POSITIVE CHARACTERISTIC REPRESENTATIONS AND DISTRIBUTION ALGEBRAS OF CERTAIN UNIPOTENT ALGEBRAIC GROUPS

MICHAEL CRUMLEY

ABSTRACT. Let G be a unipotent algebraic group over a field k of characteristic $p > 0$ gotten as a subgroup of U_n (unipotent upper triangular group) by setting certain matrix entries to zero. As is well known, in general, the representation theory of G over k is *not* well reflected by the representation theory of $\mathrm{Lie}(G)$ as it is in characteristic 0. However, in this paper we show that, for a large and relevant collection of representations of G over k, the representation theory of G is perfectly reflected by the representation theory of $\mathrm{Lie}(G)$. Along the way we give interesting results concerning the internal structure of the distribution algebras of such G in relation to their Lie algebras in characteristic $p > 0$.

1. INTRODUCTION

Let k be a field of characteristic $p > 0$, and let U_n denote the space of all $n \times n$ unipotent upper triangular matrices over k, i.e. those of the form

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

As is well known, any unipotent algebraic group G over k can be embedded in U_n for some n, and any algebraic subgroup of U_n is unipotent. In this paper we generally prefer to view a unipotent algebraic group G as a fixed subgroup of U_n for some fixed n, as virtually all of our results will rely on such an embedding. Call an algebraic subgroup G of U_n **almost upper triangular** if the defining polynomials of G are of the form

$$(x_{ij} : (i,j) \in S)$$

for some (possibly empty) subset S of $\{(i,j) : 1 \leq i < j \leq n\}$. That is, G is gotten as a subgroup of U_n simply be declaring certain matrix entries to be zero. Examples include the unipotent upper triangular groups themselves, the generalized Heisenberg groups, and any direct sum of almost upper triangular groups; see subsection 9.2 for how other such groups may be constructed. When the particular n for the embedding $G \subset U_n$ under which G is almost upper triangular is either

Date: May 2019.

2010 *Mathematics Subject Classification.* Primary 20G05, 20G15.

Key words and phrases. Generic Representation Theory, Unipotent Algebraic Groups, Additive Group, Heisenberg Group.

unimportant or understood, we shall simply say that G is almost upper triangular (without reference to n).

Denote by $\mathrm{Dist}(G)$ the distribution algebra of G, and by $\mathrm{Lie}(G)$ the Lie algebra of G. The main theorems of this paper are the following:

Theorem 1.1. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$. Then there is an ideal I_G of $Dist(G)$ such that*

$$Lie(G) \oplus Lie(G) \oplus \dots$$

embeds in $Dist(G)/I_G$, and generates $Dist(G)/I_G$ as a k-algebra.

This theorem is proved as proposition 7.6. We give an explicit set of generators for I_G (definitions 6.3 and 7.3) and place important bounds on how large it can be (definition 6.4) in section 6. From this we may prove

Theorem 1.2. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$. Let $\Phi : G \to Aut(V)$ be a finite dimensional representation of G, $\overline{\Phi} : Dist(G) \to End(V)$ the associated algebra representation. Suppose $\overline{\Phi}$ kills the ideal I_G given in theorem 1.1. Then $\overline{\Phi}$ defines Lie algebra maps*

$$\phi_0, \phi_1, \dots, \phi_m : Lie(G) \to Lie(GL(V))$$

ϕ_r commutes with ϕ_s for $r \neq s$, and $\phi_r(X)^p = 0$ for all $X \in Lie(G)$ and $r \geq 0$. Also, these Lie algebra maps completely determine the representation Φ.

This theorem is proved as proposition 7.7, where an explicit formula for these Lie algebra maps $\phi_0, \dots, \phi_m$ can be found, and proposition 7.15 gives a characterization of morphisms between such representations in terms of these Lie algebra maps. With one more condition we can be more explicit:

Theorem 1.3. *Let G be an almost upper triangular subgroup of U_n over a field k of characteristic $p > 0$. Suppose that $p \geq n$. Then any finite dimensional representation $\Phi : G \to Aut(V)$ whose induced representation $\overline{\Phi} : Dist(G) \to End(V)$ kills I_G can be written*

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

where $\phi_0, \phi_1, \dots, \phi_m : Lie(G) \to Lie(GL(V))$ are the commuting Lie algebra representations given in theorem 1.2.

This theorem is proved as proposition 7.14. Finally, we give at least one useful criterion for which theorem 1.2 applies:

Proposition 1.4. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$. Let $\Phi : G \to Aut(V)$ be a finite dimensional representation of G on a d-dimensional vector space V, and suppose that $p \geq 2d$. Then $\overline{\Phi} : Dist(G) \to End(V)$ kills I_G.*

This is proven as proposition 9.2.

We note that a version of this theorem is already known for $G = U_n$ (theorem 1.3 of [2]):

Theorem 1.5. *Let k be a field of characteristic $p > 0$, and suppose that $p \geq max(n, 2d)$.*

1. *let $\phi_0, \phi_1, \dots, \phi_m : Lie(U_n) \to Lie(GL_d)$ be a collection of Lie algebra homomorphisms such that*

(a) *$\phi_i(X)$ is a nilpotent matrix for all $X \in Lie(U_n)$ and $0 \leq i \leq m$*
(b) *For all $i \neq j$ and $X, Y \in Lie(U_n)$, $\phi_i(X)$ 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).*

Thus, the purpose of the present paper is to replace the condition $p \geq \max(n, 2d)$ in theorem 1.5 with a more general one, and to extend this result to all almost upper triangular groups.

Much investigation for this paper was done in Python and Maple. We direct the interested reader to [11] for a Python implementation for doing computations in $\mathrm{Dist}(G)$ for unipotent groups G.

2. Preliminaries

In this paper we will move freely between working with G-modules and A-comodules where A is the representing Hopf algebra of G. We recommend to the reader chapters 1 - 3 of [14] for a review of the correspondence between algebraic groups, their Hopf algebras, and modules for algebraic groups vs. comodules for their Hopf algebras. One may also consult [3] for a review of Hopf algebras in general, and [9] for a review of linear algebraic groups in general. And [10] is a good reference for all things relating to representations of algebraic groups.

Our main theorems will require a knowledge of distribution algebras, which we review here for the reader's convenience. All of what is contained in this section can be found in chapter 7 of [10], but here we will be slightly less general in order to get to what we actually need.

Let A be a Hopf algebra representing the affine group scheme G over a field k, equipped with a co-multiplication map $\Delta : A \to A \otimes A$ and co-unit $\varepsilon : A \to k$. Set $I_1 = \ker(\varepsilon)$, which is an ideal of A. A **distribution** of order $\leq n$ is a linear functional $\mu : A \to k$ such that $\mu(I_1^{n+1}) = 0$, and the space of all such is denoted $\mathrm{Dist}_n(G)$. Also set $\mathrm{Dist}_n^+(G) = \{\mu \in \mathrm{Dist}_n(G) : \mu(1) = 0\}$. Such $\mu \in \mathrm{Dist}_n^+(G)$ are said to be *without constant term*. As $I_1 \supset I_1^2 \supset \dots$, we have $\mathrm{Dist}_0(G) \subset \mathrm{Dist}_1(G) \subset \dots$, and similarly $\mathrm{Dist}_0^+(G) \subset \mathrm{Dist}_1^+(G) \subset \dots$. The **distribution algebra** of G is defined to be

$$\mathrm{Dist}(G) = \bigcup_{n \geq 0} \mathrm{Dist}_n(G)$$

and the collection of distributions without constant term is defined to be

$$\mathrm{Dist}^+(G) = \bigcup_{n \geq 0} \mathrm{Dist}_n^+(G)$$

$\mathrm{Dist}(G)$ is an algebra as follows: for $\mu, \nu \in \mathrm{Dist}(G)$, define their product $\mu\nu$ to be the composition

$$\mu\nu : A \xrightarrow{\Delta} A \otimes A \xrightarrow{\mu \otimes \nu} k \otimes k \simeq k$$

The map ε, which is a distribution of order 0, can be seen to be a multiplicative identity for this algebra. $\mathrm{Dist}(G)$ is generally non-commutative, unless Δ is co-commutative.

$\mathrm{Dist}(G)$ is a graded algebra, using the grading $\mathrm{Dist}_0(G) \subset \mathrm{Dist}_1(G) \subset \ldots$, which is to say that $\mathrm{Dist}_m(G)\mathrm{Dist}_n(G) \subset \mathrm{Dist}_{n+m}(G)$. Also, for $\mu \in \mathrm{Dist}_m(G), \nu \in \mathrm{Dist}_n(G)$,

$$[\mu, \nu] = \mu\nu - \nu\mu \in \mathrm{Dist}_{n+m-1}(G)$$

This bracket result shows $\mathrm{Dist}_1^+(G)$ to be a Lie algebra, and is in fact *the* Lie algebra of G, and we identify it as such. In case of characteristic zero, $\mathrm{Dist}(G)$ is nothing but the enveloping algebra of this Lie algebra, and we effectively get no new information from it. In positive characteristic, this is not so, and this is why Lie algebras are of less utility in characteristic $p > 0$ then they are in characteristic 0 (see section 7.10 of [10]).

Let V be a k-vector space. For any representation of G on V, i.e. a comodule $\rho : V \to V \otimes A$ over A, we get an algebra representation $\overline{\rho} : \mathrm{Dist}(G) \to \mathrm{End}(V)$ as follows: for $\mu \in \mathrm{Dist}(G)$, define $\overline{\rho}(\mu)$ to be the composition

$$\overline{\rho}(\mu) : V \xrightarrow{\rho} V \otimes A \xrightarrow{1\otimes\mu} V \otimes k \simeq V$$

or, if one likes, as the $\mathrm{Dist}(G)$-module structure

$$\mathrm{Dist}(G) \otimes V \xrightarrow{1\otimes\rho} \mathrm{Dist}(G) \otimes V \otimes A \xrightarrow{\text{twist}} \mathrm{Dist}(G) \otimes A \otimes V$$
$$\xrightarrow{\mathrm{ev}\otimes 1} k \otimes V \simeq V$$

where $\mathrm{ev} : \mathrm{Dist}(G) \otimes A \to k$ is the evaluation map $\mu \otimes a \mapsto \mu(a)$.

G-modules therefore always induce $\mathrm{Dist}(G)$-modules, but the reverse is not always so; see section 7.12 of [10] for an example of a $\mathrm{Dist}(G_a)$-module that does not correspond to a G_a-module.

If H is an affine group scheme represented by the Hopf algebra B, and if H is a subgroup of G, then we get a surjective mapping of Hopf algebras $\Omega : A \to B$. This induces an injective mapping of distribution algebras $\overline{\Omega} : \mathrm{Dist}(H) \to \mathrm{Dist}(G)$ given by

$$\overline{\Omega}(\mu) : A \xrightarrow{\Omega} B \xrightarrow{\mu} k$$

and in fact, the image of $\mathrm{Dist}(H)$ under $\overline{\Omega}$ is

$$\mathrm{Dist}(H) \simeq \{\mu \in \mathrm{Dist}(G) : \mu(I) = 0\}$$

where $I = \ker(\Omega)$.

Proposition 2.1. *When Ω is surjective, $\overline{\Omega}$ respects the order of distributions. That is,*

$$\overline{\Omega}(\mathit{Dist}_m(H)) = \mathit{Dist}_m(G) \cap \overline{\Omega}(\mathit{Dist}(H))$$

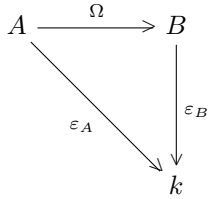
Also,

$$\overline{\Omega}(\mathit{Dist}^+(H)) = \mathit{Dist}^+(G) \cap \overline{\Omega}(\mathit{Dist}(H))$$

and

$$\overline{\Omega}(\mathit{Lie}(H)) = \mathit{Lie}(G) \cap \overline{\Omega}(\mathit{Dist}(H))$$

Proof. Let $\varepsilon_A : A \to k$, $\varepsilon_B : B \to k$ be the co-unit maps. Then by surjectivity of Ω and commutativity of



we have that $\Omega(\ker(\varepsilon_A)) = \ker(\varepsilon_B)$, and likewise, for $m \geq 0$, $\Omega(\ker(\varepsilon_A)^{m+1}) = \Omega(\ker(\varepsilon_A))^{m+1} = \ker(\varepsilon_B)^{m+1}$. For $\mu \in \mathrm{Dist}(H)$ and $m \geq 0$, consider the commutative diagram

$$\begin{array}{ccccc} A & \xrightarrow{\Omega} & B & \xrightarrow{\mu} & k \\ \uparrow{\scriptstyle \iota_A} & & \uparrow{\scriptstyle \iota_B} & & \\ \ker(\varepsilon_A)^{m+1} & \xrightarrow[\Omega']{} & \ker(\varepsilon_B)^{m+1} & & \end{array}$$

where ι_A, ι_B are the inclusion mappings, and Ω' is Ω restricted to $\ker(\varepsilon_A)^{m+1}$. The composition at the top, $\mu \circ \Omega$, is $\overline{\Omega}(\mu)$ by definition. To say that $\mu \in \mathrm{Dist}_m(H)$ is to say that $\mu \circ \iota_B = 0$, and to say that $\overline{\Omega}(\mu) \in \mathrm{Dist}_m(G)$ is to say that $\mu \circ \Omega \circ \iota_A = 0$. But these are easily see to be equivalent, by the surjectivity of Ω'. Thus

$$\overline{\Omega}(\mathrm{Dist}_m(H)) = \mathrm{Dist}_m(G) \cap \overline{\Omega}(\mathrm{Dist}(H))$$

For the second claim, simply realize that the identity $1_B \in B$ is $\Omega(1_A)$ where 1_A is the identity of A. The third claim follows from the other two, using the definition $\mathrm{Lie}(G) = \mathrm{Dist}_1^+(G)$. □

In case G is a unipotent algebraic subgroup of U_n we have

$$\mathrm{Dist}(G) \subset \mathrm{Dist}(U_n)$$

Thus, in studying $\mathrm{Dist}(U_n)$, we are in a sense studying $\mathrm{Dist}(G)$ for all unipotent algebraic groups G.

3. The distribution algebra of U_n

In this section we explicitly describe the distribution algebra $\mathrm{Dist}(U_n)$ of U_n, along with its multiplication law. We also record several products and decompositions in $\mathrm{Dist}(U_n)$ that will be needed later. Some of these results are valid in any characteristic, while others are only valid in characteristic $p > 0$; we shall be explicit about which case we are assuming.

Let k be a field of any characteristic, and let U_n denote the space of all $n \times n$ upper triangular unipotent matrices over k. We identify the Hopf algebra of U_n, called A_n, as

$$A_n = k[x_{ij} : 1 \leq i < j \leq n]$$

$$\Delta : x_{ij} \mapsto 1 \otimes x_{ij} + \sum_{k=i+1}^{j-1} x_{ik} \otimes x_{kj} + x_{ij} \otimes 1$$

$$\varepsilon : x_{ij} \mapsto 0$$

Let $u_n(\mathbb{N})$ denote the space of all $n \times n$ *strictly* upper triangular matrices with non-negative integer entries. Then For $M \in u_n(\mathbb{N})$, let x^M denote the monomial expression

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

so that $A_n = \mathrm{span}_k(x^M : M \in u_n(\mathbb{N}))$.

For $M \in u_n(\mathbb{N})$, let $\alpha(M)$ be the linear functional $A_n \to k$ which sends x^M to 1, all other $x^{M'}$ to zero. Define $|M| = \sum_{ij} m_{ij}$. Also define, for $1 \leq i < j \leq n$, $\varepsilon_{ij} = n \times n$ matrix with 1 in the $(i,j)^{\text{th}}$ entry, zeroes elsewhere.

Proposition 3.1. *Over any field k, $Dist_m(U_n) = span_k(\alpha(M) : |M| \leq m)$, and thus*

$$Dist(U_n) = span_k(\alpha(M) : M \in u_n(\mathbb{N}))$$

Also,

$$Lie(U_n) = span_k(\alpha(\varepsilon_{ij}) : 1 \leq i < j \leq n)$$

Proof. This is worked out in section 7.3 of [10] but we prove it here for convenience. Clearly $I_1 \stackrel{\text{def}}{=} \ker(\varepsilon) = \text{span}_k(x^M : |M| > 0)$, whence $I_1^{m+1} = \text{span}_k(x^M : |M| > m)$. Thus to compute $\text{Dist}_m(U_n)$ we seek functionals $A \to k$ that kill all $x^M, |M| > m$; this is clearly spanned by $\alpha(M) : |M| \leq m$, which proves the first claim, and the second easily follows.

For the last claim, by definition, $\text{Lie}(U_n) = \text{Dist}_1^+(U_n) = \{\mu \in \text{Dist}_1(U_n) : \mu(1) = 0\}$. But the only distribution in $\text{Dist}_1(U_n)$ that does *not* send 1 to 0 is $\alpha(0)$, and the result follows.

□

Next we wish to work out the multiplication in $\text{Dist}(U_n)$, valid in any characteristic. (This is to a large extent done in [2], but inside an arbitrary $\text{Dist}(U_n)$-module, and without reference to distribution algebras.) Let $\alpha(M)\alpha(N)$ denoted the product of $\alpha(M)$ and $\alpha(N)$ in $\text{Dist}(U_n)$. Then we may uniquely write

$$\alpha(M)\alpha(N) = \sum_P C_{M,N}^P \alpha(P)$$

where the summation is over all $P \in u_n(\mathbb{N})$, and $C_{M,N}^P$ is a constant (which must be zero for all but finitely many P). Obviously $C_{M,N}^P = (\alpha(M)\alpha(N))(x^P)$. By definition of the multiplication in $\text{Dist}(U_n)$, $(\alpha(M)\alpha(N))(x^P)$ is the image of x^P under the composition

$$A \stackrel{\Delta}{\longrightarrow} A \otimes A \stackrel{\alpha(M)\otimes\alpha(N)}{\longrightarrow} k \otimes k \simeq k$$

whence

$$C_{M,N}^P = (\alpha(M) \otimes \alpha(N))(\Delta(x^P))$$

with the isomorphism $k \otimes k \simeq k$ being implied.

We shall need some notation. Define the following $n \times n$ variable matrices $S_1, \ldots, S_n$. The non-zero 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, …, S_n is $(n-1)\times$ strictly upper triangular (i.e., S_n has a non-zero entry in its $(1,n)^{\text{th}}$ 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.

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

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

For example, in the case of $n = 4$, $L_{12} = s_{12}^2 + s_{13}^2 + s_{14}^2$, and $R_{34} = s_{14}^3 + s_{24}^2 + s_{34}^1$. For fixed values of the matrices $S_1, \ldots, S_n$ above, let $L(S_1, \ldots, S_n)$ be the strictly upper triangular matrix whose $(i,j)^{\text{th}}$ entry is L_{ij}, similarly for $R(S_1, \ldots, S_n)$.

