Picard-Vessiot and categorically normal extensions in differential-difference Galois theory

G. Janelidze*

Abstract

The connection between categorical and differential Galois theories established by the author (published in 1989) is extended to the context that includes difference Galois theory.

Introduction

In this paper "ring" means "commutative ring with 1" and the same applies to algebras over rings; homomorphisms of rings and algebras over rings are supposed to preserve 1.

Let us begin with what is usually called Grothendieck's formulation of the fundamental theorem of Galois theory:

Theorem 0.1. *If $K \subseteq E$ is a finite Galois field extension and G is its Galois group, then the opposite category of K-algebras A with $E \otimes_K A \approx E^n$ for some natural n is equivalent to the category of finite G-sets.* ■

*Partially supported by South African NRF and Georgian Shota Rustaveli National Science Foundation Grant DI/18/5-113/13.

Received by the editors in March 2016 - In revised form in April 2016.

Communicated by M. Gran.

2010 *Mathematics Subject Classification :* 12H10, 12H05, 18A40, 18A99.

Key words and phrases : Categorical Galois theory, Difference Galois theory, Differential Galois theory, Galois groupoid, Normal extension, Difference equation, Differential equation, Wronski determinant, Casorati determinant.

This theorem has a number of counterparts and generalizations, including Magid's most general such theorem for arbitrary (commutative) rings [18] (the first edition of that book was published in 1974). The first purely-categorical generalization, where Magid's theorem was considered as the main example, was proposed in [6], with several different versions later published as [7], [9], and [10] (see also [1] and [11]). Let us mention here the following one, which was already in [6], except that one of its requirements was simplified a bit in 1991:

Theorem 0.2. *Let $I : \mathbf{C} \to \mathbf{X}$ be a functor between categories with pullbacks that admits a right adjoint H, and let (A, f) be an I-normal extension of B, which means $f : A \to B$ be a morphism in C satisfying the following conditions:*

(a) A is I-admissible, which means that the functor $I^A : (\mathbf{C} \downarrow A) \to (\mathbf{X} \downarrow I(A))$ has a fully faithful right adjoint;

(b) the change-of-base functor $(\mathbf{C} \downarrow B) \to (\mathbf{C} \downarrow A)$ is monadic;

(c) $(A \times_B A, \mathrm{pr}_1) \approx H^A I^A (A \times_B A, \mathrm{pr}_1)$ canonically in $(\mathbf{C} \downarrow A)$ (or, equivalently under (a), $(A \times_B A, \mathrm{pr}_1)$ belongs to the replete image of H^A), where H^A is the right adjoint of I^A.

Then the I-image $I(A \times_B A)$ of the internal equivalence relation $A \times_B A$ on A in the category $\mathbf{C}$ is an internal groupoid in $\mathbf{X}$, and there is a canonical category equivalence

$$Spl_I(A, f) \sim \mathbf{X}^{I(A \times_B A)},$$

in which $Spl_I(A, f)$ is the full subcategory of $(\mathbf{C} \downarrow B)$ with objects all $S \to A$ such that $(A \times_B S, \mathrm{pr}_1) \approx H^A I^A (A \times_B S, \mathrm{pr}_1)$ (canonically or not) in $(\mathbf{C} \downarrow A)$. ∎

Being so much more general than Theorem 0.1, this theorem has many non-trivial examples far away from what A. Grothendieck considered, some of which are discussed in the above-mentioned papers and in many others they refer to. One such example was found as an answer to a question asked by S. Mac Lane:

Actually Mac Lane simply asked me whether differential Galois theory can be obtained as a special case of the categorical one, and while the complete answer would be a very long story, the first most interesting step would be to find out whether Picard-Vessiot extensions of differential fields are I-normal for some naturally chosen functor I, and to find out what kind of Galois groupoids does this present as $I(A \times_B A)$. It turned out [8] that:

- The "naturally chosen functor" I is nothing but the functor from the opposite category of differential rings to the opposite category of rings sending differential rings to their rings of constants.

- Picard-Vessiot extensions are *not* I-normal in general.

- If $K \subseteq E$ is a Picard-Vessiot extension for a (linear homogeneous) differential equation $L(y) = 0$ of degree n, and $u_1, \ldots, u_n$ are linearly independent solutions of that equation, then the differential subring of E generated by K and these solutions is still *not* an I-normal extension of K in general.

- However, in the situation above, the differential subring A of E generated by K, these solutions, and the inverse of their Wronski determinant always is an I-normal extension of K.

- According to Theorem 2, the Galois groupoid of that normal extension should be defined as the 'Hopf algebroid' of the constants of $A \otimes_K A$ (since pullbacks in the opposite category of differential rings are constructed as tensor products). The study of these Galois groupoids and their relationship with the classically defined algebraic Galois groups of Picard-Vessiot extensions has been postponed...

The paper [8] has been recommended for publication by Mac Lane himself, and some years later A. R. Magid mentioned it in [17]. Nevertheless, as far as I know no further results had been obtained in this direction.

The present paper can be considered as a difference-algebraic counterpart of [8], with many arguments copied from there. It actually introduces a generalized-differential-algebraic context in an attempt to unify the differential-algebraic and the difference-algebraic ones, although there are three important places where I was not able to find unified proofs in time (see Remark 2.2); I think it can be done though.

I am grateful to the anonymous referee for pointing out to me that the generalized differential rings I am using were originally defined by L. Guo and W. Keigher in [4] (see also [5] and references therein).

