The Cayley-Hamilton theorem for supermatrices

1993 J. Phys. A: Math. Gen. 26 L441

(http://iopscience.iop.org/0305-4470/26/9/002)

J. Phys. A: Math. Gen. 26 (1993) L441–L447. Printed in the UK

LETTER TO THE EDITOR

The Cayley–Hamilton theorem for supermatrices

L F Urrutia† and N Morales‡

† Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Circuito Exterior, CU, 04510 México, DF, and Centro de Estudios Científicos de Santiago, Casilla 16443, Santiago 9, Chile

‡ Department of Mathematics, Universidad Autónoma Metropolitana-Iztapalapa, Apartado Postal 55-534, 09340 México, D.F.

Received 27 November 1992, in final form 1 March 1993

Abstract. Starting from the expression for the superdeterminant of a supermatrix we propose a definition for the corresponding characteristic polynomial. We prove that each supermatrix satisfies its characteristic equation. In some particular cases we are able to construct polynomials of lower degree which are also shown to be nullified by the supermatrix.

In the past few years there has been a great deal of interest in 2+1 dimensional Chern–Simons (CS) theories [1], mainly in relation with conformal field theory [2, 3], link invariants [2, 4], and quantum groups [2, 5]. Pure CS theories are of topological nature and the fundamental degrees of freedom are the traces of group elements constructed as the holonomies (or Wilson lines, or integrated connections) of the gauge connection around oriented closed curves in the manifold. The observables are the expectation values of the Wilson lines which turned out to be realized as the various knot polynomials known to mathematicians [6]. Since CS theories are also exactly soluble and possess a finite number of degrees of freedom, another aspect of interest is the reduction of the initially infinite-dimensional phase space to the subspace of the true degrees of freedom. The Cayley–Hamilton theorem has played an important role in the construction of the so called skein relations, which are relevant to the calculation of expectation values [7], and also in the process of reduction of the phase space. The simplest example of the latter point arises in the discussion of the reduced phase space in a sector of anti-de Sitter gravity in 2 + 1 dimensions, which is equivalent to the Chern–Simons theory of the group $SO(2,2)$ [8]. This theory can be more easily described in terms of two copies of the group $SL(2,R)$, which is the spinorial group of $SO(2,2)$ [9]. The gauge invariant degrees of freedom associated to one genus of an arbitrary genus g two-dimensional surface turn out to be traces of any product of powers of two $SL(2,R)$ matrices M_1 and M_2, which correspond to the holonomies of the two homotopically distinct trajectories on one genus. Nevertheless, one should be able to reduce this infinite set of traces to a finite set of degrees of freedom. It is precisely at this point where the Cayley–Hamilton theorem finds its use. In the case of $SL(2,R)$ we have the Cayley–Hamilton identity $(M_1)^2 - \mathrm{Tr}(M_1)M_1 + I = 0$. By multiplying this equation by $M_2M_1^{-1}$ and taking traces we obtain the following relation among the traces

$$\mathrm{Tr}(M_2M_1^{-1}) + \mathrm{Tr}(M_1M_2) = \mathrm{Tr}(M_1)\,\mathrm{Tr}(M_2). \qquad (1)$$

By using equation (1), the general degree of freedom $\mathrm{Tr}(M_1{}^{p_1}M_2{}^{q_1}M_1{}^{p_2}M_2{}^{q_2}\cdots M_1{}^{p_n}M_2{}^{q_n}\cdots)$, for any p_i, q_i in $\mathbb{Z}$, can be shown to be reducible and can be expressed as a function of three traces only: $\mathrm{Tr}(M_1)$, $\mathrm{Tr}(M_2)$ and $\mathrm{Tr}(M_1M_2)$ [9]. A similar reduction can be performed in the case of 2 + 1 super de Sitter gravity, which is the Chern–Simons theory of the supergroup $Osp(2|1, \mathbb{C})$ [10]. The novelty here is that one is dealing with supermatrices instead of ordinary matrices. In the particular case considered, a Cayley–Hamilton identity for the supermatrices was obtained in a heuristical way and a relation analogous to (1) was derived for the supertraces. This allowed the carrying out of the reduction of the infinite dimensional phase space, this time in terms of five complex supertraces [11]. It is worth remarking that the quantization of these true degrees of freedom led to a quantum algebra structure both in the de Sitter and super de Sitter cases.

In this letter we discuss the general proof of the Cayley–Hamilton theorem for supermatrices. This is an interesting problem by itself, besides its applications in the study of the reduced phase space in Chern–Simons theories defined over a supergroup. Basically our problem is twofold: first we have to find an adequate definition of the characteristic polynomial of a supermatrix and then we have to show that indeed the supermatrix satisfies its characteristic equation. The same procedure has to be followed for the null polynomials of lower degree.