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

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

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

The variable expressions L_{12} and R_{34} are given by

$$L_{12} = s_{12}^2 + s_{13}^2 + s_{14}^2 \qquad R_{34} = s_{14}^3 + s_{24}^2 + s_{34}^1$$

and the matrices $L = L(S_1, \ldots, S_n)$ and $R = R(S_1, \ldots, S_n)$ are given by

$$L = \begin{pmatrix} 0 & s_{12}^2 + s_{13}^2 + s_{14}^2 & s_{13}^3 + s_{14}^3 & s_{14}^4 \\ & 0 & s_{23}^2 + s_{24}^2 & s_{24}^3 \\ & & 0 & s_{34}^2 \\ & & & 0 \end{pmatrix} \qquad R = \begin{pmatrix} 0 & s_{12}^1 & s_{13}^1 & s_{14}^1 \\ & 0 & s_{13}^2 + s_{23}^1 & s_{14}^2 + s_{24}^1 \\ & & 0 & s_{14}^3 + s_{24}^2 + s_{34}^1 \\ & & & 0 \end{pmatrix}$$

If $B_1 = (b_{ij}^1), \ldots, B_k = (b_{ij}^k), M = (m_{ij})$ are matrices with non-negative integer entries such that $B_1 + \ldots + B_k = M$, then the formal multinomial expression

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

is shorthand for $\prod_{ij} \binom{m_{ij}}{b_{ij}^1, \ldots, b_{ij}^k}$. 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

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

Lemma 3.2. *Let k be any field. For $P \in u_n(\mathbb{N})$*

$$\Delta(x^P) = \sum_{S_1 + \ldots + S_n = P} \binom{P}{S_1, \ldots, S_n} x^{L(S_1, \ldots, S_n)} \otimes x^{R(S_1, \ldots, S_n)}$$

where the summation runs over all $S_1 + S_2 + \ldots + S_n = P$ with non-negative integer entries of the form defined in the above paragraph.

Proof. See the proof of proposition 2.1 of [2], only replace a_{ij} with x^P, and M with P. □

Feeding this expression to $\alpha(M) \otimes \alpha(N)$, we now have the following multiplication law for $\mathrm{Dist}(U_n)$.

Proposition 3.3. *Over any field,* $\alpha(M)\alpha(N) = \sum_P C^P_{M,N}\alpha(P)$*, where*

$$C^P_{M,N} = \sum_{\substack{S_1+\cdots+S_n=P \\ L(S_1,\ldots,S_n)=M \\ R(S_1,\ldots,S_n)=N}} \binom{P}{S_1,\ldots,S_n}$$

where the summation runs over all $S_1 + S_2 + \ldots + S_n = P$ *with non-negative integer entries also satisfying* $L(S_1,\ldots,S_n) = M$ *and* $R(S_1,\ldots,S_n) = N$*, of the form defined in the above paragraph.*

As it is, this is a less-than-straightforward product to work out, although as far as the author can tell it is the tidiest description that can be given. For given M and N, one must find all assignments to the matrices $S_1,\ldots,S_n$ which simultaneously satisfy $L(S_1,\ldots,S_n) = M$ and $R(S_1,\ldots,S_n) = N$; once such a collection is found, you tack on the term $\binom{S_1+\cdots+S_n}{S_1,\ldots,S_n}\alpha(S_1+\cdots+S_n)$. No doubt this rule could be written as an explicit series of summations, with the number of summations increasing with n, but this will work for our purposes.

Note the prominence of multinomial coefficients in this multiplication law. As we shall see, it is their well-behavedness modulo a prime which to a large degree makes our main theorems possible.

Lemma 3.4. *Consider the variable expressions*

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

(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*
   (a) $s_{i,v}^{j-i+1}$ *if* $j = u$
   (b) *None otherwise*

Proof. See lemma 2.4 of [2]. □

For $g \in U_n$, let $\tau(g)$ denote the skew-transpose of g; that is, the transpose of g across the diagonal going from bottom left to top right; that is, the matrix $\tau(g)$ whose $(i,j)^{\text{th}}$ entry is

$$\tau(g)_{ij} = g_{n-j+1,n-i+1}$$

The reader can verify that this is an anti-automorphism of U_n, i.e. $\tau(gh) = \tau(h)\tau(g)$ (or if one likes, $gh = \tau(\tau(h)\tau(g))$. This in turn induces an anti-automorphism of the algebra $\text{Dist}(U_n)$ (which we also call τ), namely the linear map defined by

$$\tau(\alpha(M)) = \alpha(\tau(M))$$

Proposition 3.5. *Let k be any field. Then distributions of the form $\alpha(a\varepsilon_{ij})$, where $1 \leq i < j \leq n$ and $a \in \mathbb{N}$, generate $Dist(U_n)$ as a k-algebra. In fact, for arbitrary M, the distribution $\alpha(M)$ can be decomposed "row-wise" as*

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

and also "column-wise" as

$$\alpha(M) = \prod_{j=n}^{2} \prod_{i=1}^{j-1} \alpha(m_{ij}\varepsilon_{ij})$$

Proof. The proof of part (1) of this proposition can be gleaned from the proof of lemma 3.1 of [2]. Simply replace $\chi(M)$ with $\alpha(M)$, $\chi(m_{ij}\varepsilon_{ij})$ with $\alpha(m_{ij}\varepsilon_{ij})$, etc.

For part (2), let $M' = \tau(M)$, so that $m'_{ij} = m_{n-j+1,n-i+1}$. Write $\alpha(M')$ as in part (1) as

$$\begin{aligned}\alpha(M') &= \prod_{i=n-1}^{1} \prod_{j=i+1}^{n} \alpha(m'_{ij}\varepsilon_{ij}) \\ &= \prod_{i=n-1}^{1} \prod_{j=i+1}^{n} \alpha(m_{n-j+1,n-i+1}\varepsilon_{ij})\end{aligned}$$

If we apply τ once more, after reversing the order of the factors we obtain

$$\begin{aligned}\alpha(M) &= \tau(\alpha(M')) \\ &= \prod_{i=1}^{n-1} \prod_{j=n}^{i+1} \tau(\alpha(m_{n-j+1,n-i+1}\varepsilon_{ij})) \\ &= \prod_{i=1}^{n-1} \prod_{j=n}^{i+1} \alpha(m_{n-j+1,n-i+1}\tau(\varepsilon_{ij})) \\ &= \prod_{i=1}^{n-1} \prod_{j=n}^{i+1} \alpha(m_{n-j+1,n-i+1}\varepsilon_{n-j+1,n-i+1})\end{aligned}$$

Set $i' = n - j + 1, j' = n - i + 1$. Then we can rewrite this product as

$$\alpha(M) = \prod_{j'=n}^{2} \prod_{i'=1}^{j'+1} \alpha(m_{i',j'}\varepsilon_{i',j'})$$

which is exactly the claimed "column-wise" decomposition of $\alpha(M)$.

□

The following is worth mentioning, although we shall not need it in the sequel.

Proposition 3.6. *Let $(i_1, j_1), \ldots, (i_N, j_N)$, where $N = n(n-1)/2$, be any ordering of the set $\{(i,j) : 1 \leq i < j \leq n\}$. Then the collection of products of the form*

$$\prod_{k=1}^{N} \alpha(a_k\varepsilon_{i_k,j_k})$$

for $a_k \in \mathbb{N}$, and where $\sum a_k \leq m$, forms a basis for $Dist_m(U_n)$.

Proof. This is essentially a rework of the proof of the "easy half" of the Poincaré–Birkhoff–Witt theorem.

Let $\alpha(M)$ be a basis element of $\mathrm{Dist}_m(U_n)$; we wish to show that $\alpha(M)$ is a linear combination of products of the above form.

We note firstly that the proposition is true for at least two orderings (though one is enough for our purposes), namely the two found in the decompositions of $\alpha(M)$ given in proposition 3.5. Second, for any ordering, the proposition is true when $m = 1$, that is, when $M = \varepsilon_{ij}$ for some (i, j) (just let $a_k = 1$ when $(i_k, j_k) = (i, j)$, all other $a'_k = 0$).

We will use the well-known fact that any ordering of any set can be achieved by starting with any initial ordering, coupled with a series of adjacent transpositions. Thus, if the proposition is true for some ordering, and also true after any adjacent transposition of that ordering, by induction on the number of adjacent transpositions required, the proposition is true for all orderings.

So suppose the proposition is true for all orderings when $|M| \leq m - 1$, and suppose $|M| = m$. Note that this means that $\alpha(M) \in \mathrm{Dist}_m(U_n)$. Pick an ordering for which the proposition is true when $|M| = m$, write this ordering as $(i_1, j_1), \ldots, (i_N, j_N)$, so we may write

$$\alpha(M) = \sum_k c_k \alpha(a_{k,1}\varepsilon_{i_1,j_1}) \ldots \alpha(a_{k,N}\varepsilon_{i_N,j_N})$$

Clearly if we can show the proposition to be true for each of the products $\alpha(a_{k,1}\varepsilon_{i_1,j_1}) \ldots \alpha(a_{k,N}\varepsilon_{i_N,j_N})$, the proposition is true for M, so we can assume that $\alpha(M)$ consists of the single product

$$\alpha(M) = \alpha(a_1\varepsilon_{i_1,j_1}) \ldots \alpha(a_N\varepsilon_{i_N,j_N})$$

with $\sum a_i \leq m$.

Consider two adjacent indices, $(i_r, j_r), (i_{r+1}, j_{r+1})$, and write

$$\begin{aligned}
\alpha(M) &= \alpha(a_1\varepsilon_{i_1,j_1}) \ldots \alpha(a_r\varepsilon_{i_r,j_r})\alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}}) \ldots \alpha(m_{i_N,j_N}\varepsilon_{i_N,j_N}) \\
&= \alpha(a_1\varepsilon_{i_1,j_1}) \ldots \big(\alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}})\alpha(a_r\varepsilon_{i_r,j_r}) + [\alpha(a_r\varepsilon_{i_r,j_r}), \alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}})]\big) \\
&\quad \ldots \alpha(a_N\varepsilon_{i_N,j_N}) \\
&= \alpha(a_1\varepsilon_{i_1,j_1}) \ldots \alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}})\alpha(a_r\varepsilon_{i_r,j_r}) \ldots \alpha(a_N\varepsilon_{i_N,j_N}) \\
&\quad + \alpha(a_1\varepsilon_{i_1,j_1}) \ldots [\alpha(a_r\varepsilon_{i_r,j_r}), \alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}})] \ldots \alpha(a_N\varepsilon_{i_N,j_N})
\end{aligned}$$

where $[x, y]$ denotes the commutator $xy - yx$. Now the first summand of the last expression is exactly of the form we are looking for, so there nothing to show there. For the second summand, note that since $\alpha(a_r\varepsilon_{i_r,j_r}) \in \mathrm{Dist}_{a_r}(U_n)$ and $\alpha(a_{r+1}\varepsilon_{i_{r+1},j_{r+1}}) \in \mathrm{Dist}_{a_{r+1}}(U_n)$, their bracket belongs to $\mathrm{Dist}_{a_r+a_{r+1}-1}(U_n)$. By repeated application of the fact that $\mathrm{Dist}_t(U_n)\mathrm{Dist}_u(U_n) \subset \mathrm{Dist}_{t+u}(U_n)$, we see that the second summand belongs to $\mathrm{Dist}_s(U_n)$, where $s = a_1 + \cdots + a_r + a_{r+1} - 1 + \cdots + a_N$, which is no greater than $m - 1$. Thus by induction, this term can be written in the desired form as well.

This shows that such distributions span $\mathrm{Dist}_m(G)$. To see that they form a basis, simply note that there are exactly as many of them as the dimension of $\mathrm{Dist}_m(G)$. □

Proposition 3.7. *Let k be any field. Let $1 \leq r < s \leq n$, $1 \leq t < u \leq n$, and suppose $s \neq t$, and also that $r \neq t$ or $s \neq u$. Let a, b be any non-negative integers. Then*

$$\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{tu}) = \alpha(a\varepsilon_{rs} + b\varepsilon_{tu})$$

Proof. From proposition 3.3 we have

$$\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{tu}) = \sum_P C^P_{a\varepsilon_{rs},b\varepsilon_{tu}}\alpha(P)$$

where

$$C^P_{a\varepsilon_{rs},b\varepsilon_{tu}} = \sum_{\substack{S_1+\cdots+S_n=P \\ L=a\varepsilon_{rs} \\ R=b\varepsilon_{tu}}} \binom{P}{S_1,\ldots,S_n}$$

It falls to us then to find non-negative integer values for the entries of the variable matrices $S_1,\ldots,S_n$ such that $L = a\varepsilon_{rs}$ and $R = b\varepsilon_{tu}$ of the form defined above. That is, we seek $S_1,\ldots,S_n$ such that

$$L_{rs} = a, \text{ all other } L_{ij} = 0$$

$$R_{tu} = b, \text{ all other } R_{ij} = 0$$

We claim that there is only one such solution, namely

$$s^{s-r+1}_{rs} = a \qquad s^1_{tu} = b$$

and all other $s^k_{ij} = 0$. Examine

$$L_{rs} = \sum_{k=s}^{n} s^{s-r+1}_{rk} \qquad R_{tu} = \sum_{k=1}^{t} s^{t-u+1}_{ku}$$

Note first that by lemma 3.4, L_{rs} and R_{tu} share no variables in common. Second, note that the only variable occurring in L_{rs} which *doesn't* also occur in some R_{ij} is s^{s-r+1}_{rs}, and the only variable occurring in R_{tu} which *doesn't* occur in some L_{ij} is s^1_{tu}. This forces all of the variables occurring in L_{rs} or R_{tu} to be zero, except for these two, forcing $s^{s-r+1}_{rs} = a$, $s^1_{tu} = b$.

Thus the only P for which $C^P_{a\varepsilon_{rs},b\varepsilon_{tu}} \neq 0$ is when $P = S_1 + \cdots + S_n$ for the values of s^k_{ij} just described. These in turn give $S_{s-r+1} = a\varepsilon_{rs}$, $S_1 = b\varepsilon_{tu}$, and all other $S_k = 0$, whence

$$\begin{aligned}\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{tu}) &= \binom{a\varepsilon_{rs} + b\varepsilon_{tu}}{a\varepsilon_{rs}, b\varepsilon_{tu}}\alpha(a\varepsilon_{rs} + b\varepsilon_{tu}) \\ &= \alpha(a\varepsilon_{rs} + b\varepsilon_{tu})\end{aligned}$$

□

Proposition 3.8. *Let k be any field. Let $\varepsilon_{r_1,s_1}, \varepsilon_{r_2,s_2}, \ldots, \varepsilon_{r_m,s_m}$ be such that they are all in the same row (so $r_1 = r_2 = \cdots = r_m$) or in the same column (so $s_1 = s_2 = \cdots = s_m$), but no two of which are equal (so that for all $i \neq j$, $r_i \neq r_j$ or $s_i \neq s_j$). Let $a_1,\ldots,a_m \in \mathbb{N}$. Then*

$$\alpha(a_1\varepsilon_{r_1,s_1})\alpha(a_2\varepsilon_{r_2,s_2})\ldots\alpha(a_m\varepsilon_{r_m,s_m}) = \alpha(a_1\varepsilon_{r_1,s_1} + a_2\varepsilon_{r_2,s_2} + \cdots + a_m\varepsilon_{r_m,s_m})$$

Further, all the terms of this product commute.

Proof. Suppose they are all in the same row. Then we can decompose the right hand side of this equation "row-wise" according to proposition 3.5 and the claimed equality is immediate. In case they are all in the same column, a "column-wise" decomposition will show the same.

What remains to show is that the terms of the product commute. Consider the product $\alpha(a_i\varepsilon_{r_i,s_i})\alpha(a_j\varepsilon_{r_j,s_j})$. If they are in the same row, so that $r_i = r_j$, then since we always have $r_i < s_i$, necessarily $s_i \neq r_j$, and proposition 3.7 applies, giving

$$\alpha(a_i\varepsilon_{r_i,s_i})\alpha(a_j\varepsilon_{r_j,s_j}) = \alpha(a_i\varepsilon_{r_i,s_i} + a_j\varepsilon_{r_j,s_j})$$

But the same holds true of the reverse product $\alpha(a_j\varepsilon_{r_j,s_j})\alpha(a_i\varepsilon_{r_i,s_i})$, giving the same answer, whence they commute. If they are instead in the same column, so that $s_i = s_j$, then proposition 3.7 applies in a similar fashion, forcing them to commute. □

Proposition 3.9. *Let k be any field. For $0 \leq r < s < t \leq n$,*

$$\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{st}) = \sum_{k=0}^{min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{st} + k\varepsilon_{rt})$$

Proof. As in the proof of the proposition 3.7, we seek non-negative integer values for the s_{ij}^k such that

$$L_{rs} = a, \text{ all other } L_{ij} = 0$$

$$R_{st} = b, \text{ all other } R_{ij} = 0$$

Firstly, L_{rs} and R_{st} have exactly one variable in common, namely s_{rt}^{s-r+1} (lemma 3.4). Secondly, there is exactly one variable occurring in L_{rs} which doesn't occur in any R_{ij}, namely s_{rs}^{s-r+1}, and exactly one variable occurring in R_{st} which doesn't occur in any of the L_{ij}, namely s_{st}^1 (lemma 3.4 again). This gives

$$s_{rs}^{s-r+1} + s_{rt}^{s-r+1} = a \text{ and } s_{st}^1 + s_{rt}^{s-r+1} = b$$

with all other s_{ij}^k equal to zero. Clearly then, every non-negative integer value of s_{rt}^{s-r+1} no greater than either b or a gives a solution, and these are the only solutions. If $k = s_{rt}^{s-r+1}$ is any such value, one can check that

$$S_1 + \cdots + S_n = (a-k)\varepsilon_{rs} + (b-k)\varepsilon_{st} + k\varepsilon_{rt}$$

and that

$$\binom{S_1 + \cdots + S_n}{S_1, \ldots, S_n} = 1$$

This gives

$$\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{st}) = \sum_{k=0}^{\min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{st} + k\varepsilon_{rt})$$

as claimed.

□

The "p-digits" of an integer, i.e. the integers r_i in the p-ary decomposition $r = r_0 + r_1 p + \cdots + r_m p^m$ shall figure prominently when working in characteristic $p > 0$. We shall need the following.

Definition 3.10. Let p be a prime, and let $a, b, \ldots, z$ be non-negative integers, written in p-ary notation as

$$\begin{aligned} a &= a_0 + a_1 p + \cdots + a_m p^m \\ b &= b_0 + b_1 p + \cdots + b_m p^m \\ &\vdots \\ z &= z_0 + z_1 p + \cdots + z_m p^m \end{aligned}$$

We say that the sum $a + b + \cdots + z$ **carries** if, for some i, $a_i + b_i + \cdots + z_i \geq p$. If $a \leq b$, we say that the difference $b - a$ **borrows** if for some i, $a_i > b_i$.

