ISRAEL JOURNAL OF MATHEMATICS 96 (1996), 97–113

GRÖBNER–SHIRSHOV BASES FOR QUANTUM ENVELOPING ALGEBRAS

BY

L. BOKUT'*

Institute of Mathematics, Novosibirsk 630090, Russia
e-mail: bokut@math.nsk.su

AND

P. MALCOLMSON

Department of Mathematics, Wayne State University
Detroit MI 48202, USA
e-mail: petem@math.wayne.edu

ABSTRACT

We give a method for finding Gröbner–Shirshov bases for the quantum enveloping algebras of Drinfel'd and Jimbo, show how the methods can be applied to Kac–Moody algebras, and explicitly find the bases for quantum enveloping algebras of type A_N (for $q^8 \neq 1$).

1. Introduction

Given a free algebra F over a field k and a set of relations $S \subseteq F$, the problem of reducing a given element $f \in F$ with respect to S involves computational difficulties, mainly because the reduction procedure may need to make f more complicated before some relations of S can "take effect". The main problem

* Supported by the Russian Fund of Fundamental Research and the Soros Fund of Fundamental Research. This author wishes to thank the Department of Mathematics of Wayne State University for its hospitality.
Received June 25, 1995

is that there are implied (possibly simpler) relations which are in the ideal (S) generated by S but which are not in S.

In 1962 A. I. Shirshov ([Sh2]) introduced the notion of "composition" of polynomials in a free (even nonassociative) algebra, with the aim of determining from a set of relations a "completed" set which can be used to effectively reduce elements with respect to the relations. For commutative associative algebras this sort of relation set is now referred to as a Gröbner basis ([AL], [BW]). For the case of associative algebras the situation was also formulated by Bergman in [Be], and more recently by Mora ([Mo]) using the term Gröbner basis. We propose in general to add the name of Shirshov to the term.

To describe the compositions of Shirshov, let $k\langle X\rangle$ be a free algebra on a set X and $\langle X\rangle$ the semigroup generated by X. We need to order the monomials $\langle X\rangle$ of the free algebra so as to determine a **leading term** $\overline{f}$ for each element f. An element of $k\langle X\rangle$ will be called **monic** if the leading term has a coefficient of 1 in k. Now if f and g are monic elements of $k\langle X\rangle$ with leading terms $\overline{f}$ and $\overline{g}$, there will be a so-called **composition of intersection** if there are a and b in $\langle X\rangle$ so that $\overline{f}a = b\overline{g} = w$ with total length of $\overline{f}$ larger than that of b. We write $(f,g)_w = fa - bg$ in that case and note that the leading term $\overline{(f,g)_w} < w$. There will be a **composition of inclusion** if there are a and b in $\langle X\rangle$ so that $\overline{f} = a\overline{g}b = w$. We write $(f,g)_w = f - agb$ in that case and again note that the leading term is less than w.

Let us take now some set of relations $S \subseteq k\langle X\rangle$ (which we will assume consists of monic elements). Let us denote by (S) the ideal generated by S in its corresponding ring. If for $f,g \in S$ we have a composition $(f,g)_w = \sum \alpha_i a_i s_i b_i$, with $\alpha_i \in k$, $a_i, b_i \in \langle X\rangle$, $s_i \in S$, with $\overline{a_i s_i b_i} < w$, then we will say that the composition $(f,g)_w$ is **trivial** with respect to S. Otherwise we will need to expand S by including all nontrivial compositions (inductively) to obtain a completion S^c. If S is complete in this sense ($S^c = S$), then Shirshov's Lemma [Sh2] says that any monic element f of (S) has a **reducible** leading term $\overline{f} = a\overline{s}b$, where $s \in S$ and a and b are in $\langle X\rangle$. That Lemma also says that a **linear** basis for the factor algebra $k\langle X\rangle/(S)$ (i.e., as a vector space over k) may be obtained by taking the set of irreducible monomials in $\langle X\rangle$.

The set S will then be referred to as a **Gröbner–Shirshov basis** for the ideal (S). By abusing the definition we may also refer to S as a Gröbner–Shirshov basis for the factor algebra $k\langle X\rangle/(S)$; we remark that S provides a Hilbert-style basis for the *relations* of this algebra, rather than a linear basis for the algebra itself (though the latter may be obtained by taking irreducible monomials as above). Such a Gröbner–Shirshov basis is valuable because if S is finite (or more generally has only finitely many elements with leading terms below any given monomial), then for any $\phi \in k\langle X\rangle$ we can determine by examining only a finite sequence of leading terms whether $\phi \in (S)$.

This method (for Lie algebras) has been used to determine solvability of some word problems ([Sh2], [Sh3], [Bo2]), to prove some embedding theorems ([Bo1], [Bo4]), and more recently to give Gröbner–Shirshov bases for some finite-dimensional simple Lie algebras ([BoKl1], [BoKl2]). For associative algebras it has been used for some embedding theorems ([Bo3]) and to construct a linear basis for free metabelian algebras ([BML]). In this paper we want to indicate how to obtain a Gröbner–Shirshov basis for a quantum enveloping algebra and to find one explicitly for the case A_N.

