Polynômes de Ore en une variable

Xavier Caruso

28 décembre 2017

Résumé

Les polynômes de Ore sont une variante non commutative des polynômes classiques qui interviennent en algèbre semi-linéaire et dans l'étude des équations différentielles linéaires de la même manière que les polynômes usuels interviennent en algèbre linéaire (polynômes d'endomorphisme, polynômes caractéristiques, *etc*.) En particulier, la factorisation des polynômes de Ore est liée de très près à la réduction des endomorphismes semi-linéaires et à celle des équations différentielles linéaires.

Le but de ce cours est définir les polynômes de Ore, d'établir leurs principales propriétés arithmétiques et d'étudier leur factorisation. Nous mettons en place tout un arnesal théorique, centrée sur la notion d'algèbre d'Azumaya, qui permet, dans certains cas, d'obtenir des théorèmes de structure donnant des renseignements très précis sur la forme des factorisations des polynômes de Ore. Le cours est illustré de nombreux exemples.

Table des matières

Introduction

Il est difficile de nier que les anneaux de polynômes sont l'une des briques fondamentales — si ce n'est la brique fondamentale — des constructions en algèbre commutative ; de fait, de par sa nature, l'algèbre commutative étudie les deux opérations $+$ et $\times$ qui, lorsqu'on les met ensemble, conduit naturellement à la notion de polynômes. Les polynômes jouent également un rôle central en algèbre linéaire. Le polynôme caractéristique, typiquement, est un outil inévitable pour étudier la réduction des endomorphismes linéaires. Les propriétés de factorisation des polynômes d'endomorphisme — et notamment du polynôme caractéristique — nous renseignent sur le type de réduction auquel il faut s'attendre. Par exemple, il est bien connu qu'un endomorphisme est diagonalisable si et seulement s'il est annulé par un polynôme scindé séparable.

Au début des années 1930, le mathématicien norvégien Øystein Ore [22] a imaginé une variante « tordue » des polynômes dans laquelle la multiplication est rendue non commutative et suit la loi :

$$X \cdot a = \theta(a)X + \partial(a)$$

lorsque X désigne la variable et a est un scalaire. Dans l'écrture précédente, θ et ∂ désignent deux fonctions qui doivent être soumises à certaines contraintes de compatibilité. Étonnemment, il se trouve que les polynômes de Ore jouïssent de propriétés similaires aux polynômes usuels. Par exemple, lorsque l'on travaille sur un anneau de base qui est un corps, la division euclidienne des polynômes de Ore est bien définie et les propriétés qui en résultent classiquement (algorithme d'Euclide, existence de PGCD et de PPCM, *etc.*) s'étendent également. D'autre part, les polynômes de Ore présentent un intérêt, d'une part, car ils jouent un rôle dans certaines questions internes de l'algèbre non commutative mais également, d'autre part, car ils sont liés à l'étude des endomorphismes semi-linéaires et des équations différentielles linéaires de la même manière que les polynômes usuels étaient liés à la réduction des endomorphismes linéaires. Les polynômes de Ore s'immiscent de cette manière dans d'autres domaines centraux des mathématiques.

Les propriétés de factorisation des polynômes de Ore sont ainsi particulièrement intéressantes car elles sont liées à la réduction des endomorphismes semi-linéaires et à la décomposition des équations différentielles (ou, pour être plus précis, à celle des modules différentiels ou modules à connexions). Comprendre la mécanique de la factorisation dans les anneaux de polynômes de Ore devient, de cette manière, un enjeu important et c'est sur cette question que ce cours se focalise. Dans le cas commutatif, lorsque l'anneau de base est factoriel, on sait que la factorisation des polynômes en produit d'irréductibles est unique à l'ordre près. Ce résultat simple est malheureusement facilement mis en défaut dans le cas des polynômes de Ore. Par exemple, dans l'anneau des polynômes de Ore complexes où la multiplication est régie par la loi $Xz = \bar{z}X$ (pour $z \in \mathbb{C}$), on a l'identité :

$$X^2 - 1 = (X + \alpha) \cdot (X - \bar{\alpha})$$

pour tout nombre complexe α de module 1. Ore parvient malgré tout à démontrer un premier résultat encourageant de rigidité qui stipule que, dans le cas où l'anneau de base est un corps, toute factorisation d'un polynôme de Ore fixé en produits d'irréductibles comporte le même nombre de facteurs et, de plus, que ces facteurs peuvent être mis en correspondance.

Le résultat de Ore est intéressant mais il ne donne pas de description de l'ensemble de toutes les factorisations d'un polynôme de Ore donné *a priori*. Néanmoins, le cas du polynôme de Ore X^2-1 pris en exemple précédemment laisse espérer qu'une telle description est possible ;

en effet, dans ce cas précis, les factorisations de X^2-1 (en produit d'irréductibles unitaires) sont paramétrées par un objet géométrique tout à fait respectable, à savoir le cercle unité complexe. L'un des objectifs principaux de ce cours est de concrétiser cet espoir. Pour ce faire, nous établissons un lien fort entre, d'une part, les anneaux de Ore et, d'autre part, un autre objet incontournable de la théorie des anneaux non commutatifs : les algèbres de matrices. Ce lien se formalise *via* les notions d'algèbres simples centrales et d'algèbres d'Azumaya qui seront nos compagnons de route tout au long de ce cours. Suivant ce chemin, nous parvenons à importer les outils standard d'algèbre linéaire (comme le déterminant) dans le monde des anneaux de Ore, ce qui nous permet *in fine* de démontrer le résultat suivant énoncé ici de manière volontairement vague.

Théorème 1. *On suppose que l'anneau de base sur lesquels les polynômes de Ore sont définis est un corps. Alors, sous certaines hypothèses de « finitude », l'ensemble des diviseurs (à droite) irréductibles unitaires d'un polynôme de Ore donné est naturellement paramétré par un produit d'espaces projectifs (sur des corps non nécessairement commutatifs).*

Pour un énoncé plus précis, nous renvoyons notre lectrice aux résultats du §4.4 et, en particulier, au théorème 4.4.3. Pour l'exemple du polynôme X^2-1 considéré précédemment, le produit d'espaces projectifs promis par le théorème 1 se réduit à la droite projective réelle $\mathbb{P}^1(\mathbb{R})$ (qui est bien homéomorphe au cercle unité complexe). Soulignons également qu'il est possible d'obtenir à partir du théorème 1 une description géométrique de l'ensemble des factorisations d'un polynôme de Ore donné en produit de facteurs irréductibles unitaires ; celle-ci se fait en termes de variétés de drapeaux complets (dans des espaces vectoriels définis sur des corps non nécessairement commutatifs).

Le plan de ce cours est le suivant. Dans le §1, nous définissons les polynômes de Ore et établissons leurs propriétés arithmétiques fondamentales (division euclidienne, PGCD, PPCM, *etc*). Nous étudions également les modules sur les anneaux de Ore, ce qui nous permet de concrétiser le lien entre factorisation des polynômes de Ore et décomposition des endomorphismes semi-linéaires ou des modules à connexions. Le point culminant de cette partie est la démonstration du théorème de factorisation de Ore qui a été mentionné précédemment.

Dans le §2, nous introduisons et étudions la notion de résultant dans le cadre non-commutatif des polynômes de Ore. Nous montrons que celle-ci présente de très fortes analogies avec le cadre commutatif et, en particulier, qu'elle permet de détecter la relative primalité de polynômes de Ore et même, dans une version ramifiée, d'exprimer les coefficients des PGCD de polynômes de Ore.

Le §3 est une introduction accélérée à la théorie des algèbres simples centrales et des algèbres d'Azumaya. En aucun cas, nous ne prétendons à l'exhaustivité ; au contraire, nous nous bornons à exposer les aspects de la théorie qui nous sont utiles pour les applications que nous avons en vue. Toutefois, nous avons pris soin de n'omettre aucune démonstration, en espérant que cela puisse être agréable à notre lecteur.

Enfin, le §4, qui est le cœur de ce cours, met en application la théorie des algèbres d'Azumaya dans le cadre des polynômes de Ore. Comme expliqué précédemment, cela nous permet de ramener les questions de factorisation qui nous intéressent à des problèmes très concrets d'algèbre linéaire. Le théorème 1 découle de ces considérations.

Le cours est complété de deux courts appendices. Le premier fait le point sur la théorie des espaces vectoriels sur les corps non commutatifs (qui se trouve être très semblable à celle des espaces vectoriels sur les corps commutatifs) tandis que le second rassemble plusieurs résultats (que nous aurons besoin d'utiliser à l'occasion) sur les extensions de corps différentiels en caractéristique positive.

Tout le cours est illustré par de nombreux exemples qui, nous l'espérons, sont suffisamment variés pour donner une intuition approfondie des phénomènes étudiés.

1 L'arithmétique des polynômes de Ore

Dans cette partie, nous introduisons les anneaux de polynômes de Ore et, principalement lorsque la base est un corps, nous démontrons les premières propriétés arithmétiques de ces polynômes qui, comme dans le cas des polynômes usuels, découlent de l'existence d'une division euclidienne (à droite). Nous énonçons et démontrons également le théorème de factorisation de Ore qui est un résultat de rigidité sur la factorisation des polynômes de Ore. La démonstration de ce théorème sera, pour nous, l'occasion de faire un détour par les modules sur les anneaux de polynômes de Ore. Loin d'être anecdotique, ce détour nous permettra également de préciser les liens entre les polynômes de Ore et l'algèbre semi-linéaire et/ou différentielle.

Dans toute cette partie, la lettre $\mathfrak{A}$ désigne un anneau *commutatif* sur lequel on ne fera généralement aucune hypothèse supplémentaire. La lettre K, quant à elle, sera utilisée pour désigner un corps (commutatif).

