

$$\beta_0 = \begin{bmatrix} 0 & & \begin{matrix} 0 & 0 \\ \sqrt{2}c & 0 \\ -c & 0 \\ 0 & c \\ 0 & -\sqrt{2}c \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{matrix} & 0 \\ 0 & \begin{matrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \\ -c & 0 \\ 0 & -\sqrt{2}c \\ \sqrt{2}c & 0 \\ 0 & c \\ 0 & 0 \end{matrix} & & 0 \end{bmatrix} \qquad \beta_1 = \begin{bmatrix} 0 & & \begin{matrix} \sqrt{2}c & 0 \\ 0 & 0 \\ 0 & c \\ -c & 0 \\ 0 & 0 \\ 0 & -\sqrt{2}c \\ 0 & -1 \\ -1 & 0 \end{matrix} & 0 \\ 0 & \begin{matrix} 0 & 1 \\ 1 & 0 \\ -\sqrt{2}c & 0 \\ 0 & -c \\ 0 & 0 \\ 0 & 0 \\ c & 0 \\ 0 & \sqrt{2}c \end{matrix} & & 0 \end{bmatrix}$$

$$\beta_2 = \begin{bmatrix} 0 & & \begin{matrix} \sqrt{2}ic & 0 \\ 0 & 0 \\ 0 & ic \\ -ic & 0 \\ 0 & 0 \\ 0 & \sqrt{2}ic \\ 0 & -i \\ i & 0 \end{matrix} & 0 \\ 0 & \begin{matrix} 0 & i \\ -i & 0 \\ -\sqrt{2}ic & 0 \\ 0 & -ic \\ 0 & 0 \\ c & 0 \\ -ic & 0 \\ 0 & -\sqrt{2}ic \end{matrix} & & 0 \end{bmatrix} \qquad \beta_3 = \begin{bmatrix} 0 & & \begin{matrix} 0 & 0 \\ -\sqrt{2}c & 0 \\ -c & 0 \\ 0 & -c \\ 0 & -\sqrt{2}c \\ 0 & 0 \\ -1 & 0 \\ 0 & 1 \end{matrix} & 0 \\ 0 & \begin{matrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \\ c & 0 \\ 0 & \sqrt{2}c \\ \sqrt{2}c & 0 \\ 0 & c \\ 0 & 0 \end{matrix} & & 0 \end{bmatrix} \tag{6.3}$$

The last two solutions in eqn. (6.2) are discarded as they lead to lower order minimal equation for β_0. A closer examination[17] reveals that even the first two yield cubic minimal equation for β_0 and not of fourth degree as is required. Since the minimal equation $\beta_0^3 = \beta_0$ has distinct roots, β_0 is diagonalisable and has the diagonal form $\mathrm{diag}\{1, -1, 0\}$ They are consequently diagonalisable though not by a unitary transformation[18]. Thus, one clearly sees that by making the spin 3/2 block in β_0^2 nilpotent one does not get a solution for spin 1/2 which satisfies a minimal equation of fourth degree. Further, it can be seen that the hermitianising matrix given by Capri[16] is not correct. In fact, we shall show in a later section that a hermitianising matrix cannot exist in this case.

A justification for the equations of Capri's type has been offered by the group at Matscience[17] from algebraic considerations where it has been shown that if we project spin s from the representation which contains the maximum spin f, the Harish-Chandra condition is

$$\beta_0^{L+1} = \beta_0^{L-1}$$

where the Umezawa-Visconti condition is modified as

$$2s \leq L \leq 2f$$

The condition for unique spin[18] is

$$J^2 \beta_0^{L-1} = S(S+1)\beta_0^{L-1}$$

A similar observation has been made by Glass[35] who has explicitly demonstrated that by increasing the multiplicity of spin half components in the reducible representation one can make β_0 to satisfy a higher degree minimal equation $\beta_0^5 = \beta_0^3$ for spin 3/2 particles, a result which contradicts Umezawa-Visconti theorem. Essentially he relaxes the polynomial restriction on the K.G. divisor. Notice however we do not have any result so far to modify U.V. theorem to make higher spin equations to satisfy lower degree minimal equations except possibly the equations of Hurley whose implications for the U.V. relation have been discussed earlier in this section.

For the representation considered in the earlier section the K.G. divisor can be written as

$$\begin{aligned} d(\partial) = \chi + \alpha_\mu \partial_\mu + \alpha_{\mu\nu}\partial_\mu\partial_\nu \\ + \alpha_{\mu\nu\lambda}\partial_\mu\partial_\nu\partial_\lambda \end{aligned} \tag{6.4}$$

since the maximum spin contained is 3/2. If the field function must contain spin 3/2 the coefficient by U.V. theorem $\alpha_{\mu\nu\lambda} \neq 0$ If, on the other hand, we set $\alpha_{\mu\nu\lambda} = 0$ (the only other spin contained is 1/2), we find from eqn.(4.4) that with

$$\Lambda(\partial) = -(\beta_\mu \partial_\mu + \chi)$$

and using eqn.(4.4) for $\alpha_{\mu\nu\lambda}$ which we set equal to zero we have

$$\alpha_{\mu\nu\lambda} = \Sigma^{(p)} \beta_\mu \alpha_\nu \alpha_\lambda$$

$$= -\frac{1}{2\chi} \Sigma^{(p)} \beta_\mu (2g_{\nu\lambda} - \{\beta_\nu, \beta_\lambda\}), \qquad (6.5)$$

where $\Sigma^{(p)}$ denotes permutation of indices and expanding eqn. (6.5) one gets

$$\beta_\mu\beta_\nu\beta_\lambda + \beta_\mu\beta_\lambda\beta_\nu + \beta_\nu\beta_\mu\beta_\lambda + \beta_\nu\beta_\lambda\beta_\mu + \beta_\lambda\beta_\mu\beta_\nu + \beta_\lambda\beta_\nu\beta_\mu = -2g_{\mu\nu}\beta_\lambda + 2g_{\mu\lambda}\beta_\nu + 2g_{\nu\lambda}\beta_\mu \qquad (6.6)$$

Eqn.(6.6) admits three distinct algebras of the type of Duffin-Kemmer Patiau (DKP) obeyed by the β matrices. The first is the Duffin-Kemmer-Patiau algebra

$$\beta_\mu\beta_\nu\beta_\lambda + \beta_\lambda\beta_\nu\beta_\mu = g_{\mu\nu}\beta_\lambda + g_{\lambda\nu}\beta_\mu \qquad (6.7)$$

describing particles with spin 1 and 0. The second is the algebra obeyed by the matrices of Capri

$$\beta_\mu \{\beta_\nu, \beta_\lambda\} = 2 g_{\nu\lambda} \beta_\mu$$
$$\{\beta_\nu, \beta_\lambda\} \neq 2 g_{\nu\lambda} \qquad (6.8)$$

we find a third new algebra

$$\{\beta_\nu, \beta_\lambda\} \beta_\mu = 2 g_{\nu\lambda} \beta_\mu$$
$$\{\beta_\nu, \beta_\lambda\} \neq 2 g_{\nu\lambda} \qquad (6.9)$$

Both eqns. (6.8) and (6.9) describe particles with spin 1/2. This can be demonstrated by constructing the J^2 operator and showing that $J^2 \beta_0 = \frac{1}{2}(\frac{1}{2}+1)\beta_0$ where J^2 is the square of the generator of rotations in three dimensions. That these are the three algebras of the DKP type can be seen as follows:

since the highest non-vanishing term in $d(\partial)$, namely $a_{\mu\nu} \partial_\mu \partial_\nu$ transforms like a spin 2 object $[\partial_\mu \sim \frac{1}{2}, \frac{1}{2}]$ and since $d(\partial) \sim \psi \otimes \psi$ where ψ is the field function, ψ can either be a combination of spins 1 and 0, which yields the Duffin-Kemmer-Petiau algebra (6.7), or it can be a combination of spins 3/2 and 1/2 which yields the Capri and the new algebra. Of course, eqn. (6.6) itself describes an algebra if there are no subalgebras of the DKP type. However, the abstract algebra generated by algebraic quantities satisfying eqn. (6.6) alone is not finite and one would expect there to be an infinite number of inequivalent irreducible sets of matrices satisfying eqn. (6.6) as was shown by Bhabha[10] . Recently Hurley and Sudarshan[36] have analysed the infinite order algebra. Further, if eqns. (2.2) and (2.22) are insisted upon simultaneously, then eqn.(6.7) is the only algebra possible[10] .

It is quite clear that a representation of the new algebra given by eqn. (6.9) is furnished simply by the hermitian adjoints of the matrices of Capri. It is therefore very important that a hermitianising matrix does not exist if these two algebras are to be distinct. We shall show later that indeed the hermitianising matrix does not exist for the Capri matrices, making the two algebras inequivalent. If, on the other hand, one imposes the existence of a hermitianising matrix, the two algebras collapse into one and lead to a trivial extension of the Dirac matrices by adding zeros, making the β_μ anti-commute but $\beta_\mu^2 \neq 1$

The above situation is not altogether new since we are already used to a spin 0 particle being described by both the Klein-Gordon equation and a Duffin-Kemmer-Petiau equation. However, in the present case, both the equations are linear.

By choosing representations with higher spins and choosing different values of L, $2S \leq L \leq 2f$, and satisfying the condition $J^2 \beta_0^{L-1} = \frac{1}{2}\left(\frac{1}{2}+1\right) \beta_0^{L-1}$ one may get a hierarchy of equations describing a spin 1/2 particle. In these cases, β_0 is non-diagonalisable if $L > 2$. To what extent these equations will be physically different from the Dirac equation is yet unclear, although some preliminary calculations[18,24,25] on the evaluation of magnetic moment reveal that the content of the new equation is just the same as that of Dirac equation in a minimally coupled external electromagnetic field. (See later sections). In the next few sections we shall discuss elaborately the hierarchy of these algebras relating to a spin 1/2 particle.

7. The Representations of the New Algebras:

In this section, we discuss the realisations by matrices of the elements of the algebra satisfying the relations

$$\beta_\mu \{ \beta_\nu, \beta_\lambda \} = 2 g_{\nu\lambda} \beta_\mu \qquad \mu, \nu, \lambda = 0, 1, 2, 3 \tag{7.1}$$

and give methods for obtaining them from the Dirac matrices. The method is analogous to the method of Alladi Ramakrishnan in obtaining the higher dimensional L-matrices. We look for representations of the matrices β_μ which do not satisfy the Dirac Clifford algebra (which trivially satisfies eq. (7.1). There are 65 independent elements of the above algebra. (1, $\beta_\mu \beta_0^{\rho_0} \beta_1^{\rho_1} \beta_2^{\rho_2} \beta_3^{\rho_3}$ where $\mu = 0, 1, 2, 3$ and $\rho_\mu = 0, 1$.) and it is easily see that any set of four square matrices of the form

$$\beta_\mu = \begin{bmatrix} 0 & X_\mu \\ 0 & \gamma_\mu \end{bmatrix} \tag{7.2}$$

where γ_μ are the Dirac matrices and X_μ are arbitrary matrices with four columns satisfies eq.(7.1). This is at once obvious from the fact that

$$\beta_\mu \beta_\nu \beta_\lambda = \begin{bmatrix} 0 & X_\mu \gamma_\nu \gamma_\lambda \\ 0 & \gamma_\mu \gamma_\nu \gamma_\lambda \end{bmatrix}$$

β_μ do not anticommute unless $X = 0$ or, in general, $X_\mu \gamma_\nu + X_\nu \gamma_\mu = 0$. If further symmetry properties are required (like, for instance, parity invariance in the choice of the reducible representation) β_μ may be of the form

$$\beta_0 = \begin{bmatrix} 0 & & X_0 & 0 \\ & & 1 & 0 \\ 0 & 1 & & \\ 0 & Y_0 & & 0 \end{bmatrix} \tag{7.3a}$$

and

$$\beta_i = \begin{bmatrix} 0 & & X_i & 0 \\ & & \sigma_i & 0 \\ 0 & -\sigma_i & & \\ 0 & Y_i & & 0 \end{bmatrix} \tag{7.3b}$$

where X and Y are arbitrary matrices of two columns, 1 is a unit matrix in two dimensions and σ_i are the Pauli matrices. It may at once by inferred that the matrices of Capri are of the above form. The dimensions of X and Y depend on the choice of the reducible representation. The matrices X and Y are fixed by demanding that β_μ satisfy eq.(2.2) for Lorentz invariance. To do this, we choose β_0 in the Wild basis directly using the above prescription

$$\beta_0 = \begin{bmatrix} \begin{matrix} & \\ & 0 \\ & \end{matrix} & \begin{matrix} 0 & 0 & 0 \\ c & 0 & 0 \\ 1 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & c \\ 0 & 0 & 0 \end{matrix} & \begin{matrix} & \\ & 0 \\ & \end{matrix} \end{bmatrix} \tag{7.4}$$

where c is an arbitrary number.

This form of β_0 automatically satisfies eq. (5.11) with $E = \beta_0^{L-1}$ where L = 2, to quarantee the correct spin. All the matrices β_μ can now be written down in the original basis using the spinor equations (2.17) and (2.18) as in section 6, the non-vanishing elements being identified from eq.(7.4).

From the form of β_μ displayed, it is clear that a typical characteristic of the matrices is that there are elements only above and below the Dirac matrices. Thus we infer that the Capri matrices, which are of the above form, do not satisfy $\beta_0^4 = \beta_0^2$ as minimal equstion as required by him but satisfy the algebra given by eq. (7.1) and hence $\beta_0^3 = \beta_0$

We now proceed to consider the other algebra defined by

$$\{\beta_\nu, \beta_\lambda\}\beta_\mu = 2g_{\nu\lambda}\beta_\mu \tag{7.5}$$

By taking the hermitian adjoint of eq. (7.1) it is at once obvious that if β_μ satisfy eq.(7.1) $\beta_\mu^\dagger$ satisfy eq.(7.5). Thus the matrices which satisfy eq.(7.5) have the form

$$\beta_\mu = \begin{bmatrix} 0 & 0 \\ X_\mu & Y_\mu \end{bmatrix} \tag{7.6}$$

where X_μ are now arbitrary matrices with four rows. We note that X_μ is now on the horizontal "Dirac tail", and β_μ do not anticommute unless $X = 0$ or, in general, unless

$$Y_\mu X_\nu + Y_\nu X_\mu = 0$$

Here too, the number of elements in the algebra is 65 and can be listed as $(1, \beta_0^{\rho_0}\beta_1^{\rho_1}\beta_2^{\rho_2}\beta_3^{\rho_3}\beta_\mu)$ where $\mu = 0, 1, 2, 3.$ and $\rho_\mu = 0, 1$. Again, if some symmetry properties are required as earlier, we may take

$$\beta_0 = \begin{bmatrix} 0 & & 0 & 0 \\ & & 1 & Y_0 \\ X_0 & 1 & & \\ 0 & 0 & & 0 \end{bmatrix} \tag{7.7}$$

$$\beta_i = \begin{bmatrix} 0 & \begin{matrix} 0 & 0 \\ \sigma_i & Y_i \end{matrix} \\ \begin{matrix} X_i & -\sigma_i \\ 0 & 0 \end{matrix} & 0 \end{bmatrix} \tag{7.8}$$

where X and Y are arbitrary matrices of two rows, which are again chosen such that β_μ satisfy eq. (2.2). Using the above prescription, we choose β_0 in the Wild form as

$$\beta_0 = \begin{bmatrix} 0 & \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & 0 \end{bmatrix} \tag{7.9}$$

all the β_μ in the original basis being again determined by eqs. (2.17) and (2.18). All these matrices have elements on the horizontal "Dirac tail" and zeros elsewhere and have as their minimal equation $\beta_0^3 = \beta_0$.

The two algebras given by eqs.(7.1) and (7.5) are, in general, inequivalent. An immediate consequence of this, as we shall show presently, is that, for the Capri matrices (as also for their hermitian adjoints) a permitiaising matrix does not exist. For, if β_μ satisfy either of these algebras $\beta_\mu' = S^{-1}\beta_\mu S$ also satisfy the same algebra as can be verified by direct substitution. Thus, if β_μ satisfy eq. (7.1)

$$S^{-1}\beta_\mu S\,(S^{-1}\beta_\nu S S^{-1}\beta_\lambda S + S^{-1}\beta_\lambda S S^{-1}\beta_\nu S) = 2g_{\nu\lambda}S^{-1}\beta_\mu S$$

Therefore

$$\beta'_\mu\{\beta'_\nu, \beta'_\lambda\} = 2g_{\nu\lambda}\beta'_\mu$$

A similar statement holds for eq.(7.3). If we now demand the existence of a hermitianising matrix η, such that $\eta^{-1}\beta_\mu\eta$ the two algebras (7.1) and (7.5) become equivalent and β_μ and β^+_μ will both have to satisfy both eqs. (7.1) and (7.5). Thus, using eq. (7.1)

$$\beta_0^2\{\beta_\mu, \beta_\nu\} = 2g_{\mu\nu}\beta_0^2$$

and then using eq. (7.5) we get

$$\{\beta_\mu, \beta_\nu\} = 2g_{\mu\nu}\beta_0^2 \tag{7.10}$$

with $\beta_0^3 = \beta_0$

Thus the matrices anticommute but their squares are not unity.

The matrices of Capri satisfy (7.1) but they do not satisfy eq. (7.5) nor do they satisfy eq. (7.10). Hence, from the above argument, a hermitianising matrix cannot exist for these matrices. In fact, the matrix η given by Capri is incorrect as can be verified directly or inferred from the above arguments.

A parallel argument shows that the hermitian adjoints of the above matrices too are not hermitianisable. We shall show presently that the moment we insist on a hermitianising matrix, we obtain a trivial extension of the Dirac matrices.

The two algebras are, in general, inequivalent. As shown above, if we demand the existence of a hermitianising matrix, the two algebras will become equivalent and β_μ and β_μ^+ will have to satisfy both eqs. (7.1) and (7.5) which in turn implies that

$$\{\beta_\nu, \beta_\lambda\} = 2 g_{\nu\lambda} \beta_0^2 \qquad (7.10)$$

In this case, the dimension of the algebra reduces to 17

(β_0^2, $\beta_0^{p_0} \beta_1^{p_1} \beta_2^{p_2} \beta_3^{p_3}$ where $p_\mu = 0, 1.$). Further, the anticommutators commute with all the elements of the algebra

$$[\beta_\mu, \{\beta_\nu, \beta_\lambda\}_+]_- = 0 \qquad (7.11)$$

for all μ, ν, λ

From eqs. (7.10) and (7.11) and Pauli's Fundamental Theorem, it follows that β_μ are either trivial extensions of the Dirac matrices by adding zeros, i.e.,

$$\beta_\mu = \begin{bmatrix} 0 & 0 \\ 0 & \gamma_\mu \end{bmatrix} \qquad (7.12)$$

or all the β_μ can be reduced to such a form by a single similarity transformation.

As an example of such a set of matrices where the equivalence is not immediately obvious, consider

$$T_\mu = \beta_0^2 \beta_\mu \tag{7.13}$$

where β_μ satisfy eq. (7.1) and may be taken as the Capri matrices. Then T_μ anticommute and

$$\{T_\mu, T_\nu\} = 2 g_{\mu\nu} T_0^2 \tag{7.14}$$

Further T_μ have the form

$$T_0 = \begin{array}{c} \left[\begin{array}{cccc} 0 & & X_0 & 0 \\ & & -1 & 0 \\ 0 & -1 & & \\ 0 & Y_0 & & 0 \end{array}\right] \end{array} \begin{array}{c} (6) \\ (2) \\ (2) \\ (6) \end{array}$$

$$T_i = \left[\begin{array}{cccc} 0 & & X_0\sigma_i & 0 \\ & & -\sigma_i & 0 \\ 0 & \sigma_i & & \\ 0 & -Y_0\sigma_i & & 0 \end{array}\right] \tag{7.15a}$$

(7.15b)

(The figures on the right indicate the number of rows in the various blocks).

It is not at once obvious that they can all be brought by the same similarity transformation to the form γ_μ augmented by zeros. However, consider the transformation

$$T_\mu' = U^{-1} T_\mu U \tag{7.16}$$

where

$$U = 1 - \begin{bmatrix} 0 & X_0 & & \\ 0 & 0 & & 0 \\ & & 0 & 0 \\ & 0 & Y_0 & 0 \end{bmatrix} \tag{7.17a}$$

and

$$U^{-1} = 1 + \begin{bmatrix} 0 & X_0 & & \\ 0 & 0 & & 0 \\ & & 0 & 0 \\ & 0 & Y_0 & 0 \end{bmatrix} \tag{7.17b}$$

Then

$$T_0' = \begin{bmatrix} & & 0 & 0 \\ & 0 & -1 & 0 \\ 0 & -1 & & \\ 0 & 0 & & 0 \end{bmatrix}$$

and

$$T_i' = \begin{bmatrix} & & 0 & 0 \\ & 0 & -\sigma_i & 0 \\ 0 & -\sigma_i & & \\ 0 & 0 & & 0 \end{bmatrix}.$$

which is a trivial extension of the Dirac matrices by adding zeros. However, a hermitianising matrix exists in this case, for

$$\Gamma_\mu^{+} = \eta^{-1}\Gamma_\mu \eta$$

where

$$\eta = U\gamma_5 U^{+}$$

and

$$\gamma_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

8. **Hierarchies of Algebras Relating to Spin 1/2:**

The discussion of the previous section indicates that by adding more of other spins and then requiring that

$J^2 \beta_0^{L-1} = S(S+1)\beta_0^{L-1}$ one can have a hierarchy of algebras of the type

$$\beta_{\mu_1 \cdots \mu_K}^{(K+2)} \{\beta_\nu^{(K+2)}, \beta_\lambda^{(K+2)}\} = 2g_{\nu\lambda}\beta_{\mu_1 \cdots \mu_K}^{(K+2)} \tag{8.1a}$$

and

$$\{\beta_\nu^{(K+2)}, \beta_\lambda^{(K+2)}\}\beta_{\mu_1 \cdots \mu_K}^{(K+2)} = 2g_{\nu\lambda}\beta_{\mu_1 \cdots \mu_K}^{(K+2)} \tag{8.1b}$$

with the notation

$$\beta^{(K+2)}_{\mu_1\mu_2\mu_3} = \beta^{(K+2)}_{\mu_1}\beta^{(K+2)}_{\mu_2}\beta^{(K+2)}_{\mu_3}$$

the superscript K refers to the minimal equation satisfied by the matrices

$$\beta^{K+2} = \beta^{K}$$

The order of either algebra is $4^K 2^4 + \frac{4^K-1}{4-1}$ and the realisation of the Kth algebra can be constructed from the (K-1)th algebra as will be demonstrated below. The method is quite analogous to the one employed by Alladi Ramakrishnan[2] in the discussion of Clifford algebra.

As an example, we consider the next higher member in the hierarchy after the one considered in the previous section. Putting K = 2 in eqn.(3.1),

$$\beta^{(4)}_{\mu_1}\beta^{(4)}_{\mu_2}\{\beta^{(4)}_{\nu}, \beta^{(4)}_{\lambda}\} = 2g_{\nu\lambda}\beta^{(4)}_{\mu_1}\beta^{(4)}_{\mu_2} \tag{8.2a}$$

$$\{\beta^{(4)}_{\nu}, \beta^{(4)}_{\lambda}\}\beta^{(4)}_{\mu_1}\beta^{(4)}_{\mu_2} = 2g_{\nu\lambda}\beta^{(4)}_{\mu_1}\beta^{(4)}_{\mu_2} \tag{8.2b}$$

The order of the algebra in either case is 261. We note that eqn.(8.2a) contains the Dirac algebra and eqn.(7.1) as subalgebras, while eqn.(8.2b) contains the Dirac algebra and eq.(7.5) as subalgebras. A typical realisation of the algebra (8.2a) is the set of matrices

$$\beta_\mu^{(4)} = \begin{bmatrix} 0 & X_\mu'^{(4)} \\ 0 & \beta_\mu^{(4)} \end{bmatrix} = \begin{bmatrix} 0 & X_\mu^{(4)} & 0 \\ 0 & 0 & X_\mu^{(3)} \\ 0 & 0 & \gamma_\mu \end{bmatrix} \tag{8.3}$$

where $X_\mu^{(3)}$ is an arbitrary matrix of four columns and $X_\mu^{(4)}$ is an arbitrary matrix. The hermitian adjoint of this satisfies eq.(8.2b). We note also that a typical characteristic of the above matrices is that there are entries on the vertical (or horizontal) 'Dirac tails' and on the adjacent column (or row) respectively. Again we notice that the two algebras (8.2a) and (8.2b) are inequivalent so long as a hermitianising matrix does not exist, that is, so long as $X^{(3)}$ and $X^{(4)}$ are not zero. Following the treatment in the last section, to satisfy further symmetry properties we choose

$$\beta_0 = \left[\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & \multicolumn{2}{c}{\bigcirc} & & 0 & x_0^{(4)} & 0 & 0 \\ & & & & x_0^{(3)} & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & y_0^{(3)} & & \multicolumn{2}{c}{\bigcirc} & \\ 0 & 0 & y_0^{(4)} & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right]$$

$$\beta_i = \left[\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & \multicolumn{2}{c}{\bigcirc} & & 0 & x_i^{(4)} & 0 & 0 \\ & & & & x_i^{(3)} & 0 & 0 & 0 \\ & & & & \sigma_i & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\sigma_i & & & & \\ 0 & 0 & 0 & y_i^{(3)} & & \multicolumn{2}{c}{\bigcirc} & \\ 0 & 0 & y_i^{(4)} & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right] \qquad (8.4)$$

For getting the actual matrices, we start with the reducible representation

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

In the Wild basis we conveniently write β_0 as

$$\beta_0^{(4)} = \left[\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & 0 & & & 0 & c_2 & 0 & 0 \\ & & & & c_1 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & c_1 & & & 0 & \\ 0 & 0 & c_2 & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right] \tag{8.5}$$

where c_1 and c_2 are arbitrary constants and finally calculate the spinor form using Bhabha's prescription. These matrices have as their minimal equation $\beta_\mu^4 = \pm\beta_\mu^2$ give the correct spin 1/2, and satisfy the algebra (8.2a). The hermitian adjoints will satisfy eqn.(8.2b) but not eqn.(8.2a). Hence the algebras are distinct and a hermitianising matrix does not exist.

If we impose the existence of a hermitianising matrix on matrices satisfying eqns.(8.2a) and (8.2b), then the two algebras collapse into one giving

$$\{\beta_\nu, \beta_\lambda\} = 2g_{\nu\lambda}\beta_0^2 \tag{8.6}$$

which, with our minimal condition $\beta_0^4 = \beta_0^2$ defines a new algebra. An explicit example of these matrices can be obtained by taking

$$D = (1, \tfrac{1}{2}) \oplus 3(0, \tfrac{1}{2}) \oplus 3(\tfrac{1}{2}, 0) \oplus (\tfrac{1}{2}, 1) \tag{8.7}$$

In the Wild basis

$$\beta_0 = \begin{array}{c} \left[\begin{array}{ccccc|ccccc} & & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 & 0 & 0 \\ & & O & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & & & & & \\ 0 & 0 & 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & & & O & & \\ 0 & 0 & 0 & 0 & 0 & & & & & \\ 0 & 0 & 0 & 0 & 0 & & & & & \end{array}\right] \end{array} \begin{array}{c} (4) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (4) \end{array} \tag{8.8}$$

In the spinor form, the matrices are given by

$$\beta_0 = \left[\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & O & & & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 \\ & & & & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & 0 & & & O & \\ 0 & 0 & 0 & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right] \begin{array}{c} (6) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (2) \\ (6) \end{array} \tag{8.9a}$$

$$\beta_i = \left[\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & 0 & & & 0 & e\sigma_i & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 \\ & & & & \sigma_i & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\sigma_i & & & & \\ 0 & 0 & 0 & 0 & & 0 & & \\ 0 & 0 & -e\sigma_i & 0 & & & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right] \qquad (8.9b)$$

with the hermitianising matrix given by

$$\eta = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (8.10)$$

9. <u>Solutions of the New Equation:</u>

Let us study the solutions of the new spin half equation where the matrices satisfy the algebra eqn.(7.1)

$$(-i\beta^{\mu}\partial_{\mu} + m)\psi = 0 \tag{9.1}$$

with

$$\beta^{\mu}\{\beta^{\nu}, \beta^{\lambda}\} = 2g^{\nu\lambda}\beta^{\mu} \tag{9.2}$$

Multiplying eqn.(9.1) by β_0 and using $\beta_0^3 = \beta_0$ one gets

$$\beta_0[-i(\beta_0\partial_0 - \beta_K\partial_K) + m]\psi_1 = 0$$
$$K = 1, 2, 3 \tag{9.3}$$

where

$$\psi_1 = \beta_0^2\psi \tag{9.4}$$

Defining

$$\Gamma_K = \beta_0^2\beta_K\beta_0^2 \; ; K = 1, 2, 3.$$
$$\Gamma_0 = \beta_0 \tag{9.5}$$

We can rewrite eqn.(9.3) as

$$
\begin{aligned}
i\partial_0\psi_1 &= \Gamma_0(-i\Gamma^k\partial_k + m)\psi_1 \\
&= \Gamma_0(-i\underline{\Gamma}\cdot\underline{\partial} + m)\psi_1
\end{aligned}
\tag{9.6}
$$

We notice that (Capri[25]) ψ_1 satisfies the usual Dirac equation with the Hamiltonian given by

$$H = \Gamma_0(-i\underline{\Gamma}\cdot\underline{\partial} + m) \tag{9.7}$$

The Γ's obey the algebra

$$\{\Gamma_\mu, \Gamma_\nu\} = 2g_{\mu\nu}\Gamma_0^2 \tag{9.8}$$

The unit element is split into two primitive idempotents and $1-\beta_0^2$ as

$$1 = \beta_0^2 + (1-\beta_0^2) \tag{9.9}$$

Any operator A is therefore split as

$$
\begin{aligned}
A = {} & \beta_0^2 A \beta_0^2 + (1-\beta_0^2)A(1-\beta_0^2) \\
& + \beta_0^2 A(1-\beta_0^2) + (1-\beta_0^2)A\beta_0^2
\end{aligned}
\tag{9.10}
$$

Again by multiplying eqn. (9.1) by $(1-\beta_0^2)$ we get

$$\psi_2 = (1-\beta_0^2)\psi$$
$$= -\frac{1}{m}(1-\beta_0^2)(i\not\nabla)\psi_1$$
$$\not\nabla = \beta_k \partial_k \qquad (9.11)$$

It is interesting to note that the interdependence relation (9.11) does not involve time derivative and is reminiscent of the usual higher spin theories.