To describe the quantum enveloping algebras of Drinfel'd ([Dr]) and Jimbo ([Ji1]), let k be a field and $A = (a_{ij})$ an (integral) symmetrizable N-by-N Cartan matrix, so that $a_{ii} = 2$, $a_{ij} \le 0$ $(i \ne j)$, and there exists a diagonal matrix D with diagonal entries d_i nonzero integers such that the product DA is symmetric. Let q be a nonzero element of k so that $q^{4d_i} \ne 1$ for each i. The quantum enveloping algebra $U_q(A)$ (as in [Ya]) is generated by $4N$ elements e_i, $k_i^{\pm 1}$, f_i, subject to the following sets of relations (for $1 \le i, j \le N$):

$$\begin{aligned} K &= \{k_ik_j - k_jk_i,\ k_ik_i^{-1} - 1,\ k_i^{-1}k_i - 1,\\ &\qquad e_jk_i^{\pm 1} - q^{\pm d_ia_{ij}}k_i^{\pm 1}e_j,\ k_i^{\pm 1}f_j - q^{\pm d_ia_{ij}}f_jk_i^{\pm 1}\},\\ T &= \left\{e_if_j - f_je_i - \delta_{ij}\frac{k_i^2 - k_i^{-2}}{q^{2d_i} - q^{-2d_i}}\right\}, \end{aligned}$$

$$S^+ = \left\{\sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t e_i^{1-a_{ij}-\nu}e_je_i^\nu,\ \text{where } i \ne j,\ t = q^{2d_i}\right\},$$

and

$$S^- = \left\{\sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t f_i^{1-a_{ij}-\nu}f_jf_i^\nu,\ \text{where } i \ne j,\ t = q^{2d_i}\right\},$$

where also we use the quantum binomial

$$\begin{bmatrix} m \\ n \end{bmatrix}_t = \begin{cases} \prod_{i=1}^n \dfrac{t^{m-i+1} - t^{i-m-1}}{t^i - t^{-i}} & (\text{for } m > n > 0),\\ 1 & (\text{for } n = 0 \text{ or } m = n). \end{cases}$$

To describe this as a factor algebra of a free algebra, let us use generators x, h, y in place of e, k, f, respectively. The free algebra will be $k\langle Z\rangle$ with $Z = X \cup H \cup Y$ and $X = \{x_i\}$, $H = \{h_i^{\pm 1}\}$, $Y = \{y_i\}$. There will be relation sets K, T, S^+, S^- (in the new variables) as above.

We will prove (Theorem 2.7) that a Gröbner–Shirshov basis for $U_q(A)$ (for $q^{4d_i} \neq 1$ as above) is given by $K \cup T \cup S^{+c} \cup S^{-c}$, where $S^{+c} = (S^+)^c$ is the completion of the set S^+ in the variables x_i alone, and similarly for S^{-c}. We will then digress somewhat to describe how the general results below can be used to give an analogous result for Kac–Moody algebras. Then for the case $A = A_N$ we explicitly give the elements of $S^{\pm c}$ and so recover the same linear bases given by Rosso and Yamane ([Ro1], [Ya]). (For cases beyond A_N (e.g., $B_N, C_N, \ldots$) identities are more complicated and the compositions of Shirshov have yet to be worked out.)

2. General results

Take $F = k\langle Z\rangle$ with $Z = X \cup H \cup Y$ as above, and $\langle Z\rangle$ the set of monomials on Z. For $u \in \langle Z\rangle$ define the X-degree $\deg_X(u)$ of u by taking the X-degree of each x_i to be 1. Similarly we can define the Y-degree $\deg_Y(u)$ and the H-degree (where $\deg_H(h_i^{\pm 1}) = 1$ still). Then the total degree $\mathrm{Deg}(u)$ will be the triple $(\deg_X(u), \deg_H(u), \deg_Y(u))$ with lexicographic ordering.

We order the generators by $x_i > x_j$, $h_i > h_i^{-1} > h_j > h_j^{-1}$, and $y_i > y_j$ if $i > j$, with $x_i > h_j^{\pm 1} > y_m$ for all i, j, m. Then we order the monomials by $u > v$ if (1) $\mathrm{Deg}(u) > \mathrm{Deg}(v)$, or if (2) $\mathrm{Deg}(u) = \mathrm{Deg}(v)$ but u is greater than v in simple lexicographic ordering (for $u, v \in \langle Z\rangle$). This makes $\langle Z\rangle$ a fully ordered semigroup and for any $f \in F$ determines the leading term $\overline{f}$ and therefore $\mathrm{Deg}(f) = \mathrm{Deg}(\overline{f})$. We can write $f = \alpha_{\overline{f}}\overline{f} + f'$, with $0 \neq \alpha_{\overline{f}} \in k$ and $f' = \sum \alpha_u u$ with $u < \overline{f}$. Again, f is monic if $\alpha_{\overline{f}} = 1$.

For $f, g \in F$, $S \subseteq F$ (possible relations) and $w \in \langle Z\rangle$, let us write $f \equiv g \bmod(S, w)$ to mean that $f - g = \sum \alpha_i a_i s_i b_i$ where $\alpha_i \in k$, $a_i, b_i \in \langle Z\rangle$ and $s_i \in S$ such that $\overline{a_i s_i b_i} < w$. We will write $f \equiv g \mod(S, \deg_X = n)$ to mean the same form for $f - g$ but now requiring $\deg_X(a_i s_i b_i) < n$. Note that to say S is complete is to say that $(f, g)_w \equiv 0 \mod(S, w)$ for every $f, g \in S$.

Finally for purposes of generality let $T' = \{t_{ij} = x_i y_j - y_j x_i - \delta_{ij} p_i \mid i, j \leq N\}$ for any choice of $p_i \in k\langle H\rangle$. We will use a **differential substitution** $\partial_i = \partial(x_i \to p_i)$ defined on $k\langle X\rangle$ by $\partial_i(\alpha) = 0$ for $\alpha \in k$, $\partial_i(u x_j) = \partial_i(u) x_j + \delta_{ij} u p_i$

for $u \in \langle X \rangle$, and extending linearly to $k\langle X \rangle$.

