

## GENERIC REPRESENTATION THEORY OF THE UNIPOTENT UPPER TRIANGULAR GROUPS

Michael Crumley

Department of Mathematics, The University of Findlay, Findlay, Ohio, USA

*It is generally believed (and for the most part it is probably true) that Lie theory, in contrast to the characteristic zero case, is insufficient to tackle the representation theory of algebraic groups over prime characteristic fields. However, in this article we show that, for a large and important class of unipotent algebraic groups (namely the unipotent upper triangular groups  $U_n$ ), and under a certain hypothesis relating the characteristic  $p$  to both  $n$  and the dimension  $d$  of a representation (specifically,  $p \geq \max(n, 2d)$ ), Lie theory is completely sufficient to determine the representation theories of these groups. To finish, we mention some important analogies (both functorial and cohomological) between the characteristic zero theories of these groups and their “generic” representation theory in characteristic  $p$ .*

**Key Words:** Additive group; Generic representation theory; Heisenberg group; Unipotent algebraic groups.

**2010 Mathematics Subject Classification:** Primary: 20G05; 20G15.

### 1. INTRODUCTION

In this article, we extend a result for representations of the Additive group  $G_a$  given in [7] and the Heisenberg group  $U_3$  given in [2]. Namely, we give an intimate connection between the characteristic zero representation theory of the unipotent upper triangular groups  $U_n$ , and their characteristic  $p > 0$  theory when  $p$  greater than or equal to both  $n$  and twice the dimension of a representation. In particular, such representations are always given rise to by a commuting product of Lie algebra representations, one Lie algebra representation for each of its Frobenius layers, and conversely any such collection of Lie algebra representations induces a representation of  $U_n$  in this fashion.

The essence of what we will be proving in the present article is best illustrated by the obvious analogy between the following two theorems.

**Theorem 1.1** (See Theorem 1.1 of [2]). *Let  $k$  be a field of characteristic zero. Every representation of the Heisenberg group (i.e.,  $U_3$ ) over  $k$  is of the form*

$$e^{xX+yY+(z-xy/2)Z},$$

Received January 26, 2014. Communicated by S. Sehgal.

Address correspondence to Dr. Michael Crumley, Department of Mathematics, The University of Findlay, Findlay, OH 45840, USA; E-mail: mikecrumley@hotmail.com

where  $X, Y$ , and  $Z$  are nilpotent matrices over  $k$  satisfying  $Z = [X, Y]$  and  $[Z, X] = [Z, Y] = 0$ . Further, any such collection  $X, Y, Z$  gives a representation of  $U_3$  over  $k$  according to this formula.

**Theorem 1.2** (See Theorem 1.3 of [2]). *Let  $k$  be a field of characteristic  $p$ , and suppose  $p \geq 2d$ . Then every  $d$ -dimensional representation of the Heisenberg group over  $k$  is of the form*

$$e^{xX_0+yY_0+(z-xy/2)Z_0} e^{x^p X_1+y^p Y_1+(z^p-x^p y^p/2)Z_1} \dots e^{x^{p^m} X_m+y^{p^m} Y_m+(z^{p^m}-x^{p^m} y^{p^m}/2)Z_m},$$

where  $X_0, Y_0, Z_0, X_1, Y_1, Z_1, \dots, X_m, Y_m, Z_m$  is a collection of  $d \times d$  nilpotent matrices over  $k$  satisfying as follows:

- (1)  $[X_i, Y_i] = Z_i$  and  $[Z_i, X_i] = [Z_i, Y_i] = 0$  for every  $i$ ;
- (2) Whenever  $i \neq j$ ,  $X_i, Y_i$ , and  $Z_i$  commute with all of  $X_j, Y_j$ , and  $Z_j$ .

Further, any such collection of  $d \times d$  matrices gives a representation of the Heisenberg group over  $k$  according to the above formula.

The morale is that, for fixed  $d$  and large enough  $p$ , the  $d$ -dimensional representation theory of  $U_3$  over a field of characteristic  $p$  is in perfect analogy with the representation theory of  $U_3^\infty$  over a field of characteristic zero (countable infinite product of copies of  $U_3$ ). Further, this analogy is functorial; the characterization of morphisms between modules over  $U_3$  in characteristic zero carries over to a characterization of morphisms between modules for  $U_3$  over a field of characteristic  $p$ . Finally, this analogy is cohomological, and leads to some interesting generic cohomology results, in a sense to be discussed later.

Let  $k$  be a field, and let  $U_n$  denote the space of all  $n \times n$  upper triangular unipotent matrices over  $k$ , i.e., those of the form

$$\begin{pmatrix} 1 & x_{12} & x_{13} & \dots & x_{1n} \\ & 1 & x_{23} & \dots & x_{2n} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & x_{n-1,n} \\ & & & & 1 \end{pmatrix}.$$

Let  $\mathfrak{u}_n$  denote the Lie algebra of  $U_n$ , which we identify as the space of all  $n \times n$  strictly upper triangular matrices over  $k$ , and likewise  $\mathfrak{gl}_d$  the Lie algebra of  $GL_d$ . In what follows, if  $M(x_1, \dots, x_n)$  is a matrix whose entries are polynomials in the commuting variables  $x_1, \dots, x_n$ , then  $M(x_1, \dots, x_n)^{[m]} \stackrel{\text{def}}{=} M(x_1^m, \dots, x_n^m)$ . For example,

$$\begin{pmatrix} 1 & 2x+3y \\ 0 & 1 \end{pmatrix}^{[3]} = \begin{pmatrix} 1 & 2x^3+3y^3 \\ 0 & 1 \end{pmatrix}.$$

Note that, if  $M(x_1, \dots, x_n)$  is a representation of  $U_n$  in characteristic  $p > 0$  with corresponding comodule structure  $V \xrightarrow{\rho} V \otimes A_n$ , then  $M(x_1, \dots, x_n)^{[p]}$  has comodule structure given by the composition  $V \xrightarrow{\rho} V \otimes A_n \xrightarrow{1 \otimes [p]} V \otimes A_n$ , where  $[p]$

denotes the linear map  $A_n \rightarrow A_n$  which sends each monomial of  $A_n$  to its  $p$ th power (we argue in Lemma 5.3 that this new map does indeed give a representation of  $U_n$ ). Note that  $[p]$  is not quite the same as a Frobenius twist; we are raising the powers of the variables only, not the scalars. We can now state the main theorem of this article.

**Theorem 1.3.** *Let  $k$  be a field of characteristic  $p > 0$ , and suppose that  $p \geq \max(n, 2d)$ .*

- (1) *Let  $\phi_0, \phi_1, \dots, \phi_m : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  be a collection of Lie algebra homomorphisms such that the following statements hold:*

- (a)  $\phi_i(X)$  is a nilpotent matrix for all  $X \in \mathfrak{u}_n$  and  $0 \leq i \leq m$ ;
- (b) For all  $i \neq j$  and  $X, Y \in \mathfrak{u}_n$ ,  $\phi_i(X)$  commutes with  $\phi_j(Y)$ .

*Then the formula*

$$\Phi(g) = e^{\phi_0(\log(g))} e^{\phi_1(\log(g))^{[p]}} \dots e^{\phi_m(\log(g))^{[p^m]}}$$

*defines a valid  $d$ -dimensional representation of  $U_n$ .*

- (2) *Any valid  $d$ -dimensional representation of  $U_n$  over  $k$  is of the form given by part (1).*

Note that this is only a theorem about  $p \gg n$  and dimension; for a fixed prime  $p$  and large enough  $n$  or dimension, the analogy between characteristic zero and characteristic  $p$  completely breaks down. We direct the interested reader to Section 4 of [2] for an example of a 10 dimensional representation of  $U_3$  over the field  $\mathbb{Z}_2$  which is not of the above given form.

As an almost immediate corollary to (the arguments used to prove) our main theorem (1.3), we shall obtain the analogous result for characteristic zero fields in Section 6 (we suspect this is known under at least some assumptions, e.g., when the field is algebraically closed, but we prove it here in Section 6 for convenience).

**Theorem 1.4.** *Let  $k$  be a field of characteristic 0.*

- (1) *Let  $\phi : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  be a Lie algebra homomorphism such that  $\phi(X)$  is a nilpotent matrix for all  $X \in \mathfrak{u}_n$ . Then the formula*

$$\Phi(g) = e^{\phi(\log(g))}$$

*defines a  $d$ -dimensional representation of  $U_n$ .*

- (2) *Any valid  $d$ -dimensional representation of  $U_n$  over  $k$  is of the form given by part (1).*

We shall prove the second part of our main theorem (1.3) first, which shall amount to checking that certain matrices associated to a given representation satisfy certain relations. We shall need some terminology. Denote by  $A_n$  the representing Hopf algebra of  $U_n$ , which we identify as

$$\begin{aligned} A_n &= k[x_{ij} : 1 \leq i < j \leq n] \\ \Delta : x_{ij} &\mapsto 1 \otimes x_{ij} + \sum_{k=i+1}^{j-1} x_{ik} \otimes x_{kj} + x_{ij} \otimes 1 \\ \varepsilon : x_{ij} &\mapsto 0 \end{aligned}$$

We view a representation of  $U_n$  on a  $k$ -vector space  $V$  as a comodule over its representing Hopf algebra  $A_n$  (see Section 3.2 of [11] or Chapter 2 of [3]), i.e., as a  $k$ -linear map  $\rho : V \rightarrow V \otimes A_n$  making commutative the diagrams

$$\begin{array}{ccc} V & \xrightarrow{\rho} & V \otimes A_n \\ \rho \downarrow & & \downarrow 1 \otimes \Delta \\ V \otimes A_n & \xrightarrow[\rho \otimes 1]{} & V \otimes A_n \otimes A_n \end{array} \quad (1.1)$$

$$\begin{array}{ccc} V & \xrightarrow{\rho} & V \otimes A_n \\ & \searrow \simeq & \downarrow 1 \otimes \varepsilon \\ & & V \otimes k \end{array} \quad (1.2)$$

If we fix a basis  $\{e_1, \dots, e_m\}$  for  $V$ , then we can write  $\rho : e_j \mapsto \sum_i e_i \otimes a_{ij}$ , where  $(a_{ij})$  is the “matrix formula” for the representation in this basis. Then the diagrams above are, in equation form,

$$\Delta(a_{ij}) = \sum_k a_{ik} \otimes a_{kj}, \quad (1.3)$$

$$\varepsilon(a_{ij}) = \delta_{ij}. \quad (1.4)$$

For an  $n \times n$  strictly upper triangular matrix  $M = (m_{ij})$  with non-negative integer entries, let  $x^M$  denote the monomial expression

$$x_{12}^{m_{12}} \dots x_{1n}^{m_{1n}} x_{23}^{m_{23}} \dots x_{2n}^{m_{2n}} \dots x_{n-1,n}^{m_{n-1,n}}.$$

Then the matrix  $(a_{ij})$  can be written uniquely as

$$(a_{ij}) = \sum_M \chi(M) x^M,$$

where the summation runs over all  $n \times n$  strictly upper triangular matrices  $M$  with non-negative integer entries, and  $\chi(M)$  is an  $m \times m$  matrix with entries in  $k$ .

**Example.** Consider the representation of  $U_3$  with matrix formula

$$\left( \begin{array}{cccccc} 1 & 2x_{12} & x_{12} & 2x_{12}^2 & x_{13} & 2x_{12}x_{13} \\ & 1 & 0 & x_{12} & 0 & x_{13} \\ & & 1 & 2x_{12} & x_{23} & 2x_{12}x_{23} \\ & & & 1 & 0 & x_{23} \\ & & & & 1 & 2x_{12} \\ & & & & & 1 \end{array} \right).$$

Let  $M$  be the matrix

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then  $x^M$  is the monomial  $x_{12}x_{13}$ , and to compute  $\chi(M)$  we ask, what is the “matrix of coefficients” of this monomial in the above matrix formula? That is,

$$\chi(M) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Denote by  $\varepsilon_{ij}$  the  $n \times n$  matrix with a 1 in the  $(i, j)$ th entry, zeroes elsewhere. Then the above representation also gives

$$\chi(\varepsilon_{12}) = \begin{pmatrix} 0 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \chi(2\varepsilon_{12}) = \begin{pmatrix} 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

and  $\chi(2\varepsilon_{12} + \varepsilon_{13}) = 0$ , since the monomial  $x_{12}^2x_{13}$  never occurs in the representation. We must of course have  $\chi(M) = 0$  for all but finitely many  $M$ , since only finitely many monomials can occur in any given representation.

In what follows,  $[X, Y]$  denotes the usual matrix Lie bracket  $XY - YX$ . The following theorem implies part (2) of our main theorem, and is what we will be proving in Section 3.

