Journal of Pure and Applied Algebra 77 (1992) 169–181
North-Holland

Applications of a noncommutative jacobian matrix

Christophe Reutenauer

Département de Mathématiques, Université du Québec à Montréal, CP 8888, Suc. A, Montréal, Québec, Canada H3C 3P8

Communicated by C.A. Weibel
Received 20 February 1991

Abstract

Reutenauer, C., Applications of a noncommutative jacobian matrix, Journal of Pure and Applied Algebra 77 (1992) 169–181.

A noncommutative jacobian matrix is defined, for endomorphisms of a free associative algebra into another. The chain rules holds. Invertibility of this matrix characterizes automorphisms of the free Lie algebra. A presentation of $\mathrm{Aut}(k\langle x, y\rangle)$ and an inversion formula are given.

Introduction

We define, for an homomorphism f from a free associative algebra into another, a matrix $J(f)$, which behaves like the usual jacobian commutative matrix. The chain rule holds (Proposition 1.2), so that for an automorphism, $J(f)$ is invertible. An example shows that the converse is not true, so that there is no analogue of the jacobian conjecture. However, the main result is that for endomorphisms of the free Lie algebra, invertibility is equivalent to the invertibility of its jacobian matrix (Theorem 2.1). As a byproduct, we give a new proof of Cohn's theorem asserting that each automorphism of the free Lie algebra is a product of elementary automorphisms. Another application (Theorem 3.1) is a normal form for (continuous) automorphisms of the free associative algebra in two variables, which is an analogue for $\mathrm{Aut}(k\langle x, y\rangle)$ of Cohn's normal form for the GL_2 of a free associative algebra. A presentation of this group is also given. The same holds for $\mathrm{Aut}(k[x, y])$, because of the isomorphism between these two groups. In the last section, we give a Gröbner-like formula for the inversion of an automorphism of the algebra of noncommutative power series (Theorem 4.1).

1. Chain rule

Let X be a finite set of noncommutative variables (or *alphabet*) and k a commutative ring with unit. We denote by $k\langle X\rangle$ the algebra over k of noncommutative *polynomials* in the variables in X, and by $k\langle\langle X\rangle\rangle$ the algebra of noncommutative power *series*. Let X^* denote the free monoid over X, whose elements are called *words*. Then each series S may be viewed as an infinite linear combination over k of words

$$S = \sum_{w\in X^*} (S, w)w ,$$

where (S, w) denotes the coefficient of the word w. Similarly, a polynomial is a (finite) linear combination

$$P = \sum_{w\in X^*} (P, w)w .$$

There is a natural duality between $k\langle\langle X\rangle\rangle$ and $k\langle X\rangle$, given by

$$(S, P) = \sum_{w\in X^*} (S, w)(P, w) .$$

For any variable x, denote by $x^{-1}S$ the series defined by

$$x^{-1}S = \sum_{w\in X^*} (S, xw)w .$$

In other words, $S\mapsto x^{-1}S$ is the adjoint of $P\mapsto xP$ for the above duality. The following formula holds:

$$S = (S, 1) + \sum_{x\in X} x(x^{-1}S) , \tag{1.1}$$

where 1 is the empty word, the neutral element of $k\langle\langle X\rangle\rangle$; thus $(S, 1)$ is the *constant term* of S. Moreover,

$$x^{-1}(ST) = (x^{-1}S)T + (S, 1)(x^{-1}T) . \tag{1.2}$$

Both formulas are easily checked when S, T are words, and then extended by linearity and continuity (for the X-adic topology) to all series.

Let $k\langle\langle X\rangle\rangle$ be given the X-adic topology. If Y is another alphabet, then a continuous homomorphism $f : k\langle\langle X\rangle\rangle \to k\langle\langle Y\rangle\rangle$ is completely defined by the images $f(x)$ of the variables x in X; in this case, $f(x)$ is a *proper* series, that is, $(f(x), 1) = 0$. Define the *jacobian matrix* of f to be the $Y \times X$ matrix over $k\langle\langle Y\rangle\rangle$ defined by

$$J(f) = (y^{-1}f(x))_{y\in Y, x\in X} \in k\langle\langle Y\rangle\rangle^{Y\times X}.$$

Example 1.1. $X = Y = \{x, y\}$, $f(x) = x - yxy$, $f(y) = y$. Then,

$$J(f) = \begin{pmatrix} 1 & 0 \\ -xy & 1 \end{pmatrix}.$$

If J is a matrix and g an homomorphism, denote by J^g the matrix obtained by applying g to each entry of J.