A $(p+q)\times(p+q)$ supermatrix is a block matrix of the form

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{2}$$

where A, B, C and D are $p\times p, p\times q, q\times p, q\times q$ matrices respectively. The distinguishing feature with respect to an ordinary matrix is that the matrix elements M_{RS}, $R = (i, \alpha)$, $S = (j, \beta)$ are elements of a Grassmann algebra with the property that A_{ij} $(i, j = 1, \cdots p)$ and $D_{\alpha\beta}$ $(\alpha, \beta = 1, \cdots q)$ are even elements, while $B_{i\alpha}$ and $C_{\beta j}$ are odd elements of such algebra. Let us recall that the ordinary matrix addition and the ordinary matrix product of two supermatrices is again a supermatrix. Nevertheless, such concepts as the trace and the determinant need to be redefined, because of the odd component piece of the supermatrix. The basic invariant under similarity transformations for supermatrices is the supertrace, defined by $\mathrm{Str}\,(M) = \mathrm{Tr}(A) - \mathrm{Tr}(D)$, where the trace over the even matrices is the standard one. The generalization of the determinant, called the superdeterminant, is obtained from the supertrace by defining $\delta \ln(\mathrm{S}\det M) = \mathrm{Str}\,(M^{-1}\delta M)$. In this compact notation we are sumarizing the $(p+q)^2$ relations which give the partial derivatives of the function $\ln(\mathrm{S}\det M)$ with respect to the entries M_{RS} of the supermatrix, in terms of the elements of the inverse supermatrix M^{-1}. For example, $\partial \ln(\mathrm{S}\det M)/\partial M_{ij} = (M^{-1})_{ji}$ for the even indices i, j. These first order partial differential equations are subsequently integrated under the boundary conditions $\mathrm{S}\det I = 1$, where I is the unit supermatrix, to produce the following equivalent two forms of calculating the superdeterminant [12]

$$\mathrm{S}\det(M) = \frac{\det(A - BD^{-1}C)}{\det D} = \frac{\det A}{\det(D - CA^{-1}B)}. \tag{3}$$

Here all the matrices involved are even in the Grassmann algebra and det has its usual meaning. Let us consider now the function $h(x) = \mathrm{S}\det(xI - M)$ which could be naively taken as the analogous of the standard characteristic polinomial. Nevertheless, this function is in fact the ratio of two monic polinomials

$$h(x) = \frac{\tilde{F}(x)}{\tilde{G}(x)} = \frac{F(x)}{G(x)} \tag{4}$$

each form arising from the two alternatives (3) of calculating the superdeterminant. The explicit expressions for the numerators and denominators are

$$\tilde{F}(x) = \det(\mathrm{d}(x)(xI - A) - B\,\mathrm{adj}(xI - D)C) \qquad \tilde{G}(x) = \mathrm{d}(x)^{p+1} \tag{5a}$$

$$F(x) = a(x)^{q+1} \qquad G(x) = \det(a(x)(xI - D) - C\,\mathrm{adj}(xI - A)B) \tag{5b}$$

where we have used the basic relation $(F)^{-1} = [\det(F)]^{-1}\mathrm{adj}(F)$ valid for any even matrix F. Here $a(x) = \det(xI - A)$ and $d(x) = \det(xI - D)$.

The proposed definition of the characteristic polynomial $\mathcal{P}(x)$ for an arbitrary supermatrix is

$$\mathcal{P}(x) = \tilde{F}(x)G(x) = F(x)\tilde{G}(x) = a(x)^{q+1}\mathrm{d}(x)^{p+1}. \tag{6}$$

For notational simplicity we will not necessarily write explicitly the x-dependence on many of the polynomials considered in the sequel.

Nevertheless, and motivated by the work of [13], we have realized that there are some cases in which we can construct null polynomials of lower degree according to the factorization properties of the basic polynomials $\tilde{F}$, $\tilde{G}$, F, G. At this point it is important to observe that we do not have a unique factorization theorem for polynomials defined over a Grassmann algebra. This can be seen from the identity $x^2 = (x + \alpha)(x - \alpha)$, where α is an even Grassmann with $\alpha^2 = 0$. The construction of such null polynomials starts from finding the divisors of maximum degree of the pairs $\tilde{F}$, $\tilde{G}$, (F, G) which we denote by $r(s)$ respectively. This means that one is able to write