As a matter of notation, if we write something like "$(a+b)+c$ carries", we mean that we have computed the sum $a+b$ first, and *then* asked if the result carries with c. This is of course not the same as saying that $a+b+c$, as a sum of three separate integers, carries.

Proposition 3.11. *(Lucas' theorem) Let k be a field of characteristic $p > 0$. Write the non-negative integers $a, b, \ldots, z$ in p-ary notation, as in the previous definition. Then the multinomial coefficient*

$$\binom{a + b + \cdots + z}{a, b, \ldots, z}$$

is equal to

(1) *0 if the sum $a + b + \cdots + z$ carries;*
(2)

$$\binom{a_0 + b_0 + \cdots + z_0}{a_0, b_0, \ldots, z_0} \binom{a_1 + b_1 + \cdots + z_1}{a_1, b_1, \ldots, z_1} \cdots \binom{a_m + b_m + \cdots + z_m}{a_m, b_m, \ldots, z_m}$$

otherwise.

Proof. See theorems 14 and 15 of [4]. □

For fixed i and j, those matrices of the form $1 + r\varepsilon_{ij}$, for $r \in k$, form a subgroup of U_n isomorphic to G_a, the additive group over k. Thus, the span of those distributions of the form $\alpha(a\varepsilon_{ij})$, for $a \in \mathbb{N}$ and for fixed i and j, form a sub-algebra of $\mathrm{Dist}(U_n)$ isomorphic to $\mathrm{Dist}(G_a)$. We shall need

Proposition 3.12. *Fix $1 \leq i < j \leq n$.*

(1) *For any field k and non-negative integers $a_1, \ldots, a_m$,*

$$\alpha(a_1\varepsilon_{ij})\alpha(a_2\varepsilon_{ij}) \ldots \alpha(a_m\varepsilon_{ij}) = \binom{a_1 + \cdots + a_m}{a_1, \ldots, a_m} \alpha((a_1 + \cdots + a_m)\varepsilon_{ij})$$

(2) *Suppose k has characteristic zero, and let $r \in \mathbb{N}$. Then*

$$\alpha(r\varepsilon_{ij}) = \frac{1}{r!}\alpha(\varepsilon_{ij})^r$$

(3) *Suppose k has characteristic $p > 0$, let $r \in \mathbb{N}$, and write r in p-ary notation as $r = r_0 + r_1 p + \cdots + r_m p^m$. Then*

$$\alpha(r\varepsilon_{ij}) = \Gamma(r)^{-1}\alpha(\varepsilon_{ij})^{r_0}\alpha(p\varepsilon_{ij})^{r_1} \ldots \alpha(p^m\varepsilon_{ij})^{r_m}$$

where

$$\Gamma(r) \stackrel{defn}{=} r_0!r_1!\dots r_m!$$

Also, $\alpha(\varepsilon_{ij}), \alpha(p\varepsilon_{ij}), \dots, \alpha(p^m\varepsilon_{ij})$ *all commute and are nilpotent of order* p*.*

Proof. For (1), the case of $m = 2$, i.e. $\alpha(a_1\varepsilon_{ij})\alpha(a_2\varepsilon_{ij}) = \binom{a_1+a_2}{a_1,a_2}\alpha((a_1 + a_2)\varepsilon_{ij})$ can be found in example 7.8 of [10], and by induction the well-known identity

$$\binom{a_1 + \dots + a_m + a_{m+1}}{a_1, \dots, a_m, a_{m+1}} = \binom{a_1 + \dots + a_m}{a_1, \dots, a_m}\binom{a_1 + \dots + a_m + a_{m+1}}{a_1 + \dots + a_m, a_{m+1}}$$

shows it to be true for arbitrary m. Part (2) follows easily from part (1), using that $r!$ is always invertible in characteristic zero.

Part (3) was effectively proved in proposition 1.2 of [8], but we prove it here for convenience. By part (1) we have

$$\begin{aligned}\binom{r}{r_0, r_1p, \dots, r_mp^m}\alpha(r\varepsilon_{ij}) &= \alpha(r_0\varepsilon_{ij})\alpha(r_1p\varepsilon_{ij})\dots\alpha(r_mp^m\varepsilon_{ij}) \\ &= r_0!r_1!\dots r_m!\alpha(\varepsilon_{ij})^{r_0}\alpha(p\varepsilon_{ij})^{r_1}\dots\alpha(p^m\varepsilon_{ij})^{r_m}\end{aligned}$$

Obviously the sum $r_0 + r_1p + \dots + r_mp^m$ does not carry, so by proposition 3.11, the multinomial coefficient in front of $\alpha(r\varepsilon_{ij})$ is exactly

$$\begin{aligned}\binom{r}{r_0, r_1p, \dots, r_mp^m} &= \binom{r_0}{r_0}\binom{r_1}{r_1}\dots\binom{r_m}{r_m} \\ &= 1\end{aligned}$$

and the claimed equality is true. For the nilpotency claim, if $0 \le s < p$ we have

$$\alpha(sp^k\varepsilon_{ij}) = s!\alpha(p^k\varepsilon_{ij})^s$$

and since p doesn't divide s, $s! \neq 0$, and so also $\alpha(p^k\varepsilon_{ij})^s \neq 0$. On the other hand, again by part (1) we have

$$\begin{aligned}\alpha(p^k\varepsilon_{ij})^p &= \alpha(p^k\varepsilon_{ij})\alpha(p^k\varepsilon_{ij})\dots\alpha(p^k\varepsilon_{ij}) \\ &= \binom{p^{k+1}}{p^k, p^k, \dots, p^k}\alpha(p^{k+1}\varepsilon_{ij}) \\ &= 0\end{aligned}$$

since obviously the sum $p^k + p^k + \dots + p^k = p^{k+1}$ carries.

□

Since in any characteristic the $\alpha(a\varepsilon_{ij})$ generate $\mathrm{Dist}(U_n)$, and since, by part (3) of proposition 3.12, in characteristic $p > 0$ the $\alpha(p^m\varepsilon_{ij})$ generate all of the $\alpha(a\varepsilon_{ij})$, we have

Proposition 3.13. *Let* k *have characteristic* $p > 0$*. Then* $Dist(U_n)$ *is generated by those distributions of the form*

$$\alpha(p^m\varepsilon_{ij})$$

where $m \in \mathbb{N}$ *and* $1 \le i < j \le n$*.*

We shall need the following results about the commutator operation $[x, y] = xy - yx$ in $\mathrm{Dist}(U_n)$.

Proposition 3.14. *Let k be any field, a, b non-negative integers. Then* $[\alpha(a\varepsilon_{rs}), \alpha(b\varepsilon_{tu})] =$

$$\begin{cases} 0 & \text{if } s \neq t \text{ and } r \neq u \\ \sum_{k=1}^{min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{tu} + k\varepsilon_{ru}) & \text{if } s = t \\ -\sum_{k=1}^{min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{tu} + k\varepsilon_{ts}) & \text{if } r = u \end{cases}$$

Proof. If $s \neq t$ and $r \neq u$, then proposition 3.7 applies to both of the products $\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{tu})$ and $\alpha(a\varepsilon_{tu})\alpha(b\varepsilon_{rs})$, giving the same answer in each case, making their commutator zero.

If $s = t$, then proposition 3.9 applies to the product $\alpha(a\varepsilon_{rs})\alpha(b\varepsilon_{tu})$, whereas it is impossible for $r = u$, whence proposition 3.7 applies to the product $\alpha(b\varepsilon_{tu})\alpha(a\varepsilon_{rs})$. This gives

$$\begin{aligned} [\alpha(a\varepsilon_{rs}), \alpha(b\varepsilon_{tu})] &= \sum_{k=0}^{\min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{tu} + k\varepsilon_{ru}) - \alpha(a\varepsilon_{rs} + b\varepsilon_{tu}) \\ &= \sum_{k=l}^{\min(a,b)} \alpha((a-k)\varepsilon_{rs} + (b-k)\varepsilon_{tu} + k\varepsilon_{ru}) \end{aligned}$$

since letting $k = 0$ in the summation exactly cancels with the last term.

And if $r = u$, the same argument applies to $[\alpha(a\varepsilon_{rs}), \alpha(b\varepsilon_{tu})] = -[\alpha(b\varepsilon_{tu}), \alpha(a\varepsilon_{rs})]$. □

Two specific instances of the previous proposition will be of special interest to us and deserve recording.

Proposition 3.15. *Let k be any field, p a positive integer, and suppose $m \geq 0$. Then* $[\alpha(p^m\varepsilon_{rs}), \alpha(p^m\varepsilon_{tu})] =$

$$\begin{cases} 0 & \text{if } s \neq t \text{ and } r \neq u \\ \alpha(p^m\varepsilon_{ru}) + \sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tu} + k\varepsilon_{ru}) & \text{if } s = t \\ -\alpha(p^m\varepsilon_{st}) - \sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tu} + k\varepsilon_{ts}) & \text{if } r = u \end{cases}$$

Proposition 3.16. *Let k be any field, p any positive integer, and suppose $m \neq l$. Then* $[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{tu})] =$

$$\begin{cases} 0 & \text{if } s \neq t \text{ and } r \neq u \\ \sum_{k=1}^{min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{tu} + k\varepsilon_{ru}) & \text{if } s = t \\ -\sum_{k=1}^{min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{tu} + k\varepsilon_{ts}) & \text{if } r = u \end{cases}$$

4. Representations of Dist(U_n)

Let k be any field, V a finite dimensional k-vector space. Not all representations of Dist(U_n) correspond to representations of U_n, but the ones we care about do. In this section we show how a group representation $U_n \to \mathrm{Aut}(V)$, or G more generally, actually gives rise to an algebra representation $\mathrm{Dist}(G) \to \mathrm{End}(V)$.

Consider for example the Heisenberg group U_3 over k, i.e. all 3×3 matrices of the form

$$g = \begin{pmatrix} 1 & x_{12} & x_{13} \\ & 1 & x_{23} \\ & & 1 \end{pmatrix}$$

and the representation of U_3 given by the formula

$$\Phi(g) = \begin{pmatrix} 1 & 2x_{12} & x_{12} & 2x_{12}^2 & x_{13} & 2x_{12}x_{13} \\ & 1 & 0 & x_{12} & 0 & x_{13} \\ & & 1 & 2x_{12} & x_{23} & 2x_{12}x_{23} \\ & & & 1 & 0 & x_{23} \\ & & & & 1 & 2x_{12} \\ & & & & & 1 \end{pmatrix}$$

say in the basis $\{e_1, \ldots, e_6\}$ for V. Let M be the matrix

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

We claim that the image of $\alpha(M)$ under the associated algebra map $\overline{\Phi} : \mathrm{Dist}(U_3) \to \mathrm{End}(V)$, in this basis for V, is simply

$$\overline{\Phi}(\alpha(M)) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 2 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}$$

That is, it is simply the "matrix of coefficients" of the monomial $x^M = x_{12}x_{13}$ in the above formula for the representation. To see this, let $\rho : V \to V \otimes A_n$ be the A_n-comodule structure on V corresponding to the U_n-representation Φ on V. Then we must have, for $j = 1, \ldots, 6$ and for the "general element" $g \in U_n$,

$$\rho(e_j) = \sum_i e_i \otimes \Phi(g)_{ij}$$

By definition, the algebra representation $\overline{\Phi}$ acts on V via the composition

$$\overline{\Phi}(\alpha(M)) : V \xrightarrow{\rho} V \otimes A_n \xrightarrow{1 \otimes \alpha(M)} V \otimes k \simeq V$$

Feeding the basis element e_j to this composition gives

$$\overline{\Phi}(\alpha(M))(e_j) = \sum_i \alpha(M)(\Phi(g)_{ij}) e_i$$

Thus $\overline{\Phi}(\alpha(M))$ acts on V in this basis exactly as $\alpha(M)(\Phi(g))$, that is, as the matrix gotten by applying $\alpha(M)$ entry-wise to $\Phi(g)$. But this is exactly the above claimed matrix.

Representations always map unipotent algebraic groups to unipotent algebraic groups. Thus, up to a base change, the image of any representation of a unipotent group G is a collection of unipotent upper triangular matrices. Coupled with the observations of this section, this gives

Proposition 4.1. *For any representation of a unipotent algebraic group G on a finite dimensional vector space V, and for any $\mu \in Dist^+(G)$, the associated representation $Dist(G) \to End(V)$ maps μ to a nilpotent linear transformation.*

Proof. Recall that $\mathrm{Dist}^+(G)$ consists precisely of those distributions μ such that $\mu(1) = 0$. The image of any representation of G is, up to a base change, a collection

of unipotent upper triangular matrices (necessarily with $1's$ on the main diagonal). As argued above, the image of μ in $\mathrm{End}(V)$ is, in this basis, the matrix

$$\mu(\Phi(g))$$

If $\mu \in \mathrm{Dist}^+(G)$, then the main diagonal of $\mu(\Phi(g))$ must be zero, making it a strictly upper triangular matrix, hence nilpotent.

□

We can also see that

Proposition 4.2. *Let $\rho : V \to V \otimes B$ be a finite dimensional representation of a unipotent algebraic group G. Then ρ is completely determined by the associated algebra map $\overline{\rho} : Dist(G) \to End(V)$.*

Proof. Let G be represented by the Hopf algebra B, and let $\Omega : A_n \to B$ be the surjective Hopf algebra map corresponding to the embedding of G into U_n. Let $\omega_G = \ker(\Omega)$, so that $B \simeq A_n/\omega_G$. Let S be a set of polynomials in A_n such that $\{f + \omega_G : f \in S\}$ is a basis for B. For each $f \in S$, let $\mu_f : B \to k$ be the linear map which sends $f + \omega_G$ to 1, and for all other $g \in S$, sends $g + \omega_G$ to zero. We claim that each μ_f is an element of $\mathrm{Dist}(G)$. Write $f = \sum_i c_i x^{M_i}$, let $m = \max(|M_i|)$, and consider, as in proposition 2.1, the commutative diagram

$$\begin{array}{ccccc} A & \xrightarrow{\Omega} & B & \xrightarrow{\mu_f} & k \\ \uparrow{\scriptstyle \iota_A} & & \uparrow{\scriptstyle \iota_B} & & \\ \ker(\varepsilon_A)^{m+1} & \xrightarrow[\Omega']{} & \ker(\varepsilon_B)^{m+1} & & \end{array}$$

As in the proof of proposition 3.1, $\ker(\varepsilon_A)^{m+1} = \mathrm{span}_k(x^M : |M| > m)$; this shows that $\mu_f \circ \Omega \circ \iota_A = 0$. But then also $\mu_f \circ \iota_B \circ \Omega' = 0$, and since Ω' is surjective, $\mu_f \circ \iota_B = 0$, showing that $\mu_f(\ker(\varepsilon_B)^{m+1}) = 0$, showing that $\mu_f \in \mathrm{Dist}(G)$.

Let $\{e_1, \ldots, e_l\}$ be a basis for V, and write

$$\rho : e_j \mapsto \sum_i e_i \otimes b_{ij}$$

We would like to see that we can recover each of the b_{ij} from the (assumed to be known) algebra map $\overline{\rho} : \mathrm{Dist}(G) \to \mathrm{End}(V)$. Write each b_{ij} as

$$b_{ij} = \sum_{f \in S} c_{ij}^f (f + \omega_G)$$

with $c_{ij}^f = 0$ for all but finitely many $f \in S$. Thus, we would like to recover each c_{ij}^f. Clearly we have $\mu_f(b_{ij}) = c_{ij}^f$. Also, by definition of $\overline{\rho}$,

$$\begin{aligned}
\overline{\rho}(\mu_f)(e_j) &= (1 \otimes u_f)(\rho(e_j)) \\
&= (1 \otimes \mu_f)(\sum_i e_i \otimes b_{ij}) \\
&= \sum_i e_i \otimes \mu_f(b_{ij}) \\
&= \sum_i e_i \otimes c_{ij}^f \\
&\simeq \sum_i c_{ij}^f e_i
\end{aligned}$$

But $\overline{\rho}(\mu_f)(e_j)$ is known by assumption, hence so are each of the c_{ij}^f, and the proposition is proved.

□

5. The m^{TH} Lie algebra of U_n

Let k be any field, and consider $\text{Lie}(U_n) \subset \text{Dist}(U_n)$, which we have identified as $\text{span}_k(\alpha(\varepsilon_{ij}) : 1 \leq i < j \leq n)$. By proposition 3.14, and as it must be, the multiplication law in $\text{Dist}(U_n)$ exactly gives the Lie bracket that one expects to see in $\text{Lie}(U_n)$. That is,

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

On the other hand, consider

Definition 5.1. Let k have characteristic $p > 0$ and let $m \geq 0$. Then we define $\text{Lie}_m(U_n)$, the $m^{\textbf{th}}$ **Lie algebra** of U_n, to be the following subspace of $\text{Dist}(U_n)$:

$$\text{Lie}_m(U_n) = \text{span}_k(\alpha(p^m \varepsilon_{ij}) : 1 \leq i < j \leq n)$$

The terminology "Lie algebra" is abusive, since (proposition 3.15) for $m > 0$ it is not the case that $\text{Lie}_m(G)$ is a Lie algebra; in particular, it is *not* generally the case that

$$[\alpha(p^m \varepsilon_{rs}), \alpha(p^m \varepsilon_{tu})] = \begin{cases} \alpha(p^m \varepsilon_{ru}) & \text{if } s = t \\ -\alpha(p^m \varepsilon_{ts}) & \text{if } r = u \\ 0 & \text{otherwise} \end{cases}$$

Neither is it generally the case that (proposition 3.16) for $m \neq l$,

$$[\alpha(p^m \varepsilon_{rs}), \alpha(p^l \varepsilon_{tu})] = 0$$

As we shall see, these two complications essentially explain why, at least in the case of almost upper triangular groups, positive characteristic representation theory is so much more bothersome than in characteristic zero, and that representations in characteristic $p > 0$ do not generally correspond to Lie algebra representations as in theorem 1.2. Nonetheless, it may in fact be the case that, for a representation $U_n \to \text{Aut}(V)$, that the associated algebra representation $\text{Dist}(U_n) \to \text{End}(V)$ *does* map $\text{Lie}_m(U_n)$ onto a Lie algebra in $\text{End}(V)$, and makes (the images of) $\text{Lie}_m(U_n)$

and $\mathrm{Lie}_l(U_n)$ commute for $m \neq l$, and it is exactly when this happens that our main theorems apply.

Proposition 5.2. *Over a field k of characteristic $p > 0$, the set*

$$\bigcup_{m=0}^{\infty} Lie_m(U_n)$$

generates $Dist(U_n)$ as a k-algebra. Also, each $X \in Lie_m(U_n)$ is nilpotent of order p.

Proof. The first claim follows directly from proposition 3.13. The second follows from the additivity of the p^{th} power operation on $\mathrm{Dist}(G)$, and part (3) of proposition 3.12.