Proposition 1.2. (chain rule). *Let X, Y, Z be three alphabets and $f : k\langle\langle X\rangle\rangle \to k\langle\langle Y\rangle\rangle$, $g : k\langle\langle Y\rangle\rangle \to k\langle\langle Z\rangle\rangle$ be continuous homomorphisms. Then*

$$J(g\circ f) = J(g)J(f)^g.$$

Proof. Let $x \in X$, $z \in Z$. Then

$$\begin{aligned}
&z^{-1}(g\circ f(x)) \\
&\quad = z^{-1}(g(f(x))) \\
&\quad = z^{-1}\left(g\left[(f(x),1) + \sum_{y\in Y} y(y^{-1}f(x))\right]\right) && \text{(by (1.1))} \\
&\quad = z^{-1}\left[(f(x),1) + \sum_{y\in Y} g(y)g(y^{-1}f(x))\right] \\
&\quad = \sum_{y\in Y} [(z^{-1}g(y))g(y^{-1}f(x)) + (g(y),1)z^{-1}g(y^{-1}f(x))] && \text{(by (1.2))} \\
&\quad = \sum_{y\in Y} (z^{-1}g(y))g(y^{-1}f(x)),
\end{aligned}$$

because $(g(y),1) = 0$, g being continuous. Thus the z,x-entry of $J(g\circ f)$ is equal to the z,x-entry in the product of $J(g)$ by $J(f)^g$. □

Corollary 1.3. *If $f : k\langle\langle X\rangle\rangle \to k\langle\langle X\rangle\rangle$ is a continuous automorphism, then $J(f)$ is invertible in $k\langle\langle X\rangle\rangle^{X\times X}$.* □

This corollary (an immediate consequence of Proposition 1.2) is well known in another form: it is the analogue for $k\langle\langle X\rangle\rangle$ of the implicit function theorem.

Corollary 1.4. *If $f : k\langle X\rangle \to k\langle X\rangle$ is an automorphism, then $J(f)$ is invertible in $k\langle X\rangle^{X\times X}$.*

Proof. We may suppose that $(f(x),1) = 1$ for each letter (by multiplying f on the right by the automorphism $x \mapsto x - (f(x),1)$). Then f extends uniquely to a

continuous automorphism of $k\langle\langle X\rangle\rangle$, and so does its inverse g. Hence $J(f)J(g)^f = J(\mathrm{id}) = I_X$ (identity $X \times X$-matrix). Moreover, $J(g)J(f)^g = I_X$, thus $J(g)^f J(f) = J(g)^f J(f)^{f\circ g} = (J(g)J(f)^g)^f = I_X^f = I_X$. Hence, the inverse of $J(f)$ is $J(g)^f$. □

The analogue of the *jacobian conjecture* (i.e. the converse of Corollary 1.4) does not hold: indeed, with f as in Example 1.1, $J(f)$ is invertible in $k\langle X\rangle^{X\times X}$, but f is not an automorphism. To see it, let $a = f(x)$, $b = f(y)$. Then $x = a + bxb$, thus $x = \sum_{n\geq 0} b^n a b^n$, and x is not a polynomial in a and b.

However, we see in the next section that the converse of Corollary 1.4 holds for endomorphisms of the free Lie algebra. Note that an analogue of the jacobian conjecture in the algebra of noncommutative polynomials has been stated, and proved in the case of two variables, by Dicks and Lewin [4]. Note also that the determinant of the matrix of constant terms of $J(f)$ has been considered by Fliess [5], under the name *jacobian determinant*, in connection with an inversion problem in control theory.

2. Automorphisms of the free Lie algebra

We suppose in this section that k is a field. Recall that the *free Lie algebra* on X over k is isomorphic to the Lie subalgebra of $k\langle X\rangle$ generated by the variables x in X (see, e.g., [8, Corollary 5.3.9]). We denote it by $\mathscr{L}(X)$, and call its elements *Lie polynomials*. As $\mathscr{L}(X)$ is free, an endomorphism f of $\mathscr{L}(X)$ is completely defined by the values $f(x)$ of the variables. Note that such an endomorphism uniquely extends to a continuous endomorphism of $k\langle X\rangle$, because a Lie polynomial has no constant term. Define its jacobian matrix as in Section 1.