$$\begin{aligned} \tilde{F} &= r\tilde{f} \qquad \tilde{G} = r\tilde{g} \\ F &= sf \qquad G = sg \end{aligned} \tag{7}$$

where all polynomials are monic and also $\tilde{f}$, $\tilde{g}$, f, g are of least degree by construction. They must satisfy

$$\tilde{f}/\tilde{g} = f/g \tag{8}$$

because of the equation (4) and the expressions in (7) might be not unique. Let us observe that in the case of polynomials over the complex numbers equation (8) would imply at most $\tilde{f} = \lambda f, \tilde{g} = \lambda g$ with λ being a constant. Since we are considering polynomials over a Grassmann algebra this is not necessarily true as can be seen again in the above mentioned identify $x/(x - \alpha) = (x + \alpha)/x$, which we have rewritten in a convenient way. For each family of possible factorizations written in equation (7) we define a null polynomial $P(x)$ as

$$P(x) = \tilde{f}(x)g(x) = f(x)\tilde{g}(x) \tag{9}$$

which is clearly of less degree than $\mathcal{P}(x)$. The characteristic polynomial $\mathcal{P}(x)$ is just a particular case of these null polynomials when $r = s = 1$. In the case where $a(x)$ and $d(x)$ are coprime it can be shown that the factorization in equation (7) is unique with $f = \tilde{f}$ and $g = \tilde{g}$ [13].

Since we are interested in the Cayley–Hamilton theorem, now we have to prove that both choices (6) and (9) are in fact such that $\mathcal{P}(M) = 0$ and $P(M) = 0$ respectively. We emphasize again that the former case reduces to a particular case of the latter, so that we will concentrate only on $P(x)$. Part of our strategy to verify that $P(M) = 0$ is closely related to one of the standard methods to prove the Cayley–Hamilton theorem for ordinary matrices [14]. The starting point is the following: if for the supermatrix $(xI - M)$, with M being an $n \times n$ supermatrix independent of x, there exists a polynomial supermatrix $N(x) = N_0 x^{m-1} + N_1 x^{m-2} + .. + N_{m-1} x^0$, (where each N_k, $k = 0, \cdots, m-1$, is an $n \times n$ supermatrix) such that

$$(xI - M)N(x) = P(x)I \tag{10}$$

where $P(x) = p_0 x^n + p_1 x^{n-1} + \cdots + p_n x^0$ is a numerical polynomial over the Grassmann algebra, then $P(M) = p_0 M^n + p_1 M^{n-1} + \cdots + p_n I \equiv 0$. The proof follows by comparing the independent powers of x in equation (10). The equality of the highest power in x on both sides implies $m = n$, together with $N_0 = p_0 I$. This leads to the following set of supermatrix relations for the remaining powers

$$\begin{aligned} N_1 - p_0 M &= p_1 I \\ N_2 - M N_1 &= p_2 I \\ \cdots \quad & \cdots \\ N_{n-1} - M N_{n-2} &= p_{n-1} I \\ -M N_{n-1} &= p_n I. \end{aligned} \tag{11}$$

By multiplying on the left the k-th equation by M^{n-k} and adding up all the relations we can verify that the terms involving N_k cancel out in the LHS leaving only the term $-p_0 M^n$. The RHS is just $P(M) - p_0 M^n$ so that we have the result $P(M) = 0$ as required. In the standard case of purely even matrices $N(x)$ is just given by $N(x) = \mathrm{adj}(xI - M) = \det(xI - M)(xI - M)^{-1}$, and $P(x) = \det(xI - M)$.

In the case of a supermatrix we do not have an obvious generalization either of the polynomial matrix $\mathrm{adj}(xI - M)$ or of $\det(xI - M)$. Nevertheless, following the analogy as closely as possible we define

$$N(x) \equiv P(x)(xI - M)^{-1} \tag{12}$$

where $P(x)$ is the polynomial (9) introduced previously. The challenge now is to prove that $N(x)$, which trivially satisfies the equation (10), is indeed a polynomial matrix. In the first place we calculate $(xI - M)^{-1}$ in block form, with the result

$$(xI - M)^{-1}_{11} = ((xI - A) - B(xI - D)^{-1}C)^{-1} \tag{13a}$$

$$(xI - M)^{-1}_{12} = (xI - A)^{-1}B((xI - D) - C(xI - A)^{-1}B)^{-1} \tag{13b}$$

$$(xI - M)^{-1}_{21} = (xI - D)^{-1}C((xI - A) - B(xI - D)^{-1}C)^{-1} \tag{13c}$$

