Some improvements of a lemma of Rosenfeld

François Boulier
University of Waterloo, Symbolic Computation Group
Waterloo, Ontario, N2L 3G1
fboulier@daisy.uwaterloo.ca, boulier@lifl.fr

Submitted to IMACS'96
Technical Report CS-96-13

Abstract

We give some improvements of a lemma of Rosenfeld which permit us to optimize some algorithms in differential algebra: we prove the lemma with weaker hypotheses and we demonstrate an analogue of Buchberger's second criterion, which avoids non necessary reductions for detecting coherent sets of differential polynomials. We try also to clarify the relations between the theorems in differential algebra and some more widely known results in the Gröbner bases theory.

Keywords. Differential algebra. Rewrite systems. Buchberger's criteria. Polynomial differential equations. Rosenfeld's lemma.

1 Introduction

Stated in 1959 by Rosenfeld [Ro59, lemma, page 397], the lemma we improve in this paper can be viewed as a manifestation in differential algebra[1] of the famous Knuth–Bendix theorem [Ev51] [KB67] in term algebras[2]:

Theorem 1 *A nœtherian rewrite system is locally confluent if and only if it is confluent over its critical pairs.*

Proof. See [KB67]. □

Manifestations of theorem 1 arise in many different areas of computer algebra, providing canonical simplifiers which allow to compute in factor structures. See [BL82] for a survey. A well–known example in commutative algebra is the Gröbner basis algorithm [Bu70] which allows to compute in multivariate polynomial rings factored by their ideals. However, one should notice that the proofs of the Knuth–Bendix like theorems can not always be obtained by specializing the one of theorem 1 [BL82, page 37]: special proofs are often necessary.

[1] The reference books are [Ko73] and [Ri50]. We make precise in section 3 the notations and definitions used in this introduction.

[2] We do not recall in this paper the definitions used in the rewrite systems theory. See [BL82] for example.

Rosenfeld's lemma provides algorithmic tools for studying systems of differential polynomial equations. It is applied in the recent algorithm Rosenfeld–Gröbner [Bo94] and [BLOP95] which gathers as entry any ranking and any finite system of differential equations and computes a representation of the radical of the differential ideal generated by the system which can be used afterwards to decide membership in that ideal, through simple reductions.

Before Rosenfeld, Seidenberg demonstrated [Se56, theorem 6, page 51] a slightly weaker version of the lemma to design an elimination algorithm for polynomial PDE. Ritt proved in [Ri50, I, 12, page 30] its most basic case to study the differential algebraic variety associated with a mere algebraically irreducible ODE. Kolchin gave another version in [Ko73, III, 8, lemma 5, page 137] which applies for differential polynomial rings of characteristic non zero or with coefficients in a ring which is not necessarily a field.

In this paper

- we prove Rosenfeld's lemma (lemma 5) under weaker hypotheses than in the original version (lemma 3);
- we prove an analogue of Buchberger's second criterion [Bu79] [BW91, proposition 5.70] which avoids some computations while checking whether a system verifies the conditions of Rosenfeld's lemma.

To show the usefulness of our results, we are going to study the radical of the differential ideal generated by the following system A of $\mathbb{Q}\{u,v\}$ endowed with derivations w.r.t. to x and y. To compute this representation, we apply the idea of the Rosenfeld–Gröbner algorithm [BLOP95] and we explain which computations are avoided when our optimizations are applied.

$$A \left\{ \begin{array}{lcl} p_1 & = & v u_{xx}^3 + u_{xx}^2 + u_x \\ p_2 & = & u_{xy} \\ p_3 & = & u_{yy} + u_y^2 \\ p_4 & = & v_y \\ p_5 & = & v_{xxx} + u_{xx}^3 \end{array} \right.$$

We fix any ranking such that u_{xx}, u_{xy}, u_{yy}, v_y and v_{xxx} are the leaders of p_1, p_2, p_3, p_4 and p_5 respectively.

The system A does not satisfy the conditions of the original Rosenfeld's lemma for it is not autoreduced. The Rosenfeld–Gröbner algorithm must then reduce p_5 by p_1 applying the rule

$$u_{xx}^3 \longrightarrow -\frac{u_{xx}^2 + u_x}{v}$$

under the assumption that $v \neq 0$ and consider separately the solutions of A which also annihilate v.

Though not autoreduced, the system is differentially triangular (definition 2) and lemma 5 proves that Rosenfeld's lemma holds for such situations. Hence, the Rosenfeld–Gröbner algorithm does not need to split the system anymore.

The algorithm must check that the four Δ–polynomials below generated by A are reduced to zero by A. They are actually reduced to zero by p_2 and p_4. Rosenfeld–Gröbner does not need to split the system if it use them for reducing, since these polynomials have trivial initials and separants.

$$\left\{\begin{array}{lcl} \Delta_{12} & = & v_y u_{xx}^3 + u_{xy} \\ \Delta_{23} & = & u_y u_{xy} \\ \Delta_{13} & = & (6vu_{xx}+2)u_{xxy}^2 + \\ & & (6v_y u_{xx}^2 - 6vu_{xx}^2 u_y - 4u_{xx}u_y)u_{xxy} + \\ & & v_{yy}u_{xx}^3 - 6vu_{xx}^2 u_{xy}^2 - 4u_{xx}u_{xy}^2 + u_{xyy} \\ \Delta_{45} & = & u_{xx}^2 u_{xxy} \end{array}\right.$$

The most painful computation is the reduction of Δ_{13}. However, lemma 8 proves that this reduction is useless for Δ_{12} and Δ_{23} are both reduced to zero and the least common derivative of the leaders of p_1 and p_3 is a derivative of the leader of p_2.

Lemma 8 is an analogue of Buchberger's second criterion for differential algebra. As for Gröbner bases, given n differential polynomials, this criterion allows in the best case to perform only $n-1$ reductions instead of $n(n-1)/2$. It is also quite obvious that the avoided reductions are the most painful ones, as illustrated by the example.

