Communications in Algebra®, 41: 3174–3206, 2013

ISSN: 0092-7872 print/1532-4125 online
DOI: 10.1080/00927872.2012.683908

Taylor & Francis
Taylor & Francis Group

GENERIC REPRESENTATION THEORY OF THE HEISENBERG GROUP

Michael Crumley
Department of Mathematics, The University of Toledo, Toledo, Ohio, USA

In this article we extend a result for representations of the additive group G_a given in [4] to the Heisenberg group H_1. Namely, if p is greater than $2d$, then all d-dimensional characteristic p representations for H_1 can be factored into commuting products of representations, with each factor arising from a representation of the Lie algebra of H_1, and conversely any commuting collection of Lie algebra representations gives rise to a representation of H_1 in this fashion. In this sense, for a fixed dimension and large enough p, all representations for H_1 look generically like representations for direct powers of it over a field of characteristic zero.

The following originally appeared as Chapter 13 of the author's dissertation [1].

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

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

1. INTRODUCTION

Denote by G_a and H_1 the Additive and Heisenberg groups respectively over a fixed field k, i.e., the space of all unipotent upper triangular matrices of the form

$$\begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}$$

Throughout we prefer to think of these as affine group schemes (see [7]), i.e., as representable functors on k-algebras represented by the Hopf algebras (see [2])

$$A = k[x]$$
$$\Delta : x \mapsto x \otimes 1 + 1 \otimes x$$
$$\varepsilon : x \mapsto 0$$

Received February 3, 2011; Revised December 21, 2011. Communicated by S. Sehgal.
Address correspondence to Dr. Michael Crumley, Department of Mathematics, The University of Toledo, 2801 W. Bancroft, Mail Stop 942, Toledo, OH 43606, USA; E-mail: mikecrumley@hotmail.com

and

$$A = k[x, y, z]$$
$$\Delta : x \mapsto 1 \otimes x + x \otimes 1, \qquad y \mapsto 1 \otimes y + y \otimes 1, \qquad z \mapsto 1 \otimes z + x \otimes y + z \otimes 1$$
$$\varepsilon : x, y, z \mapsto 0,$$

respectively. Part (1) of the following is well known, and part (2) is well known at least in the case of an algebraically closed field of characteristic zero. However, we prove them both here for convenience for the case of an arbitrary characteristic zero field (see Corollary 2.2 and Theorem 5.11).

Theorem 1.1. *Let k be a field of characteristic zero.*

(1) *Every representation of G_a over k is of the form e^{xX} where X is a nilpotent matrix over k, and any nilpotent matrix X over k gives a representation of G_a according to this formula.*
(2) *Every representation of H_1 over k is of the form $e^{xX+yY+(z-xy/2)Z}$, where X, Y, and Z are nilpotent matrices over k satisfying $Z = [X, Y]$ and $[Z, X] = [Z, Y] = 0$, and any such collection X, Y, Z gives a representation of H_1 according to this formula.*

As regards the positive characteristic representation theory of G_a, the following is also known, originally due to Suslin, Friedlander, and Bendel (Proposition 1.2 of [4] and the remarks that follow). We prove it here, however, for the reader's convenience (using different notation) as part (4) of Proposition 2.6 in the next section.

Theorem 1.2. *Let k be a field of characteristic $p > 0$. Then every representation of G_a over k is of the form*

$$e^{X_0 x} e^{X_1 x^p} \dots e^{X_m p^m},$$

where $X_0, \dots, X_m$ are commuting matrices with entries in k satisfying $X_i^p = 0$. Further, any such collection of commuting, p-nilpotent matrices over k gives a representation of G_a according to the above formula.

Using part (1) of Theorem 1.1 and Theorem 1.2, we make the simple observation that, if $p \geq$ dimension, then being nilpotent and p-nilpotent are identical concepts. We see then that, for $p \gg$ dimension, the characteristic $p > 0$ representation theory of G_a and the characteristic zero theory of G_a^∞ are in perfect analogy (see Chapter 11 of [1] for an account of the representation theory of direct products). This has motivated the following theorem which is the main theorem of this article.

Theorem 1.3. *Let k be a field of characteristic $p > 0$, and suppose $p \geq 2d$. Then every d-dimensional representation of the Heisenberg group over k is of the form*

$$e^{xX_0+yY_0+(z-xy/2)Z_0} e^{x^pX_1+y^pY_1+(z^p-x^py^p/2)Z_1}$$
$$\dots e^{x^{p^m}X_m+y^{p^m}Y_m+(z^{p^m}-x^{p^m}y^{p^m}/2)Z_m},$$

where X_0, Y_0, Z_0, X_1, Y_1, $Z_1 \ldots$, X_m, Y_m, Z_m is a collection of $d \times d$ nilpotent matrices over k satisfying the following statements:

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

Further, any such collection of $d \times d$ matrices gives a representation of H_1 over k according to the above formula.

This result is perhaps more surprising than Theorem 1.2 in that, unlike modules for G_a, modules for H_1 in characteristic $p << \dim$ generally look hardly at all like representation for H_1^∞ in characteristic zero; for instance, it is not generally the case that $[X_i, Y_i] = Z_i$ for all i, nor is it the case that X_i commutes with Y_j for $i \neq j$ (see Section 6 for a counterexample). It is only when p becomes large enough with respect to dimension that these relations necessarily hold.

Our method of proof is quite elementary. We view a representation of an algebraic group over k on a vector space V as a comodule over its representing Hopf algebra (see Section 3.2 of [7] or Chapter 2 of [2]), i.e., as a k-linear map $\rho : V \to V \otimes A$ making commutative the following diagrams:

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\rho"'] & V \otimes A \arrow[d, "1\otimes\Delta"] \\ V \otimes A \arrow[r, "\rho\otimes 1"'] & V \otimes A \otimes A. \end{tikzcd}$$

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[dr, "\simeq"'] & V \otimes A \arrow[d, "1\otimes\varepsilon"] \\ & V \otimes k. \end{tikzcd}$$

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

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

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

These equations induce a combinatorial relation on certain matrices associated to a representation, which serve as necessary and sufficient conditions for them to define a representation. A systematic examination of this relation will yield our main theorem.

Section 2 is devoted to proving part (1) of Theorem 1.1 and Theorem 1.2 for the Additive group in both zero and positive characteristic; these results are not mere curiosities, as we shall in the process record several result necessary for our work with the Heisenberg group. Section 3 proves the bulk of our main theorem (1.3), modulo the Baker–Campbell–Hausdorff type formula mentioned there. In Section 4, we prove the bulk of part (2) of Theorem 1.1 for the Heisenberg group in characteristic zero, again modulo the stated formula; the arguments there are not much different at all than those in Section 3, and we shall draw upon it heavily. Section 5 finally finishes the proofs of both our main theorem (1.3) and part (2) of Theorem 1.1 by proving the veracity of the given formulas.

Sections 6 and 7 give interesting counterexamples of our main theorem when its hypotheses are not satisfied, and some observations and conjectures for future research.

2. THE ADDITIVE GROUP

Here we record several results on the Additive group which will be necessary to prove our main result on the Heisenberg group. Most of what is done here is originally due to Suslin, Friedlander, and Bendel (Proposition 1.2 of [4], as well as the remarks that follow), but it is convenient here to present most of our own proofs.

Let k be any field, and let (a_{ij}) be a representation of G_a over k, which we view as an invertible matrix with entries in $k[x]$, e.g.,

$$\begin{pmatrix} 1 & x & x^2 \\ 0 & 1 & 2x \\ 0 & 0 & 1 \end{pmatrix}.$$

Associated to this representation is, for each $r \in \mathbb{N}$, the matrix of coefficients of the monomial x^r, which we denote as $(c_{ij})^r$. In the above example, these are given by

$$(c_{ij})^0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad (c_{ij})^1 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \qquad (c_{ij})^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

with $(c_{ij})^r = 0$ for all other r. Note that, if (a_{ij}) is the matrix formula for the representation, then $(a_{ij}) = \sum_r (c_{ij})^r x^r$. In what follows, we denote the (i, j)th entry of the matrix $(c_{ij})^r$ as c_{ij}^r.

Proposition 2.1. *Let k be any field. A collection of $d \times d$ matrices $(c_{ij})^r$, $r \in \mathbb{N}$, define a representation of G_a over k if and only if $(c_{ij})^r = 0$ for all but finitely many r, $(c_{ij})^0 = 1$, and that for every r and s*

$$(c_{ij})^r (c_{ij})^s = \binom{r+s}{r} (c_{ij})^{r+s}. \tag{2.1}$$

Proof. This can be gleaned from the proof of Proposition 1.2 of [4], but we prove it here for convenience. That $(c_{ij})^r$ must vanish for all but finitely many r is obvious since the representation is algebraic, and the identity $(c_{ij})^0 = 1$ follows from Eq. (1.4). Consider then Eq. (1.3):

$$\Delta(a_{ij}) = \sum_k a_{ik} \otimes a_{kj}.$$

If we write $a_{ij} = \sum_l c_{ij}^l x^l$, we have

$$\begin{aligned}\Delta(a_{ij}) &= \sum_l c_{ij}^l \Delta(x)^l \\ &= \sum_l c_{ij}^l (1 \otimes x + x \otimes 1)^l \\ &= \sum_l c_{ij}^l \sum_{r+s=l} \binom{r+s}{l} x^r \otimes x^s\end{aligned}$$

which, if we re-arrange to a sum over distinct monomial tensors, can be written as

$$\sum_{r,s} c_{ij}^{r+s} \binom{r+s}{r} x^r \otimes x^s.$$

In particular, the coefficient of the monomial tensor $x^r \otimes x^s$ is exactly $\binom{r+s}{r} c_{ij}^{r+s}$.

On the other hand, we have

$$\begin{aligned}\sum_k a_{ik} \otimes a_{kj} &= \sum_k \left(\sum_r c_{ik}^r x^r\right) \otimes \left(\sum_s c_{kj}^s x^s\right) \\ &= \sum_{r,s} \left(\sum_k c_{ik}^r c_{kj}^s\right) x^r \otimes x^r,\end{aligned}$$

whence we see that the coefficient of $x^r \otimes x^s$ is exactly the (i, j)th entry of the matrix product $(c_{ij})^r (c_{ij})^s$. Finally, since the collection of monomial tensors $x^r \otimes x^s$ are linearly dependent in $A \otimes A$, this forces the matrix equality $(c_{ij})^r (c_{ij})^s = \binom{r+s}{r} (c_{ij})^{r+s}$, as claimed. $\square$

Corollary 2.2. *Let k have characteristic zero. Then every representation of G_a over k is of the form e^{xX}, where X is a nilpotent matrix over k. Further, any nilpotent matrix X over k gives a representation of G_a according to this formula.*

Proof. The latter claim is obvious from the identity $e^{(x+y)X} = e^{xX} e^{yX}$ for commutative variables x and y (see Lemma 5.5). Note that, since X is nilpotent, the series expansion for e^{xX} will make sense over any characteristic zero field, since it is a finite series of matrix expression having integer fraction coefficients.

For the former claim, let $(a_{ij}) = \sum_r (c_{ij})^r x^r$ be any representation, and set $X = (c_{ij})^1$. Using the fact that $\frac{1}{r!}$ is defined for all r, examination of Eq. (2.1) yields $(c_{ij})^r = \frac{1}{r!} X^r$. The necessity that $(c_{ij})^r$ vanish for large enough r forces X to be nilpotent. Then if $n+1$ is the nilpotent order of X, the matrix formula for this representation is

$$\begin{aligned}(a_{ij}) &= (c_{ij})^0 + x(c_{ij})^1 + \cdots + x^n(c_{ij})^n \\ &= 1 + xX + \cdots + \frac{x^n X^n}{n!} \\ &= e^{xX}.\end{aligned}$$

□

In the positive characteristic case, we cannot assume that $\frac{1}{r!}$ is defined for all r. We shall need some number theory.