10. Minimal Electromagnetic Coupling:

We shall show in this section that the magnetic moment of a particle satisfying the new equation is identical to that of the Dirac particle. We consider a particle of charge e in a minimally coupled electromagnetic field. We write the wave equation as

$$(-i\not{D} + m)\psi = 0 \qquad (10.1)$$

where

$$\not{D} = D^\mu \beta_\mu = D_0\beta_0 - D_k\beta_k = D_0\beta_0 - \underline{\not{D}} \qquad (10.2)$$

and

$$D^\mu = \partial^\mu + ieA^\mu \qquad (10.3)$$

Following Capri[25] we premultiply the wave equation by $\not{D}\beta_0$

to arrive at

$$i D_0 \Psi_1 = \frac{-\beta_0^2}{m}(\underline{D}^2 - \not{D}^2)\beta_0\Psi_1 + \beta_0^2(-i\not{D} + m)\beta_0\Psi_1 + \frac{ie}{4m}\Gamma_0[\Gamma_k, \Gamma_\ell]F_{k\ell}\Psi_1 \tag{10.4}$$

where

$$\Psi_1 = \beta_0^2\psi \tag{10.5}$$

$$\Gamma_\mu = \beta_0^2\beta_\mu\beta_0^2 \tag{10.6}$$

and

$$[D_\mu, D_\nu] = ieF_{\mu\nu} \tag{10.7}$$

Using

$$\beta_0^2\not{D}^2 = \beta_0^2\left(-\underline{D}^2 + \frac{ie}{4}[\Gamma_k, \Gamma_\ell]F_{k\ell}\right) \tag{10.8}$$

We finally arrive at

$$i\partial_t\Psi_1 = eA_0\Psi_1 + \Gamma_0(-i\Gamma_k D^k + m)\Psi_1 \tag{10.9}$$

or

$$i\partial_t\Psi_1 = eA_0\Psi_1 + \Gamma_0(-i\underline{\Gamma}\cdot\underline{D} + m)\Psi_1 \tag{10.10}$$

It is clear that, at the stage of eq. (10.4) an error in the numerical factor and sign in the last term led him to on additional term in eq. (10.10). However, eq. (10.10) shows that, confined to ψ_1 , the magnetic moment of the particle is just the same as for the Dirac particle.

The above calculation can actually be made much simpler using the algebra given by eq. (7.1). We merely premultiply eq. (4.1) by β_0 to get

$$\beta_0(-i\not{D}+m)\psi=0 \tag{10.11}$$

or

$$(-i\beta_0 D^\mu \beta_\mu \beta_0^2 + m\beta_0\beta_0^2)\psi=0 \tag{10.12}$$

Therefore,

$$\beta_0(-iD^\mu\beta_\mu+m)\psi_1=0 \tag{10.13}$$

Where

$$\beta_0^2\psi=\psi_1$$

Hence

$$i\partial_0\psi_1 = eA_0\psi_1+\beta_0(-iD^\kappa\beta_\kappa+m)\psi_1 \tag{10.14}$$

or

$$i\partial_t\psi_1 = eA_0\psi_1+\beta_0(-i\underline{T}\cdot\underline{D}+m)\psi_1 \tag{10.15}$$

which is identical with eq. (4.10). Hence the magnetic moment of the particle is just the same as that of the Dirac particle of mass m .

However, the interdependence relation eqn.(9.11) would be changed and it is not clear at the moment whether the problem of inconsistency pointed out in Ref.14 will exist in the case of the new spin-half field, when one tries to quantize the theory.

An independent calculation by Menon[24] shows that for the equation obtained by using the other algebra (7.5), the same value of the magnetic moment is obtained.

11. Conclusions:

It has been shown that Bhabha equations can be written for unique mass and unique spin particles. The unique mass is obtained by imposing Harish-Chandra condition and the unique spin can be projected by suitable projection operators. We have also shown that if one replaces the diagonalisability condition of β_0 by weaker block form we can get a hierarchy of equations for a spin-half particle inequivalent to the Dirac equation, arising from a hierarchy of algebras. We give representations of these algebras. We study the properties of the new-spin-half equation in detail. We show that the hermitianising matrix does not exist in some cases. This we do not feel, is so sacrosanct as the Fierz-Pauli equation studied by Fronsdal[20] has a similar property although one can construct conserved charge/ current densities. Minimally coupled radiation field does not distinguish the new equation from that of Dirac's as it leads to the same magnetic moment. It is hoped that other kinds of interactions may possibly distinguish these inequivalent descriptions of a spin-half particle.

<u>References</u>:

1. P.A.M.Dirac, Proc.Roy.Soc.,(London) <u>A117</u>, 610 (1928).
2. Alladi Ramakrishnan, L-Matrix Theory or The Grammar of Dirac Matrices, Tata-McGraw Hill, Bombay (1972).
3. P.A.M.Dirac, Proc. Roy. Soc. (London) <u>155A</u>, 447 (1935).
4. M.Fierz, Helv. Phys. Acta, <u>12</u>, 3 (1939).
   M.Fierz and W.Pauli, Proc.Roy.Soc.,(London) <u>173A</u>, 211.
5. W.Rarita and J.Schwinger, Phys.Rev. <u>60</u>, 61 (1941).
6. H.J.Bhabha, Rev. Mod. Phys. 17, 200 (1945); Proc.Ind.Acad. Sci. <u>21</u>, 241, 1945; Curr.Sci. <u>14</u>, 39 (1945).
7. E.M.Corson, Introduction to Tensors, Spinors and Relativistic Wave Equations, Blackie, London (1953).
8. Harish-Chandra, Proc. Roy. Soc. (London) <u>186A</u>,502 (1946).
9. Harish-Chandra, Phys. Rev. <u>71</u>, 793 (1947).
10. H.J.Bhabha, Rev. Mod. Phys., <u>21</u>, 451 (1949).
11. S.Weinberg, Phys. Rev. <u>133B</u>, 1318 (1964).
12. D.L.Weaver, C.L.Hammer and R.H.Good Jr., Phys.Rev. <u>135B</u> 241 (1964).
13. D.L. Pursey, Ann. Phys. <u>32</u>, 157 (1965).
14. A.S.Wightman, Proc. 5th Coral Gables Conference Eds.A.Perlmutter, C.A.Hurst: Freeman and Co.(San Francisco), 1968, pages 291-314.
15. A.Z.Capri, Phys. Rev. <u>178</u>, 2427 (1969).
16. A.Z.Capri, Phys. Rev. <u>187</u>, 1811 (1969).

17. P.S.Chandrasekaran, Nalini B.Menon and T.S.Santhanam, Progr. Theor. Phys. 47, 671 (1972).

18. A.R.Tekumalla and T.S.Santhanam, Progr.Theor.Phys. 50,992, (1973).

19. T.S.Santhanam and A.R.Tekumalla, Fort der Phys., To appear (1974).

20. C.Fronsdal, Nuovo Cim. Suppl. 9, 416 (1958).

21. H.Umezawa and A.Visconti, Nucl.Phys. 1, 348 (1956).

22. M.M.Bakri, Nuovo Cim. 51A, 864 (1967); J.Math.Phys. 10 298 (1969).

23. W.J.Hurley, Phys. Rev. D4, 3605 (1971)

24. Nalini B.Menon, Ph.D. Thesis, Madras University (1972); Progr. Theor. Phys. 50, 987 (1973).

25. A.Z.Capri, Progr. Theor. Phys. 48, 1364 (1972).

26. Z.Tokuoka, Progr. Theor. Phys. 37, 603 (1967).

27. N.D.San Gupta, Nucl.Phys. B4, 147 (1968).

28. T.S.Santhanam and P.S.Chandrasekaran, Progr.Theor.Phys. 41, 264 (1969).

29. C.Brulin and S.Hjalmers, J.Math.Phys. 5, 1368 (1964) and the references quoted.

30. T.S.Santhanam and A.R.Tekumalla, Communicated to Phys.Rev.,

31. E.Wild, Proc. Roy. Soc. (London) 191A, 253 (1947).

32. B.S.Madhava Rao, Proc.Ind.Acad.Sci. 15, 139 (1942) J.Mys.Univ. 3, 59 (1942).

33. B.S.Madhava Rao, V.R.Thiruvenkatachar and K.Venkatachaliengar, Proc.Roy.Soc. (London) 187A, 385 (1946).

34. S.N.Gupta, Phys.Rev. 95, 1334 (1954).

35. A.S.Glass, Comm. Math.Phys. 23, 176 (1971).

36. W.J.Hurley and E.C.G.Sudarshan, Univ.of Texas, Austin Preprint CPT 188, Aug. 1973.

C H A P T E R . III

RELATIVISTIC WAVE EQUATIONS FOR MASSLESS PARTICLES*+

1. Introduction:

Relativistic wave equations describing the neutrino have attracted the attention of many people in the past. Shortly after the introduction of the relativistic theory of the electron, a two-component theory was proposed by Weyl[1] for the massless particle. The Weyl equation

$$\underline{\sigma} \cdot \underline{p}\, \varphi = i \frac{\partial \varphi}{\partial t} \tag{1.1}$$

where σ_i are the three Pauli matrices and φ is a two-component spinor, is known to be not invariant under the operations of parity reversal (P) and charge conjugation (C) but invariant under time reversal (T) and the combined operation CP. Originally, this form of the theory was discarded on the grounds that it did not give a parity invariant equation. However, under the impetus of the discovery of parity non-conservation in weak interactions, the two-component theory was revived[2] because it was then evident that the former objection was not cogent.

A second way of linearising the Klein-Gordon equation to describe massless spin-half particles is to write a wave equation of the Dirac type in which the mass parameter is set equal to

---

* T.S.Santhanam and A.R.Tekumalla, Lett.Nuov.Cim. 3, 122 (1972)
\+ A.R.Tekumalla and T.S.Santhanam, Lett.Nuov.Cim. 6, 99 (1973)

zero. Thus we write

$$(\underline{\alpha}\cdot\underline{p})\psi = i\frac{\partial\psi}{\partial t} \tag{1.2}$$

where α_i are three 4 x 4 anticommuting matrices and is a 4-component spinor. The theory developed for high energy massive particles can be extended to this case.

Yet another method, and one inequivalent to the above, is to use singular idempotent linear combinations of the Dirac matrices without explicitly putting the mass parameter equal to zero. This has been known in the literature long back[3]. Recently, attention has been drawn to this fact by Tokuoka[4], Sen Gupta[5] and Santhanam and Chandrasekharan[6]. In this case, the existence of five mutually anti-commuting matrices in four dimensions is suitably exploited. Normally, in writing a Dirac type equation for massless spin ½ particles, one starts with the wave equation

$$(\underline{\alpha}\cdot\underline{p} + m\beta)\psi = i\frac{\partial\psi}{\partial t} \tag{1.3}$$

with $m = 0$. However, one can replace m here by M so that

$$H\psi \equiv (\underline{\alpha}\cdot\underline{p} + M\beta)\psi = i\frac{\partial\psi}{\partial t} \tag{1.4}$$

where

$$M = m_1\gamma_5 + m_2$$

This substitution does not jeopardise most of the relevant invariance properties and, in general, eq.(1.4) can be transformed to eq. (1.3) by a change of representation by a similarity transformation. This transformation is accomplished by

$$T = e^{\frac{1}{2}\gamma_5 \varphi} \tag{1.5}$$

where $\tan\varphi = \frac{m_1}{m_2}$,

and the transformed Hamiltonian H' is given by

$$H' = T^{-1} H T = \underline{\alpha} \cdot \underline{p} + m' \beta \tag{1.6}$$

where

$$m' = (m_2^2 - m_1^2)^{1/2}$$

Therefore eq. (1.4) can be written in the usual Dirac form

$$(\underline{\alpha} \cdot \underline{p} + \beta m') \psi' = i \frac{\partial \psi'}{\partial t} \tag{1.7}$$

where

$$\psi' = T^{-1} \psi$$

However, this transformation is not possible if $m_1^2 = m_2^2$ ie. if M is singular. In this case we get

$$[\underline{\alpha} \cdot \underline{p} + m(1 \pm \gamma_5)\beta] \psi \tag{1.8}$$

It was shown by Bhabha[3] that such equations are plausible. He pointed out that M should be of the form M = $m\beta$ where

β is idempotent, ie. $\beta^2 = \beta$, which immediately implies that β is singular unless it is unity. The latter situation corresponds to the usual Dirac equation while the former yields the above equation. It had not relevance them and thus failed to attract any attention. But later developments in both theory and experiment, suggest it is worthwhile to investigate the invariance properties of this equation. One should remember here that m does not refer to the mass of the particle. It has been interpreted by Sen Gupta as a measure of its chirality and an intrinsic property of the particle, and suggests the possibility of particles with zero rest mass but not completely polarised along their direction of motion. This property is described by the new parameter m. Another feature of this equation is that the Hamiltonian is not hermitian in the ordinary sense, which suggests an indefinite quantisation of the field[4].

Santhanam and Chandrasekharan[6] have shown in 1969 that in the case of massive spin-half particles, even if we describe it by a wave function in 2^n dimensions through an equation containing $2n + 1$ parameters, it can still be brought to the standard Dirac form involving only four anticommuting matrices by a suitable transformation. In the case of massless particles, the equation can be reduced to the form of eq. (1.8) involving five anticommuting matrices, two of them occuring in a singular idempotent combination. It is known that there are $2n + 1$ mutually anticommuting matrices of dimension $2^n \times 2^n$ which form a complete set of elements satisfying the Clifford algebra of dimension 2^n. These matrices are easily constructed using the elegant method developed by Ramakrishnan[7]. The most general wave equation of a spinor particle which makes use of all these $2n + 1$ mutually anticommuting matrices can be written as

$$H\psi \equiv \left[\underline{\alpha}\cdot\underline{p} + \alpha_0(m_4\alpha_4 + \cdots + m_{2n}\alpha_{2n} + m_0)\right]\psi = i\frac{\partial\psi}{\partial t} \tag{1.9}$$

where $\underline{\alpha}\cdot\underline{p} = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3$

and

$$\{\alpha_r, \alpha_s\} = 2\delta_{rs}\,;\quad r, s = 0, 1, \ldots, 2n$$

We have

$$H^2 = p^2 + m_0^2 - \sum_{s=4}^{2n} m_s^2 \tag{1.10}$$

and the transformation by T given by

$$T = \exp \frac{\alpha_0(m_4\alpha_4 + \cdots + m_{2n}\alpha_{2n})}{\left(\sum_{s=4}^{2n} m_s^2\right)^{\frac{1}{2}}} \tanh \frac{\left(\sum_{s=4}^{2n} m_s^2\right)^{\frac{1}{2}}}{m_0} \tag{1.11}$$

yields

$$H' = T^{-1} H T = \underline{\alpha}\cdot\underline{p} + \alpha_0 m' \tag{1.12}$$

where

$$(m')^2 = m_0^2 - \sum_{\delta=4}^{2n} m_\delta^2$$

Here again, the transformation does not exist (T becomes singular) when $m_0^2 = \sum_{\delta=4}^{2n} m_\delta^2$. Now define

$$\gamma = \frac{i\alpha_0}{m_0}(m_4\alpha_4 + \cdots + m_{2n}\alpha_{2n})$$

so that

$$\gamma^2 = I$$

and

$$[\alpha_i, \gamma] = 0$$

Now eq. (1.9) takes the form

$$[\underline{\alpha}\cdot \underline{p} + m_0\beta(1 \pm \gamma)]\psi = 0 \tag{1.13}$$

which is of the same form as eq. (1.8). $(1 \pm \gamma)$ is, of course, singular and idempotent.

More recently, Fushchich[8] has given an equation which is CP non-invariant. Here again he uses a singular idempotent linear combination of anticommuting matrices. His equation can be written as

$$\left[\underline{\alpha}\cdot\underline{p} + m\left(1 + \frac{\underline{\alpha}\cdot\underline{p}}{p}\right)\right]\psi = i\frac{\partial\psi}{\partial t} \tag{1.14}$$

In the next two sections we shall analyse the most general two-component and four-component wave equations for a massless spin-half particle and study in detail their invariance properties under the discrete transformations C, P and T. The equations of Tokuoka and Pushchich emerge as special cases of these. We shall find that some of these equations are CP and CPT non-invariant. In particular, we shall find that some of them are CP and CPT non-invariant but T (in the sense of Wigner) invariant. This is particularly relevant when people have started discussing the possibility of CPT violation in weak interactions[9].

In the fourth section we give the transformations relating the Hamiltonian in these cases to the better known Dirac and Weyl Hamiltonians and point out that this similarity does not imply an identity of the invariance properties under the discrete symmetry transformations.

2. Two - Component Equations for a Neutrino using Non-Covariant and Spherical Factorisations:

We shall first consider the two-component equations. If we restrict ourselves to the strictly covariant equations, where the matrices concerned are strictly numerical, the Weyl equation exhausts the list of all possible equations. However, as Biedenharn[10] has pointed out in connection with his two-component equation for a spin-half massive particle, the uniqueness of Dirac's factorisation of the Klein-Gordon equation hinges on the fundamental assumption that the matrices involved are independent of space-time. If this condition is dropped, it is possible to write even two-component equations for spin-half particles. This he calls the 'Dirac dichotomy' . We shall here consider the consequences of relaxing this condition for massless particle.

Let the massless spin-half particle be described by the wave equation[11]

$$(\underline{\sigma}\cdot\underline{p} + mB)\,\varphi(\underline{x},t) = i\frac{\partial\varphi}{\partial t} \tag{2.1}$$

where m is an arbitrary parameter and B is a matrix such that

$$\{\underline{\sigma}\cdot\underline{p}, B\} = 0 \tag{2.2}$$

and

$$B^2 = 0 \tag{2.3}$$

so that φ satisfies the Klein-Gordon equation with zero mass. Thus the m here is not the mass of the particle. To find the most general form for B , we make the following transformation on the three σ - matrices satisfying the Clifford algebra

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\,; \quad i,j = 1,2,3 \tag{2.4}$$

We choose the σ's as

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

Also we note that

$$\sigma_3 = i\,\sigma_1\sigma_2$$

We now construct the matrices given by

$$B_i = \Lambda_{ij}\sigma_j \quad \text{with} \quad \Lambda_{ik}\Lambda_{jk} = \delta_{ij} \tag{2.5}$$

Then, by an analogue of Pauli's theorem, the B_i constitute a set of three mutually anti-commuting matrices with the square of each being I , the two-dimensional unit matrix. We choose Λ_{ij} as the matrix

$$[\Lambda_{ij}] = \begin{bmatrix} fp_1 & fp_2 & fp_3 \\ gp_2 & -gp_1 & 0 \\ hp_3p_1 & hp_3p_2 & -h(p_1^2+p_2^2) \end{bmatrix} \tag{2.6}$$

where

$$f = \frac{1}{(p_1^2+p_2^2+p_3^2)^{\frac{1}{2}}} = \frac{1}{p}$$

$$g = \frac{1}{(p_1^2+p_2^2)^{\frac{1}{2}}}$$

$$h = \frac{1}{p(p_1^2+p_2^2)^{\frac{1}{2}}} = \frac{g}{p} = fg \tag{2.7}$$

Hence the set B_i becomes

$$B_1 = \underline{\sigma}\cdot\hat{\underline{p}}$$

$$B_2 = \frac{p_2\sigma_1 - p_1\sigma_2}{(p_1^2+p_2^2)^{\frac{1}{2}}}$$

$$B_3 = \frac{p_3p_1\sigma_1 + p_3p_2\sigma_2 - (p_1^2+p_2^2)\sigma_3}{p(p_1^2+p_2^2)^{\frac{1}{2}}}$$

$$= i\,\underline{\sigma}\cdot\hat{\underline{p}}\,B_2 \tag{2.8}$$

where

$$\underline{\sigma}\cdot\hat{\underline{p}} = \frac{\underline{\sigma}\cdot\underline{p}}{p} \quad \text{and} \quad \{B_i\, B_j\} = 2\delta_{ij}$$

In the case of 2 x 2 matrices, the above set can also be written by observing

$$(\underline{\sigma}\cdot\underline{p})(\underline{\sigma}\cdot\underline{K}) = \underline{p}\cdot\underline{K} + i\,\underline{\sigma}\cdot(\underline{p}\times\underline{K}) \tag{2.9}$$

from which we conclude $\underline{\sigma}\cdot\underline{K}$ anticommutes with $\underline{\sigma}\cdot\underline{p}$ if $\underline{K}$ is a vector orthogonal to p. And since $\underline{K}$ is a vector in a 3-dimensional space, there are two such linearly independent vectors which may be chosen as above. This method, however, cannot be used for 4 x 4 matrices.

The most general matrix which anticommutes with $\underline{\sigma}\cdot\underline{p}$ is thus

$$B = a_2B_2 + a_3B_3$$
$$= (a_2 + i a_3\,\underline{\sigma}\cdot\hat{\underline{p}})B_2 \tag{2.10}$$

If, further $B^2 = 0$, we get the condition

$$a_2^2 + a_3^2 = 0$$

or

$$a_3 = \pm i a_2 \tag{2.11}$$

Thus the most general 2-component equation is

$$H_I \varphi \equiv [\underline{\sigma} \cdot \underline{p} + m(1 + \underline{\sigma} \cdot \hat{\underline{p}})(p_2\sigma_1 - p_1\sigma_2)]\varphi = i\frac{\partial \varphi}{\partial t} \tag{2.12}$$

We observe again that

$$H^2 = p^2$$

and the arbitrary parameter m which is not related to the mass disappears in the Klein-Gordon equation satisfied by φ . The equation is C,P, CP, T_P (in the sense of Pauli) and CPT_W non-invariant. When m is real, the equation is however T_W (time reversal in the sense of Wigner) invariant, otherwise non-invariant. The details are listed in Table I.

A different equation[11] can be constructed by using a spherical factorisation[10] of the Klein-Gordan equation as has been done by Biedenharn et al in the case of a massive particle. We observe that $\underline{\sigma} \cdot \underline{p}$ anti-commutes with the operator $B_2 = (\underline{\sigma} \cdot \underline{L} + 1)$ where $\underline{L} = \underline{r} \times \underline{p}$ is the orbital angular momentum operator and

TABLE I

BEHAVIOUR OF EQUATIONS (2.12) and (2.13) UNDER DISCRETE TRANSFORMATIONS

Equation	C	P	T_W	T_P	CP	CPT_W
2.12	Not invariant	Not invariant	Invariant if m real $T_W = \sigma_2$	Not invariant	Not invariant	Not invariant
2.13	Not invariant	Not invariant	Invariant if m real $T_W = \sigma_2$	Not invariant	Not invariant	Not invariant

$\underline{r}$ and $\underline{p}$ have their usual quantum mechanical commutation relations. Further, $\underline{\sigma}\cdot\underline{p}$ also anticommutes with $B_3 = i\,\underline{\sigma}\cdot\hat{\underline{p}}\,(\underline{\sigma}\cdot\underline{L}+1)$ Thus, choosing a suitable linear combination of B_2 and B_3 as above, we construct the following equation for a spin-half massless particle.

$$H_I\varphi \equiv \left[\underline{\sigma}\cdot\underline{p} + m(1+\underline{\sigma}\cdot\hat{\underline{p}})(\underline{\sigma}\cdot\underline{L}+1)\right]\varphi = i\frac{\partial\varphi}{\partial t} \tag{2.13}$$

This equation is also C,P,CP and CPT_W non-invariant in contrast to the Weyl equation but, for m real, T_W invariant. (The details are listed in Table I). This is because of the mixed properties under the discrete transformations of the second term in both these eqs. (2.12) and (2.13). It is interesting to note that the possibility of CPT violation that has already been envisaged, although in the neutral K-decay, could as well occur in weak processes involving the neutrino[9].

3. Four-Component Equations for Massless Spin-half Particles:

We shall now analyse the most general four-component wave equations for a massless spin-half particle and study the invariance properties of these equations under the discrete transformations.

Let the particle be described by the wave equations[13]

$$H\psi \equiv (\underline{\alpha}\cdot\underline{p} + mA)\,\psi(\underline{x},t) = i\frac{\partial\psi}{\partial t} \tag{3.1}$$

where m is an arbitrary parameter and A is a matrix which anticommutes with $\underline{\alpha}\cdot\underline{p}$ and whose square is zero, so that Ψ satisfies the Klein-Gordon equation with zero mass. Thus the m here is not the mass of the particle. To find the most general form for A we make the following transformation on the five α -matrices satisfying the Clifford algebra C_4^2 , that is

$$\{\alpha_\mu, \alpha_\nu\} = 2\delta_{\mu\nu}\,; \quad \mu,\nu = 1,2,\ldots,3. \tag{3.2}$$

α_5 being given by the product of the first four α's .
We choose the α's as

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix}, \quad i = 1,2,3.$$

$$\alpha_4 = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} = \beta,$$

$$\alpha_5 = \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix} \tag{3.3}$$

Hence

$$\alpha_5 = i\gamma_5\beta$$

where

$$\gamma_5 = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

Here σ_i denote the Pauli matrices and I the 2-dimensional unit matrix. We now construct the matrices given by

$$A_\mu = \Lambda_{\mu\nu}\alpha_\nu \tag{3.4}$$

with

$$\Lambda_{\mu\nu}\Lambda_{\mu\lambda} = \delta_{\nu\lambda} \tag{3.5}$$

Then, by Pauli's theorem, the A_μ constitute a set of five mutually anticommuting matrices with the square of each being I, the 4-dimensional unit matrix. We choose Λ as the matrix

$$\Lambda = \begin{bmatrix} fp_1 & fp_2 & fp_3 & 0 & 0 \\ gp_2 & -gp_1 & 0 & 0 & 0 \\ hp_3p_1 & hp_3p_2 & -h(p_1^2+p_2^2) & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

where

$$f = \frac{1}{(p_1^2+p_2^2+p_3^2)^{\frac{1}{2}}} = \frac{1}{p}$$

$$g = \frac{1}{(p_1^2+p_2^2)^{\frac{1}{2}}}$$

$$h = \frac{1}{p(p_1^2+p_2^2)^{\frac{1}{2}}} = \frac{g}{p} \tag{3.7}$$

Hence the set A_μ becomes

$$A_1 = \underline{\alpha}\cdot\hat{p},$$
$$A_2 = \frac{\alpha_1 p_2 - \alpha_2 p_1}{(p_1^2+p_2^2)^{\frac{1}{2}}},$$
$$A_3 = \frac{\alpha_1 p_1 p_3 + \alpha_2 p_2 p_3 - \alpha_3(p_1^2+p_2^2)}{p(p_1^2+p_2^2)^{\frac{1}{2}}} = i\gamma_5 A_1 A_2,$$
$$A_4 = \alpha_4,$$
$$A_5 = \alpha_5. \tag{3.8}$$

In contrast to A_4 and A_5 , A_2 and A_3 involve the components of momentum as entries. Thus we find that there are only 8 linearly independent matrices anticommuting with $A_1 = \underline{\alpha}\cdot\hat{p}$. They may be written as

$$A_2,\ A_3,\ A_4,\ A_5,\ A_1(A_2,\ A_3,\ A_4,\ A_5)$$

The most general linear combination of these matrices may be written as

$$A = \sum_{\mu=1}^{4}(b_\mu A_\mu + c_\mu A_1 A_\mu) \tag{3.9}$$

The square of each matrix in the sum is a scalar and A_μ anticommutes with all other members of the set except $A_1 A_\nu$ where $\mu \neq \nu$. Therefore, if A^2 is to be a scalar, the coefficient of $A_\mu A_1 A_\nu$, $\mu \neq \nu$. must be zero. Thus

$$b_\mu c_\nu - b_\nu c_\mu = 0 \tag{3.10}$$

or

$$\frac{b_\mu}{c_\mu} = \frac{b_\nu}{c_\nu} = \frac{a_0}{a_1} \quad (\text{say}) \tag{3.11}$$

Thus the most general linear combination of these matrices such that its square is a scalar may be written as

$$A = (a_0 + a_1 A_1)(a_2 A_2 + a_3 A_3 + a_4 A_4 + a_5 A_5) \tag{3.12}$$

where a_i are arbitrary. If eq. (3.1) is to represent a massless particle, then

$$A^2 = 0 \tag{3.13}$$

Hence we have

$$(a_0^2 - a_1^2)(a_2^2 + a_3^2 + a_4^2 + a_5^2) = 0 \tag{3.14}$$

Therefore the parameters a_i must satisfy the relations

$$a_0^2 - a_1^2 = 0, \qquad a_2, a_3, a_4, a_5 \quad \text{arbitrary,} \tag{3.15a}$$

or

$$\sum_{r=2}^{5} a_r^2 = 0, \qquad a_0, a_1 \quad \text{arbitrary} \tag{3.15b}$$

Thus eq. (3.1) describing the massless spin-half particle can be written as

$$H_{II}\psi \equiv \left[\underline{\alpha}\cdot\underline{p} + (1 \pm \underline{\alpha}\cdot\hat{\underline{p}})\left(\sum_{r=2}^{5} a_r A_r\right)\right]\psi = i\frac{\partial\psi}{\partial t}$$

with a_1 arbitrary (3.16a)

or

$$H_{II}\psi = \left[\underline{\alpha}\cdot\underline{p} + (a_0 + a_1\underline{\alpha}\cdot\hat{\underline{p}})\left(\sum_{r=2}^{5} a_r A_r\right)\right]\psi = i\frac{\partial\psi}{\partial t}$$

with

$$\sum_{r=2}^{5} a_r^2 = 0\ ; \quad a_0, a_1 \quad \text{arbitrary.} \tag{3.16b}$$

The most interesting results come from the anticommutativity of α_4 and α_5, although the presence of the non-covariant terms may have interesting implications[10]. Some special cases of these equations have been considered in references (4,5,6,8) as outlined in the beginning of this chapter. Among the covariant equations, we find that the equations

$$\left[\underline{\alpha}\cdot\underline{p} + m\beta\,\underline{\alpha}\cdot\hat{\underline{p}}(1 \pm \gamma_5)\right]\psi = i\frac{\partial\psi}{\partial t} \tag{3.17}$$

or, more generally,

$$\left[\underline{\alpha}\cdot\underline{p} + \beta(a_0 + a_1\underline{\alpha}\cdot\hat{\underline{p}})(1 \pm \gamma_5)\right]\psi = i\frac{\partial\psi}{\partial t} \tag{3.18}$$

have not been considered earlier in reference (8). With m real, the first of these equations is P, C and T_P non-invariant but CP, T_W and CPT_W invariant. So also the equations

$$[\underline{\alpha}\cdot\underline{p} + m\beta(1 \pm \alpha\cdot\hat{p})\gamma_5]\psi = i\frac{\partial\psi}{\partial t} \tag{3.19}$$

or, more generally,

$$[\underline{\alpha}\cdot\underline{p} + \beta(1 \pm \underline{\alpha}\cdot\hat{\underline{p}})(a_4 + a_5\gamma_5)]\psi = i\frac{\partial\psi}{\partial t} \tag{3.20}$$

All these equations are particular cases of eqs.(3.16a) and (3.16b).