Last, the original version of Rosenfeld's lemma imposes to the Rosenfeld–Gröbner algorithm to put as inequations ($\neq 0$) all the initials and separants of the equations of A. With technical words: the original lemma is stated for the ideal $[A]:{H_A}^\infty$. Since system A has one non trivial separant $s_1 = 3vu_{xx}^2 + 2u_{xx}$ and one non trivial initial $i_1 = v$ (the ones of p_1) the Rosenfeld–Gröbner algorithm must consider separately three cases: $A = 0,\ s_1 \neq 0, i_1 \neq 0$ on which Rosenfeld's lemma applies, $A = 0,\ s_1 \neq 0, i_1 = 0$ and $A = 0,\ s_1 = 0$.

Our version of Rosenfeld's lemma only imposes to put as inequations the separants of the equations of A. With technical words: our version is stated for the ideal $[A]:{S_A}^\infty$. Thus we only need to consider two cases instead of three: $A = 0,\ s_1 \neq 0$ on which our version of Rosenfeld's lemma applies and $A = 0,\ s_1 = 0$.

Plan. Rosenfeld's lemma is a quite technical theorem of a mathematical theory, differential algebra, which is much less known than the Gröbner bases theory by the researchers in computer algebra. To give to this paper an audience as wide as possible, we thus recall, in section 2, some of the definitions and theorems of the Gröbner bases theory. In the following sections, we present our definitions and our results in relation with the ones of section 2. In section 3 Ritt's reduction algorithms are specified and the critical pairs which can arise, the Δ–polynomials, are defined. The new version of Rosenfeld's lemma is proven in section 4 where we state also the original one. In the last section, we consider the problem of testing the coherence. In particular, we prove the analogue of Buchberger's second criterion.

2 Gröbner bases

Most of the material of this section is borrowed from [BW91, 5, pages 218–225]. In this section R denotes a polynomial ring over a commutative field of characteristic zero.

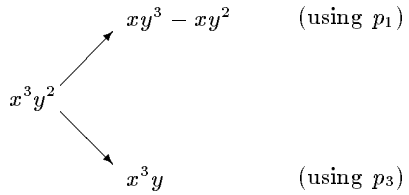
The Gröbner bases algorithms in R are based on *monomial reduction* algorithms: they interpret polynomials $p \in R$ as rewrite rules of the form "a distinguished monomial of p is rewritten into the sum of the other monomials of p". The distinguished monomial is called the leading monomial of p. It is defined by fixing a so–called admissible ordering over the set of all the terms[3] of the polynomial ring. For instance the following system of $\mathbb{Q}[x,y]$

$$G \left\{\begin{array}{l} p_1 = x^3 - xy + x, \\ p_2 = x^2 y, \\ p_3 = xy^2 - xy \end{array}\right.$$

may be viewed as a set of three rewrite rules:

$$\begin{array}{ll} p_1: & x^3 \to xy - x, \\ p_2: & x^2 y \to 0, \\ p_3: & xy^2 \to xy. \end{array}$$

Some critical pairs may occur in such rewrite systems. On the example above, we have in particular:



Now, if we denote $S_{13} = (-xy^3 + xy^2) + x^3 y$, the system G is confluent over the critical pair if the S–polynomial S_{13} is reduced to zero by G.

A formalism was adopted [BW91, 5.4, page 218] which can be viewed as a specification of the reduction algorithms and can be applied to state the theorems in the Gröbner bases theory. The use of this formalism makes the statements a little bit complicated but allows to write down the proofs much easilier. We give it in the definition below and use it to state the next lemmas, since we perform such a formalization in differential algebra.

Definition 1 *Let t be a term in the polynomial ring R. A polynomial $q \in R$ is said to have a t–representation w.r.t. a finite subset G of R if q can be written as a finite (possibly empty) sum of terms mp, where m is a monomial, $p \in G$ and the leading term of mp is less than or equal to t (according to the fixed admissible ordering). An empty sum is defined to be zero.*

Hence if a polynomial p is reducible to zero by G then p has a t–representation w.r.t. G, where t is the leading term of p.

Let p_i and p_j be two polynomials of R, with leading terms m_i and m_j. If we denote m_{ij} the least common multiple between m_i and m_j and S_{ij} the S–polynomial between p_i and p_j (we assume their leading coefficients equal to 1):

$$S_{ij} = \frac{m_{ij}}{m_i} p_i - \frac{m_{ij}}{m_j} p_j,$$

then we can state the following characterization of Gröbner bases [BW91, theorem 5.64]:

[3] Terms are power products of indeterminates of R while monomials are terms multiplied by a coefficient of the base field.

Theorem 2 *Let G be a finite set of polynomials. If for each S–polynomial S_{ij} which can be formed between any two elements of G, there exists some term $t < m_{ij}$ such that S_{ij} has a t–representation w.r.t. G then G is a Gröbner basis.*

2.1 Buchberger's criteria

For the sake of efficiency, it is a key problem to predict that some S–polynomials vanish without having to reduce them. For this purpose, Buchberger established two criteria [Bu79].

Lemma 1 *(Buchberger's first criterion)*
If p_i and p_j are two polynomials of R whose leading terms m_i and m_j have no common factor (i.e. $m_{ij} = m_i m_j$) then S_{ij} has a t–representation w.r.t. the set $\{p_i, p_j\}$ for some term $t < m_{ij}$.

See also [BW91, lemma 5.66]. Remark that in term algebras, the lemma is obvious since two rewrite rules whose heads do not overlap cannot interfere. This is *a priori* not obvious in commutative algebra: the proof in [BW91] first establishes that the subtraction $m_j p_i - m_i p_j$ does not cancel any monomial of the two polynomials $m_j p_i$ and $m_i p_j$.