An automorphism of $\mathscr{L}(X)$ is called *elementary* if either $f|_X$ defines a linear automorphism of the space $\sum_{x\in X} kx$, or if for some variable x, $f(x) = x + P$, where P is in $\mathscr{L}(X\backslash x)$, and $f(y) = y$ for any other variable y. Note that in both cases, f^{-1} is elementary too.

Theorem 2.1. *Let $f : \mathscr{L}(X) \to \mathscr{L}(X)$ be a Lie algebra endomorphism. The following conditions are equivalent:*

(i) *f is an automorphism.*

(ii) *The jacobian matrix of f is invertible in $k\langle X\rangle^{X\times X}$.*

(iii) *f is a product of elementary automorphisms.*

The equivalence of (i) and (iii) is due to Cohn [2]. For the proof of the theorem, we need two lemmas. The first one is surely well known, and is valid in any enveloping algebra. Recall that $k\langle X\rangle$ is the enveloping algebra of the Lie algebra $\mathscr{L}(X)$ (see, e.g., [8, Corollary 5.3.9]).

Lemma 2.2. *Let $\mathscr{L}$ be a Lie subalgebra of $\mathscr{L}(X)$. Then each Lie polynomial, which is in the right ideal of $k\langle X\rangle$ generated by $\mathscr{L}$, is already in $\mathscr{L}$.*

Proof. Let B be a totally ordered basis of $\mathscr{L}(X)$, containing a basis of $\mathscr{L}$, such that:

$$P,Q\in B\ ,\ P\in\mathscr{L}\ ,\ Q\notin\mathscr{L}\ \Rightarrow\ P<Q\ . \tag{2.1}$$

We claim that the polynomials

$$P_1\cdots P_n\ ,\quad n\geq 1\ ,\ P_i\in B\ ,\ P_1\in\mathscr{L}\ ,\ P_1\leq\cdots\leq P_n \tag{2.2}$$

generate linearly the right ideal $I=\mathscr{L}k\langle X\rangle$. Indeed, by the Poincaré–Birkhoff–Witt theorem, I is linearly generated by the polynomials

$$P=P_1\cdots P_n\ ,\quad n\geq 1\ ,\ P_i\in B\ ,\ P_1\in\mathscr{L}\ . \tag{2.3}$$

So it is enough to show that each polynomial (2.3) is a linear combination of polynomials (2.2); we do it by induction on $(n, t(P))$, ordered lexicographically, where $t(P)=|\{(i,j)\mid i<j \text{ and } P_i>P_j\}|$. If $n=1$ or if $t(P)=0$, then P is of the form (2.2) and there is nothing to prove. Let us assume that $t(P)\geq 1$: let i be such that $P_i>P_{i+1}$. If $i=1$, as $P_1\in\mathscr{L}$, we have by (2.1), $P_2\in\mathscr{L}$. Then $[P_1,P_2]\in\mathscr{L}$, hence $[P_1,P_2]=\sum_j\alpha_jQ_j$ $(Q_j\in\mathscr{L}\cap B,\ \alpha_j\in k)$. We have

$$\begin{aligned}P&=[P_1,P_2]P_3\cdots P_n+P_2P_1P_3\cdots P_n\\&=\sum_j\alpha_jQ_jP_3\cdots P_n+P_2P_1P_3\cdots P_n\ .\end{aligned}$$

Polynomial $Q_jP_3\cdots P_n$ is of the form (2.2) with a smaller n; $P_2P_1P_3\cdots P_n$ is of this form too, with same n and smaller $t(P)$. Hence by induction, all these polynomials are linear combinations of polynomials of the form (2.2), and so is P. Suppose now that $i\geq 2$. Then $[P_i,P_{i+1}]=\sum_k\beta_kR_k$ $(R_k\in B,\ \beta_k\in k)$ and

$$\begin{aligned}P&=P_1\cdots P_{i-1}([P_i,P_{i+1}]+P_{i+1}P_i)P_{i+2}\cdots P_n\\&=\sum_k\beta_kP_1\cdots P_{i-1}R_kP_{i+2}\cdots P_n+\sum_kP_1\cdots P_{i-1}P_{i+1}P_iP_{i+2}\cdots P_n\ .\end{aligned}$$

Then a similar but simpler inductive argument shows that P is a linear combination of polynomials of the form (2.2). This proves the claim.