□

We can give a somewhat more natural definition of $\mathrm{Lie}_m(U_n)$ as follows.

Definition 5.3. Let k be a field of characteristic $p > 0$. The **pseudo-Frobenius map** $F : A_n \to A_n$ is the k-linear map defined by

$$F(x^M) = x^{pM}$$

for all monomials $x^M \in A_n$.

Clearly F is a map of k-algebras (via the identities $(x^M + x^N)^p = x^{pM} + x^{pN}$ and $(x^M x^N)^p = x^{pM} x^{pN}$), and for $m \in \mathbb{N}$, $F^m(x^M) = x^{p^m M}$. Note that, if the Frobenius endomorphism $k \to k$ happens to be the identity mapping (i.e. if $k = \mathbb{Z}/p\mathbb{Z}$), then F is exactly the Frobenius mapping $a \mapsto a^p$, but otherwise, this is not so; for $c \in k$, we have $F(cx^M) = cx^{pM}$ by definition, whereas $(cx^M)^p = c^p x^{pM}$.

Proposition 5.4. *Over a field of characteristic $p > 0$, $F : A_n \to A_n$ is a map of Hopf algebras.*

Proof. We wish to prove commutativity of

$$\begin{array}{ccc} A_n & \xrightarrow{F} & A_n \\ \Big\downarrow{\scriptstyle \Delta} & & \Big\downarrow{\scriptstyle \Delta} \\ A_n \otimes A_n & \xrightarrow[F \otimes F]{} & A_n \otimes A_n \end{array}$$

As F and Δ are k-linear it suffices to check commutativity on the basis of monomials for A_n. Let $x^M = \prod_{1 \leq i < j \leq n} x_{ij}^{m_{ij}}$ be such a monomial. Starting with x^M at the top left, going right, and then down gives

$$\Delta(F(x^M)) = \Delta(x^{pM}) = \Delta((x^M)^p) = \Delta(x^M)^p$$

because Δ is an algebra map. Starting at the top left, going down, and then right gives

$$(F \otimes F)(\Delta(x^M))$$

Now $\Delta(x^M)$ can be written

$$\sum_i z_i (x^{M_i} \otimes x^{N_i})$$

with each z_i an integer (mod p) and x^{M_i}, x^{N_i} a monomial, whence

$$\begin{aligned}(F \otimes F)(\Delta(x^M)) &= \sum_i z_i (F(x^{M_i}) \otimes F(x^{N_i})) \\ &= \sum_i z_i (x^{pM_i} \otimes x^{pN_i})\end{aligned}$$

We also have

$$\begin{aligned}\Delta(x^M)^p &= (\sum_i z_i (x^{M_i} \otimes x^{N_i}))^p \\ &= \sum_i z_i^p (x^{pM_i} \otimes x^{pN_i})\end{aligned}$$

But by Fermat's little theorem, $z_i^p = z_i$, whence

$$\Delta(x^M)^p = \sum z_i (x^{pM_i} \otimes x^{pN_i})$$

Thus $(F \otimes F)(\Delta(x^M)) = \Delta(F(x^M))$, completing the proof. □

The dual mapping $F^* : A_n^* \to A_n^*$ given by

$$F^*(\mu)(a) = \mu(F(a))$$

thus restricts to an algebra mapping $F^* : \mathrm{Dist}(U_n) \to \mathrm{Dist}(U_n)$. For a basis element $\alpha(M) \in \mathrm{Dist}(U_n)$, it can be easily shown that

$$F^{*m}(\alpha(M)) = \begin{cases} \alpha(M/p^m) & \text{if } p^m \text{ divides } m_{ij} \text{ for all } i,j \\ 0 & \text{otherwise} \end{cases}$$

whence

$$\ker(F^{*m}) = \mathrm{span}_k(\alpha(M) : \text{some entry of } M \text{ is not divisible by } p^m)$$

Proposition 5.5. *For $m \in \mathbb{N}$,*

$$F^{*m}(Lie_m(U_n)) = Lie(U_n)$$

and in fact, F^{*m} defines an isomorphism of k-vector spaces between $Lie_m(U_n)$ and $Lie(U_n)$. Also

$$(F^{*m})^{-1}(Lie(U_n)) = Lie_m(U_n) \oplus ker(F^{*m})$$

and

$$Lie_m(U_n) = (F^{*m})^{-1}(Lie(U_n)) \cap Dist_{p^m}(U_n)$$

Proof. From the description of F^{*m} above, clearly $F^{*m}(\alpha(p^m \varepsilon_{ij})) = \alpha(\varepsilon_{ij})$ for all $1 \le i < j \le n$, and $\ker(F^{*m}) \cap \mathrm{Lie}_m(U_n) = 0$. Clearly also the map F^{*m} acts bijectively between $\mathrm{Lie}_m(U_n)$ and $\mathrm{Lie}(U_n)$; this proves the first and second claim.

For the last claim, recall that $\mathrm{Dist}_{p^m}(U_n) = \mathrm{span}_k(\alpha(M) : \sum_{ij} m_{ij} \le p^m)$. Then clearly the only way that $\alpha(M) \in \mathrm{Dist}_{p^m}(U_n)$, and $F^{*m}(\alpha(M)) \in \mathrm{Lie}(U_n)$, is if $M = p^m \varepsilon_{ij}$ for some $1 \le i < j \le n$, or if $F^{*m}(\alpha(M)) = 0$, and the result follows. □

6. The impediment to $\mathrm{Lie}_m(U_n)$ being a Lie algebra

Throughout this section let k be a field of characteristic $p > 0$.

We would like to know why, exactly, $\mathrm{Lie}_m(U_n)$ is not a Lie algebra for $m > 0$, why $\mathrm{Lie}_m(U_n)$ does not commute with $\mathrm{Lie}_l(U_n)$ for $m \neq l$, and if there are any circumstances under which we should expect them to. This is explained in this section.

Definition 6.1. Let k have characteristic $p > 0$. An **impediment of the first kind** is a distribution of either the form

$$\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{st} + k\varepsilon_{rt})$$

or

$$-\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tr} + k\varepsilon_{ts})$$

where $m > 0$. An **impediment of the second kind** is a distribution of either the form

$$\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{st} + k\varepsilon_{rt})$$

or

$$-\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{tr} + k\varepsilon_{ts})$$

Where $m, l \geq 0$ and $m \neq l$.

We note that, if $r \neq u$ and $s \neq t$, then by propositions 3.15 and 3.16, for any $m, l \geq 0$, we already have

$$[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{tu})] = 0$$

It remains to consider when either $r = u$ or $s = t$.

Proposition 6.2. *Let k be a field of characteristic $p > 0$, $m \neq l$.*

1. $[\alpha(p^m\varepsilon_{rs}), \alpha(p^m\varepsilon_{st})] - \alpha(p^m\varepsilon_{rt})$ *is an impediment of the first kind*
2. $[\alpha(p^m\varepsilon_{rs}), \alpha(p^m\varepsilon_{tr})] + \alpha(p^m\varepsilon_{ts})$ *is an impediment of the first kind*
3. $[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{st})]$ *is an impediment of the second kind*
4. $[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{tr})]$ *is an impediment of the second kind*

Proof. By propositions 3.15 and 3.16 we have the following:

$$[\alpha(p^m\varepsilon_{rs}), \alpha(p^m\varepsilon_{st})] - \alpha(p^m\varepsilon_{rt}) = \sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{st} + k\varepsilon_{rt})$$

$$[\alpha(p^m\varepsilon_{rs}), \alpha(p^m\varepsilon_{tr})] + \alpha(p^m\varepsilon_{ts}) = -\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tr} + k\varepsilon_{ts})$$

$$[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{st})] = \sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{st} + k\varepsilon_{rt})$$

$$[\alpha(p^m\varepsilon_{rs}), \alpha(p^l\varepsilon_{tr})] = -\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tr} + k\varepsilon_{ts})$$

These are all impediments of the first or second kind, proving the proposition. □

Definition 6.3. The **ideal of impediments** for $\mathrm{Dist}(U_n)$ is the ideal I in $\mathrm{Dist}(U_n)$ generated by all impediments of the first and second kind.

We do not in this paper give a characterization of I (other than to specify its generators). Rather, we shall be content to show that it is not "too big", in the sense that it intersects each $\mathrm{Lie}_m(U_n)$ trivially.

Definition 6.4. The **carrying ideal** of $\mathrm{Dist}(U_n)$ is the following subspace J of $\mathrm{Dist}(U_n)$:

$$J = \mathrm{span}_k \left(\alpha(M) : \sum_{ij} m_{ij} \text{ carries} \right)$$

That is, J consists of linear combinations of $\alpha(M)$ such that the sum of the entries in M carries (as a sum of $N = n(n-1)/2$ integers).

It is not obvious that J is an ideal in characteristic $p > 0$ (and it is certainly not in characteristic 0); we will prove this shortly. J is certainly an interesting ideal in it's own right, but the reason we are presently interested in it is because

Proposition 6.5. *Impediments of the first and second kind belong to J, whence I is contained in J (assuming J is an ideal). Further, $Lie_m(U_n) \cap J = 0$ for all m.*

Proof. The last claim is obvious from the descriptions of J and $\mathrm{Lie}_m(U_n)$. Consider an impediment of the first kind:

$$\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{st} + k\varepsilon_{rt})$$

Clearly for every value of k in this summation, the difference $p^m - k$ borrows, whence $(p^m - k) + k$ carries; thus all the terms of this summation belong to J.

Consider an impediment of the second kind:

$$\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{st} + k\varepsilon_{rt})$$

and suppose $m > l$. Clearly for every value k in this summation, the difference $p^m - k$ borrows, whence $(p^m - k) + k$ carries; thus all the terms of this summation belong to J. If instead $l > m$, a similar argument shows still all the terms belong to J. □

This proposition by no means shows that $I = J$, and in fact we suspect this not to be so (based on evidence generated in Python). However, once we have shown that J is an ideal, whence $I \subset J$, this is enough to prove our main theorem, since it must also be the case that $\mathrm{Lie}_m(U_n) \cap I = 0$, which is all we shall need.

We now prove that J is a two-sided ideal in $\mathrm{Dist}(U_n)$.

Lemma 6.6. *Let $\alpha(M)$ be a basis element of Dist(U_n), $1 \leq r < s \leq n$, $a \in \mathbb{N}$. Then*

$$\alpha(a\varepsilon_{rs})\alpha(M) = \sum_{\substack{u_s+u_{s+1}+\cdots+u_n=a \\ u_{s+1}\leq m_{s,s+1} \\ u_{s+2}\leq m_{s,s+2} \\ \vdots \\ u_n \leq m_{s,n}}} \binom{A+B}{A,B} \alpha(A+B)$$

where

$$A = u_s\varepsilon_{rs} + u_{s+1}\varepsilon_{r,s+1} + \cdots + u_n\varepsilon_{r,n}$$
$$B = M - [u_{s+1}\varepsilon_{s,s+1} + u_{s+2}\varepsilon_{s,s+2} + \cdots + u_n\varepsilon_{s,n}]$$

Proof. We work directly with the multiplication rule of Dist(U_n) (proposition 3.3) applied to the product $\alpha(a\varepsilon_{rs})\alpha(M)$:

$$\alpha(a\varepsilon_{rs})\alpha(M) = \sum_P \left(\sum_{\substack{S_1+\cdots+S_n=P \\ L(S_1,\ldots,S_n)=a\varepsilon_{rs} \\ R(S_1,\ldots,S_n)=M}} \binom{P}{S_1,\ldots,S_n} \right) \alpha(P)$$

Thus we seek solutions in the matrices $S_1, \ldots, S_n$ given in section 3 to the system of equations

$$L(S_1, \ldots, S_n) = a\varepsilon_{rs}$$
$$R(S_1, \ldots, S_n) = M$$

We claim that, once values for the variables s_{ij}^k occurring in the expression L_{rs} are fixed, i.e. those in the expression

$$L_{rs} = s_{rs}^{s-r+1} + s_{r,s+1}^{s-r+1} + \cdots + s_{r,n}^{s-r+1}$$

that the rest of the s_{ij}^k are determined; and further, that for every choice of these variables satisfying

$$\begin{aligned} L_{rs} &= a \\ s_{r,s+1}^{s-r+1} &\leq m_{s,s+1} \\ s_{r,s+2}^{s-r+1} &\leq m_{s,s+2} \\ &\vdots \\ s_{r,n}^{s-r+1} &\leq m_{s,n} \end{aligned}$$

there does indeed exist such a solution.

Assume then that values for $s_{rs}^{s-r+1}, s_{r,s+1}^{s-r+1}, \ldots, s_{r,n}^{s-r+1}$ have been fixed. We must have, for all L_{ij} except L_{rs}, $L_{ij} = 0$; and by part (1) of lemma 3.4, this forces values upon all variables occurring in the matrices $S_2, S_3, \ldots, S_n$. Thus, if there is any additional freedom to be had at all, it must be in the matrix S_1. But we claim there is none. For this we examine the condition $R(S_1, \ldots, S_n) = M$, along with the expression

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

Note that any of the variables in S_1 show up exactly once in each of the R_{ij}, and since all other values are forced, along with that of R_{ij} itself, so also is that of $s_{i,j}^1$,

hence all of S_1. Thus, we have shown that, once values for $s_{rs}^{s-r+1}, s_{r,s+1}^{s-r+1}, \ldots, s_{r,n}^{s-r+1}$ have been fixed, all other s_{ij}^k are determined.

Now we show such a solution always exists. As noted, the condition $L(S_1, \ldots, S_n) = a\varepsilon_{rs}$ leaves free the variables occurring in S_1, and a variable in S_1 shows up in each of the R_{ij}, and none show up in more than one R_{ij}. Thus, so long as the above inequalities are satisfied, one is free to choose these variables to satisfy the condition $R_{ij} = m_{ij}$, whence a solution exists.

Finally, given fixed values for $s_{rs}^{s-r+1}, s_{r,s+1}^{s-r+1}, \ldots, s_{r,n}^{s-r+1}$, our forced solution makes all of the S_j equal to zero, except S_1 and S_{s-r+1}. Then we leave it to the reader to verify that these can be written

$$\begin{aligned} S_{s-r+1} &= s_{rs}^{s-r+1}\varepsilon_{rs} + s_{r,s+1}^{s-r+1}\varepsilon_{r,s+1} + \cdots + s_{r,n}^{s-r+1}\varepsilon_{r,n} \\ S_1 &= M - [s_{r,s+1}^{s-r+1}\varepsilon_{s,s+1} + s_{r,s+2}^{s-r+1} + \cdots + s_{r,n}^{s-r+1}\varepsilon_{s,n}] \end{aligned}$$

The final step is simply the renaming

$$\begin{aligned} &u_s = s_{rs}^{s-r+1}, u_{s+1} = s_{r,s+1}^{s-r+1}, \ldots, u_n = s_{r,n}^{s-r+1} \\ &A = S_{s-r+1} \\ &B = S_1 \end{aligned}$$

This completes the proof.

□

Lemma 6.7. *Let $a_1, \ldots, a_m$, b, $c_1, \ldots, c_k$, $d_1, \ldots, d_k$, $t_0, t_1, \ldots, t_k$ be non-negative integers such that*

(1) $t_1 \leq d_1, t_2 \leq d_2, \ldots, t_k \leq d_k$
(2) $a_1 + \cdots + a_m + b + c_1 + \cdots + c_k + d_1 + \cdots + d_k$ *carries*
(3)

$$\begin{aligned} &a_1 + \cdots + a_m + (b + t_0) \\ &+ (c_1 + t_1) + \cdots + (c_k + t_k) \\ &+ (d_1 - t_1) + \cdots + (d_k - t_k) \end{aligned}$$

does not carry

Then at least one of the sums

$$\begin{aligned} &b + t_0 \\ &c_1 + t_1 \\ &c_2 + t_2 \\ &\vdots \\ &c_k + t_k \end{aligned}$$

must carry.

Proof. Suppose that (1) and (2) hold, but that none of $b + t_0$, $c_1 + t_1$, …, $c_k + t_k$ carry; we claim that (3) cannot hold, i.e. that the sum stated in (3) must in fact carry.

For an integer z, let z_i denote the i^{th} digit of z in the p-ary decomposition $z = z_0 + \cdots + z_i p^i + \cdots + z_m p^m$. If none of $b + t_0$, $c_1 + t_1$, …, $c_k + t_k$ carry, then for

all i we have $(b+t_0)_i = b_i + t_{0,i}$, $(c_1+t_1)_i = c_{1,i}+t_{1,i}, \ldots, (c_k+t_k)_i = c_{k,i}+t_{k,i}$. Then we have, for all i,

$$\begin{aligned}
&a_{1,i} + \cdots + a_{m,i} + (b+t_0)_i \\
&\quad + (c_1+t_1)_i + \cdots + (c_k+t_k)_i \\
&\quad + (d_1-t_1)_i + \cdots + (d_k-t_k)_i \\
=&a_{1,i} + \cdots + a_{m,i} + b_i + t_{0,i} \\
&\quad + c_{1,i} + t_{1,i} + \cdots + c_{k,i} + t_{k,i} \\
&\quad + (d_1-t_1)_i + \cdots + (d_k-t_k)_i
\end{aligned}$$

Suppose first that none of the differences $d_s - t_s$ borrow. Then $(d_s-t_s)_i = d_{s,i} - t_{s,i}$ for all i, which gives

$$\begin{aligned}
=&a_{1,i} + \cdots + a_{m,i} + b_i + t_{0,i} \\
&\quad + c_{1,i} + t_{1,i} + \cdots + c_{k,i} + t_{k,i} \\
&\quad + d_{1,i} - t_{1,i} + \cdots + d_{k,i} - t_{k,i} \\
=&a_{1,i} + \cdots + a_{m,i} + b_i + t_{0,i} \\
&\quad + c_{1,i} + \cdots + c_{k,i} \\
&\quad + d_{1,i} + \cdots + d_{k,i}
\end{aligned}$$

Since the sum in (2) carries by assumption, for some i, $a_{1,i} + \cdots + a_{m,i} + b_i + c_{1,i} + \cdots + c_{k,i} + d_{1,i} + \cdots + d_{k,i} \geq p$, showing that the sum in (3) does in fact carry (in the same digit that the sum in (2) carries).