$$(xI - M)^{-1}_{22} = ((xI - D) - C(xI - A)^{-1}B)^{-1} \tag{13d}$$

where the subindices 11, 12, 21 and 22 denote the corresponding $p \times p$, $p \times q$, $q \times p$, and $q \times q$ blocks respectively. Before considering in equation (12) the specific case of the previously defined $P(x)$ it will prove most convenient to rewrite the expressions (13) in terms of derivatives of the even fuctions $\tilde{F}$ and G with respect to the generic supermatrix elements A_{ij}, $B_{i\alpha}$, $C_{\alpha i}$, $D_{\alpha\beta}$. To this end we use the basic property

$$\delta \ln \det Q = \mathrm{Tr}(Q^{-1}\delta Q) \tag{14}$$

valid for any even matrix Q. The meaning of this compact notation has been already explained in the paragraph preceeding equation (3) and, *mutatis mutandis*, applies also to even matrices. From equation (5a) we can rewrite $\tilde{F}$ as

$$\tilde{F} = d^p \det((xI - A) - B(xI - D)^{-1}C). \tag{15}$$

Calculating the variation of $\ln \tilde{F}$ with respect to A_{ij} by means of equation (14) we readily obtain

$$(xI - M)^{-1}_{ij} = -\frac{1}{\tilde{F}}\frac{\partial \tilde{F}}{\partial A_{ji}} \tag{16}$$

for the 11 block of $(xI - M)^{-1}$. The analogous variation of $\ln \tilde{F}$, this time with respect to $B_{j\alpha}$, leads to

$$(xI - M)^{-1}_{\alpha j} = -\frac{1}{\tilde{F}}\frac{\partial \tilde{F}}{\partial B_{j\alpha}} \tag{17}$$

for the 21 block of $(xI - M)^{-1}$. Here we are taking the derivative with respect to an odd Grassmann number from the left in the sense that $\delta\tilde{F} \equiv \delta B_{j\alpha}(\partial \tilde{F}/\partial B_{j\alpha})$. Similar calculations starting from $\ln G$ lead to the following expressions for the remaining blocks of $(xI - M)^{-1}$

$$(xI - M)^{-1}_{i\alpha} = -\frac{1}{G}\frac{\partial G}{\partial C_{\alpha i}} \qquad (xI - M)^{-1}_{\alpha\beta} = -\frac{1}{G}\frac{\partial G}{\partial D_{\beta\alpha}}. \tag{18}$$

Now we come to the last step of our argument which consists in using the polynomial defined in equation (9) together with the corresponding factorization properties (7) and (8), to prove that $N(x) = P(x)(x - M)^{-1}$ is a polynomial matrix.

Let us consider the block-elements 11 and 21 of $P(x)(xI - M)^{-1}$ in the first place. According to the expression (16) together with (7) and (9), the first block-element can be written as

$$N_{ij} = -g\frac{\partial \tilde{f}}{\partial A_{ji}} - \frac{g\tilde{f}}{r}\frac{\partial r}{\partial A_{ji}}. \tag{19}$$

The first term of the RHS is clearly of polynomial character. In order to transform the second term we make use of the property

$$\frac{\partial \ln \tilde{G}}{\partial A_{ji}} = 0 = \frac{\partial \ln r}{\partial A_{ji}} + \frac{\partial \ln \tilde{g}}{\partial A_{ji}} \tag{20}$$

which follows from the factorization (7) of $\tilde{G}$, together with the fact that $\tilde{G}$ is just a function of $D_{\alpha\beta}$, according to equation (5a). In this way and using the relation (8) we obtain

$$N_{ij} = f\frac{\partial \tilde{g}}{\partial A_{ji}} - g\frac{\partial \tilde{f}}{\partial A_{ji}} \tag{21}$$

which leads to the conclusion that the block-matrix N_{ij} is indeed polynomial. The proof for $N_{\alpha i}$ runs along the same lines, except that now the derivatives are taken with respect to $B_{i\alpha}$ and that we have to use $\partial \ln \tilde{G}/\partial B_{i\alpha} = 0$, instead of equation (20).

The remaining terms $N_{i\alpha}$ and $N_{\alpha\beta}$ can be dealt with in analogous manner by considering the derivatives of G with respect to $C_{\alpha i}$ and $D_{\beta\alpha}$, and by replacing the condition (20) by $\partial \ln F/\partial C_{\alpha i} = 0$ and $\partial \ln F/\partial D_{\beta\alpha} = 0$ respectively. The results are again of the form (21).