Now let R be a Lie polynomial which is in I. Then we may write

$$R=\sum_{Q\in B}\alpha_QQ=\sum_P\beta_PP \tag{2.4}$$

where α_Q, β_P are in k, and where the second summation is over polynomials P of the form (2.2). By the theorem of Poincaré–Birkhoff–Witt, the increasing products $Q_1 \cdots Q_n$, $Q_i \in B$, $Q_1 \leq \cdots \leq Q_n$, are linearly independent. Hence, (2.4) implies that R is a linear combination of polynomials $Q \in B \cap \mathscr{L}$. In particular, Q is in $\mathscr{L}$. □

The next lemma is technical, but easy. Denote by $\bar{P}$ the highest homogeneous component of a polynomial P, with $\bar{P} = 0$ if $P = 0$.

Lemma 2.3. *Let $P, P_1, \ldots, P_n$ be polynomials such that $\bar{P}$ is in the Lie subalgebra generated by $\bar{P}_1, \ldots, \bar{P}_n$. Then there exists a Lie polynomial $u(x_1, \ldots, x_n)$ in $\mathscr{L}(x_1, \ldots, x_n)$ such that*

$$\deg(P - u(P_1, \ldots, P_n)) < \deg(P) .$$

Proof. We have $\bar{P} = v(\bar{P}_1, \ldots, \bar{P}_n)$ for some Lie polynomial v in $\mathscr{L}(x_1, \ldots, x_n)$. Let $v = \sum_{(d)} v_{(d)}$, where the sum is over n-tuples $(d) = (d_1, \ldots, d_n)$ and where $v_{(d)}$ is the homogeneous component of v of degree d_i in each x_i. Let $u = \sum_{(d)} v_{(d)}$, where the sum is over all (d) with $\sum_{i=1}^n d_i \deg(P_i) = \deg(P)$. By homogeneity, we have $\bar{P} = u(\bar{P}_1, \ldots, \bar{P}_n)$. Moreover, $\mathscr{L}(x_1, \ldots, x_n)$ is homogeneous, so that each $v_{(d)}$ is a Lie polynomial, hence u too. Now by construction, $u(P_1, \ldots, P_n)$ is equal to $u(\bar{P}_1, \ldots, \bar{P}_n)$ plus a polynomial of degree less than P. □

Proof of Theorem 2.1. (i) $\Rightarrow$ (ii) If f is a Lie automorphism, then it extends uniquely to an automorphism of $k\langle X\rangle$, and we use Corollary 1.4.

(ii) $\Rightarrow$ (iii) (Induction on $d(f) = \sum_{x \in X} \deg(f(x))$.) Suppose that $d(f) \leq |X|$. As no $f(x)$ is zero (otherwise $J(f)$ is not invertible) and as $f(x)$ is a Lie polynomial, we have $\deg(f(x)) \geq 1$. This implies that $\deg(f(x)) = 1$ for any variable x, and we may write $f(y) = \sum_{x \in X} \alpha_{x,y} x$ for some scalars $\alpha_{x,y}$. Then $x^{-1} f(y) = \alpha_{x,y}$, which shows that $J(f) = (\alpha_{x,y})_{x,y \in X}$. By hypothesis $J(f)$ is invertible in $k\langle X\rangle^{X \times X}$; taking constant terms, we see that $J(f)$ is invertible in $k^{X \times X}$, hence f is elementary of the first type. Suppose now that $d(f) > |X|$. Let $(P_{x,y})_{x,y}$ be the inverse of $J(f)$ in $k\langle X\rangle^{X \times X}$. Then, for any variables x, z, we have

$$\sum_{y \in X} (x^{-1} f(y)) P_{y,z} = \delta_{x,z} .$$

Multiply by x at the left, sum over all x, and note that $f(y)$ has no constant term. Hence (1.2) implies

$$\sum_{y \in X} f(y) P_{y,z} = z .$$

As $d(f) > |X|$, there is some variable x_0 such that $\deg(f(x_0)) \geq 2$; moreover,

there is some z such that $P_{x_0,y}$ is nonzero, because the matrix $(P_{x,y})$ is invertible. Hence, for this z

$$\deg\Big(\sum_{y\in X} f(y)P_{y,z}\Big) = \deg(z) = 1 < \deg(f(x_0)P_{x_0,y}) \le \sup_y\{\deg(f(y)P_{y,z})\}\,.$$