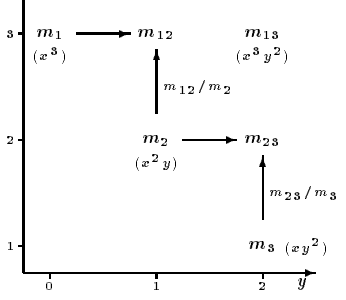
Lemma 2 *(Buchberger's second criterion)*
Let p_i, p_j and p_k be three polynomials of a subset G of R such that the leading term m_j of p_j divides the least common multiple m_{ik} of the leading terms m_i and m_k of p_i and p_k.
If S_{ij} has a t–representation w.r.t. G for some term $t < m_{ij}$ and S_{jk} has a t'–representation w.r.t. G for some $t' < m_{jk}$ then S_{ik} has a t''–representation w.r.t. G for some $t'' < m_{ik}$.

See for instance [BW91, proposition 5.70]. It is not enough to notice that $S_{ik} = \frac{m_{ik}}{m_{ij}} S_{ij} + \frac{m_{ik}}{m_{jk}} S_{jk}$ to show lemma 2 since, if p and q are two polynomials reduced to zero by a set G then $p+q$ is not necessarily reduced to zero. A more subtle analysis and the use of a concept similar to the t–representations are necessary.

We illustrate Buchberger's second criterion on the example given in the beginning of the section. One verifies easily that S_{12} and S_{23} (hence S_{13}) are all reduced to zero by $G = \{p_1, p_2, p_3\}$.

$p_1 = x^3 - xy + x,$
$p_2 = x^2 y,$
$p_3 = xy^2 - xy.$

$S_{12} = -xy^2 + xy,$
$S_{23} = x^2 y,$
$S_{13} = -xy^3 + xy^2 + x^3 y.$



3 Differential algebra

3.1 Basic preliminaries

Starting from this section, K denotes a differential field of characteristic zero endowed with a certain number of derivations denoted $\delta_1, \ldots, \delta_m$ which commute pairwise. We denote derivation operators using greek letters e.g. $\theta = \delta_1^{a_1} \cdots \delta_m^{a_m}$ where the a_i are nonnegative integers and, if v is any element of K, we denote θv the element of K obtained by differentiating it a_i times by δ_i for all $1 \leq i \leq m$. The sum of the exponents a_i is called the *order* of the operator θ. The identity operator is of order 0. All other operators are said to be *proper*. If $\theta = \delta_1^{a_1} \cdots \delta_m^{a_m}$ and $\phi = \delta_1^{b_1} \cdots \delta_m^{b_m}$ then $\theta\phi = \delta_1^{a_1+b_1} \cdots \delta_m^{a_m+b_m}$. If $a_i \geq b_i$ for $i = 1, \ldots, m$ then $(\theta/\phi) = \delta_1^{a_1-b_1} \cdots \delta_m^{a_m-b_m}$. The monoid of derivation operators is denoted Θ. If E is any subset of K, we denote ΘE the smallest subset of K stable under differentiation.

Let S be a subset of a differential ring R which contains K. We denote $K\{S\}$ the smallest differential subring of R containing K and S. We have $K[\Theta S] = K\{S\}$.

We work with differential polynomials in the differential polynomial ring $R = K\{u_1, \ldots, u_n\}$. The u_j are called *differential indeterminates* and the θu_j are called *u–derivatives*. The set of the u–derivatives is denoted ΘU.

An order $\mathcal{R}$ over ΘU is said to be a *ranking* [Ko73, I, 8, page 75] if it is total and compatible with the differentiations over the alphabet:

1. $\delta v > v$ (for all derivation δ and $v \in \Theta U$),
2. $v > w \Rightarrow \delta v > \delta w$ (for all derivation δ and $v, w \in \Theta U$).

Rankings such that $\mathrm{ord}(\theta) > \mathrm{ord}(\phi) \Rightarrow \theta v > \phi w$ (for all derivations operators θ, ϕ and all differential indeterminates v, w) are called *orderly*. Rankings such that $v > w \Rightarrow \theta v > \phi w$ (for all derivations operators θ, ϕ and all differential indeterminates v, w) are called *elimination* rankings. For more terminology, see [Ko73, page 75].

Rankings are well–orderings over ΘU [Ko73, page 75]. They are the analogue of the admissible orderings in the Gröbner bases theory. An important difference: rankings rank indeterminates while admissible orderings rank terms.

Let p be a polynomial[4] of R and $\mathcal{R}$ a ranking over ΘU. The *leader* v of p is the greatest u–derivative w.r.t. the ranking $\mathcal{R}$ which appears in p. The two conditions mentioned above imply that for each derivation operator ϕ, the leader of ϕp is ϕv. Let d be the degree of v in p. The *initial* i_p of p is the coefficient of v^d in p. The *separant* s_p of p is the initial of all the proper derivatives of p ($s_p = \partial p / \partial v$). The *rank* of a polynomial $p = i_p \cdot v^d + r$ is the polynomial v^d. The rank of a set of polynomials is the set of ranks of the elements of the set.

Let p and q be two polynomials with ranks v^d and w^e. The polynomial q is said to be *less* than p if $w < v$ or $w = v$ and $e < d$; *partially reduced* w.r.t. p if no proper derivative of v appears in q; and *reduced* w.r.t. p if q is partially reduced w.r.t. p and its degree in v is less than d.

A set of differential polynomials A is said to be *autoreduced* if each element of A is reduced w.r.t. every other element of the set. Every autoreduced set is finite [Ko73, page 77].

An autoreduced subset A of a set E of polynomials is called a *characteristic set*[5] of E if E does not contain any non zero element reduced w.r.t. A. Every set admits a characteristic set.

The proof of this, of the finiteness of autoreduced sets and of the fact that rankings are well–orderings all rely on the same argument summarized in [Ko73, 0, 17, lemma 15, page 49]. A more compact proof can be found in [Bo94, lemme 2, page 11].

[4] The definitions which we give are only valid for polynomials $p \notin K$. In this paper, we do not need to bother with the exceptions $p \in K$.