**Theorem 1.5.** *Let  $k$  be a field of characteristic  $p > 0$ , and let  $n > 0$ . If  $p \geq 2d$ , then every  $d$ -dimensional representation  $(V, \rho)$  of  $U_n$  over  $k$  satisfies the following conditions, for all  $l, m \geq 0$ , and all  $1 \leq r < s \leq n$ ,  $1 \leq t < u \leq n$ :*

- (1)  $\chi(p^m \varepsilon_{rs})$  is a nilpotent matrix:
- (2)

$$[\chi(p^l \varepsilon_{rs}), \chi(p^m \varepsilon_{tu})] = \begin{cases} 0 & \text{if } l \neq m \text{ or } [\varepsilon_{rs}, \varepsilon_{tu}] = 0 \\ \chi(p^m [\varepsilon_{rs}, \varepsilon_{tu}]) & \text{otherwise} \end{cases}$$

**Proposition 1.6.** *Theorem 1.5 implies part (2) of Theorem 1.3.*

*Proof.* Let

$$(a_{ij}) = \sum_M \chi(M) x^M$$

be a  $d$ -dimensional representation of  $U_n$  over the field  $k$  of characteristic  $p$  with  $p \geq 2d$ . For each  $l$  and  $1 \leq r < s \leq n$ , set

$$\phi_l(\varepsilon_{rs}) = \chi(p^l \varepsilon_{rs})$$

and extend each  $\phi_l$  linearly to all of  $\mathfrak{u}_n$ . We need to show, under the conclusions of Theorem 1.5, that the conclusions of part (2) of Theorem 1.3 are met:

- (1)  $\phi_l(X)$  is a nilpotent matrix for all  $l \geq 0$  and  $X \in \mathfrak{u}_n$ ;
- (2) Each  $\phi_l$  is a Lie algebra homomorphism;
- (3) For all  $l \neq m$  and  $X, Y \in \mathfrak{u}_n$ ,  $[\phi_l(X), \phi_m(Y)] = 0$ ;
- (4) The representation  $(a_{ij})$  is actually given by the formula

$$\Phi(g) = e^{\phi_0(\log(g))} e^{\phi_1(\log(g))^{[p]}} \dots e^{\phi_m(\log(g))^{[p^m]}}.$$

(1): Since the homomorphic image of any unipotent algebraic group is unipotent, and since any unipotent subgroup of  $GL_d$  is, up to conjugacy, a subgroup of  $U_d$  (see Theorem 2.4.11 of [10]), if we write

$$(a_{ij}) = 1 + \sum_{M \neq 0} \chi(M) x^M,$$

we see that  $\chi(M)$ , for every  $M \neq 0$ , is a nilpotent matrix, whence so is  $\phi_l(X)$  for any  $l$  and  $X \in \mathfrak{u}_n$ .

(2): We check that  $\phi_l$  is a Lie algebra homomorphism on the basis  $\{\varepsilon_{rs} : 1 \leq r < s \leq n\}$  for  $\mathfrak{u}_n$ . If  $[\varepsilon_{rs}, \varepsilon_{tu}] = 0$ , by conclusion (2) of Theorem 1.5 we have

$$\begin{aligned} [\phi_l(\varepsilon_{rs}), \phi_l(\varepsilon_{tu})] &= [\chi(p^l \varepsilon_{rs}), \chi(p^l \varepsilon_{tu})] \\ &= 0 \\ &= \phi_l([\varepsilon_{rs}, \varepsilon_{tu}]) \end{aligned}$$

since  $\phi_l$  is linear by construction. Otherwise, we have

$$\begin{aligned} [\phi_l(\varepsilon_{rs}), \phi_l(\varepsilon_{tu})] &= [\chi(p^l \varepsilon_{rs}), \chi(p^l \varepsilon_{tu})] \\ &= \chi(p^m [\varepsilon_{rs}, \varepsilon_{tu}]) \\ &= \phi_l([\varepsilon_{rs}, \varepsilon_{tu}]). \end{aligned}$$

(3): For  $m \neq l$ , Theorem 1.5 gives

$$[\phi_l(\varepsilon_{rs}), \phi_m(\varepsilon_{tu})] = [\chi(p^l \varepsilon_{rs}), \chi(p^m \varepsilon_{tu})] = 0.$$

(4): Finally, assuming that the formula

$$\Phi(g) = e^{\phi_0(\log(g))} e^{\phi_1(\log(g))^{[p]}} \dots e^{\phi_m(\log(g))^{[p^m]}}$$

actually is a representation of  $U_n$  (which will be proven in Theorem 5.4), we want to see that it is the same representation as  $(a_{ij}) = \sum_M \chi(M)x^M$ . Lemmas 3.1 and 3.4 given in Section 3 make it clear that it suffices to check that, for each  $r, s$ , and  $l$ , the matrix of coefficients of the monomial  $x_{rs}^{p^l}$  in the first formula is actually  $\chi(p^l \varepsilon_{rs})$ , the same as in the second formula.

We prove this first for the case of  $m = 0$ . Let  $g = 1 + \sum_{1 \leq i < j \leq n} x_{ij} \varepsilon_{ij}$  be an arbitrary element of  $U_n$ . Then

$$\begin{aligned} \log(g) &\stackrel{\text{def}}{=} \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k} (g - 1)^k \\ &= \sum_{1 \leq i < j \leq n} x_{ij} \varepsilon_{ij} + \dots \end{aligned}$$

where  $(\dots)$  denotes a sum of matrices whose monomial coefficients have length greater than 1. As  $\phi_0$  is linear, we have

$$\phi_0(\log(g)) = \sum_{1 \leq i < j \leq n} x_{ij} \phi_0(\varepsilon_{ij}) + \phi_0(\dots).$$

Then

$$\begin{aligned} \Phi(g) &= e^{\phi_0(\log(g))} \\ &= 1 + \left( \sum_{1 \leq i < j \leq n} x_{ij} \phi_0(\varepsilon_{ij}) + \phi_0(\dots) \right) + \dots \end{aligned}$$

where the last  $(\dots)$  denotes terms including higher powers of  $\phi_0(\log(g))$ . Clearly then we have that the matrix of coefficients of the terms  $x_{ij}$  are exactly  $\phi_0(\varepsilon_{ij}) = \chi(\varepsilon_{ij})$ .

For  $m > 0$ , consider

$$\begin{aligned} \Phi(g) &= e^{\phi_0(\log(g))} e^{\phi_1(\log(g))^{[p]}} \dots e^{\phi_l(\log(g))^{[p^l]}} \dots e^{\phi_m(\log(g))^{[p^m]}} \\ &= \sum_{k_0=0}^{d-1} \dots \sum_{k_l=0}^{d-1} \dots \sum_{k_m=0}^{d-1} \left( \frac{\phi_0(\log(g))^{k_0}}{k_0!} \right) \dots \left( \frac{\phi_l(\log(g))^{k_l}}{k_l!} \right)^{[p^l]} \dots \left( \frac{\phi_m(\log(g))^{k_m}}{k_m!} \right)^{[p^m]}, \end{aligned}$$

where we are justified in running the summations only to  $d - 1$  since each  $\phi_i(\log(g))$  is nilpotent of order no greater than  $d = \dim$ . And since  $d - 1 < d \leq p/2$ , we see that the only contribution to the monomial  $x_{rs}^{p^l}$  comes when all  $k_i$  are zero except  $k_l = 1$ , and thus the matrix of coefficients of the monomial  $x_{rs}^{p^l}$  is exactly  $\phi_l(\varepsilon_{rs})$ . Recalling that  $\phi_l(\varepsilon_{rs}) \stackrel{\text{def}}{=} \chi(p^l \varepsilon_{rs})$ , this proves that  $\Phi(g) = (a_{ij})$ .  $\square$

Our approach to proving Theorem 1.5 is partially inductive. Take for example the group  $U_4$ :

$$\begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ & 1 & x_{23} & x_{24} \\ & & 1 & x_{34} \\ & & & 1 \end{pmatrix}.$$

This group contains four conspicuous subgroups, namely those matrices of the form

$$\begin{pmatrix} 1 & x_{12} & x_{13} & 0 \\ & 1 & x_{23} & 0 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ & 1 & x_{23} & x_{24} \\ & & 1 & x_{34} \\ & & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & x_{12} & x_{13} & x_{14} \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix},$$

and       $\begin{pmatrix} 1 & 0 & 0 & x_{14} \\ & 1 & 0 & x_{24} \\ & & 1 & x_{34} \\ & & & 1 \end{pmatrix},$

which are isomorphic to, respectively,  $U_3$ ,  $U_3$ ,  $G_a^3$ , and  $G_a^3$  (three-fold direct product of the Additive group). Since all of these various embeddings are given by simply setting certain variables to zero, by the main result given in [2] for the Heisenberg group, and by the result given in [7] or [1] for products of the Additive group, we immediately conclude by induction that, for instance,

$$[\chi(p^m \varepsilon_{23}), \chi(p^m \varepsilon_{34})] = \chi(p^m \varepsilon_{24})$$

since the variables  $x_{23}$  and  $x_{34}$  already occur together in a subgroup isomorphic to  $U_3$ . Generalizing this to arbitrary  $n$ , by induction, the only brackets  $[\chi(p^m \varepsilon_{rs}), \chi(p^n \varepsilon_{tu})]$  that need to be checked are those that do *not* occur together in one of the above given subgroups:

- (1) When  $r = 1$  and  $u = n$  (i.e., when the first variable is in the top row, and the second is in the right-most column);
- (2) When  $\varepsilon_{rs} = \varepsilon_{1n}$ , and when  $t \neq 1$  and  $u \neq n$  (i.e., when the first variable is in the top right corner, and the second variable is in neither the top row nor the right-most column).

Restricting ourselves to these cases shall somewhat simplify our calculations.

## 2. COMBINATORICS

Our arguments for proving Theorem 1.5 are heavily combinatorial, so as we go along we shall provide examples, generally using the group  $U_4$ .

The learned reader may recognize that what we are really doing in this section is working out certain facts about the multiplicative structure of  $\text{Dist}(U_n)$ , the distribution algebra of  $U_n$  (see Chapter 7 of [9]), but we have no need for this terminology, and no understanding of distribution algebras is assumed on the part of the reader.

Let  $n$  be arbitrary, and throughout, fix a  $d$ -dimensional representation  $(a_{ij})$  for  $U_n$  over the field  $k$ , and write

$$(a_{ij}) = \sum_M \chi(M) x^M,$$

where the summation runs over all  $n \times n$  strictly upper triangular matrices  $M = (m_{ij})$  with non-negative integer entries, and  $\chi(M)$  is a  $d \times d$  matrix with entries in  $k$  for each  $M$ , and where we define

$$x^M \stackrel{\text{defn}}{=} \prod_{1 \leq i < j \leq n} x_{ij}^{m_{ij}}.$$

Define the following  $n \times n$  variable matrices  $S_1, \dots, S_n$ . The nonzero entries of  $S_k$  are written  $s_{ij}^k$ .  $S_1$  and  $S_2$  are demanded to be strictly upper triangular,  $S_3$  is strictly-strictly upper triangular,  $S_4$  is strictly-strictly-strictly upper triangular, and  $\dots, S_n$  is  $(n-1) \times$  strictly upper triangular (i.e.,  $S_n$  has a nonzero entry in its  $(1, n)$ th spot only). In other words, for each  $1 \leq i < j \leq n$  we have the variables  $s_{ij}^1$  and for each  $2 \leq k \leq n$ , and each  $1 \leq i < j \leq n$  with  $j - i \geq k - 1$ , we have the variables  $s_{ij}^k$ .

**Example.** In the case of  $n = 4$ , the matrices  $S_1, S_2, S_3, S_4$  are given by

$$S_1 = \begin{pmatrix} 0 & s_{12}^1 & s_{13}^1 & s_{14}^1 \\ & 0 & s_{23}^1 & s_{24}^1 \\ & & 0 & s_{34}^1 \\ & & & 0 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 0 & s_{12}^2 & s_{13}^2 & s_{14}^2 \\ & 0 & s_{23}^2 & s_{24}^2 \\ & & 0 & s_{34}^2 \\ & & & 0 \end{pmatrix},$$

$$S_3 = \begin{pmatrix} 0 & 0 & s_{13}^3 & s_{14}^3 \\ & 0 & 0 & s_{24}^3 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}, \quad S_4 = \begin{pmatrix} 0 & 0 & 0 & s_{14}^4 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}.$$

For  $1 \leq i < j \leq n$ , define the following variable expressions among the  $s_{ij}^k$ :

$$L_{ij} \stackrel{\text{defn}}{=} \sum_{k=j}^n s_{ik}^{j-i+1} \quad R_{ij} \stackrel{\text{defn}}{=} \sum_{k=1}^i s_{kj}^{i-k+1}$$

For example, in the case of  $n = 4$ ,  $L_{12} = s_{12}^2 + s_{13}^2 + s_{14}^2$ , and  $R_{34} = s_{14}^3 + s_{24}^3 + s_{34}^1$ .

The following notation will be useful: If  $B_1 = (b_{ij}^1), \dots, B_k = (b_{ij}^k), M = (m_{ij})$  are matrices with non-negative integer entries such that  $B_1 + \dots + B_k = M$ , then the formal multinomial expression

$$\binom{M}{B_1, \dots, B_k}$$

is shorthand for  $\prod_{ij} \binom{m_{ij}}{b_{ij}^1, \dots, b_{ij}^k}$ . In other words, it is just the product of the multinomial coefficients of the individual entries. For example,

$$\left( \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ choose } \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right)$$

equals

$$\binom{2}{1, 1, 0} \binom{1}{0, 1, 0} \binom{3}{2, 0, 1} = 6.$$

**Proposition 2.1.** *In the notation above, the matrix  $(\Delta(a_{ij}))$  is equal to*

$$\sum_M \chi(M) \left[ \sum_{S_1 + \dots + S_n = M} \binom{S_1 + \dots + S_n}{S_1, \dots, S_n} \left( \prod_{1 \leq i < j \leq n} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq n} x_{ij}^{R_{ij}} \right) \right],$$

where the first summation runs over all  $n \times n$  strictly upper triangular matrices  $M$  with non-negative integer entries, and the second runs over all  $S_1 + S_2 + \dots + S_n = M$  with non-negative integer entries of the form defined in the above paragraph.