So suppose that at least one of the differences $d_s - t_s$ does in fact borrow. By re-ordering if necessary, assume $d_k - t_k$ borrows. Let j be least such that $d_{k,j} < t_{k,j}$. Then by the "borrowing algorithm" we learned in grade school, we have $(d_k-t_k)_j = p + d_{k,j} - t_{k,j}$, whence

$$\begin{aligned}
&a_{1,j} + \cdots + a_{m,j} + (b+t_0)_j \\
&\quad + (c_1+t_1)_j + \cdots + (c_k+t_k)_j \\
&\quad + (d_1-t_1)_j + \cdots + (d_k-t_k)_j \\
=&a_{1,j} + \cdots + a_{m,j} + b_j + t_{0,j} \\
&\quad + c_{1,j} + t_{1,j} + \cdots + c_{k,j} + t_{k,j} \\
&\quad + (d_1-t_1)_j + \cdots + (d_{k-1}-t_{k-1})_j \\
&\quad + p + d_{k,j} - t_{k,j} \\
=&a_{1,j} + \cdots + a_{m,j} + b_i + t_{0,j} \\
&\quad + c_{1,j} + t_{1,j} + \cdots + c_{k,j} \\
&\quad + (d_1-t_1)_j + \cdots + (d_{k-1}-t_{k-1})_j \\
&\quad + p + d_{k,j} \\
\geq& p
\end{aligned}$$

which shows that the sum in (3) carries in its j^{th} digit.

□

Proposition 6.8. *J is a left ideal in Dist(U_n).*

Proof. Let $\alpha(M)$ be a basis element of J, so that $\sum_{ij} m_{ij}$ carries. Let $a \in \mathbb{N}$, $1 \leq r < s \leq n$. We wish to show that $\alpha(a\varepsilon_{rs})\alpha(M) \in J$.

By lemma 6.6 we can write

$$\alpha(a\varepsilon_{rs})\alpha(M) = \sum_{\substack{u_s+u_{s+1}+\cdots+u_n=a \\ u_{s+1}\leq m_{s,s+1} \\ u_{s+2}\leq m_{s,s+2} \\ \vdots \\ u_n \leq m_{s,n}}} \binom{A+B}{A,B} \alpha(A+B)$$

where

$$\begin{aligned} A &= u_s\varepsilon_{rs} + u_{s+1}\varepsilon_{r,s+1} + \cdots + u_n\varepsilon_{r,n} \\ B &= M - [u_{s+1}\varepsilon_{s,s+1} + u_{s+2}\varepsilon_{s,s+2} + \cdots + u_n\varepsilon_{s,n}] \end{aligned}$$

Fix values for $u_s, u_{s+1}, \ldots, u_n$ in this summation, and hence values for the matrices A and B. It is entirely possible (and frequently happens) that $\alpha(A+B)$ does not belong to J, i.e. that the sum $\sum_{ij}(A+B)_{ij}$ does not carry. However, we claim that, if this is the case, then the binomial coefficient $\binom{A+B}{A,B}$ must be zero.

For this we apply lemma 6.7. To use the notation of that lemma, let

$$\begin{aligned} b &= m_{rs} \\ c_1 &= m_{r,s+1}, c_2 = m_{r,s+2}, \ldots, c_k = m_{r,n} \\ d_1 &= m_{s,s+1}, d_2 = m_{s,s+2}, \ldots, d_k = m_{s,n} \\ t_0 &= u_s, t_1 = u_{s+1}, \ldots, t_k = u_n \end{aligned}$$

and let $a_1, \ldots, a_m$ be names for any of the m_{ij} not already mentioned. By assumption $\sum m_{ij}$ carries, which is to say that $a_1+\cdots+a_m+b+c_1+\cdots+c_k+d_1+\cdots+d_k$ carries, satisfying condition (2) of lemma 6.7. Further, saying that $\alpha(A+B) \notin J$, i.e. that $\sum_{ij}(A+B)_{ij}$ does not carry, can be seen to be identical to condition (3) of lemma 6.7. Thus we may invoke the conclusion of that lemma, which says that at least one of

$$\begin{aligned} &b + t_0 \\ &c_1 + t_1 \\ &c_2 + t_2 \\ &\vdots \\ &c_k + t_k \end{aligned}$$

must carry, i.e. that at least one of

$$\begin{aligned} &m_{rs} + u_s \\ &m_{r,s+1} + u_{s+1} \\ &m_{r,s+2} + u_{s+2} \\ &\vdots \\ &m_{r,n} + u_n \end{aligned}$$

must carry. And these in turn can be replaced with

$$\begin{aligned} &B_{rs} + A_{rs} \\ &B_{r,s+1} + A_{r,s+1} \\ &B_{r,s+2} + A_{r,s+2} \\ &\vdots \\ &B_{r,n} + A_{r,n} \end{aligned}$$

Thus we see that, for at least one pair of indices (i, j), $A_{ij}+B_{ij}$ must carry, and thus by proposition 3.11, the binomial coefficient $\binom{A+B}{A,B}$ must be zero. This completes the proof.

□

Proposition 6.9. *J is a two-sided ideal in Dist(U_n).*

Proof. It remains to show that J is a right ideal. Let $\alpha(M)$ be a basis element of J, $\alpha(N)$ any basis element of $\mathrm{Dist}(U_n)$. Recall the anti-automorphism τ of $\mathrm{Dist}(U_n)$ given in section 3. Clearly $\tau(J) = J$. Then we can write

$$\alpha(M)\alpha(N) = \tau(\tau(\alpha(N))\tau(\alpha(M)))$$

Since $\alpha(M) \in J$, so also $\tau(\alpha(M)) \in J$, and since J is known to be a left ideal, $\tau(\alpha(N))\tau(\alpha(M)) \in J$. Then so also $\tau(\tau(\alpha(N))\tau(\alpha(M))) = \alpha(M)\alpha(N) \in J$.

□

Now that we know that J is an ideal, necessarily $I \subset J$, whence $I \cap \mathrm{Lie}_m(U_n) = 0$. We now have

Proposition 6.10. *For all $m \geq 0$, Lie$_m(U_n)$ embeds in Dist$(U_n)/I$, and under this identification, is in fact a Lie algebra, isomorphic to Lie(U_n). Lie$_m(U_n)$ and Lie$_l(U_n)$ commute inside Dist$(U_n)/I$ for $m \neq l$, whence*

$$\mathit{Lie}_0(U_n) \oplus \mathit{Lie}_1(U_n) \oplus \ldots \simeq \mathit{Lie}(U_n) \oplus \mathit{Lie}(U_n) \oplus \ldots$$

embeds in Dist$(U_n)/I$. Finally, Lie$_0(U_n) \oplus$ Lie$_1(U_n) \oplus \ldots$ generates Dist$(U_n)/I$ as a k-algebra.

Proof. That $\mathrm{Lie}_m(U_n)$ embeds in $\mathrm{Dist}(U_n)/I$ follows from $\mathrm{Lie}_m(U_n) \cap I = 0$. All impediments to $\mathrm{Lie}(U_n)$ being a Lie algebra isomorphic to $\mathrm{Lie}(U_n)$ been killed by the modding out by I, likewise any impediment to $\mathrm{Lie}_m(U_n)$ commuting with $\mathrm{Lie}_l(U_n)$ for $m \neq l$ have been killed, and clearly still $\mathrm{Lie}_m(U_n) \cap \mathrm{Lie}_l(U_n) = 0$ in $\mathrm{Dist}(U_n)/I$; this proves the second claim. The third claim follows since

$$\bigcup_{m=0}^{\infty} \mathrm{Lie}_m(U_n)$$

generates all of $\mathrm{Dist}(U_n)$, hence obviously all of $\mathrm{Dist}(U_n)/I$.

□

7. The main theorems

Let G be a unipotent algebraic subgroup of U_n for some n over a field of characteristic $p > 0$, and let B denote the Hopf algebra of G. Then we have a surjective Hopf algebra map $\Omega : A_n \to B$; let us call the kernel of this map ω_G. We can of course think of ω_G as being generated by the "defining equations" of G as an algebraic group; that is, $g \in U_n$ is a member of G if and only if, for every polynomial $f(x_{ij} : 1 \leq i < j \leq n) \in \omega_G$, the entries of g satisfy $f = 0$. In case G is almost upper triangular (which we will be assuming later) we have

$$\omega_G = (x_{ij} : (i,j) \in S)$$

for some subset S of $\{(i,j) : 1 \leq i < j \leq n\}$.

From this we get an injective mapping of distribution algebras $\overline{\Omega} : \mathrm{Dist}(G) \to \mathrm{Dist}(U_n)$ given by, for $\mu \in \mathrm{Dist}(G)$,

$$\overline{\Omega}(\mu)(a) = \mu(\Omega(a))$$

and under this mapping we may identify

$$\mathrm{Dist}(G) = \{\mu \in \mathrm{Dist}(A_n) : \mu(\omega_G) = 0\}$$

We also have, under this identification

$$\mathrm{Lie}(G) \subset \mathrm{Lie}(U_n)$$

Henceforth we shall identify $\mathrm{Dist}(G)$, $\mathrm{Lie}(G)$, etc. with their images in $\mathrm{Dist}(U_n)$ under $\overline{\Omega}$.

Let $F : A_n \to A_n$ be the pseudo-Frobenius map described in definition 5.3, and suppose that $F(\omega_G) \subset \omega_G$. When this is the case, the map $F^* : \mathrm{Dist}(U_n) \to \mathrm{Dist}(U_n)$ restricts to a map $F^* : \mathrm{Dist}(G) \to \mathrm{Dist}(G)$. One can see this as follows. We would like to define a map $F_B : B \to B$ which makes commutative the diagram

$$\begin{array}{ccc} A_n & \xrightarrow{F} & A_n \\ \Big\downarrow{\scriptstyle \Omega} & & \Big\downarrow{\scriptstyle \Omega} \\ B & \xrightarrow[F_B]{} & B \end{array}$$

That is, via the definition

$$F_B(b) \stackrel{\mathrm{def}}{=} \Omega(F(\Omega^{-1}(b)))$$

Ω is surjective, so this map is defined, but it is not necessarily well-defined. That is, for any two $a_1, a_2 \in A_n$ such that $\Omega(a_1) = \Omega(a_2)$, we need $\Omega(F(a_1)) = \Omega(F(a_2))$. By the linearity of F and Ω, this is obviously equivalent to requiring that $\Omega(F(a)) = 0$ whenever $a \in \ker(\Omega) = \omega_G$, which is equivalent to $\Omega(F(\omega_G)) = 0$, which is equivalent to $F(\omega_G) \subset \omega_G$.

Assuming then that $F(\omega_G) \subset \omega_G$, we have the dual mapping map $F_B^* : \mathrm{Dist}(G) \to \mathrm{Dist}(G)$ given by

$$F_B^*(\mu)(b) = \mu(F_B(b))$$

We claim that, under the identification $\mathrm{Dist}(G) \subset \mathrm{Dist}(U_n)$ given by $\overline{\Omega}$, F_B^* is simply F^* restricted to $\mathrm{Dist}(G)$; that is, for all $\mu \in \mathrm{Dist}(G)$,

$$F^*(\overline{\Omega}(\mu)) = \overline{\Omega}(F_B^*(\mu))$$

To see this, for $a \in A_n$, the left hand side gives

$$\begin{aligned} F^*(\overline{\Omega}(\mu))(a) &= \overline{\Omega}(\mu)(F(a)) \\ &= \mu(\Omega(F(a))) \end{aligned}$$

whereas the right hand side gives

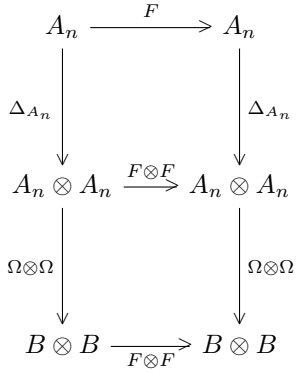
$$\begin{aligned} \overline{\Omega}(F_B^*(\mu))(a) &= F_B^*(\mu)(\Omega(a)) \\ &= \mu(F_B(\Omega(a))) \end{aligned}$$

But $\Omega(F(a)) = F_B(\Omega(a))$ by construction, whence they are equal.

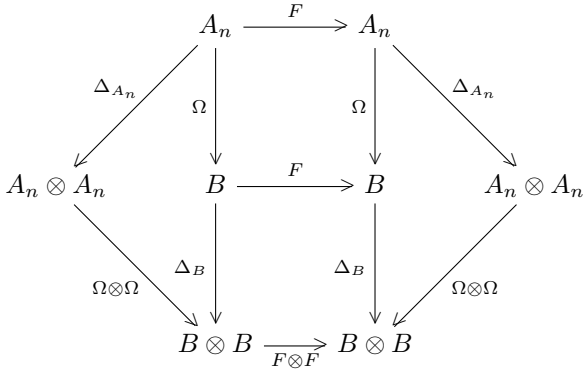
Let us then drop the subscript and refer to this map $F_B^* : \mathrm{Dist}(G) \to \mathrm{Dist}(G)$ (when it exists) as F^* as well, and similarly write F instead of F_B.

Proposition 7.1. *Supppose $F(\omega_G) \subset \omega_G$, so that $F : B \to B$ exists. Then $F : B \to B$ is a Hopf algebra map.*

Proof. Consider first



The top square commutes because F is a Hopf algebra map for A_n (proposition 5.4), and the bottom square commutes by definition of F induced on B; thus the outermost rectangle commutes as well. Consider next



Commutativity of the bottom middle square is the assertion that F is a Hopf algebra map on B, and is what we are trying to prove. The top middle square commutes by the definition of F induced on B, the left-most and right-most polygons (which are in fact identical) commute because Ω is a Hopf algebra map, and the outermost

hexagon is the outermost rectangle of the previous diagram, which commutes. Start at the top left copy of A_n; with some element chasing, one can show that

$$\begin{aligned}\Delta_B \circ F \circ \Omega &= \Delta_B \circ \Omega \circ F \\ &= (\Omega \otimes \Omega) \circ \Delta_{A_n} \circ F \\ &= (F \otimes F) \circ (\Omega \otimes \Omega) \circ \Delta_{A_n} \\ &= (F \otimes F) \circ \Delta_B \circ \Omega\end{aligned}$$

But Ω is surjective, whence $\Delta_B \circ F = (F \otimes F) \circ \Delta_B$, which shows commutativity of the bottom square, and the proposition is proved.

□

The author does not see why it should be that $F(\omega_G) \subset \omega_G$ in general. But, there is at least one common instance when this is so, and includes the case of G being almost upper triangular. Call a unipotent algebraic group G **defined over** $\mathbb{Z}/p\mathbb{Z}$ if the defining polynomials of G (i.e. generators of ω_G) can be taken to have integer (mod p) coefficients. Obviously any almost upper triangular group is defined over $\mathbb{Z}/p\mathbb{Z}$.

Proposition 7.2. *Suppose that G is defined over $\mathbb{Z}/p\mathbb{Z}$. Then $F(\omega_G) \subset \omega_G$.*

Proof. To say that G is defined over $\mathbb{Z}/p\mathbb{Z}$ means that ω_G is generated by some collection $X \subset A_n$ of polynomials with integer (mod p) coefficients. Then we may identify ω_G as

$$\omega_G = \operatorname{span}_k \left\{ x^M f : f \in X, x^M \text{ is a monomial in } A_n \right\}$$

By the k-linearity of F,

$$F(\omega_G) = \operatorname{span}_k \left\{ F(x^M f) : f \in X, x^M \text{ is a monomial in } A_n \right\}$$

For each $x^M \in A_n$ and $f \in X$, $x^M f$ is a polynomial with integer coefficients; by Fermat's little theorem, i.e. since $z^p = z$ for any integer z, and by the identity $(x+y)^p = x^p + y^p$, in fact $F(x^M f) = (x^M f)^p$. Each $x^M f$ itself is an element of ω_G, whence $F(x^M f) = (x^M f)^p$ is an element of ω_G^p, hence ω_G, since ω_G is an ideal. Thus each spanning element of $F(\omega_G)$ is an element of ω_G, which proves that $F(\omega_G) \subset \omega_G$. □

For the remainder of this section we assume that G is an almost upper triangular group over a field k of characteristic $p > 0$ with defining equations

$$\omega_G = (x_{ij} : (i,j) \in S)$$

Definition 7.3. For an almost upper triangular group G over a field k of characteristic $p > 0$, The **ideal of impediments** for G is the ideal I_G of $\operatorname{Dist}(G)$ generated by all impediments of the first and second kind which belong to $\operatorname{Dist}(G)$.

The m^{th} **Lie algebra** of G is

$$\operatorname{Lie}_m(G) = \operatorname{Lie}_m(U_n) \cap \operatorname{Dist}(G)$$

where $\operatorname{Lie}_m(U_n)$ is the m^{th} Lie algebra of $\operatorname{Dist}(U_n)$.

Again, the term "Lie algebra" is abusive here; $\operatorname{Lie}_m(G) \subset \operatorname{Dist}(G)$ is by no means generally a Lie algebra for $m > 0$.

Proposition 7.4. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$, with defining equations*

$$\omega_G = (x_{ij} : (i,j) \in S)$$

Then $Lie_m(G)$ can be identified as

$$Lie_m(G) = span_k(\alpha(p^m \varepsilon_{ij}) : (i,j) \notin S)$$

Proof. ω_G can be identified as

$$\begin{aligned}\omega_G &= \mathrm{span}_k(x^M x_{ij} : (i,j) \in S, M \in u_n(\mathbb{N})) \\ &= \mathrm{span}_k(x^{M+\varepsilon_{ij}} : (i,j) \in S, M \in u_n(\mathbb{N}))\end{aligned}$$

and $\mathrm{Dist}(G) \subset \mathrm{Dist}(U_n)$ can be identified as

$$\mathrm{Dist}(G) = \{\mu \in \mathrm{Dist}(U_n) : \mu(\omega_G) = 0\}$$

If we write $\mu = \sum_M c_M \alpha(M)$, then by the description of ω_G above, we have that $\mu \in \mathrm{Dist}(G)$ if and only if $c_M = 0$ whenever $m_{ij} \neq 0$ for some $(i,j) \in S$. This gives

$$\mathrm{Dist}(G) = \mathrm{span}_k(\alpha(M) : m_{ij} = 0 \quad \text{for all} \quad (i,j) \in S)$$

and in particular

$$\mathrm{Lie}_m(G) = \mathrm{span}_k(\alpha(p^m \varepsilon_{ij}) : (i,j) \notin S)$$

as claimed. □

We now prove that, for almost upper triangular groups G, our results in proposition 6.10 for $\mathrm{Dist}(U_n)/I$ exactly carry over to results for $\mathrm{Dist}(G)/I_G$.

Proposition 7.5. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$.*

For $m \in \mathbb{N}$,

$$F^{*m}(Lie_m(G)) = Lie(G)$$

and in fact, F^{*m} defines an isomorphism of k-vector spaces between $Lie_m(G)$ and $Lie(G)$. Also

$$(F^{*m})^{-1}(Lie(G)) = Lie_m(G) \oplus ker(F^{*m})$$

and

$$Lie_m(G) = (F^{*m})^{-1}(Lie(G)) \cap Dist_{p^m}(G)$$

Proof. Given the description of $\mathrm{Lie}_m(G)$ in proposition 7.4, the proof is identical to that of proposition 5.5 for the group $G = U_n$.

□

We can now prove theorem 1.1.