[5] This definition corresponds to Ritt's one [Ri50, I, 5, page 5] and coincides with Kolchin's when E is a differential ideal. Kolchin only defined characteristic sets for ideals [Ko73, I 10, page 81 and III, 2, page 124].

We want to consider more general sets than autoreduced ones. Definition 2 exists neither in [Ri50] nor in [Ko73].

Definition 2 *A set of differential polynomials is said to be* differentially triangular *if its elements are pairwise partially reduced and if the leaders of its elements are pairwise different.*

Every differentially triangular set is finite and every autoreduced set is differentially triangular[6]. If the leaders v and v' of two elements of a differentially triangular set A have a least common derivative w, then w is a *proper* derivative of both v and v'.

3.2 Ritt's reduction algorithms

Ritt's reduction algorithms are euclidian division algorithms, extended to differential algebra. We illustrate them on an example. Consider for instance, the elements of the following subset of $\mathbb{Q}\{u,v\}$ endowed with derivations w.r.t. x and y:

$$A \left\{ \begin{array}{lcl} p_1 & = & v u_{xx} - u_x, \\ p_2 & = & u_{yy}^2 - 1. \end{array} \right.$$

If we choose u_{xx} to be the leader of p_1 then the differential polynomials of the set A stand for the rewrite rules:

$$\begin{array}{lll} p_1: & u_{xx} \to \frac{u_x}{v}, & u_{xxy} \to -\frac{v_y u_{xx} - u_{xy}}{v}, \\ & & u_{xxyy} \to -\frac{2v_y u_{xxy} + v_{yy} u_{xx} - u_{xyy}}{v} \quad \cdots \\ p_2: & u_{yy}^2 \to 1, & u_{xyy} \to 0, \quad u_{xxyy} \to -\frac{u_{xyy}^2}{u_{yy}} \quad \cdots \end{array}$$

In practice, has a differential polynomial q to be reduced by a set A, it is multiplied by appropriate powers of the initials and the separants of the elements of A, to avoid denominators.

Many such algorithms exist [Ko73, page 77] [Ri50, I, 6, page 5] [Ma91] which may produce different results. Morrison proved [Mo95] that all Ritt's reduction algorithm terminate.

We need to introduce a few definitions and notations before to give some precise specifications of Ritt's algorithms of reduction which will play in differential algebra the role of the t–representations in the Gröbner bases theory.

Definition 3 *If $\mathfrak{b}$ is a (non necessarily differential) ideal and S is a finite subset of a ring R then $\mathfrak{b} : S^\infty$ denotes the ideal of all the elements p of R such that, for some $h \in S$, the element hp belongs to $\mathfrak{b}$.*

Definition 4 *If A is any finite subset of R we denote*
H_A the set of all the initials and separants of the elements of A and
S_A the set of all the separants of the elements of A.

Definition 5 *If A is any finite subset of R and v is any u–derivative, we denote A_v the set of the derivatives of the elements of A whose leaders are less than or equal to v:*

$$A_v = \{\theta p \mid p \in A \text{ and } \theta p \leq v\}.$$

[6] Actually, autoreduced = differentially triangular + a constraint on the degrees of the polynomials.

If $v < w$ are any u–derivatives we have $A_v \subset A_w$ and for any derivation operator θ, if $p \in (A_v)$ then $\theta p \in (A_{\theta v})$.

Let q be a differential polynomial and A be any finite subset of R. *Ritt's full reduction* consists first in a *partial reduction* (i.e. purely differential) followed by a purely *algebraic reduction*:

$$q \textit{ full–rem } A = (q \textit{ partial–rem } A) \textit{ alg–rem } A.$$

Let $r_0 = q$ *partial–rem* A. A specification of the partial reduction is:

1. r_0 is partially reduced w.r.t. all the elements of A.
2. there exists a power product h of elements of S_A such that $hq \equiv r_0 \pmod{(A_v)}$, where v is the leader of q.

Let $r_1 = r_0$ *alg–rem* A. Specification of the full reduction:

1. r_1 is reduced w.r.t. all the elements of A.
2. there exists a power product h of elements of H_A such that $hq \equiv r_1 \pmod{(A_v)}$, where v is the leader of q.

We have $q \in [A] : {H_A}^\infty$ iff $(q \textit{ full–rem } A) \in [A] : {H_A}^\infty$.

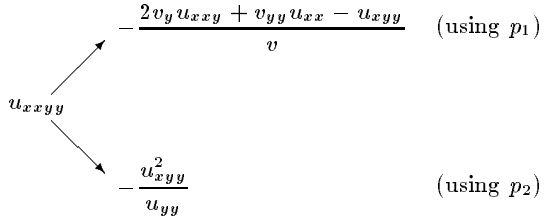
In general however, Ritt's reduction algorithm does not preserve the equivalence modulo $[A]:{H_A}^\infty$. Therefore, Rosenfeld's lemma only deals with the equivalence to zero modulo $[A] : {H_A}^\infty$ and not with normal forms modulo this ideal — while Gröbner bases do. This remark concerns also partial reductions and ideals $[A] : {S_A}^\infty$.

Moreover, even if $q \in [A] : {H_A}^\infty$ then (q *full–rem* A) is not necessarily syntactically zero. Consider for instance the set A which only contains the differential polynomial $p = (u+1)^2(u-1)$. The separant of p contains $(u+1)$ as a factor, so $(u-1) \in [A] : {H_A}^\infty$ but is irreducible by A.

This is related to the fact that every autoreduced and coherent set A is not necessarily a characteristic set of the differential ideal $[A] : {H_A}^\infty$ that it defines.

3.3 Delta–polynomials

Some critical pairs may occur between the differential polynomials of a set A, when they are considered as rewrite rules for Ritt's reduction algorithm. They only arise in systems of PDE when different rules rewrite different derivatives of a same differential indeterminate. On the example given in section 3.2 we have for instance:



Now, if we denote $\Delta_{12} = u_{yy}(2v_y u_{xxy} + v_{yy} u_{xx} - u_{xyy}) - v(u_{xyy}^2)$, the system A is confluent over the critical pair if the Δ–polynomial Δ_{12} is reduced to zero by A, using Ritt's reduction.