Eq. (3.16a) is C, T_P, CP and CPT_W non-invariant (where T_P and T_W refer to the Pauli and Wigner time reversal operators). It is P invariant if $a_3 = 0$ and T_W invariant if a_2, a_4 are real and a_3, a_5 are pure imaginary.
Eq. (3.16b), on the other hand, is both CP and CPT_W invariant with suitable choice of parameters a_i. We give in Table II the behaviour of these equations under these discrete transformations. This will be useful in the construction of equations with any desired property under the discrete transformations. It is interesting to note that in both eqs.(3.16a) and (3.16b), the Hamiltonian is non-hermitian, (though it is diagonalisable as will be shown in the last section), in contrast to the Weyl equation where it is hermitian. Hence the same remarks apply here as in section 1.

TABLE II

BEHAVIOUR OF EQUATIONS (3.16a) and (3.16b) UNDER DISCRETE TRANSFORMATIONS

EQUATION	C	P	T_ω	T_P	CP	CPT_ω
3.16a	Non-invariant	Invariant if $a_3=0$ $a_4^2+a_5^2\neq 0$ $P=\frac{(a_4+ia_5\gamma_5)\beta}{(a_4^2+a_5^2)^{1/2}}$	Invariant if a_2, a_4 real a_3, a_5 imag. $T_\omega = i\alpha_1\alpha_3$	Non - invariant	Non - Invariant	Non - Invariant
3.16b	Invariant if a_0, a_2, real a_1, a_3 imag. or vice versa with a_4, a_5 real $C=i\beta\alpha_2$; or a_4, a_5 imag. $C=i\alpha_5\alpha_2$	Invariant if $a_3=0$ $a_4^2+a_5^2\neq 0$ $P=\frac{(a_4+ia_5\gamma_5)\beta}{(a_4^2+a_5^2)^{1/2}}$	Invariant if a_0, a_1, a_2, a_4 real; a_3, a_5 imag. or vice versa $T_\omega = i\alpha_1\alpha_3$	Invariant if $a_1=0$ $a_4^2+a_5^2\neq 0$ $T_P=\frac{(a_5-ia_4\gamma_5)\beta}{(a_4^2+a_5^2)^{1/2}}$	Invariant if a_0, a_2, a_3, a_5 real; a_1, a_4 imag. or vice versa $CP = i\alpha_1\alpha_4$	Invariant if $a_1=a_3=0$ $CPT_\omega = \gamma_5$ $a_0=a_2=0$ $CPT_\omega = I$

T_ω = Wigner Time Reversal, T_P = Pauli Time Reversal.

We also notice that βA is not necessarily idempotent but is factorisable into two commuting matrices at least one of which is idempotent in the sense $X^2 = a X$. Thus for eqs. (3.16a) and (3.16b)

$$\beta A = XY \tag{3.20}$$

where, for eq. (3.16a)

$$X = (1 \mp \underline{\alpha} \cdot \hat{\underline{p}})$$

$$Y = \beta(a_2 A_2 + \cdots + a_5 A_5) \tag{3.21}$$

and X is idempotent in the above sense and commutes with Y. For eq. (3.16b)

$$X = \beta_0 (a_2 A_2 + \cdots + a_5 A_5)$$

$$Y = a_0 - a_1 \underline{\alpha} \cdot \hat{\underline{p}} \tag{3.22}$$

where X is idempotent in the sense

$$X^2 = 2a_4 X$$

and commutes with Y.

In fact a similar statement can be made for $A_i A$ where $i = 2, \ldots, 5$. $A_i A$ is factorisable into two commuting matrices at least one of which is idempotent in the sense $X^2 = a X$. Also the matrix A can be factorised into two matrices one of which has square zero. Thus for eq. (3.16a)

$$A = [(1 \pm \underline{\alpha}\cdot\hat{\underline{p}})\beta][\beta(a_2A_2 + \cdots + a_5A_5)] \tag{3.23}$$

where the first factor has square zero. For eq. (3.16b) we have

$$(a_2A_2 + \cdots + a_5A_5)^2 = 0 \tag{3.24}$$

The spherical factorisation of the Klein-Gordon equation can also be used to build, yet another class of 4-component relativistic equations for a massless spin-half particle[11] with prescribed invariance properties under C,P,T. We note again the following 4 x 4 matrices mutually anticommute, namely

$$\begin{aligned}
A_1 &= \underline{\alpha}\cdot\hat{\underline{p}}, \\
A_2 &= \gamma_5(\underline{\sigma}\cdot\underline{L} + 1)/(j + \tfrac{1}{2}), \\
A_3 &= i\underline{\alpha}\cdot\hat{\underline{p}}\,(\underline{\sigma}\cdot\underline{L} + 1)/(j + \tfrac{1}{2}), \\
A_4 &= \alpha_4 = \beta, \\
A_5 &= \alpha_5 = i\gamma_5\beta
\end{aligned} \tag{3.25}$$

where α, β and γ_5 are as defined earlier. Here we have $J = L + \frac{1}{2}$ = Total angular momentum and $J^2\psi = j(j+1)\psi$

This ensures that A_2^2 and A_3^2 act on ψ as unit operators.

Thus there are 8 linearly independent matrices anticommuting with $\underline{\alpha}\cdot\underline{p}$. They may be written as

$$A_2,\ A_3,\ A_4,\ A_5,\ A_1(A_2,\ A_3,\ A_4,\ A_5)$$

As before, the most general linear combination of these matrices such that its square is a scalar may be written as

$$A = (a_0 + a_1 A_1)(a_2 A_2 + \cdots + a_5 A_5) \tag{3.26}$$

where a_i are arbitrary. We write the equation for the massless particle as

$$(\underline{\alpha}\cdot\underline{p} + mA)\psi = i\frac{\partial\psi}{\partial t}$$

If this is to represent a spin-half particle, we require

$$A^2 = 0$$

Hence, as before

$$(a_0^2 - a_1^2)(a_2^2 + a_3^2 + a_4^2 + a_5^2) = 0 \tag{3.27}$$

Therefore the parameters must satisfy

$$a_0^2 - a_1^2 = 0\ ; \qquad a_2, \ldots, a_5 \quad \text{arbitrary} \tag{3.28}$$

or

$$\sum_{r=2}^{5} a_r^2 = 0\ ; \qquad a_0, a_1 \quad \text{arbitrary} \tag{3.29}$$

Then eq. (3.1) can be rewritten as

$$H_{\mathrm{III}}\psi \equiv [\underline{\alpha}\cdot\underline{p} + (1 \pm \underline{\alpha}\cdot\hat{\underline{p}})(a_2 A_2 + \cdots + a_5 A_5)]\psi = i\frac{\partial\psi}{\partial t} \tag{3.30a}$$

with a_i arbitrary

or

$$H_{\mathrm{III}}\psi = [\underline{\alpha}\cdot\underline{p} + (a_0 + a_1\underline{\alpha}\cdot\hat{\underline{p}})(a_2 A_2 + \cdots + a_5 A_5)]\psi = i\frac{\partial\psi}{\partial t} \tag{3.30b}$$

with

$$\sum_{r=2}^{5} a_r^2 = 0 ; \qquad a_0, a_1 \text{ arbitrary}$$

These equations supplement the ones given in the last section. We list in Table III the invariance properties of these equations under discrete transformations. The same remarks as earlier about the hermiticity of the Hamiltonian apply here too. Here also βA is factorisable into two commuting matrices at least one of which is idempotent in the sense $X^2 = aX$. And A can be factorised into two matrices one of which has square zero.

Thus we have demonstrated that wave equations can be constructed for a massless spin-half particle with any prescribed invariance properties under C,P,T. We have also shown that a

TABLE III

BEHAVIOUR OF EQUATIONS (3.30a) and (3.30b) UNDER DISCRETE TRANSFORMATIONS

Equation	C	P	T_W	CP	T_P	CPT_W
3.30a	Not Invariant	Invariant if $a_2 = 0$ $a_4^2 + a_5^2 \neq 0$ $P = \frac{a_4\alpha_4 + a_5\alpha_5}{(a_4^2 + a_5^2)^{1/2}}$ OR $a_2 = a_4 = a_5 = 0$ $P = \alpha_4$ or α_5	Invariant if a_2, a_5 real a_3, a_4 imag. $T_W = \alpha_2$ OR a_2, a_4 real a_3, a_5 imag. $T_W = i\alpha_1\alpha_3$	Not invariant	Not invariant	Not invariant
3.30b	Invariant if a_0, a_2 real a_1, a_3, a_4, a_5 imag. or vice versa, $C = i\alpha_2\alpha_5$ OR a_0, a_2, a_4, a_5 real, a_1, a_3 imag. or vice versa, $C = i\alpha_2\alpha_4$	Invariant if $a_2 = 0$ $a_4^2 + a_5^2 \neq 0$ $P = \frac{a_4\alpha_4 + a_5\alpha_5}{(a_4^2 + a_5^2)^{1/2}}$	Invariant if a_0, a_1, a_2, a_5 real a_3, a_4 imag. or vice versa $T_W = \alpha_2$ OR a_0, a_1, a_2, a_4 real; a_3, a_5 imag. or vice versa $T_W = i\alpha_1\alpha_3$	Invariant if a_0, a_4 real a_1, a_2, a_3, a_5 imag. or vice versa, $CP = \alpha_2$ OR a_0, a_5 real a_1, a_2, a_3, a_4 imag. or vice versa $CP = i\alpha_1\alpha_3$	Invariant if $a_1 = 0$ $a_4^2 + a_5^2 \neq 0$ $T_P = \frac{a_5\alpha_4 - a_4\alpha_5}{(a_4^2 + a_5^2)^{1/2}}$ OR $a_1 = a_4 = a_5 = 0$ $T_P = \alpha_4$ or α_5	Invariant if $a_1 = a_2 = 0$ $CPT_W = \gamma_5$ OR $a_1 = a_3 = 0$ $CPT_W = I$

T_W = Wigner Time Reversal, T_P = Pauli Time Reversal.

possible CP and CPT violation can well be incorporated in a T-invariant relativistic equation for a massless particle. The further implications of these equations are yet unclear although the corresponding equations of the massive particle have attracted the attention of many[10].

4. Transformation Connecting the Hamiltonians with some Elementary:

A vital feature of all the above equations is that the Hamiltonian satisfies $H^2 = p^2$ just as in the case of the Weyl equation. The eigenvalues of H are therefore given by $\pm p$ in all these cases. In fact for all these Hamiltonians, the multiplicities of the eigenvalues as also the elementary divisors are the same as can be seen from what follows. Hence the Hamiltonians in all the 4-component equations including the ones given by Sen Gupta and Fushchich are connected to each other and to the 4-component Dirac equation for a massless particle by suitable similarity transformations. Similarly the two-component equations are connected to the Weyl equation by a similarity transformation. Further, the Weyl equation itself, repeated twice along the diagonal can be related to the Dirac equation for a massless particle by a similarity transformation.

These similarity transformations are not all unitary transformations. Indeed they cannot be, for the Weyl Hamiltonian and the Dirac Hamiltonian for a massless particle are hermitian while all the others are non-hermitian and since the hermiticity of a matrix is not changed by a unitary transformation, a hermitian matrix cannot be unitarily similar to a non-hermitian matrix. Further, a salient feature of these transformations is that they are not purely numerical matrices but involve the components of momentum and hence one cannot conclude that the invariance properties of all these equations under any transformation will be the same. In fact, as will be shown presently, the Weyl Hamiltonian itself repeated twice along the diagonal can be related to the Dirac Hamiltonian for a massless particle by a similarity transformation (involving, of course, the components of momentum) and we know that while the Dirac equation is C and P invariant, the Weyl equation is not.

Let us first find the transformation connecting eqs.(3.16) to the Dirac equation for a massless particle. We wish to find S such that

$$S^{-1} H_{II} S = H_D \tag{4.1}$$

where

$$H_{II} = \underline{\alpha}\cdot\underline{p} + mA \tag{4.2}$$

and

$$H_D = \underline{\alpha}\cdot\underline{p} \tag{4.3}$$

We note that

$$H_{II}^2 = H_D^2 = p^2 \tag{4.4}$$

We observe, following the U matrix method of Ramakrishnan, that

$$H_{II}(H_{II} + H_D) = (H_{II} + H_D) H_D \tag{4.5}$$

and

$$(H_{\Pi} + H_D)^2 = 4p^2 \tag{4.6}$$

Hence, $H_{\Pi} + H_D$ is nonsingular and

$$S' = (S')^{-1} = \frac{1}{2p}(2\underline{\alpha}\cdot\underline{p} + mA) = \frac{H_{\Pi} + H_D}{2p} \tag{4.7}$$

satisfies eq. (4.1) and is therefore the desired transformation. Actually, S is arbitrary upto post-multiplication by a matrix which commutes with H_D for, if X is non singular and commutes with H_D

$$(S'X)^{-1} H_{\Pi} (S'X) = X^{-1} H_D X = H_D \tag{4.8}$$

so that S = S' X also transforms H_{Π} to H_D. Choosing X = $\underline{\alpha}\cdot\hat{\underline{p}}$, from eq. (4.7), we obtain S as

$$S = 1 + \frac{1}{2} mA\,\underline{\alpha}\cdot\underline{p} = \exp \frac{1}{2} mA\,\underline{\alpha}\cdot\underline{p} \tag{4.9a}$$

and

$$S^{-1} = 1 - \frac{1}{2} mA\,\underline{\alpha}\cdot\underline{p} = \exp - \frac{1}{2} mA\,\underline{\alpha}\cdot\underline{p} \tag{4.9b}$$

We can now extend the above procedure to obtain the transformation which diagonalises H_{Π}. For this, we shall first find the matrix which diagonalises H_D. Again, as before,

$$H_D(H_D + \beta p) = (H_D + \beta p)\beta p \tag{4.10}$$

and

$$(H_D + \beta p)^2 = 2p^2 \tag{4.11}$$

Therefore

$$U' = \frac{1}{\sqrt{2p}}\left(H_D + \beta p\right) \tag{4.12}$$

or, since again U' is arbitrary upto postmultiplication by a matrix which commutes with β , the Hamiltonian H_D is diagonalised by

$$U = \frac{1}{\sqrt{2}}(1 + \underline{\alpha}\cdot\hat{\underline{p}}\beta) = \exp\left(\underline{\alpha}\cdot\hat{\underline{p}}\beta\frac{\pi}{4}\right) \tag{4.13}$$

$$U^{-1} = \frac{1}{\sqrt{2}}(1 - \underline{\alpha}\cdot\hat{\underline{p}}\beta) = \exp\left(-\underline{\alpha}\cdot\hat{\underline{p}}\beta\frac{\pi}{4}\right) = U^{+} \tag{4.14}$$

that is

$$U^{-1} H_D U = \beta p \tag{4.15}$$

Hence, combining eqs. (4.1) and (4.15), the transformation which diagonalises H_{II} is given by

$$T^{-1} H_{II} T = \beta p \tag{4.16}$$

where

$$T = SU = \exp\left(\tfrac{1}{2} mA\, \underline{\alpha}\cdot\underline{p}\right)\exp\left(\underline{\alpha}\cdot\hat{\underline{p}}\beta\frac{\pi}{4}\right) \tag{4.17}$$

or explicitly

$$T = \frac{1}{\sqrt{2}}\left(1 + \tfrac{1}{2} mA\, \underline{\alpha}\cdot\underline{p}\right)\left(1 + \underline{\alpha}\cdot\hat{\underline{p}}\beta\right) \tag{4.18}$$

For the case of H_{III}, the transformations which connect it to H_D, and which diagonalise it will be given by the same expressions as eqs. (4.8) and (4.17), except that here A_μ and hence A are given by eq. (3.25).

We shall now derive the transformations which connect the 2-component equations to the 2-component Weyl equation. The Hamiltonian in the case of eqs.(2.12) and (2.13) is given by

$$H_I = \underline{\sigma}\cdot\underline{p} + mB \tag{4.19}$$

where the matrix B is as given by eqs.(2.10) and (2.13) respectively. The Weyl Hamiltonian is given by

$$H_W = \underline{\sigma}\cdot\underline{p} \tag{4.20}$$

As before, it can be shown that the transformation connecting H_I and H_W is given by

$$R^{-1} H_I R = H_W \tag{4.21}$$

where

$$R = 1 + \tfrac{1}{2} B\, \underline{\sigma}\cdot\underline{p} = \exp \tfrac{1}{2} B\, \underline{\sigma}\cdot\underline{p} \tag{4.22}$$

Next we shall find the transformation V which diagonalises H_W. We have

$$H_W(\underline{\sigma}\cdot\underline{p} + \sigma_3 p) = (\underline{\sigma}\cdot\underline{p} + \sigma_3 p)\,\sigma_3 p \tag{4.23}$$

with

$$(\underline{\sigma}\cdot\underline{p} + \sigma_3 p)^2 = 2p(p + p_3) \tag{4.24}$$

which shows $\underline{\sigma}\cdot\underline{p} + \sigma_3 p$ is nonsingular. Again we post-multiply by σ_3, (since the transformation is arbitrary upto post multiplication by a matrix which commutes with σ_3) so that we get

$$V^{-1} H_W V = \sigma_3 p \tag{4.25}$$

where

$$V = \left[\frac{\sigma_3 p}{2(p+p_3)}\right]^{\frac{1}{2}} (1 + \underline{\sigma}\cdot\hat{\underline{p}}\,\sigma_3) \tag{4.26}$$

and

$$V^{-1} = \left[\frac{p}{2(p+p_3)}\right]^{\frac{1}{2}} (1 + \sigma_3\,\underline{\sigma}\cdot\hat{\underline{p}}) \tag{4.27}$$

Hence, combining eqs. (4.21) and (4.26) the transformation which diagonalises H_I is given by

$$Q^{-1} H_I Q = \sigma_3 p \tag{4.28}$$

where

$$Q = RV \tag{4.29}$$

We shall now show that the Dirac Hamiltonian for a massless particle is itself similar to the Weyl Hamiltonian repeated twice along the diagonal. We know

$$P^{-1}\beta P = \begin{bmatrix} \sigma_3 & 0 \\ 0 & \sigma_3 \end{bmatrix} \tag{4.30}$$

where P is the permutation matrix

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.31}$$

Combining eqs. (4.15), (4.25), (4.30), we obtain

$$W^{-1} H_D W = H_W \tag{4.32}$$

where

$$W = U P V^{-1} \tag{4.33}$$

It has been suggested by Fushchich[13] that eqs.(2.12) and (2.13) are isometrically equivalent to the Weyl equation and eqs. (3.16) and (3.30) are isometrically equivalent to the Dirac equation for a massless particles and that all these equations have similar transformation properties under C,P,T. However, we point out that the similarity transformations connecting H_I to H_W and H_{II} and H_{III} to H_D are not unitary since H_I, H_{II}, H_{III} are non-hermitian while H_W and H_D are hermitian and a non-hermitian matrix cannot be unitarily similar to a hermitian matrix. Also these transformation matrices are not purely numerical and hence the invariance properties of these equations under the discrete transformations will not be the same. Indeed, if it were so, the equations of Fukuoka[4] and Fushchich[8] as also the Weyl equation would have the same transformation properties as the Dirac equation for a massless particle which we know, is a patently incorrect conclusion.

5. Conclusions:

Thus we have analysed the most general two-component and four-component wave equations for a massless spin-half particle using our non-covariant factorisation of the Klein-Gordon equation as also the spherical factorisation and studied their transformation properties under the discrete symmetry operations C,P,T. We have shown that it possible to write an equation with any prescribed transformation property under C,P,T. In particular, we have shown it is possible to write a CP-non invariant two-component equation. We have also given the transformations relating the Hamiltonian in these cases to the better known Dirac and Weyl Hamiltonians and shown that this similarity does not imply an identity of their invariance properties under the discrete transformations. The future implications of these equations are yet unclear although the corresponding equations for the massive particle have attracted the attention of many.

References:

1. H. Weyl, Z.Physik, **56**, 330 (1929).
2. T.D.Lee and C.N.Yang, Phys.Rev., **106**, 1671 (1957).
   L.Landau, Nucl.Phys., **3**, 127 (1957).
   A.Salam, Nuov. Cim., **5**, 299 (1957).
3. H.J.Bhabha, Rev. Mod. Phys., **21**, 451 (1949).
4. Z.Tokuoka, Progr.Theor.Phys., **37**, 603 (1967).
5. N.D.Sen Gupta, Nucl.Phys., **B4**, 147 (1967).
6. T.S.Santhanam and P.S.Chandrasekharan, Prog.Theor.Phys., **41**, 264 (1969).
7. Alladi Ramakrishnan, "L-Matrix Theory or the Grammar of Dirac Matrices", Tata-McGraw-Hill, Bombay (1972).
8. V.I. Fushchich and A.L.Grishchenko, Lett.Nuov.Cim., **4**, 927 (1970).
9. See for instance, H.Faissner, G.Kopp and P.Zerwas, Aachen Preprint, Pitha (1971).
10. L.C.Biedenharn, M.Y.Han, H.Van Dam, Phys.Rev.Lett., **27**, 1167 (1971).
11. A.R.Tekumalla and T.S.Santhanam, Lett.Nuov.Cim., **6**,99 (1973).
12. T.S.Santhanam and A.R.Tekumalla, Lett.Nuov.Cim., **3**,190 (1972).
13. V.I.Fushchich and A.G.Nikitin, Lett.Nuov.Cim., **7**, 439, (1973).

CHAPTER. IV

EQUIVALENT FORMS OF THE DIRAC HAMILTONIAN*

1. Introduction:

Several transformations yielding equivalent representations of the Dirac equation have been proposed in the past like the Foldy-Wouthuysen[1] (FW) and the Cini-Touschek[2] (CT) transformations and generalisations thereof. The FW and CT transformations are useful in the discussion of the non-relativistic and extreme relativistic limits respectively of the Dirac equation. Eriksen[3], Bhaktavatsalou[4], Saavedra[5] and de Vries[6] have given generalisations of the FW transformation. In particular, Saavedra has shown that the FW and CT Hamiltonians can be represented as orthogonal vectors in a 2-dimensional space and has given a new Hamiltonian which stands in the same relation to the Dirac Hamiltonian as CT to FW. De Vries has extended this and shown that the generalised FW transformations are connected with a 4-dimensional rotation group.

In this chapter we shall give a simple and elegant method of obtaining explicit forms of the above transformations by using the U-matrix method of Ramakrishnan[7]. We shall use this method to give explicitly the similarity transformation between any two Hamiltonians with the same elementary divisors and then give a further generalisation of the FW transformation. In section 3, we shall also show how this method can be used to obtain the transformation which connects the Dirac equation to

---

* A.R.Tekumalla, Comm. to Physica, 1974.

the non-covariant forms[9] of the Dirac equation. In section 4, we shall give a new equation for a spin-half massive particle based on a different factorisation of the Klein-Gordon equation.

2. The Foldy Wouthuysen Transformation and its Generalisations:

In general four components, and not two, are required to describe a state of positive or negative energy for a spin-half particle in the Dirac representation. The reason for this is that the Hamiltonian in the Dirac equation

$$H\psi \equiv (\underline{\alpha}\cdot\underline{p} + \beta m)\psi = i\frac{\partial\psi}{\partial t} \tag{2.1}$$

contains odd operators, namely $\underline{\alpha}$, which connect the upper and lower components of the wave function. In the FW treatment, by a canonical transformation of the wave equation, a representation is found where the Hamiltonian is an even Dirac matrix. Then the Dirac equation splits into two uncoupled equations of the Pauli type describing particles in positive and negative energy states respectively. When the particle is free, the transformation is given by Foldy and Wouthuysen[1] in a simple closed form as follows:

Let S be a hermitian operator. Then the unitary transformation

$$\begin{aligned} \psi' &= e^{iS}\psi \\ H_{FW} \equiv H' &= e^{iS} H e^{-iS} \end{aligned} \tag{2.2}$$

leaves eq.(1) in the Hamiltonian form

$$H'\psi' = i\frac{\partial\psi'}{\partial t} \tag{2.3}$$

and any physical variable whose operator representative, in the old representation is T has for its operator representative in the new representation

$$T' = e^{iS}Te^{-iS} \tag{2.4}$$

We choose, after FW, S to be of the form

$$S = -\frac{i}{2m}\beta\underline{\alpha}\cdot\underline{p}\,\omega\left(\frac{p}{m}\right) \tag{2.5}$$

where w is a real function to be determined so that H' is free of odd operators. Since S anticommutes with H,

$$\begin{aligned} H' &= e^{iS}He^{-iS} = e^{2iS}H \\ &= \left[\cos\left(\tfrac{p}{m}\omega\right) + \beta\underline{\alpha}\cdot\hat{\underline{p}}\sin\left(\tfrac{p}{m}\omega\right)\right](\underline{\alpha}\cdot\underline{p} + \beta m) \\ &= \beta\left[m\cos\left(\tfrac{p}{m}\omega\right) + p\sin\left(\tfrac{p}{m}\omega\right)\right] \\ &\qquad + \underline{\alpha}\cdot\hat{\underline{p}}\left[p\cos\left(\tfrac{p}{m}\omega\right) - m\sin\left(\tfrac{p}{m}\omega\right)\right] \end{aligned} \tag{2.6}$$

If we now set

$$\omega = \frac{m}{p}\tan^{-1}\frac{p}{m} \tag{2.7}$$

the coefficient of the $\underline{\alpha}\cdot p$ term vanishes and H' will then be free of odd operators. Thus

$$H' = \beta\left[\frac{m^2}{E} + \frac{p^2}{E}\right] = \beta E, \tag{2.8}$$

and in a representation in which β is diagonal, $H'\psi' = E'\psi'$ now has solutions where the upper components represent positive energies and the lower components negative energies. Explicitly the transformation is given by

$$\begin{aligned} e^{-iS} &= \exp\left(-\tfrac{1}{2}\beta\underline{\alpha}\cdot\hat{p}\tan^{-1}\frac{p}{m}\right) \\ &= \cos\left(\tfrac{1}{2}\tan^{-1}\frac{p}{m}\right) + \underline{\alpha}\cdot\hat{p}\beta\sin\left(\tfrac{1}{2}\tan^{-1}\frac{p}{m}\right) \\ &= \left(\frac{E+m}{2E}\right)^{\frac{1}{2}} + \underline{\alpha}\cdot\hat{p}\beta\left(\frac{E-m}{2E}\right)^{\frac{1}{2}} \\ &= \frac{\underline{\alpha}\cdot p + \beta(m+E)}{\left[2E(E+m)\right]^{\frac{1}{2}}} \end{aligned} \tag{2.9}$$

The transformed form of the operator - representatives of dynamical variables in the old and new representations is given in Table 1.

The FW representation is particularly useful for the discussion of the non-relativistic limit of the Dirac equation, since the operators representing physical quantities are in a one to one correspondence with the Pauli theory.