Proposition 7.6. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$.*

For all $m \geq 0$, $Lie_m(G)$ embeds in $Dist(G)/I_G$, and under this identification, is in fact a Lie algebra, isomorphic to $Lie(G)$. $Lie_m(G)$ and $Lie_l(G)$ commute inside $Dist(G)/I_G$ for $m \neq l$, whence

$$Lie_0(G) \oplus Lie_1(G) \oplus \ldots \simeq Lie(G) \oplus Lie(G) \oplus \ldots$$

embeds in $Dist(G)/I_G$. Finally, $Lie_0(G) \oplus Lie_1(G) \oplus \ldots$ generates $Dist(G)/I_G$ as a k-algebra.

Proof. Note that, since $\mathrm{Lie}_m(G) \subset \mathrm{Lie}_m(U_n)$ and $I_G \subset I$, $\mathrm{Lie}_m(G) \cap I_G = 0$. Given the description of $\mathrm{Lie}_m(G)$ in proposition 7.4, the proof is again identical to that of proposition 6.10 for the group $G = U_n$.

□

We can now prove theorem 1.2.

Proposition 7.7. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$.*

Let V be a d-dimensional k-vector space, $\Phi : G \to Aut(V)$ a group representation, $\overline{\Phi} : Dist(G) \to End(V)$ the induced algebra representation. Suppose $\overline{\Phi}$ kills the ideal of impediments I_G for G. For $m \in \mathbb{N}$, define a mapping $\phi_m : Lie(G) \to Lie(GL_d)$ as

$$\phi_m(x) = \overline{\Phi}\left(\left(F^{*m}|_{Lie_m(G)}\right)^{-1}(x)\right)$$

Then each ϕ_m is a Lie algebra representation, ϕ_m commutes with ϕ_l for $m \neq l$, and $\phi_m(x)^p = 0$ for all $m \in \mathbb{N}$ and $x \in Lie(G)$. Also, the maps $\phi_m : m \in \mathbb{N}$ completely determine the representation Φ.

Proof. By proposition 7.5, $F^{*m}|_{\mathrm{Lie}_m(G)}$ defines a linear isomorphism between $\mathrm{Lie}_m(G)$ and $\mathrm{Lie}(G)$; this shows that the map ϕ_m is well-defined (where we identify of course $\mathrm{Lie}(GL_d) = \mathrm{End}(V)$).

That each ϕ_m is a Lie algebra map can be seen as follows. Let $x_1, x_2 \in \mathrm{Lie}(G)$. We would like to see that

$$\phi_m([x_1, x_2]) = [\phi_m(x_1), \phi_m(x_2)]$$

Write $\phi_m(x_1) = \overline{\Phi}(y_1), \phi_m(x_2) = \overline{\Phi}(y_2)$ where $y_1, y_2 \in \mathrm{Lie}_m(G)$. Then we have, since $\overline{\Phi}$ is an algebra map and so necessarily preserves brackets

$$\begin{aligned}[\phi_m(x_1), \phi_m(x_2)] &= [\overline{\Phi}(y_1), \overline{\Phi}(y_2)] \\ &= \overline{\Phi}([y_1, y_2])\end{aligned}$$

Since $\overline{\Phi}$ factors through I_G, and since $\mathrm{Lie}_m(G)$ is a Lie algebra in $\mathrm{Dist}(G)/I_G$, necessarily $[y_1, y_2] \in \mathrm{Lie}_m(G)$ inside $\mathrm{Dist}(G)/I_G$, whence $\phi([x_1, x_2]) = \overline{\Phi}([y_1, y_2])$; thus ϕ_m is a Lie algebra map.

That ϕ_m commutes with ϕ_l for $m \neq l$ can be seen as follows. Let $x_1, x_2 \in \mathrm{Lie}(G)$. We would like to see that $\phi_m(x_1)$ and $\phi_l(x_2)$ commute. As before, let $y_1 \in \mathrm{Lie}_m(G)$, $y_2 \in \mathrm{Lie}_l(G)$ such that $\phi_m(x_1) = \overline{\Phi}(y_1)$, $\phi_l(x_2) = \overline{\Phi}(y_2)$, so that

$$[\phi_m(x_1), \phi_l(x_2)] = [\overline{\Phi}(y_1), \overline{\Phi}(y_2)]$$

But since $\overline{\Phi}$ factors through I_G, and since $\mathrm{Lie}_m(G)$ and $\mathrm{Lie}_l(G)$ commute inside $\mathrm{Dist}(G)/I_G$, $\overline{\Phi}(y_1)$ and $\overline{\Phi}(y_2)$ commute, forcing $[\overline{\Phi}(y_1), \overline{\Phi}(y_2)] = 0$, forcing $[\phi_m(x_1), \phi_l(x_2)] = 0$.

To see that $\phi_m(x)^p = 0$, note that, for $x \in \mathrm{Lie}(G)$, $\left(F^{*m}|_{\mathrm{Lie}_m(G)}\right)^{-1}(x) \in \mathrm{Lie}_m(G) \subset \mathrm{Lie}_m(U_n)$, so by proposition 5.2, $\left[\left(F^{*m}|_{\mathrm{Lie}_m(G)}\right)^{-1}(x)\right]^p = 0$, whence also $\phi_m(x)^p = \overline{\Phi}\left(\left(F^{*m}|_{\mathrm{Lie}_m(G)}\right)^{-1}(x)\right)^p = 0$, since $\overline{\Phi}$ is an algebra map.

To prove the last claim, let $y \in \mathrm{Lie}_m(G)$. Then clearly $\overline{\Phi}(y) = \phi_m(F^{*m}(y))$, showing that the image of y under $\overline{\Phi}$ can be recovered from ϕ_m. This holds for all

y in all $\mathrm{Lie}_m(G)$, and since $\mathrm{Lie}_m(G) : m \geq 0$ generates $\mathrm{Dist}(G)$, the entire algebra representation $\overline{\Phi}$ is determined by the ϕ_m. Then so also is the group representation Φ, by proposition 4.2.

□

To prove theorem 1.3, for the remainder of this section we will move to a more concrete setting. We will identify $\mathrm{Lie}(U_n)$ with the space of all strictly upper triangular $n \times n$ matrices over k, and if G is an algebraic subgroup of U_n, likewise $\mathrm{Lie}(G)$ will be a Lie-subspace of $\mathrm{Lie}(U_n)$. Note that, if $\mathrm{Nat} : U_n \to \mathrm{Aut}(k^n)$ is the natural representation of U_n, then the corresponding algebra map $\overline{\mathrm{Nat}} : \mathrm{Dist}(U_n) \to \mathrm{End}(k^n)$ exactly maps $\mathrm{Lie}(U_n) \subset \mathrm{Dist}(U_n)$ onto this matrix Lie algebra, so this identification is justified; similar remarks hold for G.

For a nilpotent $n \times n$ square matrix (or linear transformation) X with entries in a field k, define

$$\exp(X) = \sum_{i=0}^{n-1} \frac{X^i}{i!}$$

and for an $n \times n$ unipotent matrix (or linear transformation) g, say $g = 1 + X$ where X is nilpotent, define

$$\log(g) = \sum_{i=1}^{n-1} \frac{(-1)^{i+1}}{i} (g-1)^i$$

Note that these definitions only make sense when either $\mathrm{char}(k) = 0$ or $\mathrm{char}(k) = p \geq n$ (otherwise we see rational numbers with denominators divisible by $i \geq p$, which is nonsensical when $0 < \mathrm{char}(k) < n$). It is of course the case that, so long as these are defined, $\exp(\log(g)) = g$ and $\log(\exp(X)) = X$, whence log and exp define bijective correspondences between U_n and $\mathrm{Lie}(U_n)$.

The proof of the following proposition is due to Mikhail Borovoi of Tel Aviv University, and is proven as corollary A.14.

Proposition 7.8. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$, and suppose that $\log : G \to Lie(U_n)$ and $exp : Lie(G) \to U_n$ are both defined. Then $\log(g) \in Lie(G)$ for $g \in G$, and $exp(X) \in G$ for $X \in Lie(G)$.*

Proof. See appendix A, specifically corollary A.14. □

The Baker-Campbell-Hausdorff formula (see for instance section 4 of [2]) tells us that, in characteristic zero, the formal infinite series $\log(e^x e^y)$ for associative but non-commutative x and y can be written

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

where each $P_m(x, y)$ is a homogeneous polynomial with rational coefficients, and in fact, each $P_m(x, y)$ is a (rational) linear combination of nested commutators of x and y. Difficulties arise in trying to apply the Baker-Campbell-Hausdorff formula in characteristic $p > 0$, chiefly because rational numbers with denominators divisible by p can occur in the expansions for $e^X, \log(1+X)$, etc., as well as in the expansion of the Baker-Campbell-Hausdorff formula itself. But, in situations where that is not an issue, it is just as useful, as the following lemma illustrates (which is an obvious adaptation of analogous well known results for characteristic 0 fields).

Lemma 7.9. *Let k be any field, let G be an algebraic group over k, and let $\phi : Lie(G) \to Lie(GL_d)$ be a Lie algebra homomorphism. Suppose that*

1. *The series $e^{\phi(X)}$ is defined for all $X \in Lie(G)$*
2. *The series $\log(g)$ is defined for all $g \in G$, and $\log(g)$ is a member of $Lie(G)$*
3. *For all $X, Y \in Lie(G)$, $\log(e^X e^Y)$ is defined, denoted as $BCH(X, Y)$. Further, $BCH(X, Y)$ can be written uniformly (the same for all X and Y in $Lie(G)$) as a finite linear combination of brackets of X and Y, brackets of brackets of X and Y, etc.*
4. *For all $X, Y \in Lie(G)$, $\log(e^{\phi(X)}e^{\phi(Y)})$ 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 G over k.

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

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

□

Let $P_m(x, y)$ be the homogeneous polynomials in the formula

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

In [6], [7] and [5], it is proven that, under a certain mapping ϕ, $\phi(P_m(x, y)) = P_m(x, y)$ for all m, and this expression gives us an explicit way of writing each $P_m(x, y)$ as a rational linear combination of nested commutators of x and y. We shall need

Lemma 7.10. *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. See proposition 4.6 of [2].

□

Lemma 7.11. *Let G be an almost upper triangular subgroup of U_n over a field k of characteristic $p > 0$, and suppose $p \geq n$. Let $\phi : Lie(G) \to Lie(GL_d)$ be a Lie algebra homomorphism such that $\phi(X)^p = 0$ for all $X \in Lie(G)$. Then the formula*

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

defines a d-dimensional representation of G.

Proof. We shall go through the checklist of proposition 7.9 to see that they are all satisfied.

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

(2): Every $g \in G$ 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. Also, $\log(g) \in \mathrm{Lie}(G)$ by proposition 7.8.

(3): $\mathrm{Lie}(U_n)$ is a nilpotent Lie algebra of order n, and each $X \in \mathrm{Lie}(U_n)$ is itself nilpotent; the same is thus true for $\mathrm{Lie}(G) \subset \mathrm{Lie}(U_n)$. Apply part 3. of lemma 7.10.

(4): As ϕ 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 lemma 7.10.

□

For ease of notation we introduce

Definition 7.12. Let G be an almost upper triangular group over a field k of characteristic $p > 0$. Let M be a matrix with entries in the representing Hopf algebra B of G. Then $M^{[p^m]}$ is the matrix gotten by applying the map $F^m : B \to B$ entry-wise to M.

Lemma 7.13. *Let G be an almost upper triangular subgroup of U_n over a field k of characteristic $p > 0$. Suppose also that $p \geq n$. Let $\phi_0, \ldots, \phi_m : Lie(G) \to Lie(GL_d)$ be a collection of commuting Lie algebra representations such that $\phi_i(X)^p = 0$ for all $0 \leq i \leq m$ and $X \in Lie(G)$. Then the formula*

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

defines a d-dimensional representation of G.

Proof. By lemma 7.11, for each $0 \leq r \leq m$, the formula $\Phi_r(g) = e^{\phi_r(\log(g))}$ defines a representation of G, and since $F^m : B \to B$ is a Hopf algebra map, clearly also does $\Phi_r(g)^{[p^r]} = e^{\phi_r(\log(g))^{[p^r]}}$. And since ϕ_r commutes with ϕ_s for $r \neq s$, clearly also do $\Phi_r(g)^{[p^r]}$ and $\Phi_s(h)^{[p^s]}$ commute when $r \neq s$ and $g, h \in G$. Any commuting product of representations of an algebraic group is again a representation; thus the formula

$$\begin{aligned}\Phi(g) &= \Phi_0(g)\Phi_1(g)^{[p]} \ldots \Phi_m(g)^{[p^m]} \\ &= e^{\phi_0(\log(g))} e^{\phi_1(\log(g))^{[p]}} \ldots e^{\phi_m(\log(g))^{[p^m]}}\end{aligned}$$

defines a representation of G. □

We can now prove theorem 1.3.

Theorem 7.14. *Let G be an almost upper triangular subgroup of U_n over a field k of characteristic $p > 0$. Suppose also that $p \geq n$.*

Let V be a finite dimensional k-vector space. Then any representation $\Phi : G \to Aut(V)$ whose induced representation $\overline{\Phi} : Dist(G) \to End(V)$ kills I_G can be written

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

where $\phi_0, \phi_1, \dots, \phi_m : Lie(G) \to Lie(GL_d)$ are the commuting Lie algebra representations given in proposition 7.7 for Φ.

Proof. Let

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

in order to distinguish it from Φ, since we have not yet proven that they are the same. By lemma 7.13, Ψ defines a representation of G. Thus, according to proposition 7.7, Ψ defines commuting Lie algebra homomorphisms $\psi_0, \psi_1, \dots, \psi_l : \mathrm{Lie}(G) \to \mathrm{Lie}(GL_d)$ defined by, for $x \in \mathrm{Lie}(G)$,

$$\psi_r(x) = \overline{\Psi}\left(\left(F^{*r}|_{\mathrm{Lie}_r(G)}\right)^{-1}(x)\right)$$

where $\overline{\Psi} : \mathrm{Dist}(G) \to \mathrm{End}(V)$ is the algebra representation associated to Ψ. By proposition 7.7, $\phi_0, \phi_1, \dots, \phi_m$ completely determine Φ, and likewise $\psi_0, \psi_1, \dots, \psi_l$ completely determine Ψ. Thus, if we can show that $\psi_i(X) = \phi_i(X)$ for all $i \geq 0$ and $X \in \mathrm{Lie}(G)$, we are done.

Let $X \in \mathrm{Lie}(G)$, and since $X \in \mathrm{Lie}(U_n)$, write $X = \sum_{ij} c_{ij}\varepsilon_{ij}$. For $0 \leq r \leq l$ let $\mu_X^r \in \mathrm{Lie}_r(G)$ be such that $\psi_r(x) = \overline{\Psi}(\mu_X^r)$; this means that, as an element of $\mathrm{Dist}(U_n)$, $\mu_X^r = \sum_{ij} c_{ij}\alpha(p^r\varepsilon_{ij})$. As in section 4 we have

$$\psi_r(X) = \overline{\Psi}(\mu_X^r) = \mu_X^r(\Psi(g))$$

That is, $\psi_r(X)$ is the matrix $\Psi(g)$ with μ_X^r applied to it entry-wise. For $g \in G \subset U_n$ write

$$g = 1 + \sum_{ij} x_{ij}\varepsilon_{ij}$$

Consider first the case of $\Psi(g)$ being defined by a single Lie algebra map, i.e. so that $m = 0$, i.e.

$$\Psi(g) = e^{\phi_0(\log(g))}$$

Then we can write

$$\begin{aligned}\log(g) &\stackrel{\mathrm{def}}{=} \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k}(g-1)^k \\ &= \sum_{1\leq i<j\leq n} x_{ij}\varepsilon_{ij} + \sum_{|M|\geq 2} x^M\chi(M)\end{aligned}$$

where $\chi(M)$ is a matrix for each strictly upper triangular M. As ϕ_0 is linear we have

$$\phi_0(\log(g)) = \sum_{1\leq i<j\leq n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\geq 2} x^M\phi_0(\chi(M))$$

Since $\log(g) \in \mathrm{Lie}(G)$, $\log(g)^p = 0$, since $p \geq n$. Also, since $\phi_m(\log(g))$ is defined to be

$$\phi_m(\log(g)) = \overline{\Phi}\left(\left(F^{*m}|_{\mathrm{Lie}_m(G)}\right)^{-1}(\log(g))\right) \in \overline{\Phi}\left(\mathrm{Lie}_m(G)\right)$$

we see that $\phi_m(\log(g))$ is the image under $\overline{\Phi}$ of the p-nilpotent element $\left(F^{*m}|_{\mathrm{Lie}_m(G)}\right)^{-1}(\log(g))$; thus we must have $\phi_m(\log(g))^p = 0$. Then

$$\begin{aligned}\Psi(g) &= e^{\phi_0(\log(g))} \\ &= 1 + \left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right) \\ &+ \left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right)^2 /2! + \dots \\ &+ \left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right)^{p-1} /(p-1)!\end{aligned}$$

Applying μ_X^0 to this expression gives

$$\begin{aligned}\mu_X^0(\Psi(g)) &= \mu_X^0(e^{\phi_0(\log(g))}) \\ &= \mu_X^0(1) + \mu_X^0\left[\left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right)\right] \\ &+ \mu_X^0\left[\left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right)^2 /2!\right] + \mu_X^0(\dots) \\ &+ \mu_X^0\left[\left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij}) + \sum_{|M|\ge 2} x^M \phi_0(\chi(M))\right)^{p-1} /(p-1)!\right]\end{aligned}$$

Since the effect of $\mu_X^0 = \sum c_{ij}\alpha(\varepsilon_{ij})$ is to kill any x^M when $|M| \ge 2$, and also $\mu_X^0(1) = 0$, we have

$$\begin{aligned}\mu_X^0(\Psi(g)) &= \mu_X^0\left(\sum_{1\le i<j\le n} x_{ij}\phi_0(\varepsilon_{ij})\right) \\ &= \sum_{1\le i<j\le n} \mu_X^0(x_{ij})\phi_0(\varepsilon_{ij}) \\ &= \sum_{1\le i<j\le n} c_{ij}\phi_0(\varepsilon_{ij})\end{aligned}$$

But this is exactly $\phi_0(X)$, whence $\psi_0(X) = \mu_X^0(\Psi(g)) = \phi_0(X)$, and the proposition is true in the case $m = 0$.

For $m > 0$, consider

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

where we are justified in running the summations only to $p-1$ since $\phi_i(\log(g))^p = 0$ for all i and g. Thus we see that the only contribution to the monomials $x_{ij}^{p^l}$ comes when all k_i are zero except $k_l = 1$. Thus in fact

$$\mu_X^l(\Psi(g)) = \mu_X^l\left(\phi_l(\log(g))^{[p^l]}\right)$$

and, as above, this is exactly

$$\mu_X^l(\Psi(g)) = \sum_{ij} c_{ij}\phi_l(\varepsilon_{ij})$$

But again, this is exactly $\phi_l(X)$, whence $\psi_l(X) = \phi_l(X)$, and the proposition is proved.