The Δ–polynomials correspond to the S–polynomials in the Gröbner bases theory. This terminology comes from Rosenfeld who denoted them Δ_{ij} as we do. They are called *differential S–polynomials* in [Ma91] but we find this misleading since there exists generalizations[7] [Ca87] [Ol90] of Gröbner bases to the differential case, based on a monomial reduction, where this denomination fits better.

[7] However, such Gröbner bases are generally infinite.

Definition 6 *Let p_i and p_j be two differential polynomials of some differentially triangular subset A of R whose leaders $\theta_i u$ and $\theta_j u$ have some common derivatives. Denote $\theta_{ij} u$ the common derivative of $\theta_i u$ and $\theta_j u$ of least order.*

The Δ–polynomial *between p_i and p_j is*

$$\Delta_{ij} = s_j(\theta_{ij}/\theta_i)p_i - s_i(\theta_{ij}/\theta_j)p_j.$$

The set of all the Δ–polynomials which can be formed between any two elements of A is denoted $\Delta(A)$.

If θu is any derivative of $\theta_{ij} u$ we denote

$$\Delta_{ij}^{\theta} = s_j(\theta/\theta_i)p_i - s_i(\theta/\theta_j)p_j.$$

It follows from the definition of differentially triangular sets (definition 2) that the leader of any cross derivative Δ_{ij}^{θ} is less than θu.

4 Rosenfeld's lemma

Rosenfeld's lemma is an analogue in differential algebra, for Ritt's reduction algorithms of the theorem 2 for Gröbner bases in commutative algebra,

We bring two changes to the original version of Rosenfeld: we state the result for differentially triangular sets instead of autoreduced ones and we state it for the quotient $[A] : S^{\infty}$ by any set S which contains S_A instead of H_A. We compare our version with Kolchin's one.

4.1 The original version

In [Ro59, page 397] an autoreduced set A is defined to be *coherent* if, for all elements p_i and p_j of A whose leaders have a least common derivative v, there exists a u–derivative $w < v$ such that $\Delta_{ij} \in (A_w) : {H_A}^{\infty}$. Rosenfeld's original lemma can then be stated as:

Lemma 3 *(Rosenfeld's lemma, version of 1959)*
If A is an autoreduced and coherent subset of R then every differential polynomial q partially reduced w.r.t. A which belongs to $[A] : {H_A}^{\infty}$ belongs also to $(A) : {H_A}^{\infty}$.

4.2 The new version

We first generalize the definition of the coherence.

Definition 7 *Let A be any differentially triangular subset of R and S be any finite subset of R which contains S_A and which is partially reduced w.r.t. A. The set A is said to be* coherent by inverting S *if, for all elements p_i and p_j of A whose leaders have a least common derivative v, there exists a u–derivative $w < v$ such that $\Delta_{ij} \in (A_w) : S^{\infty}$.*

Assuming S to be partially reduced w.r.t. A is necessary for Rosenfeld's lemma. The sets S_A and H_A both satisfy this constraint. An autoreduced set A is coherent in the sense of Rosenfeld if it is coherent by inverting H_A in the sense of the definition 7.

The following lemma is used for proving both Rosenfeld's lemma (lemma 5) and the analogue of Buchberger's criterion (lemma 8). Seidenberg and Rosenfeld proved it inside the proofs of [Se56, theorem 6, page 51] and [Ro59, lemma, page 397]. Kolchin did not need to prove it since one of his hypotheses [Ko73, condition C3, page 136] implies it.

Lemma 4 *Let p_i and p_j be two elements of a differentially triangular set $A \subset R$ with leaders $\theta_i u$ and $\theta_j u$. Let $\theta_{ij} u$ be the least common derivative of $\theta_i u$ and $\theta_j u$. Let S be any finite subset of R which contains S_A.*

If there exists a u–derivative $v < \theta_{ij} u$ such that $\Delta_{ij} \in (A_v) : S^{\infty}$ then for every derivative θu of $\theta_{ij} u$, there exists a u–derivative $w < \theta u$ such that $\Delta_{ij}^{\theta} \in (A_w) : S^{\infty}$.

Proof. The proof is an induction on the order of (θ/θ_{ij}).

If the order is zero then Δ_{ij}^{θ} is the Δ–polynomial Δ_{ij} and the lemma is verified by hypothesis.

If the order is nonzero, we decompose $\theta = \delta\phi$ where δ is a mere derivation and (ϕ/θ_{ij}) exists and we assume (induction hypothesis) that there exists a power product h of elements of S and $v < \phi u$ such that $h\Delta_{ij}^{\phi} \in (A_v)$.

Consider the polynomial $\delta(h\Delta_{ij}^{\phi})$. The second condition satisfied by the rankings implies that it belongs to the ideal $(A_{\delta v})$ and that $\delta v < \delta\phi u = \theta u$. Multiply that polynomial by h. One obtains $(\delta h)h\Delta_{ij}^{\phi} + h^2(\delta\Delta_{ij}^{\phi})$ whose first term is in (A_v) by induction hypothesis. Since $(A_v) \subset (A_{\delta v})$ we conclude that $h^2(\delta\Delta_{ij}^{\phi})$ belongs to this latter ideal. Develop this polynomial:

$$\begin{aligned} h^2(\delta\Delta_{ij}^{\phi}) &= h^2\delta\Big\{s_j(\phi/\theta_i)p_i - s_i(\phi/\theta_j)p_j\Big\} && (1)\\ &= h^2\Big\{(\delta s_j)(\phi/\theta_i)p_i - (\delta s_i)(\phi/\theta_j)p_j\Big\} && (2)\\ &+ h^2\Delta_{ij}^{\theta}. && (3)\end{aligned}$$