LEMMA 2.1: *Let $f \in k\langle X \rangle$, $t_{ji} \in T'$. Then*

$$(f, t_{ji})_w \equiv \partial_i(f) \mod(\{f\} \cup T', w).$$

Proof: The only possible composition is that of intersection, since $f \in k\langle X \rangle$. So let $w = \overline{f} y_i = b x_j y_i$, where $f = \overline{f} + f'$ with $\overline{f} = b x_j$ and f' has only lower terms. First note that if $u = u_1 x_m u_2$ with $u_1, u_2 \in \langle X \rangle$ and $u < \overline{f}$, then $u_1 x_m y_i u_2 \equiv u_1(y_i x_m + \delta_{im} p_i) u_2 \mod(T', w)$ since $u_1 x_m y_i u_2 \leq u_1 x_m u_2 y_i < w$. This allows the y_i to be inductively moved forward in the last congruence in $(f, t_{ji})_w = f y_i - b(x_j y_i - y_i x_j - \delta_{ij} p_i) = f' y_i + b y_i x_j + \delta_{ij} b p_i \equiv y_i f + \partial_i(f) \mod(T', w)$. Now the lemma follows since $y_i \overline{f} < w$. ∎

LEMMA 2.2: *Let $f \in k\langle X \rangle$, L any subset of $k\langle Z \rangle$. Suppose that*

$$\partial_i(f) \equiv 0 \mod(L, \deg_X = \deg_X f). \tag{2.2}$$

Then for all $t_{ji} \in T'$ we have $(f, t_{ji})_w \equiv 0 \mod(\{f\} \cup T' \cup L, w)$.

Proof: By induction each term in $\partial_i(f)$ has X-degree less than $\deg_X(f) = \deg_X(w)$, so each term is $< w$ and this follows from Lemma 2.1. ∎

Again for generality purposes, let S be any set of homogeneous polynomials in $k\langle X \rangle$, with L and T' as above.

LEMMA 2.3: *Let L be any subset of $k\langle Z \rangle$ and suppose that for all $f \in S$ we have (2.2). Then for all compositions $(f, g)_u$ of elements $f, g \in S$ we have*

$$\partial_i((f, g)_u) \equiv 0 \mod(\{f, g\} \cup L, \deg_X = \deg_X((f, g)_u)). \tag{2.3}$$

Proof: We have two kinds of composition. If $(f, g)_u = fa - bg \neq 0$ with $u = \overline{f} a = b \overline{g}$, then by homogeneity all terms have the same X-degree $\deg_X u = \deg_X(fa) = \deg_X((f, g)_u)$. So then $\partial_i((f, g)_u) = \partial_i(f) a - b\partial_i(g) + f \partial_i(a) - \partial(b) g$. By (2.2) we have $\partial_i(f) a \equiv 0 \mod(L, \deg_X u = \deg_X(fa))$ and similarly for g, taking care of the first two terms. The other two are $\equiv 0 \mod(\{f, g\}, \deg_X = \deg_X u)$.

For the other composition $(f, g)_u = f - agb (\neq 0)$ with $u = \overline{f} = a \overline{g} b$. Again by homogeneity $\deg_X f = \deg_X(agb) = \deg_X((f, g)_u)$. Now a similar argument works in this case. ∎

We remark that to form the completion M^c of some set $M \subseteq k\langle Z\rangle$, it is sufficient to form $M^{(0)} = M$,

$$M^{(i)} = M^{(i-1)} \cup \{(f,g)_w \text{ nontrivial} \mid f,g \in M^{(i-1)}\},$$

and $M^c = \bigcup M^{(i)}$. As above the completion of some homogeneous set M of elements in $k\langle X\rangle$ will contain only homogeneous elements of $k\langle X\rangle$. In fact from the proof above it is easy to see that if every element of M is homogeneous in every variable $x_i \in X$, then any element of M^c will also be homogeneous in every variable x_i.

LEMMA 2.4: *Let S_0 be a homogeneous set of polynomials in $k\langle X\rangle$. Let L_0 be any subset of $k\langle Z\rangle$ and suppose that for all $f \in S_0$ we have (2.2) with $L = L_0$. Then for all $\phi \in S_0^c$ we have*

$$\partial_i(\phi) \equiv 0 \mod(S_0^c \cup L_0, \deg_X = \deg_X \phi). \tag{2.4}$$

Proof: For $\phi \in S_0^{(1)}$ the conclusion follows from Lemma 2.3 with $S = S_0$. Now for $\phi \in S_0^{(2)}$ we enlarge L_0 to $L_1 = L_0 \cup S_0^{(1)}$ and apply Lemma 2.3 with $S = S_0^{(1)}$ and $L = L_1$. Inductively for $\phi \in S_0^{(i)}$ we enlarge to $L_{i-1} = L_{i-2} \cup S_0^{(i-1)}$ and apply Lemma 2.3 with $S = S_0^{(i-1)}$ and $L = L_{i-1}$. This completes the proof since S_0^c is the union and $\bigcup L_i = S_0^c \cup L_0$. ∎

PROPOSITION 2.5: *Let T' be as above, S a set of homogeneous elements of $k\langle X\rangle$, and $L \subseteq k\langle Z\rangle$ as before. Suppose that for any $f \in S$ we have $\partial_i(f) \equiv 0 \mod(L, \deg_X = \deg_X f)$. Then for any $\phi \in S^c$ we have $(\phi, t_{ji})_w \equiv 0 \mod(S^c \cup T' \cup L, w)$.*

Proof: By Lemma 2.1 we have $(\phi, t_{ji})_w \equiv \partial_i(\phi) \mod(S^c \cup T', w)$. Using Lemma 2.4 we get the condition (2.2) with $S^c \cup L$ replacing L. Then Lemma 2.2 implies that $(\phi, t_{ji})_w \equiv 0 \mod(S^c \cup T' \cup L, w)$. ∎

Now for T' as before, we can define differential substitutions $\partial_i = \partial(y_i \to p_i)$ on $k\langle Y\rangle$ analogously as before. There will be corresponding results (2.1′)–(2.5′), which we will state because of the nonsymmetry of the ordering.

