ANNALES DE L'I. H. P., SECTION A

D. M. O'BRIEN
A. CANT
A. L. CAREY

On characteristic identities for Lie algebras

Annales de l'I. H. P., section A, tome 26, nº 4 (1977), p. 405-429
<http://www.numdam.org/item?id=AIHPA_1977__26_4_405_0>

Ann. Inst. Henri Poincaré,
Vol. XXVI, n° 4, 1977, p. 405-429.

Section A :
Physique théorique.

On characteristic identities for Lie algebras

by

D. M. O'BRIEN, A. CANT and A. L. CAREY
Department of Mathematical Physics, University of Adelaide

ABSTRACT. — Let U(L) be the universal enveloping algebra of a complex semi-simple Lie algebra L, and let V be an irreducible finite dimensional L-module. Denote by ∂ the map from U(L) into (End V) $\otimes$ U(L) defined for $x \in \mathrm{L}$ by $\partial(x) = 1 \otimes x + x \otimes 1$. In a recent paper, Kostant [*6*] has shown that for each element z of the centre of U(L) there is a unique monic polynomial p_z with coefficients from the centre of U(L) such that $p_z(\partial z) = 0$. For the case where z is the universal Casimir element, these identities have been exploited in the physical literature for many years, and were systematically analysed as « characteristic identities for Lie algebras » by Bracken and Green [*3*] and Green [*4*]. In this paper we provide a common viewpoint for these different approaches. The main aim of this paper is to establish the existence and properties of the characteristic identity for the Lie algebra of the general linear group, and to prove some results which indicate the relevance of these identities for infinite dimensional representation theory.

RÉSUMÉ. — Soient U(L) l'algèbre enveloppante d'une algèbre de Lie L semi-simple complexe, et V un L-module irréductible, de dimension finie. On note par ∂ l'application de U(L) dans (End V) $\otimes$ U(L), définie pour tout $x \in \mathrm{L}$ par $\partial(x) = 1 \otimes x + x \otimes 1$. Récemment, Kostant [*6*] a montré que, pour chaque élément z dans le centre de U(L), il existe un polynôme p_z avec coefficients appartenant au centre de U(L), tel que $p_z(\partial z) = 0$. Dans le cas particulier, où z est l'élément de Casimir, les physiciens ont utilisé depuis longtemps telles « identités caractéristiques ». On peut voir Bracken

et Green [3] et Green [4] pour une analyse systématique des « identités caractéristiques pour les algèbres de Lie ». On donne ici un traitement unifié pour ces approches divers. Le but principal de cet article est l'établissement de l'existence et des propriétés de « l'identité caractéristique » pour l'algèbre de Lie $gl(n)$. On démontrera aussi certains résultats qui indiquent l'utilité de ces identités pour la théorie des représentations de dimension infinie.

1. INTRODUCTION

Suppose that Q denotes the rational field and that $gl(n, \mathrm{Q})$ denotes the Lie algebra over Q with basis

$$\{ a^i{}_j,\ 1 \leqslant i, j \leqslant n \}$$

and Lie products

$$[a^i{}_j, a^k{}_l] = \delta^k{}_j a^i{}_l - \delta^i{}_l a^k{}_j, \qquad 1 \leqslant i, j, k, l \leqslant n . \tag{1}$$

Let U denote the universal enveloping algebra of $gl(n, \mathrm{Q})$ and Z the centre of U. Finally, let A denote the $n \times n$ matrix over U with entries $(a^i{}_j)$,

$$\mathrm{A} = \begin{pmatrix} a^1{}_1 a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 a^2{}_2 & \dots & a^2{}_n \\ \cdot \quad \cdot & & \cdot \\ \cdot \quad \cdot & & \cdot \\ \cdot \quad \cdot & & \cdot \\ a^n{}_1 a^n{}_2 & \dots & a^n{}_n \end{pmatrix}.$$

The aim of this paper is to demonstrate the following facts.

1) A satisfies an nth degree monic polynomial identity,

$$p(\mathrm{A}) = \mathrm{A}^n + c_1 \mathrm{A}^{n-1} + \dots + c_{n-1}\mathrm{A} + c_n = 0 , \tag{2}$$

whose coefficients $c_1, c_2, \dots, c_n$ are elements of Z.

2) A does not satisfy a polynomial identity of lower degree, so equation (2) is the only nth degree monic polynomial identity satisfied by A. $p(\mathrm{A})$ is called the characteristic polynomial of A and equation (2) the characteristic identity of A.

3) U can be embedded injectively in an algebra $\bar{\mathrm{U}}$ over Q in which $p(\mathrm{A})$ can be factorized:

$$p(\mathrm{A}) = \prod_{i=1}^{n} (\mathrm{A} - a_i) , \qquad a_i \in \bar{\mathrm{U}} .$$

The roots of $p(\mathrm{A})$ satisfy

$$[a_i, u] = 0, \qquad 1 \leqslant i \leqslant n ,$$

for all $u \in \mathrm{U}$.

4) $a_1, a_2, \ldots, a_n$ are distinct elements of $\bar{U}$.

5) A has a spectral representation,

$$A = \sum_{i=1}^{n} a_i E_i ,$$

where the projection matrices $E_1, E_2, \ldots, E_n$ over $\bar{U}$ satisfy

$$E_i E_j = \delta_{ij} E_i ,$$

$$\sum_{i=1}^{n} E_i = 1 .$$

6) Suppose that F is any field which satisfies

$$Q \subseteq F \subseteq C .$$

If $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in F^n$, there exists (up to isomorphism) one and only one irreducible U_F-module with highest weight λ, denoted by $V_F(\lambda)$, and $a_1, a_2, \ldots, a_n$ can be reordered so that

$$a_k v = (\lambda_k + n - k)v , \qquad \forall v \in V_F(\lambda) .$$

7) The values of the roots of the characteristic polynomial in a particular irreducible representation determine, and are determined by, the infinitesimal character of the representation.

8) Let G be any field such that $Q \subseteq G \subseteq C$. Then all the results 1) to 7) hold for the Lie algebra $gl(n, G)$ if we replace Q by G throughout. Note the following points:

a) We shall only prove 1)-7) for $gl(n, Q)$, as the extension to $gl(n, G)$ is immediate.

b) Similar results hold for all the classical subalgebras of $gl(n, G)$. (Compare Green [*4*] and Kostant [*6*].)

c) In a particular representation of $gl(n, G)$, the matrix A may be regarded as a « matrix of operators », in which case it may satisfy a polynomial identity of degree less than n. We do not pursue this point here, but it is often important in applications [*4*].

Identities of the form (2) have a distinguished history. The first person to exploit them was Dirac [*1*], who wrote down what amounts to the characteristic identity for the Lie algebra so(3, 1). This particular example is intimately connected with the problem of describing the structure of relativistically invariant wave equations. Later, it was shown by Lehrer-Ilamed [*2*] that n^2 elements chosen from the universal enveloping algebra of any Lie algebra satisfy n^2 identities. He also noted that in certain circumstances the n^2 identities can be written as a single polynomial identity of degree n for an $n \times n$ matrix, analogous to the Cayley-Hamilton identity satisfied by a matrix over a commutative algebra. The first major steps in the analysis of the general problem were taken by Bracken and Green [3] and Green [4], [5], who gave the identities for the classical Lie algebras, and employed them together with some of the results listed above to synthesize irreducible representations (over the complex field) for these algebras. The techniques devised by Bracken and Green are powerful and are quite straightforward to apply. However, the correct algebraic framework for their results needs clarification, and this is carried out in this paper.

Some recent unpublished work by Hannabuss has shown how the matrix A over $gl(n, Q)$, or the corresponding matrix over any of the classical subalgebras of $gl(n, Q)$, can be interpreted as an operator on the tensor product module $V(\lambda) \otimes V$, where $V(\lambda)$ and V are both finite dimensional. By this means we show how to connect the above results with an important recent paper by Kostant [6].

In Section 2, we give a general discussion of Kostant's results, and derive, as a corollary, the most general theorem on the existence and factorization of the characteristic identity of a complex semisimple Lie algebra. At the same time we discuss an extension of Kostant's results to real forms, and conjecture, on the basis of our analysis of $gl(n, Q)$, that they extend to reductive Lie algebras over any field F which satisfies

$$Q \subseteq F \subseteq C.$$

In the subsequent sections the results listed at the beginning of the introduction are proved. A number of conventions are employed there and we list them below.