TABLE I

TABLE OF OPERATOR REPRESENTATIVES OF DYNAMICAL VARIABLES IN OLD AND NEW REPRESENTATIONS FOR FW TRANSFORMATION

Dynamical Variable	Operator representative in Dirac representation	Operator representative in new representation with $X = \beta$	Operator representative in new representation with $X = 1$
Position	$\underline{q}$	$\underline{q}' = \underline{q} - \frac{i\beta\underline{\alpha}}{2E} + \frac{i\beta(\underline{\alpha}\cdot\underline{p})\hat{\underline{p}} - [\underline{\sigma}\times\underline{p}]}{2E(E+m)}$	$\underline{q}'' = \underline{Q}$
Momentum	$\underline{p} = i\hbar\underline{\nabla}$	$\underline{p}' = \underline{p}$	$\underline{p}'' = \underline{p}$
Hamiltonian	$H = \underline{\alpha}\cdot\underline{p} + \beta m$	$H = \beta(m^2+p^2)^{1/2} = \beta E$	$H'' = \beta E$
Orbital angular momentum	$\underline{l} = [\underline{q}\times\underline{p}]$	$\underline{l}' = [\underline{q}'\times\underline{p}]$	$\underline{l}'' = \underline{L}$
Spin angular momentum	$\underline{\sigma} = -\frac{i}{2}[\underline{\alpha}\times\underline{\alpha}]$	$\underline{\sigma}' = \underline{\sigma} + \frac{i\beta[\underline{\alpha}\times\underline{p}]}{E} - \frac{\underline{p}\times[\underline{\sigma}\times\underline{p}]}{E(E+m)}$	$\underline{\sigma}'' = \underline{\Sigma}$
Mean Position	$\underline{Q} = \underline{q} + \frac{i\beta\underline{\alpha}}{2E} - \frac{i\beta(\underline{\alpha}\cdot\underline{p})\hat{\underline{p}} + [\underline{\sigma}\times\underline{p}]}{2E(E+m)}$	$\underline{Q}' = \underline{q}$	$\underline{Q}'' = \underline{q}$
Mean orbital angular momentum	$\underline{L} = [\underline{Q}\times\underline{p}]$	$\underline{L}' = \underline{l}$	$\underline{L}'' = \underline{l}$
Mean spin angular momentum	$\underline{\Sigma} = \underline{\sigma} - \frac{i\beta[\underline{\alpha}\times\underline{p}]}{E} - \frac{[\underline{p}\times[\underline{\sigma}\times\underline{p}]]}{E(E+m)}$	$\underline{\Sigma}' = \underline{\sigma}$	$\underline{\Sigma}'' = \underline{\sigma}$

There exists another limit of equal interest, namely the ultrarelativistic limit[2] where the mass of the particle can be neglected in comparison, to its kinetic energy, that is, m can be neglected in comparison to p . A form of the Dirac equation which has this property is obtained by choosing w such that the coefficient of β vanishes and only the term in $\underset{\sim}{\alpha}\cdot\hat{p}$ remains. Thus if

$$\omega' = -\frac{m}{p}\tan^{-1}\left(\frac{m}{p}\right) \tag{2.10}$$

$$H_{CT} \equiv H'' = \underset{\sim}{\alpha}\cdot\hat{p}\left[\frac{p^2}{E}+\frac{m^2}{E}\right] = \underset{\sim}{\alpha}\cdot\hat{p}E \tag{2.11}$$

and the explicit form of e^{-iS} is given by

$$e^{-iS} = \frac{\left[\underset{\sim}{\alpha}\cdot p + \beta m + \underset{\sim}{\alpha}\cdot\hat{p}E\right]\underset{\sim}{\alpha}\cdot\hat{p}}{\left[2E(E+m)\right]^{1/2}} \tag{2.12}$$

A simple analogy between the FW transformation and two-dimensional rotations was pointed out by Saavedra[5]. He regards the Dirac representation as intermediate between the FW and CT representations. Thus the Dirac Hamiltonian can be written as

$$\begin{aligned} H_0 &= \beta E\cos\theta + \underset{\sim}{\alpha}\cdot\hat{p}E\sin\theta \\ &= H_{FW}\cos\theta + H_{CT}\sin\theta \end{aligned} \tag{2.13}$$

where

$$\theta = \tan^{-1}\frac{p}{m} \tag{2.14}$$

Thus if we consider a two-dimensional space spanned by the orthogonal vectors H_{FW} and H_{CT}, a positive rotation through an angle θ will take the axis H_{FW} to the position H which will be identical to the Dirac Hamiltonian for the above value of θ. Therefore another equivalent Hamiltonian can be written as

$$H' = -H_{FW}\sin\theta + H_{CT}\cos\theta$$
$$= -p\beta + m\underline{\alpha}\cdot\hat{p} \tag{2.15}$$

which is obtained from H_D by

$$H' = S^{-1}H_D S$$

where

$$S = \exp\frac{\pi}{4}\beta\underline{\alpha}\cdot\hat{p} \tag{2.16}$$

Thus H' stands in the same relation to H_D as H_{CT} to H_{FW}. Saavedra suggests that this Hamiltonian H' is suitable for the discussion of massless fermions, the corresponding wave functions in the new representation being directly given in terms of two-component spinors.

This connection has been extended by de Vries[6] who showed that the generalised FW transformations are connected with a four dimensional rotation group. His Hamiltonian is an arbitrary linear combination of the four Dirac matrices such that the square is E^2.

3. The U-Matrix Method and its Generalisation:

We shall now give a simple and elegant method of obtaining explicit forms of the above transformations using the U-matrix method of Alladi Ramakrishnan[7]. We shall use this method to give explicitly the similarity transformation between any two Hamiltonians with the same elementary divisors and then give a further generalisation of the FW transformation.

Let H be the Dirac Hamiltonian given by

$$H = \underset{=}{\alpha} \cdot \underset{=}{p} + \beta m, \tag{3.1}$$

and U the matrix which diagonalises it, so that

$$U^{-1} H U = \beta E \tag{3.2}$$

where β is diagonal. Since H is a linear combination of anticommuting matrices

$$H^2 = p^2 + m^2 = E^2, \quad E = +(p^2+m^2)^{\frac{1}{2}} \tag{3.3}$$

We observe after Ramakrishnan, that

$$H(H+\beta E) = (H+\beta E)\,\beta E \tag{3.4}$$

Since $(H+\beta E)$ is again a linear combination of anticommuting matrices,

$$(H+\beta E)^2 = 2E(E+m) \tag{3.5}$$

and therefore it is nonsingular and invertible. Hence

$$(H+\beta E)^{-1} H (H+\beta E) = \beta E \tag{3.6}$$

Thus $H+\beta E$ diagonalises H and is a U matrix of H. The matrix is obviously Hermitian. It can be made involutionary (self-inverse) and unitary by taking U as

$$U = \frac{H+\beta E}{[2E(E+m)]^{\frac{1}{2}}} \tag{3.7}$$

Actually, because of the degeneracy in the eigenvalues of H, there is an arbitrariness in the choice of the U-matrix. Thus we can write

$$H[(H+\beta E)X] = [(H+\beta E)X]\beta E \tag{3.8}$$

where X is any matrix which commutes with β and therefore has the block diagonal form

$$X = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \tag{3.9}$$

where X_1 and X_2 are arbitrary 2 x 2 nonsingular matrices. Thus the most general form for the U-matrix is

$$U = \frac{H+\beta E}{[2E(E+m)]^{\frac{1}{2}}} X . \tag{3.10}$$

We mention here that the Foldy-Wouthuyson transformation is merely a diagonalisation of the Dirac Hamiltonian and is therefore given by eq.(3.10). For the F.W. transformation in the conventional form, the choice of X is $X = \beta$. We give in Table I (pg.115) the transformed operators when $X = \beta$ and when $X = 1$. We observe that the complementarity between the Dirac representation and the transformed representation for the pairs of operations like position and mean position, orbital angular momentum and mean orbital angular momentum, spin angular momentum and mean spin angular momentum is exact in the case $X = 1$ whereas it is only upto a sign when $X = \beta$.

If X is chosen as $X = \underline{\Sigma} \cdot p + \Sigma_3 p$ where $\underline{\Sigma} = \begin{bmatrix} \underline{\sigma} & 0 \\ 0 & \underline{\sigma} \end{bmatrix}$ then U simultaneously diagonalises the Hamiltonian and the helicity operator.

We now generalize the above method as follows. Let H' be a matrix which is a new linear combination of the same anti-commuting matrices but with the same elementary divisors[10] as H. Then H' has the same diagonal form as H which may be taken as βE. Hence

$$(H')^2 = E^2 = H^2 \tag{3.11}$$

We shall now construct the transformation S such that

$$S^{-1} H S = H' \tag{3.12}$$

We can write

$$H'(H+H') = (H+H')H' \qquad (3.13)$$

Again, since $H + H'$ is a linear combination of anticommuting matrices, its square is a scalar and it is therefore nonsingular and invertible. Hence, except when $H = -H'$, we can write

$$(H+H')^{-1} H (H+H') = H' \qquad (3.14)$$

Further, $H + H'$ is hermitian and since

$$(H+H')^2 = 2E^2 + HH' + H'H \quad = \text{scalar}, \qquad (3.15)$$

S can be made hermitian, unitary and involutionary by taking

$$S = \frac{H+H'}{(2E^2 + HH' + H'H)^{1/2}} \qquad (3.16)$$

This result is true even when H' is not merely a linear combination of the same anticommuting matrices as H but any matrix with same elementary divisors as H. In such a case, the denominator is not a scalar but an operator which can be expanded in a power series since $H H' + H'H$ has eigenvalues strictly less than $2E^2$. Further, in view of the degeneracy in the eigenvalues of H and H' we can write

$$H(H+H')X = (H+H')XH' \qquad (3.17)$$

where X is an arbitrary nonsingular matrix which commutes with H' . Therefore, the most general choice for S is

$$S = (H+H')X \tag{3.18}$$

An alternative way of looking at this result is to note

$$HY(H+H') = Y(H+H')H'$$

where Y commutes with H.

Thus the result given by eq.(3.18) can also be written as

$$S = Y(H+H') \tag{3.18a}$$

where Y is an arbitrary non-singular matrix which commutes with H. Eq.(3.8) gives all the transformations given in Ref. (6) with suitable values of H' and X . For instance

$H' = \underline{\alpha}\cdot\hat{\underline{p}}E$, gives a $X = \underline{\alpha}\cdot\hat{\underline{p}}$ gives the CT transformation.

To further restrict the F.W. transformation, Eriksen imposes the condition

$$S^{\dagger}_{H} = H'S \tag{3.19}$$

This is guaranteed by

$$X = H' \tag{3.20}$$

However, Eriksen's condition does not uniquely determine the transformation. For instance, in Eriksen's case, for the F.W. transformation

$$X = \underline{\alpha}\cdot\hat{\underline{p}}\,\alpha_5 \tag{3.21}$$

where

$$\alpha_5 = \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix},$$

will also do.

4. A Further Generalisation of the FW Transformation:

In the earlier section, we required H' to be a linear combination of anticommuting matrices. Since there are five 4 x 4 matrices which satisfy the Clifford algebra

$$\{\alpha_\mu, \alpha_\nu\} = 2\delta_{\mu\nu} \tag{4.1}$$

such a most general Hamiltonian is actually a linear combination of five matrices

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4 = \beta, \alpha_5 = \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix} \tag{4.2}$$

Thus

$$H' = E\left[\sum_{i=1}^{4} a_i\alpha_i + \left(1 - \sum_{i=1}^{4} a_i^2\right)^{1/2}\alpha_5\right] \tag{4.3}$$

The transformation S which transforms the Dirac Hamiltonian to H' is given by

$$S = (H + H')X. \tag{4.4}$$

Following de Vries, H' can be rewritten using spherical coordinates as

$$H' = E(\alpha_1 \sin\eta \sin\xi \sin\theta \cos\varphi + \alpha_2 \sin\eta \sin\xi \sin\theta \sin\varphi$$
$$+ \alpha_3 \sin\eta \sin\xi \cos\theta + \alpha_4 \sin\eta \cos\xi \tag{4.5}$$
$$+ \alpha_5 \cos\eta)$$

By defining

$$p_x = E \sin\eta \sin\xi \sin\theta \cos\varphi, \quad p_y = E \sin\eta \sin\xi \sin\theta \sin\varphi,$$
$$p_z = E \sin\eta \sin\xi \cos\theta, \quad M = E \sin\eta \cos\xi,$$
$$\chi = E \cos\eta \tag{4.6}$$

We see that p_x, p_y, p_z, M and χ can be looked upon as the components of a vector with length E in a 5-dimensional Euclidean space and the generalized F.W. transformations are rotations in the 5-dimensional space. All the earlier transformations can be recovered from this transformation by suitable values of the angles as listed below

F W $\eta = \frac{\pi}{2}, \ \xi = 0$

G T $\eta = \frac{\pi}{2}, \ \xi = \frac{\pi}{2}$

de Vries $\eta = \frac{\pi}{2}$

New $\eta = 0$

with S being given by eq. 3.18.

The case when $\eta = 0$ is especially interesting. Here the Hamiltonian becomes

$$H' = \alpha_5 E \tag{4.7}$$

which is analogous to F.W. Hamiltonian with the difference that while β is an even operator, α_5 is an odd one.

5. Two-Component Equations for a Massive Spin ½ Particle:

In writing a linear relativistic wave equation for the electron, one begins with the Klein-Gordon equation

$$(p^\mu p_\mu - m^2)\psi(x) = 0 \tag{5.1}$$

and seeks to replace this by equations linear in all components of p. Algebraically the problem is to factorise a five term quadratic of the form

$$A^2 + B^2 + C^2 + D^2 + E^2 = 0 \tag{5.2}$$

This cannot be done over the field of complex numbers but can readily be done by using the elements of an appropriate Clifford algebra. The representation theory of Clifford algebra of any dimension and of the generalised Clifford algebra which arises in the equivalent problem of linearisation of the sum of n^{th} powers

$$A = A_1^n + \cdots + A_m^n \tag{5.3}$$

have been systematically studied by Alladi Ramakrishnan in his work on L-matrix theory[7]. The most familiar example of a Clifford algebra, the Pauli matrices, allows are to factorise at most a four term quadratic. To factorise the Klein-Gordon equation, therefore, one has to go to the next higher Clifford algebra. Thus, Dirac obtains

$$(\underline{\alpha}\cdot\underline{p} + \beta m)\psi = E\psi \tag{5.4}$$

where

$$\{\alpha_\mu, \alpha_\nu\} = 2\delta_{\mu\nu}; \quad \mu,\nu = 1,2,3,4. \tag{5.5}$$

and α_μ may be chosen as

$$\alpha_i = \begin{bmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{bmatrix}, \quad \alpha_4 = \beta = \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \tag{5.6}$$
$$i = 1,2,3.$$

As shown by Pauli, the choice of α_μ is unique upto a similarity transformation.

However, as pointed out by Biedenharn et al[9], the uniqueness of the factorisation hinges on a fundamental assumption, that the matrices entering the factorisation are to be independent of space-time, that is, they are to represent independent new degrees of freedom. If, however, this condition is dropped, Dirac's factorisation is no longer unique and Dirac's classic derivation of his equation for a massive charged spin-½ particle admits of a alternative which leads to a distinct new equation arising from an alternative factorisation of the

Klein-Gordon equation. It is possible to write even 2-component wave equations for the electron. This is termed by Biedenharn as the 'Dirac dichotomy' ie. 1) either the factorising matrices are independent of space-time, leading to Dirac's factorisation or 2) the factorising matrices are not independent of space-time.

The second alternative gives rise to the possibility of constructing a 2-component equation for the electron. For now the problem of linearising the equation

$$p^2 + m^2 = E^2, \quad p^2 + m^2 - E^2 = 0$$

may be treated as a problem of factorising a quadratic equation with only three terms. This can be done using only 2 x 2 matrices. Thus a two component equation becomes possible. Biedenharn, Han and Van Dam give one such using what they term as the spherical factorisation of the Klein-Gordon equation. Their equation can be written as

$$H_B \varphi \equiv (\underline{\sigma} \cdot \underline{p} + \eta_3 m)\varphi = E\varphi \tag{5.7}$$

where $\underline{\sigma}$ are the Pauli matrices and

$$\eta_3 = (-1)^{j(K) + \frac{1}{2}} \frac{K}{|K|}, \quad K = -(\underline{\sigma} \cdot L + 1)$$

They also show that the equation is Poincaré invariant. However, in some problems, as, for instance, when the particle is in an external Coulomb field, since a particular point and a particular Lorentz frame and singled out, Lorentz invariance is not particularly relevant. They show that applied to Coulomb potential, the equation leads to exactly the Dirac Coulomb energies. They find that particles represented by their equation possess a new dichotomic quantum number, stigma, given by the eigenvalues of

$$\varsigma \equiv m^{-1}\eta_3 \Pi_0 + \rho_1 \underline{\sigma} \cdot \underline{\Pi} \; ; \; \varsigma^2 = 1 \tag{5.8}$$

Subsequently, Good[10] showed that the stigma quantum number coincides with the relativistic parity and Biedenharn's solution of their equation amounts to solving the Dirac Coulomb problem for a definite relativistic parity.

The methods of section 3 can be used here too to find the equivalent forms of the Hamiltonian of this 2-component equation and the transformations relating them to one another. Thus, let H_B and H' given by

$$H_B = \underline{\sigma} \cdot \underline{p} + \eta_3 m \tag{5.9}$$

and

$$H' = (\underline{\sigma} \cdot \hat{\underline{p}} \cos\theta + \eta_3 \sin\theta \cos\varphi + i \underline{\sigma} \cdot \hat{\underline{p}} \eta_3 \sin\theta \sin\varphi) E \tag{5.10}$$

This form of H' is the most general linear combination of anticommuting 2×2 matrices in the spherical factorisation with eigenvalues $\pm E$. Thus H_B and H' have the same elementary divisors and are related by a similarity transformation. To find the transformation are observe

$$H_B^2 = H'^2 = E^2 \tag{5.11}$$

and

$$H_B(H_B+H') = (H_B+H')H' \tag{5.12}$$

Further

$$(H_B+H')^2 = 2E^2 + H_BH' + H'H_B = \text{a scalar} \tag{5.13}$$

and hence $H_B + H'$ is non-singular. Thus the similarity transformation connecting H_B and H' may be written as

$$H' = U^{-1}H_BU \tag{5.14}$$

where

$$U = U^{-1} = \frac{H_B+H'}{(2E^2+H_BH'+H'H_B)^{\frac{1}{2}}} \tag{5.15}$$

U is Hermitian, involutionary and hence also unitary. We also know that this transformation is arbitrary upto postmultiplication of U by a non-singular matrix X which commutes with H' but is otherwise arbitrary. However, in this case, since H' has distinct eigenvalues (+ E and -E), X can at best be a function of H' itself. If $X = H'$,

$$U = \frac{1+\frac{H_BH'}{E^2}}{\left[2+\frac{1}{E^2}(H_BH'+H'H_B)\right]^{\frac{1}{2}}} \tag{5.16}$$

The non-relativistic and extreme relativistic limits of H_B follow directly from this, the former being obtained with $H' = \eta_3 E$ and the latter when $H' = \underline{\sigma} \cdot \hat{\underline{p}} E$.

We shall now show how the Biedenharn Hamiltonian H_B in its a 4-dimensional form that is repeated twice along the diagonal can be obtained from the Dirac Hamiltonian H_D by a (unitary) similarity transformation. The method of section 2 can be used to explicitly derive the unitary transformation ~~commuting~~ connecting the non-covariant Hamiltonian of Biedenharn to the Dirac Hamiltonian, since, both of them have the same elementary divisors and both can be diagonalised by unitary transformations.

First we observe that the Necessary and sufficient condition[8], for the two Hamiltonians to be connected by a similarly transformation is that they have the same elementary divisors. For, if the elementary divisors are the same, the matrices will have the same Jordan form, say J. Thus,

$$S^{-1} H S = J , \tag{5.17}$$

$$S'^{-1} H' S' = J . \tag{5.18}$$

Therefore

$$H' = (S S'^{-1})^{-1} H (S S'^{-1}) \tag{5.19}$$

and hence, H' is similar to H.

Now we note that H_B is a hermitian operator and hence can be diagonalised by a unitary matrix. Again we observe, since

$$H_B^2 = (\underline{\sigma}\cdot\hat{p}E)^2 = (\sigma_3 E)^2 = E^2 \tag{5.20}$$

we have

$$\frac{H_B(H_B + \underline{\sigma}\cdot\hat{p}E)(\underline{\sigma}\cdot\hat{p} + \sigma_3)}{[2E(E+p)\,2(1+p_3/p)]^{1/2}}$$

$$= \frac{(H_B + \underline{\sigma}\cdot\hat{p}E)(\underline{\sigma}\cdot\hat{p} + E)\sigma_3 E}{[2E(E+p)\,2(1+p_3/p)]^{1/2}} \tag{5.21}$$

Thus

$$U_B^{-1} H_B U_B = \sigma_3 E$$

where $\sigma_3 E$ is the diagonal form of H_B and U the unitary matrix that diagonalises it being given by

$$U_B = \frac{(H_B + \underline{\sigma}\cdot\hat{p}E)(\underline{\sigma}\cdot\hat{p} + \sigma_3)}{[2E(E+p)\,2(1+p_3/p)]^{1/2}} \tag{5.22}$$

$$U_B^{-1} = U_B^{+} \tag{5.23}$$

The same expressions are valid when H_B and $\underline{\sigma}$ are treated as 4-dimensional, obtained by repeating the 2-dimensional matrices twice along the diagonal. Now coming to the Dirac Hamiltonian H_D, as shown earlier

$$\frac{H_D(H_D+\beta E)}{[2E(E+m)]^{1/2}} = \frac{(H_D+\beta E)}{[2E(E+m)]^{1/2}}\beta E \tag{5.24}$$

and therefore

$$U_D^{-1} H_D U_D = \beta E \tag{5.25}$$

where

$$U^{-1} = U^{+} = U = \frac{H_D+\beta E}{[2E(E+m)]^{1/2}} \tag{5.26}$$

Further, β and σ_3 are themselves related by a permutation matrix. Thus

$$\beta = P^{-1}\sigma_3 P \tag{5.27}$$

where

$$P = P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.28}$$

combining eqs. (5.21a), (5.25), (5.28)

$$U^{-1} H_B U = H_D \tag{5.29}$$

where U is given by

$$U = U_B^{-1} P U_D \tag{5.30}$$

which, being a product of unitary matrices, is unitary.

An alternative way of writing this transformation is as follows: We write

$$H_B(H_B+H_D)(2E^2+H_BH_D+H_DH_B)^{-1/2}$$

$$=(H_B+H_D)(2E^2+H_BH_D+H_DH_B)^{-1/2}H_D \tag{5.31}$$

where the term $(2E^2+H_BH_D+H_DH_B)^{-1/2}$ here is not a scalar. However, this expression can be expanded binomially and the series coverages for the spectral radius of $H_DH_B+H_BH_D<2E^2$. This follows from the fact that the matrix $H_B\,H_D$ (as also $H_D\,H_B$) is unitary except for normalisation and hence all its eigenvalues have absolute value E^2. Further the eigenvalues of $H_BH_D+H_DH_B$ are twice the real part of the eigenvalues of $H_B\,H_D$ and hence $\leqslant 2E^2$. The equality holds only when $H_B = \pm H_D$ which is not the case here. Thus the spectral radius of $H_B\,H_D + H_D\,H_B$ is less than $2\,E^2$ and hence the expression can be expanded binomially. In particular, it has no eigenvalue as $-2\,E^2$ and hence the denominator in eq. (5.31) is non-singular. With the further observation that $2E^2+H_BH_D+H_DH_B$ commutes with $H_B + H_D$, we find that eq. (5.31) is true and hence

$$H_D=U^{-1}H_BU \tag{5.32}$$

where

$$U^{\dagger}=U^{-1}=U=(H_B+H_D)(2E^2+H_BH_D+H_DH_B)^{-1/2} \tag{5.33}$$

However, U here is the sum of an infinite series.

6. A New Factorisation of the Klein-Gordon Equation:

We shall now give a new two-component wave equation for a massive spin-half particle arising from a different factorisation of the Klein-Gordon equation[12] distinct from the spherical factorisation of Biedenharn. Thus the problem is to linearise

$$p^2 + m^2 = E^2 \tag{6.1}$$

to yield an equation

$$H\varphi \equiv (\underline{\sigma} \cdot \underline{p} + m\underline{B})\varphi(x) = E\varphi \tag{6.2}$$

As discussed earlier in this chapter, we shall allow the matrices to be dependent on the space-time coordinates. Then we can treat the above as a problem of linearising a quadratic equation of two terms and look for a solution involving 2 x 2 matrices. To find the most general form of B, we make the following transformation on the Pauli matrices satisfying the Clifford algebra

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij}\ ; \quad i,j = 1,2,3. \tag{6.3}$$

we choose the σ's as

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{6.4}$$

where, we note

$$\sigma_3 = i\sigma_1\sigma_2$$

We now construct the matrices given by

$$B_i = \Lambda_{ij}\sigma_j \tag{6.5}$$

with

$$\Lambda_{ik}\Lambda_{jk} = \delta_{ij} \tag{6.6}$$

Then, by an analogue of Pauli's fundamental theorem, the B_i constitute a set of ~~these~~ three mutually anticommuting matrices with the square of each being I, the two-dimensional unit matrix.

We choose Λ_{ij} as the matrix

$$\Lambda_{ij} = \begin{bmatrix} fp_1 & fp_2 & fp_3 \\ gp_2 & -gp_1 & 0 \\ hp_3p_1 & hp_3p_2 & -h(p_1^2+p_2^2) \end{bmatrix} \tag{6.7}$$

where

$$f = \frac{1}{(p_1^2+p_2^2+p_3^2)^{1/2}} = \frac{1}{p}$$

$$g = \frac{1}{(p_1^2+p_2^2)^{1/2}} \tag{6.8}$$

$$h = \frac{1}{p(p_1^2+p_2^2)^{1/2}} = \frac{g}{p} = fg$$

Hence the set B_1 becomes

$$B_1 = \underline{\sigma} \cdot \hat{p}$$

$$B_2 = \frac{p_2\sigma_1 - p_1\sigma_2}{(p_1^2+p_2^2)^{1/2}}$$

$$B_3 = \frac{p_3p_1\sigma_1 + p_3p_2\sigma_2 - (p_1^2+p_2^2)\sigma_3}{p(p_1^2+p_2^2)^{1/2}}$$

$$= \frac{1}{(p_1^2+p_2^2)}\left[(p_3\underline{\sigma}\cdot\hat{p} - p\sigma_3)\right] \tag{6.9}$$

Also

$$\{B_i, B_j\} = 2\delta_{ij}$$

and

$$B_3 = iB_1B_2 = i\underline{\sigma}\cdot\hat{p}\, B_2 \tag{6.10}$$

This construction can be slightly simplified by observing

$$(\underline{\sigma}\cdot\underline{p})(\underline{\sigma}\cdot\underline{k}) = \underline{p}\cdot\underline{k} + i\underline{\sigma}\cdot(\underline{p}\times\underline{k}) \tag{6.11}$$

Since $\underline{p} \times \underline{k} = -\underline{k} \times \underline{p}$ when the elements of $\underline{p}$ and $\underline{k}$ commute, we conclude that $\underline{\sigma}\cdot\underline{k}$ anticommutes with $\underline{\sigma}\cdot\underline{p}$ if $\underline{k}$ is a vector orthogonal to $\underline{p}$. And since $\underline{k}$ is a vector in a 3-dimensional space, there are two such linearly independent vectors which may be chosen as above.

Thus a two-component equation for a massive spin-half particle may be written as

$$H\varphi \equiv (\underline{\sigma}\cdot\underline{p} + B_2 m)\varphi = i\frac{\partial\varphi}{\partial t} \tag{6.12}$$

where

$$B_2 = \frac{p_2\sigma_1 - p_1\sigma_2}{(p_1^2+p_2^2)^{1/2}} \tag{6.13}$$

We notice that the Hamiltonian here, unlike the Biedenharn Hamiltonian, commutes with $\underline{p}$.

The eigenstates vectors of this Hamiltonian (or the transformation which diagonalises this Hamiltonian) can again be easily found using the U-matrix method. Since

$$H^2 = (\sigma_3 E)^2 = E^2 \tag{6.14}$$

we have

$$H(H+\sigma_3 E) = (H+\sigma_3 E)\,\sigma_3 E \tag{6.15}$$

and

$$(H+\sigma_3 E)^2 = 2E(E+p) \tag{6.16}$$

from which it follows that $(H+\sigma_3 E)$ is non-singular and the transformation relating H to its diagonal form $\sigma_3 E$ is given by

$$U^{-1} H U = \sigma_3 E \tag{6.17}$$

where

$$U^{-1} = U^{\dagger} = U = \frac{H+\sigma_3 E}{[2E(E+p)]^{1/2}} \tag{6.18}$$

Thus the columns of U are the eigenvectors of H.

Similarly, the transformation relating H to its non-relativistic limit

$$H_{nr} = \beta_2 E \tag{6.19}$$

is given by

$$U_{nr}^{-1} H U_{nr} = H_{nr} \tag{6.20}$$

where

$$U_{nr} = \frac{E + m + \sigma \cdot p \beta_2}{[2E(E+m)]^{1/2}} \tag{6.21}$$

with

$$U^{-1} = U^{\dagger} \tag{6.22}$$

The transformation relating H to its extreme relativistic limit

$$H_{er} = \underline{\sigma} \cdot \hat{p} E \tag{6.23}$$

is given by

$$U_{er}^{-1} H U_{er} = H_{er} \tag{6.24}$$

where

$$U = \frac{E + p + m \beta_2 \underline{\sigma} \cdot \hat{p}}{[2E(E+p)]^{1/2}} \tag{6.25}$$

Again, the most general form of the equivalent Hamiltonian is given by

$$H' = (\underline{\sigma} \cdot \hat{p} \cos\theta + \beta_2 \sin\theta \cos\varphi + i \underline{\sigma} \cdot \hat{p} \beta_2 \sin\theta \sin\varphi) E \tag{6.26}$$

This form of H' is the most general linear combination of the anticommuting set given by eq. (6.9) with eigenvalues $\pm E$. H and H' have the same elementary divisors and the transformation connecting H to H' is given by

$$H' = U^{-1} H U \tag{6.27}$$

where

$$U = \left(1 + \frac{HH'}{E^2}\right)\left[2 + \frac{1}{E^2}(HH' + H'H)\right]^{-\frac{1}{2}} \tag{6.28}$$

with

$$U^{-1} = U^{+} \tag{6.29}$$

The non-relativistic and extreme relativistic limits of H follow directly from this equation where $H' = B_2 E$ and $H' = \underline{\sigma} \cdot \hat{\underline{p}} E$ respectively.