LEMMA 2.1′: *Let $f \in k\langle Y\rangle$, $t_{ij} \in T'$. Then*

$$(t_{ij}, f)_w \equiv \partial_i(f) \mod(\{f\} \cup T', w).$$

Here $w = x_i\overline{f}$ and $\deg_X w = 1$ but $\deg_X(\partial_i(f)) = 0$.

LEMMA 2.2′: *Let $f \in k\langle Y\rangle$, L any subset of $k\langle Z\rangle$. Suppose that*

$$\partial_i(f) \equiv 0 \bmod(L, \deg_X = 1). \tag{2.2′}$$

Then for all $t_{ij} \in T'$ we have

$$(t_{ij}, f)_w \equiv 0 \ \bmod(\{f\} \cup T' \cup L, w).$$

LEMMA 2.3′: *Let L be any subset of $k\langle Z\rangle$ and suppose that for all $f \in S$ we have (2.2′). Then for all compositions $(f,g)_u$ of elements $f,g \in S$ we have*

$$\partial_i((f,g)_u) \equiv 0 \bmod(\{f,g\} \cup L, \deg_X = 1). \tag{2.3′}$$

LEMMA 2.4′: *Let S_0 be a homogeneous set of polynomials in $k\langle Y\rangle$. Let L_0 be any subset of $k\langle Z\rangle$ and suppose that for all $f \in S_0$ we have (2.2′) with $L = L_0$. Then for all $\phi \in S_0^c$ we have*

$$\partial_i(\phi) \equiv 0 \ \bmod(S_0^c \cup L_0, \deg_X = 1). \tag{2.4′}$$

LEMMA 2.5′: *Let T' be as above, S a set of homogeneous elements of $k\langle Y\rangle$, and $L \subseteq k\langle Z\rangle$ as before. Suppose that for any $f \in S$ we have*

$$\partial_i(f) \equiv 0 \bmod(L, \deg_X = 1).$$

Then for any $\phi \in S^c$ we have

$$(t_{ij}, \phi)_w \equiv 0 \bmod(S^c \cup T' \cup L, w).$$

Now we return to the relation sets S^+, T, K, S^- of the introduction. We will want to prove the hypotheses of Lemmas 2.5. After we do that the set $S^{+c} \cup T \cup S^{-c}$ will have no nontrivial compositions (since there are no compositions between S^{+c} and S^{-c}).

LEMMA 2.6: *Let $f \in S^+$. Then for any l $(1 \le l \le N)$ we have*

$$\partial_l(f) \equiv 0 \ \bmod(K, \deg_X = \deg_X f).$$

Proof: Let

$$f = \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t x_i^{1-a_{ij}-\nu} x_j x_i^\nu,$$

with $t = q^{2d_i}$ and let $r_i = (t - t^{-1})^{-1}$. Of course here $p_l = (h_l^2 - h_l^{-2})r_i$ in $\partial_l = \partial(x_l \to p_l)$. The only two cases to compute are $l = i$ and $l = j$.

For $l = i$ we have

$$\begin{aligned}
\partial_i(f) =& r_i \sum_{\nu=1}^{1-a_{ij}} \sum_{\mu=0}^{\nu-1} (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-\mu-1} (h_i^2 - h_i^{-2}) x_i^\mu \\
&+ r_i \sum_{\nu=0}^{-a_{ij}} \sum_{\lambda=0}^{-a_{ij}-\nu} (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t x_i^{-a_{ij}-\nu-\lambda} (h_i^2 - h_i^{-2}) x_i^\lambda x_j x_i^\nu \\
\equiv& r_i \left(\sum_{\nu=1}^{1-a_{ij}} \sum_{\mu=0}^{\nu-1} (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t \right. \\
&\times \left. (q^{d_i(4\mu+2a_{ij})} h_i^2 - q^{-d_i(4\mu+2a_{ij})} h_i^{-2}) x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1} \right) \\
&+ r_i \left(\sum_{\nu=0}^{-a_{ij}} \sum_{\lambda=0}^{-a_{ij}-\nu} (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t \right. \\
&\times \left. (q^{4d_i(a_{ij}+\nu+\lambda)} h_i^2 - q^{-4d_i(a_{ij}+\nu+\lambda)} h_i^{-2}) x_i^{-a_{ij}-\nu} x_j x_i^\nu \right)
\end{aligned}$$

where the congruence is $\mathrm{mod}(K, \deg_X = \deg_X f)$. The coefficient of $r_i h_i^2 x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1}$ is

$$\begin{aligned}
&(-1)^\nu \sum_{\mu=0}^{\nu-1} \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t q^{4d_i\mu} q^{2d_i a_{ij}} \\
&+ (-1)^{\nu-1} \sum_{\lambda=0}^{1-a_{ij}-\nu} \left[\begin{array}{c} 1-a_{ij} \\ \nu-1 \end{array} \right]_t q^{4d_i\lambda} q^{4d_i(a_{ij}+\nu-1)} \\
=& (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu \end{array} \right]_t \frac{q^{4d_i\nu} - 1}{q^{4d_i} - 1} q^{2d_i a_{ij}} \\
&- (-1)^\nu \left[\begin{array}{c} 1-a_{ij} \\ \nu-1 \end{array} \right]_t \frac{q^{-4d_i(a_{ij}+\nu-2)} - 1}{q^{4d_i} - 1} q^{4d_i(a_{ij}+\nu-1)} \\
=& \frac{(-1)^\nu}{q^{4d_i} - 1} \left[\begin{array}{c} 1-a_{ij} \\ \nu-1 \end{array} \right]_t \left(\frac{q^{2d_i(2-a_{ij}-\nu)} - q^{2d_i(a_{ij}+\nu-2)}}{q^{2d_i\nu} - q^{-2d_i\nu}} \right. \\
&\qquad \times \left. q^{2d_i a_{ij}} (q^{4d_i\nu} - 1) - (q^{4d_i} - q^{4d_i(a_{ij}+\nu-1)}) \right) = 0.
\end{aligned}$$