We use now Cohn's weak algorithm (see [3, Proposition 2.4.2]) which asserts that, as no $f(y) = 0$, there exists a variable x and polynomials Q_y such that

$$\deg(f(x) - \sum_{y\neq x} f(y)R_y) < \deg(f(x))\,, \qquad \deg(f(y)R_y) \le \deg(f(x))\,.$$

This implies that

$$\overline{f(x)} = \sum_{y\neq x} \overline{f(y)}\,\overline{R_y}\,.$$

Now, Lemma 2.2 implies that $\overline{f(x)}$ is in the Lie subalgebra generated by the Lie polynomials $\overline{f(y)}$, $y \neq x$. Hence, by Lemma 2.3, there exists a Lie polynomial $P(y)_{y\neq x}$ in $\mathscr{L}(X\backslash x)$ such that

$$\deg(f(x) - P(f(y))_{y\neq x}) < \deg(f(x))\,. \tag{2.5}$$

Define an elementary automorphism g of the second type by

$$g(x) = x - P(y)_{y\neq x}\,, \qquad g(y) = y \quad \text{if } y \neq x\,.$$

Let $h = f\circ g$. Then we have $h(x) = f(g(x)) = f(x - P(y)_{y\neq x}) = f(x) - P(f(y))_{y\neq x}$, because f is a Lie algebra endomorphism. Moreover, if $y \neq x$, then $h(y) = f(g(y)) = f(y)$. Hence, by (2.5), $d(h) < d(f)$. Moreover, by Proposition 1.2, $J(h) = J(f)J(g)^f$, hence $J(h)$ is invertible in $K\langle X\rangle^{X\times X}$.

By induction, we conclude that h is a product of elementary automorphisms. This implies that so is f.

(iii) $\Rightarrow$ (i) This is clear. □

The proof also shows the following result, which is an analogue for free Lie algebras of a result of Nielsen in free groups (see [9, Proposition I.2.7]).

Corollary 2.4. *If n Lie polynomials generate the free Lie algebra $\mathscr{L}(x_1, \ldots, x_n)$, then they generate it freely.* □

Remarks. (1) One may define a jacobian matrix J' by using the operators $P \mapsto Px^{-1}$, symmetric to $P \mapsto x^{-1}P$. Then the chain rule also holds, and Theorem

2.1 shows that for a free Lie algebra endomorphism f, $J(f)$ is invertible if and only if $J'(f)$ is. This may be seen directly, because

$$J'(f) = {}^{t}J(f)^{\alpha} ,$$

where α is the principal anti-automorphism of $k\langle X\rangle$, sending each Lie polynomial P to $-P$.

(2) A theorem analogue to Theorem 2.1 for free groups has been proved by Birman [1]. She uses differential operators introduced by Fox [6]. Actually, these operators d_x are, up to a left–right symmetry, related to the operators $x^{-1}P$ by the formula

$$d_x = M^{-1}(x^{-1}M(P)) ,$$

for any element P of the group algebra of the free groups $F(X)$, where M is the Magnus embedding

$$kF(X) \to k\langle\langle X\rangle\rangle , \quad x \mapsto 1 + x .$$

While the operator $P \mapsto x^{-1}P$ is a φ-derivation (see Section 4) for $\varphi(P) = (P, 1)$, the operator d_x is a ψ-derivation for $\psi(P) =$ sum of the coefficients of P.

3. Automorphisms of two variables polynomial algebras

We assume that k is a field. Define, for each polynomial $a(x) \in k\langle x\rangle$, an automorphism $\pi(a)$ of $k\langle x, y\rangle$ by

$$\pi(a)(x) = y + xa(x) , \qquad \pi(a)(y) = x .$$

As $a(x)$ depends on x only, $\pi(a)$ is an automorphism of $k\langle x, y\rangle$. The following relation holds ($b \in k\langle x\rangle$):

$$\pi(a + b) = \pi(a)\pi(0)\pi(b) . \tag{3.1}$$

Indeed,

$$\begin{aligned}\pi(a)\pi(0)\pi(b)(x) &= \pi(a)\pi(0)(y + xb(x)) = \pi(a)(x + yb(y)) \\ &= y + xa(x) + xb(x) = \pi(a + b)(y) ,\end{aligned}$$

and

$$\pi(a)\pi(0)\pi(b)(y) = \pi(a)\pi(0)(x) = \pi(a)(y) = x = \pi(a + b)(y) .$$

Moreover, if $a(x) = \alpha \in k\backslash 0$, we have

$$\pi(\alpha)\pi(-\alpha^{-1})\pi(\alpha) = \delta(\alpha, -\alpha^{-1}), \tag{3.2}$$

