Algorithmic methods in skew polynomial rings

(Talk)

José Gómez-Torrecillas

Granada, july 6th, 2000

Example 1

Consider an homogeneous linear ordinary differential equation

$$a_n(t)\frac{d^n y(t)}{dt^n} + \cdots + a_1(t)\frac{dy(t)}{dt} + a_0(t)y(t) = 0,$$

where the $a_i(t)$'s are functions in some field (e.g. the field $\mathbb{C}(t)$ of rational complex functions).

Consider the linear operator

$$L = a_n(t)\frac{d^n}{dt^n} + \cdots + a_1(t)\frac{d}{dt} + a_0(t) \in \mathrm{End}_{\mathbb{C}}(\mathcal{F})$$

acting on a $\mathbb{C}$-vector space $\mathcal{F}$ of functions. Think of $\mathcal{F}$ as a (commutative) algebra of functions containing $\mathbb{C}(t)$ as a subalgebra. This allows $\mathbb{C}(t)$ act on $\mathcal{F}$ by multiplication. Thus, $\mathbb{C}(t) \cup \{d/dt\} \subseteq \mathrm{End}_{\mathbb{C}}(\mathcal{F})$ generates a $\mathbb{C}$-subalgebra, say R, of $\mathrm{End}_{\mathbb{C}}(\mathcal{F})$. Obviously, $L \in R$ and the rule

$$L \cdot y(t) = L(y(t))$$

endows $\mathcal{F}$ with a structure of a left R-module, and the equation becomes $L \cdot y(t) = 0$.

Let $y(t) \in \mathcal{F}$ be a solution of $L \cdot y(t) = 0$. Then the mapping

$$R/RF \to \mathcal{F} \qquad (r + RL \mapsto r \cdot y(t))$$

is a homomorphism of left R-modules (here,

$$RF = \{rF \mid r \in R\}$$

is the *left ideal* of R generated by F.
Conversely, every homomorphism of left R-modules $\varphi : R/RF \to \mathcal{F}$ provides a solution $y(t) := \varphi(1 + RF)$ of our differential equation. Therefore, the generator $1 + RF$ of the left R-module can be viewed as a 'generic solution' of the differential equation $L(y(t)) = 0$.

Idea: *The module R/RL contains relevant information about the differential equation. For instance, two differential equations are equivalent if and only if they have isomorphic associated left R-modules.*

More generally, a homogeneous *system* of differential linear ordinary equations is identified with a finitely generated left R–module, which is of the form R^m/K, where K is a submodule of of the free left R–module

$$R^m = R\times \overset{(m)}{\cdots} \times R$$

This gives, for instance, a safe framework to declare when a system is equivalent to an equation...

Which kind of ring is our R?

By Leibniz's rule, for every $a(t) \in \mathbb{C}(t)$,

$$d/dt \circ a(t) = a(t) \circ d/dt + \frac{da(t)}{dt} \qquad (*)$$

whence R is a non-commutative ring (commutative rings rarely appear in nature).
From the commutation relation $(*)$ we get that every operator $r \in R$ can be represented as a polynomial

$$r = a_n(t)(\frac{d}{dt})^n + \cdots a_1(t)\frac{d}{dt} + a_0(t)$$

On the other hand, the powers $(\frac{d}{dt})^n = \frac{d^n}{dt^n}$ are linearly independent over $\mathbb{C}(t)$. Therefore, R is already a non-commutative polynomial ring in the 'variable' $\frac{d}{dt}$ with coefficients in the field $\mathbb{C}(t)$. The multiplication is given by (*).

This is a fundamental example of *Ore extension* of a field (O. Ore, 1933).

The fundamental property of the ring R is the Euclidean Division Algorithm, which we will briefly recall.

Given a differential operator

$$r = a_n(t)(\frac{d}{dt})^n + \cdots + a_1(t)\frac{d}{dt} + a_0(t)$$

with $a_n(t) \neq 0$, write $\deg(r) = n$ and $\mathrm{lc}(r) = a_n(t)$.

Algorithm 1 Euclidean Division Algorithm

INPUT: $f, g \in R$ with $g \neq 0$

OUTPUT: q, r such that $f = qg + r$ and $r = 0$ or $\deg(r) < \deg(g)$.