Definition 2.3. Let p be a prime, and let $a, b, \ldots, z$ be a finite collection of non-negative integers. Write

$$\begin{aligned}a &= a_m p^m + a_{m-1}p^{m-1} + \cdots + a_1 p + a_0 \\ b &= b_m p^m + b_{m-1}p^{m-1} + \cdots + b_1 p + b_0 \\ &\vdots \\ z &= z_m p^m + z_{m-1}p^{m-1} + \cdots + z_1 p + z_0\end{aligned}$$

in p-ary notation (so that for all i, $0 \le a_i, b_i, \ldots, z_i < p$). We call the a_i the *p-digits* of a, similarly for the b_i, etc. We say that the sum $a+b+\cdots+z$ *carries* modulo p if for some i, $a_i + b_i + \cdots + z_i \ge p$ (in case $p = 10$, this coincides with the notion of carrying in the sense of grade school arithmetic). If $n = a+b+\cdots+z$, the *multinomial coefficient* $\binom{n}{a,b,\ldots,z}$ denotes, as usual, $n!/(a!b!\ldots z!)$. If $r \le n$, the *binomial coefficient* $\binom{n}{r}$ is shorthand for the multinomial coefficient $\binom{n}{r,n-r}$.

Theorem 2.4 (Lucas' Theorem)**.** *Let n and $a, b, \ldots, z$ be non-negative integers with $a+b+\cdots+z = n$, p a prime. Write $n = n_0 + n_1 p + \cdots + n_m p^m$ in p-ary notation, similarly for $a, b, \ldots, z$. Then, modulo p,*

$$\binom{n}{a,b,\ldots,z} = \begin{cases} 0 & \text{if for some } i,\ a_i + b_i + \cdots + z_i \ge p \\ \binom{n_0}{a_0, b_0, \ldots, z_0}\binom{n_1}{a_1, b_1, \ldots, z_1}\cdots\binom{n_m}{a_m, b_m, \ldots, z_m} & \text{otherwise.}\end{cases}$$

In other words, $\binom{n}{a,b,\ldots,z}$ is zero modulo p if there is some carrying in computing the sum $a+b+\cdots+z$; otherwise, it is the product of the multinomial coefficients of the individual p-digits.

Proof. See Theorems 14 and 15, [3]. □

Corollary 2.5. *Let p be a prime and n, r, and s non-negative integers.*

(1) $\binom{n}{r}$ *is nonzero if and only if every p-digit of n is greater than or equal to the corresponding p-digit of r;*
(2) $\binom{r+s}{r}$ *is nonzero if and only if there is no carrying for the sum $r+s$.*

Proof. For part (1), by Theorem 2.4, $\binom{n}{r} = \binom{n}{r, n-r}$ is nonzero if and only if there is no carrying in the sum $(n-r)+r$, which is the case if and only if the p-digits of n are all greater than or equal to those of r. Writing $\binom{r+s}{r} = \binom{r+s}{r,s}$, part (2) is likewise obvious from Theorem 2.4. □

Proposition 2.6. *Let (a_{ij}) be a representation of G_a over a field k of characteristic $p > 0$, given by the matrices $(c_{ij})^r$, and for each $r \in \mathbb{N}$ set $X_r = (c_{ij})^{p^r}$.*

(1) *The X_r commute, are nilpotent of order $\leq p$, and are zero for all but finitely many r;*
(2) *For any $r \in \mathbb{N}$, the matrix $(c_{ij})^r$ is given by*

$$(c_{ij})^r = \Gamma(r)^{-1} X_0^{r_0} X_1^{r_1} \dots X_k^{r_k},$$

where $r = r_0 + r_1 p + \cdots + r_k p^k$ is the p-ary expansion of r and $\Gamma(r) \overset{def}{=} r_0! \dots r_k!$
(3) *(a_{ij}), the matrix formula for the representation, is given by*

$$e^{X_0 x} e^{X_1 x^p} \dots e^{X_m p^m};$$

(4) *Any collection of commuting, p-nilpotent matrices $X_0, \dots, X_m$ gives a representation of G_a over k according to the formula given in* (3).

Proof. Most of this can be gleaned from the proof of Proposition 1.2 of [4], but we prove it here for convenience. All is proved by examining Eq. (2.1)

$$(c_{ij})^r (c_{ij})^s = \binom{r+s}{r} (c_{ij})^{r+s}.$$

To see that X_r and X_s commute when $r \neq s$, note that $\binom{p^r+p^s}{p_r} = 1$ by Theorem 2.4, and so

$$X_r X_s = (c_{ij})^{p^r} (c_{ij})^{p^s} = \binom{p^r + p^s}{p^r} (c_{ij})^{p^r+p^s} = (c_{ij})^{p^r+p^s}$$

and likewise

$$X_s X_r = (c_{ij})^{p^s} (c_{ij})^{p^r} = \binom{p^s + p^r}{p^s} (c_{ij})^{p^s+p^r} = (c_{ij})^{p^r+p^s}.$$

To prove the nilpotency claim, note that $\binom{kp^r}{p^r} = k$ for $k = 1, \dots, p-1$ by Theorem 2.4, and consider

$$\begin{aligned}
(c_{ij})^{p^r} &= X_r \\
(c_{ij})^{2p^r} &= \binom{2p^r}{p^r}^{-1} (c_{ij})^{p^r} (c_{ij})^{p^r} = \frac{1}{2} X_r^2 \\
(c_{ij})^{3p^r} &= \binom{3p^r}{p^r}^{-1} (c_{ij})^{2p^r} (c_{ij})^{p^r} = \frac{1}{3}\frac{1}{2} X_r^3
\end{aligned}$$

$$\vdots$$

$$(c_{ij})^{(p-1)p^r} = \frac{1}{(p-1)!} X_r^{p-1}.$$

Noting that there is carrying in computing the sum $(p^{r+1} - p^r) + p^r$, Corollary 2.5 tells us that $\binom{(p-1)p^r + p^r}{p^r} = 0$, and we have

$$\begin{aligned} 0 &= \binom{(p-1)p^r + p^r}{p^r} X_{r+1} \\ &= (c_{ij})^{(p-1)p^r} (c_{ij})^{p^r} \\ &= \frac{1}{(p-1)!} X_r^{p-1} X_r \\ &= \frac{1}{(p-1)!} X_r^p, \end{aligned}$$

forcing $X_r^p = 0$.

For claim (2), let $r = r_0 + r_1 p + \cdots + r_k p^m$ be the p-ary expansion of r. If r has exactly one nonzero p-digit, say $r = r_t p^t$, the second to last computation shows that $(c_{ij})^{r_t p^t} = \frac{1}{r_t!} X_t^{r_t}$. For arbitrary r, we proceed by induction on k, where r_k is the last nonzero p-digit of r, and noting that $\binom{r}{r_k p^k} = 1$, we have

$$\begin{aligned} (c_{ij})^r &= \binom{r}{r_k p^k} (c_{ij})^{r_0 + \cdots + r_{k-1} p^{k-1}} (c_{ij})^{r_k p^k} \\ &= \frac{1}{r_0! r_1! \ldots r_{k-1}!} X_0^{r_0} \ldots X_{k-1}^{r_{k-1}} \frac{1}{r_k!} X_k^{r_k} \\ &= \frac{1}{r_0! r_1! \ldots r_k!} X_0^{r_0} \ldots X_k^{r_k} \\ &= \Gamma(r)^{-1} X_0^{r_0} \ldots X_k^{r_k}, \end{aligned}$$

proving (2).

(3) is proved by straightforward computation, using the expression for $(c_{ij})^r$ just derived and using the commutativity of the X_i:

$$\begin{aligned} (a_{ij}) &= \sum_{r=0}^{p^{m+1}-1} (c_{ij})^r x^r \\ &= \sum_{r=0}^{p^{m+1}-1} \Gamma(r)^{-1} X_0^{r_0} X_1^{r_1} \ldots X_m^{r_m} x^{r_0 + r_1 p + \cdots + r_m p^m} \\ &= \sum_{r_0=0}^{p-1} \sum_{r_1=0}^{p-1} \cdots \sum_{r_m=0}^{p-1} \frac{1}{r_0!} \cdots \frac{1}{r_m!} X_0^{r_0} \ldots X_m^{r_m} x^{r_0} x^{r_1 p} \ldots x^{r_m p^m} \end{aligned}$$

$$
\begin{aligned}
&= \left(\sum_{r_0=0}^{p-1} \frac{1}{r_0!} X_0^{r_0} x^{r_0}\right) \left(\sum_{r_1=0}^{p-1} \frac{1}{r_1!} X_1^{r_1} x^{r_1 p}\right) \dots \left(\sum_{r_m=0}^{p-1} \frac{1}{r_m!} X_m^{r_m} x^{r_m p^m}\right) \\
&= e^{xX_0} e^{x^p X_1} \dots e^{x^{p^m} X_m},
\end{aligned}
$$

as claimed. All of the factors of course commute, since the X_i do.

Finally, let $X_0, \dots, X_m$ be any finite collection of commuting p-nilpotent matrices over k, set $X_i = 0$ for $i > m$, and assign

$$(c_{ij})^r = \Gamma(r)^{-1} X_0^{r_0} \dots X_k^{r_k}$$

as in part (2). Checking that these $(c_{ij})^r$ define a representation of G_a amounts to checking that $(c_{ij})^0 = 1$, that the $(c_{ij})^r$ are nonzero for all but finitely many r (which are both obvious), and Eq. (2.1):

$$\binom{r+s}{r} (c_{ij})^{r+s} = (c_{ij})^r (c_{ij})^s.$$

Let $r = r_0 + \dots + r_k p^k$, $s = s_0 + \dots + s_k p^k$, and suppose first that $\binom{r+s}{r} = 0$. Then by 2.5, there is some carrying in the sum $r + s$, i.e., $r_i + s_i \geq p$ for some i. Looking at the right-hand side of Eq. (2.1), in view of the given assignments we see that $X_i^{r_i+s_i}$ will occur as a factor. But X_i is nilpotent of order less than or equal to p, so the right-hand side will be zero as well. On the other hand, if $\binom{r+s}{r} \neq 0$, then there is no carrying for the sum $r + s$, that is, $r_i + s_i < p$ for every i. Then the given assignments give the same power of each X_i on either side, so it only remains to check the coefficients. This reduces to

$$\binom{r+s}{r} \Gamma(r)\Gamma(s) = \Gamma(r+s).$$

After applying Theorem 2.4 for the term $\binom{r+s}{r}$, the equality is clear. □

3. THE HEISENBERG GROUP IN LARGE POSITIVE CHARACTERISTIC

In this section we prove the bulk of our main theorem (1.3), namely that the relations stated there among the matrices X_i, Y_i, and Z_i are necessary and sufficient to determine a representation of H_1 in characteristic $p \geq 2$ dimension. We defer to Section 5 the proof that such a representation is given by the following formula:

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

For a representation of the Heisenberg group H_1 over a field k and 3-tuple (r, s, t) of non-negative integers, as in the previous section, we define the matrix

$(c_{ij})^{(r,s,t)}$ as the matrix of coefficients of the monomial $x^r y^s z^t$. For example, for the representation

$$\begin{pmatrix} 1 & 2x & x & 2x^2 & z & 2xz \\ & 1 & 0 & x & 0 & z \\ & & 1 & 2x & y & 2xy \\ & & & 1 & 0 & y \\ & & & & 1 & 2x \\ & & & & & 1 \end{pmatrix},$$

we define

$$(c_{ij})^{(0,0,0)} = \mathrm{Id}, \qquad (c_{ij})^{(1,0,1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 2 \\ & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 \\ & & & & 0 & 0 \\ & & & & & 0 \end{pmatrix}, \qquad (c_{ij})^{(1,1,1)} = 0,$$

and so forth. In what follows, we adopt the notation, for 3-tuples $\vec{r} = (r_1, r_2, r_3)$ and $\vec{s} = (s_1, s_2, s_3)$, $\vec{r} + \vec{s} \stackrel{\text{def}}{=} (r_1 + s_1, r_2 + s_2, r_3 + s_3)$. With these $(c_{ij})^{\vec{r}}$ so defined, the matrix formula for a representation of H_1 can be written $(a_{ij}) = \sum_{\vec{r}} (c_{ij})^{\vec{r}} x^{r_1} y^{r_2} z^{r_3}$, with the summation running over all 3-tuples $\vec{r}$ of non-negative integers.

