JOURNAL OF COMPUTATIONAL PHYSICS 69, 420–433 (1987)

On the Solution of Matrix Equations, Example: Application to Invariant Equations

GILLES LABONTÉ

Department of Mathematics and Computer Science, Royal Military College of Canada, Kingston, Ontario, Canada K7K 5L0

Received February 10, 1986; revised August 4, 1986

We explain a method for constructing the representations of matrix rings specified by sets of equations for generating elements. An illustration is provided by its application to the resolution of the problem of finding generalized gamma-matrices, i.e., the matrix coefficients in Lorentz invariant wave equations. A computer program, which carries out such constructions, has been written. © 1987 Academic Press, Inc.

1. INTRODUCTION

We present a constructive method for solving sets of polynomial equations in noncommuting variables. As far as we know, no such general method existed before.[1]

This method is a generalized version of one which was initially developed to solve a particular algebraic problem: that of constructing generalized Dirac Γ-algebras. The method has been described before in this particular context, and a first problem has been solved completely through this procedure in reference [1]. In this first application, it appeared clearly that this algorithm requires the keeping track of a relatively large number of vectors and their properties with respect to a relatively large number of matrices. Thus the use of a computer to help store and retrieve the data was soon considered with interest.

Because of the symbolic nature of the data and the manipulations involved, LISP was used. Initially, an interactive program was written to assist in the carrying out of the solution. However, the whole method of solution is such a constructive procedure that can be completely carried out by the computer. Indeed, a modified version of our program constructs automatically the representations of finite groups, starting from a presentation (i.e., a set of equations for the generators of the group); this work will be described elsewhere.

[1] None of the main computer algebra system, which were demonstrated at the recent (July 86) Sigsam conference in Waterloo, is able to solve such equations for want of an algorithm.

0021-9991/87 $3.00

The method, as presented in Section 2, is quite general: it can be used to solve any problem consisting of finding representations for matrices satisfying a given set of equations. In the following Section 3, we show how it can be used, in particular, to deal with the construction of any generalized Γ-algebra. A problem of this type, which was solved by other means before [3], has then been selected as an example; we list the eleven matrix equations to be solved in this case. In Section 4, the evolution of the construction of the solution, as produced by the computer, is described. A brief discussion follows in the last section.

2. The Method

Let us be given a set of algebraic equations (SEQS), to be satisfied by some matrices. We want to construct all possible representations of matrices which will satisfy these equations.

Let V stand for any vector in a representation space S for the matrices. While such a representation is being constructed, we shall let SN denote the set of known vectors, among those of S, which until otherwise established are to be considered as linearly independent. Thus at the beginning, SN may consist only of V.

Clearly, an equation between matrices holds if and only if an identity is obtained whenever each side of it is applied to any one vector of the representation space S. Let us then apply in turn each equation of SEQS on the vector V. In doing so, it will happen that the result of applying some matrix M on V is unknown. Since, however, it is certain that the resulting vector is in S, we represent it by the name MV, and add this name to the set SN. Thus, vector names are created as the need for the existence of the vectors which they represent is ascertained. A vector name as $MQRV$ will clearly indicate that this vector was generated from V, through the successive application of the matrices represented by the letters R, Q, and M.

The data pertaining to the effect of the matrices on the vectors is stored in relations of the type: "the matrix M on the vector $V =$ some sum of vectors." Such an expression will be considered trivial if it is of the particular form "M on $V = MV$."

The application of any one equations on V will either result in an identity or in a linear relation between some vectors. The first case will occur when enough is already known about the vector V that no additional property is implied by this particular equation. In the second case, one vector can be expressed as a linear combination of the others and thus, consequently, the name of this vector should be removed from SN. The rule governing the choice of which vector this will be is derived from an order relation between vectors which is described below.

We shall say that $V1 > V2$ if either the name of $V1$ has more letters than that of $V2$, or if they have the same number of letters, when the name of $V1$ comes after that of $V2$ with respect to alphabetical ordering. (We note that for our purposes the choice of the order relation for vector names of equal lengths is completely arbitrary. Any other one will do; in fact, in that case an order relation is not necessary.)

Thus, in a linear relation, the vector to be expressed in terms of the other ones shall always be the largest one of those involved in this equation. This is a crucial part of the construction method: larger vectors are always eliminated in favor of smaller ones. This ensures that the tree structure of the set SN always remains connected.