Apart from this introduction the paper has six sections, the first of which defines generalized differential algebras whose special cases are both differential and difference algebras over rings, starting from a naive motivation for that definition. Section 2 is devoted to easy adaptations of the notions of Wronski and Casorati determinants and of their relation with linear independence over constants to our context, while Section 3 calculates their derivatives/'difference'. In fact the last result of Section 3 is not used later, and it is mentioned only for those readers who might be interested in improving what I have done for the above-mentioned unification (see Remark 2.2 again). Sections 4 and 5 are devoted to two lemmas, respectively, needed to prove the main result, which is Theorem 6.2 of the last section. This theorem extends Theorem 2.4 of [8] that is the main result of that paper, first of all to include the case of difference algebras. The formulations look slightly different, but in fact Theorem 6.2 fully generalizes Theorem 2.4 of [8] due to the simplification mentioned above Theorem 0.2; the details are omitted though.

Finally, let us mention some of many further comparisons to be made:

- The next step would be to extend the categorical approach from the Picard-Vession theory to the strongly normal theory due to E. Kolchin [14] (see also [17]). Unfortunately my discussion with Kolchin in Novosibirsk in 1989 about this was too brief.

- Another desirable extension would be towards the J.-F. Pommaret approach [19], although several letters he and I send to each other many years ago did not help.
- M. Takeuchi's paper [20] published at the same year as [8] involves Hopf algebras in the Picard-Vessiot theory, and therefore should be closely related to [8].
- The much more recent paper [2] of Z. Chatzidakis, C. Hardouin, and M. Singer compares various definitions of difference Galois groups, and it would desirable to compare its constructions with our $I(A \times_B A)$; in fact the links between them might go back to the older book of M. Van der Put and M. F. Singer [21].
- Various authors mention *torsors* in differential/difference Galois theory, which also occur of course in categorical Galois theory; this might indicate easy further links with it.

1 Generalized differential algebras

Given a normed ring R, and a fixed element e in it, we could define the derivative $D(f)$ of a map $f : R \to R$ at an element r of R as

$$D(f)(r) = \lim_{x \to e} \frac{f(r+x) - f(r)}{x}, \tag{1.1}$$

which becomes the standard definition when $e = 0$, and which requires, however, some invertibility explanations even in that special case. Omitting that explanation and naively following the usual calculation of $D(fg)$, where fg is the argument-wise product of maps $f, g : R \to R$, we obtain

$$D(fg)(r) = (D(f)(r))\Big(\lim_{x \to e} g(r+x)\Big) + f(r)(D(g)(r)). \tag{1.2}$$

On the other hand, (1.1) applied to g (also naively) gives

$$\lim_{x \to e} g(r+x) = g(r) + e(D(g)(r)), \tag{1.3}$$

and so

$$D(fg)(r) = (D(f)(r))g(r) + f(r)(D(g)(r)) + e(D(f)(r))(D(g)(r)). \tag{1.4}$$

In, say, the classical case of R being the field of real numbers, $e = 0$ gives back the classical derivative, while $e \neq 0$ and having continuous f gives $D(f) = (f(r+e) - f(r))/e$, in which cases (1.4) is classical and straightforward, respectively. In the latter case $e = 1$ would make D the so-called *difference operator*. Using (1.4) as a motivation, we introduce, following L. Guo and W. Keigher [4] (who briefly mention essentially the same motivation and give the same examples as our Examples 1.2 and 1.3 below):

Definition 1.1. (a) Let R be a ring, and e an element in R. A differential (R, e)-algebra is a pair (A, D), in which A is an R-algebra, and $D : A \to A$ an R-module homomorphism with

$$D(1) = 0, \tag{1.5}$$
$$D(ab) = D(a)b + aD(b) + eD(a)D(b), \tag{1.6}$$

for all a and b in A. We shall often write just A instead of (A, D), use the same letter D for all differential (R, e)-algebras, and simply call it derivative, as in the case $e = 0$.

Example 1.2. When $e = 0$ in Definition 1.1, a differential (R, e)-algebra is the same as a differential R-algebra.

Example 1.3. When e, in Definition 1.1, is invertible, (A, D) is a differential (R, e)-algebra if and only if (A, φ) having $\varphi : A \to A$ defined by $\varphi(a) = eD(a) + a$ is a difference R-algebra (we are not requiring φ to be injective).

Example 1.4. Every R-algebra equipped with the zero derivative, that is, the derivative D defined by $D(a) = 0$ for all a in A, is a differential (R, e)-algebra. In particular, R itself equipped with the zero derivative is the initial object in the category of differential (R, e)-algebras.

Remark 1.5. (a) When $e = 0$, (1.5) follows from (1.6), since (classically) taking $a = b = 1$ in (1.6) gives $D(1) = D(1) + D(1)$. In general this implication fails: if either $e = 1$ and $D(a) = -a$ for each a, or $e = -1$ and $D(a) = a$ for each a, then (1.6) holds but (1.5) does not, unless the algebra is trivial. On the other hand, if the map $\varphi : A \to A$ defined as in Example 1.3 preserves 1, then $eD(1) = 0$, and so the invertibility of e implies (1.5).

(b) One could try to replace (1.6) with more general

$$D(ab) = cD(a)b + daD(b) + eD(a)D(b), \tag{1.7}$$

for fixed c, d (and e) in R, which would be satisfied, together with (1.5), in Example 1.4. However, together with (1.5) it would give $D(a) = D(a1) = cD(a)$ and, similarly, $D(b) = D(1b) = dD(b)$. Therefore, together with (1.5) it implies (1.6).