where $\delta(\alpha, \beta)$ is the automorphism $x \mapsto \alpha x$, $y \mapsto \beta y$ $(\alpha, \beta \in k\backslash 0)$. Indeed,

$$\begin{aligned}\pi(\alpha)\pi(-\alpha^{-1})\pi(\alpha)(x) &= \pi(\alpha)\pi(-\alpha^{-1})(y + \alpha x)\\ &= \pi(\alpha)(x + \alpha(y - \alpha^{-1}x))\\ &= \pi(\alpha)(\alpha y) = \alpha x = \delta(\alpha, -\alpha^{-1})(x),\end{aligned}$$

and

$$\begin{aligned}\pi(\alpha)\pi(-\alpha^{-1})\pi(\alpha)(y) &= \pi(\alpha)\pi(-\alpha^{-1})(x) = \pi(\alpha)(y - \alpha^{-1}x)\\ &= x - \alpha^{-1}(y + \alpha x) = -\alpha^{-1}y = \delta(\alpha, -\alpha^{-1})(y).\end{aligned}$$

Another identity is the following $(\alpha, \beta \in k\backslash 0,\ a \in k\langle x\rangle)$:

$$\pi(a^{\beta})\delta(\alpha, \beta) = \delta(\beta, \alpha)\pi(\beta^{-1}\alpha a), \tag{3.3}$$

where $a^{\beta}(x) = a(\beta x)$.

Indeed,

$$\pi(a^{\beta})\delta(\alpha, \beta)(x) = \pi(a^{\beta})(\alpha x) = \alpha y + \alpha x a^{\beta}(x),$$

$$\begin{aligned}\delta(\beta, \alpha)\pi(\beta^{-1}\alpha a)(x) &= \delta(\beta, \alpha)(y + \beta^{-1}\alpha x a(x))\\ &= \alpha y + \beta^{-1}\alpha\beta x a(\beta x) = \alpha y + \alpha x a(\beta x),\end{aligned}$$

$$\pi(a^{\beta})\delta(\alpha, \beta)(y) = \pi(a^{\beta})(\beta y) = \beta x,$$

and

$$\delta(\beta, \alpha)\pi(\beta^{-1}\alpha a)(y) = \delta(\beta, \alpha)(x) = \beta x.$$

We also have the evident identity

$$\delta(\alpha, \beta)\delta(\alpha', \beta') = \delta(\alpha\alpha', \beta\beta'). \tag{3.4}$$

Note that all these definitions and relations also make sense in $k[x, y]$, the algebra of commutative polynomials in x, y over k. Recall that an automorphism is continuous (or augmentation-preserving) if it sends each variable on a polynomial without constant term.

Theorem 3.1. (1) *The group* $\mathrm{Aut}^c(k\langle x, y\rangle)$ *of continuous automorphisms of* $k\langle x, y\rangle$ *is generated by the automorphisms* $\pi(a)$*and* $\delta(\alpha, \beta)(a \in k\langle x\rangle,\ \alpha, \beta \in k\backslash 0)$, *and the relations* (3.1)–(3.4) *form a complete set of defining relations. Each automorphism* f *of* $k\langle x, y\rangle$ *has a unique decomposition of the form*

$$f = \delta(\alpha, \beta)\pi(a_1)\cdots\pi(a_r)\,, \tag{3.5}$$

where $r \geq 0$, *and where for* $i = 2, \ldots, r-1$, a_i *is not constant.*

(2) *The same property holds for the group* $\mathrm{Aut}^c(k[x, y])$ *of continuous automorphisms of* $k[x, y]$.

Proof. Note that the second assertion follows immediately from the fact that the canonical mapping $\mathrm{Aut}(k\langle x, y\rangle \to \mathrm{Aut}(k[x, y])$ is an isomorphism (see [3, Theorem 6.9.3]). We know that $\mathrm{Aut}(k\langle x, y\rangle)$ is generated by its elementary automorphisms and the diagonal automorphisms $\delta(\alpha, \beta)$ (ibid.). This easily implies that $\mathrm{Aut}^c(k\langle x, y\rangle)$ is generated by its elementary continuous automorphisms and the $\delta(\alpha, \beta)$. Now, an elementary continuous automorphism is either of the form: (i) $x \mapsto x + ya(y)$, $y \mapsto y$, or (ii) $x \mapsto x$, $y \mapsto y + xa(x)$. In the first case it is equal to $\pi(0)\pi(a)$, while in the second it is $\pi(a)\pi(0)$. This shows that the automorphisms $\pi(a)$ and $\delta(\alpha, \beta)$ generate $\mathrm{Aut}^c(k\langle x, y\rangle)$.