Whenever an expression is found for a vector named W, this vector must be replaced by the corresponding expression everywhere in the data accumulated until then. Furthermore, all vectors with names ending with the suffixe W are affected. Indeed, if for example, it is found that $W = A + B$, then a vector named CDW should be substituted for everywhere by the vector resulting from the application of C and D on $A + B$. We remark that in doing so there can appear some new vector equations. This will be the case whenever there is, in the data, a nontrivial relation of the form "a matrix M on a vector generated from $W =$ some sum of vectors." Such equations of course are dealt with as all other equations.

Once all equations of SEQS have been applied on V, the set SN will in general contain other vectors than V. Since the equations should be true when applied on these vectors also, one should repeat with the next smallest vector of SN what has just been done with V, and so on. The construction is finished when no more new vectors are produced in this process, and when all equations have been applied on all vectors of SN.

Those vectors remaining in SN at the end of the construction are vectors which on one hand had to be generated to ensure the satisfaction of the equations by smaller vectors, and on the other hand, were never required to be linearly dependent for these equations to hold. Thus the vectors of SN form a basis for a representation of the matrices. This representation constitutes furthermore a general solution to the problem in the sense that all the possible irreducible representations containing the generating vector V are either exactly the representation constructed or are particular cases of it (i.e., representations obtained by postulating, in an ad hoc fashion, one or more additional linear relations between the vectors of SN). The following treatment of an example will help illustrate the above discussion.

3. Example of Application

3.1. *General Problem*

The description of the motion of high-energy particles in Quantum Mechanics is usually done with the help of partial differential equations which are covariant under the Poincaré group. The mathematical requirements for wave equations to agree with the Special Theory of Relativity were formulated very early in the development of Quantum Mechanics. (An excellent survey on this and the following matters has been presented by Wightman in [3]). Here is a brief description of these requirements.

It is postulated that free particles should display space-time translational invariance, and since all differential equations can be written as first order systems, all these wave equations can be written as the constant coefficients system:

$$(-i\Gamma_\mu \partial^\mu + m)\,\psi(x) = 0. \tag{3.1}$$

The variables x^0 and x^i, $i=1, 2, 3$, are, respectively, the time and space variables and $\partial^\mu = \partial/\partial x^\mu$. The constant metric tensor $g^{\mu\nu}$ is diagonal with $\mathrm{diag}(g^{\mu\nu}) = (1, -1, -1, -1)$. ψ has N components, Γ_μ and m are matrices of constants. The fact that the wave function ψ may have more than one component comes from the necessity to describe the spin which, with the mass, constitute the characteristic physical attributes of elementary particles.

The so called "standard equations" are those for which the latter physical concepts are incorporated as follows. (See, e.g., [4] or Chap. II of [3].)

(i) The free particle theory is required to be invariant for the orthochronous Lorentz group $L^\uparrow$. Thus, under a Lorentz transformation Λ, as the coordinates become $x' = \Lambda x$, the wave function ψ becomes $\psi'(x') = S(\Lambda)\,\psi(x)$, with $\Lambda \to S(\Lambda)$, a representation of $SL(2, C)$ ($SL(2, C)$, the covering group of the Lorentz group is considered because half integral values of the spin are allowed). Covariance of Eq. (3.1) is then ensured by requiring Γ_μ and m to transform respectively as a vector and a scalar, i.e.,

$$\begin{aligned} S(\Lambda)\,\Gamma_\mu S(\Lambda)^{-1} &= \Lambda_\mu^\nu \Gamma_\nu, \\ S(\Lambda)\,m S(\Lambda)^{-1} &= m. \end{aligned} \tag{3.2}$$

(ii) A free particle should have a unique mass. Thus, the matrix m can be taken simply as the positive real number m times the identity matrix. Furthermore, if p is the 4-momentum of a plane wave solution $\psi(x) = \exp(-ipx)\,\tilde{\psi}(p)$, the minimal equation for $(\Gamma \cdot p)$ will be [5],

$$(\Gamma \cdot p)^{n-2}[(\Gamma \cdot p)^2 - p^2] = 0 \tag{3.3}$$

with $n \in N$, $n \geqslant 2$. This ensures that $\tilde{\psi}(p)$ has support only on the hyperboloid $p^2 = m^2$.

(iii) A free particle should have a single spin. The solution $\tilde{\psi}(p)$ corresponding to a plane wave of momentum $p^i = 0$, $p^0 = m$, transforms under a representation of $SU(2)$. There is a single spin when this representation is irreducible; the spin being defined as the label $s = 0, \frac{1}{2}, 1, \ldots$ of this representation.