Example 1.6. When A and B are differential (R, e)-algebras, so is their tensor product $A \otimes_R B$ with the derivative given by

$$D(a \otimes b) = D(a) \otimes b + a \otimes D(b) + eD(a) \otimes D(b), \tag{1.8}$$

and, moreover, this differential (R, e)-algebra is the coproduct of A and B with coproduct injections $a \mapsto a \otimes 1$ and $b \mapsto 1 \otimes b$, as for the (ordinary) algebras. Since differential (R, e)-algebras form a variety of universal algebras, this and what we say in Example 1.4 implies that the forgetful functor from the category of differential (R, e)-algebras to the category of R-algebras preserves all colimits.

2 Wronski and Casorati determinants

Following the classical definition in differential algebra, we define the Wronski determinant (=Wronskian) $W(a_1,\ldots,a_n)$ of a finite sequence $a_1,\ldots,a_n$ of elements in a differential (R,e)-algebra A as the determinant

$$W(a_1,\ldots,a_n)=\begin{vmatrix} D^0(a_1) & \ldots & D^0(a_n)\\ \vdots & & \vdots\\ D^{n-1}(a_1) & \ldots & D^{n-1}(a_n)\end{vmatrix}, \tag{2.1}$$

where $D^0(a_i)=a_i$ and $D^{j+1}(a_i)=D(D^j(a_i))$ for all $i=1,\ldots,n$ and $j=0,\ldots,n-2$. On the other hand, following the classical definition of difference algebra, we define the Casorati determinant $C(a_1,\ldots,a_n)$ as

$$C(a_1,\ldots,a_n)=\begin{vmatrix} \varphi^0(a_1) & \ldots & \varphi^0(a_n)\\ \vdots & & \vdots\\ \varphi^{n-1}(a_1) & \ldots & \varphi^{n-1}(a_n)\end{vmatrix}, \tag{2.2}$$

where φ is defined as in Example 1.3 (even if e is not necessarily invertible). Using the obvious presentation

$$C(a_1,\ldots,a_n)=(-1)^{n-1}\varphi^{n-1}(a_1)C(a_2,\ldots,a_n)+\ldots+\varphi^{n-1}(a_n)C(a_1,\ldots,a_{n-1}), \tag{2.3}$$

the similar presentation of the Wronski determinant, other standard properties of determinants, and the induction in n, it is easy to show that

$$C(a_1,\ldots,a_n)=e^{n(n-1)/2}W(a_1,\ldots,a_n). \tag{2.4}$$

For a differential (R,e)-algebra A, the R-subalgebra

$$I(A)=\{a\in A|D(a)=0\} \tag{2.5}$$

of A will be called the R-subalgebra of constants of A, and its elements will simply be called constants (as classically); as motivated by Galois theory, we also think of $I(A)$ as the *index R-algebra* of A.

Lemma 2.1. *(cf. Lemma 3.1 in [8]) Let $a_1,\ldots,a_n$ be a finite sequence of elements in a differential (R,e)-algebra A. Then:*

(a) if $W(a_1,\ldots,a_n)$ is invertible in A, then $(a_1,\ldots,a_n)$ is linearly independent over $I(A)$;

(b) if e is either 0 or invertible, $W(a_1,\ldots,a_n)$ is invertible in A, and a is an element in A with $W(a,a_1,\ldots,a_n)=0$, then a is a linear combination of $a_1,\ldots,a_n$ over $I(A)$;

(c) if A is a field, then $(a_1,\ldots,a_n)$ is linearly dependent over $I(A)$ if and only if $W(a_1,\ldots,a_n)=0$.

Proof. (a) can be proved in the same way as for differential fields.

(b) When $e = 0$, consider the system

$$\begin{array}{rcl} D^0(a_1)x_1 + \ldots + D^0(a_n)x_n & = & D^0(a), \\ & \ldots & \\ D^{n-1}(a_1)x_1 + \ldots + D^{n-1}(a_n)x_n & = & D^{n-1}(a) \end{array} \tag{2.6}$$

of linear equations over A, and let us write

$$x_i = \Delta_i \Delta^{-1} \quad (i = 1, \ldots, n) \tag{2.7}$$

for its unique solution system given by Cramer's Rule; in particular, $\Delta = W(a_1, \ldots, a_n)$. Applying the derivative to (2.6) we obtain the system

$$\begin{array}{r} D^1(a_1)x_1 + \ldots + D^1(a_n)x_n + D^0(a_1)D^1(x_1) + \ldots + D^0(a_n)D^1(x_n) = D^1(a), \\ \ldots \\ D^n(a_1)x_1 + \ldots + D^n(a_n)x_n + D^{n-1}(a_1)D^1(x_1) + \ldots + D^{n-1}(a_n)D^1(x_n) = D^n(a), \end{array} \tag{2.8}$$

whose first $n - 1$ equations, together with (2.6), give

$$\begin{array}{rcl} D^0(a_1)D^1(x_1) + \ldots + D^0(a_n)D^1(x_n) & = & 0, \\ & \ldots & \\ D^{n-2}(a_1)D^1(x_1) + \ldots + D^{n-2}(a_n)D^1(x_n) & = & 0. \end{array} \tag{2.9}$$