1.1 Définition des anneaux de Ore

L'anneau des polynômes de Ore, introduit dans [22], est une variante non commutative de l'anneau des polynômes usuels $\mathfrak{A}[X]$; il est défini en remplaçant la multiplication classique par une multiplication « tordue » pour laquelle la variable X ne commute par avec les scalaires mais, au contraire, est régie par la loi :

$$\forall a \in \mathfrak{A}, \quad X \times a = \theta(a)X + \partial(a) \tag{1}$$

où $\theta : \mathfrak{A} \to \mathfrak{A}$ et $\partial : K \to K$ sont des fonctions sur lesquelles on va être amené à imposer des contraintes reflétant les axiomes d'anneaux.

Plus précisément, on se donne un morphisme d'anneaux $\theta : \mathfrak{A} \to \mathfrak{A}$ et une application $\partial : \mathfrak{A} \to \mathfrak{A}$ vérifiant l'axiome :

$$\forall a, b \in \mathfrak{A}, \quad \partial(a+b) = \partial(a) + \partial(b) \tag{2}$$

$$\text{et} \quad \partial(ab) = \theta(a)\partial(b) + \partial(a)b \tag{3}$$

Lorsque $\theta = \mathrm{id}$, on remarque que l'axiome (3) ci-dessus n'est autre que la loi de Leibniz ; autrement dit, on demande à ∂ d'être une dérivation. Dans le cas général, une application ∂ vérifiant les axiomes (2) et (3) est appelé une *θ-dérivation*. Notons que l'ensemble des θ-dérivations est stable par addition et par multiplication externe par les éléments de $\mathfrak{A}$; autrement dit, c'est un $\mathfrak{A}$-module.

Les applications θ et ∂ étant fixées, nous allons définir une $\mathfrak{A}$-algèbre (non commutative) $\mathfrak{A}[X, \theta, \partial]$. En tant que groupe additif, $\mathfrak{A}[X, \theta, \partial]$ n'est autre que $\mathfrak{A}[X]$. Autrement dit, $\mathfrak{A}[X, \theta, \partial]$ est l'ensemble des expressions formelles :

$$a_0 + a_1 X + a_2 X^2 + \cdots + a_{d-1} X^{d-1} + a_d X^d$$

où d est un entier variable et les a_i sont des éléments de $\mathfrak{A}$. De plus, ces expressions s'additionnent de la manière usuelle :

$$\left(\sum_i a_i X^i\right) + \left(\sum_i b_i X^i\right) = \sum_i (a_i + b_i) X^i.$$

La loi de multiplication, quant à elle, est déterminée par la loi de Ore (1). Plus précisément, on a le résultat suivant.

Proposition 1.1.1. *Il existe une unique loi de multiplication $\times$ sur $\mathfrak{A}[X, \theta, \partial]$ qui vérifie les relations :*

$$\forall a \in \mathfrak{A}, \quad X \times a = \theta(a)X + \partial(a)$$

$$\textit{et} \quad a \times X = aX$$

et coïncide avec la mutiplication de $\mathfrak{A}$ sur les polynômes constants.

Démonstration. Remarquons que, des conditions de la proposition, on peut déduire la valeur du produit $X^2 \times a$ lorsque a est un scalaire. En effet, on peut écrire :

$$\begin{aligned} X^2 \times a &= X \times (X \times a) \\ &= X \times \big(\theta(a)X + \partial(a)\big) \\ &= (X \times \theta(a)) \times X + (X \times \partial(a)) \\ &= \big(\theta^2(a)X + \partial \circ \theta(a)\big) \times X + \theta \circ \partial(a)X + \partial^2(a) \\ &= \theta^2(a)X^2 + \big(\partial \circ \theta(a) + \theta \circ \partial(a)\big)X + \partial^2(a). \end{aligned}$$

De la même manière, on obtient des formules pour $X^i \times a$ pour tout entier $i \geq 0$. De plus, si $P = \sum_i a_i X^i$ et $Q = \sum_j a_j X^j$ sont deux éléments de $\mathfrak{A}[X, \theta, \partial]$, la loi de distributivité impose :

$$P \times Q = \sum_{i,j} (a_i X^i) \times (b_j X^j) = \sum_{i,j} a_i \times X^i \times b_j \times X^j.$$

L'unicité de la loi $\times$ vérifiant les conditions de la proposition en résulte. Pour ce qui est de l'existence, il suffit de vérifier les relations suivantes :

$$\begin{aligned} (a+b) \times X &= a \times X + b \times X \quad ; \quad (ab) \times X = a \times (b \times X) \\ X \times (a+b) &= X \times a + X \times b \quad ; \quad X \times (ab) = (X \times a) \times b. \end{aligned}$$

Les égalités de la première ligne s'obtiennent directement sans calcul. La première égalité de la seconde ligne résulte de l'additivité de θ et de ∂. Enfin, pour la dernière, on écrit :

$$\begin{aligned} (X \times a) \times b &= (\theta(a)X + \partial(a)) \times b \\ &= \theta(a) \times (X \times b) + \partial(a) \times b \\ &= \theta(a) \times (\theta(b)X + \partial(b)) + \partial(a)b \\ &= \theta(a)\theta(b)X + \theta(a)\partial(b) + \partial(a)b \\ &= \theta(ab)X + \partial(ab) = X \times (ab) \end{aligned}$$

l'avant-dernière égalité résultant du fait que θ est un morphisme d'anneaux et que ∂ est une θ-dérivation. □

Remarque 1.1.2. Il suit directement des définitions que l'application évidente $\mathfrak{A} \to \mathfrak{A}[X, \theta, \partial]$ est compatible et au produit. Autrement dit $\mathfrak{A}[X, \theta, \partial]$ apparaît canoniquement comme une algèbre (non commutative) sur $\mathfrak{A}$.

Dans la suite, nous omettrons systématiquement le signe $\times$ lorsque nous écrirons des multiplications dans $\mathfrak{A}[X, \theta, \partial]$; cela ne prêtera jamais à confusion mais allègera considérablement les écritures. Au niveau de la terminologie, nous utiliserons la locution *polynôme de Ore* pour faire référence aux éléments de $\mathfrak{A}[X, \theta, \partial]$.

Nous avons vu, dans la démonstration de la proposition 1.1.1 que l'on peut théoriquement obtenir des formules closes donnant les valeurs de produits $X^n a$ lorsque n est un entier et a est un élément de $\mathfrak{A}$. Généralement, ces formules sont complexes et difficiles à écrire. Il y a toutefois deux exceptions notables. Premièrement, lorsque $\partial = 0$, la relation de commutation de Ore s'écrit simplement $Xa = \theta(a)X$, à partir de quoi, on déduit par une récurrence immédiate que $X^n a = \theta^n(a)X$ pour tout $n \geq 0$. Dans ce cas, on dispose en outre d'une formule explicite très simple donnant le produit de deux polynômes de Ore, à savoir :

$$\big(\textstyle\sum_i a_i X^i\big) \cdot \big(\sum_j b_i X^j\big) = \sum_k c_k X^k \quad \text{avec} \quad c_k = \displaystyle\sum_{i+j=k} a_i \theta^i(b_j).$$

De la même manière, lorsque $\theta = \mathrm{id}_{\mathfrak{A}}$, le produit $X^n a$ admet une écriture relativement agréable. Précisément, on démontre par récurrence sur n que :

$$X^n a = \sum_{i=0}^{n} \binom{n}{i} \partial^i(a) X^{n-i}. \tag{4}$$

où la notation $\binom{n}{i}$ désigne le coefficient binomial.

Exemple 1.1.3. Se donner un exemple d'anneau de polynômes de Ore, revient à se donner un anneau $\mathfrak{A}$ muni d'applications θ et ∂. Il s'avère, en fait, que les cas les plus intéressants sont ceux pour lesquels $\theta = \mathrm{id}_{\mathfrak{A}}$ ou $\partial = 0$.

a- On prend $\mathfrak{A} = \mathbb{C}$, $\theta = \mathrm{conj}$; cet exemple, avec lequel nous travaillerons régulièrement dans la suite de ce cours, est lié à l'étude des applications semi-linéaires sur des $\mathbb{C}$-espaces vectoriels.

b- On prend pour $\mathfrak{A} = \mathbb{F}_q$ un corps fini, pour θ le morphisme de Frobenius et $\partial = 0$. De la même manière, cet exemple est lié à l'étude des applications semi-linéaires sur des corps finis qui sont des objets qui apparaissent naturellement en théorie des nombres. De plus, cet exemple est intéressant à un second titre car l'anneau de Ore qui en résulte $\mathbb{F}_q[X, \mathrm{Frob}]$ se réinterprète en termes de polynômes linéairisés comme suit. Un polynôme linéaire à coefficients dans $\mathbb{F}_q$ (où $q = p^r$ pour un certain nombre premier p et un certain entier r) est un polynôme de la forme :

$$a_0 T + a_1 T^p + a_2 T^{p^2} + \cdots + a_d T^{p^d} \in \mathbb{F}_q[T].$$

Ces polynômes sont évidemment stables par addition mais ils le sont aussi par composition. Ils forment ainsi un anneau noté $\mathbb{F}_q[T]_{\mathrm{lin}}$ (sur lequel les opérations sont donc $+$ et $\circ$). Il se trouve qu'on a un isomorphisme (vérification simple laissée au lecteur) :

$$\begin{array}{rcl} \mathbb{F}_q[X, \mathrm{Frob}] & \xrightarrow{\sim} & \mathbb{F}_q[T]_{\mathrm{lin}} \\ a_0 + a_1 X + \cdots + a_d X^d & \mapsto & a_0 T + a_1 T^p + \cdots + a_d T^{p^d}. \end{array}$$

Une grande partie des résultats que nous démontrerons dans la suite sur les polynômes de Ore s'appliquent donc *de facto* aux polynômes linéarisés.

c- On prend $\mathfrak{A} = \mathbb{C}[t]$ ou $\mathbb{C}(t)$, $\theta = \mathrm{id}_{\mathfrak{A}}$ et $\partial = \frac{d}{dt}$. Cet exemple est lié à l'étude algébrique des équations différentielles linéaires ; plus précisément, les éléments de $\mathbb{C}[t][X, \frac{d}{dt}]$ apparaissent comme des opérateurs différentiels. Dans cet exemple, on peut formellement remplacer $\mathbb{C}$ par un autre corps, éventuellement de caractéristique positive.