The polynomials $(\phi/\theta_i)p_i$ and $(\phi/\theta_j)p_j$ have both $\phi u < \theta u$ for leaders. Denote $w = \max(\phi u, \delta v)$. The term (2) is thus in (A_w) and $\Delta_{ij}^{\theta} \in (A_w) : S^{\infty}$. Since $w < \theta u$ the lemma is proven. □

If Δ_{ij} *full–rem* $A = 0$ then there exists a $v < \theta_{ij} u$ (see lemma 7) such that $\Delta_{ij} \in (A_v) : S^{\infty}$ (assuming S contains the initials involved in the reduction) but Δ_{ij}^{θ} *full–rem* A can be different than zero. Take the following system A of $\mathbb{Q}\{u, v, w\}$ endowed with derivations w.r.t x, y and z, which is differentially triangular for any orderly ranking. The Δ–polynomial $\Delta_{12} = v_y$ is reduced to zero by A but the remainder of $\Delta_{12}^{z} = \delta_z\delta_y p1 - \delta_z\delta_x p2 = v_{yz}$ by p_4, equal to w_y, is irreducible.

$$A\begin{cases} p_1 &= u_x + v, \\ p_2 &= u_y, \\ p_3 &= v_y, \\ p_4 &= v_z + w. \end{cases}$$

Definition 8 *An ideal $[A] : S^{\infty}$ of R is called a* regular differential ideal *for a ranking $\mathcal{R}$ if A is differentially triangular w.r.t. the ranking, S is any finite subset of R which contains S_A and which is partially reduced w.r.t. A, and A is coherent by inverting S.*

Lemma 5 *(Rosenfeld's lemma, new version)*
If $[A] : S^{\infty}$ is a regular differential ideal of R then every differential polynomial q partially reduced w.r.t. A which belongs to $[A] : S^{\infty}$ belongs also to $(A) : S^{\infty}$.

Proof. Let $q \in [A] : S^{\infty}$ be a polynomial partially reduced w.r.t. A. There exists some power product h of elements of S and a u–derivative θu such that

$$hq = \sum_{\phi p_j \in A_{\theta u}} B_{j,\phi}\ \phi p_j \qquad (4)$$

Since rankings are well–orderings, we may assume that the formula (4) is such that θu is minimal. Thus θu is necessarily a derivative of the leaders $\theta_1 u, \ldots, \theta_i u$ of some (at least one) elements $p_1, \ldots, p_i$ of A, renaming the p's if necessary. We assume that $q \notin (A) : S^\infty$, hence that θu is a *proper* derivative of $\theta_1 u, \ldots, \theta_i u$, and seek a contradiction.

Denote $(\theta/\theta_i)p_i = s_i\theta_u + r$, apply on the terms $B_{j,\phi}\ \phi p_j$ of the sum (4) the substitution $\theta u \to \frac{(\theta/\theta_i)p_i - r}{s_i}$ and multiply then by some power s_i^α to erase denominators. For some $v < \theta u$ we have

$$s_i^\alpha hq = D\ (\theta/\theta_i)p_i \quad (5)$$

$$+ \sum_{j=1}^{i-1} E_j\ \Delta_{ij}^\theta \quad (6)$$

$$+ \sum_{\phi p_j \in A_v} C_{j,\phi}\ \phi p_j \quad (7)$$

The terms in the sums (6) and (7) are free of θu. Since S and q are partially reduced w.r.t. A, the substitution does not apply on hq which is also free of θu. Therefore $D = 0$.

If A is a system of ODE the sum (6) is empty and $hq \in (A_v)$. Since $v < \theta u$ we have a contradiction.

Assume A is a PDE system. Since it is coherent by inverting S, according to lemma 4, there exists a u–derivative $w < \theta u$ such that the sum (6) belongs to $(A_w) : S^\infty$. Thus the polynomial $hq \in (A_r) : S^\infty$ where $r = \max(v, w) < \theta u$. This contradiction proves the lemma. □

In [Ko73, III, 8, page 136] Kolchin modifies Rosenfeld's notion of coherence to the so–called $\mathfrak{k}$–coherence. The $\mathfrak{k}$–coherence does not contain Rosenfeld's coherence since [Ko73, condition C3, page 136] imposes a test, not only on the Δ–polynomials, but on all the cross derivatives generated by the system. Take for instance $A = \{u_x + v,\ u_y\}$ and $\mathfrak{k} = (v_y)$ for any ranking such that $u_x > v$; the Δ–polynomial v_y passes the test of condition C3 while the other cross derivatives do not. Therefore [Ko73, lemma 5, page 137] does not imply Rosenfeld's lemma.

According to [Ko73, remark, page 136] the $\mathfrak{k}$–coherence is useful for differential polynomial rings of characteristic non zero or with coefficients in a ring which is not a field. We do not know if it is algorithmic.

However, the proofs of Seidenberg [Se56, theorem 6, page 51], Rosenfeld [Ro59, lemma, page 397], Kolchin [Ko73, lemma 5, page 137] and lemma 5 involve the same arguments. Only the hypotheses change.

5 Testing the coherence

Lemma 6 *If A is a differentially triangular subset of R and S is a finite subset of R which contains S_A and which is partially reduced w.r.t A then A is coherent by inverting S if and only if, for all $p_i, p_j \in A$ whose leaders have a least common derivative $\theta_{ij}u$, we have*

$$\Delta_{ij} \text{ partial–rem } A \in (A^{ij}) : S^\infty$$

where A^{ij} denotes the set of the elements of A whose leaders are less than $\theta_{ij}u$.

Proof. Denote $D_{ij} = \Delta_{ij}$ *partial–rem* A.

The implication from right to left. If $D_{ij} \in (A^{ij}) : S^\infty$ then, by the specifications of Ritt's partial reduction algorithms, there exists some u–derivative $v < \theta_{ij}u$ (for instance the leader of Δ_{ij}) such that $\Delta_{ij} \in (A_v) : S^\infty$. Thus A is coherent by inverting S.