In addition to that we have

$$D^n(a_1)x_1 + \ldots + D^n(a_n)x_n - D^n(a) = \Delta^{-1}(D^n(a_1)\Delta_1 + \cdots + D^n(a_n)\Delta_n - D^n(a)\Delta), \tag{2.10}$$

while $W(a, a_1, \ldots, a_n) = 0$ implies

$$D^n(a)\Delta'_0 + D^n(a_1)\Delta'_1 + \ldots + D^n(a_n)\Delta'_n = 0, \tag{2.11}$$

where $\Delta'_i (i = 0, \ldots, n)$ are the suitable cofactors. It is then easy to check that the determinants Δ_i and Δ'_i coincide up to signs for each $i = 0, \ldots, n$, where Δ_0 is Δ. More precisely, those signs agree in such a way that the left-hand side of (2.11) coincides with the expression in parentheses in the right-hand side of (2.10), making it 0. Therefore, in addition to (2.9), the last equality of (2.8) gives

$$D^{n-1}(a_1)D^1(x_1) + \ldots + D^{n-1}(a_n)D^1(x_n) = 0. \tag{2.12}$$

Since $\Delta = W(a_1, \ldots, a_n)$ is invertible, putting (2.9) and (2.12) together we conclude that $D^1(x_1) = \ldots = D^1(x_n) = 0$. That is, $x_1, \ldots, x_n$ are constants. Hence, by the first equality of (2.6), a is a linear combination of $a_1, \ldots, a_n$ over $I(A)$, as desired.

When e is invertible, the invertibility of $W(a_1, \ldots, a_n)$ implies the invertibility of $C(a_1, \ldots, a_n)$, and $W(a, a_1, \ldots, a_n) = 0$ (always) implies $C(a, a_1, \ldots, a_n) = 0$, and we can argue as follows:

Consider the matrix

$$\begin{pmatrix} \varphi^0(a) & \varphi^0(a_1) & \dots & \varphi^0(a_n) \\ \vdots & \vdots & & \vdots \\ \varphi^n(a) & \varphi^n(a_1) & \dots & \varphi^n(a_n) \end{pmatrix}, \tag{2.13}$$

whose determinant is $C(a, a_1, \dots, a_n)$. Since $C(a, a_1, \dots, a_n) = 0$, we have

$$a\Delta_0 + a_1\Delta_1 + \dots + a_n\Delta_n = 0, \tag{2.14}$$

where $\Delta_0, \dots, \Delta_n$ are the suitable cofactors. Replacing the first row in (2.13) with any other row we obtain a matrix with two equal rows, and so we also have

$$\begin{aligned} \varphi^1(a)\Delta_0 + \varphi^1(a_1)\Delta_1 + \dots + \varphi^1(a_n)\Delta_n &= 0, \\ &\dots \\ \varphi^n(a)\Delta_0 + \varphi^n(a_1)\Delta_1 + \dots + \varphi^n(a_n)\Delta_n &= 0. \end{aligned} \tag{2.15}$$

Here Δ_0 is the determinant of the matrix obtained from (2.13) by removing the first row and the first column. Since φ is an R-algebra homomorphism (see Example 1.3) this gives

$$\Delta_0 = \varphi(C(a_1, \dots, a_n)), \tag{2.16}$$

and so Δ_0 is invertible. Now compare the systems

$$\begin{aligned} \varphi^1(a_1)\Delta_1\Delta_0^{-1} + \dots + \varphi^1(a_n)\Delta_n\Delta_0^{-1} &= -\varphi^1(a), \\ &\dots \\ \varphi^n(a_1)\Delta_1\Delta_0^{-1} + \dots + \varphi^n(a_n)\Delta_n\Delta_0^{-1} &= -\varphi^n(a), \end{aligned} \tag{2.17}$$

and

$$\begin{aligned} \varphi^1(a_1)\varphi(\Delta_1\Delta_0^{-1}) + \dots + \varphi^1(a_n)\varphi(\Delta_n\Delta_0^{-1}) &= -\varphi^1(a), \\ &\dots \\ \varphi^n(a_1)\varphi(\Delta_1\Delta_0^{-1}) + \dots + \varphi^n(a_n)\varphi(\Delta_n\Delta_0^{-1}) &= -\varphi^n(a), \end{aligned} \tag{2.18}$$

of equations, where (2.17) is obtained from (2.15) while (2.18) is obtained by applying φ to (2.14) and to the first $n-1$ equations of (2.15). From these systems and the invertibility of $\varphi(C(a_1, \dots, a_n))$, we obtain $\Delta_i\Delta_0^{-1} = \varphi(\Delta_i\Delta_0^{-1})$ for each $i = 1, \dots, n$. Therefore (2.14) gives the desired expression for a, namely

$$a = (-\Delta_1\Delta_0^{-1})a_1 + \dots + (-\Delta_n\Delta_0^{-1})a_n. \tag{2.19}$$

(c) follows from known results of differential and difference algebra. Specifically: for $e = 0$ it is the same as Theorem 3.7 of [13], or, equivalently, Proposition 2.8 of [16]; for e invertible it is the same as Lemma II on page 271 of [3], or, equivalently, Lemma 8.2.1 of [15]. ∎

Remark 2.2. In the case $e = 0$, the Lemma 2.1 becomes Lemma 3.1 of [8], which was briefly called an obvious generalization of Theorem 3.7 of [13] there. However, I included the proof of 2.1(b) here in order to compare it with the case of invertible e, where the proof could be similarly (briefly) called an obvious generalization of the proof of Lemma II on page 271 of [3]. The comparison shows that these 'almost generalizations' are not so obvious, and, moreover, I don't see how to unify them and what to do when e is neither 0 nor invertible. I hope someone will try harder to do that. The same applies to the calculations in the next section and to the last part of the proof of Theorem 6.2.