The coefficient of $r_i h_i^{-2} x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1}$ is

$$\begin{aligned}
&-(-1)^{\nu}\begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t \frac{q^{-4d_i\nu}-1}{q^{-4d_i}-1} q^{-2d_i a_{ij}} \\
&+(-1)^{\nu}\begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix}_t \frac{q^{4d_i(a_{ij}+\nu-2)}-1}{q^{-4d_i}-1} q^{-4d_i(a_{ij}+\nu-1)} \\
=&\frac{(-1)^{\nu}}{q^{-4d_i}-1}\begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix}_t \Bigg(-\frac{q^{2d_i(2-a_{ij}-\nu)}-q^{2d_i(a_{ij}+\nu-2)}}{q^{2d_i\nu}-q^{-2d_i\nu}} \\
&\times q^{-2d_i a_{ij}}(q^{4d_i\nu}-1)+(q^{-4d_i}-q^{-4d_i(a_{ij}+\nu-1)})\Bigg)=0.
\end{aligned}$$

For $l=j$ we have

$$\begin{aligned}
\partial_j(f) &= r_i \sum_{\nu=0}^{1-a_{ij}} (-1)^{\nu}\begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t x_i^{1-a_{ij}-\nu}(h_j^2-h_j^{-2})x_i^{\nu} \\
&\equiv r_i \sum_{\nu=0}^{1-a_{ij}} (-1)^{\nu}\begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix}_t x_i^{1-a_{ij}}(q^{2d_j a_{ji}\nu}h_j^2-q^{-2d_j a_{ji}\nu}h_j^{-2})
\end{aligned}$$

where the congruence is $\mathrm{mod}(K, \deg_X = \deg_X f)$. This last expression is zero because we have $d_j a_{ji} = d_i a_{ij}$ and also we have the general equality

$$\sum_{n=0}^{m}(-1)^n \begin{bmatrix} m \\ n \end{bmatrix}_t t^{\pm(m-1)n} = 0 \quad \text{for } m \geq 1.$$

It is easy to verify this by induction by using the equation

$$\begin{bmatrix} m \\ n \end{bmatrix}_t = t^{\pm(m-n)}\begin{bmatrix} m-1 \\ n-1 \end{bmatrix}_t + t^{\mp n}\begin{bmatrix} m-1 \\ n \end{bmatrix}_t \qquad \text{for } m > n > 1.$$

This completes the proof of Lemma 2.6. ∎

Analogously we have

LEMMA 2.6′: *Let $f \in S^-$. Then for any l ($1 \leq l \leq N$) we have*

$$\partial_l(f) \equiv 0 \ \mathrm{mod}(K, \deg_X = 1).$$

THEOREM 2.7: *Let A be a generalized Cartan matrix, $S(A) = S^+ \cup K \cup T \cup S^-$ be the Drinfel'd–Jimbo relations of $U_q(A)$. Then $S^{+c} \cup K \cup T \cup S^{-c}$ is a Gröbner–Shirshov basis for $U_q(A)$.*

Proof: There are no compositions between S^{+c} and S^{-c}, between elements of T and K, nor among elements of T. In view of Lemmas 2.5 and 2.6 (unprimed and primed) there are no nontrivial compositions between T and $S^{\pm c}$. By construction there are no nontrivial compositions among elements of $S^{\pm c}$. It is easy to check that compositions among elements of K are trivial using $K \cup T$. For example, for $f = x_j h_i^{\pm 1} - q^{\pm d_i a_{ij}} h_i^{\pm 1} x_j$, $g = h_i^{\pm 1} y_l - q^{\pm d_i a_{il}} y_l k_i^{\pm 1}$, and $w = x_j h_i^{\pm 1} y_l$ we have

$$\begin{aligned}(f,g)_w &= q^{\pm d_i a_{il}} x_j y_l h_i^{\pm 1} - q^{\pm d_i a_{ij}} h_i^{\pm 1} x_j y_l \\ &\equiv q^{\pm d_i a_{il}} y_l x_j h_i^{\pm 1} - q^{\pm d_i a_{ij}} h_i^{\pm 1} y_l x_j \\ &\equiv q^{\pm d_i (a_{ij}+a_{il})}(y_l h_i^{\pm 1} x_j - y_l h_i^{\pm 1} x_j) \\ &\equiv 0 \operatorname{mod}(K \cup T, w).\end{aligned}$$

To complete the proof, consider any $\phi \in S^{+c} \subseteq k\langle X\rangle$ and its composition with an element $g_{ji} = x_j h_i^{\pm 1} - q^{\pm d_i a_{ij}} h_i^{\pm 1} x_j$ of K with $w = \overline{\phi} h_i^{\pm 1}$. Now ϕ is homogeneous in each variable x_i by the remark following Lemma 2.3. Thus we may use the various $g_{li} \in K$ to "push" the $h_i^{\pm 1}$ in $(\phi, g_{ji})_w$ to the left, and the eventual result will have a coefficient with the same power of q on each term by homogeneity. Thus we conclude $(\phi, g_{ji})_w \equiv 0 \operatorname{mod}(S^{+c} \cup K, w)$. The analogous result for S^{-c} completes the proof. ∎

Because of the Shirshov lemma ([Sh2]) we get the general result on bases.

COROLLARY 2.8 ([Ya], [Ro2]): *For any A we have*

$$U_q(A) = k\langle Y\rangle/(S^-) \otimes k[H] \otimes k\langle X\rangle/(S^+)$$

as a k-space.