(iv) Quantum Mechanics requires that there exist a scalar product

$$(\varphi, \psi) = \int_\sigma d\sigma_\mu J^\mu \qquad \text{with } \sigma \text{ a time-like surface}$$

and $J^{\mu} = \varphi^{\dagger}\eta\Gamma^{\mu}\psi$ a conserved vector current. It should be nonnegative, in the sense $(\psi, \psi) \geqslant 0$ for all ψ which is a superposition of positive energy (i.e., $p^0 > 0$) plane waves. The Hermitianizing matrix η is such that $\eta^{\dagger} = \eta$, $(\eta\Gamma_{\mu})^{\dagger} = (\eta\Gamma_{\mu})$, $S(\Lambda)^{\dagger}\,\eta S(\Lambda) = \eta$. Its existence is linked with parity invariance; it requires the representation of $SL(2, C)$ to be self-conjugate.

It is known (see, e.g., Bhabha's Theorem 4 in reference [4]) that for such equations the representation will in fact be of the form

$$\mathscr{R} = \mathscr{R}(n_1, m_1) + \mathscr{R}(n_2, m_2) + \cdots,$$

with $\mathscr{R}(n, m)$ irreducible representation of $L^{\uparrow}$; n and m such that $n \geqslant m \geqslant 0$, being integers or half-odd integers depending on whether the physical spin is integer or half-odd integer. The existence of a matrix R, with the properties of a representative of space-time inversion can then be deduced from the form of the representation $\mathscr{R}$.

The problem of constructing wave equations with the properties described above can be stated algebraically as follows. Find 13 matrices corresponding to the generators of the infinitesimal Lorentz transformations: **J**, **N**; and the discrete symmetries: parity P and space-time inversion R; and four Γ-matrices with a Hermitianizing matrix η, these matrices being defined by the following equations. The commutations relations for **J** and **N**,

$$[J_i, J_j] = i\varepsilon_{ijk} J_k; \qquad [N_i, N_j] = -i\varepsilon_{ijk} J_k; \qquad [J_i, N_j] = i\varepsilon_{ijk} N_k \tag{3.4}$$

the commutation relations implied by Eq. (3.2)

$$[J_i, \Gamma_0] = 0; \qquad [N_i, \Gamma_0] = i\Gamma_i; \qquad [J_i, \Gamma_j] = i\varepsilon_{ijk}\Gamma_k, \tag{3.5}$$

$$[N_i, \Gamma_j] = i\delta_{ij}\Gamma_0. \tag{3.6}$$

The parity representative P satifies

$$[P, J_i] = 0; \qquad \{P, N_i\} = 0; \qquad [P, \Gamma_0] = 0 \tag{3.7}$$

with $P^2 = \varepsilon$ with $\varepsilon = +1$ for bosons, i.e., particles with integer spins and $\varepsilon = -1$ for fermions, i.e., half-odd integer spins. $\{A, B\}$ denotes the anticommutator of A and B. The space-time inversion matrix R has the properties:

$$[R, J_i] = 0; \qquad [R, N_i] = 0; \qquad \{R, \Gamma_{\mu}\} = 0; \quad R^2 = 1, \tag{3.8}$$

$$[R, P] = 0 \quad \text{for bosons} \qquad \text{and} \qquad \{R, P\} = 0 \quad \text{for fermions.}$$

The matrix η is such that η, ηJ_i, ηN_i, $\eta\Gamma_{\mu}$ are Hermitian with $(\eta P)^{\dagger} = \varepsilon(\eta P)$ and $(\eta R)^{\dagger} = \varepsilon(\eta R)$, with ε as above. Equation (3.3) gives rise to the Harish-Chandra conditions [5]:

$$\sum_{p(\mu,\nu,\rho,\ldots,\lambda)} [\Gamma_{\mu}\Gamma_{\nu} - g_{\mu\nu}]\,\Gamma_{\rho}\cdots\Gamma_{\lambda} = 0 \tag{3.9}$$

in which the last factor has $(n-2)$ terms, and the summation runs over all permutations p of the indices. The number "n" is called the Harish-Chandra number of the equation. (It is well known [6] that n must be larger than 2 when the spin is larger than $\frac{1}{2}$.)