3 The derivative of the Wronski determinant

Similarly to differential and difference equations, let us consider an equation

$$L(y) = a_0 D^n(y) + \ldots + a_n D^0(y) = 0 \quad (\text{with } a_0 = 1 \text{ and } n \neq 0) \tag{3.1}$$

of degree n with coefficients $a_0, \ldots, a_n$ in a differential (R, e)-algebra A (where (R, e) is as above). Let us suppose that $u_1, \ldots, u_n$ are (some of) solutions of this equation in A. According to (2.4.2) of [16], we have

$$e = 0 \Rightarrow D(W(u_1, \ldots, u_n)) = -a_1 W(u_1, \ldots, u_n), \tag{3.2}$$

which, in the case of fields, is also mentioned as Lemma 6.3 in [13]. We will need to calculate $D(W(u_1, \ldots, u_n))$ also in the case of invertible e, for we observe:

(a) For any $a \in A$, we have

$$D(a) = e^{-1}(\varphi(a) - a), \tag{3.3}$$

where $\varphi : A \to A$ is the R-algebra homomorphism defined in Example 1.3, and this implies

$$D^i(a) = e^{-i} \sum_{j=0}^{i} (-1)^j \binom{i}{j} \varphi^{i-j}(a), \tag{3.4}$$

for each natural i.

(b) The equation (3.1) rewritten in terms of φ becomes

$$b_0 \varphi^n(y) + \ldots + b_n \varphi^0(y) = 0, \tag{3.5}$$

where

$$b_{n-k} = \sum_{i=k}^{n} (-1)^{i-k} e^{-i} \binom{i}{i-k} a_{n-i}. \tag{3.6}$$

In particular,

$$b_0 = e^{-n}, b_n = \sum_{i=0}^{n} (-1)^i e^{-i} a_{n-i}. \tag{3.7}$$

and (3.5) implies

$$\varphi^n(y) = -e^n b_1 \varphi^{n-1}(y) - \ldots - e^n b_n \varphi^0(y) \text{ with } e^n b_n = \sum_{i=0}^{n} (-1)^i e^{n-i} a_{n-i}. \tag{3.8}$$

(c) Consider the Casorati determinant $\varphi(C(u_1, \ldots, u_n)) = C(\varphi(u_1), \ldots, \varphi(u_n))$. Since $u_1, \ldots, u_n$ are solutions of the equation (3.1), we can replace each $\varphi^n(u_i)$ $(i = 1, \ldots, n)$ in the matrix that defines the determinant $C(\varphi(u_1), \ldots, \varphi(u_n))$ with a linear combination of $\varphi^j(u_i)(j = 0, \ldots, n-1)$ according to (3.8). After that comparing the last row with other rows we can remove all multiples of $\varphi^j(u_i)$ with $j \neq 0$, and, according to the second equality of (3.8), the last row becomes

$$\left(\left(\sum_{i=0}^{n} (-1)^{i+1} e^{n-i} a_{n-i} \right) u_1, \ldots, \left(\sum_{i=0}^{n} (-1)^{i+1} e^{n-i} a_{n-i} \right) u_n \right). \tag{3.9}$$

Next, moving this last row up to make it the first row, we shall change the sign of our determinant $n-1$ times, after which it becomes nothing but the Casorati determinant $C(u_1,\ldots,u_n)$ multiplied by $\sum_{i=0}^{n}(-1)^{n+i}e^{n-i}a_{n-i}$. That is,

$$\varphi(C(u_1,\ldots,u_n)) = C(\varphi(u_1),\ldots,\varphi(u_n)) = \left(\sum_{i=0}^{n}(-1)^{n+i}e^{n-i}a_{n-i}\right)C(u_1,\ldots,u_n)$$
$$= (1-ea_1+e^2a_2-\ldots+(-1)^n e^n a_n)C(u_1,\ldots,u_n). \quad (3.10)$$

Now, using (3.3), (2.4), the fact $\varphi : A \to A$ is an R-algebra homomorphism, and (3.10), we calculate:

$$\begin{aligned}
&D(W(u_1,\ldots,u_n)) = e^{-1}(\varphi(W(u_1,\ldots,u_n)) - W(u_1,\ldots,u_n))\\
&= e^{-1}(\varphi(e^{-n(n-1)/2}C(u_1,\ldots,u_n)) - W(u_1,\ldots,u_n))\\
&= e^{-1}(e^{-n(n-1)/2}C(\varphi(u_1),\ldots,\varphi(u_n)) - W(u_1,\ldots,u_n))\\
&= e^{-1}\left(e^{-n(n-1)/2}\left(\sum_{i=0}^{n}(-1)^{n+i}e^{n-i}a_{n-i}\right)C(u_1,\ldots,u_n) - W(u_1,\ldots,u_n)\right)\\
&= e^{-1}\left(\left(\sum_{i=0}^{n}(-1)^{n+i}e^{n-i}a_{n-i}\right)W(u_1,\ldots,u_n) - W(u_1,\ldots,u_n)\right)\\
&= \left(\sum_{i=0}^{n-1}(-1)^{n+i}e^{n-i-1}a_{n-i}\right)W(u_1,\ldots,u_n).
\end{aligned}$$

That is, we obtain:

Proposition 3.1. *When e is invertible, for any n solutions $u_1,\ldots,u_n$ of the equation (3.1), we have:*

$$D(W(u_1,\ldots,u_n)) = \left(\sum_{i=1}^{n}(-1)^{i}e^{i-1}a_{i}\right)W(u_1,\ldots,u_n)$$
$$= (-a_1+ea_2-e^2a_3+\ldots+(-1)^n e^{n-1}a_n)W(u_1,\ldots,u_n). \quad (3.11)$$

4 Admissibility Lemma

For a fixed pair (R,e) as in Definition 1.1, consider the adjunction

$$(I,H,\eta,\varepsilon) : \mathbf{C} \to \mathbf{X}, \quad (4.1)$$

in which:

- $\mathbf{C}$ is the opposite (=dual) category of differential (R,e)-algebras;
- $\mathbf{X}$ is the opposite category of R-algebras;
- $I : \mathbf{C} \to \mathbf{X}$ is the functor defined by (2.5);

- $H : \mathbf{X} \to \mathbf{C}$ is the functor defined by

$$H(X) = X \text{ equipped with the zero derivative} \tag{4.2}$$

(see Example 1.4), and we shall simply write $H(X) = X$;

- for each object A in $\mathbf{C}$, $\eta_A : A \to HI(A)$ is the inclusion map $I(A) \to A$;

- for each object X in $\mathbf{X}$, $\varepsilon_X : IH(X) \to X$ is the identity map of X.