Exemple 1.1.4. Dans $\mathbb{C}[X, \mathrm{conj}]$, calculons le produit $(X-a)(X-b)$ pour des nombres complexes a et b. On trouve :

$$\begin{aligned} (X-a)(X-b) &= X^2 - Xb - aX + ab \\ &= X^2 - \big(a + \bar{b}\big)X + ab. \end{aligned}$$

En particulier, pour $b = -\bar{a}$, on trouve :

$$(X-a)(X+\bar{a}) = X^2 - |a|^2.$$

Si c est un nombre réel strictement positif, le polynôme $X^2 - c$ possède ainsi de nombreuses factorisations dans $\mathbb{C}[X, \mathrm{conj}]$. Nous étudierons en détails ces phénomènes dans la suite de ce cours.

Notion de degré. La notion de degré, classique dans le cadre des polynômes commutatifs, s'étend sans difficultés aux polynômes de Ore.

Définition 1.1.5. Le *degré* d'un polynôme de Ore $P = a_0 + a_1 X + \cdots + a_d X^d \in \mathfrak{A}[X, \theta, \partial]$ est le plus grand entier i pour lequel $a_i \neq 0$.

La proposition suivante est immédiate.

Proposition 1.1.6. *Pour $P, Q \in \mathfrak{A}[X, \theta, \partial]$, on a :*

$$\deg(P+Q) \leq \max(\deg P, \deg Q) \tag{5}$$
$$\deg(PQ) \leq \deg P + \deg Q. \tag{6}$$

De plus, l'inégalité (5) est une égalité dès que $\deg(P) \neq \deg Q$ et l'inégalité (6) est une égalité dès que $\mathfrak{A}$ est intègre et que le morphisme θ est injectif.

Notons que lorsque l'anneau de base $\mathfrak{A}$ est un corps, le morphisme θ est nécessairement injectif. Ainsi l'égalité $\deg(PQ) = \deg P + \deg Q$ vaut inconditionnellement dans ce contexte.

1.2 Deux isomorphismes entre anneaux de polynômes de Ore

Les anneaux de Ore que nous venons de définir sont reliés entre eux par différents isomorphismes qui, dans certains cas, s'avérent très utiles pour réduire le champ d'étude. Dans ce numéro, nous détaillons deux tels isomorphismes. Le premier est ce que l'on appelle parfois le *twist d'Hilbert*. Il est particulièrement utile dans le cas l'anneau de base $\mathfrak{A}$ est un corps car il permet alors de se ramener soit à $\theta = \mathrm{id}$, soit à $\partial = 0$, ce qui peut être très commode dans certains cas. Le second isomorphisme que nous présentons permet d'interpréter, sous certaines hypothèses, l'anneau opposé d'un anneau de Ore comme un nouvel anneau de Ore associé à de nouveaux paramètres. Cette isomorphisme est particulièrement agréable pour « passer les propriétés de gauche à droite » et *vice-et-versa*.

1.2.1 Twist d'Hilbert

Le twist d'Hilbert est un changement de variable affine qui a pour effet de modifier la dérivation jusqu'à l'annuler dans certains cas favorables. En guise de préliminaire, remarquons qu'étant donné un endomorphisme d'anneaux $\theta : \mathfrak{A} \to \mathfrak{A}$, l'application $\partial = \theta - \mathrm{id}_{\mathfrak{A}}$ est une θ-dérivation de K. En effet, ∂ est clairement additive comme différence de deux fonctions additives et, de plus, pour $a, b \in \mathfrak{A}$, on a :

$$\partial(ab) = \theta(ab) - ab = \theta(a)\theta(b) - ab = \theta(a)(\theta(b) - b) + (\theta(a) - a)b = \theta(a)\partial(b) + \partial(a)b.$$

Proposition 1.2.1. *Soient ∂_1 et ∂_2 deux dérivations de K liés par la formule $\partial_1 = \partial_2 - c \cdot (\theta - \mathrm{id}_K)$. Alors on a un isomorphisme :*

$$\begin{array}{ccc} \mathfrak{A}[X_1, \theta, \partial_1] & \xrightarrow{\sim} & \mathfrak{A}[X_2, \theta, \partial_2] \\ X_1 & \mapsto & X_2 + c. \end{array}$$

Démonstration. Il suffit de vérifier que, pour $a \in \mathfrak{A}$, on a l'égalité :

$$(X_2 + c) \cdot a = \theta(a)(X_2 + c) + \partial_1(a)$$

ce qui est immédiat. □

Proposition 1.2.2. *On suppose qu'il existe $a \in \mathfrak{A}$ tel que $\theta(a) - a$ soit inversible dans $\mathfrak{A}$. Alors toute θ-dérivation sur $\mathfrak{A}$ est de la forme $c \cdot (\theta - \mathrm{id}_K)$ pour un $c \in \mathfrak{A}$. En particulier, $\mathfrak{A}[X, \theta, \partial] \simeq \mathfrak{A}[X, \theta, 0]$ pour toute dérivation $\partial : K \to K$.*

Remarque 1.2.3. En d'autres termes, la proposition affirme que, sous l'hypothèse d'existence d'un élément a tel que $\theta(a) - a$ est inversible, le $\mathfrak{A}$-module des θ-dérivations sur $\mathfrak{A}$ est libre de rang 1, engendré par $(\theta - \mathrm{id}_K)$.

Démonstration de la proposition 1.2.2. On pose $\partial_0 = \theta - \mathrm{id}_K$, on sait que c'est une θ-dérivation. Soit ∂ une deuxième θ-dérivation. Pour $x \in \mathfrak{A}$, on peut écrire :

$$\begin{aligned} \partial(ax) &= \theta(a)\partial(x) + \partial(a)x \\ &= \partial(xa) = \theta(x)\partial(a) + \partial(x)a. \end{aligned}$$

On en tire $\partial(x) = \frac{\partial(a)}{\partial_0(a)} \cdot \partial_0(x)$. Par conséquent, $\partial = c\partial_0$ avec $c = \frac{\partial(a)}{\partial_0(a)}$. La deuxième assertion de la proposition résulte de la première en vertu de la proposition 1.2.1. □

Lorsque $\mathfrak{A} = K$ est un corps, l'existence d'un élement $a \in \mathfrak{A}$ tel que $\theta(a) - a$ est inversible est assurée dès lors que θ n'est pas l'identité. Ainsi, sous ces hypothèses, la proposition 1.2.2 assure que $K[X, \theta, \partial] \simeq K[X, \theta, 0]$ pour toute θ-dérivation ∂. Ainsi, lorsque $\mathfrak{A} = K$ est un corps, on a l'alternative suivante : soit $\theta = \mathrm{id}_K$, soit $\theta \neq \mathrm{id}_K$ et on peut alors supposer, quitte à appliquer un twist d'Hilbert, que $\partial = 0$. Autrement dit, lorsque $\mathfrak{A} = K$, on peut séparer l'étude des anneaux de Ore $K[X, \theta, \partial]$ en deux catégories : celles des anneaux de Ore « purement endomorphiques » $K[X, \theta, 0]$ et celles des anneaux « purement différentiels » $K[X, \mathrm{id}_K, \partial]$. Dans la suite de ce cours, on utilisera les notations simplifiées $K[X, \theta]$ pour $K[X, \theta, 0]$ et $K[X, \partial]$ pour $K[X, \mathrm{id}_K, \partial]$; cela ne prêtera pas à confusion.

L'anneau opposé de $\mathfrak{A}[X, \theta, \partial]$

Rappelons que l'opposé d'un anneau non commutatif $(A, +, \times)$ est l'anneau $(A, +, \&)$ où la multiplication $\&$ est définie par $a\&b = b \times a$. Dans la suite, on notera A^{op} l'anneau opposé de A. Le but de ce paragraphe est de montrer que, lorsque θ est bijectif, l'anneau opposé de $\mathfrak{A}[X, \theta, \partial]$ est encore un anneau de polynômes de Ore.

Lemme 1.2.4. *On suppose que θ est bijectif. Soit $\partial : \mathfrak{A} \to \mathfrak{A}$ une θ-dérivation, alors $\partial \circ \theta^{-1}$ est une (θ^{-1})-dérivation.*

Démonstration. Il s'agit d'une simple vérification. Posons $d = \partial \circ \theta^{-1}$. Il est clair que d est additif. Par ailleurs, pour $a, b \in \mathfrak{A}$, on a :

$$d(ab) = d(ba) = \partial(\theta^{-1}(b)\theta^{-1}(a)) = b\, d(a) + d(b)\, \theta^{-1}(a)$$

ce qui conclut. □

Proposition 1.2.5. *On suppose que θ est bijectif. Alors, l'application :*

$$\begin{array}{rcl} \mathfrak{A}[X, \theta, \partial]^{op} & \longrightarrow & \mathfrak{A}[X^{op}, \theta^{-1}, -\partial \circ \theta^{-1}] \\ P = \sum_i a_i X^i & \mapsto & P^{op} = \sum_i (X^{op})^i a_i \end{array}$$

est un isomorphisme qui préserve le degré. De plus, pour tout $P \in \mathfrak{A}[X, \theta, \partial]$, on a $(P^{op})^{op} = P$.