The positivity condition on the scalar product implies

$$v^{\dagger}\eta\Gamma_0^n(\Gamma_0+1)v \geqslant 0 \qquad \text{for all } v \text{ in } C^N. \tag{3.10}$$

The unicity of the physical spin s implies that the plane wave solutions with $p_0=m$ and $p_i=0$, have only $(2s+1)$ nonzero components and the angular momentum $\mathbf{J}^2$ has value $s(s+1)$ on such solutions. This implies

$$[\mathbf{J}^2-s(s+1)]\,\Gamma_0^n(\Gamma_0+1)=0. \tag{3.11}$$

3.2. *Motivation and Partial Solution*

It is a well-known result for anyone who has encountered such relativistic wave equations that for spins larger than one, all the presently known equations are unstable when minimally coupled to an external magnetic field, and become stable only when a gravitational field is added [7]. It is worthwhile then to try and find out whether all possible equations manifest this property. If this were the case, one could very well deduce that the principles of Special Relativity and Quantum Mechanics imply either the nonexistence of truly elementary particles of spins larger than one or that certain interactions are not possible in the absence of other ones. This theoretical aspect, together with the fact that there exist in nature particles of spin $\frac{3}{2}$, and possibly some with still higher spins which need to be described, are the main reasons to try and construct such wave equations.

Some methods have been proposed (see, e.g., the references given in Sect. 2 of Chap. II of [3]) to find all possible Γ's, when given a representation of the Lorentz group. However, the method which has proven most efficient in constructing such equations is that described in references [8] and [9]. The latter solves the problem in the form: given the two free parameters physical spin s and Harish-Chandra degree n, find all possible Γ-algebras. It uses a representation in which the spin matrix $\mathbf{J}^2$ is diagonal, and the discrete symmetries P and R are readily incorporated. For example, in the case of fermions, in this representation,

$$J_i=\begin{bmatrix} s_i & & & \\ & s_i & & \\ & & l_i & \\ & & & l_i \end{bmatrix} \qquad \text{with} \quad l_i=\begin{bmatrix} l_i^1 & & & \\ & l_i^2 & & \\ & & \ddots & \\ & & & l_i^q \end{bmatrix}$$

in which s_i correspond to an irreducible representation of $SU(2)$ for the physical spin s, and $l_i^k = l^k \times s_i^k$ with s_i^k corresponding to an irreducible representation of $SU(2)$ with label s^k; this being an "auxiliary spin." It can be that the tth one would have the same value as the physical spin; the index 0 or "t" will be used to characterize that one, and the indices $(t+1)$ and $(t-1)$ will also sometimes be more simply replaced by $+$ or $-$. The matrices for the discrete symmetries are

$$P = \pm i \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \end{bmatrix}, \qquad R = \begin{bmatrix} 0 & 1 & & \\ 1 & 0 & & \\ & & 0 & 1 \\ & & 1 & 0 \end{bmatrix},$$

$$\eta = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & \theta & \\ & & & -\theta \end{bmatrix} \quad \text{with} \quad \theta = \begin{bmatrix} \theta^1 & & & \\ & \theta^2 & & \\ & & \ddots & \\ & & & \theta^q \end{bmatrix},$$

$$\Gamma_0 = \begin{bmatrix} 1 & & & \\ & -1 & & \\ & & N & \\ & & & -N \end{bmatrix} \quad \text{with} \quad N = \begin{bmatrix} n^1 & & & \\ & n^2 & & \\ & & \ddots & \\ & & & n^q \end{bmatrix},$$

$$N_i = \begin{bmatrix} 0 & i\lambda s_i & 0 & -\alpha_i^\dagger \theta \\ i\lambda s_i & 0 & -\alpha_i^\dagger \theta & 0 \\ 0 & \alpha_i & 0 & iv_i \\ \alpha_i & 0 & iv_i & 0 \end{bmatrix},$$

the Γ_i being obtained as $i[\Gamma_0, N_i]$, according to one of the Eqs. (3.5). λ is an arbitrary real parameter; the column vector $\alpha_k = \text{coln}(0, 0,..., \alpha^- \times (\delta_k^0)^\dagger, \alpha^0 \times s_k, \alpha^+ \times \delta_k^0, 0,..., 0)$ and v_k is a tridiagonal matrix, the jth row of which is

$$(0,..., 0, A^{(j-1)} \times (\delta_k^j)^\dagger, v^j \times s_k^j, B^{(j+1)} \times \delta_k^j, 0,..., 0)$$

with δ_k^j standard intertwining matrices for the generators of $SU(2)$ such that