Since the category $\mathbf{C}$ has pullbacks, for every object A in $\mathbf{C}$, we have the adjunction

$$(I^A, H^A, \eta^A, \varepsilon^A) : (\mathbf{C} \downarrow A) \to (\mathbf{X} \downarrow I(A)), \tag{4.3}$$

induced by the adjunction (4.1). This adjunction is involved in the Galois theory of the adjunction (4.1) (see [11] and references therein), and, in particular, A is I-admissible in the sense of [6] when

$$\varepsilon^A : I^A H^A \to 1_{(\mathbf{X}\downarrow I(A))} \tag{4.4}$$

is an isomorphism. This definition of admissibility, formulated for a general adjunction between the categories with pullbacks, in the present case gives:

Observation 4.1. A is I-admissible if and only if for every $I(A)$-algebra X the map

$$X \to I(A \otimes_{I(A)} X) \tag{4.5}$$

defined by $x \mapsto 1 \otimes x$ is bijective.

This observation copies similar observation in [8], and, moreover, the proof of Lemma 4.2 below copies its special case from [8].

Lemma 4.2. *If $I(A)$ is a field, then A is I-admissible.*

Proof. Since $I(A)$ is a field, the canonical map $I(A) \otimes_{I(A)} X \to A \otimes_{I(A)} X$ is injective, and therefore so is the map (4.5). To prove the surjectivity we will use any basis B of X considered as a vector space over $I(A)$, and the fact that it makes $A \otimes_{I(A)} X$ a free A-module over the family $(1 \otimes b)_{b\in B}$. Every element c of $A \otimes_{I(A)} X$ can be written as $c = \sum_{i=1}^{n} a_i \otimes b_i$ with $a_i \in A$ and $b_i \in B$, and when c is in $I(A \otimes_{I(A)} X)$ we have

$$0 = D(c) = \sum_{i=1}^{n} D(a_i) \otimes b_i + \sum_{i=1}^{n} a_i \otimes D(b_i) + \sum_{i=1}^{n} eD(a_i) \otimes D(b_i) = \sum_{i=1}^{n} D(a_i) \otimes b_i.$$

Since $A \otimes_{I(A)} X$ a free A-module over the family $(1 \otimes b)_{b\in B}$ this implies that each a_i is a constant, and so $c = \sum_{i=1}^{n} 1 \otimes a_i b_i = 1 \otimes \sum_{i=1}^{n} a_i b_i$ is the image of $\sum_{i=1}^{n} a_i b_i$ under the map (4.5). ■

5 Injectivity Lemma

Returning to the induced adjunction (4.3), consider the natural transformation

$$\eta^A : 1_{(\mathbf{C}\downarrow A)} \to H^A I^A. \tag{5.1}$$

For an object (A', α) in $(\mathbf{C} \downarrow A)$, the (A', α)-component $(\eta^A)_{(A',\alpha)}$ of η^A is the map

$$A \otimes_{I(A)} I(A') \to A' \tag{5.2}$$

defined by $a \otimes a' \mapsto \alpha(a)a'$.

Lemma 5.1. *If A is a field, then each component of η^A is an injective map.*

Proof. Since A is a field, so is $I(A)$, which easily follows from (1.5) and (1.6). Therefore $A \otimes_{I(A)} I(A')$ is a free module over $I(A')$; moreover, if B is a basis of A considered as a vector space over $I(A)$, then the family $(b \otimes 1)_{b\in B}$ is a basis of that free module. It follows that in order to prove injectivity of the map (5.2) it suffices to prove that if $(a_1, \ldots, a_n)$ is linearly independent in A over $I(A)$, then $(\alpha(a_1), \ldots, \alpha(a_n))$ is linearly independent in A' over $I(A')$. However, excluding the case of empty B, that follows from (c) and (a) of Lemma 2.1, since $\alpha(W(a_1, \ldots, a_n)) = W(\alpha(a_1), \ldots, \alpha(a_n))$ and α must be injective. ∎

6 Picard-Vessiot extensions

The following definition copies its classical differential- and difference-algebraic special cases:

Definition 6.1. Let K and E be differential (R, e)-algebras that are fields with $K \subseteq E$. We say that E is a Picard-Vessiot extension of K for an equation (3.1) (of degree n) with coefficients in K, if $I(E) = I(K)$ and there exist solutions $u_1, \ldots, u_n \in E$ of that equation, which are linearly independent over K and generate E in the following sense: there is no proper differential (R, e)-subalgebra of E that is a field containing K and $u_1, \ldots, u_n$.