1) F will denote a subfield of the complex field C which includes the rational field Q.

2) The statement « K is an algebra » will imply that K is associative and has an identity element.

3) If K is an algebra over Q, the statement « V is a K_F-module » will imply that V is considered as a vector space over F.

4) Unless explicitly stated otherwise, Roman indices will take values $1, 2, \ldots, n$ and the summation convention will apply to repeated Roman indices.

2. THE APPLICATION OF KOSTANT'S WORK TO POLYNOMIAL IDENTITIES

Our notation follows that of Humphreys [7]. Let L be a complex semisimple Lie algebra of rank l, let U(L) be the universal enveloping algebra of L, and let Z(L) be the centre of U(L). Select a Cartan subalgebra H of L, with dual space H*, and let Φ denote the set of roots of L relative to H. Let $\Delta = \{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ be a base of Φ. We denote the set of positive (negative) roots relative to Δ by $\Phi^+(\Phi^-)$, and take δ to be half the sum of the positive roots. Let $\Lambda^+ \subset H^*$ be the set of dominant integral linear functions on H. Finally, let W denote the Weyl group.

We are interested in the structure of the L-module $V(\lambda) \otimes V$, where $V(\lambda)$ is the finite dimensional irreducible L-module with highest weight $\lambda \in \Lambda^+$ and V is any L-module (possibly infinite dimensional) which admits an infinitesimal character χ. This means that, for every $z \in Z(L)$,

$$zv = \chi(z)v, \qquad v \in V,$$

where

$$\chi : Z(L) \to C$$
$$z \mapsto \chi(z)$$

is an algebra homomorphism. This condition on V is not very restrictive because Dixmier [8] has shown that every irreducible L-module admits a character. A well-known theorem proved by Harish-Chandra [7] asserts that any character χ is equal to the character χ_μ of an irreducible L-module $V(\mu)$ with highest weight $\mu \in H^*$, However, μ is only determined up to linkage, where $\mu, \mu' \in H^*$ are linked, written $\mu \sim \mu'$, if there is some $\sigma \in W$ such that $\mu + \delta = \sigma(\mu' + \delta)$.

We denote the representations of L afforded by $V(\lambda)$ and V by π_λ and π respectively, and we let

$$\{ \nu_1, \nu_2, \ldots, \nu_m \}$$

be the set of distinct weights of $V(\lambda)$.

Kostant considers the algebra

$$(\text{End } V(\lambda)) \otimes U(L)$$

and defines

$$\partial : U(L) \to (\text{End } V(\lambda)) \otimes U(L)$$

by

$$x \mapsto \partial x = 1 \otimes x + \pi_\lambda(x) \otimes 1 \qquad (\text{for } x \in L).$$

He extends ∂ to a homomorphism of algebras and proves the following theorems.

THEOREM 4.9 [6]. — Let $z \in Z(L)$. Then there exists a unique monic polynomial $p_z(x) \in Z(L)[x]$ such that

$$p_z(\partial z) = 0. \tag{3}$$

THEOREM 5.1 [6]. — Let $z \in Z(L)$ and put $\tilde{z} = (\pi_\lambda \otimes \pi)(z)$. The operator $\tilde{z}$ in $V(\lambda) \otimes V$ satisfies the identity

$$\prod_{i=1}^{m} (\tilde{z} - \chi_{\nu_i + \mu}(z)) = 0. \tag{4}$$

In order to relate these results to previous work on characteristic identities, we take the special case $z = c_L$, the universal Casimir element of L. We know that [7]

$$\chi_\lambda(c_L) = (\lambda, \lambda + 2\delta), \tag{5}$$

where we have used the bilinear form (,) induced on H* by the Killing form of L. The identity (4) then becomes

$$\prod_{i=1}^{m} (\tilde{c}_L - (v_i + \mu, v_i + \mu + 2\delta)) = 0. \tag{6}$$

Following Hannabuss (unpublished), we now define the operator A_V on $V(\lambda) \otimes V$ by

$$A_V = \frac{1}{2}(\tilde{c}_L - \pi_\lambda(c_L) \otimes 1 - 1 \otimes \pi(c_L)). \tag{7}$$

It is straightforward to verify that A_V commutes with the L-action on $V(\lambda) \otimes V$. Furthermore, by (6),

$$\prod_{i=1}^{m} (2A_V + (\lambda, \lambda + 2\delta) + (\mu, \mu + 2\delta) - (v_i + \mu, v_i + \mu + 2\delta)) = 0,$$

which reduces to

$$\prod_{i=1}^{m} (A_V + \frac{1}{2}(\lambda, \lambda + 2\delta) - \frac{1}{2}(v_i, v_i + 2(\mu + \delta))) = 0. \tag{8}$$

This is the polynomial identity which appears in the work of Bracken and Green [3], Green [4] and Hannabuss, under the assumption that V is a finite dimensional irreducible L-module with highest weight μ. However Kostant's results show that the identity (8) is valid for any V which admits a character.

Similarly, we can consider the operator

$$A_L = \frac{1}{2}(\partial c_L - \pi_\lambda(c_L) \otimes 1 - 1 \otimes c_L) \tag{9}$$

as an element of $(\text{End } V(\lambda)) \otimes U(L)$. Then, using (3), we obtain

$$p_{c_L}(2A_L + \pi_\lambda(c_L) \otimes 1 + 1 \otimes c_L) = 0. \tag{10}$$

We call (10) the characteristic identity. When π_λ is the natural representation of a classical Lie algebra, A_L is essentially the matrix over the Lie algebra considered by Bracken and Green.

We see that Kostant's results link the various accounts of polynomial identities and show, in particular, that the characteristic identity does not depend on the choice of basis of L. This is so because c_L is independent of the choice of basis. Nevertheless, it is of interest to write A_L in terms of a basis $\{ x_1, x_2, \ldots, x_k \}$ ($k = \dim L$), of L and its dual $\{ x^1, x^2, \ldots, x^k \}$ with respect to Killing form of L. With

$$c_L = \sum_i x_i x^i ,$$

we obtain

$$A_L = \frac{1}{2} \sum_i (\pi_\lambda(x^i) \otimes x_i + \pi_\lambda(x_i) \otimes x^i) . \tag{11}$$

For the particular case where $L = gl(n, Q)$, we choose the basis $\{ a^i{}_j \}$ and take for π_λ the $n \times n$ matrix representation in which

$$\pi_\lambda(a^i{}_j) = - E^j{}_i ,$$

where $E^i{}_j$ is the $n \times n$ matrix whose only non-zero element is equal to one and lies at the intersection of the ith row and the jth column. We find

$$A_L = \sum_{i,j} a^i{}_j E^i{}_j .$$

This matrix agrees with the matrix A introduced previously. (Note that in this case the number m of distinct weights of π_λ coincides with the rank n of $gl(n, Q)$.)

Now suppose that L_R is a real form of L. The universal enveloping algebra $U(L_R)$ of L_R and its centre $Z(L_R)$ are naturally embedded in U(L) and Z(L) respectively. It is clear from the proof of theorem 4.9 in [6] that, when $z \in Z(L_R)$, the coefficients in the polynomial p_z are also elements of $Z(L_R)$. Thus, Kostant's theorem can be used to derive characteristic identities for the real forms of complex semi-simple Lie algebras.

On the basis of the results derived for $gl(n, Q)$ in this paper, we conjecture that Kostant's results apply to reductive Lie algebras over any subfield of the complex field which includes the rational field.

Henceforth we shall consider only the case of $gl(n, Q)$ and adhere to the notational conventions established in the introduction.

3. EXISTENCE OF THE CHARACTERISTIC IDENTITY

Lehrer-Ilamed [2] has shown that the elements of any 3×3 matrix over a non-commutative algebra satisfy nine identities. His proof can be generalized so that it applies to any $n \times n$ matrix. The starting point is the classical Cayley-Hamilton theorem.