Our first step is to work out the 'fundamental relation' for H_1, analogous to Eq. (2.1) for G_a.

Proposition 3.1. *Let k be any field. A collection $(c_{ij})^{\vec{r}}$ of matrices over k defines a representation of H_1 if and only if they are zero for all but finitely many $\vec{r}$, satisfy $(c_{ij})^{(0,0,0)} = Id$, and for all 3-tuples $\vec{s}$ and $\vec{t}$*

$$(c_{ij})^{\vec{s}}(c_{ij})^{\vec{t}} = \sum_{l=0}^{min(s_1,t_2)} \binom{s_1 + t_1 - l}{t_1} \binom{s_2 + t_2 - l}{s_2} \binom{s_3 + t_3 + l}{s_3, t_3, l} (c_{ij})^{\vec{s}+\vec{t}+(-l,-l,l)}. \quad (3.1)$$

Proof. The first statement is immediate since the representation is algebraic, and the second follows from Eq. (1.4), namely $\varepsilon(a_{ij}) = \delta_{ij}$. For the third, we examine Eq. (1.3), namely $\Delta(a_{ij}) = \sum_k a_{ik} \otimes a_{kj}$:

$$\begin{aligned}
\Delta(a_{ij}) &= \Delta\left(\sum_{\vec{r}} c_{ij}^{\vec{r}} x^{r_1} y^{r_2} z^{r_3}\right) = \sum_{\vec{r}} c_{ij}^{\vec{r}} \Delta(x)^{r_1} \Delta(y)^{r_2} \Delta(z)^{r_3} \\
&= \sum_{\vec{r}} c_{ij}^{\vec{r}} (x \otimes 1 + 1 \otimes x)^{r_1} (y \otimes 1 + 1 \otimes y)^{r_2} (z \otimes 1 + x \otimes y + 1 \otimes z)^{r_3} \\
&= \sum_{\vec{r}} c_{ij}^{\vec{r}} \left[\left(\sum_{k_1 + l_1 = r_1} \binom{k_1 + l_1}{k_1} x^{k_1} \otimes x^{l_1} \right) \left(\sum_{k_2 + l_2 = r_2} \binom{k_2 + l_2}{k_2} y^{k_2} \otimes y^{l_2} \right) \right.
\end{aligned}$$

$$\times\left(\sum_{k_3+l_3+m_3=r_3}\binom{k_3+l_3+m_3}{k_3,\,l_3,\,m_3}x^{l_3}z^{k_3}\otimes y^{l_3}z^{m_3}\right)\Bigg]$$
$$=\sum_{\vec{r}}c_{ij}^{\vec{r}}\sum_{\substack{k_1+l_1=r_1\\k_2+l_2=r_2\\k_3+l_3+m_3=r_3}}\left[\binom{k_1+l_1}{k_1}\binom{k_2+l_2}{k_2}\binom{k_3+l_3+m_3}{k_3,\,l_3,\,m_3}x^{k_1+l_3}y^{k_2}z^{k_3}\otimes x^{l_1}y^{l_2+l_3}z^{m_3}\right].$$

We seek to write this expression as a sum over distinct monomial tensors, i.e., in the form

$$\sum_{\vec{s},\vec{t}}\chi\left(\vec{s},\vec{t}\right)x^{s_1}y^{s_2}z^{s_3}\otimes x^{t_1}y^{t_2}z^{t_3},$$

where the summation runs over all possible pairs of 3-tuples, and $\chi(\vec{r},\vec{s})$ is a scalar for each such pair. Thus, for fixed $\vec{s}$ and $\vec{t}$, we seek non-negative integer solutions to the system of equations

$$\begin{array}{cc} k_1+l_3=s_1 & l_1=t_1\\ k_2=s_2 & l_2+l_3=t_2\\ k_3=s_3 & m_3=t_3\end{array}.$$

Once one chooses l_3 all other values are determined, so we parameterize by l_3. For fixed $l_3=l$, its contribution to the coefficient $\chi\left(\vec{s},\vec{t}\right)$ is

$$\binom{s_1+t_1-l}{t_1}\binom{s_2+t_2-l}{s_2}\binom{s_3+t_3+l}{s_3,\,t_3,\,l}c_{ij}^{(s_1+t_1-l,s_2+t_2-l,s_3+t_3+l)},$$

and in order for such an l to induce a solution, it is necessary and sufficient that it be no larger than either s_1 or t_2, whence we can sum the above expression over all $0\le l\le\min(s_1,t_2)$ to obtain

$$\chi\left(\vec{s},\vec{t}\right)=\sum_{l=0}^{\min(s_1,t_2)}\binom{s_1+t_1-l}{t_1}\binom{s_2+t_2-l}{s_2}\binom{s_3+t_3+l}{s_3,\,t_3,\,l}c_{ij}^{\vec{s}+\vec{t}+(-l,-l,l)}.$$

As for the right-hand side of Eq. (1.3), one easily computes, as in the proof of Proposition 2.1, that

$$\sum_k a_{ik}\otimes a_{kj}=\sum_{\vec{s},\vec{t}}\left(\sum_k c_{ik}^{\vec{s}}c_{kj}^{\vec{t}}\right)x^{\vec{s}}\otimes x^{\vec{t}},$$

and upon matching coefficients for the basis of monomial tensors, we have

$$\sum_k c_{ik}^{\vec{s}}c_{kj}^{\vec{t}}=\sum_{l=0}^{\min(s_1,t_2)}\binom{s_1+t_1-l}{t_1}\binom{s_2+t_2-l}{s_2}\binom{s_3+t_3+l}{s_3,\,t_3,\,l}c_{ij}^{\vec{s}+\vec{t}+(-l,-l,l)},$$

for every $\vec{s}, \vec{t}, i$, and j, i.e.,

$$(c_{ij})^{\vec{s}}(c_{ij})^{\vec{t}} = \sum_{l=0}^{\min(s_1,t_2)} \binom{s_1+t_1-l}{t_1}\binom{s_2+t_2-l}{s_2}\binom{s_3+t_3+l}{s_3, t_3, l}(c_{ij})^{\vec{s}+\vec{t}+(-l,-l,l)},$$

for every $\vec{s}$ and $\vec{t}$. □

Definition 3.2. Let k have characteristic $p > 0$, and (a_{ij}) a representation of H_1 over k given by the matrices $(c_{ij})^{\vec{r}}$, $\vec{r} \in \mathbb{N}^3$. For a non-negative integer m, define

$$X_m \stackrel{\text{def}}{=} (c_{ij})^{(p^m,0,0)}$$
$$Y_m \stackrel{\text{def}}{=} (c_{ij})^{(0,p^m,0)}$$
$$Z_m \stackrel{\text{def}}{=} (c_{ij})^{(0,0,p^m)}$$

and

$$X_{(m)} \stackrel{\text{def}}{=} (c_{ij})^{(m,0,0)}$$
$$Y_{(m)} \stackrel{\text{def}}{=} (c_{ij})^{(0,m,0)}$$
$$Z_{(m)} \stackrel{\text{def}}{=} (c_{ij})^{(0,0,m)}.$$

Note that H_1 contains three copies of the Additive group, namely, those matrices of the form

$$\begin{pmatrix} 1 & x & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ & 1 & y \\ & & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 0 & z \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Then by Propositions 2.1 and 2.6, the following must hold for all r and s:

$$\begin{aligned} X_{(r)}X_{(s)} &= \binom{r+s}{r} X_{(r+s)} \\ X_r^p &= 0 \\ X_r X_s &= X_s X_r \\ X_{(r)} &= \Gamma(r)^{-1} X_0^{r_0} \dots X_m^{r_m}. \end{aligned} \tag{3.2}$$

Identical statements hold if we replace X with Y or Z.

We first observe the following lemma.

Lemma 3.3. *A representation of H_1 over a field k of characteristic $p > 0$ is completely determined by the matrices $X_0, Y_0, Z_0, \dots, X_m, Y_m, Z_m$.*

Proof. We work with the fundamental relation for H_1, Eq. (4.2):

$$(c_{ij})^{\vec{r}}(c_{ij})^{\vec{s}} = \sum_{l=0}^{\min(r_1,s_2)} \binom{r_1+s_1-l}{s_1}\binom{r_2+s_2-l}{r_2}\binom{r_3+s_3+l}{r_3,s_3,l}(c_{ij})^{(r_1+s1-l,r_2+s_2-1,l)}.$$

We have

$$\begin{aligned} Y_{(m)}X_{(n)} &= (c_{ij})^{(0,m,0)}(c_{ij})^{(n,0,0)} \\ &= \sum_{l=0}^{0}\binom{n-l}{n}\binom{m-l}{m}\binom{l}{l}(c_{ij})^{(n-l,m-l,l)} \\ &= (c_{ij})^{(n,m,0)}. \end{aligned}$$

A similar computation shows

$$\begin{aligned} Y_{(m)}X_{(n)}Z_{(k)} &= (c_{ij})^{(n,m,0)}(c_{ij})^{(0,0,k)} \\ &= (c_{ij})^{(n,m,k)}, \end{aligned}$$

whence, for any 3-tuple (n, m, k),

$$(c_{ij})^{(n,m,k)} = Y_{(m)}X_{(n)}Z_{(k)}.$$

But each of $Y_{(m)}$, $X_{(n)}$, and $Z_{(k)}$ are determined by the X_i, Y_i, and Z_i, respectively (Eq. (3.2)), whence so is the entire representation. □

As mentioned in the introduction, there is no reason to suspect that modules for H_1 in characteristic $p > 0$ bear any resemblance to modules for H_1^∞ in characteristic zero, unless p is large enough with respect to dimension.

Lemma 3.4. *Suppose that p is greater than or equal to twice the dimension of a representation, and that the sum $r + s$ carries modulo p (see Definition 2.3). Then at least one of $P_{(r)}$ or $Q_{(s)}$ must be zero, where P and Q can be any of X, Y, or Z (see Definition 3.2).*

Proof. Write $r = r_m p^m + r_{m-1}p^{m-1} + \cdots + r_1 p + r_0$, $s = s_m p^m + s_{m-1}p^{m-1} + \cdots + s_1 p + s_0$ as in Definition 2.3. Since the X_i, Y_i, and Z_i are all nilpotent, they are nilpotent of order less than or equal to the dimension of the representation, which we assume is no greater than $p/2$. Since the sum $r + s$ carries, we have $r_i + s_i \geq p$ for some i, whence, say, $r_i \geq p/2$. Then

$$P_{(r)} = \Gamma(r)^{-1}P_0^{r_0}\ldots P_i^{r_i}\ldots P_i^{r_m}$$

is zero, since $P_i^{r_i}$ is. □

We now prove the necessity of the relations given in Theorem 1.3.