The implication from left to right. Assume A coherent by inverting S. According to the definition of A^{ij}, the specifications of Ritt's partial reduction algorithms and the fact that the leader of Δ_{ij} is less than $\theta_{ij}u$, we see that $D_{ij} \in (A_v^{ij}) : S^\infty$ for some $v < \theta_{ij}u$ and is partially reduced w.r.t. A^{ij}.

Now, A^{ij} is not necessarily coherent by inverting S for there may exist some $p_k, p_\ell \in A^{ij}$, whose leaders have a least common derivative $\theta_{k\ell}u' > \theta_{ij}u$. However, the proof of lemma 5 still applies: set $q = \Delta_{ij}$ and $\theta u = \theta_{ij}u$; since $\theta_{k\ell}u' > \theta u = \theta_{ij}u$, no cross derivative between p_k and p_ℓ can arise in the sum (6) of the proof of lemma 5. Following this proof, we see that there exists some $h \in S$ such that $hD_{ij} \in (A^{ij})$. Thus $D_{ij} \in (A^{ij}) : S^\infty$. □

Lemma 6 is algorithmic since one can compute a Gröbner basis of $(A^{ij}) : S^\infty$. Compute for instance a Gröbner basis of the set $A_i \cup \{Xh - 1\}$, where h is the product of the elements of S and X is a new indeterminate, for any admissible ordering which eliminates X. The set of the polynomials of the basis which are free of X is a Gröbner basis of $(A^{ij}) : S^\infty$.

Remark that we *cannot* simplify lemma 6 as "*A is coherent by inverting S if and only if, for all $\Delta_{ij} \in \Delta(A)$ we have Δ_{ij}* partial–rem $A \in (A) : S^\infty$" — though this is very tempting ! The implication from left to right is true (by Rosenfeld's lemma) but its converse is wrong, the following example shows.

The system A of $\mathbb{Q}\{t, u, v, w\}$ endowed with derivations w.r.t. x and y is differentially triangular for any elimination ranking such that $t > u > v > w$. It generates only one Δ–polynomial $\Delta_{23} = v_y - w_x$ which belongs to $(A) : {S_A}^\infty$.

$$A \begin{cases} p_1 = t^2 + v_y + w_x, \\ p_2 = u_x + v, \\ p_3 = u_y + w, \\ p_4 = (v_y - w_x)(v_y + w_x). \end{cases}$$

However, the differential ideal $[A]{:}{S_A}^\infty$ contains polynomials partially reduced w.r.t. A which do not belong to $(A) : {S_A}^\infty$. Take for instance the cross derivative $\Delta_{23}^x = \delta_x\delta_y p_2 - \delta_x^2 p_3 = v_{xy} - w_{xx}$. Reducing it by A we get $D_{23}^x = w_{xx}v_y - w_x w_{xx}$ which does not belong to $(A) : {S_A}^\infty$. Thus A is not coherent by inverting S_A.

This example shows also that, even for the special case $\mathfrak{k} = (A)$, Kolchin's [Ko73, condition C3, page 136] does not reduce to a test over the Δ–polynomials.

Lemma 7 *If A is a differentially triangular subset of R such that Δ_{ij}* full–rem *$A = 0$ for all $\Delta_{ij} \in \Delta(A)$ then A is coherent by inverting any finite subset S of R, partially reduced w.r.t. A and which contains both S_A and the initials of the elements of A involved in the algebraic part of the reductions.*

Proof. If Δ_{ij} *full–rem* $A = 0$ then $\Delta_{ij} \in (A_v) : S^\infty$ where $v < \theta_{ij}u$ denotes the leader of Δ_{ij}. □

Lemma 7 is useful for practical purposes but only gives a sufficient condition. The set A below is coherent by inverting S_A for any ranking such that $u_x > v$: the second condition verified by rankings implies $u_{xy} > v_y$ and, since the separant of p_3 contains $v_y + 1$ as a factor, $\Delta_{12} = v_y \in (A_{v_y}) : {S_A}^\infty$. However, v_y is irreducible by A.

$$A \begin{cases} p_1 = u_x + v, \\ p_2 = u_y, \\ p_3 = v_y(v_y + 1)^2 \end{cases}$$

5.1 The analogue of Buchberger's criteria

Buchberger's first criterion has no equivalence for Ritt's reduction in differential algebra. Actually, the leaders of two differential polynomials which have some common derivatives are never disjoint: they share at least the same differential indeterminate.

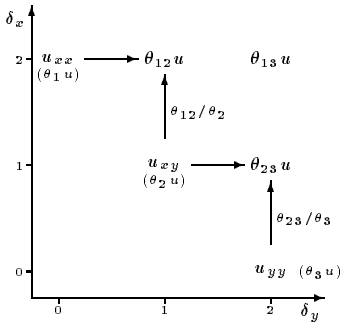
The following lemma is an analogue in differential algebra and for Ritt's reduction algorithms of the lemma 2.

Lemma 8 *Let p_i, p_j and p_k be three differential polynomials of some differentially triangular subset A of R whose leaders $\theta_i u$, $\theta_j u$ and $\theta_k u$ have least common derivatives denoted $\theta_{ij}u$, $\theta_{jk}u$ and $\theta_{ik}u$. Let S be any finite subset of R which contains S_A.*

If there exists u–derivatives $v < \theta_{ij}u$ and $w < \theta_{jk}u$ such that $\Delta_{ij} \in (A_v) : S^\infty$ and $\Delta_{jk} \in (A_w) : S^\infty$ and if $\theta_{ik}u$ is a derivative of $\theta_j u$ then there exists a u–derivative $r < \theta_{ik}u$ such that $\Delta_{ik} \in (A_r) : S^\infty$.