THEOREM. — Suppose that K is a commutative algebra over Q. If

$$B = (b^i{}_j) = \begin{pmatrix} b^1{}_1 b^1{}_2 \dots b^1{}_n \\ b^2{}_1 b^2{}_2 \dots b^2{}_n \\ \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \\ b^n{}_1 b^n{}_2 \dots b^n{}_n \end{pmatrix}$$

is any $n \times n$ matrix over K, then B satisfies

$$B^n + b_1 B^{n-1} + \dots + b_{n-1} B + b_n = 0,$$

where

$$b_1 = - \text{trace } B$$

and

$$kb_k = - \text{trace } B^k - \sum_{j=1}^{k-1} b_j \text{ trace } B^{k-j}, \qquad 1 < k \leqslant n.$$

This result is well known (Greub [9]), but there are several points which should be noted.

1) b_k is a polynomial with rational coefficients in trace B, trace B^2, ..., trace B^k. More precisely,

$$kb^k + \text{trace } B^k \in Q\,[\text{trace } B, \text{trace } B^2, \dots, \text{trace } B^{k-1}], \tag{12}$$

where $Q[x_1, x_2, \dots, x_{k-1}]$ denotes the ring of polynomials over Q in indeterminates $x_1, x_2, \dots, x_{k-1}$.

2) Each term $b_k B^{n-k}$ is a homogeneous polynomial of degree n in the elements of B.

3) The single matrix identity provides n^2 identities between the elements of B. The (i, j) identity can be written

$$\left\{ \begin{matrix} i \\ j \end{matrix} \middle| \begin{matrix} j_1 j_2 \dots j_n \\ i_1 i_2 \dots i_n \end{matrix} \right\} b^{i_1}{}_{j_1} b^{i_2}{}_{j_2} \dots b^{i_n}{}_{j_n} = 0. \tag{13}$$

In this identity,

$$\left\{ \begin{matrix} i \\ j \end{matrix} \middle| \begin{matrix} j_1 j_2 \dots j_n \\ i_1 i_2 \dots i_n \end{matrix} \right\} = \sum_{m=0}^{n} \varepsilon_m \begin{pmatrix} j_1 j_2 \dots j_m \\ i_1 i_2 \dots i_m \end{pmatrix} \theta_m \begin{pmatrix} i \\ j \end{pmatrix|} \begin{matrix} j_{m+1} j_{m+2} \dots j_n \\ i_{m+1} i_{m+2} \dots i_n \end{matrix} \Bigg),$$

where

$$\varepsilon_0 = 1,$$

$$\varepsilon_m \begin{pmatrix} j_1 j_2 \dots j_m \\ i_1 i_2 \dots i_m \end{pmatrix} = \frac{(-1)^m}{m!} \det \begin{pmatrix} \delta^{j_1}{}_{i_1} \delta^{j_2}{}_{i_1} \dots \delta^{j_m}{}_{i_1} \\ \delta^{j_1}{}_{i_2} \delta^{j_2}{}_{i_2} \dots \delta^{j_m}{}_{i_2} \\ \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \\ \delta^{j_1}{}_{i_m} \delta^{j_2}{}_{i_m} \dots \delta^{j_m}{}_{i_m} \end{pmatrix}, \qquad 1 \leqslant m \leqslant n,$$

and

$$\theta_n\begin{pmatrix} i \\ j \end{pmatrix} = \delta^i{}_j ,$$

$$\theta_m\left(\begin{array}{c|c} i & j_{m+1}j_{m+2} \ \dots \ j_n \\ j & i_{m+1}i_{m+2} \ \dots \ i_n \end{array}\right) = \delta^i_{i_{m+1}}\delta^{j_{m+1}}_{i_{m+2}} \dots \delta^{j_{n-1}}_{i_n}\delta^{j_n}_j , \qquad 0 \leqslant m < n .$$

The precise form of

$$\left\{\begin{array}{c|c} i & j_1j_2 \ \dots \ j_n \\ j & i_1i_2 \ \dots \ i_n \end{array}\right\}$$

is not essential for the arguments which follow; it is sufficient to know that for each (i, j) it is an array of rational numbers.

Equation (13) possesses a simple generalization which holds when the assumption that K should be commutative is relaxed.

Suppose that S_1 is a finite dimensional vector space over Q with basis

$$\{ s_1, s_2, \dots, s_m \} , \qquad m \geqslant 1 .$$

Set

$$S_0 = Q ,$$
$$S_p = S_1 \otimes S_1 \otimes \dots \otimes S_1 \ (p \text{ factors}, \ p \geqslant 1) ,$$

and construct

$$T = \bigoplus_{p=0}^{\infty} S_p ,$$

the tensor algebra based on S_1. Define

$$\langle s_{i_1}s_{i_2} \ \dots \ s_{i_p} \rangle = \frac{1}{p!}\sum_{\pi} s_{i_{\pi(1)}}s_{i_{\pi(2)}} \ \dots \ s_{i_{\pi(p)}} ,$$

where

$$1 \leqslant i_q \leqslant m , \qquad 1 \leqslant q \leqslant p ,$$

and the summation is over all permutations of the set $\{ 1, 2, \dots, p \}$. If $v_1, v_2, \dots, v_p$ are arbitrary vectors in S_1 with

$$v_q = \sum_{i_q=1}^{m} \beta_q^{i_q} s_{i_q} , \qquad \beta_q^{i_q} \in Q , \qquad 1 \leqslant q \leqslant p ,$$

define

$$\langle v_1v_2 \ \dots \ v_p \rangle = \sum_{i_1=1}^{m}\sum_{i_2=1}^{m} \dots \sum_{i_p=1}^{m} \beta_1^{i_1}\beta_2^{i_2} \dots \beta_p^{i_p} \langle s_{i_1}s_{i_2} \ \dots \ s_{i_p} \rangle .$$

It is obvious that $\langle v_1v_2 \ \dots \ v_p \rangle$ is independent of the order in which $v_1v_2, \dots, v_p$ are prescribed.

THEOREM 1. — If $B = (b^i{}_j)$ is an $n \times n$ matrix whose elements are chosen from S_1, possibly with repetitions, then

$$\left\{ \begin{array}{c|c} i & j_1 j_2 \ldots j_n \\ j & i_1 i_2 \ldots i_n \end{array} \right\} \langle b^{i_1}{}_{j_1} b^{i_2}{}_{j_2} \ldots b^{i_n}{}_{j_n} \rangle = 0 . \tag{14}$$

Proof. — Set $S^0 = Q$ and $S^1 = S_1$. The pth symmetric power of S^1, denoted S^p, is the subspace of T with basis

$$\{ \langle s_{i_1} s_{i_2} \ldots s_{i_p} \rangle , \qquad 1 \leqslant i_1 \leqslant i_2 \leqslant \ldots \leqslant i_p \leqslant m \} .$$

Construct the graded vector space

$$S = \bigoplus_{p=0}^{\infty} S^p .$$

Each element of S, say (ω_p), may be regarded as an infinite sequence

$$(\omega_p) = (\omega_0, \omega_1, \omega_2, \ldots)$$

with $\omega_p \in S_p (p = 0, 1, 2, \ldots)$ and only finitely many ω_p non-zero. Addition and scalar multiplication in S are performed componentwise. Define a bilinear map $\times : S^p \times S^q \to S^{p+q}$ by

$$\langle s_{i_1} s_{i_2} \ldots s_{i_p} \rangle \times \langle s_{i_{p+1}} s_{i_{p+2}} \ldots s_{i_{p+q}} \rangle = \langle s_{i_1} s_{i_2} \ldots s_{i_{p+q}} \rangle$$

and then extend $\times$ to a bilinear map $\times : S \times S \to S$ with the further definition

$$(w^{(1)}_{p_1}) \times (w^{(2)}_{p_2}) = \left(w_p = \sum_{p_1 + p_2 = p} w^{(1)}_{p_1} \times w^{(2)}_{p_2} \right) .$$

Equipped with this product, S is a commutative algebra over Q. The matrix B is a matrix over S, provided $b^i{}_j = \langle b^i{}_j \rangle$ is identified with

$$(0, \langle b^i{}_j \rangle, 0, 0, \ldots) ,$$

so the Cayley-Hamilton theorem asserts that

$$\left\{ \begin{array}{c|c} i & j_1 j_2 \ldots j_n \\ j & i_1 i_2 \ldots i_n \end{array} \right\} \langle b^{i_1}{}_{j_1} \rangle \times \langle b^{i_2}{}_{j_2} \rangle \times \ldots \times \langle b^{i_n}{}_{j_n} \rangle = 0 . \tag{15}$$

Since

$$\langle b^{i_1}{}_{j_1} \rangle \times \langle b^{i_2}{}_{j_2} \rangle \times \ldots \times \langle b^{i_n}{}_{j_n} \rangle = \langle b^{i_1}{}_{j_1} b^{i_2}{}_{j_2} \ldots b^{i_n}{}_{j_n} \rangle ,$$

equation (14) follows directly from (15) and the proof of the theorem is complete.