Proposition 3.5. *Suppose p is greater than or equal to twice the dimension of a representation. Then the following relations hold*:

(1) $[Z_n, X_m] = [Z_n, Y_m] = 0$ *for every n and m*;
(2) $[X_m, Y_m] = Z_m$, *for every m*;
(3) $[X_n, Y_m] = 0$, *for every $n \neq m$.*

Proof. To prove (2), consider Eq. (3.1) applied to $X_m Y_m$:

$$\begin{aligned} X_m Y_m &= (c_{ij})^{(p^m,0,0)}(c_{ij})^{(0,p^m,0)} \\ &= Y_m X_m + \left(\sum_{l=1}^{p^m-1} \binom{p^m-l}{0} \binom{p^m-l}{0} \binom{l}{l} Z_{(l)} Y_{(p^m-l)} X_{(p^m-l)} \right) + Z_m. \end{aligned}$$

For every $0 < l < p^m$, there is clearly some carrying in computing the sum $(p^m - l) + l$, so by Lemma 3.4 the summation term $Z_{(l)} Y_{(p^m-l)} X_{(p^m-l)}$ is always zero, since at least one of $Z_{(l)}$ or $Y_{(p^m-l)}$ is zero. This gives $Z_m = [X_m, Y_m]$ as claimed.

To prove (3), consider Eq. (3.1) applied to $X_m Y_n$ for $m \neq n$:

$$\begin{aligned} X_m Y_n &= (c_{ij})^{(p^m,0,0)}(c_{ij})^{(0,p^n,0)} \\ &= Y_n X_m + \left(\sum_{l=1}^{\min(p^n,p^m)} \binom{p^m-l}{0} \binom{p^n-l}{0} \binom{l}{l} Z_{(l)} Y_{(p^n-l)} X_{(p^m-l)} \right). \end{aligned}$$

In case $m < n$, for every value of l in the above summation, the sum $(p^n - l) + l$ carries, forcing at least one of $Z_{(l)}$ or $Y_{(p^n-l)}$ to be zero, forcing every term in the summation to be zero. A similar statement holds in case $n < m$. This proves $X_m Y_n = Y_n X_m$, as claimed.

(1) is in fact true without any hypothesis on the characteristic. To prove it, apply Eq. (3.1) to $X_n Z_m$ and $Z_m X_n$, for which you get the same answer, and the same can be done to show $[Y_m, Z_n] = 0$. $\square$

We have shown thus far that, if $p \geq 2d$, every d-dimensional representation in characteristic p is given by a finite sequence X_i, Y_i, Z_i of $d \times d$ matrices over k satisfying, the following equations:

(1) The X_i, Y_i, and Z_i are all nilpotent;
(2) $Z_i = [X_i, Y_i]$ for every i;
(3) $[X_i, Z_i] = [Y_i, Z_i] = 0$ for every i;
(4) For every $i \neq j$, X_i, Y_i, Z_i all commute with X_j, Y_j, Z_j.

We now show sufficiency of these relations.

Lemma 3.6. *Let k be a field of arbitrary characteristic, and let X, Y, and Z be matrices over k satisfying $Z = [X, Y]$ and $[Z, X] = [Z, Y] = 0$. Then for any $m, n \in \mathbb{N}$,*

$$X^n Y^m = \sum_{l=0}^{min(n,m)} l! \binom{n}{l} \binom{m}{l} Z^l Y^{m-l} X^{n-l}.$$

Proof. We proceed by a double induction on n and m. If n or m is zero, the result is trivial, and if $n = m = 1$, the equation is $XY = YX + Z$, which is true by assumption. Consider then $X^n Y$, and by induction suppose that the claim is true for $X^{n-1}Y$, i.e., that $X^{n-1}Y = YX^{n-1} + (n-1)ZX^{n-2}$. Then using the relation $XY = Z + YX$ and X commuting with Z, we have

$$\begin{aligned} X^n Y &= X^{n-1}XY \\ &= X^{n-1}(Z + YX) \\ &= ZX^{n-1} + (X^{n-1}Y)X \\ &= ZX^{n-1} + (YX^{n-1} + (n-1)ZX^{n-2})X \\ &= nZX^{n-1} + YX^n, \end{aligned}$$

and so the claim is true when $m = 1$. Now suppose that $m \leq n$, so that $\min(n, m) = m$. Then

$$\begin{aligned} X^n Y^m &= (X^n Y)Y^{m-1} \\ &= (YX^n + nZX^{n-1})Y^{m-1} \\ &= Y(X^n Y^{m-1}) + nZ(X^{n-1}Y^{m-1}), \end{aligned}$$

which by induction is equal to

$$\begin{aligned} &= Y\left(\sum_{l=0}^{m-1} l!\binom{n}{l}\binom{m-1}{l} Z^l Y^{m-1-l} X^{n-l}\right) \\ &\quad + nZ\left(\sum_{l=0}^{m-1} l!\binom{n-1}{l}\binom{m-1}{l} Z^l Y^{m-1-l} X^{n-1-l}\right) \\ &= \sum_{l=0}^{m-1} l!\binom{n}{l}\binom{m-1}{l} Z^l Y^{m-l} X^{n-l} \\ &\quad + \sum_{l=0}^{m-1} nl!\binom{n-1}{l}\binom{m-1}{l} Z^{l+1} Y^{m-1-l} X^{n-1-l} \\ &= Y^m X^n + \sum_{l=1}^{m-1} l!\binom{n}{l}\binom{m-1}{l} Z^l Y^{m-l} X^{n-l} \\ &\quad + \sum_{l=1}^{m} n(l-1)!\binom{n-1}{l-1}\binom{m-1}{l-1} Z^l Y^{m-l} X^{n-l}, \end{aligned}$$

where, in the last step, we have chopped off the first term of the first summation and shifted the index l of the second summation. If we chop off the last term of the second summation, we obtain

$$= Y^m X^n + \sum_{l=1}^{m-1} l!\binom{n}{l}\binom{m-1}{l} Z^l Y^{m-l} X^{n-l}$$

$$+\sum_{l=1}^{m-1} n(l-1)!\binom{n-1}{l-1}\binom{m-1}{l-1} Z^l Y^{m-l} X^{n-l}$$

$$+\, n(m-1)!\binom{n-1}{m-1}\binom{m-1}{m-1} Z^m X^{n-m};$$

then, upon merging the summations, we have

$$= Y^m X^n + \sum_{l=1}^{m-1}\left[l!\binom{n}{l}\binom{m-1}{l} + n(l-1)!\binom{n-1}{l-1}\binom{m-1}{l-1}\right] Z^l Y^{m-l} X^{n-l}$$

$$+\, n(m-1)!\binom{n-1}{m-1}\binom{m-1}{m-1} Z^m X^{n-m}$$

$$= Y^m X^n + \sum_{l=1}^{m-1}\left[l!\binom{n}{l}\binom{m-1}{l} + n(l-1)!\binom{n-1}{l-1}\binom{m-1}{l-1}\right] Z^l Y^{m-1} X^{n-l}$$

$$+\, m!\binom{n}{m}\binom{m}{m} Z^m X^{n-m}$$

$$= Y^m X^n + \sum_{l=1}^{m-1}\left[l!\binom{n}{l}\binom{m}{l}\right] Z^l Y^{m-l} X^{n-l}$$

$$+\, m!\binom{n}{m}\binom{m}{m} Z^m X^{n-m}$$

$$= \sum_{l=0}^{m} l!\binom{n}{l}\binom{m}{l} Z^l Y^{m-l} X^{n-l}.$$

This proves the case of $m \leq n$, and the case of $n \leq m$ is hardly any different, and left to the reader. $\square$

Theorem 3.7. *Let k be a field of characteristic $p > 0$, and suppose $p \geq 2d$. Let X_i, Y_i, and Z_i be a finite sequence of $d \times d$ matrices over k satisfying the following statements:*

(1) *The X_i, Y_i, and Z_i are all nilpotent;*
(2) *$Z_i = [X_i, Y_i]$ for every i;*
(3) *$[X_i, Z_i] = [Y_i, Z_i] = 0$ for every i;*
(4) *For every $i \neq j$, X_i, Y_i, Z_i all commute with X_j, Y_j, Z_j.*

Let $n = n_m p^m + n_{m-1}p^{m-1} + \cdots + n_1 p + n_0$, and assign

$$X_{(n)} = \Gamma(n)^{-1} X_0^{n_0} \ldots X_m^{n_m},$$

and similarly for $Y_{(n)}$ and $Z_{(n)}$. Set

$$(c_{ij})^{(n,m,k)} = Z_{(k)} Y_{(m)} X_{(n)}.$$

Then these assignments define a valid d-dimensional representation of H_1 over k.

Proof. The first two conditions of Proposition 3.1 are immediate. Then for arbitrary $n, m, k, r, s, t \in \mathbb{N}$, the equation we must verify is

$$(c_{ij})^{(n,m,k)}(c_{ij})^{(r,s,t)} = \sum_{l=0}^{\min(n,s)} \binom{n+r-l}{r}\binom{m+s-l}{m}\binom{k+t+l}{k,t,l}(c_{ij})^{(n+r-l,m+s-l,k+t+l)}$$

which, with the given assignments and assumptions, can be written as

$$Z_{(k)}Z_{(t)}Y_{(m)}X_{(n)}Y_{(s)}X_{(r)} = \sum_{l=0}^{\min(n,s)} \binom{n+r-l}{r}\binom{m+s-l}{m}\binom{k+t+l}{k,t,l} Z_{(k+t+l)}Y_{(m+s-l)}X_{(n+r-l)}.$$

Eq. (3.2) gives the identities

$$Z_{(k)}Z_{(t)} = \binom{k+t}{t} Z_{(k+t)}$$

$$Y_{(m)}Y_{(s-l)} = \binom{m+s-l}{m} Y_{(m+s-l)}$$

$$X_{(n-l)}X_{(r)} = \binom{n+r-l}{r} X_{(n+r-l)},$$

so we can rewrite our equation as

$$\binom{k+t}{t} Z_{(k+t)}Y_{(m)}X_{(n)}Y_{(s)}X_{(r)} = \sum_{l=0}^{\min(n,s)} \binom{k+t+l}{k,t,l} Z_{(k+t+l)}Y_{(m)}Y_{(s-l)}X_{(n-l)}X_{(r)}.$$

First suppose that the sum $k+t$ carries. In this case, the equation is true, since the left-hand side binomial coefficient vanishes, and the right-hand side multinomial coefficient vanishes for every l, causing both sides to be zero (see Theorem 2.4). We assume then that $k+t$ does not carry, so we can divide both sides by $\binom{k+t}{t}$ to yield

$$Z_{(k+t)}Y_{(m)}X_{(n)}Y_{(s)}X_{(r)} = \sum_{l=0}^{\min(n,s)} \binom{k+t+l}{l} Z_{(k+t+l)}Y_{(m)}Y_{(s-l)}X_{(n-l)}X_{(r)}.$$

Now apply $\binom{k+t+l}{l}Z_{(k+t+l)} = Z_{(k+t)}Z_{(l)}$:

$$Z_{(k+t)}Y_{(m)}X_{(n)}Y_{(s)}X_{(r)} = \sum_{l=0}^{\min(n,s)} Z_{(k+t)}Z_{(l)}Y_{(m)}Y_{(s-l)}X_{(n-l)}X_{(r)}.$$

We have $Z_{(k+t)}$ in the front and $X_{(r)}$ in the rear of both sides, so it suffices to show

$$Y_{(m)}X_{(n)}Y_{(s)} = \sum_{l=0}^{\min(n,s)} Z_{(l)}Y_{(m)}Y_{(s-l)}X_{(n-l)},$$

and since $Y_{(m)}$ commutes with $Z_{(l)}$, we can move it to the front of the right-hand side, and then take it off both sides, so it suffices to show

$$X_{(n)}Y_{(m)} = \sum_{l=0}^{\min(n,m)} Z_{(l)}Y_{(m-l)}X_{(n-l)}, \tag{3.3}$$

where we have replaced s with the more traditional m.