Démonstration. Désignons par f l'application $\mathfrak{A}[X, \theta, \partial]^{\mathrm{op}} \to \mathfrak{A}[X^{\mathrm{op}}, \theta^{-1}, -\partial \circ \theta^{-1}]$ définie dans l'énoncé de la proposition. Clairement f est additif et préserve le degré. Grâce à l'additivité et la distributivité, il suffit, pour établir que f est un morphisme d'anneaux, de démontrer que $f(aX \cdot bX) = f(bX) \cdot f(aX)$ pour a et b dans $\mathfrak{A}$. On calcule :

$$f(aX \cdot bX) = f\big(a\theta(b)X^2 + a\partial(b)X\big) = (X^{\mathrm{op}})^2 \cdot a\theta(b) + X^{\mathrm{op}} \cdot a\partial(b) \tag{7}$$

$$f(bX) \cdot f(aX) = X^{\mathrm{op}} b \cdot X^{\mathrm{op}} a = X^{\mathrm{op}} \cdot (bX^{\mathrm{op}}) \cdot a. \tag{8}$$

Remarquons de plus que $X^{\mathrm{op}}\theta(b) = bX^{\mathrm{op}} - \partial(b)$. Ainsi $bX^{\mathrm{op}} = X^{\mathrm{op}}\theta(b) + \partial(b)$. En insérant cette dernière égalité dans (8), on obtient $f(bX) \cdot f(aX) = (X^{\mathrm{op}})^2 \cdot \theta(b)a + X^{\mathrm{op}} \cdot \partial(b)a$. Il ne reste plus maintenant qu'à comparer avec (7) pour en déduire l'égalité souhaitée.

Si $\theta^{\mathrm{op}} = \theta^{-1}$ et $\partial^{\mathrm{op}} = -\partial \circ \theta^{-1}$, on remarque que $(\theta^{\mathrm{op}})^{\mathrm{op}} = (\theta^{\mathrm{op}})^{-1} = \theta$ et $(\partial^{\mathrm{op}})^{\mathrm{op}} = -\partial^{\mathrm{op}} \circ (\theta^{\mathrm{op}})^{-1} = \partial$. Soit $f^{\mathrm{op}} : \mathfrak{A}[X^{\mathrm{op}}, \theta^{\mathrm{op}}, \partial^{\mathrm{op}}]^{\mathrm{op}} \to \mathfrak{A}[X, \theta, -\partial]$ le morphisme d'anneaux construit à partir de l'anneau de Ore $\mathfrak{A}[X^{\mathrm{op}}, \theta^{\mathrm{op}}, \partial^{\mathrm{op}}]$. La composée $f^{\mathrm{op}} \circ f$ est un endomorphisme d'anneaux de $\mathfrak{A}[X, \theta, \mathrm{op}]$ qui fixe les constantes ainsi que la variable X. On en déduit que $f^{\mathrm{op}} \circ f = \mathrm{id}$. De même, on vérifie que $f \circ f_{\mathrm{op}} = \mathrm{id}$. Il est résulte que f est bijectif et que $(P^{\mathrm{op}})^{\mathrm{op}}$ comme annoncé dans l'énoncé de la proposition. □

1.3 Division euclidienne et conséquences

On suppose dans ce numéro que $\mathfrak{A}$ est un corps et on le note K. L'anneau de Ore $K[X, \theta, \partial]$ est alors muni d'une division euclidienne à droite qui, de même que dans le cas des polynômes classiques, est un outil pour étudier les propriétés arithmétiques des polynômes de Ore. Le résultat précis s'énonce comme suit.

Proposition 1.3.1 (Division euclidienne à droite)**.** *Soient $A, B \in K[X, \theta, \partial]$ avec B non nul. Il existe un unique couple $(Q, R) \in K[X, \theta, \partial]^2$ tel que $A = QB + R$ et $\deg R < \deg B$.*

Démonstration. Commençons par démontrer l'unicité. Pour cela, on suppose qu'il existe deux écritures $A = Q_1B + R_1 = Q_2B + R_2$ vérifiant les conditions de la proposition. À partir de là, on obtient directement $(Q_1 - Q_2)B = R_2 - R_1$. Or $\deg(R_2 - R_1) < \deg B$, tandis que $\deg(Q_1 - Q_2)B = \deg(Q_1 - Q_2) + \deg B$ grâce à l'injectivité de θ (résultant, elle-même, du fait que l'anneau de base K est un corps). Ces deux égalités ne peuvent avoir lieu que si $Q_1 = Q_2$ et, par suite, $R_1 = R_2$. L'unicité est démontrée.

Pour l'existence, on procède par récurrence sur la différence $\deg A - \deg B$. Si cette différence est négative, c'est-à-dire si $\deg A < \deg B$, la division euclidienne de A par B s'écrit $A = 0 \times B + A$. Sinon, notons n (resp. m) le degré de A (resp. de B) et a_n (resp. b_m) son coefficient dominant. Considérons le polynôme $P = A - \lambda X^{n-m} B$ avec $\lambda = a_n\, \theta^{n-m}(\frac{1}{b_m})$. Il est construit de manière à être de degré au plus $n-1$. L'hypothèse de récurrence s'applique donc au couple (P, B) : il existe des polynômes de Ore $Q', R' \in K[X, \theta, \partial]$ tels que $P = Q'B + R'$ et $\deg R' < \deg B$. Revenant à la définition de P, on trouve $A = (\lambda X^{n-m} + Q')B + R'$ et l'existence de la division euclidienne à droite est démontrée. □

Notez bien, cependant, qu'il n'existe pas de division euclidienne à gauche en général : étant donné $A, B \in K[X, \theta, \partial]$ avec $B \neq 0$, on ne peut pas toujours écrire $A = BQ + R$ avec $\deg R < \deg B$. Par exemple, dans l'anneau de Ore $\mathbb{R}(t)[X, \theta : t \mapsto t^2]$ la division euclidienne à gauche de tX par X n'existe pas. En effet, une telle division euclidienne s'écrirait nécessairement $tX = Xq(t) + r(t)$ avec $q, r \in \mathbb{R}(t)$ et, en identifiant les coefficients en X, on obtiendrait $t = \theta(q(t)) = q(t^2)$, ce qui n'est manifestement pas possible. Toutefois, la division euclidienne à gauche existe toutefois lorsque θ est bijectif. Cela résulte du second isomorphisme du §1.2.

La propriété suivante est immédiate mais elle nous sera bien utile à plusieurs reprises dans la suite de ce cours.

Proposition 1.3.2. *Soient $A, B \in K[X, \theta, \partial]$ avec $\deg A \geq \deg B$ et B non nul. Alors, le quotient Q de la division euclidienne à droite de A par B a pour degré $\deg Q = \deg A - \deg B$.*

Cas particuliers et exemples

En pratique, la division euclidienne dans les anneaux de Ore se pose de la même manière que dans un anneau de polynômes commutatif. Par exemple, dans l'anneau $\mathbb{R}(t)[X, \frac{d}{dt}]$, la division à droite de $X^3 + (3 - t^2)X^2 + tX$ par $X - t^2$ se pose de la manière suivante :

$$
\begin{array}{r|l}
X^3+(3{-}t^2)X^2+\ tX+(3t{+}1) & X-t^2 \\
\cline{2-2}
-\ \big(X^3-\ \ t^2X^2-4tX-\ \ 2\big) & X^2+3X+(t^2+5t) \\
\cline{1-1}
3X^2+5tX+(3t{+}3) & \\
-\ \big(3X^2-t^2X-\ 2t\ \ \big) & \\
\cline{1-1}
(t^2{+}5t)X\ +(5t{+}3) & \\
-\ \big((t^2{+}5t)X\ -(t^4{+}5t^3)\big) & \\
\cline{1-1}
t^4{+}5t^3{+}5t{+}3 &
\end{array}
$$

Le quotient est $X^2+3X+(t^2{+}5t)$ tandis que le reste est $t^4{+}5t^3{+}5t{+}3$.

Étant donné que $\theta = \mathrm{id}_K$ est bijectif dans cet exemple, il est également possible d'effectuer la division euclidienne à gauche des polynômes précédents. On obtient, ce faisant, le calcul différent suivant :

$$
\begin{array}{r|l}
X^3+(3{-}t^2)X^2+\ \ tX+(3t{+}1) & X-t^2 \\
\cline{2-2}
-\ \big(X^3-\ \ t^2X^2\ \big) & \\
\cline{1-1}
3X^2+\ \ tX+(3t{+}1) & X^2+3X+(3t^2+t) \\
-\ \big(3X^2-3t^2X\ \big) & \\
\cline{1-1}
(3t^2{+}t)X\ +(3t{+}1) & \\
-\ \big((3t^2{+}t)X\ -(-3t^4{-}t^3{+}6t{+}1)\big) & \\
\cline{1-1}
3t^4{+}t^3{+}3t &
\end{array}
$$

qui conduit également à un résultat différent : le quotient, à présent, est $X^2+3X+(3t^2{+}t)$ et le reste $3t^4{+}t^3{+}3t$.

Lorsque $\partial = 0$, il existe une formule simple donnant le reste de la division euclidienne à droite par un polynôme de Ore de degré 1 comme le précise le lemme suivant.