Before we prove the proposition, let us illustrate with an example what it is actually saying, again using the case of  $n = 4$ . Using  $(a_{ij}) = \sum_M \chi(M)x^M$ , we have

$$\begin{aligned} \Delta(a_{ij}) &= \sum_M \chi(M) \Delta(x^M) \\ &= \sum_M \chi(M) \Delta(x_{12})^{m_{12}} \Delta(x_{23})^{m_{23}} \Delta(x_{34})^{m_{34}} \Delta(x_{13})^{m_{13}} \Delta(x_{24})^{m_{24}} \Delta(x_{14})^{m_{14}} \\ &= \sum_M \chi(M) (1 \otimes x_{12} + x_{12} \otimes 1)^{m_{12}} (1 \otimes x_{23} + x_{23} \otimes 1)^{m_{23}} (1 \otimes x_{34} + x_{34} \otimes 1)^{m_{34}} \\ &\quad (1 \otimes x_{13} + x_{13} \otimes x_{23} + x_{12} \otimes x_{23} + x_{13} \otimes 1)^{m_{13}} (1 \otimes x_{24} + x_{24} \otimes x_{34} + x_{23} \otimes x_{34} + x_{24} \otimes 1)^{m_{24}} \\ &\quad (1 \otimes x_{14} + x_{14} \otimes x_{24} + x_{12} \otimes x_{24} + x_{13} \otimes x_{34} + x_{14} \otimes 1)^{m_{14}} \end{aligned}$$

and by repeated application of the binomial/multinomial theorem

$$\begin{aligned} &= \sum_M \chi(M) \\ &\quad \left( \sum_{s_{12}^1 + s_{12}^2 = m_{12}} \binom{m_{12}}{s_{12}^1, s_{12}^2} (1 \otimes x_{12})^{s_{12}^1} (x_{12} \otimes 1)^{s_{12}^2} \right) \\ &\quad \left( \sum_{s_{23}^1 + s_{23}^2 = m_{23}} \binom{m_{23}}{s_{23}^1, s_{23}^2} (1 \otimes x_{23})^{s_{23}^1} (x_{23} \otimes 1)^{s_{23}^2} \right) \\ &\quad \left( \sum_{s_{34}^1 + s_{34}^2 = m_{34}} \binom{m_{34}}{s_{34}^1, s_{34}^2} (1 \otimes x_{34})^{s_{34}^1} (x_{34} \otimes 1)^{s_{34}^2} \right) \\ &\quad \left( \sum_{s_{13}^1 + s_{13}^2 + s_{13}^3 = m_{13}} \binom{m_{13}}{s_{13}^1, s_{13}^2, s_{13}^3} (1 \otimes x_{13})^{s_{13}^1} (x_{12} \otimes x_{23})^{s_{13}^2} (x_{13} \otimes 1)^{s_{13}^3} \right) \end{aligned}$$

$$\left( \sum_{s_{24}^1 + s_{24}^2 + s_{24}^3 = m_{24}} \binom{m_{24}}{s_{24}^1, s_{24}^2, s_{24}^3} (1 \otimes x_{24})^{s_{24}^1} (x_{23} \otimes x_{34})^{s_{24}^2} (x_{24} \otimes 1)^{s_{24}^3} \right)$$

$$\left( \sum_{s_{14}^1 + s_{14}^2 + s_{14}^3 + s_{14}^4 = m_{14}} \binom{m_{14}}{s_{14}^1, s_{14}^2, s_{14}^3, s_{14}^4} (1 \otimes x_{14})^{s_{14}^1} (x_{12} \otimes x_{24})^{s_{14}^2} (x_{13} \otimes x_{34})^{s_{14}^3} (x_{14} \otimes 1)^{s_{14}^4} \right)$$

and, condensing

$$= \sum_M \chi(M)$$

$$\left( \sum_{s_{12}^1 + s_{12}^2 = m_{12}} \binom{m_{12}}{s_{12}^1, s_{12}^2} x_{12}^{s_{12}^2} \otimes x_{12}^{s_{12}^1} \right)$$

$$\left( \sum_{s_{23}^1 + s_{23}^2 = m_{23}} \binom{m_{23}}{s_{23}^1, s_{23}^2} x_{23}^{s_{23}^2} \otimes x_{23}^{s_{23}^1} \right)$$

$$\left( \sum_{s_{34}^1 + s_{34}^2 = m_{24}} \binom{m_{34}}{s_{34}^1, s_{34}^2} x_{34}^{s_{34}^2} \otimes x_{34}^{s_{34}^1} \right)$$

$$\left( \sum_{s_{13}^1 + s_{13}^2 + s_{13}^3 = m_{13}} \binom{m_{13}}{s_{13}^1, s_{13}^2, s_{13}^3} x_{12}^{s_{13}^2} x_{13}^{s_{13}^3} \otimes x_{13}^{s_{13}^1} x_{23}^{s_{13}^2} \right)$$

$$\left( \sum_{s_{24}^1 + s_{24}^2 + s_{24}^3 = m_{24}} \binom{m_{24}}{s_{24}^1, s_{24}^2, s_{24}^3} x_{23}^{s_{24}^2} x_{24}^{s_{24}^3} \otimes x_{24}^{s_{24}^1} x_{34}^{s_{24}^2} \right)$$

$$\left( \sum_{s_{14}^1 + s_{14}^2 + s_{14}^3 + s_{14}^4 = m_{14}} \binom{m_{14}}{s_{14}^1, s_{14}^2, s_{14}^3, s_{14}^4} x_{12}^{s_{14}^2} x_{13}^{s_{14}^3} x_{14}^{s_{14}^4} \otimes x_{14}^{s_{14}^1} x_{24}^{s_{14}^2} x_{34}^{s_{14}^3} \right).$$

With the variable matrices  $S_1, S_2, S_3, S_4$  so previously defined, we can collapse all of these summations to

$$\sum_M \chi(M) \left[ \sum_{S_1 + S_2 + S_3 + S_4 = M} \binom{M}{S_1, S_2, S_3, S_4}$$

$$x_{12}^{s_{12}^2 + s_{13}^2 + s_{14}^2} x_{23}^{s_{23}^2 + s_{24}^2} x_{34}^{s_{34}^2} x_{13}^{s_{13}^3 + s_{14}^3} x_{24}^{s_{24}^3} x_{14}^{s_{14}^4}$$

$$\otimes$$

$$x_{12}^{s_{12}^1} x_{23}^{s_{23}^1 + s_{13}^1} x_{34}^{s_{34}^1 + s_{24}^1 + s_{14}^1} x_{13}^{s_{13}^1} x_{24}^{s_{24}^1 + s_{14}^1} x_{14}^{s_{14}^1} \right],$$

which the reader can verify by hand that can be written

$$\sum_M \chi(M) \left[ \sum_{S_1+S_2+S_3+S_4=M} \binom{S_1+S_2+S_3+S_4}{S_1, S_2, S_3, S_4} \left( \prod_{1 \leq i < j \leq 4} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq 4} x_{ij}^{R_{ij}} \right) \right].$$

We now prove Proposition 2.1 for arbitrary  $n$ .

*Proof.* Compute

$$\begin{aligned} (\Delta(a_{ij})) &= \sum_M \chi(M) \Delta(x^M) \\ &= \sum_M \chi(M) \prod_{1 \leq i < j \leq n} \Delta(x_{ij})^{m_{ij}} \\ &= \sum_M \chi(M) \prod_{1 \leq i < j \leq n} (1 \otimes x_{ij} + \sum_{k=i+1}^{j-1} x_{ik} \otimes x_{kj} + x_{ij} \otimes 1)^{m_{ij}} \\ &= \sum_M \chi(M) \prod_{1 \leq i < j \leq n} \left( \sum_{\substack{s_{ij}^1 + \dots + s_{ij}^{j-i+1} = m_{ij} \\ s_{ij}^1, \dots, s_{ij}^{j-i+1}}} \binom{s_{ij}^1 + \dots + s_{ij}^{j-i+1}}{s_{ij}^1, \dots, s_{ij}^{j-i+1}} \right. \\ &\quad \left. (1 \otimes x_{ij})^{s_{ij}^1} \prod_{k=i+1}^{j-1} (x_{ik} \otimes x_{kj})^{s_{ij}^{k-i+1}} (x_{ij} \otimes 1)^{s_{ij}^{j-i+1}} \right) \\ &= \sum_M \chi(M) \prod_{1 \leq i < j \leq n} \left( \sum_{\substack{s_{ij}^1 + \dots + s_{ij}^{j-i+1} = m_{ij} \\ s_{ij}^1, \dots, s_{ij}^{j-i+1}}} \binom{s_{ij}^1 + \dots + s_{ij}^{j-i+1}}{s_{ij}^1, \dots, s_{ij}^{j-i+1}} \right. \\ &\quad \left. x_{i,i+1}^{s_{ij}^2} x_{i,i+2}^{s_{ij}^3} \dots x_{ij}^{s_{ij}^{j-i+1}} \otimes x_{ij}^{s_{ij}^1} x_{i+1,j}^{s_{ij}^2} \dots x_{j-1,j}^{s_{ij}^{j-i}} \right) \\ &= \sum_M \chi(M) \sum_{\substack{s_{12}^1 + s_{12}^2 = m_{12} \\ s_{23}^1 + s_{23}^2 = m_{23}}} \prod_{1 \leq i < j \leq n} \left( \sum_{\substack{s_{ij}^1 + \dots + s_{ij}^{j-i+1} \\ s_{ij}^1, \dots, s_{ij}^{j-i+1}}} \binom{s_{ij}^1 + \dots + s_{ij}^{j-i+1}}{s_{ij}^1, \dots, s_{ij}^{j-i+1}} \right) x_{i,i+1}^{s_{ij}^2} x_{i,i+2}^{s_{ij}^3} \dots x_{ij}^{s_{ij}^{j-i+1}} \\ &\quad \vdots \\ &\quad \substack{s_{n-1,n}^1 + s_{n-1,n}^2 = m_{n-1,n} \\ s_{13}^1 + s_{13}^2 + s_{13}^3 = m_{13} \\ s_{24}^1 + s_{24}^2 + s_{24}^3 = m_{24}} \\ &\quad \vdots \\ &\quad \substack{s_{n-2,n}^1 + s_{n-2,n}^2 + s_{n-2,n}^3 = m_{n-2,n} \\ s_{1,n}^1 + s_{1,n}^2 + \dots + s_{1,n}^n = m_{1,n}} \\ &\quad \otimes x_{ij}^{s_{ij}^1} x_{i+1,j}^{s_{ij}^2} \dots x_{j-1,j}^{s_{ij}^{j-i}} \\ &= \sum_M \chi(M) \sum_{S_1+\dots+S_n=M} \binom{M}{S_1, \dots, S_n} \prod_{1 \leq i < j \leq n} x_{i,i+1}^{s_{ij}^2} x_{i,i+2}^{s_{ij}^3} \dots x_{ij}^{s_{ij}^{j-i+1}} \\ &\quad \otimes x_{ij}^{s_{ij}^1} x_{i+1,j}^{s_{ij}^2} \dots x_{j-1,j}^{s_{ij}^{j-i}}. \end{aligned}$$

Our task now is to write the expression  $\prod_{1 \leq i < j \leq n} x_{i,i+1}^{s_{ij}^2} x_{i,i+2}^{s_{ij}^3} \cdots x_{ij}^{s_{ij}^{j-i+1}} \otimes x_{ij}^{s_{ij}^1} x_{i+1,j}^{s_{ij}^2} \cdots x_{j-1,j}^{s_{ij}^{j-i}}$  in the form

$$\left( \prod_{1 \leq i < j \leq n} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq n} x_{ij}^{R_{ij}} \right)$$

and to figure out what  $L_{ij}$  and  $R_{ij}$  are. To compute  $L_{ij}$  we ask, in which factors, for  $1 \leq u < k \leq n$ , does the term  $x_{ij}$  occur in the left tensor slot, and what is its exponent? For this to happen, it is necessary and sufficient that  $u = i$  and  $j \leq k$ , and in this case the exponent of  $x_{ij}$  is  $s_{i,k}^{j-i+1}$ . Summing this expression over all  $k$  greater than or equal to  $j$  gives

$$L_{ij} = \sum_{k=j}^n s_{ik}^{j-i+1}$$

as claimed. To compute  $R_{ij}$  we ask, for which  $1 \leq k < u \leq n$  does  $x_{ij}$  occur in the right tensor slot, and what is its exponent? For this to happen, it is necessary and sufficient that  $u = j$ , and that  $k \leq i$ , and in this case the exponent of  $x_{ij}$  is  $s_{k,j}^{i-k+1}$ . Summing this expression over all  $k$  less than or equal to  $i$  gives

$$R_{ij} = \sum_{k=1}^i s_{k,j}^{i-k+1},$$

as claimed. This completes the proof.  $\square$

**Proposition 2.2.** *The matrix  $(\sum_k a_{ik} \otimes a_{kj})$  is equal to*

$$\sum_{M,N} \chi(M)\chi(N)x^M \otimes x^N,$$

where the summation runs over all  $n \times n$  strictly upper triangular matrices with non-negative integer entries, and  $\chi(M)\chi(N)$  is the usual product of matrices.

*Proof.* We have that  $a_{ik} = \sum_M \chi(M)_{ik} x^M$ , and  $a_{kj} = \sum_N \chi(N)_{kj} x^N$ . Compute

$$\begin{aligned} \sum_k a_{ik} \otimes a_{kj} &= \sum_k \left( \sum_M \chi(M)_{ik} x^M \right) \otimes \left( \sum_N \chi(N)_{kj} x^N \right) \\ &= \sum_{M,N} \left( \sum_k \chi(M)_{ik} \chi(N)_{kj} \right) x^M \otimes x^N. \end{aligned}$$

Realizing that  $\sum_k \chi(M)_{ik} \chi(N)_{kj}$  is nothing more than the  $(i, j)$ th entry of the matrix product  $\chi(M)\chi(N)$ , this completes the proof.  $\square$

Recalling Eq. 1.3 and 1.4, by Propositions 2.1 and 2.2, we have the following proposition.