3. Digression into Kac–Moody algebras

We take a short digression here to demonstrate that the previous lemmas may be applied to the classical (nonquantized) case of Kac–Moody algebras. Recall here that, as above, $A = (a_{ij})$ is a symmetrizable Cartan matrix. Then the universal enveloping algebra $U(A)$ of the Kac–Moody algebra $\mathcal{G}(A)$ is given by free generators $Z = X \cup H \cup Y$ and relations as follows (for $1 \le i,j \le N$):

$$K = \{h_i h_j - h_j h_i,\ x_j h_i - h_i x_j + d_i a_{ij} x_i,\ h_i y_j - y_j h_i + d_i a_{ij} y_j\},$$

$$T = \{x_i y_j - y_j x_i - \delta_{ij} h_i\},$$

$$S^+ = \left\{ \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} x_i^{1-a_{ij}-\nu} x_j x_i^\nu, \text{ where } i \ne j \right\},$$

$$S^- = \left\{ \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} y_i^{1-a_{ij}-\nu} y_j y_i^\nu, \text{ where } i \neq j \right\},$$

and here we use ordinary binomial coefficients ([Ka], (0.3.1)).

To give a Gröbner–Shirshov basis for $U(A)$ we want to prove that $S^{+c} \cup K \cup T \cup S^{-c}$ is complete under compositions. This will follow the same outline as above.

Let us choose the ordering as before, and define differential substitutions $\partial_i = \partial(x_i \to h_i)$ as before.

LEMMA 3.1: *Let $f \in S^+$. Then for any l $(1 \leq l \leq N)$ we have*

$$\partial_l(f) \equiv 0 \operatorname{mod}(K, \deg_X = \deg_X f).$$

Proof: Let

$$f = \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} x_i^{1-a_{ij}-\nu} x_j x_i^\nu.$$

The only two cases to compute are $l = i$ and $l = j$.

For $l = i$ we have

$$\begin{aligned}
\partial_i(f) = &\sum_{\nu=1}^{1-a_{ij}} \sum_{\mu=0}^{\nu-1} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-\mu-1} h_i x_i^\mu \\
&+ \sum_{\nu=0}^{-a_{ij}} \sum_{\lambda=0}^{-a_{ij}-\nu} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} x_i^{-a_{ij}-\nu-\lambda} h_i x_i^\lambda x_j x_i^\nu \\
\equiv &\left(\sum_{\nu=1}^{1-a_{ij}} \sum_{\mu=0}^{\nu-1} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} \right. \\
&\left. \times\ (h_i - d_i(2(-a_{ij} - \mu) + a_{ij})) x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1} \right) \\
&+ \left(\sum_{\nu=0}^{-a_{ij}} \sum_{\lambda=0}^{-a_{ij}-\nu} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} \right. \\
&\left. \times\ (h_i - 2d_i(-a_{ij} - \nu - \lambda)) x_i^{-a_{ij}-\nu} x_j x_i^\nu \right)
\end{aligned}$$

where the congruence is $\operatorname{mod}(K, \deg_X = \deg_X f)$. The coefficient of

$h_i x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1}$ is

$$(-1)^\nu \sum_{\mu=0}^{\nu-1} \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} + (-1)^{\nu-1} \sum_{\lambda=0}^{1-a_{ij}-\nu} \begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix}$$

$$=(-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} \nu - (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix} (2-a_{ij}-\nu)$$

$$=0.$$

The coefficient of $d_i x_i^{1-a_{ij}-\nu} x_j x_i^{\nu-1}$ is

$$-(-1)^\nu \sum_{\mu=0}^{\nu-1} \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} (-a_{ij}-2\mu)$$

$$-(-1)^{\nu-1} \sum_{\lambda=0}^{1-a_{ij}-\nu} \begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix} 2(1-a_{ij}-\nu-\lambda)$$

$$=(-1)^{\nu+1} \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} \frac{\nu(-2a_{ij}-2\nu+2)}{2}$$

$$+(-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu-1 \end{bmatrix} \frac{(2-a_{ij}-\nu)2(1-a_{ij}-\nu)}{2}$$

$$=0.$$

For $l=j$ we have

$$\partial_j(f) = \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} x_i^{1-a_{ij}-\nu} h_j x_i^\nu$$

$$\equiv \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} (h_j - 2d_j a_{ji}(1-a_{ij}-\nu)) x_i^{1-a_{ij}}$$

where the congruence is $\mathrm{mod}(K, \deg_X = \deg_X f)$. Clearly the coefficient of $h_j x_i^{1-a_{ij}}$ is zero, while for the coefficient of $x_i^{1-a_{ij}}$ we get

$$\sum_{\nu=0}^{1-a_{ij}} -(-1)^\nu \begin{bmatrix} 1-a_{ij} \\ \nu \end{bmatrix} 2d_j a_{ji}(1-a_{ij}-\nu)$$

$$=-2d_j a_{ji}(1-a_{ij}) \sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{bmatrix} -a_{ij} \\ \nu \end{bmatrix}$$

$$=0.$$

This completes the proof of Lemma 3.1. ∎

Analogously for $\partial_i = \partial(y_i \to h_i)$ we have

LEMMA 3.1′: *Let $f \in S^-$. Then for any l $(1 \le l \le N)$ we have*

$$\partial_l(f) \equiv 0 \ \mathrm{mod}(K, \deg_X = 1).$$

Now just as for the quantum case the earlier Lemmas (from both sections) will imply

THEOREM 3.2: *Let A be a generalized Cartan matrix, $U(A)$ the universal enveloping algebra of the Kac–Moody algebra $\mathcal{G}(A)$ and $S(A) = S^+ \cup K \cup T \cup S^-$ the relations of $U(A)$ as above. Then $S^{+c} \cup K \cup T \cup S^{-c}$ is a Gröbner–Shirshov basis for $U(A)$.*