As a corollary of this theorem, the elements of any matrix B over the universal enveloping algebra of any Lie algebra over Q satisfy the n^2 identifies (14). To prove this, simply let S_1 be the subspace spanned by the elements of B. For the matrix A over U, this result can be sharpened.

THEOREM 2. — Set $t_k = \text{trace } A^k$. The matrix A over U satisfies the polynomial identity

$$p(A) = A^n + c_1 A^{n-1} + \dots + c_{n-1}A + c_n = 0\,, \tag{16}$$

in which

$$\begin{aligned} &c_1 + t_1 \in Q\,,\\ &2c_2 + t_2 \in Q[t_1]\,,\\ &\dots\dots\dots\dots\\ &nc_n + t_n \in Q[t_1, t_2, \dots, t_{n-1}]\,. \end{aligned} \tag{17}$$

$p(A)$ is called the characteristic polynomial of A.

Proof. — Set

$$b_1 = -\,t_1$$
$$kb_k = -\,t_k - \sum_{j=1}^{k-1} b_j t_{k-j}\,, \qquad 1 < k \leqslant n\,.$$

Thus, $b_1, b_2, \dots, b_n$ are the coefficients which appear in the Cayley-Hamilton theorem and certainly satisfy (17). According to theorem 1, the elements of A satisfy

$$\left\{ \begin{array}{c|c} i & j_1 j_2 \dots j_n \\ j & i_1 i_2 \dots i_n \end{array} \right\} \sum_\pi a^{i_{\pi(1)}}{}_{j_{\pi(1)}} a^{i_{\pi(2)}}{}_{j_{\pi(2)}} \dots a^{i_{\pi(n)}}{}_{j_{\pi(n)}} = 0\,, \tag{18}$$

where π denotes a permutation of $\{ 1, 2, \dots, n \}$. Alternatively,

$$\sum_{m=0}^{n} \varepsilon_m \begin{pmatrix} j_1 j_2 \dots j_m \\ i_1 i_2 \dots i_m \end{pmatrix} \langle\, a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \dots a^{i_m}{}_{j_m} (A^{n-m})^i{}_j \,\rangle = 0\,. \tag{19}$$

The commutation relations (1) permit (19) to be cast in the form

$$h^i{}_j + q^i{}_j = 0\,,$$

where

$$h^i{}_j = h(A)^i{}_j\,,$$
$$h(A) = A^n + b_1 A^{n-1} + \dots + b_{n-1}A + b_n\,,$$

and $q^i{}_j$ is the sum of all those terms produced by reordering the factors in (18). $h^i{}_j$ is a homogeneous polynomial in the elements of A with degree n, but $q^i{}_j$ is not homogeneous and its degree is $(n - 1)$ at most. Write

$$q^i{}_j = \sum_{m=0}^{n} (q_m)^i{}_j\,,$$

where $(q_m)^i{}_j$ denotes the contribution to $q^i{}_j$ from

$$\varepsilon_m \begin{pmatrix} j_1 j_2 \dots j_m \\ i_1 i_2 \dots i_m \end{pmatrix} \langle\, a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \dots a^{i_m}{}_{j_m} (A^{n-m})^i{}_j \,\rangle\,, \qquad 0 \leqslant m \leqslant n\,. \tag{20}$$

$(q_m)^i{}_j$ is a sum of monomials of the form

$$1 \text{ or } a^{i_1}{}_{j_1}a^{i_2}{}_{j_2} \ldots a^{i_p}{}_{j_p}, \qquad 1 \leqslant p < n.$$

The coefficient of each monomial ensures that one superscript is equal to i, one subscript is equal to j, and each of the remaining superscripts is contracted with one of the remaining subscripts. Thus $(q_m)^i{}_j$ is the (i, j) element of a polynomial in A with coefficients in $Q[t_1, t_2, \ldots, t_n]$;

$$(q_m)^i{}_j = (q_m(A))^i{}_j, \qquad 0 \leqslant m \leqslant n.$$

It follows from (20) that $q_m(A)$ has degree $(n - m)$ at most, so

$$q_m(A) = \sum_{k=m}^{n} d_{mk}A^{n-k}.$$

The only term in (19) which involves A^n has been absorbed into $h(A)$, so

$$d_{00} = 0.$$

Furthermore,

$$d_{mk} \in \begin{cases} Q, & \text{if } k = 1, \\ Q[t_1, t_2, \ldots, t_{k-1}], & \text{if } k > 1, \end{cases}$$

because $(q_m)^i{}_j$ has degree $(n - 1)$ at most when regarded as a polynomial in the elements of A. If

$$c_k = b_k + \sum_{m=0}^{k} d_{mk},$$

then A satisfies (16). Finally, it is clear that $c_1, c_2, \ldots, c_n$ satisfy (17).

A similar proof can be devised for the case in which the matrix A is replaced by the matrix $A_L \in (\text{End } V(\lambda)) \otimes U$ appropriate to a tensor representation of $gl(n, Q)$ in $V(\lambda)$. Naturally the proof is more complicated, and so will not be given here. However, Green [*4*] does treat this extension.

4. FACTORIZATION OF THE CHARACTERISTIC POLYNOMIAL

The characteristic polynomial of A cannot be split into linear factors in the algebra of $n \times n$ matrices over U. Nevertheless, it is in factorized form that $p(A)$ is most useful for the construction of irreducible $gl(n, Q)$-modules. In this section we shall injectively embed U in an algebra $\bar{U}$ over Q so that $p(A)$ factorizes in the algebra of $n \times n$ matrices over $\bar{U}$.

LEMMA 1. — If $q(A)$ is any polynomial in A with rational coefficients, then

$$[q(A)^i{}_j, a^k{}_l] = \delta^k{}_j q(A)^i{}_l - \delta^i{}_l q(A)^k{}_j.$$

Proof. — The result is obvious when the degree of $q(\mathrm{A})$ is equal to one. Induction on the degree of $q(\mathrm{A})$ establishes the result generally.

LEMMA 2.

$$\mathrm{Z} = \mathrm{Q}[t], \qquad \text{where} \qquad t = (t_1, t_2, \ldots, t_n).$$

Proof. — According to lemma 1,

$$\begin{aligned}[t_m, a^k{}_l] &= [(\mathrm{A}^m)^i{}_i, a^k{}_l] \\ &= \delta^k{}_i(\mathrm{A}^m)^i{}_l - \delta^i{}_l(\mathrm{A}^m)^k{}_i \\ &= 0.\end{aligned}$$

Thus, $t_m \in \mathrm{Z}$ and so $\mathrm{Q}[t] \subseteq \mathrm{Z}$.

If $z \in \mathrm{Z}$, then z cannot have any free indices. Every fully contracted element of U can be reduced to a polynomial in $t_1, t_2, \ldots$ by repeated use of the commutation relations. Because A satisfies an nth order polynomial identity with coefficients in $\mathrm{Q}[t]$, the elements $t_{m+1}, t_{m+2}, \ldots$ must also lie in $\mathrm{Q}[t]$. Hence, $\mathrm{Z} \subseteq \mathrm{Q}[t]$.

Now, Z is an integral domain, because Q is a field. The field of quotients of Z can be identified with $\mathrm{Q}(t)$, the field of rational functions of t with rational coefficients. Let $\bar{\mathrm{Z}}$ denote the minimal extension field of $\mathrm{Q}(t)$ in which the polynomial $p(x)$ splits into linear factors:

$$p(x) = x^n + c_1 x^{n-1} + \ldots + c_{n-1}x + c_n \tag{21}$$

$$= \prod_{i=1}^{n}(x - a_i), \qquad a_i \in \bar{\mathrm{Z}}. \tag{22}$$

Finally, set $a = (a_1, a_2, \ldots, a_n)$.