**Proposition 2.3.** *The matrix formula  $(a_{ij}) = \sum_M \chi(M)x^M$  is a representation of  $U_n$  if and only if the following statements hold:*

- (1)  $\chi(M) = 0$  for all but finitely many  $M$ ;
- (2)  $\chi(0) = Id$ ;
- (3)

$$\begin{aligned} & \sum_{M,N} \chi(M)\chi(N)x^M \otimes x^N \\ &= \sum_M \chi(M) \left[ \sum_{S_1+\dots+S_n=M} \binom{S_1 + \dots + S_n}{S_1, \dots, S_n} \left( \prod_{1 \leq i < j \leq n} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq n} x_{ij}^{R_{ij}} \right) \right]. \end{aligned} \quad (2.1)$$

In particular, the matrix product  $\chi(M)\chi(N)$  must be exactly the coefficient of the monomial tensor  $x^M \otimes x^N$  on the right-hand side of Eq. 2.1.

The left-hand side of Eq. 2.1 is a sum over distinct monomial tensors, and so the coefficient of  $x^M \otimes x^N$  is exactly  $\chi(M)\chi(N)$ . The right-hand side of this equation, on the other hand, is *not* a sum over distinct monomial tensors, and for the most part it is a nontrivial exercise figuring out exactly what the coefficient of  $x^M \otimes x^N$  is for arbitrary  $M$  and  $N$ . Nonetheless, for the  $M$  and  $N$  we are interested in, we shall be able to compute exactly what the coefficient of the monomial tensor  $x^M \otimes x^N$  is on the right-hand side, and in so doing compute  $\chi(M)\chi(N)$ , and in so doing prove Theorem 1.5.

As an example of how we will be using this equation, consider again the case  $n = 4$ , and consider the matrix product  $\chi(\varepsilon_{14})\chi(\varepsilon_{13})$ . We claim that it is always equal to  $\chi(\varepsilon_{14} + \varepsilon_{13})$ . To show this, consider the equation

$$\begin{aligned} & x_{12}^{s_{12}^2+s_{13}^2+s_{14}^2} x_{23}^{s_{23}^2+s_{24}^2} x_{34}^{s_{34}^2} x_{13}^{s_{13}^3+s_{14}^3} x_{24}^{s_{24}^3} x_{14}^{s_{14}^4} \\ & \otimes \\ & x_{12}^{s_{12}^1} x_{23}^{s_{23}^1+s_{13}^2} x_{34}^{s_{34}^1+s_{24}^2+s_{14}^3} x_{13}^{s_{13}^1} x_{24}^{s_{24}^1+s_{14}^2} x_{14}^{s_{14}^1} \\ & = x_{14} \otimes x_{13}. \end{aligned}$$

Since the variables  $s_{ij}^k$  are demanded to be non-negative integers, clearly there is only one collection of matrices  $S_1, S_2, S_3, S_4$  in non-negative integer entries that gives a solution to this equation, namely  $s_{14}^4 = 1, s_{13}^1 = 1$ , and all other  $s_{ij}^k = 0$ . Thus the coefficient of  $x_{14} \otimes x_{13}$  on the right-hand side of Eq. 2.1 is exactly

$$\begin{aligned} \binom{S_1 + S_2 + S_3 + S_4}{S_1, S_2, S_3, S_4} \chi(S_1 + S_2 + S_3 + S_4) &= \binom{\varepsilon_{13} + 0 + 0 + \varepsilon_{14}}{\varepsilon_{13}, 0, 0, \varepsilon_{14}} \chi(\varepsilon_{13} + \varepsilon_{14}) \\ &= \chi(\varepsilon_{13} + \varepsilon_{14}). \end{aligned}$$

The coefficient of  $x_{14} \otimes x_{13}$  on the left-hand side of Eq. 2.1 is of course  $\chi(\varepsilon_{14})\chi(\varepsilon_{13})$ , whence they are equal.

**Lemma 2.4.** Consider the variable expressions

$$L_{ij} = \sum_{k=j}^n s_{ik}^{j-i+1} \quad R_{ij} = \sum_{k=1}^i s_{kj}^{i-k+1}.$$

- (1) Each of the variables  $s_{ij}^k$  occur at most once in any of the  $L_{ij}$ , and the only variables that do not occur in any of the  $L_{ij}$  are those of the form  $s_{ij}^1$  (i.e., those variables occurring in the matrix  $S_1$ )
- (2) Each of the variables  $s_{ij}^k$  occur at most once in any of the  $R_{ij}$ , and the only variables that do not occur in any of the  $R_{ij}$  are those of the form  $s_{i,k}^{k-i+1}$ , for  $k > i$  (i.e., those on the super-diagonal of  $S_2$ , on the super-super-diagonal of  $S_3$ , ..., on the  $(1, n)^{th}$  entry of  $S_n$ )
- (3) The variables that occur in both of the expressions  $L_{ij}$  and  $R_{uv}$  are exactly as follows:
  - (a)  $s_{i,v}^{j-i+1}$  if  $j = u$ ;
  - (b) None otherwise.

*Proof.* (1) For the variable  $s_{uv}^w$  to occur in  $L_{ij}$ , we must have a  $k$  with  $j \leq k \leq n$  such that  $s_{ik}^{j-i+1} = s_{uv}^w$ . This forces  $i = u$ ,  $k = v$ , and  $w = j - i + 1$ , i.e.,  $j = w + u - 1$ . Thus there is only one possible choice for  $i$ ,  $j$ , and  $k$ .

Now suppose  $w = 1$ . Then there is no  $L_{ij}$  in which  $s_{uv}^w$  occurs, since this forces  $j = u = i$ , and of course  $i$  is always demanded to be less than  $j$ . On the other hand, if  $2 \leq w \leq n$ , set  $i = u$ ,  $k = v$ , and  $j = w + u - 1$ ; we claim that  $s_{uv}^w$  occurs as the  $k$ th term of  $L_{ij}$ . To prove this, we need to verify that  $1 \leq i < j \leq n$ , and that  $j \leq k \leq n$ . That  $1 \leq i$  is obvious from  $1 \leq u$ . Obviously, also  $k \leq n$  since  $v \leq n$ . Recall that, by assumption of the form of the variable  $s_{uv}^w$ ,  $v - u \leq w - 1$ , i.e.,  $u \leq v - w + 1$ . Then

$$j = w + u - 1 \leq w + (v - w + 1) - 1 = v \leq n$$

and

$$j \leq v = k$$

as required.

(2) For the variable  $s_{uv}^w$  to occur in  $R_{ij}$ , we must have a  $k$  with  $1 \leq k \leq i$  such that  $s_{kj}^{i-k+1} = s_{uv}^w$ . This forces  $j = v$ ,  $k = u$ , and  $w = i - k + 1$ , i.e.,  $i = w + u - 1$ . Thus there is only one possible choice for  $i$ ,  $j$ , and  $k$ .

Now suppose that  $s_{uv}^w$  is of the form  $s_{i,k}^{k-i+1}$ , for  $k > i$ , that is, so that  $v - u = w - 1$ . Then this forces

$$j - k = v - u = w - 1 = i - k + 1 - 1 = i - k$$

so that  $j = i$ , which is impossible. On the other hand, if  $s_{uv}^w$  is not of this form, we can assume that  $v - u > w - 1$ . Then set  $i = w + u - 1$ ,  $j = v$ ,  $k = u$ . We claim that  $s_{uv}^w$  occurs as the  $k$ th term of  $R_{ij}$ . We must verify that  $1 \leq i < j \leq n$ , and that  $1 \leq k \leq i$ .

That  $j \leq n$  is obvious from  $v \leq n$ ,  $1 \leq k$  since  $1 \leq u$ , and  $1 \leq i$  since  $1 \leq w + u - 1$ , i.e.,  $2 \leq w + u$ .  $i < j$  is equivalent to  $w + u - 1 < v$ , which is equivalent to  $v - u > w - 1$ , and  $k \leq i$  is equivalent to  $u \leq w + u - 1$ , which is equivalent to  $1 \leq w$ .

(3) To see which variables occur in both  $L_{ij}$  and  $R_{uv}$ , we seek  $k_0$  and  $k_1$  with  $j \leq k_0 \leq n$  and  $1 \leq k_1 \leq u$  such that

$$s_{i,k_0}^{j-i+1} = s_{k_1,v}^{u-k_1+1},$$

which forces  $i = k_1$ ,  $v = k_0$ , and  $j - i + 1 = u - k_1 + 1 = u - i + 1$ , i.e.,  $j = u$ . Thus we must have  $j = u$  to have any variables in common. If this is the case, then  $1 \leq i < j = u < v \leq n$ , whence we have  $1 \leq k_1 < j = u < k_0 \leq n$ , as required, giving us the variable  $s_{iv}^{j-i+1}$ .  $\square$

We close this section with one more combinatorial lemma.

**Lemma 2.5.** *Let  $Y = (y_{ij})$  be an  $n \times n$  strictly upper triangular matrix with zeroes on its top row, and let  $Z = (z_{ij})$  be an  $n \times n$  strictly upper triangular matrix with zeroes in all rows except perhaps the top, both with non-negative integer entries. Then the system*

$$\left( \prod_{1 \leq i < j \leq n} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq n} x_{ij}^{R_{ij}} \right) = x^Y \otimes x^Z$$

has exactly one solution in the variables  $s_{ij}^k$ :

- (1)  $s_{1j}^1 = z_{1j}$  for all  $2 \leq j \leq n$ ;
- (2)  $s_{ij}^{j-i+1} = y_{ij}$  for all  $2 \leq i < j \leq n$ ;
- (3) All other  $s_{ij}^k = 0$ .

**Remark 2.6.** In the case of  $n = 5$ , a useful heuristic to remember this solution is

$$\begin{pmatrix} 0 & s_{12}^1 & s_{13}^1 & s_{14}^1 & s_{15}^1 \\ & 0 & s_{23}^2 & s_{24}^3 & s_{25}^4 \\ & & 0 & s_{34}^2 & s_{35}^3 \\ & & & 0 & s_{45}^2 \\ & & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & z_{12} & z_{13} & z_{14} & z_{15} \\ & 0 & y_{23} & y_{24} & y_{25} \\ & & 0 & y_{34} & y_{35} \\ & & & 0 & y_{45} \\ & & & & 0 \end{pmatrix} = Y + Z$$

with any  $s_{ij}^k$  not mentioned assumed to be zero.

*Proof.* The key fact is that the variables mentioned in part (1) of the above do not occur in any of the  $L_{ij}$ , and those given in part (2) do not occur in any of the  $R_{ij}$  (see Lemma 2.4).

We prove first that the assignments given actually are a solution. We need to verify that

$$R_{1,j} = z_{1,j} \quad \text{for all } 2 \leq j \leq n$$

and

$$L_{i,j} = y_{i,j} \quad \text{for all } 2 \leq i < j \leq n$$

with all other  $L_{ij}$ ,  $R_{ij} = 0$ . Consider

$$R_{ij} = \sum_{k=1}^i s_{kj}^{i-k+1} = s_{1,j}^i + s_{2,j}^{i-1} + \cdots + s_{ij}^1.$$

We see firstly that  $R_{1,j} = s_{1,j}^1 = z_{1,j}$ . By Lemma 2.4, no variable of the form given in part (2) of the proposition occur in any  $R_{ij}$ , and if  $i > 1$ , no variable of the form given in part (1) may occur either. This gives  $R_{ij} = 0$  for  $i > 1$ , as required.

Consider now

$$L_{ij} = \sum_{k=j}^n s_{ik}^{j-i+1} = s_{ij}^{j-i+1} + s_{i,j+1}^{j-i+1} + \cdots + s_{in}^{j-i+1}.$$

Note firstly that no variables of the form given part (1) occur in any  $L_{ij}$  (Lemma 2.4). If  $i = 1$ , then we have

$$L_{1j} = s_{1,j}^j + s_{1,j+1}^j + \cdots + s_{1,n}^j.$$

In order for a variable of the form given in (2) to occur in this expression, we must have  $j = 1 + j + 1 - 1$ , which is absurd; thus  $L_{1j} = 0$ . In case  $i > 1$ , we see there is exactly one term of the form given in (2) occurring, namely the first term,  $s_{i,j}^{j-i+1}$ , which is equal to  $y_{ij}$  by fiat. This gives  $L_{ij} = y_{ij}$  for  $i > 1$ , as required.

We now argue that this is the only possible solution. First, the equations  $R_{1,j} = s_{1,j}^1 = z_{1,j}$  force the assignments given in part (1). Second, if  $i \geq 2$ , then the variable  $s_{ij}^{j-i+1}$  occurs in  $L_{ij}$ , and this is the *only* variable of this form occurring in  $L_{ij}$ . Third, if  $i \geq 2$ , then every other variable occurring in  $L_{ij}$ , not being of the form given in (2), occurs in some  $R_{uv}$  (Lemma 2.4), necessarily with  $i = u > 1$ ; and since  $R_{uv} = 0$  for all  $u > 1$ , this makes them zero. Thus, for  $i \geq 2$ ,  $L_{ij} = y_{ij}$  has only one possible non-zero variable, namely  $s_{ij}^{j-i+1}$ ; this forces  $s_{ij}^{j-i+1} = y_{ij}$  for all  $i \geq 2$ .