In chapter III we constructed the most general linear combination of linearly independent matrices which anticommute with $\underline{\alpha} \cdot \underline{p}$ in the spherical factorisation as well as our non-covariant factorisation[11,12]. The equation obtained there can be extended to a massive particle. Let the massive spin-half particle be described by the wave equation

$$H\psi \equiv (\underline{\alpha} \cdot \underline{p} + mA)\psi = \frac{i\partial\psi}{\partial t} \tag{6.30}$$

where m is here the rest mass of the particle. The most general matrix A which anticommutes with $\underline{\alpha} \cdot \underline{p}$ and whose square is a scalar was found to be of the form

$$A = (a_0 + a_1 A_1)\left(\sum_{i=2}^{5} a_i A_i\right) \tag{6.31}$$

If the wave function is to satisfy the Klein-Gordon equation with rest mass m , we get the condition $A^2 = 1$ or

$$(a_0^2 - a_1^2)\left(\sum_{i=2}^{5} a_i^2\right) = 1 \tag{6.32}$$

We get different equations according as we choose the set A_i as given by eq.(3.8) or eq.(3.25) of chapter III. In this case, however, the Hamiltonian is not necessarily non-hermitian. In any case, whether H is hermitian or otherwise, it is diagonalised by

$$U^{-1} H U = \beta E \tag{6.33}$$

where

$$U = \frac{(H + \underline{\alpha}\cdot\vec{p}\, E)\,(\underline{\alpha}\cdot\hat{p} + \beta)}{[2E(E+p)\cdot 2]^{1/2}} \tag{6.34}$$

When the Hamiltonian is hermitian, this transformation is unitary.

7. Conclusions:

We have given a simple and elegant method of obtaining explicit forms of the Foldy-Wouthuysen transformation and its generalisations using the U-matrix method of Alladi Ramakrishnan and shown that the transformation is not unique. We give a new generalisation of the FW transformation using the five 4 x 4 matrices satisfying the Clifford algebra. We have also shown how this method can be used to obtain the transformation which connects the Dirac equation to the non-covariant forms of the Dirac equation. We then give a two component equation for a massive spin-half particle using a new non-covariant factorisation of the Klein-Gordon equation (distinct from the spherical factorisation advocated by Biedenharn and others).

<u>References:</u>

1. L.L.Foldy and S.A.Wouthuysen, Phys.Rev., <u>78</u>, 29 (1950).
2. M.Cini and B.Touschek, Nuov.Cim., <u>7</u>, 422 (1958).
3. E.Eriksen, Phys.Rev., <u>111</u>, 1011 (1958).
4. M.Bhaktavatsalou, Nuov.Cim., <u>25</u>, 964 (1962).
5. I.Saavedra, Nucl.Phys., <u>B2</u>, 690 (1967).
6. E.De Vries, Physica, <u>43</u>, 45 (1969).
7. Alladi Ramakrishnan, "L-Matrix Theory or the Grammar of Dirac Matrices", Tata McGraw-Hill, Bombay (1972).
8. A.R.Tekumalla, Communicated to Physica.
9. L.C.Biedenharn and M.Y.Han, Phys.Rev.Lett., <u>27</u>, 1167 (1971).
10. J.H.Wilkinson, "The Algebraic Eigenvalue Problem", Clarendon, Oxford (1965) p.12.
11. T.S.Santhanam and A.R.Tekumalla, Lett.Nuov.Cim., <u>3</u>, 190 (1971).
12. A.R.Tekumalla and T.S.Santhanam, Lett.Nuov.Cim., <u>6</u>, 99 (1973).

CHAPTER. V

UNITARY FOLDY WOUTHUYSEN TRANSFORMATIONS FOR PARTICLES OF ARBITRARY SPIN*

1. Introduction:

A few years ago Weaver Hammer and Good[1] gave a Hamiltonian formulation of the theory of a free massive particle with arbitrary spin. Their object was to give a description without auxiliary conditions on the wave function and also with simple Lorentz transformation properties. The technique they employed was essentially to start with the rest frame wave functions (which are eigen functions of the spin operator) and use a generalised Foldy-Wouthuysen (FW) transformation. One difficulty, as was realised earlier, is that except in the case of spin ½, the transformation is not unitary and consequently leads to a problem in constructing orthonormal wave functions in the laboratory system and in transforming various operators from the FW representation to the laboratory system. More recently, Weaver[2] succeeded in constructing a unitary spin 1 FW operator. Recently there has been some misconception in the literature[3] on the existence of a unitary FW transformation for spin greater than 1. We shall here first establish the existence of a unitary FW transformation for arbitrary spin and then explicitly calculate it. We show that this transformation is not unique and find the condition that reduces it to the FW transformation in the conventional form for spin-half and to Weaver's transformation

---

* A.R.Tekumalla and T.S.Santhanam, Lett.Nuov.Cim., to appear 1974

for spin one. The explicit calculation of the unitary FW transformation in this case involves the evaluation of matrix functions $f(z)$ where z is related to the spin matrices. We give methods for evaluating $f(z)$ in general and in particular, when $f(z)$ is on odd or even function of z. We then use these methods for also writing the extreme relativistic form of the Hamiltonian and the (unitary) transformation leading to it.

2. The Method of Weaver, Hammer and Good:

The technique employed by Weaver, Hammer and Good is as follows: the particle is described by a wave function which is the basis of a $(0,s) \oplus (s,0)$ representation of the Lorentz group. They set up the properties of the system in the rest frame and then make a Lorentz transformation to the laboratory frame. In the rest frame the Hamiltonian is identified as βm. They then exploit the connection between the Lorentz transformation and the FW transformation to obtain the wave function for a particle of mass m and spin s with energy E and momentum $\underline{p}$ in the laboratory frame using the generalised FW operator. The Hamiltonian in the FW representation is taken as $H_{FW} = \beta E$

The Hamiltonian in the laboratory system is therefore given by

$$H = S \beta E S^{-1} \tag{2.1}$$

where S is given by

$$S = \exp \epsilon z \theta \tag{2.2}$$

where

$$Z = S\underline{\alpha}\cdot p, \ \theta = \tanh^{-1}\frac{p}{E}, \ \epsilon = \pm 1 \tag{2.3}$$

$$\underline{\alpha} = \frac{1}{S}\begin{bmatrix} \underline{S} & 0 \\ 0 & -\underline{S} \end{bmatrix}, \quad \beta = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \tag{2.4}$$

and $\underline{S}$ are the 2 s + 1 dimensional spin matrices satisfying the usual angular momentum commutation relations, I being the 2-dimensional unit matrix. ϵ acting on in rest frame can be replaced by β and the operator S finally takes the explicit form

$$S = \cosh Z\theta - \beta \sinh Z\theta \tag{2.5}$$

and

$$S^{-1} = (\cosh Z\theta - \beta \sinh Z\theta) \operatorname{sech} 2Z\theta$$
$$= S^{+} \operatorname{sech} 2Z\theta = \operatorname{sech} 2Z\theta \, S^{+}$$

so that

$$H = \beta \operatorname{sech} 2Z\theta + \tanh 2Z\theta \tag{2.6}$$

The functions $\cosh Z\theta$, $\sinh Z\theta$ and H have been given in terms of the spin polynomials by Weber and Williams.

Some of the spin ½ properties still apply here. For instance, β anticommutes with $\underline{\alpha}$ and γ_5 where γ_5 is defined as

$$\gamma_5 = \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} \tag{2.7}$$

Also, $\underline{\alpha}$ commutes with γ_5 and $\gamma^2 = \beta^2 = 1$. Further, since sech $2z\theta$ involves only even powers of z , it commutes with β and hence with S and S^+ . Again, since

$$H^\dagger = \operatorname{sech} 2z\theta S\beta E S^\dagger = S\beta E S^\dagger \operatorname{sech} 2z\theta = H \tag{2.8}$$

it follows that the Hamiltonian is hermitian. However, the factor sech $2z\theta$ is not a scalar except in the case of spin $\frac{1}{2}$. Hence the operator S is not unitary except for spin $\frac{1}{2}$. The expression for H, S and S^{-1} for spin $\frac{1}{2}$, 1 and 3/2 are given in the table.

3. Existence of a Unitary FW Transformation for Arbitrary Spin:

We now prove the following:

THEOREM 1: For any spin s, there exists a unitary transformation $H = U\beta E U^\dagger$ and the unitary matrix U is given by $U = SX$ where S is given by eq. (2.5) and X is such that a) it commutes with β and b) satisfies the condition $X X^+ = \operatorname{sech} 2z\theta$

PROOF: Since the Hamiltonian given by

$$H = S\beta E S^{-1} \tag{3.1}$$

is Hermitian, it follows that there exists a unitary matrix which diagonalises it, that is

$$U_1 H U_1^\dagger = \gamma_5 E \tag{3.2}$$

Similarly, since βE is also Hermitian and has the same eigenvalues as H , we have

TABLE SHOWING H, S and S^{-1}

Spin	H	S	S^{-1}
1/2	$\alpha\cdot p + \beta m$	$\dfrac{E+m+\alpha\cdot p\beta}{[2m(E+m)]^{1/2}}$	$\dfrac{(E+m-\alpha\cdot p\beta)\,m^{1/2}}{E\,[2(E+m)]^{1/2}}$
1	$\dfrac{[(2E^2-m^2)\beta + 2E\alpha\cdot p - 2(\alpha\cdot p)^2\beta]E}{2E^2-m^2}$	$\dfrac{[m(E+m) + (E+m)\alpha\cdot p\beta + (\alpha\cdot p)^2]}{m(E+m)}$	$\dfrac{[(E+m)(2E^2-m^2) - m(E+m)\alpha\cdot p\beta - (2E+m)(\alpha\cdot p)^2]}{(E+m)(2E^2-m^2)}$
3/2	$\dfrac{[(9E^2-7m^2)m\beta + 2(13E^2-10m^2)\alpha\cdot p - 9m(\alpha\cdot p)^2\beta - 18(\alpha\cdot p)^3]}{2(4E^2-3m^2)}$	$\dfrac{[-(E+m)(E-5m) - (E-13m)\alpha\cdot p\beta + 9(\alpha\cdot p)^2 + \frac{9(\alpha\cdot p)^3\beta}{E+m}]}{[32m^3(E+m)]^{1/2}}$	$\dfrac{[(E+m)(18E^2-mE-13m^2) - (54E^2-Em-41m^2)\alpha\cdot p\beta - 9(2E+m)(\alpha\cdot p)^2 + 9\frac{(6E+5m)}{E+m}(\alpha\cdot p)^3\beta]\,m^{1/2}}{[32(E+m)]^{1/2}\,E(4E^2-3m^2)}$

The expressions for U are given in the text, Pg. 150, 162-165.

$$U_2^{+} \beta E U_2 = \gamma_5 E \tag{3.3}$$

where U_2 is again unitary. Hence

$$H = U \beta E U^{+} \tag{3.4}$$

where

$$U = U_1^{+} U_2^{+} \tag{3.5}$$

which, since the product of two unitary matrices is also unitary, proves the existence of the required unitary transformation U.

We now proceed to obtain an explicit form for U. From eq. (3.1) it is clear that S is not unique but is arbitrary upto post-multiplication by a matrix X which is non-singular and which commutes with β but is otherwise arbitrary. For, if X commutes with β,

$$SX \beta E (SX)^{-1} = S \beta E S^{-1} = H \tag{3.6}$$

which shows that SX also transforms βE to H by a similarity transformation. We now choose X such that

$$U = SX \tag{3.7}$$

is unitary. Therefore

$$SX (SX)^{+} = S X X^{+} S^{+} = 1$$

whence

$$XX^{+} = S^{-1} S^{+-1} = (S^{+} S)^{-1} \tag{3.8}$$

$$\text{or } XX^{+} = \operatorname{sech} 2Z\theta, \tag{3.9}$$

which is the required condition. We mention here in passing that

$$\operatorname{sech} 2z\theta = \frac{1}{2E}\{H,\beta\}$$

An alternative formulation of Theorem 1 is possible by observing that if Y is a non-singular matrix that commutes with H,

$$YS\,\beta E(YS)^{-1} = YHY^{-1} = H \tag{3.10}$$

Thus S is arbitrary upto pre-multiplication by a matrix Y which commutes with H. Similarly we can show that S^{-1} is arbitrary upto pre-multiplication by a matrix X which commutes with β or post-multiplication by a matrix Y which commutes with H.

We now consider the solution of eq.(3.9). We observe that the left hand side of eq.(3.9) is necessarily hermitian and since the right hand side is also hermitian, solutions to this equation exist. This is at once obvious when we consider the diagonal form of sech $2z\theta$. For, let the matrix sech $2z\theta$ be diagonalised by an operator O so that

$$O^{-1}(\operatorname{sech} 2z\theta)\,O = D \tag{3.11}$$

where D is the diagonal form of sech $2z\theta$ so that it can be written as

$$D_{ij} = d_i\,\delta_{ij} \tag{3.12}$$

where d_i are all real. Now D can be factorised in an infinite number of ways into two matrices $D_1 D_2$ such that $D_1 = D_2^{+} = D_2^{*}$. Thus D, can be written as

$$(D_1)_{kj} = (d_k)^{\frac{1}{2}} e^{iq_k} \tag{3.13}$$

Then the matrix

$$X = O^{-1} D_1 O \tag{3.14}$$

satisfies eq.(3.9). We note here that since O is unitary and D_1 diagonal, X is a normal matrix. It is obvious that X is not unique. We shall presently consider the most general solution for X in terms of α, β .

We may mention here that the problem of finding U may be solved (though more laboriously) even at the stage of eq.(3.2) by finding U_1 and then U_2 . First we note that SU_2 diagonalises H though not unitarily. Now the eigenvectors belonging to $+E$ and $-E$ can be separately orthogonalised. Further H being Hermitian any eigenvector belonging to $+E$ is orthogonal to any eigenvector belonging to $-E$. Therefore the resulting set is completely orthogonalised and gives a unitary U_1 . However, this method is too laborious.

4. Explicit Form of the Unitary FW Transformation:

In general it is not necessary to diagonalise the operator sech $2s\theta$. We shall first find a particular solution of equation (3.9) for any spin s and then give the most general solution for X and hence for U . We shall then show that the choice $X = X^+$ is of special significance as this choice yields the FW operator in the conventional form as obtained by Foldy and Wouthuysen for spin ½ and Weaver for spin 1.

THEOREM 2 : A possible form of X true for any spin is given by

$$X = (\cosh Z\theta + \alpha_5 \sinh Z\theta)\,\mathrm{sech}\,2Z\theta \tag{4.1}$$

where

$$\alpha_5 = i\beta\gamma_5 = \begin{bmatrix} 0 & -iI \\ iI & 0 \end{bmatrix} \tag{4.2}$$

Proof: Since z as well as α_5 anticommute with β, it follows that X commutes with β. Thus, this choice of X satisfies condition (a) of Theorem 1. Again, with this choice of X, since $\alpha_5^2 = 1$

$$XX^{\dagger} = (\cosh Z\theta + \alpha_5 \sinh Z\theta)\cdot(\cosh Z\theta - \alpha_5 \sinh Z\theta)(\mathrm{sech}\,2Z\theta)^2$$

$$= (\cosh 2Z\theta)(\mathrm{sech}\,2Z\theta)^2$$

Therefore

$$XX^{\dagger} = \mathrm{sech}\,2Z\theta \tag{4.3}$$

and X also satisfies condition (b) of Theorem 1 which proves Theorem 2. Thus a possible form of the unitary FW transformation for any spin is given by

$$U = (\cosh Z\theta - \beta \sinh Z\theta)\cdot(\cosh Z\theta + \alpha_5 \sinh Z\theta)\,\mathrm{sech}\,2Z\theta \tag{4.4}$$

We shall now consider the most general solution of the equation $U\beta E U^{\dagger} = H$. From eq.(3.6) we found that the operator S which transforms βE to H is arbitrary upto post-multiplication by a non-singular matrix which commutes with β but is otherwise arbitrary. We shall first find the most general matrix which commutes with β. Thus we desire the most general solution of

$$X\beta = \beta X \tag{4.5}$$

Since β is hermitian, it is diagonalised by a unitary matrix

$$T^{-1}\beta T = -\gamma_5 \tag{4.6}$$

where

$$T = T^{\dagger} = T^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{4.7}$$

Therefore

$$\beta = T(-\gamma_5)T \tag{4.8}$$

and eq. (4.5) can be rewritten as

$$XT(-\gamma_5)T = T(-\gamma_5)TX$$

or

$$(TXT)\gamma_5 = \gamma_5(TXT)$$

or

$$X'\gamma_5 = \gamma_5 X' \tag{4.9}$$

where

$$X' = T X T \tag{4.10}$$

is the most general matrix which commutes with γ_5. X' obviously has the form

$$X' = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \tag{4.11}$$

where X_1 and X_2 are any arbitrary square matrices each of dimension $2s + 1$. Further X' is nonsingular if and only if X_1 and X_2 are non-singular. Therefore X is given by

$$X = T X' T$$

or

$$X = \frac{1}{2}\begin{bmatrix} X_1 + X_2 & X_1 - X_2 \\ X_1 - X_2 & X_1 + X_2 \end{bmatrix} \tag{4.12}$$

Thus the most general matrix which transforms βE to H is S X where X is given by eq.(4.12). If the transformation is to be unitary, X must further satisfy the condition $X X^+ =$ sech $2 z \theta$. Alternatively, if we find one transformation U which is already unitary and which transforms βE to H (as in eq. (4.4)), we can write the most general unitary transformation U_G which takes βE to H as

$$U_G = U X . \tag{4.13}$$

where X is given by eq.(4.12) with X_1 and X_2 both unitary.

The solution of eq.(3.9) with X hermitian seems to be of special interest. With the additional condition $X = X^+$, the transformation $U = S X$ leads, (as will be shown in the examples later) in the case of spin $\frac{1}{2}$ to the FW transformation in the conventional form as obtained by Foldy and Wouthuysen while in the case of spin 1 it leads to the unitary transformation obtained by Weaver. We shall give here the solutions for spin 3/2 also ~~and spin 2 ... half-integral and~~

Set $X = X^+$ in eq. (3.9). Then

$$X^2 = \operatorname{sech} 2 z \theta \tag{4.14}$$

or

$$X = (\operatorname{sech} 2 z \theta)^{\frac{1}{2}} \tag{4.15}$$

The quantity on the right is multivalued since each eigenvalue can take positive or negative values. However we shall consider the solution in terms of a convenient expansion in even powers of z, since we desire X to commute with β. However, eq.(4.15) as it stands, is not useful as the operator z occurs in the argument of the function while we desire the solution as a polynomial in z of finite degree. This polynomial will of course be of degree less than 2 s + 1 as z obeys a characteristic equation of degree 2 s + 1 . We shall therefore now give a method of obtaining any function of z , (in particular, sech 2 z θ) as a polynomial in z .

5. Evaluation of Functions of z :

Since S are the spin matrices, S.p satisfies a characteristic equation of degree 2 s + 1 . Thus

$$(\underset{\sim}{S}\cdot p - s)(\underset{\sim}{S}\cdot p - s+1)\cdots(\underset{\sim}{S}\cdot p + s) = 0, \tag{5.1}$$

that is

$$[(\underset{\sim}{S}\cdot p)^2 - s^2][(\underset{\sim}{S}\cdot p)^2 - (s-1)^2]\cdots[(\underset{\sim}{S}\cdot p)^2 - \tfrac{1}{4}] = 0 \tag{5.2a}$$

when s is half-integral and

$$[(\underset{\sim}{S}\cdot p)^2 - s^2][(\underset{\sim}{S}\cdot p)^2 - s^2]\cdots[\underset{\sim}{S}\cdot p] = 0 \tag{5.2b}$$

when s is integral, Since

$$z = -\gamma_5 \begin{bmatrix} \underset{\sim}{S}\cdot p & 0 \\ 0 & \underset{\sim}{S}\cdot p \end{bmatrix} \tag{5.3}$$

and

$$z^2 = \begin{bmatrix} (\underset{\sim}{S}\cdot p)^2 & 0 \\ 0 & (\underset{\sim}{S}\cdot p)^2 \end{bmatrix} \tag{5.4}$$

the minimal equation of z is given by

$$(z^2 - s^2)[z^2 - (s-1)^2]\cdots[z^2 - \tfrac{1}{4}] \tag{5.5a}$$

if s is half-integral and

$$[z^2 - s^2][z^2 - (s-1)^2]\cdots[z] = 0 \tag{5.5b}$$

if s is integral, the last equations being obtained by multiplying eq. (5.2b) by the nonsingular matrix $-\gamma_5$. Thus z satisfies a minimal equation of degree $2s + 1$ and hence any function $f(z)$ that is given as a polynomial in z can be expressed as a polynomial of degree at most 2s. Now let the form of $f(z)$ be given by

$$f(z) = a_1 z^{2s} + a_2 z^{2s-1} + \cdots + a_{2s+1} \qquad (5.6)$$

where the coefficients a_i are to be determined. Since $f(z)$ satisfies eq.(5.6), all the eigenvalues of z also satisfy eq. (5.6). And since there are $2s + 1$ distinct eigenvalues we get a system of $2s + 1$ linearly independent equations. Thus, if we denote the eigenvalues by $\lambda_1, \lambda_2, \ldots, \lambda_{2s+1}$, we get

$$\begin{aligned} f(\lambda_1) &= a_1\lambda_1^{2s} + a_2\lambda_1^{2s-1} + \cdots + a_{2s+1} \\ f(\lambda_2) &= a_1\lambda_2^{2s} + \cdots \cdots + a_{2s+1} \\ &\cdots\cdots\cdots\cdots \\ f(\lambda_{2s+1}) &= a_1\lambda_{2s+1}^{2s} + \cdots\cdots + a_{2s+1} \end{aligned} \qquad (5.7)$$

or more concisely

$$\underline{f(\lambda)} = V\underline{a} \qquad (5.8)$$

where $\underline{f(\lambda)}$ and $\underline{a}$ are the vectors

$$\underline{f(\lambda)} = \begin{bmatrix} f(\lambda_1) \\ f(\lambda_2) \\ \vdots \\ f(\lambda_{2s+1}) \end{bmatrix}, \quad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{2s+1} \end{bmatrix} \qquad (5.9)$$

and V is the Vandermonde matrix

$$V = \begin{bmatrix} \lambda_1^{2s} & \lambda_1^{2s-1} & \dots & 1 \\ \lambda_2^{2s} & \lambda_2^{2s-1} & \dots & 1 \\ \vdots & & & \\ \lambda_{2s+1}^{2s} & \lambda_{2s+1}^{2s-1} & \dots & 1 \end{bmatrix} \tag{5.10}$$

As the Vandermonde is non-singular, the solution of eq.(5.8) can be written as

$$\underline{a} = V^{-1}\underline{f(\lambda)} \tag{5.11}$$

and $f(z)$ is given by eq.(5.6) or formally

$$f(Z) = \underline{Z}^T\underline{a} = \underline{Z}^T V^{-1}\underline{f(\lambda)} \tag{5.12}$$

where $\underline{Z}^T$ is the row vector

$$Z^T = [Z^{2s}, Z^{2s-1}, \dots, Z^0] \tag{5.13}$$

and the elements of the vector are matrices.

The constants A_i can also be separately written in closed form by suitably interpreting eq.(5.11) as will be shown in Chapter VI. Thus,

$$a_i = \det \begin{bmatrix} \lambda_1^{2s} & \lambda_2^{2s} & \cdots & \lambda_{2s+1}^{2s} \\ \lambda_1^{2s-1} & \lambda_2^{2s-1} & \cdots & \lambda_{2s+1}^{2s-1} \\ \vdots & & & \\ \lambda_1^{2s-i+1} & \cdots & \cdots & \cdots \\ f(\lambda_1) & f(\lambda_2) & \cdots & f(\lambda_{2s+1}) \\ \lambda_1^{2s-i-1} & \cdots & \cdots & \cdots \\ \vdots & & & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \div \det V \tag{5.14}$$

where det. V , the determinant of the Vandermonde matrix is given by

$$\det V = \prod_{\substack{i,j=1 \\ i<j}}^{2s+1} (\lambda_i - \lambda_j) \tag{5.15}$$

When the function is odd or even, the evaluation is much simpler. Thus, for an even function of z (as for instance, $\cosh z\theta$ or $\operatorname{sech} z\theta$), we have

$$f(z) = \sum_{i=1}^{2s+1} a_i z^{2s+1-i}$$

$$f(-z) = \sum a_i (-1)^{2s+1-i} z^{2s+1-i} \tag{5.16}$$

with

$$f(z) = f(-z) \tag{5.17}$$

so that

$$\sum a_i \left[1-(-1)^{2s+1-i}\right] z^{2s+1-i} = 0 \tag{5.18}$$

Since the z^k are all linearly independent,

$a_i = o$ for i odd if s is half integer

$a_i = o$ for i even if s is integer (5.19)

Thus $f(z)$ is given by

$$f(z) = b_1 z^{2\{s\}} + b_2 z^{2\{s\}-2} + \cdots + b_{2\{s\}+1} \tag{5.20}$$

where $\{s\}$ denotes the integral part of s. The set of independent equations corresponding to eq.(5.7) is reduced by half and we get only $\{s\}+1$ equations. Thus

$$f(\lambda_1) = b_1 \lambda_1^{2\{s\}} + b_2 \lambda_1^{2\{s\}-2} + \cdots + b_{2\{s\}+1}$$

$$\vdots$$

$$f(\lambda_{\{s\}+1}) = b_1 \lambda_{\{s\}+1}^{2\{s\}} + \cdots + b_{2\{s\}+1} \tag{5.21}$$

where the λ_i now run only over the positive and o values, the last equation being for $f(\frac{1}{2})$ or $f(o)$, according as s is half-integral or integral. As before, we write

$$\underline{f(\lambda)} = V\underline{b} \tag{5.22}$$

where V is now given by

$$V = \begin{bmatrix} \lambda_1^{2\{s\}} & \lambda_1^{2\{s\}-2} & \cdots & 1 \\ \lambda_2^{2\{s\}} & \lambda_2^{2\{s\}-2} & \cdots & 1 \\ \vdots & & & \\ \lambda_{\{s\}+1}^{2\{s\}} & \cdots & \cdots & 1 \end{bmatrix} \tag{5.23}$$

(which is the Vandermonde matrix of $\lambda_1^2, \lambda_2^2, \ldots, \lambda^2_{\{s\}+1}$ **)**

so that

$$\underline{b} = V^{-1} f(\lambda_1) \tag{5.24}$$

The elements of b are given by

$$b_i = \det \begin{bmatrix} (\lambda_1^2)^{\{s\}} & (\lambda_2^2)^{\{s\}} & \cdots & (\lambda^2_{\{s\}+1})^{\{s\}} \\ (\lambda_1^2)^{\{s\}-1} & (\lambda_2^2)^{\{s\}-1} & \cdots & (\lambda^2_{\{s\}+1})^{\{s\}-1} \\ \vdots & & & \\ (\lambda_1^2)^{\{s\}-i+1} & \cdots & \cdots & \cdot \\ f(\lambda_1) & f(\lambda_2) & \cdots & f(\lambda_{\{s\}+1}) \\ (\lambda_1^2)^{\{s\}-i-1} & \cdots & \cdots & \\ \vdots & & & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \div \det V \tag{5.25}$$

where

$$\det V = \prod_{i<j} (\lambda_i^2 - \lambda_j^2)$$

Similarly for an odd function (as $\sinh z\theta$ or $\tanh z^{\theta}$)

we have,

$$f(z) = \sum_{i=1}^{2s+1} a_i z^{2s+1-i} \tag{5.26}$$

$$f(z) = -f(-z) = -\sum a_i (-1)^{2s+1-i} z^{2s+1-i} \tag{5.27}$$

so that

$$\sum a_i [1 - (-1)^{2s-i}] z^{2s+1-i} = 0 \tag{5.28}$$

since the z^k are all linearly independent,

$$\begin{aligned} &a_i = 0 \quad \text{for } i \text{ even if } s \text{ is half integer} \\ &a_i = 0 \quad \text{for } i \text{ odd if } s \text{ is integer} \end{aligned} \tag{5.29}$$

Thus $f(z)$ is given by

$$f(z) = \sum_{i=1}^{\{s+\frac{1}{2}\}} c_i z^{2\{s+\frac{1}{2}\}+1-2i} \tag{5.30}$$

The set of equations corresponding to eq.(5.7) now read

$$f(\lambda_k) = \sum_{i=1}^{\{s+\frac{1}{2}\}} c_i \lambda_k^{2\{s+\frac{1}{2}\}+1-2i} \tag{5.31}$$

where the λ_k now run over the strictly positive eigenvalues of z only, the last equation being for $f(\frac{1}{2})$ or $f(1)$, the number of equations being $\{s + \frac{1}{2}\}$. As before, we write,

$$\underline{f}(\lambda) = V\underline{c} \tag{5.32}$$

where V is now given by

$$V = \begin{bmatrix} \lambda_1^{2\{s+\frac{1}{2}\}-1} & \lambda_1^{2\{s+\frac{1}{2}\}-3} & \cdots & \lambda_1 \\ \lambda_2^{2\{s+\frac{1}{2}\}-1} & & \cdots & \lambda_2 \\ \vdots & & & \\ \lambda_{\{s+\frac{1}{2}\}}^{2\{s+\frac{1}{2}\}-1} & & \cdots & \lambda_{\{s+\frac{1}{2}\}} \end{bmatrix} \tag{5.33}$$

so that

$$\underline{c} = V^{-1}\underline{f}(\lambda) \tag{5.34}$$

6. Explicit Solution when X is Hermitian:

We shall now use the results of the last section to find a Hermitian X which yields the unitary matrix $U = S\,X$. We require

$$XX^{+} = \operatorname{sech} 2Z\theta$$

Let X be hermitian,

$$X = (\operatorname{sech} 2Z\theta)^{\frac{1}{2}} = f(Z) \text{ say,}$$

We note that $f(z)$ is an even function of z. Also

$$\cosh\theta = \frac{E}{m}, \ \sinh\theta = \frac{p}{m}, \ \tanh\theta = \frac{p}{E}, \ \operatorname{sech}\theta = \frac{m}{E}$$

Spin $\frac{1}{2}$:

$$f(Z) = (\operatorname{sech} 2Z\theta)^{\frac{1}{2}}$$

We write

$$f(Z) = b_1$$

Since

$$\lambda_1 = \frac{1}{2}$$

$$f(\lambda) = f(\tfrac{1}{2}) = (\operatorname{sech}\theta)^{\frac{1}{2}} = \left(\frac{m}{E}\right)^{\frac{1}{2}}$$

Therefore

$$U = SX = \frac{E + m + \underline{\alpha}\cdot\underline{p}\beta}{[2m(E+m)]^{\frac{1}{2}}}\left(\frac{m}{E}\right)^{\frac{1}{2}}$$

$$= \frac{E + m + \underline{\alpha}\cdot\underline{p}\beta}{[2E(E+m)]^{\frac{1}{2}}}$$

which is identical to the transformation obtained by Foldy and Wouthuysen for spin $\frac{1}{2}$.