COROLLARY 3.3 ([Ka]): *For any A we have*

$$U(A) = U^-(A) \otimes k[H] \otimes U^+(A)$$

as a k-space. Equivalently, as k-spaces

$$\mathcal{G}(A) = \mathcal{G}^-(A) \otimes H \otimes \mathcal{G}^+(A).$$

4. The case of A_N

Let $q^8 \neq 1$ and

$$A = A_N = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix}.$$

We want to find a Gröbner–Shirshov basis for $U_q^+(A_N) = k\langle X\rangle/(S^+)$, in the notation of Section 2. To more easily describe these we introduce some new variables (defined by Jimbo [Ji2], or see [Ya]) $\tilde{X} = \{x_{ij} \mid 1 \le i < j \le N+1\}$ which generate $U_q^+(A_N)$ and are to be interpreted as being defined inductively by

$$x_{ij} = \begin{cases} x_i, & j = i+1, \\ q x_{i,j-1} x_{j-1,j} - q^{-1} x_{j-1,j} x_{i,j-1}, & j > i+1. \end{cases}$$

The set $\tilde{S}^+$ of relations on these will be the Jimbo relations ((1) in Section 3 of [Ya]), which are indexed by the ways in which one pair can be less than another.

For this we recall from [Ya] the notation:

$$\begin{aligned}
C_1 =&\{((i,j),(m,n)) \mid i = m < j < n\},\\
C_2 =&\{((i,j),(m,n)) \mid i < m < n < j\},\\
C_3 =&\{((i,j),(m,n)) \mid i < m < j = n\},\\
C_4 =&\{((i,j),(m,n)) \mid i < m < j < n\},\\
C_5 =&\{((i,j),(m,n)) \mid i < j = m < n\},\\
C_6 =&\{((i,j),(m,n)) \mid i < j < m < n\}.
\end{aligned}$$

Then the set $\tilde{S}^+$ of Jimbo relations consists of:

$$\begin{aligned}
&x_{mn}x_{ij} - q^{-2}x_{ij}x_{mn} && \text{if } ((i,j),(m,n)) \in C_1 \cup C_3,\\
&x_{mn}x_{ij} - x_{ij}x_{mn} && \text{if } ((i,j),(m,n)) \in C_2 \cup C_6,\\
&x_{mn}x_{ij} - x_{ij}x_{mn} + (q^2 - q^{-2})x_{in}x_{mj} && \text{if } ((i,j),(m,n)) \in C_4,\\
&x_{mn}x_{ij} - q^2x_{ij}x_{mn} + qx_{in} && \text{if } ((i,j),(m,n)) \in C_5.
\end{aligned}$$

Let F^+ be the free algebra on $\tilde{X}$ and order the variables by $x_{ij} < x_{mn}$ if $(i,j) < (m,n)$ (lexicographically). As before we order the monomials $\langle\tilde{X}\rangle$ by length first, and lexicographically for words of the same length. We will obtain:

THEOREM 4.1: *For $q^8 \neq 1$ the set $\tilde{S}^+$ of Jimbo relations is a Gröbner–Shirshov basis for $U_q^+(A_N) = F^+/(\tilde{S}^+)$.*

Thus $\tilde{S}^+$ can be interpreted (in the new variables) as the completion S^{+c} of the original S^+. Because of the composition lemma we then get:

COROLLARY 4.2 ([Ya]): *For $q^8 \neq 1$ a linear basis for the algebra $U_q^+(A_N)$ consists of elements $x_{i_1j_1}x_{i_2j_2}\cdots x_{i_kj_k}$, with $(i_1,j_1) \leq (i_2,j_2) \leq \cdots \leq (i_k,j_k)$ and $k \geq 0$.*

Using similar relations for similar y_{ij}, together with Corollary 2.8, we obtain:

COROLLARY 4.3 ([Ya]): *For $q^8 \neq 1$ a linear basis for the algebra $U_q(A_N)$ consists of elements*

$$y_{m_1n_1}y_{m_2n_2}\cdots y_{m_ln_l}h_1^{s_1}h_2^{s_2}\cdots h_N^{s_N}x_{i_1j_1}x_{i_2j_2}\cdots x_{i_kj_k},$$

with $(m_1,n_1) \leq (m_2,n_2) \leq \cdots \leq (m_l,n_l)$, $(i_1,j_1) \leq (i_2,j_2) \leq \cdots \leq (i_k,j_k)$, $k,l \geq 0$ and each $s_i \in \mathbf{Z}$.

For the proof of Theorem 4.1 we need to check triviality of all compositions of elements of $\tilde{S}^+$. In any such composition of intersection $(f,g)_w$ we will have

$w = x_{mn}x_{ij}x_{kl}$, $\overline{f} = x_{mn}x_{ij}$ and $\overline{g} = x_{ij}x_{kl}$, with $(k,l) < (i,j) < (m,n)$. Since there are so many cases we adopt the notation from [Ya], that

$$f = x_{mn}x_{ij} - \varepsilon_{ijmn}x_{ij}x_{mn} + y_{ijmn}$$

(and similarly for g), where

$$\varepsilon_{ijmn} = \begin{cases} 1 & \text{for } ((i,j),(m,n)) \in C_2 \cup C_4 \cup C_6, \\ q^{-2} & \text{for } ((i,j),(m,n)) \in C_1 \cup C_3, \\ q^2 & \text{for } ((i,j),(m,n)) \in C_5; \end{cases}$$

$$y_{ijmn} = \begin{cases} 0 & \text{for } ((i,j),(m,n)) \in C_1 \cup C_2 \cup C_3 \cup C_6, \\ (q^2 - q^{-2})x_{in}x_{mj} & \text{for } ((i,j),(m,n)) \in C_4, \\ qx_{in} & \text{for } ((i,j),(m,n)) \in C_5. \end{cases}$$