Finally, let  $s_{ij}^k$  be any variable not of the form given in either (1) or (2). Certainly, it occurs somewhere among the  $L_{ij}$  or  $R_{ij}$ . If this happens to be any of the nonzero  $L$ 's or  $R$ 's, we have already argued that  $s_{ij}^k$  must be zero; and if it is *not* among these  $L$ 's or  $R$ 's, then that particular  $L$  or  $R$  must itself be zero, forcing  $s_{ij}^k = 0$ . This completes the proof.  $\square$

### 3. THE MAIN THEOREM: NECESSITY

We can now put Eq. 2.1 to work in proving Theorem 1.5. Recall that  $\varepsilon_{ij}$  is the  $n \times n$  matrix with a 1 in its  $(i, j)$ th entry, zeroes elsewhere. Let  $M$  be an arbitrary  $n \times n$  strictly upper triangular matrix with non-negative integer entries. To begin, we need a factorization for  $\chi(M)$  in terms of those matrices of the form  $\chi(r\varepsilon_{ij})$ .

**Lemma 3.1.** *If  $M = (m_{ij})$ , then*

$$\chi(M) = \prod_{i=n-1}^1 \prod_{j=i+1}^n \chi(m_{ij}\varepsilon_{ij}).$$

**Remark 3.2.** The notation “ $\prod_{i=n-1}^1$ ” is not a typo; we are listing the factors in reverse order for convenience. For example, in the case of  $n = 4$ , we are saying that

$$\chi(M) = \chi(m_{34}\varepsilon_{34})\chi(m_{23}\varepsilon_{23})\chi(m_{24}\varepsilon_{24})\chi(m_{12}\varepsilon_{12})\chi(m_{13}\varepsilon_{13})\chi(m_{14}\varepsilon_{14}).$$

*Proof.* We proceed by induction on  $n$ . If  $n = 2$  the above equation is  $\chi(M) = \chi(m_{12}\varepsilon_{12}) = \chi(M)$ , which is obvious. Let  $Y$  and  $Z$  be the  $n \times n$  matrices

$$Y = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & m_{23} & m_{24} & \dots & m_{2n} \\ 0 & m_{34} & \dots & m_{3n} \\ \ddots & \ddots & \vdots \\ 0 & m_{n-1,n} \\ 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & m_{12} & m_{13} & m_{14} & \dots & m_{1n} \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \ddots & \ddots & \ddots & \vdots \\ 0 & 0 \\ 0 & 0 \end{pmatrix},$$

i.e., the  $n \times n$  matrices with the top row of  $M$  deleted, and *all but* the top row deleted. Using the second embedding of  $U_{n-1}$  into  $U_n$  described at the end of section 1, we conclude by induction that

$$\begin{aligned} \chi(Y) &= \chi(m_{n-1,n}\varepsilon_{n-1,n})\chi(m_{n-2,n-1}\varepsilon_{n-2,n-1})\chi(m_{n-2,n}\varepsilon_{n-2,n}) \dots \chi(m_{23}\varepsilon_{23}) \dots \chi(m_{2n}\varepsilon_{2n}) \\ &= \prod_{i=n-1}^2 \prod_{j=i+1}^n \chi(m_{ij}\varepsilon_{ij}) \end{aligned}$$

And by using the first embedding  $G_a^{n-1}$  into  $U_n$  described at the end of section 1, we have

$$\begin{aligned} \chi(Z) &= \chi(m_{12}\varepsilon_{12}) \dots \chi(m_{1n}\varepsilon_{1n}) \\ &= \prod_{j=2}^n \chi(m_{1j}\varepsilon_{1j}). \end{aligned}$$

Now consider the matrix product  $\chi(Y)\chi(Z)$ ; we claim it is equal to  $\chi(Y + Z) = \chi(M)$ . Using Eq. 2.1, we seek solutions to the system

$$\left( \prod_{1 \leq i < j \leq n} x_{ij}^{L_{ij}} \right) \otimes \left( \prod_{1 \leq i < j \leq n} x_{ij}^{R_{ij}} \right) = x^Y \otimes x^Z.$$

By Lemma 2.5, this system has exactly one solution, and we leave it to the reader that this solution gives

$$\binom{S_1 + \dots + S_n}{S_1, \dots, S_n} = 1$$

and

$$S_1 + \dots + S_n = M.$$

This gives

$$\begin{aligned}\chi(M) &= \chi(Y)\chi(Z) \\ &= \left( \prod_{i=n-1}^2 \prod_{j=i+1}^n \chi(m_{ij}\varepsilon_{ij}) \right) \left( \prod_{j=2}^n \chi(m_{1j}\varepsilon_{1j}) \right) \\ &= \prod_{i=n-1}^1 \prod_{j=i+1}^n \chi(m_{ij}\varepsilon_{ij}).\end{aligned}$$

This completes the proof.  $\square$

For any  $1 \leq i < j \leq n$ , those matrices of the form  $1 + x_{ij}\varepsilon_{ij}$  form a subgroup of  $U_n$  isomorphic to the Additive group  $G_a$ , and we shall use this fact repeatedly. We shall need the following concerning the representation theory of the Additive group  $G_a$  over prime characteristic fields. In what follows we shall identify  $G_a$  as  $U_2$ , i.e., the collection of all matrices of the form

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}.$$

The following proposition is originally due to A. Suslin, E. M. Friedlander, and C. P. Bendel [7] and can be found in the proof of Proposition 1.2 of [7]. However, our notation differs markedly from theirs, and so the reader may wish to see Proposition 2.6 of [2] instead.

**Proposition 3.3.** *Let  $k$  be a field of characteristic  $p > 0$ , and let  $(a_{ij}) = \sum_{n \in \mathbb{N}} \chi(n\varepsilon_{12})x^n$  be any representation of  $G_a$  over  $k$ . Set  $X_0 = \chi(\varepsilon_{12})$ ,  $X_1 = \chi(p\varepsilon_{12}), \dots, X_m = \chi(p^m\varepsilon_{12})$ , with  $X_i = 0$  for  $i > m$ .*

(1)  $(a_{ij})$  can be written as

$$e^{xX_0} e^{x^p X_1} \dots e^{x^{p^m} X_m}.$$

The  $X_i$  all commute and are nilpotent of order no greater than  $p$ . Further, any finite collection of such matrices defines a representation according to this formula.

(2) Let  $r \in \mathbb{N}$ , and write  $r = r_0 + r_1 p + \dots + r_m p^m$  in  $p$ -ary notation. Then

$$\chi(r\varepsilon_{12}) = \Gamma(r)^{-1} \chi(\varepsilon_{12})^{r_0} \chi(p\varepsilon_{12})^{r_1} \dots \chi(p^m\varepsilon_{12})^{r_m},$$

where  $\Gamma(r) \stackrel{\text{def}}{=} r_0! r_1! \dots r_m!$ .

**Lemma 3.4.** *Let  $k$  be a field of characteristic  $p > 0$ ,  $(a_{ij}) = \sum_M \chi(M)x^M$  a representation of  $U_n$  over  $k$ . For  $r \in \mathbb{N}$ , write  $r = r_0 + r_1 p + \dots + r_m p^m$  in  $p$ -ary notation. Then for any  $1 \leq i < j \leq n$  and  $r \in \mathbb{N}$ ,*

$$\chi(r\varepsilon_{ij}) = \Gamma(r)^{-1} \chi(\varepsilon_{ij})^{r_0} \chi(p\varepsilon_{ij})^{r_1} \dots \chi(p^m\varepsilon_{ij})^{r_m},$$

where  $\Gamma(r) \stackrel{\text{def}}{=} r_0!r_1!\dots r_m!$ . Further, all of the factors of this product commute, and all of the  $\chi(p^m \varepsilon_{ij})$  are nilpotent of order no greater than  $p$ .

*Proof.* For fixed  $i$  and  $j$ , the collection of those matrices in  $U_n$  of the form  $1 + x_{ij}\varepsilon_{ij}$  form a subgroup of  $U_n$  isomorphic to the additive group  $G_a$ . This embedding is given by the Hopf algebra map  $A_n \rightarrow k[x]$  which sends  $x_{ij}$  to  $x$  and all other  $x_{rs}$  to zero. Apply Proposition 3.3.  $\square$

The next lemma, while simple, is the crucial fact which allows us to prove the main theorem of this article. It also illustrates why we do *not* suspect a result analogous to our main theorem to hold for non-unipotent algebraic groups (i.e., groups with a Lie algebra containing non-nilpotent matrices).

**Definition 3.5.** Let  $a$  and  $b$  be non-negative integers,  $p$  a prime. Write

$$\begin{aligned} r &= r_m p^m + r_{m-1} p^{m-1} + \dots + r_1 p + r_0 \\ s &= s_m p^m + s_{m-1} p^{m-1} + \dots + s_1 p + s_0 \end{aligned}$$

in  $p$ -ary notation (so that for all  $i$ ,  $0 \leq r_i, s_i < p$ ). We say that the sum  $r + s$  **carries** modulo  $p$  if for some  $i$ ,  $r_i + s_i \geq p$ .

**Lemma 3.6.** Let  $k$  be a field of characteristic  $p > 0$ ,  $(a_{ij}) = \sum_M \chi(M)x^M$  a  $d$ -dimensional representation of  $U_n$  over  $k$ , and suppose that  $p \geq 2d$ . Let  $r, s \in \mathbb{N}$ , and suppose that the sum  $r + s$  carries modulo  $p$ . Then for any  $i, j, u, v$ , at least one of  $\chi(r\varepsilon_{ij})$  or  $\chi(s\varepsilon_{uv})$  must be zero.

*Proof.* Write  $r = r_0 + r_1 p + \dots + r_m p^m$  in  $p$ -ary notation, similarly for  $s$ , and since  $r + s$  carries, let  $r_t + s_t \geq p$ . Then at least one of  $r_t$  or  $s_t$  is greater than or equal to  $p/2$ , say  $r_t$ . By Lemma 3.4, write

$$\chi(r\varepsilon_{ij}) = \Gamma(r)^{-1} \chi(\varepsilon_{ij})^{r_0} \chi(p\varepsilon_{ij})^{r_1} \dots \chi(p^t\varepsilon_{ij})^{r_t} \dots \chi(p^m\varepsilon_{ij})^{r_m}$$

Then  $\chi(r\varepsilon_{ij})$  is zero, since  $\chi(p^t\varepsilon_{ij})^{r_t}$  is zero, since  $\chi(p^t\varepsilon_{ij})$  is nilpotent of order no greater than  $d \leq p/2 \leq r_t$ .  $\square$

We are now ready to prove Theorem 1.5. As stated in the introduction, the four embeddings mentioned there give us a great deal for free; all that needs to be checked is the bracket  $[\chi(p^m\varepsilon_{ij}), \chi(p^n\varepsilon_{tu})]$  when  $(i, j)$  is in the top row and  $(t, u)$  in the right-most column, or when  $(i, j) = (1, n)$  and  $(t, u)$  is in neither the top row nor the right-most column. We break it down into cases.

Recall that the Lie algebra of  $U_n$ , which we identify as  $\mathfrak{u}_n = \text{span}_k(\varepsilon_{ij} : 1 \leq i < j \leq n)$ , has the following Lie bracket:

$$[\varepsilon_{rs}, \varepsilon_{tu}] = \begin{cases} \varepsilon_{ru} & \text{if } s = t \\ -\varepsilon_{ts} & \text{if } r = u \\ 0 & \text{otherwise.} \end{cases}$$

For the rest of this section, we assume that  $k$  is a field of characteristic  $p$ , that  $p \geq 2d$ , and that the  $\chi(M)$  correspond to a  $d$ -dimensional representation of  $U_n$  over  $k$ .