Spin 1 :

It is simpler to calculate U = S X directly. Thus

$$U = F(Z) = (\cosh Z\theta - \beta \sinh Z\theta)(\operatorname{sech} 2Z\theta)^{\frac{1}{2}}$$

$$= \cosh Z\theta (\operatorname{sech} 2Z\theta)^{\frac{1}{2}} - \beta \sinh Z\theta\, (\operatorname{sech} Z\theta)^{\frac{1}{2}}$$

$$= f_e(Z) - \beta f_o(Z), \text{ say}$$

where $f_e(Z)$ is an even function and $f_o(Z)$ is an odd function. Further

$$f_e(Z) = \cosh Z\theta (\operatorname{sech} 2Z\theta)^{\frac{1}{2}} = \left(\frac{1+\cosh 2Z\theta}{2} \operatorname{sech} 2Z\theta\right)^{\frac{1}{2}}$$

$$= \left(\frac{1+\operatorname{sech} 2Z\theta}{2}\right)^{\frac{1}{2}}$$

Similarly

$$f_o(Z) = \left(\frac{1-\operatorname{sech} 2Z\theta}{2}\right)^{\frac{1}{2}}$$

We write

$$f_e(Z) = b_1 Z^2 + b_2$$

$$\underline{f_e(\lambda)} = \begin{bmatrix} \left(\frac{1+\operatorname{sech} 2Z\theta}{2}\right)^{\frac{1}{2}} \\ 1 \end{bmatrix} = \begin{bmatrix} E(E^2+p^2)^{-\frac{1}{2}} \\ 1 \end{bmatrix}$$

$$V_e = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad V_e^{-1} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

Therefore

$$\underline{b} = V_e^{-1} \underline{f_e(\lambda)} = \begin{bmatrix} E(E^2+p^2)^{-\frac{1}{2}} - 1 \\ 1 \end{bmatrix}$$

and

$$f_e(Z) = Z^2\left[E(E^2+p^2)^{-\frac{1}{2}} - 1\right] + 1$$

Similarly

$$f_o(Z) = c_1 Z$$

$$V = V^{-1} = 1, \quad f_o(\lambda) = \left(\frac{1-\operatorname{sech} 2Z\theta}{2}\right)^{\frac{1}{2}} = \frac{p}{(p^2+E^2)^{\frac{1}{2}}}$$

Therefore

$$\underline{c} = V^{-1} \underline{f(\lambda)} = \frac{p}{(p^2+E^2)^{\frac{1}{2}}}$$

and

$$f_o(Z) = \frac{p}{(p^2+E^2)^{\frac{1}{2}}} Z$$

Thus

$$U = f_e(Z) - \beta f_o(Z)$$

$$= 1 - \beta Z p (E^2+p^2)^{-\frac{1}{2}} + \left[E(E^2+p^2)^{-\frac{1}{2}} - 1\right] Z^2$$

which is identical with Weaver's result.

Spin 3/2:

Here again

$$U = F(Z) = f_e(Z) - \beta f_o(Z)$$

$$f_e(Z) = b_1 Z^2 + b_2$$

$$\underline{f_e(\lambda)} = \begin{bmatrix} \left(\frac{1+\operatorname{sech} 3\theta}{2}\right)^{\frac{1}{2}} \\ \left(\frac{1+\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}} \end{bmatrix}$$

$$V = \begin{bmatrix} (3/2)^2 & 1 \\ (1/2)^2 & 1 \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} 1 & -1 \\ -\frac{1}{4} & \frac{9}{4} \end{bmatrix}$$

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = V^{-1} \underline{f_e(\lambda)} = \begin{bmatrix} \frac{1}{2}\left[\left(\frac{1+\operatorname{sech}3\theta}{2}\right)^{\frac{1}{2}} - \left(\frac{1+\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}}\right] \\ \frac{1}{8}\left[-\left(\frac{1+\operatorname{sech}3\theta}{2}\right)^{\frac{1}{2}} + 9\left(\frac{1+\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}}\right] \end{bmatrix}$$

Similarly

$$f_o(Z) = c_1 Z^3 + c_2 Z$$

$$\underline{f_o(\lambda)} = \begin{bmatrix} \left(\frac{1-\operatorname{sech}3\theta}{2}\right)^{\frac{1}{2}} \\ \left(\frac{1-\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}} \end{bmatrix}$$

$$V = \begin{bmatrix} (3/2)^3 & 3/2 \\ (1/2)^3 & 1/2 \end{bmatrix}, \quad V^{-1} = \frac{2}{3}\begin{bmatrix} \frac{1}{2} & -3/2 \\ -\left(\frac{1}{2}\right)^3 & (3/2)^3 \end{bmatrix}$$

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = V^{-1} f(\lambda) = \begin{bmatrix} \frac{1}{3}\left(\frac{1-\operatorname{sech} 3\theta}{2}\right)^{\frac{1}{2}} - \left(\frac{1-\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}} \\ -\frac{1}{12}\left(\frac{1-\operatorname{sech} 3\theta}{2}\right)^{\frac{1}{2}} + \frac{9}{4}\left(\frac{1-\operatorname{sech}\theta}{2}\right)^{\frac{1}{2}} \end{bmatrix}$$

so that

$$U = c_1 Z^3 + b_1 Z^2 + c_2 Z + b_2$$

7. The Extreme Relativistic Limit:

We shall now consider the extreme relativistic limit of the Hamiltonian for any spin. First we shall investigate the form of the Hamiltonian in this limit and give an explicit form of the Hamiltonian in this limit and give an explicit form for it and then give the unitary transformation which takes the laboratory Hamiltonian to this form. We shall find that there is a difference between the case of half-integral and integral spin because of the presence of zero as an eigenvalue of the spin matrices in the latter case.

The Hamiltonian is given by

$$H = S \beta E S^{-1} \tag{7.1}$$

where S is given by

$$S = \cosh Z\theta - \beta \sinh Z\theta \tag{7.2}$$

so that

$$H = (\beta \operatorname{sech} 2Z\theta + \tanh 2Z\theta) E \tag{7.3}$$

We approach the extreme relativistic limit by ~~taking~~ letting

$p \to E$ so that $m \to 0$ and $\theta \to \infty$ (7.4)

we shall find the limiting values of the functions sech $n z \theta$ tanh $n z \theta$ as also cosh $n z \theta$ and sinh $n z \theta$ in the limit $p \to E$ when $n \neq 0$. Two cases arise according as the spin is half integral or integral.

Half-Integral spin:

Since sech $n z \theta$ is an even function of z,

$$f(z) = \operatorname{sech} n z \theta = b_1 z^{2s-1} + b_2 z^{2s-3} + \cdots \qquad (7.5)$$

$$\underline{f(\lambda)} = \operatorname{sech} n \lambda \theta = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} \qquad (7.6)$$

the null vector. Therefore

$$\underline{b} = V^{-1} \underline{f(\lambda)} = 0$$

and

$$\operatorname{sech} n z \theta = 0 \qquad (7.7)$$

Similarly, we can show that cosh $n z \theta \to \infty$. Again, since tanh $n z \theta$ is an odd functions of z,

$$\tanh n z \theta = f(z) = c_1 z^{2s} + c_2 z^{2s-2} + \cdots \qquad (7.8)$$

$$\underline{f(\lambda)} = \tanh n \lambda \theta = \begin{bmatrix} 1 \\ 1 \\ \cdot \\ 1 \end{bmatrix} \qquad (7.9)$$

and

$$\underline{c} = V^{-1} \underline{f(\lambda)} \qquad (7.10)$$

The diagonal form Λ of $\tanh n z \theta$ is given by

$$\Lambda = \mathrm{diag}\{\underbrace{1,\cdots 1}_{\frac{2s+1}{2}\text{ times}}, \underbrace{-1\cdots -1}_{2s+1\text{ times}}, \underbrace{1,\cdots 1}_{\frac{2s+1}{2}\text{ times}}\} \tag{7.11}$$

We note that as $\theta \to \infty$, $\tanh^2 n z \theta = 1$. Similarly, we can show that $\sinh n z \theta \to \infty$. Thus, in the extreme relativistic limit the Hamiltonian for half-integral spin is given by

$$H_e = \tanh 2Z\theta \tag{7.12}$$

or, equivalently,

$$H_e = \tanh Z\theta, \theta \to \infty \tag{7.13}$$

where $\tanh z\theta$ is given by eq.(7.8)

<u>Integral Spin</u>:

$$\mathrm{Sech}\, nZ\theta = f(Z) = b_1 Z^{2s} + b_2 Z^{2s-2} + \cdots \tag{7.14}$$

$$\underline{f}(\lambda) = \mathrm{sech}\, n\lambda\theta = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{7.15}$$

Therefore

$$\underline{b} = V^{-1}\underline{f(\lambda)} \tag{7.16}$$

A more elegant form can be given to $\mathrm{sech}\, n z \theta$ using the Sylvester interpolation formula. Thus

$$\mathrm{Sech}\, 2Z\theta = \sum_i (\mathrm{Sech}\, n\lambda_i\theta) P_i \tag{7.17}$$

where P_i are orthogonal idempotent (projection) operators, satisfying $P_i P_j = \delta_{ij} P_i$, $\sum_i P_i = 1$ and explicitly given by

$$P_i = \prod_{j \neq i} \frac{z - \lambda_j}{\lambda_i - \lambda_j} \tag{7.18}$$

As $\theta \to \infty$ all the terms in eq.(7.17) vanish except the one corresponding to $\lambda_i = 0$. Thus,

$$\lim_{\theta \to \infty} \operatorname{sech} n z \theta = P_0 = \prod_{\lambda_j = 0} \frac{z - \lambda_j}{-\lambda_j} \tag{7.19}$$

which is just the characteristic polynomial of z divided by the product of all eigenvalues except o and with a factor of z missing. The diagonal form Λ of sech n z θ is given by

$$\Lambda = \operatorname{diag}\{\underbrace{0, \ldots 0}_{s \text{ times}}, 1, \underbrace{0, \ldots, 0}_{2s \text{ times}}, 1, \underbrace{0, \ldots 0}_{s \text{ times}}\} \tag{7.20}$$

We note that sech n z θ is idempotent and is a good projection operator. Similarly, for tanh n z θ

$$\tanh n z \theta = f(z) = c_1 z^{2s-1} + c_2 z^{2s-3} + \cdots \tag{7.21}$$

$$\underline{f(\lambda)} = \tanh n z \theta = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{7.22}$$

and

$$\underline{c} = V^{-1} \underline{f(\lambda)} \tag{7.23}$$

The diagonal form Λ of $\tanh n z \theta$ is given by

$$\Lambda = \text{diag}\{\underbrace{1,\dots,1}_{8\text{ times}}, 0, \underbrace{-1,\dots,-1}_{28\text{ times}}, 0, \underbrace{1,\dots,1}_{8\text{ times}}\} \tag{7.24}$$

We note again that $\tanh^2 n z \theta$ is not unity but is idempotent and is a good projection operator. Further

$$\tanh^2 n z\theta \; \text{sech}\, n z\theta = 0$$

or

$$\tanh^2 n z\theta + \text{sech}\, n z\theta = 1 \tag{7.25}$$

The Hamiltonian in the extreme relativistic limit for a particle of integral spin is therefore given by

$$H_e = \beta\, \text{sech}\, 2z\theta + \tanh 2z\theta \tag{7.26}$$

or equivalently

$$H_e = \beta\, \text{sech}\, z\theta + \tanh z\theta, \quad \theta \to \infty \tag{7.27}$$

where the functions on the right hand side are given by eqs. (7.14) and (7.21). Here, unlike the case of half integral spin, the first term does not vanish.

8. The Transformations Leading to the Extreme Relativistic Form or the Generalised CT Transformation:

With the above expressions for the functions of z, we can now obtain the (unitary) operator transforming the laboratory Hamiltonian H to the extreme relativistic limit of the Hamiltonian H_e. Since the unitary operator U transforms βE

to H, it is sufficient to find a unitary transformation connecting βE and H_e . For, since

$$H = U\beta E U^{-1} \tag{8.1}$$

if we find a unitary matrix C such that

$$H_e = C\beta E C^{-1} \tag{8.2}$$

then

$$H_e = CU^{-1}HUC^{-1} \tag{8.3}$$

or

$$H_e = RHR^{-1} \tag{8.4}$$

where the unitary matrix R

$$R = CU^{-1} \tag{8.5}$$

is the operator transforming H to H_e . Since the product of two unitary matrices is unitary, R is unitary. We shall now find the operator C for half-integral and integral spin.

<u>Half integral spin:</u>

For half-integral spin, C can be very easily found by using the U-matrix method of Ramakrishnan, dealt with in detail in Chapter IV. Here

$$H_e = \tanh Z\theta \; , \; \theta \to \infty \tag{8.6}$$

$$H_e^2 = (\beta E)^2 = E^2 \tag{8.7}$$

Therefore

$$H_e(H_e + \beta E) = (H_e + \beta E)\beta E \tag{8.8}$$

Further, since H_e anti commutes with β

$$(H_e + \beta E)^2 = 2E^2 \tag{8.9}$$

so that $H_e + \beta E$ is non singular. Therefore from eq.(8.8)

$$H_e = C \beta E C^{-1} \tag{8.10}$$

where

$$C = C^{-1} = C^{\dagger} = \frac{1}{\sqrt{2}\,E}(H_e + \beta E) \tag{8.11}$$

C being hermitian as well as involutionary and hence unitary. The operator C can also be written in exponential form. We observe again that C is arbitrary upto post-multiplication a matrix X which is non singular and which commutes with β for

$$C X \beta E X^{-1} C^{-1} = C \beta E C = H_e \tag{8.12}$$

We choose, $X = \beta$. Then eq. (8.11) yields

$$C' = \frac{1}{\sqrt{2}}\left(1 + \frac{H_e}{E}\beta\right) = \exp\left(\frac{H_e}{E}\beta\omega\right), \text{ say,} \tag{8.13}$$

where C' is still unitary. Then, since $\left(\frac{H_e}{E}\beta\right)^2 = -1$

$$\exp\left(\frac{H_e}{E}\beta\omega\right) = \cos\omega + \frac{H_e}{E}\beta\sin\omega \tag{8.14}$$

$$= \frac{1}{\sqrt{2}}\left(1 + \frac{H_e\beta}{E}\right)$$

whence

$$\omega = \tan^{-1} = \pi/4$$

Thus

$$C' = e^{\frac{H_e}{E}\beta\pi/4} = \exp\left[(\tanh z\theta)\beta\pi/4\right] \quad \theta \to \infty \tag{8.15}$$

This method however cannot be used for integral spin since H_e and β do not anticommute.

An alternative method, which can be used for integral spin also, is to start with the non-unitary FW transformation used by Weaver, Hammer and Good. Thus

$$H = S\beta E S^{-1} \tag{8.16}$$

As $\theta \to \infty$, $H \to H_e$ but the right hand side is not defined as $S \to \infty$. However as discussed earlier in this chapter, there is a freedom in choosing S upto post-multiplication by a matrix X which is nonsingular and which commutes with β. Further, to make the transformation unitary, we required $X X^+ = \text{sech}\ 2z\theta$. If X is chosen Hermitian, $X = (\text{Sech}\ 2z\theta)^{\frac{1}{2}}$ and

$$C = SX = (\cosh z\theta - \beta\sinh z\theta)(\text{sech}\ z\theta)^{\frac{1}{2}} \tag{8.17}$$

This freedom of post-multiplying by X can be exploited here, for the matrix now longer becomes infinite but on closer examination using the expression for f (z), is found to be well defined and non-singular even in the limit as $\theta \to \infty$.

Let us write

$$C = \cosh Z\theta\,(\operatorname{sech} 2Z\theta)^{\frac{1}{2}} - \beta \sinh Z\theta\,(\operatorname{sech} 2Z\theta)^{\frac{1}{2}}$$

$$\equiv fe(Z) - \beta\, fo(Z) \tag{8.18}$$

$$fe(Z) = b_1 Z^{2S-1} + b_2 Z^{2S-3} + \cdots \tag{8.19}$$

$$\underline{fe(\lambda)} = \frac{\cosh \lambda\theta}{(\cosh^2 \lambda\theta + \sinh^2 \lambda\theta)^{\frac{1}{2}}} = \frac{1}{(1 + \tanh^2 \lambda\theta)^{\frac{1}{2}}}$$

$$= \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \vdots \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{8.20}$$

so that

$$\underline{b} = V^{-1}\,\underline{fe(\lambda)} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad fe(Z) = \frac{1}{\sqrt{2}} \tag{8.21}$$

since for every b_i except the last, the i^{th} row of the determinant in eq.(5.25) is $\frac{1}{\sqrt{2}}$ times the last row and hence all the b_i vanish except for the last one. Similarly, if

$$fo(Z) = C_1 Z^{2S} + C_2 Z^{2S-2} + \cdots \tag{8.22}$$

$$\underline{fo(\lambda)} = \frac{\sinh \lambda\theta}{(\cosh^2 \lambda\theta + \sinh^2 \lambda\theta)^{\frac{1}{2}}}$$

$$= \frac{1}{(1 + \coth^2 \lambda\theta)^{\frac{1}{2}}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \vdots \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{8.23}$$

and, by comparison with eq. (7.9) we find

$$f_0 = \frac{1}{\sqrt{2}} \tanh z\theta \tag{8.24}$$

Thus we find

$$c = \frac{1}{\sqrt{2}} \left[1 + \frac{H_e}{E} \beta \right] \tag{8.25}$$

which is identical with the result obtained by the earlier method in eq.(8.13).

The evaluation of $f_e(z)$ and $f_o(z)$ can also be done by examining their diagonal forms and comparing with the diagonal forms of the functions $\tanh z\theta$, $\operatorname{sech} z\theta$, etc. evaluated earlier, since all these functions are simultaneously diagonalisable.

<u>Integral Spin:</u>

We shall here use the second of the two methods in the last section. Thus

$$\begin{aligned} c = Sx &= \cosh z\theta \, (\operatorname{sech} 2z\theta)^{\frac{1}{2}} \\ &\qquad - \beta \sinh z\theta \, (\operatorname{sech} 2z\theta)^{\frac{1}{2}} \\ &= f_e(z) - \beta f_o(z), \quad \text{say} \end{aligned} \tag{8.26}$$

If

$$f_e(z) = b_1 z^{2s} + b_2 z^{2s-2} + \cdots \tag{8.27}$$

$$f_e(\lambda) = \frac{\cosh\lambda\theta}{(\cosh^2\lambda\theta + \sinh^2\lambda\theta)^{\frac{1}{2}}}$$

$$= \frac{1}{(1+\tanh^2\lambda\theta)^{\frac{1}{2}}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ \vdots \\ \frac{1}{\sqrt{2}} \\ 1 \end{bmatrix} \tag{8.28}$$

and

$$\underline{b} = V^{-1} f_e(\lambda) = V^{-1} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} + V^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1-\sqrt{2} \end{bmatrix} \tag{8.29}$$

and comparing with eqs.(7.15) and (7.22)

$$f(Z) = \frac{1}{\sqrt{2}} + \left(1 - \frac{1}{\sqrt{2}}\right) \operatorname{sech} Z\theta \tag{8.30}$$

Similarly, if

$$f_o(Z) = C_1 Z^{2S-1} + C_2 Z^{2S-3} + \cdots \tag{8.31}$$

$$f_o(\lambda) = \frac{\sinh\lambda\theta}{(\cosh^2\lambda\theta + \sinh^2\lambda\theta)^{\frac{1}{2}}}$$

$$= \frac{1}{(1+\coth^2\lambda\theta)^{\frac{1}{2}}} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \vdots \\ \frac{1}{\sqrt{2}} \end{bmatrix} \tag{8.32}$$

and

$$\underline{b} = V^{-1} f_o(\lambda) = V^{-1} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \tag{8.33}$$

Therefore, by comparison with eq.(7.22)

$$f_0(Z) = \frac{1}{\sqrt{2}} \tanh Z\theta \qquad (8.34)$$

Thus

$$\begin{aligned} C &= f_e - \beta f_0 \\ &= \frac{1}{\sqrt{2}}\left[1 + \operatorname{sech} Z\theta - \beta \tanh Z\theta\right] \\ &= \frac{1}{\sqrt{2}}\left[1 - \sqrt{2}\operatorname{sech} Z\theta - \beta H_e\right] \end{aligned} \qquad (8.35)$$

Using eq. (7.25) this reduces to

$$C = \exp\left[(\tanh Z\theta)\beta\frac{\pi}{4}\right], \theta \to \infty \qquad (8.36)$$

It can be verified directly that C is unitary. For spin one, this yields

$$C = \exp Z\beta\pi/4$$

which is identical with the transformation for extreme relativistic limit obtained by Weaver. Thus eq. (8.36) is the unitary transformation taking β E to the extreme relativistic limit of H and is the appropriate generalisation of Weaver's result for spin 1.

9. <u>Conclusions</u>:

We have shown that for the Hamiltonian proposed by Weaver, Hammer and Good for a particle of arbitrary spin can be transformed to the form βE by a unitary transformation. We give methods for obtaining this transformation U from the transformation S used by Weaver et al. We find the additional condition which reduces this transformation to the FW transformation in the conventional form for spin $\frac{1}{2}$ and to Weaver's transformation for spin 1. We give methods for evaluating $f(z)$ in general, and in particular when it is an odd or even function of z. We use these methods to find the extreme relativistic form of the Hamiltonian and the unitary transformation leading to it. This leads to a generalisation for arbitrary spin of the corresponding result of Weaver for spin 1.

<u>References</u>:

1. D.L.Weaver, C.L.Hammer and R.H.Good, Jr., Phys.Rev. <u>135</u>, B241 (1964). For a detailed discussion on the subject see R.H.Good, Jr. 'Description of Particle with any spin and with internal symmetry', Matscience Report 62.
2. D.L. Weaver, Nuov. Cim. <u>53A</u>, 665 (1968).
3. A.K.Nagpal and R.S.Kaushal, Lett.Nuov.Cim. <u>9</u>,391 (1974).
4. T.A.Weber and S.A.Williams, Jour.Math.Phys., <u>6</u>, 1980 (1965).
   S.A.Williams, J.P.Draayer and T.A.Weber, Phys.Rev., <u>152</u>, 1207 (1966).
5. A.R.Tekumalla and T.S.Santhanam, Unitary Foldy Wouthuysen Transformation for arbitrary spin, to appear Lett.Nuov. Cim., (1974).

CHAPTER. VI

MATRIX METHODS AND APPLICATIONS

In this Chapter, we shall deal with some applications of Matrix theory to physical problems.

A. QUANTUM MECHANICS IN FINITE DIMENSIONS*

1. Introduction:

Weyl[1] has shown that the Schroedinger representation for the momentum operator is a necessary consequence of Heisenberg's commutation relation. He proves this using the ray representations of the Abelian group of rotations. In proving this, he uses an ingenious limiting process to go from finite rotations in ray space to a 2-parameter continuous group. More recently, Alladi Ramakrishnan[2] and his collaborators have studied exhaustively the representation theory of Generalised Clifford Algebra which immediately furnishes the ray representations of the Abelian group of rotations.

Here we derive, by limiting to the case of finite dimensions, the explicit expression for the commutator $[Q,P]$ where Q and P are the position and momentum operators respectively. We show that by going to the limit of continuous parametrisation (valid as the dimension goes to infinity) we recover the standard Heisenberg commutation relations.

We believe that this work will open up the possibility of

---

* T.S.Santhanam and A.R.Tekumalla, Comm.to Int.Jour.Theor. Phys., (1974).

studying quantum mechanics in finite dimensions.

2. Weyl's form of the Heisenberg Relations:

Suppose A and B are two elements of the Abelian group of unitary rotations on a ray space so that

$$AB = \omega BA \tag{2.1}$$

where ω is a primitive n^{th} root of unity. By iteration we get

$$A^k B^\ell = \omega^{k\ell} B^\ell A^k \tag{2.2}$$

from which it follows that A^n commutes with B and B^n commutes with A and if the representation is irreducible, it follows from Schur's lemma that

$$A^n = I, \quad B^n = I \tag{2.3}$$

We take the following representations for A and B[1]

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdot & 0 \\ 0 & 0 & 1 & \cdot & 0 \\ & & & & \\ 0 & 0 & 0 & \cdot & 1 \\ 1 & 0 & 0 & \cdot & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & & & \\ & \omega & & 0 \\ & & \omega^2 & \\ 0 & & & \omega^{n-1} \end{bmatrix} \tag{2.4}$$

The interesting properties of the algebra satisfied by operators like A and B which is a generalisation of the usual Clifford algebra have been systematically studied by Alladi Ramakrishnan[2] and collaborators. If one identifies

$$A = e^{i\xi P} \quad \text{and} \quad B = e^{i\eta Q} \tag{2.5}$$

where ξ, η are arbitrary real parameters, then it follows that eqn.(2.1) is the Weyl form of the Heisenberg commutation relation $[Q, P] = i$, if we allow power series expansion of operator exponentials (which is justified if A is bounded but not otherwise)[3]. Weyl takes the limit $n \to \infty$ such that $\xi \eta n = 2\pi$ to show that

$$P = -i\frac{\partial}{\partial q} \tag{2.6}$$

3. <u>Case of Finite Dimensions</u>:

We now solve eqn.(2.5) for P and Q by taking logarithms. We solve the problem that, given

$$e^{i\xi P} = A, \quad e^{i\eta Q} = B \tag{3.1}$$

where ξ and η are arbitrary real parameters and

$$AB = \omega BA, \qquad \omega^n = 1 \tag{3.2}$$

to compute the commutator $[Q, P]$ and show that

$$[Q, P] = i \tag{3.3}$$

in the continuous limit.

We take A, B as in eq.(2.4). The diagonal form of A is given by B, for

$$S^{-1} A S = B \tag{3.4}$$

where S is the sylvester matrix

$$S = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{n-1} \\ \cdot & & & & \\ \cdot & & & & \\ 1 & \omega^{n-1} & \omega^{n-2} & \cdots & \omega \end{bmatrix} \tag{3.5}$$

$$S^{-1} = S^{\dagger} \tag{3.6}$$

We then have, taking logarithms

$$i\xi P = \log A = S \log B \, S^{-1} \tag{3.7}$$

$$i\eta Q = \log B \tag{3.8}$$

where log B is given by

$$\log B = \log\omega \begin{bmatrix} 0 & & & \\ & 1 & & 0 \\ & & 2 & \\ 0 & & & n-1 \end{bmatrix} \tag{3.9}$$

Since A and B are diagonalizable and non-singular it follows[4] that log A and log B exist. Elementwise, labelling the rows and columns from 0 to n-1,

$$B_{rs} = \omega^{r} \delta_{rs} \tag{3.10}$$

$$S_{rs} = \frac{1}{\sqrt{n}} \omega^{rs}, \quad (S^{-1})_{rs} = \frac{1}{\sqrt{n}} \omega^{-rs} \tag{3.11}$$

$$(\log B)_{rs} = (\log\omega)\, r\, \delta_{rs} \tag{3.12}$$

We now compute the commutator

$$[i\eta Q, i\xi P] = [\log B, S(\log B)S^{-1}]$$

We have

$$[i\eta Q, i\xi P]_{rs} = \{(\log B)S(\log B)S^{-1}\}_{rs} - \{S(\log B)S^{-1}(\log B)\}_{rs}$$

$$= \sum_{t,u,v=0}^{n-1} \{(\log B)_{rt} S_{tu} (\log B)_{uv} (S^{-1})_{vs} - S_{rt} (\log B)_{tu} (S^{-1})_{uv} (\log B)_{vs}\}$$

$$= \sum_{t,u,v} \frac{(\log \omega)^2}{n} \left(r\delta_{rt}\,\omega^{tu}\,u\,\delta_{uv}\,\omega^{-vs} - \omega^{rt}\,t\,\delta_{tu}\,\omega^{-uv}\,v\,\delta_{vs}\right)$$

$$= \frac{(\log\omega)^2}{n} \sum_{u=0}^{n-1} \omega^{u(r-s)}\,u(r-s)$$

Therefore

$$[i\eta Q, i\xi P]_{rs} = \frac{(\log\omega)^2}{n}(r-s)\sum_{u=0}^{n-1} u\,\omega^{u(r-s)} \qquad \textbf{(3.13)}$$

If $\omega^{r-s} = x = 1$

$$[i\eta Q, i\xi P]_{rs} = (\log\omega)^2 (r-s)\,\frac{n-1}{2} \qquad \textbf{(3.14)}$$

If $\omega^{r-s} = x \neq 1$, then, since $x^{n} = 1$, we have

$$\sum_{u=0}^{n-1} u x^{u} = \frac{n}{x-1} \tag{3.15}$$

and hence

$$[i\eta Q, i\xi P]_{rs} = \frac{(\log w)^2}{} \frac{r-s}{(\omega^{r-s}-1)} \tag{3.16}$$

Thus we have,

$$[Q, P]_{rs} = \frac{(s-r)(\log w)^2}{n\xi\eta} \frac{n(n-1)}{2}$$
$$\text{when } \omega^{r-s} = 1,$$

and

$$[Q, P]_{rs} = \frac{(s-r)(\log w)^2}{n\xi\eta} \frac{n}{(\omega^{r-s}-1)}$$
$$\text{when } \omega^{r-s} \neq 1$$

We notice that since n is finite we could choose $\xi = \eta = 1$ and that $[Q,P]$ is strictly off diagonal and hence is trace free.