INITIALIZATION: $q := 0, r := f$

WHILE $r \neq 0$ **AND** $\deg(g) \leq \deg(r)$ **DO**

$q := q + \mathrm{lc}(r)\mathrm{lc}(g)^{-1}(d/dt)^{\deg(r)-\deg(g)}$

$r := r - \mathrm{lc}(r)\mathrm{lc}(g)^{-1}(d/dt)^{\deg(r)-\deg(g)}g$

Some Consequences

• Division algorithm allows to decide if a linear differential operator $P \in R$ is a multiple of a given linear differential operator L. This is an easy example of effective solving a non-commutative 'membership problem', that is, to decide if P belongs to the left ideal RL generated by L.

• Effective computations (e.g., highest left or right common factors, least left or right common multiples, Bézout identity...) can be done on R (O. Ore, Ann. Math. 34(1933), 480-512).

• Elementary row and column transformations allow to compute normal forms for matrices with entries in R (which leads to the structure theorem for modules R^n/K) (N. Jacobson, The Theory of Rings, Amer. Math. Soc, 1943).

• Effective methods to solve systems of linear differential equations can be developed (M. Bronstein and M. Petkovsek, 'An introduction to pseudo-linear algebra', 1998).

The abstract setting: Systems of linear equations over a non-commutative ring

Let $\mathcal{F}$ be a vector space over a field k and $R \subseteq \mathrm{End}_k(\mathcal{F})$ a subalgebra (under composition) of linear operators. Thus, $\mathcal{F}$ is a left R-module whose elements play the rôle of 'functions'.

Definition: A system of linear equations over R is a left R-module R^m/I, where I is a finitely generated submodule of the free left R-module $R^m = R\times \overset{(m)}{\cdots} \times R$.

Thus, a central problem is the computational treatment of (finitely presented) modules over R.

Let us focus our attention in a very elementary problem: Given $r_1, \ldots, r_s \in R$ and $r \in R$, is r linearly dependent of $r_1, \ldots, r_s$? In other words, we want to solve the equation

$$r = g_1 r_1 + \cdots + g_s r_s \qquad (g_1, \ldots, g_s \in R)$$

This is the 'membership problem', i.e., is r an element of the left ideal $Rr_1 + \cdots + Rr_s$ generated by $r_1, \ldots, r_s$?

Example 2

Let $A = \mathbb{C}[t_1, \ldots, t_n]$ be the ring of polynomials in the (commuting) variables $t_1, \ldots, t_n$. Let $\mathcal{F}$ be an algebra of functions containing A such that the partial derivates $\partial_i = \frac{\partial}{\partial t_i}$ 'make sense' on $\mathcal{F}$. Consider R the subring of $\mathrm{End}_{\mathbb{C}}(\mathcal{F})$ generated by A and $\partial_1, \ldots, \partial_n$. So, every element of R is a linear differential operator with polynomial coefficients acting on $\mathcal{F}$ (this is the n-th complex Weyl algebra).

The product in R is built (from Leibniz's rule) on the commutation relations $\partial_i t_i - t_i \partial_i = 1$ (think of t_i as operators!) and any other pair among the t_i's and ∂_j's commute.

An interpretation from the point of view of the theory of differential equations of the membership problem here is to know whether a given linear differential equation in a system dropped because it is 'linearly dependent' of the rest.

We cannot expect an algorithm to solve the membership problem for any ring R. Some basic algorithms can be developed for certain non-commutative polynomial rings.

Let $x_1, \ldots, x_n \in R$. A *standard monomial* in $x_1, \ldots, x_n$ is an element of R of the form

$$\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

where $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is a multi-index with non-negative integer components.

Now assume that R contains a field $\mathbf{k}$ (or, more generally, a skew-field). Assume that the standard monomials $\{\mathbf{x}^\alpha \mid \alpha \in \mathbb{N}^n\}$ form a basis of R as a left vector space over $\mathbf{k}$. This means that every element $0 \neq f \in R$ can be uniquely represented as a polynomial

$$f = \sum_{\alpha \in \mathcal{N}(f)} a_\alpha \mathbf{x}^\alpha, \qquad (a_\alpha \in \mathbf{k} \setminus \{0\})$$

where $\mathcal{N}(f) \subset \mathbb{N}^n$ is a finite subset, called *Newton diagram* of f.

Now, we need a 'good' notion of leading term. Order $\mathbb{N}^n$ by means of an *admissible order* $\preceq$, that is, a total order relations satisfying

- $\mathbf{0} = (0, \ldots, 0) \preceq \alpha$ for every $\alpha \in \mathbb{N}^n$,
- $\alpha \preceq \beta$ implies $\alpha + \gamma \preceq \beta + \gamma$.