Now we begin to replace the $X'_{(i)}s$ with their definitions in terms of the $X'_i s$, similarly for Y and Z, so that the left-hand side of Eq. (3.3) is

$$[\Gamma(n)\Gamma(m)]^{-1}\, X_0^{n_0}\ldots X_k^{n_k}Y_0^{m_0}\ldots Y_k^{m_k},$$

and since everything commutes except X_i and Y_j when $i=j$, we can write

$$[\Gamma(n)\Gamma(m)]^{-1}\,(X_0^{n_0}Y_0^{m_0})\ldots(X_k^{n_k}Y_k^{m_k}).$$

Moving all coefficients to the right, we must show

$$(X_0^{n_0}Y_0^{m_0})\ldots(X_k^{n_k}Y_k^{m_k}) = \Gamma(n)\Gamma(m)\sum_{l=0}^{\min(n,m)} Z_{(l)}Y_{(m-l)}X_{(n-l)}. \tag{3.4}$$

We proceed by induction on k, the maximum number of p-digits of either m or n. If $k=0$, the equation is

$$\begin{aligned}
X_0^{n_0}Y_0^{m_0} &= n_0!m_0!\sum_{l=0}^{\min(n_0,m_0)} Z_{(l)}Y_{(m_0-l)}X_{(n_0-l)} \\
&= \sum_{l=0}^{\min(n_0,m_0)} \frac{n_0!m_0!}{(m_0-l)!(n_0-l)!l!}Z_0^lY_0^{m_0-l}X_0^{n_0-l} \\
&= \sum_{l=0}^{\min(n_0,m_0)} l!\binom{n_0}{l}\binom{m_0}{l}Z_0^lY_0^{m_0-l}X_0^{n_0-l},
\end{aligned}$$

which is true by Lemma 3.6 applied to X_0, Y_0, and Z_0. Now suppose that Eq. (3.4) holds when n and m have no more than $k-1$ digits. Let $n = n_{k-1}p^{k-1}+\cdots+n_0$, and let $n' = n_kp^k + n_{k-1}p^{k-1}+\cdots+n_0$, and similarly for m. Then by induction, we have

$$\begin{aligned}
&(X_0^{n_0}Y_0^{m_0})\ldots(X_k^{n_k}Y_k^{m_k}) \\
&= \left[(X_0^{n_0}Y_0^{m_0})\ldots(X_{k-1}^{n_{k-1}}Y_{k-1}^{m_{k-1}})\right](X_k^{n_k}Y_k^{m_k}) \\
&= \left(\Gamma(n)\Gamma(m)\sum_{l=0}^{\min(n,m)} Z_{(l)}Y_{(m-l)}X_{(n-l)}\right) \\
&\quad\times\left(\sum_{l'=0}^{\min(n_k,m_k)} l'!\binom{n_k}{l'}\binom{m_k}{l'}Z_k^{l'}Y_k^{m_k-l'}X_k^{n_k-l'}\right) \\
&= n_k!\Gamma(n)m_k!\Gamma(m)\sum_{l,l'}\left[\left(\frac{Z_{(l)}Z_k^{l'}}{l'!}\right)\left(\frac{Y_{(m-l)}Y_k^{m_k-l'}}{(m_k-l')!}\right)\left(\frac{X_{(n-l)}X_k^{n_k-l'}}{(n_k-l')!}\right)\right].
\end{aligned}$$

Note that these divisions are valid, since for every value of l' in the summation, $l' \le m_k, n_k < p$. Note also that, since $l \le p^{k-1}$ and $l' < p$ for all values of l, l' in the summation, Theorem 2.4 gives that $\binom{l+l'p^k}{l} = 1$ for all such l and l'. For similar reasons, we have $\binom{(m-l)+(m_k-l')p^k}{m-l} = \binom{(n-l)+(n_k-l')p^k}{n-l} = 1$. Then we have the identities

$$\begin{aligned}
n_k!\Gamma(n) &= \Gamma(n') \\
m_k!\Gamma(m) &= \Gamma(m') \\
\frac{Z_{(l)}Z_k^{l'}}{l'!} &= Z_{(l)}Z_{(l'p^k)} \\
&= \binom{l+l'p^k}{l} Z_{(l+l'p^k)} \\
&= Z_{(l+l'p^k)} \\
\frac{Y_{(m-l)}Y_k^{m_k-l'}}{(m_k-l')!} &= Y_{(m-l)}Y_{((m_k-l')p^k)} \\
&= \binom{(m-l)+(m_k-l')p^k}{m-l} Y_{((m+m_kp^k)-(l+l')p^k)} \\
&= Y_{(m'-(l+l'p^k))},
\end{aligned}$$

and similarly

$$\frac{X_{(n-l)}X_k^{n_k-l'}}{(n_k-l')!} = X_{(n'-(l+l'p^k))}.$$

These substitutions transform the right-hand side of our equation into

$$= \Gamma(n')\Gamma(m') \sum_{l,l'} Z_{(l+l'p^k)}Y_{(m'-(l+l'p^k))}X_{(n'-(l+l'p^k))}.$$

But, if we look at the summation limits of $l = 0 \ldots \min(n, m)$ and $l' = 0 \ldots \min(n_k, m_k)$, we see that it is really a single summation running from 0 to $\min(n', m')$, with $l + l'p^k$ as the summation variable. That is,

$$= \Gamma(n')\Gamma(m') \sum_{l=0}^{\min(n',m')} Z_{(l)}Y_{(m'-l)}X_{(n'-l)},$$

which proves Eq. (3.4). This completes the proof. $\square$

The main theorem of this article, Theorem 1.3, is now proved, modulo the fact that, for $\mathrm{char}(k) \ge 2d$, such a collection X_i, Y_i, Z_i of $d \times d$ matrices generate a representation according to the formula

$$e^{xX_0+yY_0+(z-xy/2)Z_0}e^{x^pX_1+y^pY_1+(z^p-x^py^p/2)Z_1} \\ \ldots e^{x^{p^m}X_m+y^{p^m}Y_m+(z^{p^m}-x^{p^m}y^{p^m}/2)Z_m}.$$

We shall prove this, and its characteristic zero analogue, in Section 5.

4. THE HEISENBERG GROUP IN CHARACTERISTIC ZERO

In this section we prove the bulk of part (2) of Theorem 1.1, namely that the relations stated there among the matrices X, Y, and Z are necessary and sufficient to determine a representation of H_1 in characteristic zero. We defer to Section 5 the proof that such a representation is given by the formula

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

We mention again that this result is (at least) known for the case of an algebraically closed characteristic zero field. Though we have chosen to present this result in a section separate to that of the previous section dealing with large positive characteristic, the characteristic zero result is an almost direct corollary to (the arguments for) the large positive characteristic result. We shall therefore reference analogous results and proofs from Section 3 often, taking care to justify that they apply equally well to the characteristic zero setting.

Let k be a field of characteristic zero. As in the previous section for (the matrix formula for) a representation of H_1 over k and $(r, s, t) \in \mathbb{N}^3$, define $(c_{ij})^{(r,s,t)}$ to be the matrix of coefficients of the monomial $x^r y^s z^t$.

Set $X = (c_{ij})^{(1,0,0)}$, $Y = (c_{ij})^{(0,1,0)}$, $Z = (c_{ij})^{(0,0,1)}$. Note that, in the notation of the previous section, $X \stackrel{\text{def}}{=} X_{(1)}$, $Y \stackrel{\text{def}}{=} Y(1)$, $Z \stackrel{\text{def}}{=} Z_{(1)}$. We shall have no need in this section for the notation $X_m \stackrel{\text{def}}{=} X_{(p^m)}$, since p is irrelevant here.

Note that H_1 contains three copies of the additive group G_a, namely, those matrices of the form

$$\begin{pmatrix} 1 & x & 0 \\ & 1 & 0 \\ & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ & 1 & y \\ & & 1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 0 & z \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Then the proof of Corollary 2.2 gives, for all $r \in \mathbb{N}$,

$$(c_{ij})^{(r,0,0)} = \frac{1}{r!}X^r \qquad (c_{ij})^{(0,r,0)} = \frac{1}{r!}Y^r \qquad (c_{ij})^{(0,0,r)} = \frac{1}{r!}Z^r, \tag{4.1}$$

and that each of X, Y, and Z must be nilpotent. Recall also, for 3-tuples $\vec{s} = (s_1, s_2, s_3)$ and $\vec{t} = (t_1, t_2, t_3)$, the 'fundamental relation' for H_1

$$(c_{ij})^{\vec{s}}(c_{ij})^{\vec{t}} = \sum_{l=0}^{\min(s_1,t_2)} \binom{s_1+t_1-l}{t_1}\binom{s_2+t_2-l}{s_2}\binom{s_3+t_3+l}{s_3, t_3, l}(c_{ij})^{\vec{s}+\vec{t}+(-l,-l,l)} \tag{4.2}$$

(see Proposition 3.1) which is valid in any characteristic.

Theorem 4.1. *A representation of H_1 over k is completely determined by the assignments X, Y, and Z. Necessarily, $Z = [X, Y]$, each of X and Y must commute with Z, and X, Y, and Z must all be nilpotent. Further, any X, Y, and Z satisfying these relations defines a representation of H_1 over k.*

Proof. We work with the fundamental relation for H_1, Eq. (4.2):

$$(c_{ij})^{\vec{r}}(c_{ij})^{\vec{s}} = \sum_{l=0}^{\min(r_1,s_2)} \binom{r_1+s_1-l}{s_1}\binom{r_2+s_2-l}{r_2}\binom{r_3+s_3+l}{r_3,\, s_3,\, l}(c_{ij})^{(r_1+s1-l,r_2+s_2-1,l)}.$$

We have

$$(c_{ij})^{(0,m,0)}(c_{ij})^{(n,0,0)} = \sum_{l=0}^{0}\binom{n-l}{n}\binom{m-l}{m}\binom{l}{l}(c_{ij})^{(n-l,m-l,l)},$$

which says that

$$(c_{ij})^{(n,m,0)} = \frac{1}{m!n!}Y^m X^n.$$

Using again the fundamental relation, we also have

$$(c_{ij})^{(n,m,0)}(c_{ij})^{(0,0,k)} = (c_{ij})^{(n,m,k)},$$

which together with the last equation gives

$$(c_{ij})^{(n,m,k)} = \frac{1}{n!m!k!}Y^m X^n Z^k.$$

Thus, all of the $(c_{ij})^{\vec{r}}$ are determined by X, Y, and Z, according to the above formula. Further,

$$XY = \sum_{l=0}^{1}\binom{1-l}{0}\binom{1-l}{0}\binom{l}{l}(c_{ij})^{(1-l,1-l,l)} = YX + Z,$$

and so $Z = [X, Y]$ as claimed. Each of Y and X must commute with Z, for if we apply the fundamental relation to each of XZ and ZX, in each case we obtain $(c_{ij})^{(1,0,1)}$, showing $XZ = ZX$, and an identical computation shows $YZ = ZY$. And by the remarks preceding the statement of this theorem, each of X, Y, and Z must be nilpotent.