Lemme 1.3.3. *Soient $a, a_1, \ldots, a_n \in K$. Dans l'anneau $K[X,\theta]$, le reste de la division euclidienne à droite de $a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ par $X - a$ est :*

$$\sum_{i=0}^{n} a_i\, a\, \theta(a)\, \cdots\, \theta^{i-1}(a).$$

Démonstration. On procède par récurrence sur n. Lorsque $n = 0$, il n'y a rien à démontrer. Sinon, écrivons $A = a_nX^n + \cdots + a_1X + a_0$ et, comme dans la démonstration du théorème de la division euclidienne (*cf* proposition 1.3.1), posons :

$$P = A - a_nX^{n-1}(X-a) = (a_{n-1} + a_n\theta^{n-1}(a))X^{n-1} + a_{n-2}X^{n-2} + \cdots + a_1X + a_0.$$

Clairement P a même reste que A dans la division euclidienne à droite par $X - a$. On conclut alors en utilisant l'hypothèse de récurrence. □

En particulier, lorsque a est de la forme $a = \frac{\theta(b)}{b}$ (pour $b \in K$), on obtient le corollaire suivant.

Corollaire 1.3.4. *Soient $b, a_1, \ldots, a_n \in K$ avec $b \neq 0$. Dans l'anneau $K[X,\theta]$, le reste de la division euclidienne à droite de $a_nX^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ par $X - \frac{\theta(b)}{b}$ est $\frac{1}{b}\sum_{i=0}^{n} a_i\theta^i(b)$.*

Il est intéressant de remarquer que ce dernier corollaire admet un analogue dans le cadre différentiel dont l'énoncé s'obtient, de manière relativement troublante, en remplaçant formellement chaque instance de θ par ∂.

Lemme 1.3.5. *Soient $g, f_1, \dots, f_n \in K$ avec $g \neq 0$. Dans l'anneau $K[X, \partial]$, le reste de la division euclidienne à droite de $f_n X^n + f_{n-1} X^{n-1} + \cdots + f_1 X + f_0$ par $X - \frac{\partial(g)}{g}$ est $\frac{1}{g} \sum_{i=0}^{n} f_i \partial^i(g)$.*

Démonstration. Il suffit de montrer que, pour tout entier $n \geq 0$, le reste de la division euclidienne de X^n par $X - \frac{\partial(g)}{g}$ est égal à $\frac{\partial^n(g)}{g}$. On procède par récurrence sur n. Pour $n = 0$, l'assertion est clairement vraie. Supposons à présent qu'elle le soit pour l'entier n, c'est-à-dire que l'on ait une écriture de la forme $X^n = P_n(X) \cdot \left(X - \frac{\partial(g)}{g}\right) + \frac{\partial^n(g)}{g}$ pour un certain polynôme de Ore $P_n(X) \in K[X, \partial]$. En multipliant l'égalité précédente par X à gauche, on obtient la relation :

$$\begin{aligned} X^{n+1} &= X \cdot P_n(X) \cdot \left(X - \tfrac{\partial(g)}{g}\right) + X \cdot \tfrac{\partial^n(g)}{g} \\ &= X \cdot P_n(X) \cdot \left(X - \tfrac{\partial(g)}{g}\right) + \tfrac{\partial^n(g)}{g} X + \tfrac{\partial^{n+1}(g)}{g} - \tfrac{\partial^n(g)}{g} \cdot \tfrac{\partial(g)}{g} \\ &= \left(X \cdot P_n(X) + \tfrac{\partial^n(g)}{g}\right) \cdot \left(X - \tfrac{\partial(g)}{g}\right) + \tfrac{\partial^{n+1}(g)}{g} \end{aligned}$$

d'où on déduit l'assertion souhaitée au rang $n+1$. □

Remarque 1.3.6. L'ensemble des applications additives de K dans lui-même pour les lois d'addition et de composition. Notons-le $\mathrm{End}_{\mathbb{Z}}(K)$. Pour $\bullet = \theta$ ou ∂, on dispose d'une application d'*évaluation* :

$$\begin{array}{rcl} K[X, \bullet] & \xrightarrow{\mathrm{ev}_\bullet} & \mathrm{End}_{\mathbb{Z}}(K) \\ P(X) = a_0 + a_1 X + \cdots + a_d X^d & \mapsto & P(\bullet) = a_0 + a_1 \bullet + \cdots + a_d \bullet^d . \end{array}$$

qui, aussi bien dans le cas d'un endomorphisme que dans celui d'un dérivation, s'avère être un morphisme d'anneaux. De plus, le corollaire 1.3.4 (dans le cas d'un endomorphisme) et le lemme 1.3.5 (dans le cas d'une dérivation) font le lien entre ce morphisme d'évaluation et le reste dans la division euclidienne à droite par un polynôme de Ore de degré 1.

Principalité des anneaux de polynômes de Ore

Au niveau structurel, l'existence d'une division euclidienne admet les mêmes conséquences que dans le cadre commutatif. Si A est un anneau, on rappelle qu'un idéal à gauche (resp. à droite) $\mathcal{I}$ de A est un sous-groupe additif de A vérifiant la condition supplémentaire suivante : pour tout $a \in A$ et tout $x \in \mathcal{I}$, le produit ax (resp. xa) appartiennent à $\mathcal{I}$.

Corollaire 1.3.7. *Tout idéal à gauche de $K[X, \theta, \partial]$ est principal,* i.e. *de la forme $K[X, \theta, \partial]\, P$ pour un certain polynôme de Ore $P \in K[X, \theta, \partial]$.*

Démonstration. Soit $\mathcal{I}$ un idéal à gauche de $K[X, \theta, \partial]$. On considère un élément $B \in \mathcal{I}$ non nul de degré minimal. Soit A dans $\mathcal{I}$. On écrit la division euclidienne de A par B : on a $A = QB + R$ avec $\deg R < \deg B$. De l'égalité $R = A - QB$, on déduit que $R \in \mathcal{I}$. Par la condition de minimalité sur le degré, il en résulte que $R = 0$. Ainsi $A = QB$. L'idéal $\mathcal{I}$ est donc l'idéal principal engendré par B. □

Remarque 1.3.8. Lorsque θ est bijectif, il existe une division euclidienne à gauche sur $K[X, \theta, \partial]$, à partir de quoi on déduit comme ci-dessus que $K[X, \theta, \partial]$ est aussi principal à droite.

De même que dans le cadre commutatif, l'existence d'une division euclidienne entraîne aussi l'existence de PGCD, de PPCM ainsi que la validité d'un théorème de Bézout. Il faut néanmoins faire attention à ne pas mélanger la gauche et la droite. Précisément, fort d'une division à droite, on peut définir un PGCD à droite (noté RGCD) et un PPCM à gauche (noté LLCM) comme suit.

Définition 1.3.9. Étant donnés $A, B \in K[X, \theta, \partial]$, on note $\text{RGCD}(A, B)$ et $\text{LLCM}(A, B)$ les deux polynômes unitaires pour lesquels :

$$K[X, \theta, \partial]\, P + K[X, \theta, \partial]\, B = K[X, \theta, \partial]\, \text{RGCD}(A, B)$$
$$K[X, \theta, \partial]\, P \cap K[X, \theta, \partial]\, B = K[X, \theta, \partial]\, \text{LLCM}(A, B).$$

On dispose en outre d'une variante non commutative de l'algorithme d'Euclide permettant de calculer des RGCD et des LLCM. Pour l'expliquer considérons deux polynômes de Ore $A, B \in K[X, \theta, \partial]$ avec $B \neq 0$. Posons $R_0 = A$, $R_1 = B$ et, pour $i \geq 2$, tant que $R_i \neq 0$, définissons R_{i+1} comme le reste de la division euclidienne de R_{i-1} par R_i.

Proposition 1.3.10. *Avec les notations précédentes, si n est le plus grand entier pour lequel $R_n \neq 0$, on a $\text{RGCD}(A, B) = \frac{1}{\text{lc}(R_n)} \cdot R_n$ où $\text{lc}(R_n)$ désigne le coefficient dominant de R_n.*

Démonstration. Soit $i \in \{1, \ldots, n\}$. De l'écriture $R_{i+1} = R_{i-1} - Q_i R_i$ (pour un $Q_i \in K[X, \theta, \partial]$), on déduit que les idéaux à gauche de $K[X, \theta, \partial]$ engendrés par R_{i-1} et R_i, d'une part, et par R_i et R_{i+1}, d'autre part, coïncident. Par récurrence, il s'ensuit que l'idéal engendré par $R_0 = A$ et $R_1 = B$ est égal à l'idéal engendré par R_n (puisque $R_{n+1} = 0$). La proposition en découle. □

De même que dans le cas commutatif, on peut étendre l'algorithme d'Euclide pour calculer en même temps les coefficients de Bézout. Pour ce faire, on conserve la suite R_i définie précédemment mais on définit de plus de nouvelles suites de polynômes de Ore $(U_i)_{1 \leq i \leq n}$ et $(V_i)_{1 \leq i \leq n}$ par les formules récurrentes suivantes :

$$U_0 = 1, \quad V_0 = 0, \quad U_1 = 0, \quad V_1 = 1$$
$$U_{i+1} = U_{i-1} - Q_i U_i, \quad V_{i+1} = V_{i-1} - Q_i V_i$$

où Q_i désigne le quotient de la division euclidienne de R_{i-1} par R_i. Il est immédiat de vérifier par récurrence que $U_i A + V_i B = R_i$ pour tout entier i. Ainsi, en posant $U = \frac{1}{\text{lc}(R_n)} \cdot U_n$ et $V = \frac{1}{\text{lc}(R_n)} \cdot V_n$, on obtient la relation de Bézout :

$$UA + VB = \text{RGCD}(A, B).$$

On dispose de plus, de même que dans le cas commutatif, d'une formule pour les degrés des polynômes U_i et V_i (qui vaut donc également, en particulier, pour les coefficients de Bézout U et V), donnée par le lemme suivant :