**Proposition 3.7.** *For any  $l, m, t$ , and  $u$ ,*

$$\chi(p^l \varepsilon_{tn}) \chi(p^m \varepsilon_{1u}) = \binom{p^l \varepsilon_{tn} + p^m \varepsilon_{1u}}{p^l \varepsilon_{tn}, p^m \varepsilon_{1u}} \chi(p^l \varepsilon_{tn} + p^m \varepsilon_{1u}).$$

*Proof.* We seek solutions to  $L_{tn} = s_{tn}^{n-t+1} = p^l$ ,  $R_{1u} = s_{1u}^1 = p^m$ , and all other  $L_{ij}, R_{ij} = 0$ . Notice first that the variable  $s_{tn}^{n-t+1}$  does not occur in any of the  $R_{ij}$ , and that  $s_{1u}^1$  does not occur in any of the  $L_{ij}$ , and that neither occur in any more  $L$ 's or  $R$ 's (Lemma 2.4), so there is no conflict in setting  $L_{tn} = p^l$ ,  $R_{1u} = p^m$ ; that is, we have at least one solution. But also by Lemma 2.4, every other variable occurs in at least one of the other  $L_{ij}$  or  $R_{ij}$ , forcing them to be zero, giving us exactly one solution, namely,  $s_{tn}^{n-t+1} = p^l$ ,  $s_{1u}^1 = p^m$ , and all other  $s_{ij}^k = 0$ . This gives us that the coefficient of the monomial tensor  $x_{tn}^{p^l} \otimes x_{1u}^{p^m}$  is exactly

$$\binom{p^l \varepsilon_{tn} + p^m \varepsilon_{1u}}{p^l \varepsilon_{tn}, p^m \varepsilon_{1u}} \chi(p^l \varepsilon_{tn} + p^m \varepsilon_{1u}),$$

and so (by Eq. 2.1), the proposition is proved.  $\square$

**Proposition 3.8.** *If  $t \neq u$ , then*

$$\chi(p^l \varepsilon_{1u}) \chi(p^m \varepsilon_{tn}) = \binom{p^l \varepsilon_{1u} + p^m \varepsilon_{tn}}{p^l \varepsilon_{1u}, p^m \varepsilon_{tn}} \chi(p^l \varepsilon_{1u} + p^m \varepsilon_{tn}).$$

*Proof.* We again examine

$$L_{1u} = \sum_{k=u}^n s_{1,k}^u \quad \text{and} \quad R_{tn} = \sum_{k=1}^t s_{k,n}^{t-k+1}.$$

We seek solutions to  $L_{1u} = p^l$ ,  $R_{tn} = p^m$ , and all other  $L_{ij}, R_{ij} = 0$ . We claim there is exactly one solution, namely,  $s_{1u}^u = p^l$ ,  $s_{tn}^1 = p^m$ , and all other  $s_{ij}^k = 0$ . First,  $L_{1u}$  and  $R_{tn}$  share no variables in common. Second, the only variable occurring in  $L_{1u}$  which *does not* also occur in some  $R_{ij}$  is  $s_{1u}^u$ , and the only variable occurring in  $R_{tn}$  which *does not* occur in some  $L_{ij}$  is  $s_{tn}^1$  (Lemma 2.4). This forces all of the variables occurring in  $L_{1u}$  or  $R_{tn}$  to be zero, except for these two, forcing  $s_{1u}^u = p^l$ ,  $s_{tn}^1 = p^m$ . This gives that the coefficient of the monomial tensor  $x_{1u}^{p^l} \otimes x_{tn}^{p^m}$  in Eq. 2.1 is exactly

$$\binom{p^l \varepsilon_{1u} + p^m \varepsilon_{tn}}{p^l \varepsilon_{1u}, p^m \varepsilon_{tn}} \chi(p^l \varepsilon_{1u} + p^m \varepsilon_{tn}),$$

which proves the proposition.  $\square$

**Corollary 3.9.** *If  $t \neq u$ , then for any  $l$  and  $m$ ,*

$$[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{tn})] = 0.$$

*Proof.* By Propositions 3.7 and 3.8,  $\chi(p^l \varepsilon_{1u}) \chi(p^m \varepsilon_{tn}) = \chi(p^m \varepsilon_{tn}) \chi(p^l \varepsilon_{1u})$ .  $\square$

**Proposition 3.10.** *For any  $1 \leq t < u \leq n$  and any  $l$  and  $m$ ,*

$$[\chi(p^l \varepsilon_{1n}), \chi(p^m \varepsilon_{tu})] = 0$$

*Proof.* Note that  $L_{1n} = s_{1n}^n$  and that  $R_{1n} = s_{1n}^1$ , both consisting of a single variable. We leave it to the reader then to compute the products  $\chi(p^l \varepsilon_{1n})\chi(p^m \varepsilon_{tu})$  and  $\chi(p^m \varepsilon_{tu})\chi(p^l \varepsilon_{1n})$ , to see that each have exactly one solution, and these solutions give the same answer for both.  $\square$

Thus far, we have not used at all the assumption that  $p \geq 2d$ . This will change as we consider the last case, namely,  $[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{un})]$ .

**Proposition 3.11.** *For any  $l, m$ , and  $1 < u \leq n$ ,*

$$[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{un})] = \sum_{k=1}^{\min(p^l, p^m)} \chi((p^m - k)\varepsilon_{un})\chi((p^l - k)\varepsilon_{1u})\chi(k\varepsilon_{1n}).$$

*Proof.* We shall compute the product  $\chi(p^l \varepsilon_{1u})\chi(p^m \varepsilon_{un})$ . Thus we seek solutions to

$$L_{1u} = p^l \text{ and } R_{un} = p^m,$$

with all other  $L_{ij}$ ,  $R_{ij}$  equal to zero. First,  $L_{1u}$  and  $R_{un}$  have exactly one variable in common, namely,  $s_{1u}^u$  (Lemma 2.4). Secondly, there is exactly one variable occurring in  $L_{1u}$  which does not occur in any  $R_{ij}$ , namely,  $s_{1u}^u$ , and exactly one variable occurring in  $R_{un}$  which does not occur in any of the  $L_{ij}$ , namely,  $s_{un}^1$  (Lemma 2.4 again). This gives

$$s_{1u}^u + s_{1n}^u = p^l \text{ and } s_{un}^1 + s_{1n}^u = p^m,$$

with all other  $s_{ij}^k$  equal to zero. Clearly then, every non-negative integer value of  $s_{1n}^u$  no greater than either  $p^l$  or  $p^m$  gives a solution, and these are the only solutions. If  $k = s_{1n}^u$  is any such value, its contribution to the coefficient of  $x_{1u}^{p^l} \otimes x_{un}^{p^m}$  in Eq. 2.1 is

$$\binom{k\varepsilon_{1n} + (p^l - k)\varepsilon_{1u} + (p^m - k)\varepsilon_{un}}{k\varepsilon_{1n}, (p^l - k)\varepsilon_{1u}, (p^m - k)\varepsilon_{un}} \chi(k\varepsilon_{1n} + (p^l - k)\varepsilon_{1u} + (p^m - k)\varepsilon_{un}).$$

Note that it is impossible for any of the tuples  $(1, n)$ ,  $(n, u)$ , or  $(u, n)$ , to be equal, so the above multinomial coefficient is exactly 1. Thus we can write

$$\chi(p^l \varepsilon_{1u})\chi(p^m \varepsilon_{un}) = \sum_{k=0}^{\min(p^l, p^m)} \chi(k\varepsilon_{1n} + (p^l - k)\varepsilon_{1u} + (p^m - k)\varepsilon_{un})$$

and by Lemma 3.1 we can further write

$$\chi(p^l \varepsilon_{1u})\chi(p^m \varepsilon_{un}) = \sum_{k=0}^{\min(p^l, p^m)} \chi((p^m - k)\varepsilon_{un})\chi((p^l - k)\varepsilon_{1u})\chi(k\varepsilon_{1n}).$$

Finally, note that the value  $k = 0$  gives  $\chi(p^m \varepsilon_{un})\chi(p^l \varepsilon_{1u})$ , so we can write

$$\chi(p^l \varepsilon_{1u})\chi(p^m \varepsilon_{un}) - \chi(p^m \varepsilon_{un})\chi(p^l \varepsilon_{1u}) = \sum_{k=1}^{\min(p^l, p^m)} \chi((p^m - k)\varepsilon_{un})\chi((p^l - k)\varepsilon_{1u})\chi(k\varepsilon_{1n}),$$

which proves the proposition.  $\square$

**Theorem 3.12.** *If  $p \geq 2d$ , then*

$$[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{un})] = \begin{cases} 0 & \text{if } l \neq m \\ \chi(p^m \varepsilon_{1n}) & \text{otherwise} \end{cases}.$$

*Proof.* By the previous proposition, examine

$$[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{un})] = \sum_{k=1}^{\min(p^l, p^m)} \chi((p^m - k)\varepsilon_{un})\chi((p^l - k)\varepsilon_{1u})\chi(k\varepsilon_{14}).$$

First suppose that  $l \neq m$ , say that  $l < m$ . Then for every value  $k$  of this summation, there is clearly some ‘‘carrying’’ in computing the sum  $(p^m - k) + k$ . Then by Lemma 3.6, at least one of  $\chi((p^m - k)\varepsilon_{un})$  or  $\chi(k\varepsilon_{1n})$  is always zero, forcing every term in the summation to be zero. This gives  $[\chi(p^l \varepsilon_{1u}), \chi(p^m \varepsilon_{un})] = 0$ ; an analogous proof holds in case  $l > m$ .

Now suppose that  $l = m$ . Then for the same reason every term in the above summation is zero, *except* for the last term  $k = p^m$ , since  $(p^m - p^m) + p^m$  does not carry. This gives

$$[\chi(p^m \varepsilon_{1u}), \chi(p^m \varepsilon_{un})] = \chi(0\varepsilon_{un})\chi(0\varepsilon_{1u})\chi(p^m \varepsilon_{1n}).$$

Recalling that  $\chi(0) = \text{Id}$ , this completes the proof.  $\square$

Theorem 1.5 is now proved, and hence also part (2) of our main theorem, Theorem 1.3.

#### 4. THE BAKER–CAMPBELL–HAUSDORFF FORMULA

The proof of part (1) of our main theorem (Theorem 1.3) will require some understanding of the Baker–Campbell–Hausdorff formula. For the reader’s convenience, we briefly review here the calculation of said formula as computed by E. B. Dynkin for the case of characteristic zero in [5], [6], and [4]. The principal result of course is that, for not-necessarily-commutative variables  $x$  and  $y$  in an associative algebra over a characteristic zero field, the series  $\log(e^x e^y)$  can be expressed as a formal infinite series of brackets of  $x$  and  $y$ , brackets of brackets of  $x$  and  $y$ , etc., all with rational coefficients.

The details are important to us; we will claim later that these arguments apply just as well to the characteristic  $p > 0$  setting, under some additional assumptions ( $p$ -nilpotency of a Lie algebra principal among them) that disallow the appearance of any rational numbers with denominators divisible by  $p$ .

In what follows we shall follow Dynkin in using the less cumbersome notation  $x \circ y = xy - yx$  for the commutator operator instead of  $[x, y]$ . As  $\circ$  is by no means associative, when we write  $x_1 \circ x_2 \circ \dots \circ x_n$ , we shall take it to be *left-nested*; e.g.,  $x_1 \circ x_2 \circ x_3 \circ x_4 \stackrel{\text{def}}{=} ((x_1 \circ x_2) \circ x_3) \circ x_4$ .

**Lemma 4.1.** *Define the formal infinite series*

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad \log(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k$$

*Then  $\log(e^x e^y)$  can be written*

$$\sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1!q_1! \dots p_k!q_k!} x^{p_1} y^{q_1} \dots x^{p_k} y^{q_k},$$

*where the summation runs over all tuples of non-negative integers  $(p_1, q_1, \dots, p_k, q_k)$  with the property that  $p_i + q_i \neq 0$  (the length  $2k$  of the tuples vary arbitrarily).*

*Proof.* The trick is in collecting the terms correctly. Write

$$e^x e^y = \sum_{r=0}^{\infty} \sum_{s=0}^r \frac{x^s y^{r-s}}{s!(r-s)!},$$

and hence

$$\log(e^x e^y) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \left( \sum_{r=1}^{\infty} \sum_{s=0}^r \frac{x^s y^{r-s}}{s!(r-s)!} \right)^k.$$

An induction argument shows that, for fixed  $k$ ,

$$\left( \sum_{r=1}^{\infty} \sum_{s=0}^r \frac{x^s y^{r-s}}{s!(r-s)!} \right)^k = \sum_{\substack{(p_1, q_1, \dots, p_k, q_k) \\ p_i + q_i \neq 0}} \frac{1}{p_1!q_1! \dots p_k!q_k!} x^{p_1} y^{q_1} \dots x^{p_k} y^{q_k}.$$

Whence, summing over all  $k$ , we get

$$\sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1!q_1! \dots p_k!q_k!} x^{p_1} y^{q_1} \dots x^{p_k} y^{q_k}.$$

□

From this expression, we write

$$\log(e^x e^y) = \sum_{m=1}^{\infty} P_m(x, y), \tag{4.1}$$

where  $P_m(x, y)$  is a homogeneous polynomial of degree  $m$ ; for example,  $P_1(x, y) = x + y$ ,  $P_2(x, y) = \frac{1}{2}(xy - yx)$ , and  $P_3(x, y) = \frac{1}{12}x^2y - \frac{1}{6}xyx + \frac{1}{12}xy^2 + \frac{1}{12}y^2x + \frac{1}{12}yx^2 - \frac{1}{6}yxy$ .

Let  $k$  be a field of characteristic zero, and let  $R$  be the free associative (and non-commutative) algebra over  $k$  on the generators  $x$  and  $y$ . Define a linear mapping  $\phi$  from  $R$  to itself as follows: it sends the monomial  $x_1 x_2 \dots x_n$  to

$$\frac{1}{n} x_1 \circ x_2 \circ \dots \circ x_n$$

(here the  $x_i$  can be either of  $x$  or  $y$ ). For example,  $\phi(x^2 y + x + 2xyx) = \frac{1}{3}x \circ x \circ y + x + \frac{2}{3}x \circ y \circ x$ .

**Proposition 4.2.** *Any bracket expression can be written as a linear combination of left-nested bracket expressions, all of length no greater than the original, and with no new coefficients up to perhaps a negation.*

*Proof.* Clearly, it suffices to prove that, if  $P$  and  $Q$  themselves are left-nested bracket expressions, then  $P \circ Q$  can be written as a linear combination of left-nested bracket expressions, of lengths no greater than that of  $P$  plus that of  $Q$ . Let  $\Omega(n)$  be the statement “if  $P$  is a left-nested bracket of length  $m$  ( $m$  arbitrary), and if  $Q$  is a left-nested bracket of length  $\leq n$ , then  $P \circ Q$  can be written as a linear combination of left-nested brackets of length no greater than  $n + m$ .” Certainly,  $\Omega(1)$  is true. Now suppose that  $\Omega(n)$  is true, and let  $P$  be any left-nested bracket,  $Q \circ x$  a left-nested bracket expression of length  $n + 1$ , so necessarily  $Q$  is left-nested of length  $n$ . Then (by Jacobi and anticommutativity)

$$\begin{aligned} P \circ (Q \circ x) &= -Q \circ (x \circ P) - x \circ (P \circ Q) \\ &= (x \circ P) \circ Q + (P \circ Q) \circ x \\ &= -(P \circ x) \circ Q + (P \circ Q) \circ x. \end{aligned}$$

$P \circ x$  is left-nested, and since  $\Omega(n)$  is true,  $-(P \circ x) \circ Q$  is a linear combination of left-nested brackets of lengths no greater than the original. The same can be said of  $P \circ Q$ , whence (by linearity of  $\circ$ )  $(P \circ Q) \circ x$  is a linear combination of left-nested brackets, whence  $P \circ (Q \circ x)$  is also a linear combination of left-nested brackets, of lengths no greater than the original.  $\square$

**Proposition 4.3.** *If  $x_1 \circ x_2 \circ \dots \circ x_n$  is any left-nested bracket expression, then  $\phi(x_1 \circ x_2 \circ \dots \circ x_n) = x_1 \circ x_2 \circ \dots \circ x_n$ .*

*Proof.* See [6, pp. 31–32]  $\square$

The proposition predicts, for example, that  $\phi(x \circ y \circ z) = \phi(xyz - yxz - zxy + zyx) = \frac{1}{3}x \circ y \circ z - \frac{1}{3}y \circ x \circ z - \frac{1}{3}z \circ x \circ y + \frac{1}{3}z \circ y \circ x = x \circ y \circ z$ .