LEMMA 3. — $\bar{\mathrm{Z}} = \mathrm{Q}(a)$.

Proof. — Jacobson [*10*] shows that $\bar{\mathrm{Z}} = \mathrm{Q}(t, a)$. It is obvious that

$$\mathrm{Q}(a) \subseteq \mathrm{Q}(t, a),$$

so it is only necessary to prove the opposite inclusion. Each of the coefficients $c_1, c_2, \ldots, c_n$ in (21) can be expressed as a symmetric function of $a_1, a_2, \ldots, a_n$, denoted by

$$c_k = c_k(a) \in \mathrm{Q}(a).$$

Theorem 1 asserts that

$$c_1 + t_1 \in \mathrm{Q} \subseteq \mathrm{Q}(a).$$

Thus,

$$t_1 \in \mathrm{Q}(a).$$

Similarly,

$$2c_2 + t_2 \in \mathrm{Q}[t_1] \subseteq \mathrm{Q}(a),$$

so

$$t_2 \in \mathrm{Q}(a).$$

The argument can be repeated to prove

$$t_k \in \mathrm{Q}(a), \qquad 1 \leqslant k \leqslant n.$$

Hence,

$$\mathrm{Q}(t, a) \subseteq \mathrm{Q}(a),$$

and so

$$\bar{\mathrm{Z}} = \mathrm{Q}(a).$$

Although the polynomial $p(x)$ can be factorized in $\bar{\mathrm{Z}}$, it is not correct to say that $p(\mathrm{A})$ can also be factorized, because a meaning has not yet been assigned to products ua_i and a_iu, $u \in \mathrm{U}$. However, this is easily remedied. Let

$$\bar{\mathrm{U}} = \bar{\mathrm{Z}} \otimes_{\mathrm{Z}} \mathrm{U}, \tag{23}$$

where the tensor product is that of Z-modules. The maps

$$\phi : \mathrm{U} \to \bar{\mathrm{U}} \qquad \text{and} \qquad \psi : \bar{\mathrm{Z}} \to \bar{\mathrm{U}}$$
$$u \to 1 \otimes u \qquad\qquad z \to \bar{z} \otimes 1$$

are injections and, once $\bar{\mathrm{U}}$ is endowed with the obvious multiplication, they become injective algebra homomorphisms. The factorization of $p(\mathrm{A})$ now follows simply.

THEOREM 3. — The matrix A, considered as a matrix over $\bar{\mathrm{U}}$, satisfies

$$\prod_{i=1}^{n} (\mathrm{A} - a_i) = 0. \tag{24}$$

$a_1, a_2, \ldots, a_n$ are called the eigenvalues of A.

5. RELATIONSHIP BETWEEN EIGENVALUES OF A AND HIGHEST WEIGHTS

Let F be any field satisfying $\mathrm{Q} \subseteq \mathrm{F} \subseteq \mathrm{C}$, and let H be the Cartan subalgebra of $gl(n, \mathrm{Q})$ with basis

$$\{ a^i{}_i \mid 1 \leqslant i \leqslant n \}.$$

DEFINITION. — Suppose that V is an irreducible $\mathrm{U_F}$-module which contains a vector v_0 such that

$$a^i{}_j v_0 = 0, \qquad i < j,$$
$$a^i{}_i v_0 = \lambda_i v_0, \qquad \lambda_i \in \mathrm{F} \qquad \text{(no summation)}.$$

v_0 is called a maximal vector of highest weight λ, where λ is the linear function from H to F given by $a^i{}_i \to \lambda_i$. We shall write $\mathrm{V_F}(\lambda)$ for such a V.

The main result of this section is the following theorem, whose proof will be given in a sequence of lemmas.

THEOREM 4. — Suppose $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in F^n$. Up to isomorphism there is only one irreducible U_F-module $V_F(\lambda)$ with highest weight λ. The eigenvalues of A can be reordered so that

$$a_k v = \alpha_k v \,, \qquad \text{for all} \qquad v \in V_F(\lambda) \,,$$

where

$$\alpha_k = \lambda_k + n - k \,. \tag{25}$$

Note that the expression (25) is compatible with the expression which would be expected on the basis of Kostant's results, namely,

$$\alpha_k = \frac{1}{2}(\nu_k, \nu_k + 2(\lambda + \delta)) - \frac{1}{2}(\lambda, \lambda + 2\delta) \,.$$

Here, $\nu_1, \nu_2, \ldots, \nu_n$ are the weights of the contragredient of the natural representation of $gl(n, Q)$.

Let us begin with the case F = C. Harish Chandra has proven the existence of an irreducible U_C-module $V_C(\lambda)$ with highest weight λ [*11*].

LEMMA 4. — Suppose that v_0 is the maximal vector of $V_C(\lambda)$ and that $s(x)$ is a polynomial in $\bar{Z}[x]$.

1) $[s(A)^i{}_j, a^k{}_l] = \delta^k{}_j s(A)^i{}_l - \delta^i{}_l s(A)^k{}_j \,.$

2) $s(A)^i{}_j v_0 = 0 \qquad \text{if } i < j \,.$

3) $s(A)^i{}_i v_0 = \sigma_i v_0 \,, \qquad \sigma_i \in C \qquad \text{(no summation)}.$

Proof. — All three results are obvious if the degree of $s(x)$ is one. Induction on the degree of $s(x)$ then establishes the results generally.

We now set

$$s_j(k) = \left[\prod_{i=j}^{n}(A - a_i)\right]^k_{\# k} \qquad \text{(no summation)}.$$

Lemma 4 shows that v_0 is an eigenvector of $s_j(k)$ with some eigenvalue $\sigma_j(k) \in C$. Because C is algebraically closed and $V_C(\lambda)$ is irreducible, the eigenvalues of A are represented by scalars;

$$a_k v = \alpha_k v \,, \qquad \text{for all} \qquad v \in V_C(\lambda) \,,$$

where $\alpha_k \in C$.

LEMMA 5. — The eigenvalues $\sigma_j(k)$ of $s_j(k)$ satisfy the difference equation,

$$\sigma_j(k) = \sigma_{j+1}(k)(\lambda_k - \alpha_j + n - k) - \sigma_{j+1}(k+1) - \ldots - \sigma_{j+1}(n) \,,$$

and the following boundary conditions,

$$\sigma_n(k) = \lambda_k - \alpha_n \,,$$

$$\sigma_j(n) = \prod_{i=j}^{n}(\lambda_n - \alpha_i) \,.$$

Proof. — We have

$$s_j(k)v_0 = \left[\prod_{i=j+1}^{n}(\mathrm{A} - a_i)\right]^k_{\#l}(a^l_{\ k} - a_j\delta^l_{\ k})v_0 .$$

Since $a^l_{\ k}v_0 = 0$ for $l < k$, the summation over l can be restricted to the range $l \geqslant k$. Thus,

$$s_j(k)v_0 = \sum_{l \geqslant k}\left\{ (a^l_{\ k} - a_j\delta^l_{\ k})\left[\prod_{i=j+1}^{n}(\mathrm{A} - a_i)\right]^k_{\#l} - s_{j+1}(k) - s_{j+1}(l) \right\} v_0 ,$$

where part (1) of Lemma 4 has been used to commute the factors. Since

$$\left[\prod_{i=j+1}^{n}(\mathrm{A} - a_i)\right]^k_{\#l} v_0 = 0 , \qquad k < l ,$$

it follows that

$$s_j(k)v_0 = (a^k_{\ k} - a_j + n - k)s_{j+1}(k)v_0 - s_{j+1}(k+1)v_0 - \ \dots \ - s_{j+1}(n)v_0 .$$

Hence,

$$\sigma_j(k) = \sigma_{j+1}(k)(\lambda_k - \alpha_j + n - k) - \sigma_{j+1}(k+1) - \ \dots \ - \sigma_{j+1}(n) .$$

It is clear that

$$\sigma_n(k) = \lambda_k - \alpha_n . \tag{26}$$

Also, $\sigma_j(n)$ satisfies the difference equation

$$\sigma_j(n) = \sigma_{j+1}(n)(\lambda_n - \alpha_j)$$

which has the solution

$$\sigma_j(n) = \prod_{i=j}^{n}(\lambda_n - \alpha_i) ,$$

consistent with (26).