Lemme 1.3.11. *Pour tout entier $i \in \{2, 3, \ldots, n+1\}$, on a*

$$\deg U_i = \deg B - \deg R_{i-1} \quad \textit{et} \quad \deg V_i = \deg A - \deg R_{i-1}.$$

Démonstration. Sans nuire à la généralité, on peut supposer $\deg A \geq \deg B$. Nous allons montrer le résultat par récurrence sur i. Lorsque $i = 2$, on a $U_2 = 1$, $V_2 = -Q_1$ et $R_1 = B$, à partir de quoi les formules annoncées se vérifient sans peine. Pour $i \in \{3, \ldots, n\}$, il suit de l'hypothèse de récurrence que :

$$\begin{aligned} \deg U_{i-1} &= \deg B - \deg R_{i-2} && \text{si } i > 3 \\ &= -\infty && \text{si } i = 3 \end{aligned}$$

$$\begin{aligned} \deg Q_i U_i &= \deg Q_i + \deg U_i \\ &= \deg R_{i-1} - \deg R_i + \deg B - \deg R_{i-1} = \deg B - \deg R_i. \end{aligned}$$

On en déduit que $\deg U_{i-1} < \deg Q_i U_i$ et, par suite, que le degré de $U_{i+1} = U_{i-1} - Q_i U_i$ est égal à celui de $Q_i U_i$, c'est-à-dire à $\deg B - \deg R_i$. On démontre exactement de la même manière que $\deg V_{i+1} = \deg A - \deg R_i$, ce qui termine la récurrence. □

L'algorithme d'Euclide permet également de calculer le LLCM. En effet, au rang $i = n+1$, la relation $U_iA + V_iB = R_i$ conduit à l'égalité $U_{n+1}A = -V_{n+1}B$. Ainsi cette quantité commune apparaît comme un multiple à gauche commun de A et de B. Il s'avère en fait que celle-ci est bel et bien le LLCM (à un facteur de renormalisation près) comme le précise la proposition suivante.

Proposition 1.3.12. *Avec les notations précédentes, on a :*

$$\text{LLCM}(A,B) = \frac{U_{n+1}A}{lc(U_{n+1}A)} = \frac{V_{n+1}B}{lc(V_{n+1}B)}.$$

Démonstration. Pour la démonstration, nous avons besoin d'anticiper légèrement sur des résultats à venir et admettre temporairement la proposition suivante (qui sera démontrée au §1.4, voir proposition 1.4.5 et la remarque qui la suit).

Proposition 1.3.13. *Pour $A, B \in K[X,\theta,\partial]$ avec $(A,B) \neq (0,0)$, on a :*

$$\deg \text{RGCD}(A,B) + \deg \text{LLCM}(A,B) = \deg A + \deg B.$$

Fort de ce résultat, on peut argumenter comme suit. Tout d'abord, on remarque que, par le lemme 1.3.11, on a $\deg U_{n+1} = \deg B - \deg R_n$ et donc $\deg U_{n+1} = \deg B - \deg \text{RGCD}(A,B)$ d'après la proposition 1.3.10. Par suite :

$$\deg(U_{n+1}A) = \deg A + \deg B - \deg \text{RGCD}(A,B) = \deg \text{LLCM}(A,B) \tag{9}$$

la dernière égalité résultant de la proposition 1.3.13. Comme $U_{n+1}A$ est un multiple commun à A et B, il est aussi un multiple de $\text{LLCM}(A,B)$ et l'inégalité (9) implique donc que $U_{n+1}A$ et $\text{LLCM}(A,B)$ sont égaux à multiplication près par un scalaire non nul. □

1.4 Modules sur les anneaux de polynômes de Ore

L'étude des modules sur les anneaux $\mathfrak{A}[X,\theta,\partial]$ est intéressante à plusieurs titres. Premièrement, elle permet de faire le pont entre les polynômes de Ore, d'une part, et l'algèbre semi-linéaire ou la théorie des équations différentielles linéaire, d'autre part. Elle donne ainsi aux anneaux de Ore une dimension nouvelle. Deuxièmement, l'introduction de méthodes algébriques issues de la théorie des modules permet de clarifier certaines constructions (comme nous l'avons déjà entrevu dans le cas des RGCD et LLCM) et, plus généralement, offre un nouveau regard et une nouvelle intuition, souvent très fructueux, sur les polynômes de Ore.

À toutes fins utiles, on commence par rappeler la définition suivante.

Définition 1.4.1. Soit R un anneau non nécessairement commutatif. Un *R-module à gauche* est un groupe commutatif M muni d'une multiplication externe $R \times M \to M$, $(a,x) \mapsto a \cdot x$ vérifiant les axiomes suivants :

- $a \cdot (x+y) = ax + ay$,
- $(a+b) \cdot x = ax + by$,
- $1 \cdot x = x$,
- $a \cdot (bx) = (ab) \cdot x$

pour $a, b \in R$ et $x, y \in M$.

Étant donné un anneau $\mathfrak{A}$ commutatif, on rappelle que la donnée d'un $\mathfrak{A}[X]$-module est équivalente à la donnée d'un $\mathfrak{A}$-module muni d'un endomorphisme $\mathfrak{A}$-linéaire, ce dernier endomorphisme correspondant à la multiplication par X. De la même manière, la donnée d'un module à gauche sur $\mathfrak{A}[X,\theta,\partial]$ est équivalente à la donnée d'un $\mathfrak{A}$-module M muni d'une application additive $f : M \to M$ telle que :

$$\forall a \in \mathfrak{A},\ \forall x \in M, \quad f(ax) = \theta(a)f(x) + \partial(a)x. \tag{10}$$

De la même manière que les polynômes interviennent dans les théorèmes de réduction des applications linéaires, les polynômes de Ore peuvent être utilisés pour étudier les endomorphismes vérifiant l'axiome (10). Lorsque $\partial = 0$, l'axiome (10) est celui des applications semi-linéaires tandis que lorsque $\theta = \mathrm{id}_K$, c'est celui des dérivations/connexions.

Dans le cas général, une application additive vérifiant l'axiome (10) s'appelle une *transformation pseudo-linéaire*. Ces transformations ont été introduites et étudiées par Jacobson dans [15] ; nous y reviendrons plus en détails au §**??**.

Modules cycliques. Afin de préciser encore davantage les liens entrevus ci-dessus et d'établir des résultats concrets dans cette direction, nous nous proposons d'étudier en détails l'exemple des modules cycliques. À partir de maintenant, *nous supposons pour simplifier que l'anneau de base $\mathfrak{A}$ est un corps et on le note K.*

Définition 1.4.2. Un $K[X, \theta, \partial]$-module à gauche est dit *cyclique* s'il est engendré par un unique élément.

Soit M un $K[X, \theta, \partial]$ un module cyclique. Soit $x \in M$ un générateur. L'application linéaire naturelle $f_x : K[X, \theta, \partial] \to M$, $A \mapsto Ax$ est alors surjective. De plus, son noyau est un idéal à gauche, il est donc de la forme $K[X, \theta, \partial]P$ pour un certain polynôme de Ore P, éventuellement nul. Il s'ensuit que f_x induit un isomorphisme :

$$M \simeq M_P \quad \text{avec} \quad M_P = K[X, \theta, \partial]/K[X, \theta, \partial]\, P.$$

Autrement dit, tout module cyclique est de la forme M_P pour un certain polynôme de Ore P. Le lemme suivant est une évidence mais il nous sera très utile.

Lemme 1.4.3. *Tout $K[X, \theta, \partial]$-module contient un module cyclique.*

Démonstration. Il suffit de considérer le sous-module engendré par n'importe quel élément. □

Les propositions qui suivent font le lien entre la factorisation dans $K[X, \theta, \partial]$ et la décomposition des modules cycliques.

Proposition 1.4.4. *Soit $P \in K[X, \theta, \partial]$. La fonction $A \mapsto M_A$ réalise une bijection de l'ensemble des diviseurs à droite unitaires de P dans l'ensemble des quotients de M_P.*

*De plus, si P s'écrit $P = BA$ (avec $A, B \in K[X, \theta, \partial]$), le noyau de la projection canonique $M_P \to M_A$ est canoniquement isomorphe à M_B (*via *la multiplication à droite par A) de sorte que l'on a une suite exacte :*

$$0 \to M_B \to M_P \to M_A \to 0. \tag{11}$$

Démonstration. Soit M un quotient de M_P. Soit $f : K[X, \partial, \theta] \to M_P \to M$ la composée des deux projections canoniques. Le noyau de f est un idéal de $K[X, \partial, \theta]$ qui contient P. D'après le corollaire 1.3.7, il est de la forme $K[X, \partial, \theta]A$ pour un certain polynôme de Ore A. De plus, A est uniquement déterminé si on demande qu'il soit unitaire. Enfin, le fait que $P \in \ker f$ indique que A est un diviseur à droite de P. On a ainsi démontré, qu'en tant que quotient de M_P, M est isomorphe à M_A. La première assertion de la proposition en résulte.

Pour démontrer la seconde assertion, notons tout d'abord que, d'après ce qui vient d'être fait, nous savons déjà que le noyau de la projection $M_P \to M_A$ est égal au sous-module $M_P\, A$. Ainsi il suffit de démontrer que la multiplication à droite par A induit un isomorphisme $M_B \to M_P\, A$. Cela résulte du fait que l'application $K[X, \partial, \theta] \to K[X, \partial, \theta]$, $R \mapsto RA$ réalise une bijection $K[X, \partial, \theta] \to K[X, \partial, \theta]\, A$ qui envoie $K[X, \partial, \theta]\, B$ sur $K[X, \partial, \theta]\, P$ et induit ainsi un isomorphisme au niveau des quotients. □

Proposition 1.4.5. *Soient $A, B \in K[X,\theta,\partial]$. On a une suite exacte :*

$$0 \to M_{\mathrm{LLCM}(A,B)} \xrightarrow{f} M_A \oplus M_B \xrightarrow{g} M_{\mathrm{RGCD}(A,B)} \to 0 \tag{12}$$

où le morphisme f est induit par les projections canoniques et g provient de l'application $(S,T) \mapsto S - T$.