Theorem 6.2. *Let E be a Picard-Vessiot extension of K for an equation (3.1) with $u_1, \ldots, u_n$ as in Definition 6.1, $w = W(u_1, \ldots, u_n)$, and let A be the differential (R, e)-subalgebra of E generated by $K \bigcup \{u_1, \ldots, u_n, w^{-1}\}$. If e is either 0 or invertible in R, then A is a normal extension of K in the sense of categorical Galois theory (see Theorem 0.2), that is, the following conditions hold:*

- *(a) A is I-admissible;*
- *(b) the change-of-base functor $(\mathbf{C} \downarrow K) \to (\mathbf{C} \downarrow A)$ is monadic;*
- *(c) $(A \times_K A, \mathrm{pr}_1) \approx H^A I^A (A \times_K A, \mathrm{pr}_1)$ canonically in $(\mathbf{C} \downarrow A)$ (or, equivalently under (a), $(A \times_K A, \mathrm{pr}_1)$ belongs to the replete image of H^A).*

Proof. (a) follows from Lemma 4.2, and (b) follows from the corresponding result for rings, which itself follows from the fact that K is a field (see e.g. [12] for details, although this is a reformulation of a very old observation of A. Grothendieck).

(c): The pullback $A \times_K A$ in $\mathbf{C}$ is nothing the tensor product $A \otimes_K A$, and the canonical morphism

$$(\eta^A)_{(A\times_K A,\mathrm{pr}_1)} : (A \times_K A, \mathrm{pr}_1) \to H^A I^A (A \times_K A, \mathrm{pr}_1) \tag{6.1}$$

in $(\mathbf{C} \downarrow A)$ is the map

$$A \otimes_{I(A)} I(A \otimes_K A) \to A \otimes_K A \tag{6.2}$$

defined by $a \otimes \left(\sum_{i=1}^{k} b_i \otimes c_i\right) \mapsto \sum_{i=1}^{k} ab_i \otimes c_i$. We have to prove that this map, which we will simply denote by f, is injective and surjective.

Injectivity: Consider the commutative diagram

$$\begin{array}{ccc} A \otimes_{I(A)} I(A \otimes_K A) & \to & A \otimes_K A \\ \downarrow & & \downarrow \\ E \otimes_{I(E)} I(E \otimes_K A) & \to & E \otimes_K A, \end{array} \tag{6.3}$$

in which the top arrow is f, the bottom arrow is defined similarly, and the vertical arrows are maps induced by the inclusion map $A \to E$. Since $I(A) = I(E)$ and K are fields, the vertical arrows are injective, while the injectivity of the bottom arrow follows Lemma 5.1. It follows that f is also injective.

Surjectivity: It suffices to show that, for every $a \in A$, the image of f contains the element $1 \otimes a$ of $A \otimes_K A$. Moreover, since the set of all $a \in A$ for which $1 \otimes a$ is in the image of f is (obviously) a differential (R, e)-subalgebra of A, it suffices to do that for all $a \in \{u_1, \ldots, u_n, w^{-1}\}$.

Since $u_1, \ldots, u_n$ are solutions of the equation (3.1), we can write

$$a_1 D^{n-1}(u_i) + \ldots + a_n D^0(u_i) = -D^n(u_i) \quad (i = 1, \ldots, n). \tag{6.4}$$

In particular, this gives us, for each fixed $i = 1, \ldots, n$, the system

$$\begin{array}{rcl} a_1 D^{n-1}(1 \otimes u_i) + \ldots + a_n D^0(1 \otimes u_i) & = & -D^n(1 \otimes u_i), \\ a_1 D^{n-1}(u_1 \otimes 1) + \ldots + a_n D^0(u_1 \otimes 1) & = & -D^n(u_1 \otimes 1), \\ & \cdots & \\ a_1 D^{n-1}(u_n \otimes 1) + \ldots + a_n D^0(u_n \otimes 1) & = & -D^n(u_n \otimes 1) \end{array} \tag{6.5}$$

of equations in $A \otimes_K A$, where a_j is used for $1 \otimes a_j = a_j \otimes 1$. As follows from these equations, the last row of the matrix used to define the Wronski determinant $W(1 \otimes u_i, u_1 \otimes 1, \ldots, u_n \otimes 1)$ is a linear combination of previous rows. Therefore $W(1 \otimes u_i, u_1 \otimes 1, \ldots, u_n \otimes 1) = 0$. On the other hand, $W(u_1 \otimes 1, \ldots, u_n \otimes 1)$ is invertible in $A \otimes_K A$. It follows now from Lemma 2.1(b) that $1 \otimes u_i$ is a linear combination of $u_1 \otimes 1, \ldots, u_n \otimes 1$ over $I(A \otimes_K A)$. Therefore $1 \otimes u_i$ is in the image of f.

It remains to prove that $1 \otimes w^{-1}$ is in the image of f. Since $1 \otimes w^{-1} = (w \otimes w^{-1})(w^{-1} \otimes 1)$, it suffices to prove that $D(w \otimes w^{-1}) = 0$. When $e = 0$ this is done [8] as follows:

$$D(w \otimes w^{-1}) = D(w) \otimes w^{-1} + w \otimes D(w^{-1}) = \\ (-a_1 w) \otimes w^{-1} + w \otimes (-D(w)w^{-2}) = 0,$$

where the second equality follows from (3.2) and from $D(w)w^{-1} + wD(w^{-1}) = 0$, while the third one follows from $-D(w)w^{-2} = a_1 w^{-1}$, which is also a consequence of (3.2). If e is invertible, we will write c for the Casorati determinant $C(u_1, \ldots, u_n)$ and we use (3.10) in the form $\varphi(c) = ac$, where $a = 1 - ea_1 + e^2 a_2 - \ldots + (-1)^n e^n a_n \in K$. In this case we have

$$\begin{aligned} & D(w \otimes w^{-1}) = D(c \otimes c^{-1}) \quad \text{(by (2.4))} \\ & = e^{-1}(\varphi(c \otimes c^{-1}) - c \otimes c^{-1}) \quad \text{(by definition of } \varphi \text{ in Example 1.3)} \\ & = e^{-1}(\varphi(c) \otimes \varphi(c)^{-1} - c \otimes c^{-1}) \quad \text{(since } \varphi \text{ is an } R\text{-algebra homomorphism)} \\ & = e^{-1}(ac \otimes (ac)^{-1} - c \otimes c^{-1}) \quad \text{(since } \varphi(c) = ac) \\ & = 0 \quad \text{(since } a \text{ is in } K), \end{aligned}$$

as desired. ∎

Remark 6.3. Since K is a field and R can be replaced with its canonical image in K, our condition on e in Theorem 6.2 is in fact irrelevant.