We now prove that the commutation relation given by eqn. (3.18) does indeed yield the Heisenberg commutation relation in the limit as $n \to \infty$. We begin with eqn.(3.13), relabelling the rows and columns from $-\left(\frac{n-1}{2}\right)$ to $\left(\frac{n-1}{2}\right)$ and replace the sum by an integral, that is, we let the matrix index take continuous values. Thus the sum

$$-\frac{(\log\omega)^2}{n\xi\eta}\sum_{u=0}^{n-1} u(r-s)\,\omega^{u(r-s)}$$

reduces in the limit as $n \to \infty$ to the integral[5,6].

$$-\frac{(\log\omega)^2}{n\xi\eta}\int_{-\infty}^{+\infty} u(r-s)\exp\left(\frac{2\pi i}{n}u(r-s)\right)du$$

$$= -\left(\frac{2\pi i}{n}\right)^2\frac{1}{n\xi\eta}(r-s)\int u\exp\left(\frac{2\pi i}{n}u(r-s)\right)du$$

$$= 2\pi(r-s)\int\frac{u}{n}\exp\left(2\pi i(r-s)\frac{u}{n}\right)d(u/n)$$

$$= -i(r-s)\frac{d}{d(r-s)}\int\exp\left(2\pi i(r-s)\frac{u}{n}\right)d\left(\frac{u}{n}\right)$$

$$= -i(r-s)\,\delta'(r-s)$$

$$= i\,\delta(r-s)$$

where we have used $\operatorname{Lim} n\xi\eta = 2\pi$ as $n \to \infty$.

This completes the proof. It should be remembered that in the limit as n approaches infinity continuously, we are taking only the principal value of $\log\omega$ as this gives the correspondence to the Heisenberg commutation relation.

4. Conclusions:

We have calculated the commutator $[Q,P]$ when the space on which the operators act is finite. We elevate the commutator for finite n to what we call 'Finite Dimensional Quantum Mechanics'. It turns out that the operator is strictly off-diagonal for finite n.

This implies no uncertainty and no zero point energy if these concepts have any meaning for finite n . Of course, in the limiting case as n approaches infinity continuously, the commutator becomes strictly diagonal and reduces to a multiple of the Dirac delta function, thus restoring the Heisenberg commutation relations.

<u>References</u>:

1. H. Weyl, "Theory of Groups and Quantum Mechanics", Dover (1931) p.272.280.
2. A.Ramakrishnan, "L-matrix theory or the Grammar of Dirac Matrices", Tata-McGraw Hill, Bombay (1972).
3. P.Carter, "Proceedings of Symposia in Pure Mathematics", Vol.9, Algebraic Groups and Discontinuous Subgroups, American Mathematical Society (1966) pp.361-383.
4. Gantmacher, "Matrix Theory", Vol.1, Chelsea, New York (1959) p.99.
5. P.A.M.Dirac, Proc. Roy. Soc., <u>A113</u>, 621 (1927)
6. W.Heisenberg, W. "The Physical Principles of the Quantum Theory", Dover (1931) pp.129.

B. SEQUENCE SEGREGATION AND THE SYLVESTER MATRIX*

The Sylvester matrix has been recently shown[1] to be very useful in a variety of problems ranging from elementary particle physics to problems in electrical engineering. We shall give here a simple property of the Sylvester matrix which is very useful in the analysis of electrical networks.

J.E. Parton[2] has given a matrix analysis of symmetrical component networks for N-phase systems. For a 2N-terminal network, ignoring zero sequence component, it has been shown that when N is prime only one network is needed for sequence segregation but when N contains factors an additional network is required for each different factor. In this note[3] we point out that the above result follows from a very simple property of the Sylvester Matrix.

The Sylvester Matrix of the N^{th} order associated with is given by

$$S_N(\omega) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega^{N-1} & \omega^{N-2} & \cdots & \omega \\ \vdots & & & & \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \end{bmatrix} \tag{1}$$

where ω is a primitive N^{th} root of unity. We denote by $\{S_N\}$ the collection of N-1 S-matrices obtained from

* T.S.Santhanam and A.R.Tekumalla, Matrix and Tensor Quarterly, **21**, 121 (1971).

$S_N(\omega)$ by replacing ω by $\omega^2, \omega^3, \ldots, \omega^{N-1}$. We observe the interesting property that when N is prime, the whole collection $\{S_N\}$ can be obtained from a single member by permutation of the rows. On the other hand, when N contains factors, the collection $\{S_N\}$ splits up into sub-collections $\{S_N^{(1)}\}$, $\{S_N^{(2)}\}$, ..., $\{S_N^{(K)}\}$, where K is equal to the number of factors of N. The members of any one sub-collection are connected by permutation of rows while sub-collections $\{S_N^{(i)}\}$, $\{S_N^{(j)}\}$, $i \neq j$ are not connected by such a permutation.

The output voltages $V_{10}, \ldots, V_{N0}$ are related to the symmetrical components $V_{11}, \ldots V_{N-1,1}$ of the input voltage by the equation

$$\begin{bmatrix} V_{10} \\ V_{20} \\ \cdot \\ V_{N0} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & \ldots & t_{1N} \\ t_{21} & t_{22} & \ldots & t_{2N} \\ & \cdot & & \\ t_{N1} & t_{N2} & \ldots & t_{NN} \end{bmatrix} \tag{2}$$

where we assume the zero sequence component to be 0. Symbolically this can be written as

$$\underline{V}_o = T\, S\, \underline{V}_s \tag{3}$$

where $\underline{V}_o$ is the output voltages vector, T the transfer matrix, S the Sylvester matrix and $\underline{V}_s$ the symmetrical components vector. The ordered symmetry requirements of the network elements enable us to express the transfer matrix as

$$T = \sum_{i=1}^{N} t_{1i} P^{i-1} \tag{4}$$

where C is the cyclic permutation matrix given by

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5}$$

with $C^N = I$.

Since the symmetry conditions imply that we need to know only the output voltage of the 1st phase we have the equation

$$V_{10} = [t_{11} \; t_{12} \cdots t_{1N}] \, S \underline{V}_s = \underline{t}^T S \underline{V}_s \tag{6}$$

where

$$t^T = [\, t_{11}, t_{12} \cdots t_{1N} \,]$$

we denote

$$\underline{t}^T S = \underline{U}^T, \quad U_0 = \sum_{i=1}^{N} t_{1i} \tag{7}$$

Then, by transposition, remembering S is symmetric,

$$S\underline{t} = \underline{U} \tag{8}$$

If the M^{th} order sequence component is required at the output the conditions for isolating it are obtained by setting

$$U_1 = \cdots = U_{N-1} = 0 \text{ except that } U_M \neq 0 \tag{9}$$

Hence the M^{th} order segregation can be obtained from

1st order segregation by permuting suitably the elements of the vector U . So the point is, two different orders of network are obtained by reconnection if and only if the corresponding S-matrices in the collection $\{S_N\}$ are connected by a permutation of the rows.

By our earlier analysis, it is therefore clear that only a single network is needed in the case when N is prime since the entire collection $\{S_N\}$ are related by permutations. On the other hand, if N contains factors some members of the collection $\{S_N\}$ become singular and hence there exists no permutation which takes it from $S_N(\omega)$.

For instance, for N = 3, the collection $\{S_3\}$ consists of the two matrices

$$S_3(\omega) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega^2 & \omega \\ 1 & \omega & \omega^2 \end{bmatrix}, \quad S_3(\omega^2) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix} \tag{10}$$

They are obviously connected by a permutation of the rows, since

$$S_3(\omega^2) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} S_3(\omega) \tag{11}$$

On the other hand, when N = 4, the collection $\{S_4\}$ consists of the three matrices

$$S_4(\omega) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega^3 & \omega^2 & \omega \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & \omega & \omega^2 & \omega^3 \end{bmatrix}, \quad S_4(\omega^2) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & 1 & 1 & 1 \\ 1 & \omega^2 & 1 & \omega^2 \end{bmatrix}, \quad S_4(\omega^3) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega \\ 1 & \omega^2 & 1 & \omega^2 \\ 1 & \omega^3 & \omega^2 & \omega \end{bmatrix} \tag{12}$$

The members of the subcollection consisting of the matrices $S_4(\omega)$, $S_4(\omega^3)$ are connected by a permutation since

$$S_4(\omega^3) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} S_4(\omega) \tag{13}$$

However, $S_4(\omega^2)$ cannot be so connected. Indeed $S_4(\omega^2)$ is singular while $S_4(\omega)$ is non-singular. Therefore, in this case, we need two independent networks. In general, it is easy to see that when N is nonprime, the collection $\{S_N\}$ breaks up into subcollections and we need as many independent networks as there are subcollections not connected by a permutation of the rows.

Further, in each independent network, only $\frac{N}{M}$ elements are distinct as can be seen from Parton's analysis. We find that these elements can be chosen to be identical with $\frac{N}{M}$ elements in the 1st order network.

References:

1. Alladi Ramakrishnan, "L-Matrix Theory or the Grammar of Dirac Matrices", Tata McGraw Hill, Bombay, 1972.

2. J.E.Parton, "The matrix and tensor Quarterly", Vol.20,3,1970.

3. T.S. Santhanam and A.R. Tekumalla, Matrix and Tensor Quarterly, 21, 121 (1971).

C. TRACE OF PRODUCTS OF CLIFFORD ELEMENTS*

1. Introduction:

Several formulae exist for the evaluation of the traces of Pauli and Dirac matrices and products of their linear combinations[1]. These are very useful especially when one works with perturbation theory to higher orders. In particular, Caianiello and his collaborators[1] and Chisholm[2] have identified them in terms of Pfaffians. Recently, Alladi Ramakrishnan and his collaborators[3] have initiated studies on Generalized Clifford Algebra and its possible applications. In this note,[6] we study the trace properties of products of linear combinations of generalized Clifford elements which yields a possible generalization of the concept of a pfaffian. We point out a lack of uniqueness in such a definition which is inherent in ordered commutation relations.

2. Trace Properties of Products of Linear Combinations of Clifford Elements:

The Clifford algebra C_n^2 is defined as consisting of the set of elements $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ satisfying

$$\{\alpha_i, \alpha_j\} = (\alpha_i\alpha_j + \alpha_j\alpha_i) = 2\delta_{ij} \qquad i,j = 1, 2, \ldots, n. \tag{1}$$

The set of elements defined by

$$A_{p_1 p_2 \cdots p_n} = \alpha_1^{p_1}\alpha_2^{p_2}\cdots\alpha_n^{p_n} \tag{2}$$

---

* A.R.Tekumalla and T.S.Santhanam, Matrix and Tensor Qly. 23, 99 (1973).

with p_i being an integer modulus 2 constituting a set of $2^{2\{\frac{n}{2}\}}$ linearly independent elements with a product defined by eqn.(1) form an algebra C_n^2. It is known[4] that when $n=2\nu$ this algebra is isomorphic to the matrix ring in 2^ν dimensions. When $n=2\nu+1$ the algebra is isomorphic to the direct sum of two matrix rings one obtained from the other by changing the signs of all the matrices. The last element $\alpha_{2\nu+1}$ can be expressed as the product

$$\alpha_{2\nu+1} = i^{\nu} \alpha_1 \alpha_2 \cdots \alpha_{2\nu} \tag{3}$$

Consider now the linear combination

$$L^{(\mu)} = \sum_{i=1}^{n} a_i^{(\mu)} \alpha_i \tag{4}$$

From eqns.(1) and (2) it is clear that all the elements of C_n^2 except the unit matrix are traceless and $A_{p_1 p_2 \cdots p_\nu}$ will be the unit element iff the p_i are even. Let us first treat the case when $n=2\nu$. Then, using eqn.(1) we have the following result. When k is odd, say $2k+1$

$$T = Tr\, L^{(1)} L^{(2)} \cdots L^{(2k+1)} = 0 \tag{5}$$

and when k is even, say $2k$

$$T = Tr\, L^{(1)} \cdots L^{(2k)} = m^{\nu} (12 \cdots 2k) \tag{6}$$

where $(1\,2 \cdots 2k)$ is a pfaffian defined by

$$(1\,2 \cdots 2k) = \sum_{p} (-1)^p (i_1 i_2)(i_3 i_4) \cdots (i_{2k-1} i_{2k}) \tag{7}$$

where the summation is over the permutations of $i_1, i_2, \ldots, i_{2k}$ such that $i_1 < i_3 < \cdots < i_{2k-1}$ and $i_1 < i_2$, $i_3 < i_4, \ldots$ etc and p is the parity of the permutation with respect to the original ordering $1\,2 \ldots 2k$. The bracket (12) denotes the scalar product

$$(12) = a_1^{(1)} a_1^{(2)} + a_2^{(1)} a_2^{(2)} + \cdots \tag{8}$$

This result is well known in the case of Dirac and Pauli matrices and we find it is true for the elements of $C_{2\nu}^2$.

Let us now consider the case when n is odd, say $n = 2\nu+1$. Define a linear combination

$$L^{(\mu)} = \sum_{i=1}^{2\nu+1} a_i^{(\mu)} \alpha_i \tag{9}$$

and the trace

$$T = \operatorname{Tr} L^{(1)} L^{(2)} \cdots L^{(k)}$$

By virtue of eqn.(1) and (3), the trace is given by eqn.(6) when k is even. However, for k odd, the corresponding eq.(5) is not in general true. The case of special interest is thus when the number of generators n is odd, the number of factors k is odd and $L^{(\mu)}$ contain all the $2\nu+1$ generators in the linear combination. In this case we find

$$\operatorname{Tr} L^{(1)} \cdots L^{(2k+1)} = 0 \tag{10}$$

when $k < \nu$ and

$$\mathrm{Tr}\, L^{(1)} \cdots L^{(2\nu+1)} = m^{\nu}\{1\,2 \cdots 2\nu+1\} \qquad (11)$$

where $\{1\,2 \cdots 2\nu+1\}$ is the determinant defined below. Now the trace of $2k+1$ factors can be reduced recursively by using the above eqn. to yield

$$\mathrm{Tr}\, L^{(1)} \cdots L^{(2k+1)} = m^{\nu}(1\,2 \cdots 2k+1\}_{2\nu+1} \qquad (12)$$

where the new bracket is expanded as

$$(1\,2 \cdots 2k+1\}_{2\nu+1} = \sum_{p} (-1)^{p} \{i_1 i_2 \cdots i_{2\nu+1}\}(i_{2\nu+2} \cdots i_{2k+1}) \qquad (13)$$

where

$$\{i_1 i_2 \cdots i_{2\nu+1}\} = \mathrm{Det}\begin{vmatrix} a_1^{i_1} & a_1^{i_2} & \cdots & a_1^{i_{2\nu+1}} \\ a_2^{i_1} & a_2^{i_2} & \cdots & a_2^{i_{2\nu+1}} \\ \vdots & & & \\ a_{2\nu+1}^{i_1} & a_{2\nu+1}^{i_2} & \cdots & a_{2\nu+1}^{i_{2\nu+1}} \end{vmatrix}$$

and $(i_{2\nu+2} \cdots i_{2k+1})$ is a pfaffian. and the summation is over all permutations of $i_1, i_2, \ldots, i_{2k+1}$ such that $i_1 < i_2 < \cdots < i_{2\nu+1}$ and $i_{2\nu+2} < i_{2\nu+3} < \cdots < i_{2k+1}$ and $i_1 < i_{2\nu+2}$ and the signature factor $(-1)^{p}$ is the parity of the permutation $i_1 \cdots i_{2k+1}$ with respect to the original permutation $1\,2 \cdots 2k+1$. Eq.(13) can also be written (with a few number of terms especially when $k \gg \nu$)

as

$$(12\ldots 2k+1\}_{2\nu+1} = \{12\ldots 2\nu+1\}(2\nu+2,\ldots 2k+1)$$

$$+\sum_{P}(-1)^{P}(i_1 i_2)(i_3\ldots i_{2k+1}\}_{2\nu+1} \quad (14)$$

where the sum is over all permutations i_1, i_2 of $1, 2, \ldots, 2\nu+1$ such that $i_1 < i_2$ keeping $i_3 < i_4 < \cdots < i_{2k+1}$.

All the results given above in eq.5,6,11,12 can now be combined in the eq.

$$\mathrm{Tr}\, L^{(1)} L^{(2)} \cdots L^{(k)} = m^{\nu} (12\ldots k\}_{2\nu+1}$$

where the bracket now means

$$(12\ldots k\}_{2\nu+1} \equiv (12\ldots k)$$

when k is even and it is expanded as in eqs.(13) or (14) when k is odd.

3. Generalized Clifford Elements and Their Trace Properties:

The set of elements $e_1, \ldots, e_n$ satisfying the ordered commutation relation

$$e_i^m = 1,\ \forall i \quad \text{and} \quad e_i e_j = \omega e_j e_i,\ i<j \quad (15)$$

$$\text{and} \quad i, j = 1, 2, \ldots, n$$

where ω is a primitive m^{th} root of unity fowm the base elements of the Generalized Clifford algebra C_n^m, the ordinary Clifford algebra being obtained as a special case when $m=2$. The set of elements $e_1^{r_1} e_2^{r_2} \dots e_n^{r_n}$ consisting of m^n elements which are linearly independent with r_i = integer mod. m form the algebra C_n^m with product defined by eq.(15). Again in this case, it has been shown[5] that when n is even, $C_{2\nu}^m$ is isomorphic to the matrix ring of dimension m^ν and when n is odd, $C_{2\nu+1}^m$ reduces to m copies of $C_{2\nu}^m$ consisting of elements $\omega^i e_1^{r_1} e_2^{r_2} \dots e_{2\nu}^{r_{2\nu}}$, $i=0,1,\dots,m-1$. Even in this case, the last generator is obtained as a product of the other 2ν as

$$e_{2\nu+1} = \omega^j e_1^{m-1} e_{\nu+1} e_2^{m-1} e_{\nu+2} \dots e_\nu^{m-1} e_{2\nu} \qquad (16)$$

All the elements of C_n^m except the unit element are traceless and the unit element is obtained if $r_1, r_2, \dots, r_n = 0$ mod. m

Let us now specialise on the case $n=2\nu$. Define the linear combination

$$\mathcal{L}^{(\mu)} = \sum_{i=1}^{2\nu} a_i^{(\mu)} e_i \qquad (17)$$

and

$$T = \mathrm{Tr}\, \mathcal{L}^{(1)} \mathcal{L}^{(2)} \dots \mathcal{L}^{(k)} \qquad (18)$$

Then

$$T = \sum_P \prod_{i=1}^{K} a_{\mu_i}^{(i)} \, \mathrm{Tr} \prod_{i=1}^{K} e_{\mu_i} \qquad (19)$$

where the summation is over the permutations of the subindices. This can be rewritten as

$$T = \mathrm{Tr} \sum C_{l_1 \ldots l_{2\nu}} e_1^{l_1} \ldots e_{2\nu}^{l_{2\nu}} \tag{20}$$

where l_i satisfy

$$l_1 + l_2 + \ldots + l_{2\nu} = K \tag{21}$$

The only terms in the expression which give a nonvanishing trace occur when

$$l_1, l_2, \ldots, l_{2\nu} = 0 \text{ mod. } m \tag{22}$$

Thus

$$T = m^\nu \sum_{\{l\}} C_{l_1 \ldots l_{2\nu}} \tag{23}$$

where l_i satisfy eq.(21) and (22) and

$$C_{l_1 \ldots l_{2\nu}} = \sum_P \left[\prod_{i_1=1}^{l_1} a_1^{(i_1)} \prod_{i_2=l_1+1}^{l_1+l_2} a_2^{(i_2)} \ldots \prod_{K-l_{2\nu}+1}^{K} a_{2\nu}^{i_{2\nu}} \right] \omega^{T_p} \tag{24}$$

where the sum is over all permutations of the subindices and T_p is the number of adjacent transpositions required to take it to the completely ordered form.

From eq.(21) and (22), it immediately follows that T = 0 when $K \neq 0 \text{ mod } m$ (25). The above expression for T can be recast in the pfaffian-like form

$$T = m^\nu (12 \cdots K)_m \tag{26}$$

where

$$(12\ldots K)_m = \sum_p \Delta p \prod_{j=0}^{K-1} (i_{jm+1}\, i_{jm+2} \ldots i_{jm+m})$$

where $i_{jm+1} < \ldots < i_{jm+m}$

and $i_1 < i_{m+1} < \ldots < i_{K-m+1}$

and

$$(12\ldots m)_m = \sum_{i=1}^{2\nu} a_i^{(1)} a_i^{(2)} \ldots a_i^{(m)} \tag{27}$$

The phase factor is a function of ω and is not uniquely fixed and can be calculated for each of the terms in the final expansion on the basis of the prescription given in eq.(24). This is because the e_i obey on ordered commutation relation and not a cyclic one. In this sense, eq.(26) can only be interpreted as giving the various terms in the expansion while the explicit ω factor has to be worked out in each case. However eq.(27) may be considered a kind of generalization of the well known pfaffian.

The case when $2\nu+1$ reduces to the above form when the number of factors k is a multiple of m . However, in general, for arbitrary k we have the result

$$T = \sum_{\{l\}} C_{l_1 l_2 \ldots l_{2\nu+1}}$$

where $\{l\}$ is chosen such that

$$l_1 - l_{2\nu+1} = 0 \mod m$$
$$l_2 - l_{2\nu+1} = 0 \mod m$$
$$l_\nu - l_{2\nu+1} = 0 \mod m$$

and

$$l_{\nu+1} + l_{2\nu+1} = 0 \mod m$$
$$l_{\nu+2} + l_{2\nu+1} = 0 \mod m$$
$$l_{2\nu} + l_{2\nu+1} = 0 \mod m.$$

References:

1. E.R.Caianiello, Nuov. Cim., 13, 637 (1959).

2. J.S.R.Chisholm, Nuov. Cim., 30, 426 (1963).

3. Alladi Ramakrishnan, "L-Matrix Theory or the Grammar of Dirac Matrices," Tata McGraw Hill, Bombay (1972).

4. See, for example, H.Boerner, Representation of Groups (North-Holland, Amsterdam, 1963), Chap. 8

5. Alladi Ramakrishnan, T.S.Santhanam and P.S.Chandrasekaran, J.Math.Phys. Sci. (Madras) 3, 307 (1969).

6. A.R.Tekumalla and T.S. Santhanam, Matrix and Tensor Quarterly. 23, 99 (1973).

D. AN APPLICATION OF THE REARRANGEMENT OPERATION *

In connection with a New Approach to Matrix Theory, Alladi Ramakrishnan[1] defined a new geometrical operation on a matrix which he termed the rearrangement operation. We shall here[2] use this operation to find the matrix which diagonalises a complex circulant.

The matrix is here looked upon not as rows and columns but as diagonals. In a rearrangement, the principal diagonal is placed in the first column, the first superdiagonal and $(N-1)^{th}$ subdiagonal in the second column, the second superdiagonal and the $(N-2)^{th}$ subdiagonal in the third column and so on. Thus for example, if A^R is the rearranged form of the matrix A,

$$A^R = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \\ a_{20} & a_{21} & a_{22} \end{bmatrix}^R = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{11} & a_{12} & a_{10} \\ a_{22} & a_{20} & a_{21} \end{bmatrix} \qquad (1)$$

The operation can be repeated and we denote by A^{R_K} the matrix obtained by successive rearrangement k times. The original matrix is restored after N rearrangements where N is the dimension of the matrix.

The operation may also be performed as follows: Leave the first row unchanged, perform a cyclic permutation of the elements of the first row once, second row twice, etc. Elementwise, A^{R_K} can be written as

* A.R. Tekumalla, MATSCIENCE Preprint, 1974.

$$\left(A^{R_k}\right)_{ij} = A_{i,\, j+ki} \tag{2}$$

where the rows and columns are labelled from 0 to N-1.

Let C and B be the matrices

$$C = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & & & \\ & \omega & & \\ & & \omega^2 & \\ & & & \ddots & \\ & & & & \omega^{N-1} \end{bmatrix} \tag{3}$$

which satisfy $C^N = B^N = 1$, $CB = \omega\, BC$, where is a primitive N^{th} root of unity. It is well known that the matrix C (or in fact any circulant, which is a linear combination of powers of C) is diagonalised by the Sylvester matrix S

$$S = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ \vdots & & & & \\ 1 & \omega^{N-1} & \omega^{N-2} & \cdots & \omega \end{bmatrix}, \qquad S^{-1} = S^{+} = S^{*} \tag{4}$$

We define a complex circulant as a linear combinations of powers of CB (or CB^l; $l \neq 0, \mathrm{mod}\, N$) where

$$CB = \begin{bmatrix} 0 & \omega & 0 & \cdots & 0 \\ 0 & 0 & \omega^2 & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & \omega^{N-1} \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad CB^l = \begin{bmatrix} 0 & \omega^l & 0 & \cdots & 0 \\ 0 & 0 & \omega^{2l} & \cdots & 0 \\ \vdots & & & & \\ 0 & 0 & 0 & \cdots & \omega^{l(N-1)} \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{5}$$

we now prove the following:

THEOREM: The matrix CB^l (and hence any complex circulant consisting of a linear combination of its powers) is diagonalised by the matrix obtained by a repeated rearrangement of the Sylvester matrix $\frac{N-1}{2}\ell$ times, that is by $S^{R_{\frac{N-1}{2}\ell}}$

Proof: Elementwise, the above matrices can be written as

$$C_{ij} = \delta_{i+1,j} \quad B_{ij} = \omega^j \delta_{ij}; \quad (CB^\ell)_{ij} = \omega^{\ell j}\delta_{i+1,j}$$

$$S_{ij} = \frac{1}{\sqrt{N}}\omega^{ij}, \quad S^{R_K}_{ij} = \frac{1}{\sqrt{N}}\omega^{i(j+Ki)}$$

It can also be verified that S^{R_K} is unitary for all k and $((S^{R_K})^{-1})_{ij} = \frac{1}{\sqrt{N}}\omega^{-j(i+Kj)}$. Now

$$\left[\left(S^{R_{\frac{N-1}{2}\ell}}\right)^{-1} CB^\ell \left(S^{R_{\frac{N-1}{2}\ell}}\right)\right]_{ij} = \sum_{p,q}\frac{1}{N}\omega^{-p(i+\frac{N-1}{2}\ell p)}\omega^{\ell q} \times \delta_{p+1,q}\,\omega^{q(j+\frac{N-1}{2}\ell q)} \tag{6}$$

Replacing q by $p+1$ the exponent reduces to

$$-p(i+\tfrac{N-1}{2}\ell p) + (p+1)(\ell+j+\tfrac{N-1}{2}\ell p + \tfrac{N-1}{2}\ell)$$
$$= p(-i+j+\tfrac{N+1}{2}\ell + \tfrac{N-1}{2}\ell) + j + \tfrac{N+1}{2}\ell$$

are remembering $\omega^{\ell N} = 1$, we have

$$\left[\left(S^{R_{\frac{N-1}{2}\ell}}\right)^{-1} CB^\ell \left(S^{R_{\frac{N-1}{2}\ell}}\right)\right]_{ij} = \omega^{\frac{N+1}{2}\ell}\,\omega^j\,\delta_{ij} \tag{7}$$

which is diagonal. The jth column of $S^{R_{\frac{N-1}{2}\ell}}$ is the eigen vector of CB^ℓ corresponding to eigen value $\omega^{\frac{N+1}{2}\ell+j}$

An especially interesting case arises when N is even and l is odd (say 1). Then $\frac{N-1}{2}$ is half-integral and the rearrangement has to be done a half integral number of times. This can be looked upon as the cyclic permutation of the i^{th} row by the half integral number is $\frac{N-1}{2}$ times where by a half permutation we mean the process in which the j^{th} element is replaced not by the succeeding $(j+1)^{th}$ element but the intermediate $(j+\frac{1}{2})^{th}$ element which is supposed to be the geometric mean of the j^{th} and $(j+1)^{th}$ elements. Alternatively it may be defined as in eq.(2).

We may mention here that for N odd, N distinct integral rearrangement are possible and each of these diagonalises one matrix of the set CB^l; $l=0,1,\ldots,N-1$. For N even however, the integral rearrangements are not all distinct, the $(\frac{N}{2}+K)^{th}$ rearrangement being just the K^{th} rearrangement with the columns permuted (and therefore their columns are not two distinct sets of vectors and cannot diagonalise two non-commuting matrices). However, when we include the semi rearrangements, we get the N distinct matrices which diagonalise the set CB^l.

Illustrations:

N odd

i) $N=3$

$$S = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}, \quad S^{R_1} = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & 1 & 1 \\ \omega & \omega^2 & 1 \\ \omega & 1 & \omega^2 \end{bmatrix}, \quad S^{R_2} = \begin{bmatrix} 1 & 1 & 1 \\ \omega^2 & 1 & \omega \\ \omega^2 & \omega & 1 \end{bmatrix}$$

S, S^{R_1}, S^{R_2} diagonalise C, CB, CB^2 respectively

ii) $N=5$

C, CB, CB^2, CB^3, CB^4 are diagonalised by $S, S^{R_2}, S^{R_4}, S^{R_1}, S^{R_3}$ respectively

N even

$N=4$

$$S = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}, \quad S^{R_{1/2}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ i^{1/2} & i^{3/2} & i^{5/2} & i^{7/2} \\ -1 & 1 & -1 & 1 \\ i^{1/2} & i^{7/2} & i^{5/2} & i^{3/2} \end{bmatrix}$$

$$S^{R_{3/2}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ i^{3/2} & i^{5/2} & i^{7/2} & i^{1/2} \\ -1 & 1 & -1 & 1 \\ i^{3/2} & i^{1/2} & i^{7/2} & i^{5/2} \end{bmatrix}$$

CB is diagonalised by $S^{R_{3/2}}$.

References:

1. Alladi Ramakrishnan, "L-matrix theory or the Grammar of Dirac Matrices", Tata McGraw-Hill, Bombay (1972) (See Appendix I).