We note now that the ordering of $a_1, a_2, \dots, a_n$ is arbitrary, because these quantities appear as the roots of a polynomial equation.

LEMMA 6. — We may reorder $a_1, a_2, \dots, a_n$ so that

$$\sigma_j(k) = 0, \qquad 1 \leqslant j \leqslant k .$$

Furthermore, if $\sigma_j(j-1)$ is non-zero for $k < j < n$, then

$$\lambda_j = \alpha_j - n + j , \qquad k \leqslant j \leqslant n .$$

Proof. — Because $s_1(\mathrm{A}) = 0$, $\sigma_1(k) = 0$ for all k.

Now,

$$\sigma_1(n) = \prod_{i=1}^{n}(\lambda_n - \alpha_i) = 0\,,$$

so λ_n must be equal to one of the scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$. Relabel $a_1, a_2, \ldots, a_n$ so that $\lambda_n = \alpha_n$. Then,

$$\sigma_j(n) = \prod_{i=j}^{n}(\lambda_n - \alpha_i) = 0\,, \qquad 1 \leqslant j \leqslant n\,.$$

Now use induction. Suppose that $a_1, a_2, \ldots, a_n$ have been reordered so that $\sigma_j(k) = 0$ for $1 \leqslant j \leqslant k$. Since

$$\sigma_j(k-1) = \sigma_{j+1}(k-1)(\lambda_{k-1} - \alpha_j + n - k + 1) - \sigma_{j+1}(k) - \ldots - \sigma_{j+1}(n)\,,$$

we have, for $1 \leqslant j \leqslant k - 1$,

$$\sigma_j(k-1) = \sigma_{j+1}(k-1)(\lambda_{k-1} - \alpha_j + n - k + 1)\,.$$

Thus,

$$\sigma_1(k-1) = \sigma_k(k-1)\prod_{i=1}^{k-1}(\lambda_{k-1} - \alpha_i + n - k + 1) = 0\,,$$

and so either $\sigma_k(k-1) = 0$, or

$$\prod_{i=1}^{k-1}(\lambda_{k-1} - \alpha_i + n - k + 1) = 0\,,$$

or both. In the first case, we can say nothing. In the second case, relabel the roots $a_1, a_2, \ldots, a_{k-1}$ so that

$$\lambda_{k-1} = \alpha_{k-1} - n + k - 1\,.$$

Since

$$\sigma_j(k-1) = \sigma_k(k-1)\prod_{i=j}^{k-1}(\lambda_{k-1} - \alpha_i + n - k + 1)\,,$$

we have in either case that

$$\sigma_j(k-1) = 0\,, \qquad 1 \leqslant j \leqslant k - 1\,.$$

Note that the relabelling of $a_1, a_2, \ldots, a_{k-1}$ does not alter the labelling of $a_k, a_{k+1}, \ldots, a_n$. Hence, the lemma is established.

LEMMA 7. — Suppose that $V_C(\lambda)$ is such that $\sigma_k(k-1)=0$ but $\sigma_j(j-1)\neq 0$

for any $j > k$. Then there is a polynomial $q_k \in C[x_{k-1}, x_k, \ldots, x_n]$, where $x_{k-1}, x_k, \ldots, x_n$ are indeterminates, for which

$$q_k(\lambda_{k-1}, \lambda_k, \ldots, \lambda_n) = 0.$$

If C^n is equipped with the Zariski topology, the set of $\lambda \in C^n$ such that $\sigma_j(j-1) \neq 0$ for any j is dense in C^n.

Proof.—$\sigma_k(k-1)$ is a polynomial in $\alpha_k, \alpha_{k+1}, \ldots, \alpha_n$ and $\lambda_{k-1}, \lambda_k, \ldots, \lambda_n$ with coefficients in C. Because $\sigma_j(j-1) \neq 0$ for $k < j$, lemma 6 asserts that

$$\alpha_j = \lambda_j + n - j, \qquad j \geqslant k.$$

Hence, $\sigma_k(k-1)$ can be rewritten as a polynomial in $\lambda_{k-1}, \lambda_k, \ldots, \lambda_n$. Denote this polynomial by q_k. Let I_k denote the ideal in $C[x_1, x_2, \ldots, x_n]$ generated by q_k and let M_k denote the closed subset of C^n determined by I_k. Then $\bigcup_{k=1}^{n} M_k$ is also closed. Its complement M_0. which is open, is necessarily non-empty, for otherwise we would cover C^n by closed sets, none of which is equal to C^n, contradicting a well known property of the Zariski topology [7]. M_0 consists of those elements of C^n for which $\sigma_j(j-1) \neq 0$ for any j. Any non-empty open subset of C^n is dense in the Zariski topology.

Define an equivalence relation on F^n as follows. If $\alpha, \beta \in F^n$, write $\alpha \sim \beta$ if and only if

$$\alpha_k = \beta_{\pi(k)}, \qquad 1 \leqslant k \leqslant n,$$

where π is a permutation of the set $\{1, 2, \ldots, n\}$. Let $\hat{\alpha}$ denote the equivalence class which contains α, and let $\hat{F}^n$ denote the set of all equivalence classes. Lastly, if g is a symmetric polynomial function on F^n, define $\hat{g} : \hat{F}^n \to F$ by $g(\alpha) = \hat{g}(\hat{\alpha})$. Note that the symmetry of g is needed if $\hat{g}$ is to be well defined.

We have constructed a map $f_C : C^n \to \hat{C}^n$ where $f_C(\lambda) = \hat{\alpha}$, by taking $\lambda \in C^n$ to the uniquely determined module $V_C(\lambda)$ on which

$$a_k v = \alpha_k v, \qquad \text{for all} \qquad v \in V_C(\lambda),$$

where $\alpha_k \in C$, and $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ uniquely determines $\hat{\alpha}$. We have shown that $\alpha_1, \alpha_2, \ldots, \alpha_n$ can be reordered so that

$$\alpha_k = \lambda_k + n - k$$

on a (Zariski) dense subset of C^n. The next step will be to equip $\hat{C}^n$ with a suitable topology and show that f_C is continuous.

Let $\hat{C}[x]$ denote the ring of symmetric polynomials in indeterminates $x = (x_1, x_2, \ldots, x_n)$. If $\hat{I}$ is any ideal in $\hat{C}[x]$, let

$$\hat{M}(\hat{I}) = \{\hat{\alpha} \in \hat{C}^n \mid \hat{g}(\hat{\alpha}) = 0, \qquad \forall g \in \hat{I}\}.$$

We impose a topology on $\hat{C}^n$ by declaring the sets $\hat{M}(\hat{I})$ to be closed.

LEMMA 8. — f_C is continuous.

Proof. — Consider the following polynomial in C[x, y],

$$\prod_k (y - x_k) = y^n + s_1 y^{n-1} + \ldots + s_{n-1} y + s_n ,$$

where

$$s_k = s_k(x) \in C[x] .$$

It is well known that every symmetric polynomial in $x_1, x_2, \ldots, x_n$ can rearranged as a polynomial in $s_1(x), s_2(x), \ldots, s_n(x)$. Thus $\hat{C}[x] = C[s(x)]$ where $s = (s_1, s_2, \ldots, s_n)$. Theorem 2 shows that

$$s_1(a) + t_1 \in Q ,$$

$$ks_k(a) + t_k \in Q[t_1, t_2, \ldots, t_{k-1}] , \qquad 1 < k \leqslant n ,$$

so $C[s(a)] = C[t]$. Hence $\hat{C}[a] = C[t]$. If $g \in \hat{C}[a]$, let g' denote the same polynomial regarded as an element of $C[t]$.