To summarize, we have proved that $N(x) \equiv P(x)(xI - M)^{-1}$ is a polynomial matrix, for any $P(x)$ defined in equation (9). Then it follows immediately that $(xI - M)N(x) = P(x)I$ and by using the statement after equation (10) we obtain the desired result $P(M) = 0$. The same conclusion holds for the characteristic polynomial $\mathcal{P}(x)$ defined in equation (6). This concludes the proof of the Cayley–Hamilton theorem for supermatrices.

Before closing we give two simple examples of null polynomials of minimun degree for supermatrices, which belong to the case when the factorization (7) is unique. The first one corresponds to an arbitrary $(1+1) \times (1+1)$ supermatrix with elements $A_{11} = \mathbf{a}$, $B_{11} = \alpha$, $C_{11} = \beta$ and $D_{11} = \mathbf{d}$. The result is [15]

$$P(x) = (\mathbf{a} - \mathbf{d})x^2 - (\mathbf{a}^2 - \mathbf{d}^2 + 2\alpha\beta)x + \mathbf{ad}(\mathbf{a} - \mathbf{d}) + (\mathbf{a} + \mathbf{d})\alpha\beta. \tag{22}$$

The second example has to do with $(2+1) \times (2+1)$ supermatrices M which belong to the supergroup $Osp(1|2; \mathbb{C})$. In this case the null polynomial is [11]

$$P(x) = x^3 - (2 + \mathrm{Str} M)(x^2 - x) - 1. \tag{23}$$

A detailed version of this work will be presented elsewhere [16].

The work of both authors has been partially supported by the grant DGAPA-UNAM - IN100691. LFU also acknowledges support from the project CONACyT-0758-E9109. LFU thanks Dr R Weder for making [13] available to him. He also thanks Dr H Waelbroeck for suggesting some improvements in the presentation.

References

[1] For a review see for example
Birmingham D, Blau M, Rakowski M and Thompson G 1991 *Phys. Rep.* **209** 129
[2] Witten E 1989 *Commun. Math. Phys.* **121** 351; 1989 *Nucl. Phys.* B **322** 629
[3] Moore G and Seiberg N 1989 *Phys. Lett.* **220B** 422
Elitzur S, Moore G, Schwimmer A and Seiberg N 1989 *Nucl. Phys.* **326** 108
Labastida J M F and Ramallo A V 1989 *Phys. Lett.* **227B** 92
Frölich J and King C 1989 *Commun. Math. Phys.* **126** 167
Boss M and Nair V P 1970 *Mod. Phys. Lett.* A **5** 959
[4] Atiyah M F 1988 *The Mathematical Heritage of Hermann Weyl, Proc. Symp. Pure Math.* **48** ed R Wells (Providence, RI: American Mathematical Society)
Cotta-Ramusino P, Guadagnini E, Martellini M and Mintchev M 1990 *Nucl. Phys.* B **330** 557
Guadagnini E, Martellini M and Mintchev M 1989 *Phys. Lett.* **228B** 489; 1990 *Nucl. Phys.* B **330** 575
[5] Guadagnini E, Martellini M and Mintchev M 1990 *Phys. Lett.* **235B** 275; 1990 *Nucl. Phys.* B **336** 581
[6] Jones V 1985 *Bull AMS* **12** 103; 1989 *Pacific J. Math.* **137** 311
Freyd P, Yetter D, Hoste J, Lickorish W, Millet K and Ocneanu A 1985 *Bull AMS* **12** 239
Kauffman L 1987 *Topology* **26** 395
[7] Horne J H 1990 *Nucl. Phys.* B **334** 669
[8] Witten E 1988/89 *Nucl. Phys.* B **311** 46
[9] Nelson J E, Regge T and Zertuche F 1990 *Nucl. Phys.* B **339** 516
[10] Koehler K, Mansouri F, Vaz C and Witten L 1990 *Mod. Phys. Lett.* A **5** 935; 1990 *Nucl. Phys.* B **341** 167; 1991 *Nucl. Phys.* B **348** 373
[11] Urrutia L F, Waelbroeck H and Zertuche F 1992 *Mod. Phys. Lett.* A **7** 2715
[12] See for example De Witt B 1984 *Supermanifolds* (Cambridge: Cambridge University Press)

[13] Kobayashi Y and Nagamachi S 1990 *J. Math. Phys.* **31** 2726

[14] See for example Nering E D 1970 *Linear Algebra and Matrix Theory* (New York: Wiley) 2nd edn

[15] Urrutia L F, Waelbroeck H and Zertuche F *preprint ICN-UNAM 1992*; The Algebra of Supertraces for 2+1 Super de Sitter Gravity, to be published in *Proc. Workshop on Harmonic Oscillators, Maryland, March 1992*

[16] Urrutia L F and Morales N work in progress