2. A.R. Tekumalla, Matscience Preprint, 1974.

E. A THEOREM ON THE LIMITING PROPERTY OF A STOCHASTIC MATRIX*

1. Introduction:

Let R be an n x n matrix whose non diagonal elements are all positive that is $r_{ij} \geq 0 \; ; \; i \neq j$ and whose diagonal elements are equal to their respective column sums so that each column adds up to o, that is $\sum_i r_{ij} = 0$ for all j. We shall refer to this as a column stochastic matrix. It is well known that[1]

$$\lim_{t \to \infty} e^{Rt} \underline{\pi} = \underline{\pi}_0 \qquad (1)$$

where $\underline{\pi}$ is an arbitrary probability vector (that is, its elements are all ≥ 0 and odd upto 1) and $\underline{\pi}_0$ (also a probability vector) is the eigenvector of R corresponding to eigenvalue o. (We assume here that the eigenvalue o occurs with multiplicity 1 as for instance, when R is irreducible). This is the so called ergodic property of a stochastic matrix.

Recently, Professor Alladi Ramakrishnan suggested it would be interesting to study the approach to ergodicity that is to study how the elements of $\underline{\pi}$ approach the elements of the stationary vector $\underline{\pi}_0$ as $t \to \infty$: whether the elements of $\underline{\pi}$ approach the stationary values monotonically or whether they can cross it.

We shall here answer this question and establish that

---

* A.R.Tekumalla, presented at Matscience Conference on Probability Theory and Stochastic Processes, Bangalore, 1973.

when the eigenvalues of R are all real then after sufficiently large t, the stationary (limiting) values are reached monotonically and when the eigenvalues of R are complex we have a damped oscillation about the limiting value.

2. <u>The Ergodic Property of a Stochastic Matrix</u>:

Frist we shall investigate how the elements of e^{Rt} behave as $t \to \infty$ and then draw the necessary conclusions. We shall use a modification of the Sylvester interpolation procedure for functions of matrices by which a matrix can be expressed as a linear combination of mutually orthogonal idempotent matrices of rank 1.

The matrix R can be diagonalised (or Jordanised) by a similarity transformation. Hence

$$\Lambda = U^{-1} R U \tag{2}$$

or

$$R = U \Lambda U^{-1} \tag{3}$$

where Λ is the diagonal (or Jordan) form of R and U is the matrix which diagonalises R . We shall consider here the case when R is diagonalisable (because the non-diagonalisable case, while it is slightly more complicated, can be dealt with in a similar manner and yields the same results). In this case, Λ has the form

$$\Lambda = \mathrm{Diag}\,\{\lambda_1, \lambda_2, \ldots, \lambda_n\} \tag{4}$$

where λ_i are the eigenvalues of R . Since R is a stochastic matrix, one eigenvalue say $\lambda_1 = 0$ and $\lambda_2, \lambda_3, \ldots$ all have real parts negative. The U matrix can be written as

$$U = (\underline{u}_1 \ \underline{u}_2 \ \ldots \ \underline{u}_n) \tag{5}$$

where $\underline{u}_K$ is a column vector which is an eigenvector of R corresponding to eigenvalue λ_K . Further

$$U^{-1} = V = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \tag{6}$$

where v_K are the corresponding row eigenvectors $\underline{u}_K$ and v_K satisfy

$$v_i \underline{u}_j = \delta_{ij} \tag{7}$$

and $\underline{u}_i v_j$ are all n x n matrices of rank 1 whose range is the vector $\underline{u}_i$. Hence eq.(3) can be rewritten as

$$R = (\underline{u}_1 \ \underline{u}_2 \ \ldots \ \underline{u}_n) \left\{ \begin{bmatrix} \lambda_1 & 0 & \\ & 0 & \\ & & \ddots \end{bmatrix} + \begin{bmatrix} 0 & \lambda_2 & 0 \\ & & \ddots \\ & & \end{bmatrix} + \ldots \right\} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

or

$$\tag{8}$$

$$R = \sum_K \lambda_K \underline{u}_K v_K \tag{9}$$

Note that $\underline{u}$ is a column and v is a row so that $\underline{u}v$ is an n x n matrix of rank 1.

Now exponentiating eq.(3)

$$e^{Rt} = Ue^{\Lambda t}U^{-1} \tag{10}$$

and using the same procedure

$$e^{Rt} = \sum_{K} e^{\lambda_K t} \underline{u}_K v_K \tag{11}$$

Since R is column stochastic and irreducible, $\lambda_1 = 0$ and all the other λ_K have real parts $-ve$. Further,

$$v_1 = [1\ 1\ \cdots\ 1] \tag{12}$$

and

$$\underline{u}_1 = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \tag{13}$$

where $a_i \geq 0$ and $\sum a_i = 1$. That is $\underline{u}_1$ is a probability vector. Therefore eq.(11) gives

$$\lim_{t\to\infty} e^{Rt} = \underline{u}_1 v = \begin{bmatrix} a_1 & a_1 & \cdots & a_1 \\ a_2 & a_2 & \cdots & a_2 \\ \vdots & & & \\ a_n & a_n & \cdots & a_n \end{bmatrix} \tag{14}$$

Thus as $t \to \infty$, e^{Rt} approaches a stationary matrix and

$$(e^{Rt})_{ij} \Rightarrow a_i \tag{15}$$

These results can easily be generalised to row stochastic and doubly stochastic matrices.

III. Approach to Ergodicity of a Stochastic Matrix:

Now we come to the more interesting problem of how the elements of e^{Rt} (where R is column stochastic) approach the limiting values.

THEOREM 1:

If R has all its eigenvalues real, then, at sufficiently large t, the elements $(e^{Rt})_{ij}$ approach the values a_i monotonically as $t \to \infty$. (We note here in particular that when R is symmetric all its eigenvalues are real).

Proof: First we note that at t = o,

$$e^{Rt} = 1$$

that is the diagonal elements are 1 and off diagonal elements, o.

Now going back to eq.(11)

$$(e^{Rt})_{ij} = \sum_{k=1}^{n} e^{\lambda_k t} (u_k)_i (v_k)_j \tag{16}$$

Since λ_k are all real and R is real, u_k and v_k can be chosen real. Further, since R is stochastic, $\lambda_1 = 0$ and other λ_k −ve. Hence,

$$(e^{Rt})_{ij} = a_i + \sum_{k=2}^{n} e^{\lambda_k t} (u_k)_i (v_k)_j \tag{17}$$

At $t = o$, the K^{th} term is $(\underline{u}_K)_i (v_K)_j$. As t increases, since λ_K are all $-ve$, $e^{\lambda_K t}(\underline{u}_K)_i (v_K)_j$ goes monotonically to o. Hence $(e^{Rt})_{ij}$ is a sum of n terms ~~are~~ one of which is a_i and all the others monotically approach o. By superposing these terms, we find that the elements of $(e^{Rt})_{ij}$ after a sufficiently large t approach the limiting values a_i monotically (and do not cross the stable value again). However when the initial values of the terms in eq.(17) are not all of the same sign, the element may cross the stable value a finite number of times.

THEOREM 2:

When the eigenvalues of R are complex, the elements of e^{Rt} approach the limiting values as damped oscillations about the limiting values.

Proof: Again we start with the expansion

$$e^{Rt} = \sum_K e^{\lambda_K t} \underline{u}_K v_K \tag{18}$$

where $\lambda_K, \underline{u}_K, v_K$ may now be complex. However, we know that e^{Rt} is real, since R and t are real. So we shall rewrite the equation to make it manifestly real.

First we note that since R is real, its characteristic equation has real coefficients and therefore imaginary roots will enter in pairs. Further, if $\underline{u}_K$ is an eigen vector corresponding to the eigen value λ_K, that is

$$R\underline{u}_k = \lambda_k \underline{u}_k \tag{19}$$

then, since R is real,

$$R\underline{u}_k^* = \lambda_k^* \underline{u}_k^* \tag{20}$$

that is $\underline{u}_k^*$ is also an eigenvector of R corresponding to λ_k^*. Here the star denotes just complex conjugation. A similar statement holds for the row eigenvectors v_k

Hence

$$\begin{aligned} e^{Rt} &= \frac{1}{2}\sum_{k=1}^{n}\left(e^{\lambda_k t}\underline{u}_k v_k + e^{\lambda_k^* t}\underline{u}_k^* v_k^*\right) \\ &= \sum_{k=1}^{n} \operatorname{Re} e^{\lambda_k t}\underline{u}_k v_k \end{aligned} \tag{21}$$

Elementwise,

$$\left(e^{Rt}\right)_{ij} = a_i + \sum_{k=2}^{n} \operatorname{Re} e^{\lambda_k t} (\underline{u}_k)_i (v_k)_j \tag{22}$$

where, of course, a_i is real. We now write

$$\lambda_k = -x_k + i y_k \tag{23}$$

$$(\underline{u}_k)_i = r_{ik} e^{i\varphi_{ik}} \tag{24}$$

$$(v_k)_j = r'_{kj} e^{i\varphi'_{kj}} \tag{25}$$

where x_k is now positive and $x_k, y_k, r_{ik}, r'_{kj}, \varphi_{ik}, \varphi'_{kj}$ are all real. Then

$$(e^{Rt})_{ij} = a_i + \sum_{k=2}^{n} \mathrm{Re}\, e^{-x_k t} r_{ik} r'_{kj} e^{i(yt + \varphi_{ik} + \varphi'_{kj})}$$

$$= a_i + \sum_{k=2}^{n} r_{ik} r'_{kj} e^{-x_k t} \cos(y_k t + \varphi_{ik} + \varphi'_{kj}) \qquad (26)$$

Hence $(e^{Rt})_{ij}$ is a sum of one constant term a_i and $n-1$ terms each of which is a damped oscillation about the mean value o . The cosine factor causes each term to oscillate with a period $2\pi/y_k$ and the exponential term is a damping factor which makes the term finally o. The superposition of these oscillating terms is strictly analogous to the phenomenon of beats in accoustics. The result is, in general, an amplitude modulated wave and when this is superposed on the constant term, the result is an amplitude modulated oscillation about a_i . Thus the limiting value a_i is crossed an infinite number of times.

IV. <u>Action on an Arbitrary Probability Vector</u>:

Again, using the same decomposition,

$$(e^{Rt}\Pi)_i = (e^{Rt})_{ij}\Pi_j = \sum_j \sum_k e^{\lambda_k t} (u_k)_i (v_k)_j \Pi_j$$

$$= a_i + \sum_{j=1}^{n} \sum_{k=2}^{n} e^{\lambda_k t} (u_k)_i (v_k)_j \Pi_j \qquad (27)$$

which is again a linear combination of terms which monotonically approach a_i or oscillate about a_i according as the

λ_K are real or complex and hence, by the same argument as above $(e^{Rt}\underline{\Pi})_i$ monotonically approaches a_1 after a sufficiently large t or oscillates about a_1 according as λ_K are all real or complex.

We may also rewrite the last equation as

$$(e^{Rt}\underline{\Pi})_i = \sum_K e^{\lambda_K t}(v_K, \underline{\Pi})(\underline{u}_K)_i$$

$$= a_i + \sum_{K=2}^{n} \mathrm{Re}\, e^{\lambda_K t}(v_K, \underline{\Pi})(\underline{u}_K)_i \qquad (28)$$

where the inner product $(v_K, \underline{\Pi})$ is a number. Since the monotonicity or the periodicity of the elements depends on $e^{\lambda_K t}$ [$(v_K, \underline{\Pi})$ at best contributes a phase factor and decides the amplitude)], we arrive again at the same conclusions.

F. GENERALIZED LUCAS POLYNOMIALS AND FUNCTIONS OF MATRICES*

1. Introduction:

In a recent paper, Barakat and Baumann[1] indicated the importance of generalized lucas polynomials in a variety of physical problems[2,3] and suggested that it is desirable to obtain them in a closed form. In Barakat's notation these polynomials are conveniently defined through a set of difference equations given by

$$U_{n+N}^{(N)}(a_1, a_2, \ldots, a_N) = \sum_{i=1}^{N} (-1)^{i-1} a_i U_{n+N-i} \qquad (A_N)$$

together with the N-initial conditions

$$U_i^{(N)} = \delta_{N-1,i} \; ; \quad i = 0, 1, \ldots, N-1 \qquad (B_N)$$

where δ_{ij} is the Kronecker symbol. Here, in fact, we solve the problem even when the difference equation satisfies a set of arbitrary initial conditions

$$U_i^{(N)} = b_i \; ; \quad i = 0, 1, \ldots, N-1 \qquad (C_N)$$

Further, Barakat and Baumann show that the M^{th} power of on N x N matrix is related to its $(N-1)^{th}$ to 0^{th} powers as follows

---

* I.V.V.Raghavacharyulu and A.R.Tekumalla, Jour.Math.Phys., **13**, 321 (1972)

$$X^M = \left[U_M^{(N)}\right] X^{N-1} + \left[-a_2 U_{M-1}^{(N)} + a_3 U_{M-2}^{(N)} - \cdots \pm a_N U_{M-N+1}^{(N)}\right] X^{N-2} + \left[a_3 U_{M-1}^{(N)} - a_4 U_{M-2}^{(N)} + \cdots \pm a_N U_{M-N+2}\right] X^{N-3} + \cdots + \left[\pm a_N U_{M-1}^{(N)}\right] I \tag{1.1}$$

we rewrite the term in the i^{th} parenthesis as $A_M^N(J)$ so that

$$X^M = \sum_{J=1}^{N} A_M^{(N)}(J)\, X^{N-J} \tag{1.2}$$

and give expressions in closed form for $A_M^{(N)}(J)$ which are linear translates of the polynomials defined above. We also give a different expression for any function of X when its eigenvalues are known.

The usual method[4] of solving eq.(A_N) is by the method of generating functions, making use of the roots of the characteristic equation

$$F(x) \equiv x^N - a_1 x^{N-1} + \cdots \pm a_N = 0 \tag{D_N}$$

Obviously, it is difficult to solve for the roots of eq.(D_N) in terms of the coefficients $(a_i | i = 1, 2, \ldots, N)$.

So, here the solutions of eqs. (A_N) and (C_N) are obtained in terms of the coefficients $(a_i | i = 1, 2, \ldots, N)$ themselves directly without solving the characteristic equation.

In section 2, we summarise conveniently the available information about the explicit form of $U^{(2)}(a_1, a_2)$ and $U^{(3)}(a_1, a_2, a_3)$ and extend the generalization of $V^{(2)}(a_1, a_2)$ to $V^{(N)}(a_1, a_2, \ldots, a_N)$ and give explicit expressions in closed form for $U^{(N)}_m(a_1, a_2, \ldots, a_N)$ and $V^{(N)}_n(a_1, a_2, \ldots, a_N)$. In section 3, we prove that indeed the generalized Lucas U and V polynomials have the form given in section 2. We also give expressions in closed form for $A^N_M(J)$ and expression for any function of X when its eigen values are known.

2. Lucas Polynomials for N = 2,3 :

When N = 2, the Lucas polynomials[5,6,7] are obtained as the solution of difference eq. (A_2) satisfying the initial conditions eq. (B_2) . The solution for eqs. (A_2) and (B_2) is

$$U^{(2)}_{n+1}(a_1, a_2) = \sum_\nu (-1)^\nu \binom{n-\nu}{\nu} a_1^{n-2\nu} a_2^\nu , \tag{2.1}$$

the series terminating when the exponent of a_1 or a_2 ~~terms~~ turns negative.

The general term of eq.(3.1) can be more conveniently written as

$$(-1)^{n-\Sigma\lambda} \frac{(\lambda_1+\lambda_2)!}{\lambda_1!\,\lambda_2!} a_1^{\lambda_1} a_2^{\lambda_2} \tag{2.2}$$

where $\lambda_1 + 2\lambda_2 = n$.

For N = 2 the Lucas $V^{(2)}(a_1, a_2)$ polynomials are defined by the difference equation A_2 satisfying the initial conditions and $V_0^{(2)} = 2, V_1^{(2)} = a_1$.

For N = 3, one has to solve the difference eq.(A_3) satisfying the initial conditions (B_3). The possible general solution suggested in closed form by Barakat and Baumann for $U^{(3)}$ is

$$U_{n+2}^{(3)}(a_1, a_2, a_3) = U_{n+1}^{(2)}(a_1, a_2) + \sum_{K=1} \sum_{K=l} (-1)^{l-K} \binom{l}{k} \binom{n-l-K}{l} a_1^{n-2l-K} a_2^{l-K} a_3^{K} \tag{2.3}$$

again with the understanding that all the exponents are ≥ 0

Making use of eq.(2.1) the general term of eq.(2.3) can be written as

$$(-1)^{n-\Sigma\lambda} \frac{(\lambda_1 + \lambda_2 + \lambda_3)!}{\lambda_1! \lambda_2! \lambda_3!} a_1^{\lambda_1} a_2^{\lambda_2} a_3^{\lambda_3} \tag{2.4}$$

where $\lambda_1 + 2\lambda_2 + 3\lambda_3 = n$.

Assuming the correctness of eqs. (2.1) and (2.3), it is easy to guess the general solution of the difference equations of generalized Lucas Polynomials. Indeed, we have the following:

THEOREM 1:

The general solution to the difference equation

$$U^{(N)}_{n+N}(a_1, a_2, \cdots, a_N) = a_1 U^{(N)}_{n+N-1} - a_2 U^{(N)}_{n+N-2} + \cdots \pm a_N U^{(N)}_{n}$$

satisfying the initial conditions

$$U_0^{(N)} = U_1^{(N)} = \cdots = U_{N-2}^{(N)} = 0, \quad U_{N-1}^{(N)} = 1,$$

is

$$U^{(N)}_{n+N-1}(a_1, a_2, \cdots a_N) = h_n(a_1, a_2, \cdots, a_N, \cdots)$$

with $a_{N+1} = a_{N+2} = \cdots = 0$ (2.5)

where h_n are the symmetric functions called homogeneous[9] product sums of weight n.

THEOREM 2:

The general solution of the difference equation

$$V^{(N)}_{n+N}(a_1, a_2, \cdots, a_N) = a_1 V^{(N)}_{n+N-1} - a_2 V^{(N)}_{n+N-2} + \cdots \pm a_n V^{(N)}_{n}$$

satisfying the initial conditions $V_0 = N$, $V_i = S_i;\ i = 1, 2, \cdots, N-1$ is $V^{(N)}_n(a_1, a_2, \cdots, a_N) = S_n(a_1, a_2, \cdots, a_N, \cdots)$

~~is~~ with $a_{N+1} = a_{N+2} = \cdots = 0$ (2.6)

where S_n is the one part symmetric function.

The expressions for $U^{(N)}_n$ and $V^{(N)}_n$ given by eqs. (2.5) and (2.6) respectively enable us to obtain, simply, the

Lucas polynomials. The first few h_n and s_n (n = 1 to 6) are given in Ref. 8 from which all the results that have been given by Barakat and Baumann for N = 3(8), 4(9) and 5(10) are easily obtained. Hence, in tabulating $U_n^{(N)}$ and $V_n^{(N)}$ it is sufficient to give h_n and s_n only.

The theorems are proved in the next section.

3. Solutions of Difference Equations:

We know from the well known connection between the symmetric functions[8] a_n and h_n that

$$\frac{1}{1-a_1x+a_2x^2-\dots\pm a_rx^r\mp\dots} = 1+h_1x+h_2x^2+\dots \quad (3.1)$$

where, in the denominator on the left hand side, we can assume, without loss of generality, an infinite number of terms. Expanding the left hand side by the multinomial theorem[8], we obtain (see note)

$$h_n = \sum_{\lambda_1,\lambda_2,\dots}(-1)^{n-\Sigma\lambda}\frac{(\Sigma\lambda)!}{\lambda_1!\,\lambda_2!\cdots\lambda_n!}a_1^{\lambda_1}a_2^{\lambda_2}\cdots a_n^{\lambda_n} \quad (3.2)$$

where $\sum i\lambda_i = n$ (Note: $(-1)^{n-\Sigma\lambda} = (-1)^{\lambda_2+\lambda_4+\cdots}$)

The generating function for solution of (A_N) with (B_N) is $f(x)=\sum_{n=0}^{\infty}U_nx^n = x^{N-1}/(1-a_1x+a_2x^2-\cdots)$

Hence, the general solution of the difference eq.(A_N) with initial condition B_N can be written down by ~~taking the~~ coefficient of x^n in $x^{N-1}/(1-a_1x+a_2x^2-\dots\pm a_Nx^N)$

In particular, as the solution for the Lucas polynomials we obtain

$$U_{n+N-1}^{(N)} = \sum_{\lambda_1,\lambda_2,\ldots,\lambda_N} (-1)^{n-\Sigma\lambda} \frac{(\Sigma\lambda)!}{\lambda_1!\lambda_2!\cdots\lambda_N!} a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_N^{\lambda_N} \tag{3.3}$$

and $\sum i\lambda_i = n$ which proves theorem 2.

In general, the generating function[4] for the solution of eqn.(A_N) with (C_N) is given by

$$f(x) = \sum_{n=0}^{\infty} U_n x^n = \frac{W_0 + W_1 x + \cdots + W_{N-1} x^{N-1}}{1 - a_1 x + a_2 x^2 - \cdots \pm a_N x^N} \equiv \frac{n(x)}{p(x)} \tag{3.4}$$

where $W_n = b_n - a_1 b_{n-1} + \cdots \pm a_n b_0$, for $n = 0, 1, \ldots, N-1$. and $W_n = 0$ for $n \geq N$

We now proceed to generalize the Lucas V polynomials to N^{th} order. Note that eq. (3.4) gives the solution to eq.(A_N) with any initial conditions (C_N). In particular, for $N = 2$ and $V_0 = 2$, $V_1 = a_1$ we get Lucas $V^{(2)}$ polynomials which are identifiable as the one part symmetric functions S_n with $a_3 = a_4 = \cdots = 0$. We now define the generalized Lucas $V^{(N)}$ polynomials as the solution of eq.(A_N) with the initial conditions

$$V_0 = N; V_i = S_i ; i = 1, \ldots, N-1 \tag{3.5}$$

Solving for W_n making use of the initial conditions eq.(3.5) we obtain that

$m(x) = F'(\frac{1}{x})\, x^{N-1}$ where $F'(x)$ is the derivative of $F(x)$. Further, we know that

$$\frac{F'(x)}{F(x)} = \frac{1}{x-\alpha_1} + \frac{1}{x-\alpha_2} + \cdots + \frac{1}{x-\alpha_N}$$

where $(\alpha_i \mid i=1,2,\ldots,N)$ are the roots of $F(x)=0$.
Hence, in a well known way we obtain

$$f(x) = \frac{1}{x}\,\frac{F'(1/x)}{F(1/x)}$$

giving S_n as the solution of the difference equation. It is not difficult to see that the general expression for the generalized Lucas $V_n^{(N)}$ polynomials in closed form is

$$V_n^{(N)} = \sum_{\substack{\lambda_1,\lambda_2,\ldots,\lambda_N \\ \sum i\lambda_i = n}} (-1)^{n-\sum\lambda}\left(\frac{(\sum\lambda-1)!\,n}{\lambda_1!\,\lambda_2!\cdots\lambda_N!}\right) a_1^{\lambda_1} a_2^{\lambda_2}\cdots a_N^{\lambda_N}$$

with (3.6)

and again we can write a simple algorithm to find $V^{(N)}$

$$V_n^{(N)}(a_1,a_2,\ldots,a_N) = S_n(a_1,a_2,\ldots,a_{N+1},\ldots)$$

with $a_{N+1} = a_{N+2} = \cdots = 0$.

THEOREM 7:

If $\alpha_1, \alpha_2, \ldots, \alpha_N$ are the roots of the equation

$x^N - a_1 x^{N-1} + \cdots = 0$ then

$$U_n^N = \begin{vmatrix} \alpha_1^n & \alpha_2^n & \cdots & \alpha_N^n \\ \alpha_1^{N-2} & \alpha_2^{N-2} & \cdots & \alpha_N^{N-2} \\ \vdots & & & \\ \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ 1 & 1 & \cdots & 1 \end{vmatrix} \div \begin{vmatrix} \alpha_1^{N-1} & \alpha_2^{N-1} & \cdots & \alpha_N^{N-1} \\ \alpha_1^{N-2} & \alpha_2^{N-2} & \cdots & \alpha_N^{N-2} \\ \vdots & & & \\ \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ 1 & 1 & \cdots & 1 \end{vmatrix}$$

and $V_n^{(N)} = \alpha_1^n + \alpha_2^n + \cdots + \alpha_N^n$

When repeated roots occur we use d'Hospital's rule to simplify these expressions. $U_n^{(N)}$ and $V_n^{(N)}$ obviously satisfy the initial conditions (BN) and (3·5) respectively.

and for .

The first part of the theorem follows immediately from the expression

$$\frac{1}{p(x)} = \sum_{i=1}^{N} \frac{\alpha_i^{N-1}}{\prod_{j=1}^{\prime N}(\alpha_i - \alpha_j)} \frac{1}{1-\alpha_i x}$$

by expanding them in power series and collecting the coefficient of x^{n-N+1}. The second part is obvious. When N = 2, the above expressions reduce to

5

$$U_n^{(2)} = \frac{\alpha_1^n - \alpha_2^n}{\alpha_1 - \alpha_2}; \quad V_n = \alpha_1^n + \alpha_2^n$$

We now give an expression for any function of X when the eigen values of X are known. From the analysis of chapter 5, it is clear that if α_i are the eigenvalues of X, then

$$f(X) = a_1 X^{N-1} + a_2 X^{N-2} + \cdots + a_N$$

where

$$a_i = \det \begin{vmatrix} \alpha_1^{N-1} & \alpha_2^{N-1} & \cdots & \alpha_N^{N-1} \\ \alpha_1^{N-2} & \alpha_2^{N-2} & \cdots & \alpha_N^{N-2} \\ \vdots & & & \\ \alpha_1^{N-i+1} & \alpha_2^{N-i+1} & \cdots & \alpha_N^{N-i+1} \\ f(\alpha_1) & f(\alpha_2) & \cdots & f(\alpha_N) \\ \alpha_1^{N-i-1} & \alpha_2^{N-i-1} & \cdots & \alpha_N^{N-i-1} \\ \vdots & & & \\ \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ 1 & 1 & \cdots & 1 \end{vmatrix} \div \det \begin{vmatrix} \alpha_1^{N-1} & \alpha_2^{N-1} & \cdots & \alpha_N^{N-1} \\ \alpha_1^{N-2} & \alpha_2^{N-2} & \cdots & \alpha_N^{N-2} \\ \vdots & & & \\ \alpha_1 & \alpha_2 & \cdots & \alpha_N \\ 1 & 1 & \cdots & 1 \end{vmatrix}$$

where the determinant in the numerator is obtained by replacing the i^{th} row of the ~~denominatior~~ denominator by the indicated values. The same comment as above applies when there are repeated eigenvalues. We now give without proof expressions in closed form for the linear translates of generalised Lucas polynomials $A_M^N(J)$ defined by eq.(1.2)

$$A_M^{(N)}(J) = \sum (-1)^{\sum i\lambda_i + \sum \lambda_i} \frac{\left(\sum_{i=1}^{N} \lambda_i - 1\right)!}{\lambda_1!\lambda_2!\cdots\lambda_N!} \sum_{j=J}^{N} \lambda_j a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_N^{\lambda_N}$$

$$\sum_{i=1}^{N} i\lambda_i = M - N + J$$

The solution of the difference equation (A_N) with arbitrary initial conditions (C_N) can be written in terms of these coefficient $A_M^{(N)}(J)$ as

$$U_m^{(N)} = \sum_{J=1}^{N} A_M^{(N)}(J)\, b_{N-J}$$

A further extension of the Lucas V polynomials is possible by writing

$$V_{n,m}^{N} = S_{n+m,\,m} \tag{3.7}$$

with $a_{N+1} = a_{N+2} = \cdots = 0$, and $S_{\omega,m}$ are the 2 part, 3 part etc. symmetric functions of given weight, given by

$$S_{\omega,m} = \left\{ \begin{array}{l} \sum (-1)^{\omega+m-\Sigma\lambda-1} \\ \times \dfrac{(\Sigma\lambda-1)!\left[\lambda_m + \binom{m+1}{1}\lambda_{m+1} + \cdots\right]}{\lambda_1!\,\lambda_2!\cdots\lambda_\omega!} \\ \times\, a_1^{\lambda_1} a_2^{\lambda_2} \cdots a_\omega^{\lambda_\omega} \end{array} \right.$$

where $\Sigma i\lambda_i = \omega$ and $\omega \geq m$.

$V_{n,m}^{(N)}$ satisfied the difference equation

$$V_{n,m}^{(N)}(a_1, a_2, \ldots, a_N) = a_1 V_{n-1,m} + \cdots \pm a_N V_{n-N,m}^{(N)}$$

with the initial conditions, $V_{i,m}^{(N)} = S_{i+m,m};\ i = 1, 2, \ldots, N$

For $m = 1$, we get the generalized Lucas $V^{(N)}$ polynomials with the initial conditions $V_i^{(N)} = S_i;\ i = 1, 2, \ldots, N$.

Further it is very well known that h_n, S_n could be extended to what are called Schur functions. Similar considerations as are developed in the paper can be generalized to apply to these functions too.

Reference:

1. R.Barrat and E.Baumann, J.Math.Phys., 10, 1474 (1969).
2. V.Mower, IRE Trans.Circuit Theory, 11, 232 (1964).
3. S.Basil Math. Mag. 37, 83 (1964).
4. Martin sen, Elementary Combinatorial Analysis, Gordon & Breac New York, 1969.
5. E.Lucas Theories des Nombres (Gauthier-Villars, Paris, 1891).
6. A.Herpin Compt. Rend. Acad. Sci., Paris, 225, 17 (1947).
7. R.Baraka J.Math.Phys., 43, 325 (1964).
8. P.A.Macma n, Combinatory Analysis, (Chelsea Publishing Company New York, 1960).
9. D.E.Littlewood, The theory of group characters and matrix representations of groups (Clarendon, Oxford, 1950) 2nd Ed.

10. I.V.V.Raghavacharyulu and A.R.Tekumalla, Jour.Math.Phys., 13, 321 (1