An important property of admissible orders is that they are good orders in the sense that every non empty subset of $\mathbb{N}^n$ has a first element (**Dickson's Lemma**). Thus, induction proofs are available here and, from the computational point of view, recursive algorithms can be built.

Notice that, for $n > 1$, there are uncountably infinitely many admissible orders on $\mathbb{N}^n$.

A basic example is the lexicographical order $\leq_{lex}$ with

$$\epsilon_1 <_{lex} \epsilon_2 <_{lex} \cdots <_{lex} \epsilon_n,$$

where $\epsilon_i = (0, \ldots, \underset{(i)}{1}, \ldots, 0)$.

A family of such orders is given as follows: Let $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{R}^n$ with $w_i \geq 0$ for $i = 1, \ldots, n$. Define the $\mathbf{w}$–weighted degree lexicographical order $\preceq_{\mathbf{w}}$ by

$$\alpha \preceq_{\mathbf{w}} \beta \begin{cases} \langle \mathbf{w}, \alpha \rangle < \langle \mathbf{w}, \beta \rangle \\ \qquad or \\ \langle \mathbf{w}, \alpha \rangle = \langle \mathbf{w}, \beta \rangle \quad and \quad \alpha <_{lex} \beta \end{cases}$$

where $\langle -, - \rangle$ denotes the usual dot product in $\mathbb{R}^n$.

Let $0 \neq f \in R$. For each admissible order $\preceq$ on $\mathbb{N}^n$, define the *exponent* of f as

$$\exp(f) = \max_{\preceq} \mathcal{N}(f)$$

This exponents will play the rôle of leading terms of f. A major difference between our non-commutative polynomials and the usual commutative polynomial is that in our setting the product of two standard monomials is NOT a monomial. However, the fundamental algebraic algorithms developed for commutative multivariable polynomials (Division Algorithm, Buchberger's Algorithm...) can be adapted to a large class of non-commutative polynomial rings due to the following result:

Theorem (Bueso, Gómez-Torrecillas, Lobillo, 1998)

Let R contain an skew-field $\mathbf{k}$ and operators $x_1, \ldots, x_n$ such that the standard monomials $\mathbf{x}^\alpha$, $\alpha \in \mathbb{N}^n$ form a basis of R as a left vector space over $\mathbf{k}$. The following statements are equivalent for an admissible order $\preceq$ on $\mathbb{N}^n$:

(i) $\exp(fg) = \exp(f) + \exp(g)$ for every $f, g \in R \setminus \{0\}$;

(ii) (a) for every $\alpha, \beta \in \mathbb{N}^n$, there is $0 \neq q_{\alpha\beta} \in \mathbf{k}$ and $p_{\alpha\beta} \in R$ with $\exp(()p_{\alpha\beta}) \prec \alpha + \beta$ such that

$$\mathbf{x}^\alpha \mathbf{x}^\beta = q_{\alpha\beta}\mathbf{x}^{\alpha+\beta} + p_{\alpha\beta}$$

(b) for every $\alpha \in \mathbb{N}^n$ and every $0 \neq a \in \mathbf{k}$ there are $0 \neq a^\alpha \in \mathbf{k}$ and $p_{a\alpha} \in R$ with $\exp(p_{a\alpha}) \prec \alpha$ such that

$$\mathbf{x}^\alpha a = a^\alpha \mathbf{x}^\alpha + p_{a\alpha}$$

(next page contains one more equivalent statement)

(iii) (a) for every $1 \leq i < j \leq n$

$$x_j x_i = q_{ij} x_i x_j + p_{ij} \qquad (0 \neq q_{ij} \in \mathbf{k}, \exp(p_{ij}) \prec \epsilon_i + \epsilon_j)$$

(b) for every $i = 1, \ldots, n$ and every $0 \neq a \in \mathbf{k}$,

$$x_i a = a_j x_i + p_{ia} \qquad (0 \neq a_i \in \mathbf{k}, \exp(p_{ia}) \prec \epsilon_i)$$

A ring R satisfying the equivalent conditions of this theorem is called a left Poincaré-Birkhoff-Witt (PBW, for short) ring (with respect to $x_1, \ldots, x_n$, $\mathbf{k}$ and $\preceq$).