$$s_i^j \delta_k^j - \delta_k^j s_i^{(j+1)} = i\varepsilon_{ikl} \delta_l^j.$$

With this decomposition of the matrices, the problem is reduced to that of constructing matrices acting in the auxiliary spins vector subspace. Let then $S = S^1 + S^2 + \cdots + S^q$ denote this space. One then has to find matrices θ^j, n^j, v^j, acting within S^j and matrices A^j and B^j mapping, respectively, from $S^{(j-1)}$ to S^j and from $S^{(j+1)}$ to S^j, together with three vectors α^-, α^0, α^+, respectively, in the subspaces S^-, S^0, and S^+. These unknowns must satisfy a set of algebraic relations derived from the above Eqs. (3.4)–(3.11). These relations are listed in [9] for the general case, and in the next section for an example of a particular family of equations.

As a final remark concerning the general problem, the following two cricial properties of the vectors α^-, α^0, α^+ should be mentioned. First, the three of them cannot all be null. Otherwise the Γ's would be in block diagonal form, and thus reducible, while one needs be concerned only with irreducible algebras. Second, all vectors in the representation space must be generated from one of these three vectors, through the action of polynomials of the matrices v, n, etc., mentioned above. (One can easily prove that such vectors always form an invariant space for the Γ-algebra.) Thus the representation space S will have the structure of a tree of vectors, stemming from α^-, α^0, and α^+, and can then be constructed as described in Section 2.

3.3. *Particular Case*

We list below the equations to be solved in the construction of all spin $\frac{3}{2}$ equations involving auxiliary spins $\frac{1}{2}$ and $\frac{3}{2}$, with arbitrary multiplicities. This problem will be solved afterwards as an example of application of our method of solution of polynomial equations in noncommuting variables. We note that it has been solved previously by somewhat different methods [2].

There are two subspaces to be constructed: S^- for auxiliary spins $\frac{1}{2}$, and S^0 for auxiliary spins $\frac{3}{2}$. The matrices to be constructed are v^-, n^-, θ^-, v^0, n^0, θ^0, A^- and B^0; these will be hereafter represented, respectively, by the letters E, G, J, F, H, K, A, B. The vectors α^- of S^- and α^0 of S^0, will be represented respectively by V and W. We note that there is, in this example, no vector α^+ since there is no auxiliary spin $\frac{5}{2}$ sector to the vector space.

There appear in the equations to be solved certain constants, defined as follows

$$c^- = \tfrac{4}{3}(\alpha^-)^\dagger\theta^-\alpha^-, \qquad d^- = \tfrac{4}{3}(\alpha^-)^\dagger\theta^- n^-\alpha^-,$$
$$c^0 = (\alpha^0)^\dagger\theta^0\alpha^0, \qquad d^0 = (\alpha^0)^\dagger\theta^0 n^0\alpha^0.$$

We shall use for them the notation:

$$CV = \tfrac{3}{4}c^-, \qquad CGV = \tfrac{3}{4}d^-, \qquad CW = c^0, \qquad CHW = d^0. \tag{3.12}$$

The reason behind this choice will become apparent below. It is furthermore given that these constants are related as

$$L^2 = 1 - \tfrac{8}{9}CV - CW, \tag{3.13}$$

$$CHW = \tfrac{2}{9} - 2L^2 - CW, \tag{3.14}$$

$$CGV = \tfrac{1}{2} - CV \tag{3.15}$$

in which L stands for a real constant λ, an unknown parameter at this point.

There are four "projectors" involved in the equations; they are the particular matrices

$$M^- = \alpha^-(\alpha^-)^\dagger\theta^-, \quad M^0 = \alpha^0(\alpha^0)^\dagger\theta^0, \quad \alpha^0(\alpha^-)^\dagger\theta^- \quad \text{and} \quad \alpha^-(\alpha^0)^\dagger\theta^0.$$

They will be respectively denoted by M, N, P, and Q. When these matrices are applied on a vector, they always produce a vector parallel to α or α^0. Thus, their action can be represented as:

$$M \quad \text{on VECT} = (\text{CVECT})\, V \quad \text{with CVECT} = (\alpha\ )^{\dagger}\theta\ \ \text{VECT}, \tag{3.16}$$

$$N \quad \text{on VECT} = (\text{CVECT})\, W \quad \text{with CVECT} = (\alpha^0)^{\dagger}\theta^0\ \text{VECT}, \tag{3.17}$$

$$P \quad \text{on VECT} = (\text{CVECT})\, W \quad \text{with CVECT as in Eq. (3.16)}, \tag{3.18}$$

$$Q \quad \text{on VECT} = (\text{CVECT})\, V \quad \text{with CVECT as in Eq. (3.17)}. \tag{3.19}$$