□

To finish this section, we give a characterization of morphisms between representations of G for which proposition 7.7 applies.

Proposition 7.15. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$.*

Let $\Phi : G \to Aut(V)$,$\Psi : G \to Aut(W)$ be representations of G on the finite dimensional vector spaces V and W such that $\overline{\Phi}$ and $\overline{\Psi}$ both kill I_G. Let $\phi_0, \phi_1, \ldots$, $\psi_0, \psi_1, \ldots$ be the Lie algebra representations given in proposition 7.7 for Φ and Ψ. Then the linear map $f : V \to W$ is a morphism of G-modules if and only if, for all $i \geq 0$ and $X \in Lie(G)$, the following commutes:

$$\begin{array}{ccc} V & \xrightarrow{\phi_i(X)} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\psi_i(X)]{} & W \end{array}$$

Proof. Let $X \in \mathrm{Lie}(G)$, and as in the proof of theorem 7.14, for $r \in \mathbb{N}$, let $\mu_X^r \in \mathrm{Lie}_r(G)$ be such that $F^{*r}(\mu_X^r) = X$. This means that

$$\phi_r(X) = \overline{\Phi}(\mu_X^r) = \mu_X^r(\Phi(g))$$
$$\psi_r(X) = \overline{\Psi}(\mu_X^r) = \mu_X^r(\Psi(g))$$

To say that f is a morphism of G-modules is to say that the following commutes for all $g \in G$:

$$\begin{array}{ccc} V & \xrightarrow{\Phi(g)} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\Psi(g)]{} & W \end{array}$$

Then since this diagram commutes "generically" for all $g \in G$, so also it commutes upon applying μ_X^r:

$$\begin{array}{ccc} V & \xrightarrow{\mu_X^r(\Phi(g))} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\mu_X^r(\Psi(g))]{} & W \end{array}$$

which is of course the diagram

$$\begin{array}{ccc} V & \xrightarrow{\phi_r(X)} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\psi_r(X)]{} & W \end{array}$$

and the forward implication is proved. For the reverse, as in the proof of proposition 4.2, pick a basis $\{h + \omega_G : h \in S\}$ for B, where $S \subset A_n$, and let $\mu_h \in \mathrm{Dist}(G)$ send $h + \omega_G$ to 1, all other $h' + \omega_G$ to zero. Then we can write

$$\Phi(g) = \sum_{h \in S} (h + \omega_G)\mu_h(\Phi(g))$$
$$\Psi(g) = \sum_{h \in S} (h + \omega_G)\mu_h(\Psi(g))$$

The set $\{\mu_X^r : X \in \mathrm{Lie}(G), r \in \mathbb{N}\}$ generates all of $\mathrm{Dist}(G)$ (proposition 5.2 and definition 7.3), and

$$\begin{array}{ccc} V & \xrightarrow{\mu_X^r(\Phi(g))} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\mu_X^r(\Psi(g))]{} & W \end{array}$$

always commutes; this implies that

$$\begin{array}{ccc} V & \xrightarrow{\mu_h(\Phi(g))} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\mu_h(\Psi(g))]{} & W \end{array}$$

commutes for all $h \in S$, which implies that

$$\begin{array}{ccc} V & \xrightarrow{\Phi(g)} & V \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ W & \xrightarrow[\Psi(g)]{} & W \end{array}$$

commutes, and the proposition is proved.

□

8. Example

Let G be an almost upper triangular group over a field k of characteristic $p > 0$. We would like to know if the algebra representation $\overline{\Phi} : \mathrm{Dist}(G) \to \mathrm{End}(V)$ associated to a given representation $\Phi : G \to \mathrm{Aut}(V)$ in fact kills I_G, i.e. if any of our main theorems apply to it. This tends to be easy to do in practice. For ease of viewing, let us write elements g of the Heisenberg group $G = U_3$ as

$$g = \begin{pmatrix} 1 & x & z \\ & 1 & y \\ & & 1 \end{pmatrix}$$

Consider the following representation of G in characteristic $p = 2$:

$$\Phi(g) = \begin{pmatrix} 1 & x & y & z & x^2 & xy & xz & y^2 & yz & z^2 \\ & 1 & 0 & y & 0 & y & z+xy & 0 & y^2 & 0 \\ & & 1 & 0 & 0 & x & 0 & 0 & z & 0 \\ & & & 1 & 0 & 0 & x & 0 & y & 0 \\ & & & & 1 & 0 & y & 0 & 0 & y^2 \\ & & & & & 1 & 0 & 0 & y & 0 \\ & & & & & & 1 & 0 & 0 & 0 \\ & & & & & & & 1 & 0 & 0 \\ & & & & & & & & 1 & 0 \\ & & & & & & & & & 1 \end{pmatrix}$$

(This is a certain finite dimensional slice of the regular representation of G.) One could go through all possible impediments to see if they are killed under $\overline{\Phi}$, but the easiest thing to do is to just find the images of each $\mathrm{Lie}_m(G)$ in $\mathrm{End}(V)$, see if they do in fact map to Lie algebras (with bracket given by what we expect them to be, i.e. mimicking that of $\mathrm{Lie}(G)$), and to see if they commute.

Let us illustrate. For each m, let

$$\begin{aligned} X_m &= \alpha(p^m \varepsilon_{12}) \\ Y_m &= \alpha(p^m \varepsilon_{23}) \\ Z_m &= \alpha(p^m \varepsilon_{13}) \end{aligned}$$

so that $\{X_m, Y_m, Z_m\}$ is a basis for $\mathrm{Lie}_m(G)$ inside $\mathrm{Dist}(G)$. As in section 4, the image of, for example Y_m under $\overline{\Phi}$ is simply the "matrix of coefficients" for the

monomial y^{p^m} in this representation; for instance,

$$\overline{\Phi}(Y_1) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & & & & 0 & 0 & 0 & 0 & 0 & 1 \\ & & & & & 0 & 0 & 0 & 0 & 0 \\ & & & & & & 0 & 0 & 0 & 0 \\ & & & & & & & 0 & 0 & 0 \\ & & & & & & & & 0 & 0 \\ & & & & & & & & & 0 \end{pmatrix}$$

since this is the matrix of coefficients for $y^{p^1} = y^2$. One can check by hand that, for instance, $[\overline{\Phi}(X_1), \overline{\Phi}(Y_1)] = \overline{\Phi}(Z_1)$, and that $[\overline{\Phi}(X_1), \overline{\Phi}(Z_1)] = [\overline{\Phi}(Y_1), \overline{\Phi}(Z_1)] = 0$, giving us exactly the Lie bracket we would like to see on $\overline{\Phi}(\mathrm{Lie}_1(G))$. However, we also compute that $[\overline{\Phi}(X_0), \overline{\Phi}(Y_1)] \neq 0$, whence the images of $\mathrm{Lie}_0(G)$ and $\mathrm{Lie}_1(G)$ in $\mathrm{End}(V)$ do *not* commute, whence the ideal of impediments I_G for G is *not* killed by $\overline{\Phi}$, whence none of our main theorems apply to this particular representation (in characteristic $p = 2$ that is; for characteristic 3 and higher, it is trivially the case that $\overline{\Phi}(I_G) = 0$).

On the other hand, suppose that we do find that $\overline{\Phi}(I_G) = 0$ for a given representation Φ, and suppose further that $p \geq n = 3$. One first computes generically that

$$\log(g) = \begin{pmatrix} 0 & x & z - xy/2 \\ & 0 & y \\ & & 0 \end{pmatrix}$$

One next defines Lie algebra representations $\phi_m : \mathrm{Lie}(G) \to \mathrm{Lie}(GL_d)$ via, for example,

$$\phi_m(X) = \overline{\Phi}(X_m)$$

where

$$X = \begin{pmatrix} 0 & 1 & 0 \\ & 0 & 0 \\ & & 0 \end{pmatrix}$$

and similarly for Y and Z. Using that each ϕ_m is linear, we have that

$$\begin{aligned} \phi_m(\log(g)) &= \phi_m(xX + yY + (z - xy/2)Z) \\ &= x\phi_m(X) + y\phi_m(Y) + (z - xy/2)\phi_m(Z) \end{aligned}$$

Then one has, by theorem 1.3, that this representation necessarily takes the form

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

9. Further directions

9.1. **Extending (or not) the main theorems to other unipotent groups.** Here are the author's thoughts on possibly how (or if) these results might be extended to other unipotent groups.

First, let us discuss exactly why the apparently special class of *almost upper triangular* groups was chosen as the topic of this paper. In short, it is because it gives us proposition 7.4. There we are automatically given that, as defined, $\mathrm{Lie}_m(G)$ is linearly isomorphic to $\mathrm{Lie}(G)$ (just as it is for $G = U_n$), and that, upon modding

out by the ideal of impediments for G (directly derived from those for U_n), $\mathrm{Lie}_m(G)$ in fact becomes Lie algebra isomorphic to $\mathrm{Lie}(G)$, just as it is for U_n, and the rest of our results follow more-or-less straightforwardly.

Let us discuss then why this line of reasoning fails when G is *not* almost upper triangular. Take for example the group

$$G = \left\{ \begin{pmatrix} 1 & x & x^2 \\ & 1 & 2x \\ & & 1 \end{pmatrix} \right\}$$

which happens to be isomorphic to the additive group $G_a = U_2$. Our main theorems of course apply to this group, because G is (isomorphic to) an almost upper triangular group. But because of the specific embedding $G \subset U_3$ used here, our arguments in this paper do not immediately lend themselves to recovering our main theorems for this group.

Let us illustrate. The defining equations of G (as a subgroup of U_3) are

$$\omega_G = (2x_{12} - x_{23}, x_{12}^2 - x_{13})$$

Using arguments similar to that in the proof of proposition 7.4, we can show that, as a sub-algebra of $\mathrm{Dist}(U_n)$,

$$\mathrm{Dist}(G) = \left\{ \sum_M c_M \alpha(M) : 2c_{M+\varepsilon_{12}} - c_{M+\varepsilon_{23}} = 0, c_{M+2\varepsilon_{12}} - c_{M+\varepsilon_{13}} = 0 \quad \forall M \in u_3(\mathbb{N}) \right\}$$

from whence we can show

$$\mathrm{Lie}(G) = \mathrm{span}_k(\alpha(\varepsilon_{12}) + 2\alpha(\varepsilon_{23}))$$

But in fact, for $\mathrm{char}(k) = p \geq 3$ and $m > 0$,

$$\mathrm{Lie}_m(G) = 0$$

To see this, let $\mu = c_{12}\alpha(p^m\varepsilon_{12}) + c_{13}\alpha(p^m\varepsilon_{13}) + c_{23}\alpha(p^m\varepsilon_{23})$ be an arbitrary element of $\mathrm{Lie}_m(U_3)$. In order for μ to be an element $\mathrm{Lie}_m(G)$, we must have $\mu(\omega_G) = 0$. In particular, we must have

$$\begin{aligned} \mu((2x_{12} - x_{23})x_{23}^{p^m-1}) &= -c_{23} = 0 \\ \mu((2x_{12} - x_{23})x_{12}^{p^m-1}) &= 2c_{12} = 0 \\ \mu((x_{12}^2 - x_{13})x_{13}^{p^m-1}) &= -c_{13} = 0 \end{aligned}$$

forcing $\mu = 0$, showing that proposition 7.4 definitely does *not* apply here, and thus making $\mathrm{Lie}_m(G)$ obviously *not* isomorphic to $\mathrm{Lie}(G)$, making all of the proofs for our main theorems inapplicable to this group. The culprit here is clear; it is because the defining polynomials of G involved more than one x_{ij} variable at a time, thus the reason we restrict to almost upper triangular groups.

An obvious thing to try at this point would be to alter definition 7.3, giving us different notions of $\mathrm{Lie}_m(G)$ and I_G, and in such a way that theorem 1.1 can be recovered for groups such as G. But it is not clear how to do this. To further illustrate, let $B = A_3/\omega_G$ be the representing Hopf algebra of G, and consider the isomorphism $G \simeq G_a$, which is given by the Hopf algebra map $\Omega : B \to A_2$ defined by

$$\Omega : x_{12} \mapsto x_{12}, x_{23} \mapsto 2x_{12}, x_{13} \mapsto x_{12}^2$$

This induces an isomorphism $\overline{\Omega} : \mathrm{Dist}(G_a) \to \mathrm{Dist}(G)$. $\mathrm{Lie}_m(G_a)$ is simply $\mathrm{span}_k(\alpha(p^m \varepsilon_{12}))$, and we can ask what the image of $\mathrm{Lie}_m(G_a)$ is under $\overline{\Omega}$, which yields

$$\overline{\Omega}(\mathrm{Lie}_m(G_a)) = \mathrm{span}_k(\mu)$$

where $\mu = \sum_M c_M \alpha(M)$ is the distribution whose coefficients c_M satisfy

$$c_M = 2^{m_{23}} \text{ if } m_{12} + m_{23} + 2m_{13} = p^m$$

and 0 otherwise. Thus, to extend our theorems to G, one expects to be able to recover in some natural way this particular distribution as a basis for $\mathrm{Lie}_m(G)$ *independent of* our isomorphism $G_a \simeq G$. But, again, we do not see how to do this.

The most persuasive argument the author can muster for the possibility of extending our theorems to other unipotent groups is Lucas' theorem (proposition 3.11) which, while simple, seems to be a profound number-theoretic fact about multinomial coefficients modulo a prime (and without which none of the results of this paper would be possible). For any unipotent G, we always have that $\mathrm{Dist}(G) \subset \mathrm{Dist}(U_n)$, and by proposition 3.3, multinomial coefficients are always intrinsic to multiplication in $\mathrm{Dist}(G)$. Perhaps this is enough to replace definitions 7.3, 6.3, and 6.4 with ones that work for more general unipotent G; perhaps not.

9.2. **Isomorphism classes of almost upper triangular groups.** We have little idea at this point of how many (up to isomorphism) almost upper triangular subgroups of U_n there are for a given n, but at first glance there seem to be quite a lot. If $G \subset U_n$ is almost upper triangular with defining polynomials

$$\omega_G = (x_{ij} : (i,j) \in S)$$

then there is a simple criterion on S for ω_G to in fact form a group, just by looking at the group closure requirement. If $g = 1 + \sum_{ij} a_{ij}\varepsilon_{ij}$, $h = 1 + \sum_{ij} b_{ij}\varepsilon_{ij}$ are two elements of G, their product is

$$\begin{aligned} gh &= 1 + \sum_{ij}(a_{ij} + b_{ij})\varepsilon_{ij} + \sum_{(r,s),(t,u)} a_{rs}b_{tu}\varepsilon_{rs}\varepsilon_{tu} \\ &= 1 + \sum_{ij}(a_{ij} + b_{ij})\varepsilon_{ij} + \sum_{(r,s),(t,u)} a_{rs}b_{tu}\delta_{st}\varepsilon_{ru} \\ &= 1 + \sum_{ij}(a_{ij} + b_{ij})\varepsilon_{ij} + \sum_{r,k,u} a_{rk}b_{ku}\varepsilon_{ru} \end{aligned}$$

We require that, whenever $(i,j) \in S$ (meaning that always $a_{ij} = b_{ij} = 0$), that the $(i,j)^{\text{th}}$ entry of gh also be always zero. The $(i,j)^{\text{th}}$ entries of 1 and $\sum_{i,j}(a_{ij}+b_{ij})\varepsilon_{ij}$ are automatically zero, so we need only consider the $(i,j)^{\text{th}}$ entry of

$$\sum_{r,k,u} a_{rk}b_{ku}\varepsilon_{ru}$$

which is

$$\sum_{k=i+1}^{j-1} a_{ik}b_{kj}$$

Thus we require that at least one of a_{ik} or b_{kj} always be zero for every k in this summation, which gives

Proposition 9.1. *Over any field k, the subset S of $\{(i,j) : 1 \leq i < j \leq n\}$ defines an almost upper triangular group if and only if S satisfies the following property: whenever $(i,j) \in S$, then for all $i < k < j$, either $(i,k) \in S$ or $(k,j) \in S$.*

Let $S(n)$ = number of subsets of $\{(i,j) : 1 \leq i < j \leq n\}$ having the property given in proposition 9.1. A brute force search in Maple gives the following first several values of $S(n)$:

$$S(1) = 1$$
$$S(2) = 2$$
$$S(3) = 7$$
$$S(4) = 40$$
$$S(5) = 357$$
$$S(6) = 4824$$
$$S(7) = 96428$$

But again, this is without identifying isomorphism classes, and upon doing so, these values would no doubt be vastly smaller (but by what factor we could not now predict). In any case, with some combinatorics, the classification problem for almost upper triangular groups appears not to be out of reach.

9.3. **Conditions under which the theorems hold.** All of the results of this paper depend on the condition that G be almost upper triangular, and the condition that the ideal of impediments for G dies under a given representation. The author himself is wondering exactly how often and under what conditions I_G might die for a given representation of such G. But, we know of at least one instance where this must be so.

Let G be an almost upper triangular subgroup of U_n, generated by the polynomials

$$(x_{ij} : (i,j) \in S)$$

Proposition 9.2. *Let G be an almost upper triangular group over a field k of characteristic $p > 0$. Let $\Phi : G \to Aut(V)$ be a finite dimensional representation of G on a d-dimensional vector space V, and suppose that $p \geq 2d$. Then $\overline{\Phi} : Dist(G) \to End(V)$ kills I_G.*

Proof. Recall from definition 6.1 that impediments of the first and second kind in $\mathrm{Dist}(U_n)$ are those distributions of one of the following forms:

(1) $\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{st} + k\varepsilon_{rt})$
(2) $-\sum_{k=1}^{p^m-1} \alpha((p^m-k)\varepsilon_{rs} + (p^m-k)\varepsilon_{tr} + k\varepsilon_{ts})$
(3) $\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{st} + k\varepsilon_{rt})$
(4) $-\sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{tr} + k\varepsilon_{ts})$

for appropriate choices of m, l, r, s, t. We claim that the ones that belong to $\mathrm{Dist}(G)$, i.e. those that generate I_G, are precisely those, in (1) and (3), where $(r,s), (s,t) \notin S$, and in (2) and (4), where $(r,s), (t,r) \notin S$. We argue this for (1); the proof for (2),(3) and (4) is similar.

As in the proof of proposition 7.4 we have

$$\mathrm{Dist}(G) = \mathrm{span}_k(\alpha(M) : m_{ij} = 0 \quad \text{for all} \quad (i,j) \in S)$$

If $(r,s),(s,t) \notin S$, then by proposition 9.1, necessarily $(r,t) \notin S$. This forces every term in the summation in (1) to belong to $\mathrm{Dist}(G)$, forcing (1) to belong to $\mathrm{Dist}(G)$. Conversely, if say $(r,s) \in S$, then at least one term in the summation ($k=1$ for instance) is not of the form given by our description of $\mathrm{Dist}(G)$, and hence this sum cannot belong to $\mathrm{Dist}(G)$.