We must now show sufficiency of the given relations. Let X, Y, and Z be any three nilpotent matrices satisfying $Z = XY - YX$, with each of X and Y commuting with Z. We need to show that the fundamental relation, Eq. (4.2), is always satisfied if we assign

$$(c_{ij})^{(n,m,k)} = \frac{1}{n!m!k!}Z^k Y^m X^n.$$

Since each of X, Y, and Z are nilpotent, $(c_{ij})^{\vec{r}}$ will vanish for all by finitely many $\vec{r}$, and $(c_{ij})^{(0,0,0)} = 1$, as required. The fundamental relation, with these assignments, reduces to the following equation (after shuffling all coefficients to the right-hand side and some cancellation):

$$Z^{r_3+s_3} Y^{r_2} X^{r_1} Y^{s_2} X^{s_1} = \sum_{l=0}^{\min(r_1,s_2)} l! \binom{r_1}{l} \binom{s_2}{l} Z^{r_3+s_3+l} Y^{r_2+s_2-l} X^{r_1+s_1-l}.$$

Each term in the summation has the term $Z^{r_3+s_3}$ in the front and X^{s_1} in the rear, and so does the left-hand side. So it suffices to show

$$Y^{r_2} X^{r_1} Y^{s_2} = \sum_{l=0}^{\min(r_1,s_2)} l! \binom{r_1}{l} \binom{s_2}{l} Z^l Y^{r_2+s_2-l} X^{r_1-l}$$

and since Y commutes with Z, we can write

$$Y^{r_2} X^{r_1} Y^{s_2} = \sum_{l=0}^{\min(r_1,s_2)} l! \binom{r_1}{l} \binom{s_2}{l} Y^{r_2} Z^l Y^{s_2-l} X^{r_1-l}.$$

We can now remove Y^{r_2} from the front of either side, so it suffices to show

$$X^n Y^m = \sum_{l=0}^{\min(n,m)} l! \binom{n}{l} \binom{m}{l} Z^l Y^{m-l} X^{n-l}$$

(where we have replaced r_1 and s_2 with n and m, respectively). But this is the statement of Lemma 3.6, which is true in any characteristic. ☐

5. BAKER–CAMPBELL–HAUSDORFF FORMULA FOR THE HEISENBERG GROUP IN ZERO AND LARGE POSITIVE CHARACTERISTIC

In this section, we finish the proof of part (2) of Theorem 1.1 and the proof of Theorem 1.3.

Proposition 5.1.

(1) *Let k have characteristic zero. Let $M(x, y, z)$ be the matrix formula for a representation of H_1 given by the matrices X, Y, and Z as in Theorem* 4.1. *Then*

$$M(x, y, z) = e^{xX+yY+(z-xy/2)Z}.$$

(2) *Let k have characteristic $p > 0$, and suppose $p \geq 2d$. Let $M(x, y, z)$ be the matrix formula for a d-dimensional representation of H_1 given by the matrices $X_0, Y_0, Z_0, \dots, X_m, Y_m, Z_m$ as in Theorem* 3.7. *Then*

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

Our line of proof parallels almost exactly that given in Section 3.1 of [5], where the Baker–Campbell–Hausdorff formula is derived for the Heisenberg group as a

Lie group, with only minor alterations needed for the case of a positive characteristic field (of sufficiently large characteristic when compared to the dimension of a representation). The main difference is that we replace the notion of derivative of an analytic function with 'formal derivative' of polynomials over a field.

For the remainder, by a polynomial $f(t)$, we shall mean a polynomial in the commuting variable t with coefficients which are matrix expressions among the matrices X, Y, and Z over a given field; for example, $f(t) = XY + 2(Z - YX)t + \frac{Y}{2}t^2 - t^3$.

Definition 5.2. The *formal derivative*, or just *derivative*, of the polynomial

$$f(t) = M_n t^n + M_{n-1}t^{n-1} \cdots + M_1 t + M_0,$$

where the M_i are matrix expressions among X, Y, and Z, is the polynomial

$$f'(t) = nM_n t^{n-1} + (n-1)M_{n-1}t^{n-2} + \cdots + M_1.$$

To begin, we note that the following facts hold just as well for formal differentiation as they do for standard differentiation.

Lemma 5.3. *Let $f(t)$, $g(t)$ be polynomials, and suppose that the field is either of characteristic zero, or of positive characteristic greater than the degrees of both $f(t)$ and $g(t)$.*

(1) (*Product rule*) $(fg)' = f'g + fg'$;
(2) (*Uniqueness of antiderivatives*) *If $f'(t) = g'(t)$, and if $f(0) = g(0)$, then $f(t) = g(t)$;*
(3) (*Uniqueness of solutions to first-order linear differential equations*) *Let M be some matrix expression among X, Y, and Z. If $f'(t) = Mf(t)$, and if $g'(t) = Mg(t)$, and if $f(0) = g(0)$, then $f(t) = g(t)$.*

Remark 5.4. In positive characteristic, the assumption that $\mathrm{char}(k) > \text{degree}$ is essential. For example, in characteristic 2, the derivatives of the polynomials t^2 and 0 are both zero, and they are both zero when evaluated at $t = 0$, but they are obviously not themselves equal.

Proof. (1) is true even without any hypothesis on the characteristic. Let $f = \sum_{k=0}^{m} a_k t^k$, $g = \sum_{k=0}^{m} b_k t^k$, where the a_i, b_i are matrix expressions in X, Y, and Z. Then

$$\begin{aligned}
(fg)' &= \left[\left(\sum_{k=0}^{m} a_k t^k\right)\left(\sum_{l=0}^{m} b_l t^l\right)\right]' \\
&= \left[\sum_{r=0}^{2m}\left(\sum_{k+l=r} a_k b_l\right) t^r\right]' \\
&= \sum_{r=0}^{2m} r\left(\sum_{k+l=r} a_k b_l\right) t^{r-1}
\end{aligned}$$

$$= \sum_{r=1}^{2m} r \left(\sum_{k+l=r} a_k b_l \right) t^{r-1}$$
$$= \sum_{r=0}^{2m-1} (r+1) \left(\sum_{k+l=r+1} a_k b_l \right) t^r$$

and

$$f'g + fg' = \left(\sum_{k=0}^{m} k a_k t^{k-1} \right) \left(\sum_{l=0}^{m} b_l t^l \right) + \left(\sum_{k=0}^{m} a_k t^k \right) \left(\sum_{l=0}^{m} l b_l t^{l-1} \right)$$
$$= \sum_{k,l=0}^{m} (k a_k b_l)\, t^{k-1+l} + \sum_{k,l=0}^{m} (l a_k b_l)\, t^{k-1+l}$$
$$= \sum_{k,l=0}^{m} (k+l)(a_k b_l) t^{k+l-1}$$
$$= \sum_{r=0}^{2m-1} \left(\sum_{k+l-1=r} (k+l) a_k b_l \right) t^r$$
$$= \sum_{r=0}^{2m-1} (r+1) \left(\sum_{k+l=r+1} a_k b_l \right) t^r,$$

which proves (1).

For (2), let f and g be as before. To say that $f' = g'$ is to say that $na_n = nb_n, (n-1)a_{n-1} = (n-1)b_{n-1}, \ldots, a_1 = b_1$, and to say that $f(0) = g(0)$ is to say that $a_0 = b_0$. Under the given hypotheses all of $n, n-1, \ldots, 1$ are invertible, which forces $a_n = b_n$, $a_{n-1} = b_{n-1}$, $\ldots a_1 = b_1$, and $a_0 = b_0$, whence $f = g$. This proves (2).

For (3), suppose $f' = Mf$ and $g' = Mg$. Then by matching coefficients for the various powers of t, this forces the equalities

$$\begin{array}{cc} Ma_n = 0 & Mb_n = 0 \\ na_n = Ma_{n-1} & nb_n = Mb_{n-1} \\ (n-1)a_{n-1} = Ma_{n-2} & (n-1)b_{n-1} = Mb_{n-2} \\ \vdots & \vdots \\ 2a_2 = Ma_1 & 2b_2 = Mb_1 \\ a_1 = Ma_0 & b_1 = Mb_0. \end{array}$$

The assumption $f(0) = g(0)$ again forces $a_0 = b_0$. Noting again that all of $n, n-1, \ldots, 1$ are invertible, we can work backwards to see that $a_1 = Ma_0 = Mb_0 = b_1$, that $a_2 = \frac{1}{2}Ma_1 = \frac{1}{2}Mb_1 = b_2, \ldots, a_n = \frac{1}{n}Ma_{n-1} = \frac{1}{n}Mb_{n-1} = b_n$, and whence $f = g$. This proves (3). □

Lemma 5.5. *Let X and Y be commuting nilpotent matrices over a field k such that k is of characteristic zero, or of positive characteristic greater than or equal to the dimension of X and Y. Then*:

(1) $(e^{tX})' = e^{tX} X$;

(2) $(e^{t^2X})' = e^{t^2X}(2tX)$;
(3) $e^X e^Y = e^{X+Y}$.

Remark 5.6. In characteristic $p > 0$, the assumption that $\mathrm{char}(k) \geq$ dimension $\geq$ {nilpotent order of X or Y} is essential for the statement of the lemma to even make sense, since otherwise any of the above series may contain fractions with denominators divisible by p. For (3), note that the right hand side is defined; if X and Y commute, they can be put in simultaneous upper triangular form, whence $X + Y$ is also nilpotent.

Proof. Let $n + 1$ be the dimension of X and Y. For (1), compute

$$\begin{aligned}(e^{tX})' &= \left(1 + tX + \frac{t^2X^2}{2!} + \cdots + \frac{t^nX^n}{n!}\right)' \\ &= X + \frac{2tX^2}{2!} + \cdots + \frac{nt^{n-1}X^n}{n!} \\ &= X\left(1 + tX + \cdots + \frac{t^{n-1}X^{n-1}}{(n-1)!} + \frac{t^nX^n}{n!}\right) \\ &= Xe^{tX}.\end{aligned}$$

Note that, in the second to last expression, we are justified in tacking on the term $\frac{t^nX^n}{n!}$ since multiplication by X will annihilate it anyway. This proves (1).

For (2), compute again

$$\begin{aligned}(e^{t^2X})' &= \left(1 + t^2X + \frac{t^4X^2}{2!} + \cdots + \frac{t^{2n}X^n}{n!}\right)' \\ &= 2tX + \frac{4t^3X^2}{2!} + \cdots + \frac{2nt^{2n-1}X^n}{n!} \\ &= 2tX\left(1 + \frac{2t^2X}{2!} + \cdots + \frac{nt^{2(n-1)}X^n}{n!}\right) \\ &= 2tX\left(1 + t^2X + \frac{t^4X^2}{2!} + \cdots + \frac{t^{2(n-1)}X^{n-1}}{(n-1)!} + \frac{t^{2n}X^n}{n!}\right) \\ &= 2tXe^{t^2X},\end{aligned}$$

where again, in the second to last expression, we are justified in tacking on the term $\frac{t^{2n}X^n}{n!}$ since multiplication by X will annihilate it anyhow. This proves (2).