We show that each automorphism, when written as a product of the generators, may be brought into the form (3.5), by using only the relations (3.1)–(3.4). This is done following Cohn's method for $\mathrm{GE}_2(R)$ [3, 2.7]. By (3.1), we have $\pi(0)^2 = 1$, hence

$$\pi(a)^{-1} = \pi(0)\pi(-a)\pi(0)\,. \tag{3.6}$$

This relation, together with (3.4), allows to eliminate all negative powers of the generators $\pi(a)$ and $\delta(\alpha, \beta)$. Now, (3.3) allows to bring all the $\delta(\alpha, \beta)$ at the beginning, and (3.4) gives then the desired form (3.5), but without the conditions on the a_i's. If $a_i = 0$ for $i \neq 1, r$, then one uses (3.1) to shorten the form (3.5). Suppose that $a_i = \alpha \in k\backslash 0$ for some $i \neq 1, r$. We have by (3.6),

$$\begin{aligned}
&\pi(a + \alpha^{-1})^{-1}\pi(a)\pi(\alpha)\pi(b)\pi(b + \alpha^{-1})^{-1} \\
&\quad = \pi(0)\pi(-a-\alpha^{-1})\pi(0)\pi(a)\pi(\alpha)\pi(b)\pi(0)\pi(-b-\alpha^{-1})\pi(0) \\
&\quad = \pi(0)\pi(-\alpha^{-1})\pi(\alpha)\pi(-\alpha^{-1})\pi(0) \quad \text{by (3.1)} \\
&\quad = \pi(0)\delta(-\alpha^{-1}, \alpha)\pi(0) \quad \text{by (3.2)} \\
&\quad = \pi(0)\pi(0)\delta(\alpha, -\alpha^{-1}) \quad \text{by (3.3)} \\
&\quad = \delta(\alpha, -\alpha^{-1})\,.
\end{aligned}$$

Thus we have

$$\pi(a)\pi(\alpha)\pi(b) = \pi(a + \alpha^{-1})\delta(\alpha, -\alpha^{-1})\pi(b + \alpha^{-1}) . \tag{3.7}$$

This shows that the form (3.4) may be shortened, by using (3.7), and then (3.3) to bring $\delta(\alpha, -\alpha^{-1})$ at the beginning.

We show now that the form (3.5) is uniquely determined by the automorphism f. This will end the proof.

We know by [3, Proposition 2.8.2] that each matrix J in $\mathrm{GL}_2(k\langle x, y\rangle)$ has a unique decomposition as

$$J = [\alpha, \beta]P(a_1)\cdots P(a_r) , \tag{3.8}$$

where $[\alpha, \beta] = \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix}\right)$, $P(a) = \left(\begin{smallmatrix} a & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and where for $i \neq 1, r$, a_i is not constant. Note that $J(\delta(\sigma, \beta)) = [\alpha, \beta]$ and $J(\pi(a)) = P(a)$. With f as in (3.5), define

$$f_i = \delta(\alpha, \beta)\pi(a_1)\cdots\pi(a_i)$$

for $i = 0, \ldots, r$. Then, by Proposition 1.2, we have

$$\begin{aligned} J(f) &= J(\delta(\alpha, \beta))\prod_{i=1}^{r} J(\pi(a_i))^{f_{i-1}} \\ &= [\alpha, \beta]\prod_{i=1}^{r} P(a_i)^{f_{i-1}} \\ &= [a, \beta]P(f_0(a_1))P(f_1(a_2))\cdots P(f_{r-1}(a_r)) . \end{aligned}$$

This is form (3.8) for the matrix $J(f)$; hence the knowledge of f determines that of α, β, $f_0(a_1), \ldots, f_{r-1}(a_r)$. Now, recursively, one determines a_1, f_1, a_2, f_2, and so on, until a_r. Hence, the form (3.5) is unique. □