There are five equations between vectors given in the data; they are

$$EV = (5L)\, V + 5BW \tag{3.20}$$

$$EGV = -(4L)\, V - (3L)\, GV - 3BW - 3GBW - BHW, \tag{3.21}$$

$$AGV = -\tfrac{1}{2}AV - \tfrac{1}{2}HAV \tag{3.22}$$

$$FW = -(L)\, W - \tfrac{8}{9}AV \tag{3.23}$$

$$FHW = -(L)\, W + \tfrac{4}{9}AV + \tfrac{4}{9}HAV. \tag{3.24}$$

There are eleven matrix equations:

$$G^2 = 0, \tag{3.25}$$

$$E^2 - 1 - \tfrac{16}{9}BA - \tfrac{16}{9}M = 0, \tag{3.26}$$

$$AE - 5FA - 5P = 0, \tag{3.27}$$

$$4BHA + 3BAG + 3GBA + 4M + 3MG + 3GM + \tfrac{9}{8}EGE - \tfrac{9}{8}G = 0, \tag{3.28}$$

$$P + AGE + FHA + \tfrac{3}{5}AEG + \tfrac{3}{5}HAE = 0, \tag{3.29}$$

$$H^2 = 0, \tag{3.30}$$

$$F^2 - 1 + \tfrac{8}{9}AB + N = 0, \tag{3.31}$$

$$EB - 5BF - 5Q = 0, \tag{3.32}$$

$$9FHF - 4ABH - 4HAB + 7H + 9N = 0, \tag{3.33}$$

$$2AGB + ABH + HAB - H = 0, \tag{3.34}$$

$$Q + EGB + BHF + \tfrac{3}{5}GEB + \tfrac{3}{5}EBH = 0. \tag{3.35}$$

We note that among these equations, two: Eqs. (3.32) and (3.35) are the hermitian conjugate of two others: Equations (3.27) and (3.29) since the matrices A and B are in fact related as

$$JB = A^{\dagger}K. \tag{3.36}$$

We recall that the matrices J and K are square, hermitian, and have inverses.

4. The Solution

One of the interesting features of the chosen example is that the representation space S is constituted of two subspaces S and S^0, and that there are two sets of equations: some of which hold on S (Eqs. (3.25) to (3.29)) and some on S^0 (Eqs. (3.30) to (3.35)). There are also two possibly linearly independent generating vectors: $V \in S$ and $W \in S^0$, and it is known that one of them has to be non null.

Another interesting feature of this problem is that there are some vector equations given initially on top of the matrix equations. These equations should of course be treated as if they had been obtained from the application of an equation on a vector. Thus, for example, Eq. (3.20) is taken to mean that vectors named BW and EV must exist in S, while the largest vector: EV, is a linear combination of the smaller vectors V and BW. This equation also expresses the result of the action of matrix E on vector V. According to the vector equations given, the set of possibly linearly independent vectors is initially:

$$SN = (SN1, SN2)$$

with

$$SN1 = (V, BW, GV, GBW) \qquad (4.1)$$

$$SN2 = (W, AV, HW, HAV).$$

$SN1$ contains the vectors of S and $SN2$ those of S^0.

We now let the construction of the solution proceed, as described in Section 2. It begins with the successive application of Eqs. (3.25)–(3.29) on the generating vector V. When this will have been done, because in the present case there are many (i.e., two) subspaces to be constructed simultaneously, the smallest vector (i.e., W) in the next subspace shall be subjected to the Eqs. (3.30)–(3.35). Then all equations having been used once, one should come back to the first subspace and apply all Eqs. (3.25)–(3.29) on the next smallest vector, etc. Here are the successive results obtained through the use of this algorithm.

When Eq. (3.25) is applied on V, there results G on $GV = 0$, and no new vector is added to $SN1$. When Eq. (3.26) is applied on V, there results the relation $EBW = E$ on $BW = (\frac{1}{5} - 5L^2 + \frac{16}{45}CV)V - (5L)BW + \frac{16}{45}BAV$, and a new vector name: BAV is added to $SN1$. Similarly, Eq. (3.27) leads to

$$FAV = -(CV)W + (L)AV + ABW$$

and ABW is added to $SN2$. Equation (3.28) yields an expression for $GBAV$ and the addition of the new vector names BAV, BHW, $BHAV$, and $EGBW$ to $SN1$. Equation (3.29) yields an expression for $HABW$ and the addition of $ABHW$, $AGBW$, and $FHAV$ to $SN2$.