For (3), we shall prove that $e^{tX}e^{tY} = e^{t(X+Y)}$ as polynomials; evaluating at $t = 1$ will give the desired result. By (3) of Lemma 5.3, since they are equal when evaluated at $t = 0$, it is enough to show that they satisfy the same differential equation:

$$\begin{aligned}\left(e^{tX}e^{tY}\right)' &= \left(e^{tX}\right)' e^{tY} + e^{tX}\left(e^{tY}\right)' \\ &= Xe^{tX}e^{tY} + e^{tX}Ye^{tY} \\ &= e^{tX}e^{tY}(X + Y)\end{aligned}$$

and

$$\left(e^{t(X+Y)}\right)' = e^{t(X+Y)}(X+Y).$$

This completes the proof. □

Lemma 5.7 (Baker–Campbell–Hausdorff Formula for the Heisenberg Group). *Let X and Y be nilpotent matrices over a field, commuting with their nilpotent commutator Z. If the field is either of characteristic zero, or of positive characteristic larger than twice the dimension of the matrices, then*

$$e^X e^Y = e^{X+Y+\frac{1}{2}Z}.$$

Remark 5.8. Note firstly that the matrix $X + Y + \frac{1}{2}Z$ is guaranteed to be nilpotent. Under the given hypotheses, according to either Theorem 3.7 or Theorem 4.1, X, Y, and Z define a representation of H_1. As the homomorphic image of any unipotent algebraic group is also unipotent, and as any unipotent algebraic group can be embedded in a closed subgroup, for some n, the group of all $n \times n$ unipotent upper triangular matrices over k (see Theorem 2.4.11 of [6]), X, Y, and Z can be simultaneously put in strictly upper triangular form. It is obvious then that any linear combination of X, Y, and Z is also nilpotent.

In characteristic $p > 0$, the stated assumptions are again necessary for the statement of the lemma to even make sense. So long as $p \geq 2$dimension, p will be greater than the nilpotent orders of each of X, Y, and $X + Y + \frac{1}{2}Z$, forcing the terms of these series to vanish before fractions with denominators divisible by p can occur. Note also that the theorem makes sense even in characteristic 2; in this case the given assumptions force dimension to be 0 or 1, whence $Z = [X, Y] = 0$, and whence the fraction $\frac{1}{2}$ on the right-hand side of the above does not appear.

Proof. We shall prove something stronger, namely that

$$e^{tX}e^{tY} = e^{tX+tY+\frac{t^2}{2}Z}$$

as polynomials; evaluating at $t = 1$ will give the desired result.

The proof proceeds exactly as in Theorem 3.1 of [5] for the Lie group case. Since Z commutes with both X and Y, we can rewrite the above equation as

$$e^{tX}e^{tY}e^{\frac{-t^2}{2}Z} = e^{t(X+Y)}.$$

Denote by $A(t)$ the left-hand side of this equation, $B(t)$ the right-hand side. These are both equal to 1 when evaluated at $t = 0$, so by (3) of Lemma 5.3 it suffices to show that they both satisfy the same linear differential equation. Working first with $A(t)$, using the iterated product rule we have

$$\begin{aligned} A'(t) &= e^{tX}Xe^{tY}e^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}Ye^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}e^{\frac{-t^2}{2}Z}(-tZ) \\ &= e^{tX}e^{tY}\left(e^{-tY}Xe^{tY}\right)e^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}Ye^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}e^{\frac{-t^2}{2}Z}(-tZ). \end{aligned}$$

We claim that $e^{-tY}Xe^{tY}$ is equal to $X + tZ$. They are both equal to X when evaluated at $t = 0$, and $(X + tZ)' = Z$, so it suffices to show by part 2 of Lemma 5.3 that the derivative of $e^{-tY}Xe^{tY}$ is equal to Z:

$$\begin{aligned}\left(e^{-tY}Xe^{tY}\right)' &= e^{-tY}(-YX)e^{tY} + e^{-tY}XYe^{tY}\\ &= e^{-tY}(XY - YX)e^{tY}\\ &= e^{-tY}Ze^{tY}\\ &= Z,\end{aligned}$$

as required. Thus

$$\begin{aligned}A'(t) &= e^{tX}e^{tY}\,(X + tZ)\,e^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}Ye^{\frac{-t^2}{2}Z} + e^{tX}e^{tY}e^{\frac{-t^2}{2}Z}(-tZ)\\ &= e^{tX}e^{tY}e^{\frac{-t^2}{2}Z}(X + tZ + Y - tZ)\\ &= e^{tX}e^{tY}e^{\frac{-t^2}{2}Z}(X + Y)\\ &= A(t)(X + Y).\end{aligned}$$

To finish then, it suffices to show that $B'(t) = B(t)(X + Y)$; but this is obvious by part (1) of Lemma 5.5. This completes the proof. □

Proposition 5.9. *Let X, Y, and Z be matrices over a field k such that the following statements hold*:

(1) *k is of characteristic zero, or of positive characteristic at least twice the dimensions of X, Y, and Z;*
(2) *$Z = [X, Y]$, and $[X, Z] = [Y, Z] = 0$;*
(3) *X, Y and Z are all nilpotent.*

Then

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

is a representation of H_1 over k.

Remark 5.10. See Remark 5.8 to see that this expression is always defined under the given hypotheses.

Proof. We need to verify that the multiplication of H_1 is preserved by this formula, i.e., that for all x, y, z, r, s, t, the following matrix equality is true:

$$e^{xX+yY+(z-xy/2)Z}e^{rX+sY+(t-rs/2)Z} = e^{(x+r)X+(y+s)Y+(z+xs+t-(x+r)(y+s)/2)Z}.$$

If X and Y are nilpotent and commute with their nilpotent commutator Z, then $xX + yY$ and $rX + sY$ are also nilpotent, and also commute with their nilpotent commutator $(xs - yr)Z$. So Lemma 5.7 applies:

$$e^{xX+yY}e^{rX+sY} = e^{(x+r)X+(y+s)Y+((xs-yr)/2)Z}.$$

Recalling that $e^R e^S = e^{R+S}$ whenever R and S commute (Proposition 5.5), the left-hand side of our first equation can be written as

$$\begin{aligned} e^{xX+yY} e^{rX+sY} e^{(z+t-(xy+rs)/2)Z} &= e^{(x+r)X+(y+s)Y+((xs-yr)/2)Z} e^{(z+t-(xy+rs)/2)Z} \\ &= e^{(x+r)X+(y+s)Y+(z+xs+t-(x+r)(y+s)/2)Z}, \end{aligned}$$

which proves the proposition. □

Theorem 5.11. *Let k be a field of characteristic zero, and let $M(x, y, z)$ be the matrix formula for a representation of H_1 over k given by the matrices X, Y, and Z, as in Theorem* 4.1. *Then*

$$M(x, y, z) = e^{xX+yY+(z-xy/2)Z}.$$

Proof. Proposition 5.9 guarantees that the right-hand side of the above is actually a representation, and Theorem 4.1 tells us that a representation of H_1 over k is completely determined by the assignments $(c_{ij})^{(1,0,0)}$, $(c_{ij})^{(0,1,0)}$, and $(c_{ij})^{(0,0,1)}$. Thus, we need only check that, for the right-hand side, the matrix of coefficients of the monomial x is actually X, that for y it is actually Y, and that for z it is actually Z. Examine

$$\begin{aligned} e^{xX+yY+(z-xy/2)Z} = 1 &+ xX + yY + (z - xy/2)Z \\ &+ (xX + yY + (z - xy/2)Z)^2 \\ &+ \dots \\ &+ \frac{1}{n!}(xX + yY + (z - xy/2)Z)^n \end{aligned}$$

(where $n+1$ is the nilpotent order of $xX + yY + (z - xy/2)Z$). Clearly the monomials x, y, and z all show up exactly once in this series expansion, namely in the term $xX + yY + (z - xy/2)Z$, and X, Y, and Z are their respective coefficient matrices. This proves the theorem. □

Lemma 5.12. *Suppose $M(x, y, z)$ is the matrix formula for a representation of H_1 over a field k of characteristic $p > 0$. Then so is $M(x^p, y^p, z^p)$.*

Remark 5.13. Going from $M(x, y, z)$ to $M(x^p, y^p, z^p)$ is not quite the same as applying a frobenius twist. We are raising the powers of the variables only, not the scalars.

Proof. Let $M(x, y, z)$ correspond to the comodule map $\rho : V \to V \otimes A$, where V is the k-vector space being acted upon by $M(x, y, z)$ and $A = k[x, y, z]$ is the representing Hopf algebra of H_1. Then if $M(x^p, y^p, z^p)$ is indeed a representation, its corresponding comodule map is the composition $V \xrightarrow{\rho} V \otimes A \xrightarrow{1\otimes[p]} V \otimes A$, where $[p] : A \to A$ is the k-linear map which carries the monomial $x^r y^s z^t$ to $(x^r y^s z^t)^p$. Thus, we need to check that this new map is indeed a valid A-comodule structure on V, i.e., that it satisfies Diagram (1.1).

Consider

$$\begin{tikzcd} V \arrow[r, "\rho"] \arrow[d, "\rho"'] & V \otimes A \arrow[r, "1\otimes[p]"] \arrow[d, "1\otimes\Delta"] & V \otimes A \arrow[d, "1\otimes\Delta"] \\ V \otimes A \arrow[r, "\rho\otimes 1"'] \arrow[d, "1\otimes[p]"'] & V \otimes A \otimes A \arrow[r, "1\otimes[p]\otimes[p]"] \arrow[d, "1\otimes 1\otimes[p]"] & V \otimes A \otimes A \arrow[d, "1"] \\ V \otimes A \arrow[r, "\rho\otimes 1"'] & V \otimes A \otimes A \arrow[r, "1\otimes[p]\otimes 1"'] & V \otimes A \otimes A. \end{tikzcd}$$

The outermost rectangle is Diagram (1.1) applied to $V \xrightarrow{\rho} V \otimes A \xrightarrow{1\otimes[p]} V \otimes A$, and is what we are trying to prove commutes. Commutativity of the top left square is Diagram (1.1) applied to ρ and commutes by assumption, and commutativity of the bottom left and bottom right squares is obvious. What remains to check is the top right square, i.e.,

$$\begin{tikzcd} A \arrow[r, "[p]"] \arrow[d, "\Delta"'] & A \arrow[d, "\Delta"] \\ A \otimes A \arrow[r, "[p]\otimes[p]"'] & A_{\otimes} A. \end{tikzcd}$$

As $[p]$ and Δ are k-linear, it suffices to check commutativity on the basis of monomials for A. Let $x^r y^s z^t$ be such a monomial, and denote by $F: A \to A$ the Frobenius mapping (that raises *everything* to its pth power, not just the variables). Clearly, $F(x^r y^s z^t) = [p](x^r y^s z^t)$. Further, since $\Delta(x^r y^s z^t)$ is a linear combination of monomial tensors with *integer* coefficients (this is crucial), and since $z^p = z \bmod p$ for any integer, we also have that $(F \otimes F)(\Delta(x^r y^s z^t)) = ([p] \otimes [p])(\Delta(x^r y^s z^t))$. As F is known to satisfy the above diagram in place of $[p]$, this completes the proof. $\square$

Theorem 5.14. *Let k be a field of characteristic $p > 0$, assume that $p \geq 2d$, and let $M(x, y, z)$ be the matrix formula for a d-dimensional representation of H_1 over k given by the matrices $X_0, Y_0, Z_0, X_1, Y_1, Z_1, \ldots, X_m, Y_m, Z_m$, as in Proposition 3.5. Then*

$$M(x, y, z) = e^{xX_0+yY_0+(z-xy/2)Z_0}\, e^{x^pX_1+y^pY_1+(z^p-x^py^p/2)Z_1} \\ \ldots e^{x^{p^m}X_m+y^{p^m}Y_m+(z^{p^m}-x^{p^m}y^{p^m}/2)Z_m}.$$

Proof. We argue first that the right-hand side actually is a representation. By Propositions 3.5 and 5.9, for each j, the formula

$$e^{xX_j+yY_j(z-xy/2)Z_j}$$

defines a representation of H_1. Then by Lemma 5.12 so also does

$$e^{x^{p^j}X_j+y^{p^j}Y_j(z^{p^j}-x^{p^j}y^{p^j}/2)Z_j}$$

define a representation. Finally, since X_i, Y_i, Z_i commute with all of X_j, Y_j, Z_j when $i \neq j$ (Proposition 3.5), the matrix product

$$\begin{aligned} M(x, y, z) = {} & e^{xX_0+yY_0+(z-xy/2)Z_0} e^{x^pX_1+y^pY_1+(z^p-x^py^p/2)Z_1} \\ & \dots e^{x^{p^m}X_M+y^{p^m}Y_m+(z^{p^m}-x^{p^m}y^{p^m}/2)Z_m} \end{aligned}$$

is a commuting product of representations, which is also a representation.