We now argue that $\overline{\Phi}(I_G) = 0$; it is enough to prove that $\overline{\Phi}$ sends each generator of I_G to zero. We prove this for generators of the form (3); the proof for (1), (2) and (4) is similar. Let

$$\begin{aligned} x &= \sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs} + (p^l-k)\varepsilon_{st} + k\varepsilon_{rt}) \\ &= \sum_{k=1}^{\min(p^m,p^l)} \alpha((p^m-k)\varepsilon_{rs})\alpha((p^l-k)\varepsilon_{st})\alpha(k\varepsilon_{rt}) \end{aligned}$$

be such a distribution, where say $m > l$. Note that, by our description of $\mathrm{Dist}(G)$, each of the individual terms $\alpha((p^m-k)\varepsilon_{rs}), \alpha((p^l-k)\varepsilon_{st}), \alpha(k\varepsilon_{rt})$ also belong to $\mathrm{Dist}(G)$, whence $\overline{\Phi}$ is defined for these terms. Then we have

$$\overline{\Phi}(x) = \sum_{k=1}^{\min(p^m,p^l)} \overline{\Phi}(\alpha((p^m-k)\varepsilon_{rs}))\overline{\Phi}(\alpha((p^l-k)\varepsilon_{st}))\overline{\Phi}(\alpha(k\varepsilon_{rt}))$$

We claim that, for all k in this summation, at least one of $\overline{\Phi}(\alpha((p^m-k)\varepsilon_{rs}))$ or $\overline{\Phi}(\alpha(k\varepsilon_{rt}))$ must be zero, forcing $\overline{\Phi}(x) = 0$. Write, in p-ary notation,

$$\begin{aligned} p^m - k &= a_0 + a_1 p + \cdots + a_m p^m \\ k &= b_1 + b_1 p + \cdots + b_m p^m \end{aligned}$$

Clearly $(p^m-k)+k$ carries since $p^m - k$ borrows. Thus, for some i, $a_i + b_i \geq p$, and so at least one of a_i or b_i must be greater than or equal to $p/2 \geq d$; say $a_i \geq d$. By proposition 3.12 we can write

$$\alpha((p^m-k)\varepsilon_{rs}) = \Gamma(p^m-k)^{-1}\alpha(\varepsilon_{rs})^{a_0}\alpha(p\varepsilon_{rs})^{a_1}\ldots\alpha(p^m\varepsilon_{rs})^{a_m}$$

and likewise

$$\overline{\Phi}(\alpha((p^m-k)\varepsilon_{rs})) = \Gamma(p^m-k)^{-1}\overline{\Phi}(\alpha(\varepsilon_{rs}))^{a_0}\ldots\overline{\Phi}(\alpha(p^i\varepsilon_{rs}))^{a_i}\ldots\overline{\Phi}(\alpha(p^m\varepsilon_{rs}))^{a_m}$$

But by proposition 4.1, $\overline{\Phi}(\alpha(p^i\varepsilon_{rs}))$ is a nilpotent linear transformation, of nilpotent order necessarily no greater than $d \leq a_i$, whence $\overline{\Phi}(\alpha(p_i\varepsilon_{rs}))^{a_i} = 0$. This forces $\overline{\Phi}(\alpha((p^m-k)\varepsilon_{rs})) = 0$, forcing $\overline{\Phi}(x) = 0$, forcing $\overline{\Phi}(I_G) = 0$, proving the proposition.

□

Whether or not there are other useful criteria on a representation which force $\overline{\Phi}(I_G) = 0$ is an open question.

9.4. **The classification problem.** We do not at all suggest that the work done here might lead to a full classification of representations of almost upper triangular groups; these results only apply to the case when I_G vanishes, and in many (most?) instances, this is simply not the case. Nonetheless, when in fact I_G vanishes, we have

Proposition 9.3. *Let G be an almost upper triangular subgroup of U_n over a field k of characteristic $p > 0$, and suppose $char(k) = p \geq n$. Then, for the case of d-dimensional representations of G which kill I_G, the classification problem for such representations reduces to:*

"Find all collections of commuting Lie algebra representations $\phi_0, \phi_1, \ldots, \phi_m$: $Lie(G) \to Lie(GL_d)$ such that $\phi_i(X)^p = 0$ for all $i \geq 0$ and $X \in Lie(G)$, up to a simultaneous base change of k^d."

Proof. By proposition 7.7 and lemma 7.13, there is a bijective correspondence between such representations and such collections of Lie algebra homomorphisms. Two such collections $\phi_0, \ldots, \phi_m$ and $\psi_0, \ldots, \psi_m$ define isomorphic representations if and only if they are simultaneously conjugate in $\mathrm{End}(k^d)$, by proposition 7.15. □

We have no idea how hard this problem might be.

9.5. **The ideals I and J.** Other than to specify its generators, we do not have a constructive description for the ideal of impediments I for $\mathrm{Dist}(U_n)$; for instance, we have not identified a basis for I, and we are not sure if I and J (the carrying ideal) are equal. This was investigated at some length in Python (see [11] for the implementation), and conjectures were formed (and disproved). However, we can at least say that I is none of the following (at the time seemingly plausible) candidates:

Proposition 9.4. *Let k have characteristic $p > 0$. None of the following are ideals of $Dist(U_n)$:*

1. $span_k(\alpha(M) : M$ *carries in one of its rows or its columns*$)$
2. $span_k(\alpha(M) :$ *some two entries of M in the same row or column carries*$)$
3. $span_k(\alpha(M) :$ *some two entries of M carry*$)$

Proof. The author disproved all three of these with counterexamples which he will not bother the reader with. □

Lastly, concerning the carrying ideal J of $\mathrm{Dist}(U_n)$. The algebra $\mathrm{Dist}(U_n)/J$ can of course be identified as

$$\mathrm{span}_k(\alpha(M) + J : \sum_{ij} m_{ij} \text{ does } \textit{not} \text{ carry })$$

with the multiplication law inherited from $\mathrm{Dist}(U_n)$

$$(\alpha(M) + J)(\alpha(N) + J) = \alpha(M)\alpha(N) + J$$

This is an interesting algebra in its own right, and one can ask if a tidy multiplication law analogous to that for $\mathrm{Dist}(U_n)$ in proposition 3.3 can be given (possibly involving remainders mod p or some such thing). We have no answer to this question at present.

10. Acknowledgements

The author wishes to thank Mikhail Borovoi of Tel Aviv University for appendix A, in particular his proof of proposition 7.8.

Appendix A. Exponential and logarithm in almost upper triangular unipotent groups

by Mikhail Borovoi

Let U_n denote the algebraic group of all $n \times n$ upper triangular matrices with units on the main diagonal over a field k of positive characteristic $p \geq n$ or of characteristic 0. In this appendix we show that if $G \subseteq U_n$ is an almost upper triangular k-subgroup (see the Introduction or Definition A.11 below), then $\log(g) \in \mathrm{Lie}(G)$ for all $g \in G(k)$, and $\exp(X) \in G(k)$ for all $X \in \mathrm{Lie}(G)$. Note that the similar assertions are not true for an arbitrary k-subgroup of U_n over a field k of positive characteristic; see [12] for a counter-example.

Let $\mathfrak{u}_n$ denote the tangent space of U_n at 1, that is, the space of $n \times n$ upper triangular matrices over k with zeros on the main diagonal. If $X, Y \in \mathfrak{u}_n$, we write XY for the product of X and Y as matrices. Then $[X, Y] = XY - YX$. Note that $X^n = 0$ for all $X \in \mathfrak{u}_n$.

Let $\bar{k}$ be a fixed algebraic closure of k. Let $W \subseteq \mathfrak{u}_n$ be a vector k-subspace. We write $W(\bar{k})$ for $W \otimes_k \bar{k}$.

Definition A.1. A vector subspace $W \subseteq \mathfrak{u}_n$ is called *multiplicatively closed* if we have

$$XY \in W \text{ for all } X, Y \in W.$$

Note that any multiplicatively closed subspace of $\mathfrak{u}_n$ is a Lie subalgebra.

Example A.2. For $a, b \in k$, consider the 3×3 matrix

$$M(a,b) = \begin{pmatrix} 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}$$

Let $W = \{M(a,b) : a, b \in k\}$. An easy calculation shows that

$$M(a,b) \cdot M(a',b') = M(0, aa').$$

Thus W is a multiplicatively closed subspace of $\mathfrak{u}_3$.

Lemma A.3. *Let $W \subseteq \mathfrak{u}_n$ be a vector k-subspace. Then W is multiplicatively closed if and only if the $\bar{k}$-subspace $W(\bar{k}) \subseteq \mathfrak{u}_n(\bar{k})$ is multiplicatively closed.*

Proof. Consider the map

$$\psi \colon W \times W \to \mathfrak{u}_n/W, \quad (X, Y) \mapsto XY \bmod W.$$

The map ψ is bilinear, and hence, it is identically zero on $W \times W$ if and only if it is identically zero on $W(\bar{k}) \times W(\bar{k})$. The lemma follows. □

Construction A.4. Write $I_n = \{(i,j) : 1 \leq i < j \leq n\}$. Let Γ be a subset of I_n. For $(i,j) \in I_n$, let E_{ij} denote the $n \times n$ matrix with the matrix element $e_{ij} = 1$ and all other matrix elements equal to zero. We denote by W_Γ the vector k-subspace of $\mathfrak{u}_n$ generated by the matrices E_{ij} for $(i,j) \in \Gamma$. Then the subspace $W_\Gamma \subseteq \mathfrak{u}_n$ is given by the equations $(x_{ij} = 0 : (i,j) \in S)$, where $S = I_n \smallsetminus \Gamma$.

Definition A.5. A subset of $\Gamma \subseteq I_n$ is called *closed* if for any $1 \leq i < k < j \leq n$ such that $(i,k) \in \Gamma$ and $(k,j) \in \Gamma$, we have $(i,j) \in \Gamma$.

Remark A.6. With any pair $(i,j) \in I_n$ one may associate a positive root $\alpha_{ij} = \varepsilon_i - \varepsilon_j \in R(A_{n-1})$; see Bourbaki [1, Plate I, formulas (II)]. Then a subset $\Gamma \subseteq I_n$ is closed if and only if the set of positive roots $\{\alpha_{ij} : (i,j) \in \Gamma\}$ is closed in $R(A_{n-1})$; cf. Bourbaki [1, VI.1.7, Definition 4(i)].

Lemma A.7. *Let $\Gamma \subseteq I_n$. The vector subspace $W_\Gamma \subseteq \mathfrak{u}_n$ of Construction A.4 is multiplicatively closed if and only if Γ is closed.*

Proof. Assume that Γ is closed. Since $(E_{ik} : (i,k) \in \Gamma)$ is a basis of W_Γ, it suffices to check that for any $(i,k),(l,j) \in \Gamma$ we have $E_{ik}E_{lj} \in W_\Gamma$. If $k \neq l$, then $E_{ik}E_{lj} = 0$. If $k = l$, then $i < k < j$ and $E_{ik}E_{kj} = E_{ij}$. Since Γ is closed, we have $(i,j) \in \Gamma$ and hence, $E_{ij} \in W_\Gamma$. Thus W_Γ is multiplicatively closed.

Conversely, assume that W_Γ is multiplicatively closed. Let $1 \leq i < k < j \leq n$ be such that $(i,k),(k,j) \in \Gamma$. Then $E_{ik}\,, E_{kj} \in W_\Gamma$. Since W_Γ is multiplicatively closed, we have $E_{ik}E_{kj} \in W_\Gamma$. Since $E_{ik}E_{kj} = E_{ij}$, we see that $E_{ij} \in W_\Gamma$, and hence, $(i,j) \in \Gamma$. Thus Γ is closed. □

Lemma A.8. *Let $W \subseteq \mathfrak{u}_n$ be a vector k-subspace. Write*

$$1 + W = \{1 + X : X \in W\}.$$

Then the subset $1 + W \subseteq U_n(k)$ is a subgroup if and only if W is multiplicatively closed.

Proof. Assume that $1+W$ is a subgroup. Let $X, Y \in W$. Then $1+X$, $1+Y \in 1+W$, hence

$$(1 + X)(1 + Y) = 1 + X + Y + XY \in 1 + W.$$

Since $1 + X + Y \in 1 + W$, we conclude that $XY \in W$. Thus W is multiplicatively closed.

Conversely, assume that W is multiplicatively closed. Clearly, $1 \in 1 + W$. If $1 + X \in 1 + W$, $1 + Y \in 1 + W$, where $X, Y \in W$, then

$$(1 + X)(1 + Y) = 1 + X + Y + XY \in 1 + W,$$

because $X, Y, XY \in W$.

Now let $1 + X \in W$. Set

$$Z = -X + X^2 - \cdots + (-1)^{n-1}X^{n-1} \in W.$$

Then

$$(1 + X)(1 + Z) = 1 - X^n = 1,$$

where $1 + Z \in 1 + W$. Thus $1 + W$ is a subgroup of $U_n(k)$. □

Construction A.9. Let $\Gamma \subseteq I_n$ and $W_\Gamma \subseteq \mathfrak{u}_n$ be as in Construction A.4. Consider the k-subvariety G_Γ of U_n given by the equations

$$(g_{ij} = 0 : (i,j) \in S),$$

where $S = I_n \smallsetminus \Gamma$. Then $G_\Gamma(k) = 1 + W_\Gamma$ and $G_\Gamma(\bar{k}) = 1 + W_\Gamma(\bar{k})$.

Lemma A.10. *The k-subvariety $G_\Gamma \subseteq U_n$ of Construction A.9 is an algebraic k-subgroup if and only if the subset Γ of I_n is closed in the sense of Definition A.5.*

Proof. The geometrically reduced k-subvariety $G_\Gamma \subseteq U_n$ is an algebraic k-subgroup if and only if $G_\Gamma(\bar{k})$ is a subgroup of $U_n(\bar{k})$. By Lemma A.8, the subset $G_\Gamma(\bar{k}) = 1 + W_\Gamma(\bar{k}) \subseteq U_n(\bar{k})$ is a subgroup if and only if $W_\Gamma(\bar{k})$ is multiplicatively closed. By Lemma A.7, $W_\Gamma(\bar{k})$ is multiplicatively closed if and only if Γ is closed. □

Definition A.11. An algebraic k-subgroup $G \subseteq U_n$ is called *regular* or *almost upper triangular* if $G = G_\Gamma$ for some closed subset $\Gamma \subseteq I_n$.

Note that G_Γ is actually defined over the prime subfield of k.

Remark A.12. A result similar to Lemma A.10 is known for regular subgroups G_Γ of a maximal connected unipotent subgroup U of any split semisimple k-group, at least when $k = \mathbb{C}$. Namely, the regular subgroups of U bijectively correspond to the closed subsets $\Gamma \subseteq R_+$ of the set of positive roots R_+. See [13], Section 6.1.1, Proposition 1.1 on page 183.

Proposition A.13. *Let $W \subseteq \mathfrak{u}_n$ be a multiplicatively closed k-subspace over a field k of positive characteristic $p \geq n$ or of characteristic 0. Consider the algebraic k-subgroup G of U_n with the set of $\bar{k}$-points $1+W(\bar{k})$ (where $1+W(\bar{k})$ is a subgroup of $U_n(\bar{k})$ by Lemma A.3 and Lemma A.8). Then*

(i) $\log(g) \in \mathrm{Lie}(G)$ *for all* $g \in G(k)$;
(ii) $\exp(X) \in G(k)$ *for all* $X \in \mathrm{Lie}(G)$.

Proof. Note that $G(k) = 1 + W$ and that $\mathrm{Lie}(G) = W$.

We prove (i). We have $g = 1 + Y$, where $Y \in W$. Since W is multiplicatively closed, we have $Y^m \in W$ for all $m \geq 1$, and hence,

$$\log(g) = \sum_{i=1}^{n-1} (-1)^{i-1} \frac{Y^i}{i} \in W = \mathrm{Lie}(G).$$

We prove (ii). Let $X \in \mathrm{Lie}(G) = W$. Since W is multiplicatively closed, we have $X^m \in W$ for all $m \geq 1$, and hence,

$$\exp(X) = 1 + \sum_{i=1}^{n-1} \frac{X^i}{i!} \in 1 + W = G(k). \qquad \square$$

Corollary A.14. *Let $G = G_\Gamma \subseteq U_n$ be an almost upper triangular group over a field k of positive characteristic $p \geq n$ or of characteristic 0. Then* $\log(g) \in \mathrm{Lie}(G)$ *for all* $g \in G(k)$ *and* $\exp(X) \in G(k)$ *for all* $X \in \mathrm{Lie}(G)$.

References

[1] Nicolas Bourbaki. *Lie Groups and Lie Algebras, Chapters 4-6*. Springer-Verlag, Berlin, 2002.

[2] Michael Crumley. Generic representation theory of the unipotent upper triangular groups. *Communications in Algebra*, 44(8):3349–3382, 2016.

[3] Nastasescu Dascalescu and Raianu. *Hopf Algebras: An Introduction*. Pure and Applied Mathematics. Marcel Dekker, New York, 2001.

[4] Leonard Eugene Dickson. The analytic representation of substitutions on a power of a prime number of letters with a discussion of the linear group. *The Annals of Mathematics*, 11(1):65–120, 1896 - 1897.

[5] E. B. Dynkin. 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, 1949.

[6] E. B. Dynkin. *Lie Groups*, chapter Normed Lie Algebras and Analytic Groups, pages 481–485. Translations, Series One. American Mathematical Society, Providence, Rhode Island, 1962.

[7] E. B. Dynkin. *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, 2000.

[8] A Suslin E M Friedlander and C P Bendel. Infinitesimal 1-parameter subgroups and cohomology. *Journal of the AMS*, 10(3):693–728, July 1997.

[9] James E. Humphreys. *Linear Algebraic Groups*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1981.
[10] Jens Carsten Jantzen. *Representations of Algebraic Groups*, volume 131 of *Pure and Applied Mathematics*. Academic Press, Orlando, FL, 1987.
[11] Michael Crumley (`https://github.com/mikecrumley`). Computations in dist(g) for g unipotent. Github. URL:https://github.com/mikecrumley/Computations-in-Dist-G-for-G-Unipotent.
[12] Will Sawin (`https://mathoverflow.net/users/18060/will-sawin`). Exponential/logarithm for unipotent algebraic groups. MathOverflow. URL:https://mathoverflow.net/questions/326810 (version: 2019-04-08).
[13] A L Onishchik V V Gorbatsevich and E B Vinberg. *Structure of Lie Groups and Lie Algebras,* Lie Groups and Lie Algebras III, Encyclopedia of Mathematical Sciences, volume 41. Springer-Verlag, Berlin, 1994.
[14] William C. Waterhouse. *Introduction to Affine Group Schemes*. Graduate Texts in Mathematics. Springer-Verlag, New York, 1979.

Email address: `mikecrumley@hotmail.com`