Suppose that $\hat{M}(\hat{I})$ is a closed set in $\hat{C}^n$ and that $M \subseteq C^n$ is its preimage under f_C. This, if $\hat{\alpha} \in \hat{M}(\hat{I})$, there exists $\lambda \in M$ such that

$$a_k v = \alpha_k v , \qquad \text{for all} \qquad v \in V_C(\lambda) .$$

If $g \in \hat{I}$, then

$$\begin{aligned} 0 &= g(\alpha) v_0 , \\ &= g(a) v_0 , \\ &= g'(t) v_0 . \end{aligned}$$

By direct calculation we see that

$$t_1 v_0 = a^i{}_i v_0 = \sum_i \lambda_i v_0 ,$$

$$t_2 v_0 = a^i{}_j a^j{}_i v_0 = \sum_i \lambda_i(\lambda_i + n + 1 - 2i) v_0 ,$$

and in general

$$t_k v_0 = \tau_k(\lambda) v_0 ,$$

where $\tau_k(x) \in C[x]$. Hence, if $\tau = (\tau_1, \tau_2, \ldots, \tau_n)$,

$$g'(\tau(\lambda)) = 0 .$$

Let I denote the ideal in C[x] generated by the set

$$\{ g'(\tau(x)) \mid g \in \hat{I} \} .$$

We have established that, if $\lambda \in M$, then

$$h(\lambda) = 0 , \qquad \text{for all} \qquad h \in I ,$$

so M is closed in the Zariski topology of C^n. Thus, f_C is continuous.

COROLLARY. — The roots $a_1, a_2, \ldots, a_n$ of the characteristic identity may be reordered so that

$$\lambda_k = \alpha_k - n + k\,, \qquad 1 \leqslant k \leqslant n\,, \qquad \forall \lambda \in \mathrm{C}^n\,.$$

Proof. — Define

$$\begin{aligned} g_{\mathrm{C}} : \mathrm{C}^n &\to \hat{\mathrm{C}}^n \\ \lambda &\mapsto \hat{\alpha}\,, \qquad \text{where} \qquad \alpha = (\lambda_1 + n - 1, \lambda_2 + n - 2, \ldots, \lambda_n)\,. \end{aligned}$$

It is clear that g_{C} is continuous in the Zariski topology. g_{C} coincides with f_{C} on an open (and therefore dense) subset of C^n, denoted by M. Let $h_{\mathrm{C}} = g_{\mathrm{C}} - f_{\mathrm{C}}$ and suppose $h_{\mathrm{C}} \neq 0$. The set

$$\begin{aligned} \mathrm{N} &= \{ \lambda \in \mathrm{C}^n \mid h_{\mathrm{C}}(\lambda) \neq 0 \} \\ &= h_{\mathrm{C}}^{-1}(\hat{\mathrm{C}}^n \backslash \{ 0 \}) \end{aligned}$$

is open in C^n, because $\{ 0 \}$ is closed in $\hat{\mathrm{C}}^n$ and h_{C} is continuous. Thus, we have written C^n as the union of two disjoint, non-empty, open subsets, namely,

$$\mathrm{C}^n = \mathrm{M} \cup \mathrm{N}\,.$$

Since

$$\mathrm{M} = \mathrm{C}^n \backslash \mathrm{N} \qquad \text{and} \qquad \mathrm{N} = \mathrm{C}^n \backslash \mathrm{M}\,,$$

M and N are also closed. However, C^n cannot be covered by two closed subsets, neither of which is C^n itself [7]. Consequently, N is empty and f_{C} coincides with g_{C} on C^n.

We have therefore established theorem 4 for the case F = C. We shall now show that it remains true for any field F which satisfies $\mathrm{Q} \subseteq \mathrm{F} \subseteq \mathrm{C}$.

LEMMA 9. — For any $\lambda \in \mathrm{F}^n$, there exists up to isomorphism only one irreducible $\mathrm{U_F}$-module with highest weight λ, denoted $\mathrm{V_F}(\lambda)$. The maximal vector v_0 with weight λ is unique. For any $z \in \mathrm{Z}$,

$$zv = \zeta v\,, \qquad \text{for all} \qquad v \in \mathrm{V_F}(\lambda)\,,$$

where $\zeta \in \mathrm{F}$.

Proof. — Let $\mathrm{V_C}(\lambda)$ denote the irreducible $\mathrm{U_C}$-module with highest weight λ. $\mathrm{V_C}(\lambda)$ is spanned by the set

$$\mathrm{B} = \{ v_0, a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \ldots a^{i_p}{}_{j_p} v_0 \mid i_q < j_q\,, \qquad 1 \leqslant q \leqslant p \}\,.$$

Because $\lambda \in \mathrm{F}^n$, all weights of $\mathrm{V_C}(\lambda)$ lie in F^n. Let $\mathrm{V_F}(\lambda)$ denote the vector space over F spaned by B. It is clear that $\mathrm{V_F}(\lambda)$ is a $\mathrm{U_F}$-module and is irreducible. The uniqueness of $\mathrm{V_F}(\lambda)$ follows from the uniqueness of $\mathrm{V_C}(\lambda)$.

The maximal vector in $\mathrm{V_C}(\lambda)$ is unique. Suppose v_0 and v_0' are both maximal vectors in $\mathrm{V_F}(\lambda)$. Extend the field of scalars from F to C. Both v_0 and v_0' remain maximal vectors and so $v_0 = v_0'$.

Suppose $z \in Z$ and $zv_0 \neq 0$. Since

$$a^i_i(zv_0) = \lambda_i(zv_0), \qquad \text{(no summation)},$$

$$a^i_j(zv_0) = z(a^i_j v_0) = 0, \qquad i < j,$$

zv_0 is a maximal vector with highest weight λ. Hence,

$$zv_0 = \zeta v_0, \qquad \zeta \in F,$$

where the case $\zeta = 0$ may now be included. Because $V_F(\lambda)$ is an irreducible U_F-module, the commuting ring of $V_F(\lambda)$ is a division ring, so $(z - \zeta)$ is either invertible or zero. The first possibility is obviously not true, so

$$zv = \zeta v \qquad \text{for all} \qquad v \in V_F(\lambda).$$

The proof of theorem 4 can now be completed. Define a map f_F from F^n to $\hat{C}^n$ by the sequence

$$f_F : \lambda \mapsto V_F(\lambda) \mapsto V_C(\lambda) \mapsto \hat{\alpha}$$

which is to be understood as follows. λ uniquely determines $V_F(\lambda)$. As in lemma 9, $V_F(\lambda)$ is obtained from $V_C(\lambda)$. On $V_C(\lambda)$

$$a_k v = \alpha_k v, \qquad \text{for all} \qquad v \in V_C(\lambda).$$

Finally, $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ uniquely determines $\hat{\alpha}$. Consequently,

$$f_F(\lambda) = f_C(\lambda).$$

Hence, $a_1, a_2, \ldots, a_n$ can be reordered so that

$$\alpha_k = \lambda_k + n - k \in F, \qquad \text{for all} \qquad \lambda \in F^n.$$

It follows from a result of Dixmier [8] that, if V is any irreducible U_C-module, then Z is represented by scalars in V. Because C is algebraically closed, $\bar{Z}$ will also be represented by scalars in V. This property is so useful that it is advantageous to isolate those $\bar{U}_F$-modules in which it continues to hold even when $F \neq C$.

DEFINITION. — Suppose that V is an irreducible $\bar{U}_F$-module and that π is the representation of $\bar{U}$ afforded by V. V will be called a scalar $\bar{U}_F$-module if

$$\pi(\bar{Z}) \subseteq F.1_V$$

(where 1_V is the identity operator on V).

This definition will prove useful in the next section.

6. CHARACTERS AND THE SPECTRAL REPRESENTATION OF A

If V is a scalar $\bar{U}_F$-module, then

$$a_k v = \alpha_k v, \qquad \text{for all} \qquad v \in V,$$

where $\alpha_k \in F$. Theorem 4 shows that, if V has a maximal vector of highest weight λ, then the roots $a_1, a_2, \ldots, a_n$ may be reordered so that

$$\lambda_k = \alpha_k - n + k.$$

This suggests perhaps that the numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$, suitably ordered, might generalize the labelling of irreducible $\bar{U}_F$-modules by highest weights, because $a_1, a_2, \ldots, a_n$ are always defined and, as will be shown below, are distinct elements of $\bar{Z}$. However, lemma 10 below shows that $\alpha_1, \alpha_2, \ldots, \alpha_n$ provide essentially the same information as the character on V.

LEMMA 10. — Every scalar $\bar{U}_F$-module V admits a character χ. If

$$a_k v = \alpha_k v, \qquad \text{for all} \qquad v \in V,$$

then $\hat{\alpha}$ uniquely determines χ and, conversely, is uniquely determined by χ.