Equations (3.30)–(3.35) are then applied on W. Equation (3.30) yields simply $HHW = 0$ and Eq. (3.31) an identity. Equation (3.32), however, yields

$BAV = (1 - CV)V$. Thus BAV must be substituted for in all previous vector equations. This involves on one hand a direct substitution in the expression for EBW, and on the other hand the consideration of a "new" vector equation expressing the content of the equation obtained above for $GBAV$. In the latter equation $GBAV$ is replaced by $(1 - CV)GV$ and the relation yields now an expression for $EGBW$, the largest vector it now contains. The remaining equations will yield in turn expressions for $FHAV$, $AGBW$, $EBHW$. Thus, when all equations have been applied on W,

$$SN1 = (V, BW, GV, BHW, GBW, BHAV),$$

$$SN2 = (W, AV, HW, HAV, ABW, ABHW).$$

The next vector to be treated is BW of $SN1$. The equations which are not vector identities, together with the incorporation of the new results in the accumulated data will yield here expressions for: $BABW$, $FABW$, $BABHW$, $FABHW$. Only one new vector has been added to SN in this process: $ABHAV$ is added to $SN1$.

The vector then considered is AV in $SN2$. The application of all equations on AV yields an expression for $ABHAV$ and $EBHAV$, as well as an equation between two constants: $CBW = CAV$. The other equations are identities. Note that the only change in SN is the removal of $ABHAV$. The application of the equations on GV in $SN1$ give an expression for $GBHAV$ and $HABHW$. $GBHW$ and $EGBHW$ are added to $SN1$ and $AGBHW$ to $SN2$.

When the vector HW of $SN2$ is next treated, an expression for $AGBHW$ is obtained and then Eq. (3.34) leads to the equation $(L)\,AV + (3/4 - CV)\,W = 0$. Two branches in the solution must be considered at this point. The rules given in Section 2 require that AV be expressed in terms of W, but the coefficient L may be null. Thus a copy of the environment must be saved to allow to consider later the case $L = 0$, after the case $L \neq 0$ has been completely treated. When $L \neq 0$, $AV = K(W)$ with $K = (CV - 3/4)/L$ follows; the substitution of this result in the accumulated data will lead to the situation where the result of the action of all the matrices on each vector of $SN1$ and $SN2$ is known. There just remains to verify that identities always follow when the matrix equations are applied on the vectors of SN which have not already been treated. When this has been done, a solution is known to have been obtained. In the present case, the solution obtained is:

$$SN1 = (V, GV, BHW, GBHW); \qquad SN2 = (W, HW).$$

The action of the matrices E, G, A on $SN1$ is described as:

$$EV = \tfrac{5}{3}\varepsilon V, \qquad \text{with} \quad \varepsilon = \pm 1,$$

$$EGV = (-\tfrac{4}{3}\varepsilon GV)\,V - \varepsilon GV - BHW,$$

$$EBHW = \tfrac{5}{3}\varepsilon BHW,$$

$$EGBHW = \tfrac{16}{9}(1 - CV)^2 V - \tfrac{4}{3}\varepsilon(1 - CV)\,BHW - \varepsilon GBHW,$$

$$
\begin{aligned}
GV &= (1)\, GV, \\
GGV &= 0, \\
GBHW &= (1)\, GBHW, \\
GGBHW &= 0, \\
AV &= \tfrac{3}{4}\varepsilon W, \\
AGV &= -\tfrac{3}{8}\varepsilon W - \tfrac{3}{8}\varepsilon HW, \\
ABHW &= (1 - CV)\, W + HW, \\
AGBHW &= \tfrac{1}{2}(CV - 1)\, HW.
\end{aligned}
$$

The action of F, H, and B on $SN2$ is

$$
\begin{aligned}
FW &= \tfrac{1}{3}\varepsilon(1 - 4CV)\, W, \\
FHW &= \tfrac{4}{3}\varepsilon(1 - CV)\, W + \tfrac{1}{3}\varepsilon HW, \\
HW &= (1)\, HW, \\
HHW &= 0, \\
BW &= \tfrac{4}{3}\varepsilon(1 - CV)\, V, \\
BHW &= (1)\, BHW.
\end{aligned}
$$

The constants are $L = \varepsilon(\tfrac{4}{3}CV - 1)$, $CW = \tfrac{16}{9}CV(1 - CV)$, and $CHW = -\tfrac{16}{9}(CV - 1)^2$, etc. These can be seen to give the matrices J and K. The only arbitrary parameter remaining is CV, so the solution obtained corresponds to a one parameter family of representations.