Démonstration. Il est clair que la composée $g \circ f$ s'annule. Le fait que f soit injectif résulte de la définition du LLCM. Soit un couple $(S,T) \in M_A \oplus M_B$ dont l'image par g s'annule. Soit $\hat{S}$ (resp. $\hat{T}$) un représentant de S (resp. de T) dans $K[X,\theta,\partial]$. La condition $g(S,T) = 0$ indique qu'il existe $Q \in K[X,\theta,\partial]$. tel que $\hat{T} = \hat{S} + Q \cdot \mathrm{RGCD}(A,B)$ avec Par le théorème de Bézout, on sait qu'il existe $U, V \in K[X,\theta,\partial]$ tels que $UA + VB = \mathrm{RGCD}(A,B)$. Formons le polynôme $P = \hat{S} + QUA$. Son image dans M_A est clairement S tandis que l'égalité $P = \hat{S} + Q \cdot (\mathrm{RGCD}(A,B) - VB) = \hat{T} - QVB$ montre que son image dans M_B est T. La suite (12) est donc exacte au milieu. Il ne reste plus qu'à vérifier la surjectivité de g mais elle est évidente étant donné que la projection $M_A \to M_{\mathrm{RGCD}(A,B)}$ est déjà surjective. □

Remarque 1.4.6. Un corollaire immédiat de la proposition 1.4.5 est la proposition 1.3.13 qui avait été laissée en suspens.

Morphismes entre modules cycliques. Si M et M' sont deux $K[X,\theta,\partial]$-modules, notons $\mathrm{Hom}_{K[X,\theta,\partial]}(M,M')$ l'ensemble des morphismes $K[X,\theta,\partial]$-linéaires de M dans M'. Nous nous proposons ci-dessous de décrire l'espace $\mathrm{Hom}_{K[X,\theta,\partial]}(M_P, M_{P'})$ lorsque P et P' sont des polynômes de Ore que nous supposons non nuls.

En vertu de la cyclicité, un morphisme $K[X,\theta,\partial]$-linéaire $f : M_P \to M_{P'}$ est entièrement déterminée par l'image de la classe de $1 \in M_P$. Précisément, si Q est un polynôme de Ore tel que $f(1) = Q \pmod{M_{P'}}$, alors on a $f(A) = AQ \pmod{M_{P'}}$ pour tout $A \in M_P$. Toutefois, tout polynôme de Ore Q ne définit pas un morphisme $f : M_P \to M_{P'}$: il y a des conditions sur Q pour cela. Concrètement, pour que la multiplication à droite par Q définisse un morphisme de M_P dans $M_{P'}$, il faut et il suffit que l'on ait l'inclusion entre idéaux $K[X,\theta,\partial]PQ \subset K[X,\theta,\partial]P'$. Autrement dit, il faut et il suffit que P' divise PQ à droite, c'est-à-dire que :

$$\exists\, Q' \in K[X,\theta,\partial], \quad PQ = Q'P'. \tag{13}$$

Par ailleurs, on remarque que clairement deux polynômes de Ore Q_1 et Q_2 définissent le même morphisme si et seulement s'ils sont congrus modulo P'. Autrement dit, $\mathrm{Hom}_{K[X,\theta,\partial]}(M_P, M_{P'})$ s'identifie canoniquement au sous-ensemble de $M_{P'}$ formé des polynômes Q pour lesquels la condition (13) est vérifiée. Dans le cas particulier où $P = P'$, l'ensemble des endomorphismes de M_P forme à l'évidence un anneau (pour la composition). On l'appelle l'*eigenring* de P.

Proposition 1.4.7. *Soient $Q \in K[X,\theta,\partial]$ vérifiant la condition (13) et $f : M_P \to M_{P'}$ le morphisme qui lui est associé. Alors :*

(i) f est surjectif si et seulement si $\mathrm{RGCD}(P',Q) = 1$,

(ii) f est bijectif si et seulement si $\mathrm{RGCD}(P',Q) = 1$ et $\deg P = \deg Q$.

Démonstration. Soit $\pi : K[X,\theta,\partial] \to M_{P'}$ la projection canonique. Il suit des définitions que l'image de f est le sous-module de $M_{P'}$ engendré par Q. Ainsi $\pi^{-1}(\mathrm{im} f)$ est l'idéal de $K[X,\theta,\partial]$ engendré par P' et Q, *i.e.* $\pi^{-1}(\mathrm{im} f) = K[X,\theta,\partial]\,\mathrm{RGCD}(P',Q)$. Or, puisque π est surjectif, f l'est également si et seulement si $\pi^{-1}(\mathrm{im} f) = K[X,\theta,\partial]$. Le (i) de la proposition s'ensuit.

Pour ce qui concerne le (ii), il s'agit de montrer que f est bijectif si et seulement s'il est bijectif et l'égalité des dimensions $\dim_K M_P = \dim_K M_{P'}$ est avérée. Au vu de cela, la proposition résulte de ce que $\dim_K M_P = \deg P$ et $\dim_K M_{P'} = \deg P'$. □

Définition 1.4.8. Soient $P, P' \in K[X, \theta, \partial]$ deux polynômes de Ore. On dit que P et P' sont *associés* lorsque les $K[X, \theta, \partial]$-modules cycliques M_P et $M_{P'}$ sont isomorphes (abstraitement).

Il est évident, au vu de la définition, que la relation « être associée » est une relation d'équivalence. Par ailleurs, la description explicite de $\mathrm{Hom}(M_P, M_{P'})$ établie précédemment entraîne que deux polynômes de Ore P et P' sont associés si et seulement s'il existe $Q, Q' \in K[X, \theta, \partial]$ tels que :

(i) $\deg P = \deg P'$,

(ii) $PQ = Q'P'$,

(iii) $\mathrm{RGCD}(P', Q) = 1$.

En outre, quitte à remplacer Q par le reste de sa division euclidienne à droite par P', on peut toujours supposer que $\deg Q < \deg P' = \deg P$. Cette inégalité entraîne que $\deg Q' < \deg P$ à son tour. La reformulation précédente ne fait pas apparaître aussi clairement les propriétés de la relation « être associé » mais, en contrepartie, elle a l'avantage d'être concrète et de pouvoir être utilisée avec efficacité sur des exemples explicites.

Exemple 1.4.9. Examinons ce que devient la relation « être associé » dans le cadre commutatif, c'est-à-dire lorsque $\theta = \mathrm{id}_K$ et $\partial = 0$. Considérons donc P et P' deux polynômes associés (au sens de la définition 1.4.8) dans l'anneau $K[X]$. Par la condition (ii) de la reformulation, le polynôme P' doit diviser PQ, ce qui, en vertu de la condition (iii) et du lemme de Gauss, entraîne que P' divise P. Par la condition (i), on déduit alors que $P = aP'$ pour un certain $a \in K$, $a \neq 0$. Dans le cadre commutatif, être associé au sens de la définition 1.4.8 correspond donc à la définition usuelle d'être associé dans un anneau de polynômes.

Exemple 1.4.10. Examinons à quelle condition les polynômes $X - a$ et $X - b$ sont similaires dans $\mathbb{C}[X, \mathrm{conj}]$. Cela se produit si et seulement s'il existe $c, d \in \mathbb{C}$ tels que $(X - a)c = d(X - b)$, soit encore $\bar{c}X - ac = dX + db$. En identifiant les coefficients, on obtient les conditions $d = \bar{c}$ et $ac = db$, ce qui se réécrit $ca = \bar{c}b$. L'existence d'un tel c est équivalente à l'égalité des nombres $|a| = |b|$. En conclusion, les polynômes $X - a$ et $X - b$ sont similaires dans $\mathbb{C}[X, \mathrm{conj}]$ si et seulement si $|a| = |b|$.

Comme première application de la notion de polynômes associés, énonçons un raffinement de la proposition 1.4.4.

Proposition 1.4.11. *Soit $P \in K[X, \theta, \partial]$. On suppose que P se factorise comme suit : $P = BA = A'B'$ avec $A, B, A', B' \in K[X, \theta, \partial]$ tels que $\deg A = \deg A'$ et $\mathrm{RGCD}(A, B') = 1$. Alors la suite exacte (11) (associée à la factorisation $P = BA$) est scindée.*

Démonstration. Les hypothèses entraînent que A et A' sont associés, c'est-à-dire que $M_A \simeq M_{A'}$. D'autre part, la multiplication par B' induit un morphisme de $M_{A'}$ dans M_P et donc, par composition avec l'isomorphisme précédent, elle induit un morphisme $s : M_A \to M_P$. On vérifie alors que la composée de s avec la projection $M_P \to M_A$ est un isomorphisme, ce qui permet de conclure. □

1.5 Du théorème de Jordan–Hölder au théorème de factorisation de Ore

Le théorème de Jordan–Hölder est un théorème de structure général des modules de longueur finie sur un anneau non nécessairement commutatif, qui admet des variantes dans le cadre des groupes finis ou des objets d'une catégorie abélienne. Pour simplifier l'exposition, nous ne le présentons ici que dans le cas particulier des modules à gauche sur l'anneau de Ore $K[X, \theta, \partial]$ où K est un corps. Nous commençons par une définition classique.