Proof. — Because V is a scalar U_F-module, $z.v = \chi(z)v$ for all $v \in V$ where $\chi(z) \in F$, implying that V admits the character $\chi : Z \to F$ given by $z \to \chi(z)$. As t_k is a symmetric polynomial function of a, $\hat{\alpha}$ uniquely determines

$$\chi(t) = (\chi(t_1), \chi(t_2), \ldots, \chi(t_n)),$$

which in turn uniquely determines χ because $Z = Q[t]$. Conversely, the coefficients of the characteristic polynomial of A are elements of Z, so χ uniquely determines $\hat{\alpha}$.

We now proceed to a discussion of some results of Green [5] for which the algebraic framework developed above seems appropriate.

LEMMA 11. — The eigenvalues $a_1, a_2, \ldots, a_n$ are distinct elements of $\bar{Z}$.

Proof. — Suppose the contrary, that $a_i = a_j$ for some $i \neq j$. There exists an irreducible U_C-module $V_C(\lambda)$ whose highest weight λ satisfies

$$\lambda_i - i \neq \lambda_j - j. \tag{27}$$

Theorem 4 asserts that

$$\begin{aligned} 0 &= (a_i - a_j)v \\ &= ((\lambda_i + n - i) - (\lambda_j + n - j))v, \qquad \text{for all} \qquad v \in V_C(\lambda), \end{aligned}$$

which contradicts (27).

This result allows a spectral representation of A to be written down immediately. Define the following $n \times n$ matrices over $\bar{U}$:

$$E_i = \prod_{j \neq i} (A - a_j)/(a_i - a_j).$$

Note that these matrices are well defined because $a_i \neq a_j$ if $i \neq j$.

THEOREM 5. — The matrices $E_1, E_2, \ldots, E_n$ satisfy:

1) $$E_i E_j = \delta_{ij} E_j ;$$

2) $$\sum_{i=1}^{n} E_i = 1 .$$

The matrix A has the spectral representation

3) $$A = \sum_{i=1}^{n} a_i E_i .$$

Proof. — 1) It follows trivially from the identity (24) that

$$AE_j = a_j E_j .$$

If $q(A)$ is a polynomial in A with coëfficients in $\bar{Z}$, it is elementary to establish by induction on the degree of q that

$$q(A)E_j = q(a_j)E_j .$$

In particular,

$$\begin{aligned} E_i E_j &= \prod_{k \neq i} (A - a_k)/(a_i - a_k) E_j \\ &= \prod_{k \neq i} (a_j - a_k)/(a_i - a_k) E_j \\ &= \delta_{ij} E_j . \end{aligned}$$

2) The Lagrange interpolation polynomial of degree ($n = 1$) to the constant polynomial 1 in $\bar{Z}[x]$ is

$$q(x) = \sum_{i=1}^{n} \prod_{j \neq i} (x - a_j)/(a_i - a_j) .$$

The interpolation is exact because the degree of the constant polynomial is zero and therefore certainly less than n. Thus,

$$1 = q(x) .$$

The steps in the proof of the ancient result still hold in $\bar{Z}[A]$, provided 1 is reinterpreted as the unit matrix over $\bar{Z}$. Thus,

$$1 = \sum_{i=1}^{n} E_i .$$

3) The spectral representation of A follows simply from (28) by multiplication by A:

$$\begin{aligned} A &= A \sum_{i=1}^{n} E_i \\ &= \sum_{i=1}^{n} a_i E_i . \end{aligned}$$

THEOREM 6. — A does not satisfy any identity of the form

$$q(A) = 0, \tag{29}$$

in which

$$q(x) \in Z[x] \quad \text{and} \quad \text{degree } q(x) < n. \tag{30}$$

Hence, 2) is the only nth order monic polynomial identity satisfied by A.

Proof. — Suppose that (29) and (30) hold. Since

$$q(A) = \sum_{i=1}^{n} q(a_i)E_i = 0$$

and the spectral projections are mutually annihilating,

$$q(a_i)E_i = 0,$$

and so

$$q(a_i) = 0, \qquad 1 \leqslant i \leqslant n.$$

Thus, $a_1, a_2, \ldots, a_n$ are all roots in $\bar{Z}$ of the polynomial equation

$$q(x) = 0.$$

However, $a_1, a_2, \ldots, a_n$ cannot satisfy any polynomial equation with coefficients in Z of degree less than n, because they are distinct, so this contradicts the initial assumption.

If $s(x)$ is an nth order monic polynomial in $Z[x]$ and

$$s(A) = 0,$$

then

$$(p - s)(A) = 0.$$

Since degree $(p - s)(x) < n$, it follows from the first part of the theorem that $p(x) = s(x)$.

7. APPLICATIONS

In this final section we shall briefly outline some areas in which the results of this paper can be applied.

Firstly, we are interested in the explicit construction, by algebraic methods, of infinite dimensional modules for the classical Lie algebras.

Secondly, Kostant has already demonstrated the importance of an analysis of the structure of L-modules (L reductive or semi-simple) of the form $V(\lambda) \otimes V$, where $V(\lambda)$ is finite dimensional with highest weight λ, and V is infinite dimensional but admits a character. For the case of the Lie algebra of the Lorentz group, recent work by Bracken (private communication) analysing indecomposable modules of the form $V(\lambda) \otimes V$ has relied heavily on characteristic identity techniques.

Thirdly, in the case where V is realised as a function space, the elements of $V(\lambda) \otimes V$ are « wave functions » and we may investigate the existence of L-invariant wave equations. We define a wave operator to be a first order differential operator D on $V(\lambda) \otimes V$ which commutes whith the L-module action, that is,

$$D \in \mathrm{End}_L\,(V(L) \otimes V).$$

It is clear that A_V always belongs to $\mathrm{End}_L\,(V(L) \otimes V)$, and, in certain cases, is in fact a wave operator. For example, when L is the Lie algebra of the de Sitter group and $V(\lambda)$ is the spin representation of L, this fact was exploited by Hannabuss [12] to derive the Dirac equation in de Sitter space. Consequently, we expect A_V and its characteristic identity to be of general interest in the theory of L-invariant wave equations.

Finally, characteristic identities have been applied to state labelling and reduction problems [13].

ACKNOWLEDGMENTS

The authors would like to thank Dr. K. C. Hannabuss for making available an unpublished manuscript which proved invaluable in the early stages of this work. We also thank Professor H. S. Green for many helpful suggestions. One of us (A. L. C.) acknowledges the financial support of a Rothmans Fellowship.

REFERENCES

[1] P. A. M. DIRAC, *Proc. Roy. Soc.*, t. **A 155**, 1936, p. 447-459.
[2] Y. LEHRER-ILAMED, *Bull. Res. Counc. Israël*, t. **5 A**, 1956, p. 197.
[3] A. J. BRACKEN and H. S. GREEN, *J. Math. Phys.*, t. **12**, 1971, p. 2099-2106.
[4] H. S. GREEN, *J. Math. Phys.*, t. **12**, 1971, p. 2106-2113.
[5] H. S. GREEN, to appear in *J. Aust. Math. Soc. B*, 1976.
[6] B. KOSTANT, *J. Funct. Anal.*, t. **20**, 1975, p. 257-285.
[7] J. E. HUMPHREYS, *Introduction to Lie Algebras and Representation Theory*. Springer-Verlag, New York, Heidelberg, Berlin, 1972.
[8] See lemma 2.2 in J. LEPOWSKY, *Trans. Amer. Math. Soc.*, t. **176**, 1973, p. 1-49.
[9] W. GREUB, *Multilinear Algebra*. Springer-Verlag, New York, Heidelberg, Berlin, 1967.
[10] N. JACOBSON, *Lectures in Abstract Algebra, III*. Van Nostrand, Reinhold, New York, 1964.
[11] G. WARNER, *Harmonic Analysis on Semisimple Lie Groups I*. Springer-Verlag, New York, Heidelberg, Berlin, 1972.
[12] K. C. HANNABUSS, *J. Phys. A*, Ser. 2, t. **2**, 1969, p. 274-277.
[13] H. S. GREEN, C. A. HURST and Y. ILAMED, to appear in *J. Math. Phys.*

(Manuscrit reçu le 15 juillet 1976).