The notion of PBW ring embodies that of solvable polynomial ring developed by A. Kandri-Rody and V. Weispfenning (J. Symb. Comp. 9, 1990, 1-26) and H. Kredel (Phd. Univ. Passau, 1992).

Specific examples of left PBW rings are the Weyl algebras, and, more generally, differential operator rings, the universal enveloping algebras of finite-dimensional Lie algebras or most relevant examples of quantized algebras.

The theory of Gröbner bases has been developed for left modules over a left PBW ring. To simplify, let us consider the case of a left ideal I of R (that, is, I is an additive subgroup of R such that $rb \in I$ for every $b \in I$.) The point here is that the set

$$\mathrm{Exp}(I) = \{\exp(f) \mid f \in I\} \subseteq \mathbb{N}^n$$

is an stable subset of $\mathbb{N}$ in the sense that

$$\alpha + \mathrm{Exp}(I) \subseteq \mathrm{Exp}(I)$$

for every $\alpha \in \mathbb{N}^n$. It follows from Dickson's Lemma that every left ideal posses a Gröbner basis in the sense of the following definition.

Definition A Gröbner basis for a left ideal I of R is a finite subset $G = \{g_1, \ldots, g_t\} \subseteq I$ such that

$$\mathrm{Exp}(I) = \bigcup_{i=1}^{n} (\exp(g_i) + \mathbb{N}^n)$$

Idea: *All relevant information concerning with I is contained in a Gröbner basis G.*

In order to make effective the former idea, we need to develop some fundamental algorithms. In a left PBW ring R, every non-zero element has a standard representation

$$f = a_\alpha \mathbf{x}^\alpha + \sum_{\beta \prec \alpha} c_\beta \mathbf{x}^\beta,$$

where $\alpha = \exp(f)$ and $0 \neq a_\alpha \in \mathbf{k}$ is called the *leading coefficient*; notation $a_\alpha = \mathrm{lc}(f)$. The monomial $a_\alpha \mathbf{x}^\alpha$ is called *leading monomial*; notation $a_\alpha \mathbf{x}^\alpha = \mathrm{lm}(f)$.

Assume a left ideal I of R given by a finite set of generators $F = \{f_1, \ldots, f_s\}$. By the Division Algorithm, given $f \in R$ we can compute $h_1, \ldots, h_s, r \in R$ (with r 'smaller than the f_i's in certain sense) such that

$$f = h_1 f_1 + \cdots + h_s f_s + r$$

Call r the *remainder* of the division of f by F; notation $r = \overline{f}^F$. One could expect that $f \in I$ if and only if $r = 0$. But this is not the case unless F is a Gröbner basis for I:

Proposition Let G be a Gröbner basis for a left ideal I of R and $f \in R$. Then $f \in I$ if and only the remainder of the division of f by G is zero. In particular, every Gröbner basis of I is a set of generators and, hence, every left ideal of a left PBW ring is finitely generated.

Thus, if we are able to compute a Gröbner basis for I from the given set of generators F, then we can solve effectively the membership problem for I.

Algorithm 2 Multivariable Division Algorithm

INPUT: $f, f_1, \ldots, f_s \in R$ with $f_i \neq 0$ $(1 \leq i \leq s)$

OUTPUT: $h_1, \ldots, h_s, r$ such that $f = h_1 f_1 + \cdots + h_s f_s + r$
and $r = 0$ or $\mathcal{N}(r) \cap \bigcup_{i=1}^{s} (\exp(f_i) + \mathbb{N}^n) = \emptyset$ and $\max_{\preceq}\{\exp(h_i) + \exp(f_i)\} = \exp(f)$.

INITIALIZATION: $h_1 := 0, \ldots, h_s := 0, r := 0, g := f$

WHILE $g \neq 0$ **DO**

IF there exists i such that $\exp(g) \in \exp(f_i) + \mathbb{N}^n$

THEN

choose i minimal such that $\exp(g) \in \exp(f_i) + \mathbb{N}^n$

$a_i = \mathrm{lc}(g)(\mathrm{lc}(f_i)^{\exp(g)-\exp(f_i)} q_{\exp(g)-\exp(f_i),\exp(f_i)})^-$

$h_i := h_i + a_i x^{\exp(g)-\exp(f_i)}$

$g := g - a_i x^{\exp(g)-\exp(f_i)} f_i$

ELSE

$r := r + \mathrm{lm}(g)$

$g := g - \mathrm{lm}(g)$

Fortunately, we are in the right framework to import Buchberger's Algorithm (whith some changes, of course).