Proof. We have the relation $s_j \Delta_{ik} = s_k \Delta_{ij}^{\theta_{ik}} + s_i \Delta_{jk}^{\theta_{ik}}$. The proof follows from lemma 4 and the fact that $s_j \in S$. □

$$
\begin{array}{lcl}
p_1 & = & v u_{xx}^3 + u_{xx}^2 + u_x, \\
p_2 & = & u_{xy}, \\
p_3 & = & u_{yy} + u_y^2, \\
p_4 & = & v_y, \\
p_5 & = & v_{xxx} + u_{xx}^3, \\
\\
\Delta_{12} & = & v_y u_{xx}^3 + u_{xy}, \\
\Delta_{23} & = & u_y u_{xy}.
\end{array}
$$



The picture illustrates lemma 8 over the example given in introduction. Both Δ_{12} and Δ_{23} are reduced to zero by A. Therefore $\Delta_{12} \in (A_v) : {S_A}^\infty$ (where v denotes the leader of Δ_{12}) and $\Delta_{23} \in (A_{u_{xy}}) : {S_A}^\infty$ and $v < \theta_{12}u = u_{xxy}$ and $u_{xy} < \theta_{23}u = u_{xyy}$, as summarized in lemma 7. The picture shows also that $\theta_{13}u = u_{xxyy}$ is a derivative of $\theta_2 u = u_{xy}$. Thus there exists a u–derivative $w < \theta_{13}u$ such that $\Delta_{13} \in (A_w) : S^\infty$.

Remark that lemma 8 does not prove that Δ_{13} is reduced to zero by A, though it actually is — and has always been for all the examples we have ever tried. We have only proven (after reducing Δ_{45} to zero) that A is coherent by inverting S_A whence Δ_{13} *partial–rem* $A \in (A) : {S_A}^\infty$, by Rosenfeld's lemma. See the remark following lemma 7.

Very recently, we have been aware that Morrison defined [Mo95] a coherence for non triangular systems. For instance, a system Σ which contains a coherent (in our sense) differentially triangular subset A which reduces $\Sigma \setminus A$ to zero is coherent in her sense. However [Mo95] does not contain any algorithm to decide if a system is coherent in this sense or not.

Lemma 8 holds also for non differentially triangular systems in a special case: if the leader of p_1 (or p_3) is a derivative of the leader of p_2 and if the initial of p_1 is equal to its separant. Remark that such restrictions do not exist for Gröbner bases: in lemma 2, some leading terms of the p's may be multiple of some others.

6 Conclusion

We have demonstrated Rosenfeld's lemma under stronger hypotheses: our version does not impose any condition on the degrees of the polynomials. We have also proven an analogue of Buchberger's second criterion. This permits to redesign some algorithms in differential algebra for a better efficiency and, in particular, to apply in this field a part of the reflection put in the Gröbner bases theory. Implementations of these results are being developed in MAPLE.

ACKNOWLEDGEMENTS. The author would like to thank François Ollivier who suggested him this research problem, Michel Petitot for many helpful discussions, in particular about the rewrite systems theory, an anonymous referee of a former version of this article for its fruitful remarks and Sally Morrison who sent me a draft version of her paper.

References

[Bo94] F. Boulier.– *Étude et implantation de quelques algorithmes en algèbre différentielle* (Thèse de l'Université des Sciences et Technologies de Lille, (1994))

[Bo95] F. Boulier.– *An analogue of Buchberger's second criterion in differential algebra* (Publication interne de l'Université des Sciences et Technologies de Lille IT–95–268 (jan. 1995))

[BLOP95] F. Boulier, D. Lazard, F. Ollivier, M. Petitot.– *Representation for the radical of a finitely generated differential ideal* (Proceedings of ISSAC95 (jul. 1995), 158–166)

[Bu70] B. Buchberger.– *An Algorithm for Finding a Basis for the Residue Class Ring of a Zero–Dimensional Polynomial Ideal (German)* (Ph. D. Thesis. Math. Inst. Univ. of Innsbruck, Austria 1965, and Aequationes Math. **4/3** (1970), 374–383)

[Bu79] B. Buchberger.– *A criterion for detecting unnecessary reductions in the construction of Gröbner bases* (EUROSAM'79, An International Symposium on Symbolic and Algebraic Manipulation, Springer LNCS **72**, 3–21)

[BL82] B. Buchberger, R. Loos.– *Algebraic Simplification* (Computer Algebra – Symbolic and Algebraic Computation, B. Buchberger, G. Collins, R. Loos eds. Springer Verlag, Wien–New York (1982), 11–43)

[BW91] T. Becker, V. Weispfenning.– *Gröbner Bases: a computational approach to commutative algebra* (Graduate Texts in Mathematics **141**, Springer Verlag (1991))

[Ca87] G. Carrà–Ferro.– *Gröbner bases and differential ideals* (notes de AAECC5, Menorca, Spain, Springer Verlag (1987), 129–140)

[Ev51] T. Evans.– *The Word Problem for Abstract Algebras* (J. London Math. Soc. **26** (1951) 64–71)

[KB67] D. E. Knuth, P. B. Bendix.– *Simple Word Problems in Universal Algebras* (OXFORD 67, 263–298)

[Ko73] E. R. Kolchin.– *Differential Algebra and Algebraic Groups* (Academic Press, New York (1973))

[Ma91] E. Mansfield.– *Differential Gröbner Bases* (Ph. D. thesis, University of Sydney, (1991))

[Mo95] S. Morrison.– *Pseudo–Reduction, Second Preliminary Draft* (private communication, (dec. 1995))

[Ol90] F. Ollivier.– *Le problème de l'identifiabilité structurelle globale : approche théorique, méthodes effectives et bornes de complexité* (Thèse de doctorat, École Polytechnique (1990))

[Ri50] J. F. Ritt.– *Differential Algebra* (Amer. Math. Soc, New York (1950))

[Ro59] A. Rosenfeld.– *Specializations in differential algebra* (Trans. Amer. Math. Soc. **90** (1959), 394–407)

[Se56] A. Seidenberg.– *An elimination theory for differential algebra* (Univ. California Publ. Math. (N.S.) (1956), 31–65)