Corollary 6.4. *As follows from Theorem 6.2, the equivalence*

$$Spl_I(A, f) \sim \mathbf{X}^{I(A \times_B A)},$$

of Theorem 0.2 applies to differential (R, e)-algebras as follows. We take:

- *I to be the functor from the opposite category of differential (R, e)-algebras to the opposite category of R-algebras sending differential (R, e)-algebras to their R-algebras of constants. Accordingly $\mathbf{X}$ becomes the opposite category of R-algebras.*
- *$f : A \to B$ to be the inclusion map $K \to A = K \cup \{u_1, \ldots, u_n, w^{-1}\}$, in the notation of Theorem 6.2, considered as a morphism $A \to K$ in the opposite category of differential (R, e)-algebras.*

Then:

- *The internal Galois groupoid $I(A \times_B A)$ in $\mathbf{X}$ has the R-algebra of constants of A as its object-of-objects and the R-algebra of constants of $A \otimes_K A$ as its object-of-morphisms. In particular, if A and K have the same constants, then $I(A \times_B A)$ is the Hopf algebra of constants of $A \otimes_K A$.*
- *$Spl_I(A, f)$ is the opposite category of all differential (R, e)-algebras whose tensor product with A is obtained by freely adding constants to A, that is, obtained as $A \otimes_{I(A)} X$, where X is an $I(A)$-algebra and both $I(A)$ and X are considered as differential (R, e)-algebras with the zero derivative.*

References

[1] F. Borceux and G. Janelidze, *Galois Theories*, Cambridge Studies in Advanced Mathematics **72** (Cambridge University Press, 2001).

[2] Z. Chatzidakis, C. Hardouin, and M. Singer, *On the definitions of difference Galois groups*, Model theory with applications to algebra and analysis 1, London Mathematical Society Lecture Note Series **349**, Cambridge University Press (2008) pp. 73–109.

[3] R. M. Cohn, *Difference algebra*, Interscience Tracts in Pure and Applied Mathematics **17** (Interscience Publishers, 1965).

[4] L. Guo and W. Keigher, *On differential Rota-Baxter algebras*, Journal of Pure and Applied Algebra **212** (2008) pp. 522–540.

[5] L. Guo, W. Keigher and S. Zhang, *Monads and distributive laws for Rota-Baxter and differential algebras*, Advances in Applied Mathematics **72** (2016), pp. 139–165.

[6] G. Janelidze, *Magid's theorem in categories*, Bulletin of the Georgian Academy of Sciences **114 (3)** (1984) pp. 497–500 (in Russian).

[7] G. Janelidze, *The fundamental theorem of Galois theory*, Math. USSR Sbornik **64 (2)** (1989) pp. 359–374.

[8] G. Janelidze, *Galois theory in categories: the new example of differential fields*, Proc. Conf. Categorical Topology in Prague 1988, World Scientific (1989) pp. 369–380.

[9] G. Janelidze, *Pure Galois theory in categories*, Journal of Algebra **132** (1990), pp. 270–286.

[10] G. Janelidze, *Precategories and Galois theory*, Lecture Notes in Mathematics **1488** (1991) pp. 157–173.

[11] G. Janelidze, *Categorical Galois theory: revision and some recent developments*, Galois Connections and Applications, Kluwer Academic Publishers B.V. (2004) pp. 139–171.

[12] G. Janelidze and W. Tholen, *Facets of Descent III: Monadic descent for rings and algebras*, Applied Categorical Structures **12 (5-6)** (2004) pp. 461–477.

[13] I. Kaplansky, *An introduction to differential algebra*, 2nd Ed., Paris, Hermann, 1976.

[14] E. R. Kolchin, *Differential algebra and algebraic groups*, Academic Press, New York and London, 1973.

[15] A. Levin, *Difference Algebra*, Algebra and Applications **8**, Springer, 2008.

[16] A. R. Magid, *Lectures on differential Galois theory*, University Lecture Series **7**, 1994.

[17] A. R. Magid, Review on “Selected works of Ellis Kolchin” by H. Bass, A. Buium, P. J. Cassidy (eds), Bulletin of the American Mathematical Society **37 (3)** (2000) pp. 337–342.

[18] A. R. Magid, *The separable Galois theory of commutative rings*, 2nd Ed., Pure and Applied Mathematics (Boca Raton), CRC Press, Boca Raton, FL, 2014.

[19] J.-F. Pommaret, *Differential Galois theory*, Mathematics and its Applications **15**, Gordon & Breach Science Publishers, New York, 1983.

[20] M. Takeuchi, *A Hopf algebraic approach to the Picard-Vessiot theory*, Journal of Algebra **122** (1989) pp. 481–509.

[21] M. van der Put and M. F. Singer, *Galois theory of difference equations*, Lecture Notes in Mathematics **1666** (Springer, Berlin, 1997).

Department of Mathematics and Applied Mathematics,
University of Cape Town,
Rondebosch 7701 (South Africa)
email : George.Janelidze@uct.ac.za