By Lemma 3.3, a representation of H_1 over k is completely determined by the X_i, Y_i, and Z_i. Thus, to prove the theorem it suffices to show that, for the right-hand side, for each j, the matrix of coefficients of the monomial x^{p^j} is actually X_j, that for y^{p^j} it is actually Y_j, and that for z^{p^j} it is Z_j.

For fixed j, examine the individual factor

$$\begin{aligned} e^{x^{p^j}X_j+y^{p^j}Y_j+(z^{p^j}-x^{p^j}y^{p^j}/2)Z_j} = {} & 1 + x^{p^j}X_j + y^{p^j}Y_j + (z^{p^j} - x^{p^j}y^{p^j}/2)Z_j \\ & + (x^{p^j}X_j + y^{p^j}Y_j + (z^{p^j} - x^{p^j}y^{p^j}/2)Z_j)^2 \\ & + \dots \\ & + \frac{1}{(d-1)!}(x^{p^j}X_m + y^{p^j}Y_j + (z^{p^j} - x^{p^j}y^{p^j}/2)Z_j)^{d-1}. \end{aligned}$$

Notice first that, with respect to this representation alone, X_j is indeed the coefficient matrix of the monomial x^{p^j}, Y_j that for y^{p^j}, and Z_j that for z^{p^j} (a similar argument is made in the proof of the previous theorem for the characteristic zero case, i.e., when $j = 0$). But just as importantly, notice also that, since we are assuming $p \geq 2d$, for any $i \neq j$, the monomials x^{p^i}, y^{p^i}, and z^{p^i} never occur in this series expansion of the jth factor. What this says is that there is no possibility of one factor 'spilling over' into another factor. Thus the coefficient matrix of, e.g., x^{p^j}, for the factor $e^{x^{p^j}X_j+y^{p^j}Y_j+(z^{p^j}-x^{p^j}y^{p^j}/2)Z_j}$, which we know to be X_j, is actually the coefficient matrix of x^{p^j} for the *entire* representation

$$e^{xX_0+yY_0+(z-xy/2)Z_0} e^{x^pX_1+y^pY_1+(z^p-x^py^p/2)Z_1} \dots e^{x^{p^m}X_M+y^{p^m}Y_m+(z^{p^m}-x^{p^m}y^{p^m}/2)Z_m}.$$

This completes the proof. $\square$

The main theorem of this article, Theorem 1.3, is now proved.

6. COUNTEREXAMPLES AND SHARPNESS

If $\mathrm{char}(k) = p$ is not large enough with respect to dimension, conditions (1) and (2) of Theorem 1.3 do not necessarily hold. Here we give two such examples.

Let k be the finite field $\mathbb{Z}_2$, and let V be the 10-dimensional sub-coalgebra of the Hopf algebra $A = k[x, y, z]$ given by

$$V = \mathrm{span}_k(1, x, y, z, x^2, xy, xz, y^2, yz, z^2),$$

i.e., the span of all monomials of degree no greater than 2, endowed with the structure of a right A-comodule $\rho : V \to V \otimes A$ given by the restriction of Δ to V. The corresponding representation, in this ordered basis, has matrix formula

$$M = \begin{pmatrix}
1 & x & y & z & x^2 & xy & xz & y^2 & yz & z^2 \\
 & 1 & 0 & y & 0 & y & z+xy & 0 & y^2 & 0 \\
 & & 1 & 0 & 0 & x & 0 & 0 & z & 0 \\
 & & & 1 & 0 & 0 & x & 0 & y & 0 \\
 & & & & 1 & 0 & y & 0 & 0 & y^2 \\
 & & & & & 1 & 0 & 0 & y & 0 \\
 & & & & & & 1 & 0 & 0 & 0 \\
 & & & & & & & 1 & 0 & 0 \\
 & & & & & & & & 1 & 0 \\
 & & & & & & & & & 1
\end{pmatrix}.$$

Let X_0 denote the matrix of coefficients of the monomial x in M, i.e., the 10×10 matrix with $1's$ in its $(1, 2), (3, 6)$, and $(4, 7)$ entries, and $0's$ elsewhere. Similarly define Y_1 as the matrix of coefficients of the monomial y^2. Then one can check by hand that $X_0 Y_1 - Y_1 X_0 \neq 0$, which contradicts condition (2) of Theorem 1.3.

For a counterexample to condition (1) of Theorem 1.3, again let $k = \mathbb{Z}_2$, and consider the span of all monomials of degree no greater than 3

$$V = \mathrm{span}_k(1, x, y, z, x^2, xy, xz, y^2, yz, z^2, x^3, x^2y, x^2z, xy^2, xyz, xz^2, y^3, y^2z, yz^2, z^3).$$

This is a 20-dimensional representation, whose matrix formula would not fit on this page, but the intrepid reader can verify by hand that $[X_1, Y_1] \neq Z_1$.

One may ask the following question: Is the condition $p \leq 2\mathrm{dim}$ sharp? That is, given any pair (p, d) with $p < 2d$, does there exists a d-dimensional representation of H_1 over a field of characteristic p such that at least one of conditions (1) and (2) of Theorem 1.3 do not hold? The answer is no, at least for small dimensions. For example, we claim that any 2-dimensional representation over *any* positive characteristic field necessarily satisfies these conditions; in particular, even when $\mathrm{char}(k) = 2$. As any representation of a unipotent algebraic group is upper-triangular, and since all of X_i, Y_i, and Z_i are nilpotent, we can take them all to be scalar multiples of

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

whence X_i, Y_j, and Z_k obviously all commute for all i, j and k. We claim further that $Z_m = 0$ for all m. To see this, consider

$$\begin{aligned} X_m Y_m &= (c_{ij})^{(p^m,0,0)}(c_{ij})^{(0,p^m,0)} \\ &= Y_m X_m + \left(\sum_{l=1}^{p^m-1} \binom{p^m - l}{0} \binom{p^m - l}{0} \binom{l}{l} Z_{(l)} Y_{(p^m-l)} X_{(p^m-l)} \right) + Z_m \end{aligned}$$

(see the proof of Proposition 3.5) which, since $X_m Y_m = Y_m X_m$, can be written as

$$0 = Z_m + \left(\sum_{l=1}^{p^m-1} Z_{(l)} Y_{(p^m-l)} X_{(p^m-l)} \right)$$

which, in view of Eq. (3.2) can be written as

$$0 = Z_m + \left(\sum_{l=1}^{p^m-1} \Gamma(l)^{-1} \Gamma(p^m - l)^{-2} Z_0^{l_0} \dots Z_{m-1}^{l_{m-1}} Y_0^{r_0} \dots Y_{m-1}^{r_{m-1}} X_0^{r_0} \dots X_{m-1}^{r_{m-1}} \right)$$

where the r_i are the p-digits of $p^m - l$. Let $0 < l < p^m$. Then the sum $(p^m - l) + l$ carries, whence, for some i, l_i, and r_i are both nonzero. Then the summation term

$$Z_0^{l_0} \dots Z_{m-1}^{l_{m-1}} Y_0^{r_0} \dots Y_{m-1}^{r_{m-1}} X_0^{r_0} \dots X_{m-1}^{r_{m-1}}$$

vanishes, since this entire product commutes and $Z_i^{l_i} Y_i^{r_i} = 0$. Thus the entire summation is zero, whence so is Z_m. This shows that every positive characteristic 2-dimensional representation of H_1 satisfies conditions (1) and (2) of Theorem 1.3, even when $\operatorname{char}(k) = 2 < 2$ dimension $= 4$.

7. FURTHER DIRECTIONS

The last paragraph of the last section shows that the condition $p \geq 2d$ is not sharp. However, we suspect that this condition is *asymptotically* sharp, in one of the following senses.

Conjecture 7.1. There exists a prime q such that, for every prime $p \geq q$, there exists a $\frac{p+1}{2}$-dimensional module for H_1 over a field of characteristic p which does satisfy not at least one of conditions (1) and (2) of Theorem 1.3.

Or, perhaps the weaker one, as follows.

Conjecture 7.2. For arbitrarily large primes p, there exists a $\frac{p+1}{2}$-dimensional module for H_1 over a field of characteristic p which does not satisfy at least one of conditions (1) and (2) of Theorem 1.3.

But more importantly, the author strongly suspects that results analogous to Theorems 1.2 and 1.3 for the Additive and Heisenberg groups should apply to a much wider class of unipotent algebraic groups. The author has in fact proved this result for all of the so-called generalized Heisenberg groups (though this is not in print), and at the very least believes that a proof for all of the unipotent upper triangular groups, the most important class of unipotent groups, will soon appear in a sequel.

Perhaps a more interesting question, and one which has not been at all addressed in this article, is not *that* this phenomenon might be true for unipotent groups in general, but rather *why* it might be true of unipotent groups. An algebraic group is unipotent if and only if, over a field of characteristic $p > 0$, it is of exponent a power of p. The author believes that it is exactly this property of unipotent groups which makes this theorem true, along with perhaps a clever appeal to the compactness theorem for first-order logic (but in what language?). The author again hopes that such insights will be forthcoming in a sequel.

ACKNOWLEDGMENTS

The author would like to thank his thesis advisor, Paul Hewitt, under whose direction and advice these results originally appeared. Thanks also to Dave Hemmer, for his thoughtful advice and encouragement, and to Chris Bendel, for his generous reading of this manuscript and thoughtful suggestions.

REFERENCES

[1] Crumley, M. (2010). *Ultraproducts of Tannakian Categories and Generic Representation Theory of Unipotent Algebraic Groups*. Ph.D. dissertation, The University of Toledo, Department of Mathematics.

[2] Nastasescu, S., Dascalescu, C., Raianu, S. (2001). Hopf algebras: An introduction. *Pure and Applied Mathematics*. New York: Marcel Dekker.

[3] Dickson, L. E. (1896/1897). 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.

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

[5] Hall, B. C. (2003). Lie groups, lie algebras, and representations: An elementary introduction. *Graduate Texts in Mathematics*. New York: Springer-Verlag.

[6] Springer, T. A. (1981). Linear algebraic groups. *Progress in Mathematics*. Vol. 9. Boston: Birkhäuser.

[7] Waterhouse, W. C. (1979). Introduction to affine group schemes. *Graduate Texts in Mathematics*. New York: Springer-Verlag.