Since no other requirement of linear dependence comes from the matrix equations, the above set of vector equations, in which the vectors of SN are considered linearly independent, describe a representation of the matrices.

We note that when the case $L = 0$ is examined, two possibilities arise: one of them leads to a solution of the same form as the one described above in which L would have the particular value 0. The other possibility is for the vector W to be null. This gives rise to a different Γ-algebra which is well known as the Fierz–Pauli or Rarita–Schwinger spin $\tfrac{3}{2}$ algebra [10].

5. Discussion

We have described an algorithm for constructing all representations of matrices satisfying a given set of algebraic equations. According to this method, the representation space is built up as a connected tree, stemming from generating vectors in such a way that the properties of all the matrices are obtained in the process, while the satisfaction of the given algebraic matrix equations is guaranteed.

The representations so constructed constitute a general solution to the problem in the sense that all possible irreducible representations are either exactly the representations produced or a particular case of these, obtained by postulating additional relations between the vectors, or the matrices. This comes from the fact that, during the construction, vectors are added only when the need for their existence is ascertained; and linear relations between them are established only as necessary conditions for one of the matrix equations to hold true.

To illustrate the application of this method, we have shown in Sections 3 and 4, how it can be used to find representations of generalized Dirac Γ-algebras. The solution of one particular such problem has been presented so as to show the steps through which the computer program actually goes through. This allows one to see how the construction evolves to reach the solution.

This example proves to be well suited to demonstrate the possibilities of the method. It shows how it can be adapted to deal with having a multisubspaces structure of the vector space, with some matrices mapping in the same subspace and some in other subspaces. This example also illustrates how some supplementary vector relations, as Eqs. (3.20)–(3.24), not necessarily implied by the set of matrix equations can be very naturally incorporated in the data. Finally, it shows how to deal, at least up to a certain point, with an equation as Eq. (3.36), which is not algebraic, i.e., which does not involve only matrix polynomials. As was remarked after Eq. (3.36), what has been done in this case is simply to add the hermitian conjugate of all equations which were not self-hermitian. Although this proved sufficient here, it will not be so in general; and one will have to find a way to translate the content of such "nonalgebraic" equations in purely algebraic properties of the matrices.

As mentioned earlier a program already exists which automatically goes through all the steps of the construction, except for the simplification of the symbolic constants in the coefficients of the vectors (many of which occur in the above examples). The program was initially designed as an interactive one partly to allow the user to do this kind of simplification by hand (although all numerical coefficients were automatically dealt with). We did not find it urgent at this point to write a function which would deal with this aspect of the calculations because there exist already excellent programs to do that. One other reason for writing an interactive program was to have the possibility of asking any question about the data so as to follow better the evolution of the solution.

The same program has then been applied to the finding of representations of finite groups, given a presentation. In this case, it was easy to automatized all steps of the construction and a specialized version of the program has been written expressly to produce the regular representation of groups. Work on this subject will be reported elsewhere.

A completely automatic version of the program will also be written shortly for the construction of representations of Lie algebras and of generalized Dirac Γ-algebras, starting only from the set of matrix equations.

References

1. G. Labonté, *Nuovo Cimento A* **80**, 127 (1986).
2. G. Labonté, *Nuovo Cimento A* **80**, 77 (1984).
3. A. S. Wightman, *Proceedings of the Ettore Majorana International School of Mathematical Physics of Erice, July 1977* (Springer Verlag, New York, 1978).
4. H. J. Bhabha, *Rev. Mod. Phys.* **21**, 451 (1949).
5. Harish-Chandra, *Phys Rev.* **71**, 793 (1947).
6. E. Wild, *Proc. Roy, Soc. London A* **191**, 253 (1947).
7. S. Deser and B. Zumino, *Phys. Lett.* **38**, 1433 (1977).
8. G. Labonté, *Nuovo Cimento A* **72**, 173 (1982).
9. G. Labonté, *Nuovo Cimento A* **75**, 416 (1983).
10. W. Rarita and J. Schwinger, *Phys. Rev.* **60**, 61 (1941).