In order to display our Left Gröbner algorithm, we need the following notation: Given $\alpha = (\alpha_1, \ldots, \alpha_n), \beta = (\beta_1, \ldots, \beta_n) \in \mathbb{N}^n$, write

$$\alpha \vee \beta = (\max\{\alpha_1, \beta_1\}, \ldots, \max\{\alpha_n, \beta_n\}) \in \mathbb{N}^n$$

Definition Let f and g be non-zero elements of R and let $\exp(f) = \alpha$ resp. $\exp(g) = \beta$. With $\gamma = \alpha \vee \beta$, set $a = \mathrm{lc}(x^{\gamma-\alpha} f)^{-1}$ and $b = \mathrm{lc}(x^{\gamma-\beta} g)^{-1}$. We then define the *left S-polynomial* of f and g to be

$$\mathrm{SP}(f, g) = a x^{\gamma-\alpha} f - b x^{\gamma-\beta} g$$

Algorithm 3 Left Gröbner Basis Algorithm

INPUT: $F = \{f_1, \ldots, f_s\} \subseteq L$ with $f_i \neq 0$ $(1 \leq i \leq s)$

OUTPUT: $G = \{g_1, \ldots, g_t\}$, a Gröbner basis for $Rf_1 + \cdots + Rf_s)$.

INITIALIZATION: $G := F$, $B := \{\{f, g\};\ f \neq g \in G\}$

WHILE $B \neq \emptyset$ **DO**

Choose any $\{f, g\} \in B$

$B := B \setminus \{\{f, g\}\}$

$h' := \mathrm{SP}(f, g)$

$h := \overline{\mathrm{SP}(f, g)}^{G}$

IF $h \neq 0$ **THEN**

$B := B \cup \{\{p, h\};\ p \in G\}$

$G := G \cup \{h\}$

Let us consider $F = \{f_1, f_2\}$, where

$$f_1 = xy - y,\ f_2 = y^2 - x \in R,$$

where $R = \mathbb{Q}\{x, y;\ yx = xy + 1,\ \preceq_{deglex}\}$ and let $f = xy^2$.

INITIALIZATION: $p_1^{(0)} := 0$, $p_2^{(0)} := 0$, $r_0 := 0$, $h_0 := xy^2$.

First pass through the WHILE loop:

$$\begin{aligned}
(1,1) &= \exp(f_1) \preceq^n \exp(h_0) = (1,2) \\
p_1^{(1)} &= p_1^{(0)} + y = y \\
p_2^{(1)} &= p_2^{(0)} + 0 = 0 \\
h_1 &= h_0 - yf_1 = y^2 - y \\
r_1 &= r_0 + 0 = 0
\end{aligned}$$

Second pass through the WHILE loop:

$$
\begin{aligned}
(1,1) &= \exp(f_1) \not\leq^n \exp(h_1) = (0,2) \\
(0,2) &= \exp(f_2) \leq^n \exp(h_1) = (0,2) \\
p_1^{(2)} &= p_1^{(1)} + y = y \\
p_2^{(2)} &= p_2^{(1)} + 1 = 1 \\
h_2 &= h_1 - f_2 = x - y \\
r_2 &= r_1 + 0 = 0
\end{aligned}
$$

Third pass through the WHILE loop:

$$
\begin{aligned}
(1,1) &= \exp(f_1) \not\leq^n \exp(h_2) = (0,1) \\
(0,2) &= \exp(f_2) \not\leq^n \exp(h_2) = (0,1) \\
p_1^{(3)} &= p_1^{(2)} + 0 = y \\
p_2^{(3)} &= p_2^{(2)} + 0 = 1 \\
h_3 &= h_2 + y = x \\
r_3 &= r_2 - y = -y
\end{aligned}
$$

Fourth pass through the WHILE loop:

$$\begin{aligned}
(1,1) &= \exp(f_1) \not\leq^n \exp(h_3) = (1,0) \\
(0,2) &= \exp(f_2) \not\leq^n \exp(h_3) = (1,0) \\
p_1^{(4)} &= p_1^{(3)} + 0 = y \\
p_2^{(4)} &= p_2^{(3)} + 0 = 1 \\
h_4 &= h_3 - x = 0 \\
r_4 &= r_3 + x = x - y
\end{aligned}$$

The WHILE loop stops, and we get

$$^F\overline{f} = x - y$$

and

$$f = yf_1 + f_2 + (x - y).$$