**Proposition 4.4.** *The homogeneous polynomials  $P_i(x, y)$  in Eq. 4.1 can all be written as linear combinations (with only rational coefficients) of nested commutators of  $x$  and  $y$ .*

*Proof.* See [4].  $\square$

This last proposition, along with Propositions 4.2 and 4.3 gives the following Proposition.

**Proposition 4.5.** *If  $P_i(x, y)$  is any of the homogeneous polynomials in Eq. 4.1, then  $\phi(P_i(x, y)) = P_i(x, y)$ .*

This proposition gives us not only the assurance that each  $P_i(x, y)$  can be written as a rational linear combination of bracket expressions in  $x$  and  $y$ , but gives us an explicit method for doing so. For example,  $P_2(x, y) = \frac{1}{2}xy - \frac{1}{2}yx$ , and if we were so dull as to not realize it, we merely apply  $\phi$  to yield

$$P_2(x, y) = \phi(P_2(x, y)) = \frac{1}{4}x \circ y - \frac{1}{4}y \circ x = \frac{1}{4}x \circ y + \frac{1}{4}x \circ y = \frac{1}{2}x \circ y.$$

With a view towards proving part (1) of Theorem 1.3, we make the following simple observations concerning all of this.

**Proposition 4.6.** *Let  $p$  be a prime.*

- (1) *If  $m < p$ , then  $P_m(x, y)$  contains no coefficients whose denominators are divisible by  $p$ .*
- (2) *If  $m < p$ , then  $\phi(P_m(x, y))$  also contains no coefficients whose denominators are divisible by  $p$ .*
- (3) *Let  $X$  and  $Y$  be members of a nilpotent matrix Lie algebra over a field  $k$  of characteristic  $p$ , of nilpotent order no greater than  $p$ , and suppose that  $X$  and  $Y$  themselves are nilpotent of order no greater than  $p$ . Then*

$$\log(e^X e^Y) = \sum_{i=1}^{p-1} P_i(x, y).$$

*Proof.* From the description of the series for  $\log(e^X e^Y)$  given in Eq. 4.1 we glean

$$P_m(x, y) = \sum \frac{(-1)^{k-1}}{k} \frac{1}{p_1! q_1! \dots p_k! q_k!} x^{p_1} y^{q_1} \dots x^{p_k} y^{q_k},$$

where the summation takes over all tuples of non-negative integers, of whatever length, with the property that  $p_i + q_i > 0$ , and that  $\sum_i p_i + q_i = m$ . As  $m < p$ , clearly we can have no tuple with some  $p_i$  or  $q_i$  divisible by  $p$ , and therefore, neither can we have any  $k \geq p$ , and therefore, claim (1) is established.

For claim (2), simply realize that the only difference in the coefficients for  $P_m(x, y)$  and the equivalent expression  $\phi(P_m(x, y))$  is a multiplication by a factor of  $\frac{1}{m}$  for each coefficient, which again cannot contribute any denominators divisible by  $p$ .

For (3), since  $X$  and  $Y$  are nilpotent of order  $\leq p$ , both the series  $e^X$  and  $e^Y$  are well-defined. Further, since  $X$  and  $Y$  belong to a nilpotent Lie algebra of order  $\leq p$ , any bracket expressions among  $X$  and  $Y$  of length  $p$  or more vanish; accordingly, as  $P_m(X, Y) = \phi(P_m(X, Y))$  consists solely of bracket expressions of length  $m$ , we conclude that  $P_m(X, Y)$  vanishes for all  $m \geq p$ , whence  $\log(e^X e^Y) = \sum_{i=1}^{p-1} P_i(x, y)$ .  $\square$

## 5. THE MAIN THEOREM: SUFFICIENCY

In this section we prove part (1) of Theorem 1.3.

To start, we would like to know when a given Lie algebra homomorphism  $\phi : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$ , over whatever field, can be lifted to a representation of  $U_n$ . For this, we ask how much of the Baker–Campbell–Hausdorff formula do we need to actually be true in this setting. The following proposition is an obvious adaptation of analogous facts for characteristic zero fields.

**Proposition 5.1.** *Let  $k$  be any field, and let  $\phi : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  be a Lie algebra homomorphism. Suppose as follows:*

- (1) *The series  $e^{\phi(X)}$  “makes sense” for all  $X \in \mathfrak{u}_n$ ;*
- (2) *The series  $\log(g)$  “makes sense” for all  $g \in U_n$ , and  $\log(g)$  is a member of  $\mathfrak{u}_n$ ;*
- (3) *For all  $X, Y \in \mathfrak{u}_n$ ,  $\log(e^X e^Y)$  exists, denoted as  $BCH(X, Y)$ . Further,  $BCH(X, Y)$  can be written uniformly (the same for all  $X$  and  $Y$  in  $\mathfrak{u}_n$ ) as a finite linear combination of brackets of  $X$  and  $Y$ , brackets of brackets of  $X$  and  $Y$ , etc.;*
- (4) *For all  $X, Y \in \mathfrak{u}_n$ ,  $\log(e^{\phi(X)} e^{\phi(Y)})$  exists and can be written uniformly as  $BCH(\phi(X), \phi(Y))$  as in (3).*

Then the formula  $\Phi(g) \stackrel{\text{defn}}{=} e^{\phi(\log(g))}$  defines a  $d$ -dimensional representation of  $U_n$  over  $k$ .

*Proof.* Let  $g, h \in U_n$ , and by (2), let  $X = \log(g)$ ,  $Y = \log(h)$ . Then

$$\begin{aligned} \Phi(gh) &= \Phi(e^X e^Y) \\ &= \Phi(e^{BCH(X, Y)}) \quad \text{by (3)} \\ &= e^{\phi(BCH(X, Y))} \quad \text{by definition of } \Phi \\ &= e^{BCH(\phi(X), \phi(Y))} \quad \text{because } \phi \text{ preserves brackets} \\ &= e^{\phi(X)} e^{\phi(Y)} \quad \text{by (4)} \\ &= \Phi(e^X) \Phi(e^Y) \quad \text{by definition of } \Phi \\ &= \Phi(g) \Phi(h). \end{aligned}$$

□

See Chapter 3 of [8] for more on this (for the Lie group case).

**Theorem 5.2.** *Let  $k$  be a field of characteristic  $p > 0$ , and suppose  $p \geq \max(n, d)$ . Let  $\phi : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  be a Lie algebra homomorphism such that  $\phi(X)$  is a nilpotent matrix for each  $X \in \mathfrak{u}_n$ . Then the formula*

$$\Phi(g) = e^{\phi(\log(g))}$$

*defines a  $d$ -dimensional representation of  $U_n$ .*

*Proof.* We shall go through the checklist of Proposition 5.1 to see that they are all satisfied.

(1): The expression  $e^{\phi(X)}$  makes sense since  $\phi(X)$  is nilpotent and of order no greater than  $d \leq p$  (that is, the series expansion for  $e^{\phi(X)}$  terminates before getting to see denominators divisible by  $p$ ).

(2): Every  $g \in U_n$  is unipotent, whence  $g - 1$  is nilpotent of order no greater than  $n \leq p$ , whence the series  $\log(g) = \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k}(g - 1)^k$  likewise terminates before denominators divisible by  $p$  occur.

(3):  $\mathfrak{u}_n$  is a nilpotent Lie algebra of order  $n$ , and each  $X \in \mathfrak{u}_n$  is itself nilpotent. Apply part 3 of Proposition 4.6.

(4): As  $\phi$  is a Lie algebra homomorphism, its image is also a nilpotent Lie algebra, of nilpotent order no greater than  $n \leq p$ . Again apply part 3 of Proposition 4.6.  $\square$

Recall that, if  $M(x_1, \dots, x_n)$  is a matrix with entries taken from the algebra  $k[x_1, \dots, x_n]$  in the commuting indeterminates  $x_1, \dots, x_n$ , then  $M^{[m]}$  denotes the matrix  $M(x_1^m, \dots, x_n^m)$ .

**Lemma 5.3.** Suppose  $(a_{ij})$  is the matrix formula for a representation of  $U_n$  over a field  $k$  of characteristic  $p$ , with comodule map  $V \xrightarrow{\rho} V \otimes A_n$ . Then  $(a_{ij})^{[p]}$  is also a representation of  $U_n$ , with comodule map given by the composition  $V \xrightarrow{\rho} V \otimes A_n \xrightarrow{1 \otimes [p]} A_n$ , where  $A_n \xrightarrow{[p]} A_n$  is the linear map which carries each monomial to its  $p$ th power.

*Proof.* Consider

$$\begin{array}{ccccc}
V & \xrightarrow{\rho} & V \otimes A_n & \xrightarrow{1 \otimes [p]} & V \otimes A \\
\downarrow \rho & & \downarrow 1 \otimes \Delta & & \downarrow 1 \otimes \Delta \\
V \otimes A_n & \xrightarrow{\rho \otimes 1} & V \otimes A_n \otimes A_n & \xrightarrow{1 \otimes [p] \otimes [p]} & V \otimes A_n \otimes A_n \\
\downarrow 1 \otimes [p] & & \downarrow 1 \otimes 1 \otimes [p] & & \downarrow 1 \\
V \otimes A_n & \xrightarrow{\rho \otimes 1} & V \otimes A_n \otimes A_n & \xrightarrow{1 \otimes [p] \otimes 1} & V \otimes A \otimes A .
\end{array}$$

Commutativity of the outermost rectangle is the assertion that the map  $V \xrightarrow{\rho} V \otimes A_n \xrightarrow{1 \otimes [p]} V \otimes A_n$  is a valid  $A_n$ -comodule structure on  $V$ , and it is what we are trying to prove. Commutativity of the top left square follows since  $\rho$  is a comodule map, and commutativity of the bottom left and bottom right squares is obvious. What remains to check is the top right square, i.e.,

$$\begin{array}{ccc}
 A_n & \xrightarrow{[p]} & A_n \\
 \Delta \downarrow & & \downarrow \Delta \\
 A_n \otimes A_n & \xrightarrow{[p] \otimes [p]} & A_n \otimes A_n .
 \end{array}$$

As  $[p]$  and  $\Delta$  are  $k$ -linear, it suffices to check commutativity on the basis of monomials for  $A_n$ . Let  $M = \prod_{1 \leq i < j \leq n} x_{ij}^{r_{ij}}$  be such a monomial, and denote by  $F: A_n \rightarrow A_n$  the Frobenius mapping (that raises *everything* to its  $p$ th power, not just the variables). Clearly,  $F(M) = [p](M)$ . Further, since  $\Delta(M)$  is a linear combination of monomial tensors with *integer* coefficients, and since  $z^p = z \bmod p$  for any integer  $z$ , we also have that  $(F \otimes F)(\Delta(M)) = ([p] \otimes [p])(\Delta(M))$ . As  $F$  is known to make the above diagram commutative, this completes the proof.  $\square$

**Theorem 5.4.** *Suppose that  $p \geq \max(n, d)$ , and let  $k$  be a field of characteristic  $p$ . Let  $\phi_0, \dots, \phi_m$  be a collection of Lie algebra representations  $\mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  such that the following statements hold:*

- (1)  $\phi_i(X)$  is a nilpotent matrix for all  $i$  and all  $X \in \mathfrak{u}_n$ ;
- (2) For all  $i \neq j$  and  $X, Y \in \mathfrak{u}_n$ ,  $\phi_i(X)$  commutes with  $\phi_j(Y)$ .

For  $g \in U_n$  and  $0 \leq i \leq m$ , define

$$\Phi_i(g) = e^{\phi_i(\log(g))}.$$

Then the formula

$$\Phi(g) = \Phi_0(g)\Phi_1(g)^{[p]} \dots \Phi_m(g)^{[p^m]}$$

defines a representation of  $U_n$  over  $k$ .

**Proof.** Theorem 5.2 guarantees that each  $\Phi_i$  is a representation, and the previous lemma says that so is  $\Phi_i(g)^{[p]}$ . As  $\log(g)$  is an element of  $\mathfrak{u}_n$  and as  $\phi_i, \phi_j$  commute when  $i \neq j$ , so do  $\Phi_i(g)$  and  $\Phi_j(g)$ , and hence so do  $\Phi_i(g)^{[p^i]}$  and  $\Phi_j(g)^{[p^j]}$ . Any commuting product of representations of an algebraic group is again a representation, whence  $\Phi$  is a representation of  $U_n$ .  $\square$

Part (1) of Theorem 1.3 is now proved. The main theorem of this article is now proved.