Then

$$\begin{aligned}(f,g)_w = &- \varepsilon_{ijmn}x_{ij}x_{mn}x_{kl} + y_{ijmn}x_{kl} + \varepsilon_{klij}x_{mn}x_{kl}x_{ij} - x_{mn}y_{klij} \\ \equiv &\varepsilon_{ijmn}\varepsilon_{klmn}y_{klij}x_{mn} + \varepsilon_{ijmn}x_{ij}y_{klmn} + y_{ijmn}x_{kl} \\ &- \varepsilon_{klij}\varepsilon_{klmn}x_{kl}y_{ijmn} - \varepsilon_{klij}y_{klmn}x_{ij} - x_{mn}y_{klij},\end{aligned}$$

where the congruence is $\mathrm{mod}(\tilde{S}^+, w)$.

This is the same as the polynomial in (4.5) of [Ya] (with $z_\Psi = 1$). We may therefore appeal to his proof to conclude the argument for Theorem 4.1. ∎

References

[AL] W. Adams and P. Loustaunou, *An Introduction to Gröbner Bases*, Graduate Studies in Mathematics Vol. 3, American Mathematical Society, 1994. MR 95g:13025.

[Be] G. M. Bergman, *The diamond lemma for ring theory*, Advances in Mathematics **29** (1978), 178–218. MR 81b:16001.

[Bo1] L. A. Bokut', *Imbedding of Lie algebras into algebraically closed Lie algebras*, Algebra and Logic **1** (1962), 47–53.

[Bo2] L. A. Bokut', *Unsolvability of the equality problem and subalgebras of finitely presented Lie algebras*, Mathematics of the USSR-Izvestiya **6** (1972), no. 6, 1153–1199. (In Russian: *Izvestiya Akademii Nauk SSSR. Seriya Matematicheskaya* **36** (1972), no. 6, 1173–1219.).

[Bo3] L. A. Bokut', *Imbedding into simple associative algebras*, Algebra and Logic **15** (1976), 117–142. MR 58#22167.

[Bo4] L. A. Bokut', *On algebraically closed and simple Lie algebras*, Trudy Steklov Math. **148** (1978), 30–42. MR 82c:17009..

[BML] L. A. Bokut' and L. G. Makar-Limanov, *A basis of a free metabelian associative algebra*, Sibirskii Mathematicheskii Zhurnal **32** (1991), no. 6, 12–18. MR 93e:16035.

[BoKl1] L. A. Bokut' and A. A. Klein, *Serre relations and Gröbner–Shirshov bases for simple Lie algebras I*, International Journal of Algebra and Computation, to appear.

[BoKl2] L. A. Bokut' and A. A. Klein, *Serre relations and Gröbner–Shirshov bases for simple Lie algebras II*, International Journal of Algebra and Computation, to appear.

[BoKu] L. A. Bokut' and G. P. Kukin, *Algorithmic and Combinatorial Algebras*, Kluwer, Amsterdam, 1994. MR95i:17002.

[BW] T. Becker and V. Weispfenning, *Gröbner Bases. A Computational Approach to Commutative Algebra*, Graduate Texts in Mathematics Vol. 141, Springer-Verlag, Berlin, 1993. MR 95e:13018.

[Dr] V. G. Drinfel'd, *Hopf algebras and the quantum Yang–Baxter equation*, Doklady Adakemii Nauk SSSR **283** (1985), no. 5, 1060–1064. MR 87h:58080.

[Ji1] M. Jimbo, *A q-difference analogue of $U(G)$ and the Yang–Baxter equation*, Letters in Mathematical Physics **10** (1985), no. 1, 63–69. MR 86k:17008.

[Ji2] M. Jimbo, *A q-analogue of $U(gl(N+1))$, Hecke algebra, and the Yang–Baxter equation*, Letters in Mathematical Physics **11** (1986), no. 3, 247–252. MR 87k:17011.

[Ka] V. G. Kac, *Infinite-dimensional Lie Algebras*, third edition, Cambridge University Press, Cambridge, 1990. MR 92k:17038.

[Mo] F.Mora, *Groebner bases for non-commutative polynomial rings*, in *Proc. AAECC-3*, Lecture Notes in Computer Science **229** (1986), 353–362.

[Ro1] M. Rosso, *An analogue of the P. B. W. theorem and the universal R-matrix for $U_h sl(N+1)$*, Communications in Mathematical Physics **124** (1989), no. 2, 307–318. MR 90h:17019.

[Ro2] M. Rosso, *Analogues de la forme de Killing et du theoreme d'Harish–Chandra pour les groupes quantiques*, Annales Scientifiques de l'École Normale Supérieure **23** (1990), no. 3, 445–467. MR 93e:17026.

[Sh1] A. I. Shirshov, *On free Lie rings*, Matematicheskii Sbornik **45** (1958), no. 2, 113–122.

[Sh2] A. I. Shirshov, *Some algorithmic problems for ϵ-algebras*, Sibirskii Matematicheskii Zhurnal **3** (1962), no. 1, 132–137. MR 32#1222.

[Sh3] A. I. Shirshov, *Some algorithmic problems for Lie algebras*, Sibirskii Matematicheskii Zhurnal **3** (1962), no. 2, 292–296.

[Sh4] A. I. Shirshov, *On a hypothesis of the theory of Lie algebras*, Sibirskii Matematicheskii Zhurnal **3** (1962), no. 2, 297–301.

[Ya] I. Yamane, *A Poincaré–Birkhoff–Witt theorem for the quantized universal enveloping algebra of type A_N*, Publications of the Research Institute for Mathematical Sciences of Kyoto University **25** (1989), no. 3, 503–520. MR 91a:17016.