4. A Gröbner-like inversion formula

A φ-derivation of a ring R is a linear mapping $D : R \to R$ such that $D(uv) = D(u)v + \varphi(u)D(v)$, for any u, v in R, where $\varphi : R \to R$ is a fixed homomorphism. We consider here $R = k\langle\langle X\rangle\rangle$ with $\varphi(S) = (S, 1)$, for any series S. For any letter x in X, the mapping

$$S \mapsto x^{-1}S$$

is a φ-derivation, as (1.2) shows. It is easily checked that for each family of series $(S_x)_{x\in X}$, there exists one and only one φ-derivation D sending each variable x

onto S_x. This φ-derivation D is given by the formula

$$D(S) = \sum_{x \in X} S_x(x^{-1}S) .$$

Theorem 4.1. *Let f be a continuous automorphism of $k\langle\langle X\rangle\rangle$, with jacobian matrix J. Define for each variable x, a φ-derivation D_x by the formula*

$$D_x(S) = \sum_{y \in X} J^{-1}_{x,y} (y^{-1}S) .$$

Then for each word $w = x_1 \cdots x_n$ $(x_i \in X)$ and each series S, one has

$$(f^{-1}(S), w) = (D_{x_n} \circ \cdots \circ D_{x_1}(S), 1) .$$

Proof. Denote by d_x the φ-derivation $S \mapsto x^{-1}S$. Then $d_x = f^{-1} \circ D_x \circ f$. Indeed, the right-hand member is a φ-derivation:

$$\begin{aligned} f^{-1} \circ D_x \circ f(ST) &= f^{-1}[D_x \circ f(S).f(T) + (f(S), 1).D_x \circ f(T)] \\ &= f^{-1} \circ D_x \circ f(S).T + (S, 1).f^{-1} \circ D_x \circ f(T) , \end{aligned}$$

because $f(S)$ and S have the same constant term, f being continuous. Moreover, for any variable z,

$$f^{-1} \circ D_x \circ f(z) = f^{-1}\left[\sum_{y \in X} J^{-1}_{x,y}(y^{-1}f(z)) \right] = f^{-1}(\delta_{x,z}) = \delta_{x,z} ,$$

because J is the matrix $(y^{-1}f(z))_{y,z \in X}$. Thus d_x and $f^{-1} \circ D_x \circ f$ are two φ-derivations which take the same value on each variable z. So $d_x = f^{-1} \circ D_x \circ f$.

Now, let w and S as in the theorem. We have (since f^{-1} preserves constant terms)

$$\begin{aligned} (D_{x_n} \circ \cdots \circ D_{x_1}(S), 1) &= (f^{-1} \circ D_{x_n} \circ \cdots \circ D_{x_1}(S), 1) \\ &= (f^{-1} \circ D_{x_n} \circ \cdots \circ D_{x_1} \circ f[f^{-1}(S)], 1) \\ &= (d_{x_n} \circ \cdots \circ d_{x_1}[f^{-1}(S)], 1) \\ &= (x_n^{-1} \cdots x_1^{-1}[f^{-1}(S)], 1) \\ &= (f^{-1}(S), x_1 \cdots x_n) \\ &= (f^{-1}(S), w) . \qquad \square \end{aligned}$$

Note added in proof

V. Shpilrain also obtained the jacobian characterization of automorphisms of the free Lie algebra of Theorem 2.1 (communication at ICM 1990, Kyoto).

References

[1] J.S. Birman, An inverse function theorem for free groups, Proc. Amer. Math. Soc. 41 (1973) 634–638.
[2] P.M. Cohn, Subalgebras of free associative algebras, Proc. London Math. Soc. 14 (1964) 618–632.
[3] P.M. Cohn, Free Rings and their Relations (Academic Press, New York, 2nd ed., 1985).
[4] W. Dicks and J. Lewin, A jacobian conjecture for free associative algebras, Comm. Algebra 10 (1982) 1285–1306.
[5] M. Fliess, On the concept of derivatives and Taylor expansions for nonlinear input–output systems, in: Proceedings of the 22nd IEEE Conference on Decision and Control, San Antonio, Texas (1983) 643–646.
[6] R.H. Fox, Free differential calculus, Ann. of Math. 57 (1953) 547–560.
[7] W. Gröbner, Über die Darstellung von implizit gegebenen Funktionen mittels Lie-Reihen und Verallgemeinerung der Lagrangeschen Reihen, Monatsh. Math. 66 (1962) 129–139.
[8] M. Lothaire, Combinatorics on Words (Addison-Wesley, Reading, MA, 1983).
[9] R.C. Lyndon and P.E. Schupp, Combinatorial Group Theory (Springer, Berlin, 1977).