## 6. CHARACTERISTIC ZERO

In this section, we prove Theorem 1.4, the characteristic zero analogue of our main theorem (1.3). This is an almost direct corollary to the arguments used to prove Theorem 1.3, and we shall reference them often.

Throughout this section, let  $n$  and  $d$  be arbitrary, and fix an arbitrary  $d$ -dimensional representation of  $U_n$  over a field  $k$  of characteristic zero, with matrix formula

$$(a_{ij}) = \sum_M \chi(M) x^M$$

The following proposition is the characteristic zero analogue of Theorem 1.5, and implies part (2) of Theorem 1.4.

**Theorem 6.1.** *Let  $k$  be a field of characteristic 0,  $n > 0$ . Then every  $d$ -dimensional representation  $(V, \rho)$  of  $U_n$  over  $k$  satisfies, for all  $1 \leq r < s \leq n$ ,  $1 \leq t < u \leq n$ :*

- (1)  $\chi(\varepsilon_{rs})$  is a nilpotent matrix;
- (2)

$$[\chi(\varepsilon_{rs}), \chi(\varepsilon_{tu})] = \begin{cases} 0 & \text{if } [\varepsilon_{rs}, \varepsilon_{tu}] = 0 \\ \chi([\varepsilon_{rs}, \varepsilon_{tu}]) & \text{otherwise.} \end{cases}$$

The following proposition is the characteristic zero analogue of Proposition 1.6.

**Proposition 6.2.** *Theorem 6.1 implies part (2) of Theorem 1.4.*

*Proof.* Going through the statement and proof of Proposition 1.6, we see that if we consider only the case  $l = m = 0$  (i.e., when there is only one Lie algebra homomorphism  $\phi_0$  to consider), the proof goes through just as well when the field is of characteristic zero.  $\square$

The following is a textbook exercise, and is the characteristic zero analogue of Proposition 3.3.

**Proposition 6.3.** *Let  $k$  have characteristic zero. Then every representation of  $G_a$  over  $k$  is of the form  $e^{xX}$ , where  $X$  is a nilpotent matrix over  $k$ . Further, any nilpotent matrix  $X$  over  $k$  gives a representation of  $G_a$  according to this formula.*

*Proof.* See for example Corollary 2.2 of [2].  $\square$

We note here that all of the results in Section 2 depend in no way on the characteristic of the field, and so they apply to the current setting as well.

The following proposition is the characteristic zero analogue the Proposition 3.7.

**Proposition 6.4.** *For any  $t$  and  $u$ ,*

$$\chi(\varepsilon_m)\chi(\varepsilon_{1u}) = \binom{\varepsilon_{tn} + \varepsilon_{1u}}{\varepsilon_m, \varepsilon_{1u}} \chi(\varepsilon_{tn} + \varepsilon_{1u}).$$

*Proof.* Going through the statement and proof of Proposition 3.7, letting  $l = m = 0$  gives the desired result.  $\square$

The following proposition is the characteristic zero analogue of Proposition 3.8.

**Proposition 6.5.** *If  $t \neq u$ , then*

$$\chi(\varepsilon_{1u})\chi(\varepsilon_{tn}) = \begin{pmatrix} \varepsilon_{1u} + \varepsilon_{tn} \\ \varepsilon_{1u}, \varepsilon_{tn} \end{pmatrix} \chi(\varepsilon_{1u} + \varepsilon_{tn}).$$

*Proof.* Going through the statement and proof of Proposition 3.8, letting  $l = m = 0$  gives the desired result.  $\square$

The following corollary is the characteristic zero analogue of Proposition 3.10.

**Corollary 6.6.** *If  $t \neq u$ , then,*

$$[\chi(\varepsilon_{1u}), \chi(\varepsilon_{tn})] = 0$$

*Proof.* By Propositions 6.4 and 6.5,  $\chi(\varepsilon_{1u})\chi(\varepsilon_{tn}) = \chi(\varepsilon_{tn})\chi(\varepsilon_{1u})$ .  $\square$

The following proposition is the characteristic zero analogue of Proposition 3.10.

**Proposition 6.7.** *For any  $1 \leq t < u \leq n$ ,*

$$[\chi(\varepsilon_{1n}), \chi(\varepsilon_{tu})] = 0$$

*Proof.* See the statement and proof of Proposition 3.10, letting  $l = m = 0$ .  $\square$

Theorem 6.1 is now proved, and hence part (2) of Theorem 1.4.

We now prove part (1) of Theorem 1.4. We note that the entirety of the results of Section 4 are valid in characteristic zero (except for Proposition 4.6, which is irrelevant in this setting). Proposition 5.1 is as well valid in characteristic zero.

The following theorem is the characteristic zero analogue of Theorem 5.2.

**Theorem 6.8.** *Let  $k$  be a field of characteristic 0. Let  $\phi : \mathfrak{u}_n \rightarrow \mathfrak{gl}_d$  be a Lie algebra homomorphism such that  $\phi(X)$  is a nilpotent matrix for each  $X \in \mathfrak{u}_n$ . Then the formula*

$$\Phi(g) = e^{\phi(\log(g))}$$

*defines a  $d$ -dimensional representation of  $U_n$ .*

*Proof.* Keeping in mind that the Baker–Campbell–Hausdorff formula of Section 4 is always valid in characteristic zero, and that there is no reason for the series mentioned in the statement of Proposition 5.1 there not to “make sense,” this is an immediate corollary of Proposition 5.1.  $\square$

Part (1) of Theorem 1.4 is now proved. Theorem 1.4 is now proved.

## 7. FUNCTORIAL ANALOGIES WITH DIRECT PRODUCTS IN CHARACTERISTIC ZERO

The representation theory of direct products of algebraic groups, over any field, can be summed up as follows (assuming of course that the representation theories of the constituent groups have been summed up).

**Theorem 7.1.** *Let  $k$  be any field,  $G, H$  algebraic groups over  $k$ ,  $A$  the representing Hopf algebra of  $G$ , and  $B$  the representing Hopf algebra of  $H$ .*

- (1) *Let  $(c_{ij})$  be the matrix formula for a representation of  $G \times H$ , with  $c_{ij} \in A \otimes B$ , and denote by  $(a_{ij})$  and  $(b_{ij})$  the induced representations on  $G$  and  $H$ , respectively, via the canonical embeddings  $G, H \rightarrow G \times H$ . Then  $(c_{ij})$  can be factored into the commuting product*

$$(a_{ij} \otimes 1)(1 \otimes b_{ij}).$$

- (2) *Any commuting pair of representations for  $G$  and  $H$  yields a representation for  $G \times H$  according to the above formula.*

Further,  $\phi : V \rightarrow W$  is a morphism for the representations  $V$  and  $W$  for  $G \times H$  if and only if it is both a morphism between  $V$  and  $W$  restricted to  $G$ , and for  $V$  and  $W$  restricted to  $H$ .

*Proof.* See for example Chapter 11 of [1]. □

On the other hand, we have the following theorem.

**Theorem 7.2.** *Let  $k$  be a field of characteristic  $p$ , and suppose  $p \geq \max(n, 2d)$ . Let  $\Phi, \Psi$  be representations of  $U_n$  on the vector spaces  $V, W$  respectively, of dimension no greater than  $d$ , and so necessarily of the form*

$$\begin{aligned}\Phi(g) &= \Phi_0(g)\Phi_1(g)\dots\Phi_m(g) \\ \Psi(g) &= \Psi_0(g)\Psi_1(g)\dots\Psi_m(g)\end{aligned}$$

as in Theorem 1.3. Then the linear map  $\phi : V \rightarrow W$  is a morphism between  $\Phi$  and  $\Psi$  if and only if it is a morphism between  $\Phi_0$  and  $\Psi_0$ , between  $\Phi_1$  and  $\Psi_1$ , ..., and between  $\Phi_m$  and  $\Psi_m$ .

*Proof.* See Theorem 10.1.1 of [1]. □

Thus we see that our analogy between the large characteristic representation theory of  $U_n$  and the representation theory of  $U_n^\infty$  in characteristic zero is far from superficial.

We can say even more. Denote by  $\text{Rep}_k U_n$  the category of finite dimensional representations of the algebraic group  $U_n$  over the field  $k$ . If  $k_i$  is any sequence of fields of strictly increasing positive characteristic, denote by  $\prod_H \text{Rep}_{k_i} U_n$  the “height-restricted ultraproduct” of the categories  $\text{Rep}_{k_i} U_n$  (see Section 7.2 and Chapter 14 of [1]). In the author’s dissertation [1] it is proven that  $\prod_H \text{Rep}_{k_i} U_3$  is itself the category representations for some affine group scheme over some field.

**Theorem 7.3.** *Let  $k_i$  be a sequence of fields of strictly increasing positive characteristic. Then  $\prod_H \text{Rep}_{k_i} U_n$  is tensorially equivalent, as neutral tannakian categories, to  $\text{Rep}_{\prod k_i} U_n^\infty$ , where  $U_n^\infty$  denotes a countable direct product of copies of  $U_n$ , and  $\prod k_i$  is the ultraproduct of the fields  $k_i$  (which is necessarily a field of characteristic zero).*

*Proof.* This is proven for the case of  $U_3$  (there called  $H_1$ ) in Theorem 14.0.6 of [1]. Given the main theorem of the present article, the reader should be able to convince himself that the proof given there applies equally well to  $U_n$  for any  $n$ .  $\square$

Finally, we have the following generic cohomology result.

**Theorem 7.4.** *Let  $G = U_n$ , let  $h \in \mathbb{N}$ , and let  $M$  and  $N$  be modules for  $G$  over  $\mathbb{Z}$ . Suppose that the computation  $\dim \text{Ext}_{G^h(k)}^{1,h}(M, N) = m$  (height-restricted cohomology; see Definition 15.2.1 of [1]) is both finite and the same for any characteristic zero field  $k$ . Then if  $k_i$  is any sequence of fields of increasing positive characteristic, the computation*

$$\dim \text{Ext}_{G(k_i)}^{1,h}(M, N)$$

*stabilizes to  $m$  for large enough  $i$ .*

*Proof.* See the proof of Theorem 15.2.3 of [1], where again it is proven for the case of  $U_2$  and  $U_3$ , but with Theorem 7.3 in hand it applies equally well to every  $U_n$ .  $\square$

Generic cohomology results for  $\text{Ext}^n$ ,  $n > 1$ , have been so far elusive; we hope they will be forthcoming in the future (see Section 15.3 of [1] for its apparent difficulty).

## 8. FURTHER DIRECTIONS

Our proof of part (2) of Theorem 1.3 relied essentially on the apparent nice and orderly nature of the representing Hopf algebra  $A_n$  of  $U_n$ ; on the other hand, the proof of part (1) given in Section 5, so far as we can tell, did not, and seems to apply equally well to almost any unipotent algebraic group (except for perhaps the proof of Lemma 5.3, which required that  $\Delta$  of a monomial have only integer coefficients). The author hesitates to make any bold claim concerning this, but it is certainly worth pursuing.

Assuming, then, that the arguments used to prove part (1) Theorem 1.3 apply equally well to an arbitrary unipotent algebraic group, should there not exist an argument to prove (some version of) part (2) as well? Further, Theorem 7.3 makes it clear that, while on the surface our main theorem is about the internal structure of certain  $U_n$ -modules, at heart it is really a *categorical* result; that is, it is as much a statement about the ambient category  $\text{Rep}_k U_n$  as it is about the internal structure of any particular  $U_n$ -module. This leads us to believe that a purely categorical/model theoretic proof of Theorem 7.3 should exist, one which would hopefully apply equally well to an arbitrary unipotent algebraic group, and would give us much of what we need as far as questions of large scale cohomology are concerned. We hope that such insights will be forthcoming in a later article.

**REFERENCES**

- [1] Crumley, M. (2010). *Ultraproducts of Tannakian categories and generic representation theory of unipotent algebraic groups*. Ph.D. thesis, The University of Toledo, Department of Mathematics.
- [2] Crumley, M. (2013). Generic representation theory of the heisenberg group. *Communications in Algebra* 41(8):3174–3206.
- [3] Dascalescu, S., Nastasescu, C., Raianu, S. (2001). *Hopf Algebras: An Introduction*. Pure and Applied Mathematics. New York: Marcel Dekker.
- [4] Dynkin, E. B. (1949). On the representation of the series  $\log(e^x e^y)$  for non-commutative  $x$  and  $y$  by commutators. *Mat. Sbornik (Russian)* 25(67):155–162.
- [5] Dynkin, E. B. (1962). *Lie Groups*, chapter Normed Lie Algebras and Analytic Groups, pages 481–485. Translations, Series One. American Mathematical Society, Providence, Rhode Island.
- [6] Dynkin, E. B. (2000). *Selected Papers of E. B. Dynkin*, chapter Calculation of the coefficients in the Campbell-Hausdorff formula, pages 31–35. American Mathematical Society, Providence, Rhode Island.
- [7] Suslin, A., Friedlander, E. M., Bendel, C. P. (1997). Infinitesimal 1-parameter subgroups and cohomology. *Journal of the AMS* 10(3):693–728.
- [8] Hall, B. C. (2003). *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. New York: Springer-Verlag.
- [9] Jantzen, J. C. (1987). *Representations of Algebraic Groups*, volume 131 of *Pure and Applied Mathematics*. Orlando, FL: Academic Press.
- [10] Springer, T. A. (1981). *Linear Algebraic Groups*. Volume 9 of *Progress in Mathematics*. Boston: Birkhauser.
- [11] Waterhouse, W. C. (1979). *Introduction to Affine Group Schemes*. Graduate Texts in Mathematics. New York: Springer-Verlag.