Définition 1.5.1. Un module à gauche sur $K[X, \theta, \partial]$ est dit *simple* s'il n'admet pas de sous-module non trivial.

Clairement, tout module simple est cyclique. De plus, d'après la proposition 1.4.4, le module M_P est cyclique si et seulement si P est irréductible dans l'anneau $K[X,\theta,\partial]$. Le théorème de Jordan–Hölder s'énonce comme suit.

Théorème 1.5.2 (Jordan–Hölder)**.** *Soit M un $K[X,\theta,\partial]$-module qui est de dimension finie comme K-espace vectoriel. Alors :*

1. *M admet une suite de composition,* i.e. *une suite de la forme :*
$$0 = M_0 \subset M_1 \subset \cdots \subset M_n = M$$
où tous les quotients M_i/M_{i-1} ($1 \leq i \leq n$) sont simples.
2. *Tout suite de composition de M a la même longueur et, de plus, les quotients M_i/M_{i-1} sont toujours les mêmes à réordonnement près.*

Remarque 1.5.3. Les quotient M_i/M_{i-1} qui, d'après le théorème de Jordan–Hölder, sont canoniquement associés à M s'appelle les *composants de Jordan–Hölder* de M.

Pour démontrer le théorème de Jordan–Hölder, nous aurons besoin du lemme suivant.

Lemme 1.5.4. *Soit M un module à gauche sur $K[X,\theta,\partial]$. Soient A et B deux sous-modules différents de M tels que les quotients M/A et M/B soient tous les deux simples. Alors :*

i) $M/(A\cap B) \simeq M/A \times M/B$,

ii) $A/(A\cap B) \simeq M/B$,

iii) $B/(A\cap B) \simeq M/A$.

Démonstration. Quitte à remplacer M par $M/(A\cap B)$ et A et B par leurs images dans ce quotient, on peut supposer que $A\cap B = 0$.

Considérons le morphisme $f : A\times B \to M, (a,b) \mapsto a+b$. Si un couple (a,b) est dans le noyau de f, il est nécessairement de la forme $(x,-x)$ où x appartient à la fois à A et B. Comme on a supposé $A\cap B = 0$, on en déduit que f est injective. Par ailleurs, le conoyau de f apparaît comme un quotient de M/A. Comme ce dernier module est supposé simple, on est dans l'alternative suivante : soit $\operatorname{coker} f = 0$ (*i.e* f est surjective), soit $\operatorname{coker} f = M/A$ (*i.e.* $\operatorname{im} f = A$). Le deuxième cas n'est cependant pas possible. En effet, si l'image de f était exactement A, on en déduirait que B est inclus dans A. De plus, cette inclusion devrait être stricte puisque A et B sont supposés différents. Mais alors M/A apparaîtrait comme un quotient strict de M/B, ce qui est exclu par simplicité de M/B. En conclusion, f est surjective et est donc un isomorphisme, *i.e.* $M \simeq A\times B$. Les trois propriétés du lemme s'en déduisent aisément. □

Démonstration du théorème 1.5.2. On raisonne par récurrence sur la dimension de M vu comme espace vectoriel sur K. La première assertion, à savoir l'existence d'une suite de composition, est facile : si M est simple, il n'y a rien à faire. Sinon, il existe par définition un sous-module strict S de M, de sorte que l'hypothèse de récurrence s'applique à S et à $T = M/S$: il existe deux suites de composition

$$0 = S_0 \subset S_1 \subset \cdots \subset S_s = S \quad \text{et} \quad 0 = T_0 \subset T_1 \subset \cdots \subset T_t = T.$$

Si $\pi : M \to T$ désigne la projection canonique, la suite d'inclusions

$$0 = S_0 \subset S_1 \subset \cdots \subset S_s = S = \pi^{-1}(T_0) \subset \pi^{-1}(T_1) \subset \cdots \subset \pi^{-1}(T_t) = M$$

est alors une suite de composition pour M.

Montrons à présent la seconde assertion. Considérons pour cela deux suites de composition de M :

$$0 = A_0 \subset A_1 \subset \cdots \subset A_r = M$$
$$\text{et} \quad 0 = B_0 \subset B_1 \subset \cdots \subset B_s = M.$$

Il s'agit de démontrer que $r = s$ et que les quotients A_i/A_{i-1} $(1 \leq i \leq r)$ et B_j/B_{j-1} $(1 \leq j \leq s)$ sont les mêmes à réordonnement près.

Supposons dans un premier temps que $A_{r-1} = B_{s-1}$ et appelons N ce module commun. L'hypothèse de récurrence appliquée à N nous apprend alors que $r - 1 = s - 1$, et donc $r = s$, et que les quotients A_i/A_{i-1} $(1 \leq i \leq r-1)$ et B_j/B_{j-1} $(1 \leq j \leq s-1)$ sont les mêmes à réordonnement près. Comme de plus, clairement $A_r/A_{r-1} = M/N = B_s/B_{s-1}$, on conclut.

Il reste donc à traiter le cas où $A_{r-1} \neq B_{s-1}$. Pour alléger les notations, posons $A = A_{r-1}$, $B = B_{s-1}$ et $C = A \cap B$. Considérons une suite de composition pour C (dont l'existence est assurée par l'hypothèse de récurrence) :

$$0 = C_0 \subset C_1 \subset \cdots \subset C_t = C.$$

Celle-ci nous permet de construire une nouvelle suite de composition pour A, à savoir :

$$0 = C_0 \subset C_1 \subset \cdots \subset C_t \subset C_{t+1} = A.$$

En effet, par le lemme 1.5.4, le quotient A/C est isomorphe à M/B et est donc simple. Par l'hypothèse de récurrence, on déduit que $r = t + 1$ et que les familles de modules :

$$\big(A_1/A_0, \ldots, A_{r-1}/A_{r-2}\big) \quad \text{et} \quad \big(C_1/C_0, \ldots, C_t/C_{t-1}, A/C\big)$$

sont les mêmes à réordonnement près. De la même manière, on démontre que $s = t + 1$ et que les familles de modules : que $r = t + 1$ et que les familles de modules :

$$\big(B_1/B_0, \ldots, B_{s-1}/B_{s-2}\big) \quad \text{et} \quad \big(C_1/C_0, \ldots, C_t/C_{t-1}, B/C\big)$$

sont également les mêmes à réordonnement près. Il en résulte ainsi déjà que $r = s$. Par ailleurs, le lemme 1.5.4 nous indique que $A/C \simeq M/B$ et $B/C \simeq M/A$. On en déduit que les familles

$$\big(A_1/A_0, \ldots, A_{r-1}/A_{r-2}, M/A\big) \quad \text{et} \quad \big(B_1/B_0, \ldots, B_{s-1}/B_{s-2}, M/B\big)$$

sont isomorphes à réordonnement près, comme voulu. □

Application à la factorisation des polynômes de Ore. Il est bien connu que lorsque K est un corps, l'anneau des polynômes usuels $K[X]$ est factoriel : tout polynôme de $K[X]$ admet une unique factorisation, à l'ordre près, en produit de polynômes unitaires irréductibles. Dans l'anneau des polynômes de Ore, cette propriété simple d'unicité ne vaut plus. En effet, nous avons déjà vu dans l'exemple 1.1.4 que dans $\mathbb{C}[X, \mathrm{conj}]$, on a $X^2 - c^2 = (X - a)(X + \bar{a})$ pour tout nombre complexe a de module c. Il existe toutefois un résultat de structure des factorisations, dû à Ore lui-même.

Définition 1.5.5. Un polynôme de Ore $P \in K[X, \theta, \partial]$ est dit *irréductible* s'il ne peut s'écrire sous la forme $P = AB$ avec $A, B \in K[X, \theta, \partial]$ et $\deg A, \deg B > 0$.

Théorème 1.5.6 (Ore)**.** *Soit P dans $K[X, \theta, \partial]$. Alors :*

1. *P admet une factorisation en produits de polynômes de Ore irréductibles*
2. *si $P = A_1 A_2 \cdots A_n = B_1 B_2 \cdots B_m$ sont deux telles factorisations, alors $n = m$ et il existe une permutation σ de $\{1, \ldots, n\}$ telle que A_i soit associé à $B_{\sigma(i)}$ pour tout i.*

Démonstration. Il s'agit d'une reformulation du théorème de Jordan–Hölder *via* le dictionnaire de la proposition 1.4.4. □

Exemple 1.5.7. Reprenons l'exemple 1.1.4 dans lequel on a établi la factorisation suivante dans $\mathbb{C}[X, \mathrm{conj}]$

$$X^2 - c^2 = (X - a)(X + \bar{a}) \quad \text{si } |a| = c.$$

On a vu, par ailleurs, dans l'exemple 1.4.10 que les polynômes $X - a$ et $X - b$ sont associés dès lors que $|a| = |b|$. On vérifie ainsi le théorème de Ore sur cet exemple particulier.

2 Résultants et sous-résultants

Dans le cadre des polynômes usuels, les résultants et les sous-résultants sont un outil classique qui offre une vision purement algébrique et entièrement mécanique pour étudier la relative primalité et la divisibilité des polynômes [2, §4.2], [7], [25, §4.1]. Ce point de vue est particulièrement intéressant car il est valable sur n'importe anneau de base et se comporte bien par spécialisation. Il est ainsi, par exemple, particulièrement adapté pour étudier le comportement de polynômes dont les coefficients dépendent d'un ou plusieurs paramètres.

Vers le début des années 1990, Berkovich–Tsirulik [3] et Chardin [5] ont remarqué que la théorie des résultants et sous-résultants pouvait s'étendre à l'anneau des opérateurs différentiels $k(t)[X, \frac{d}{dt}]$. Quelques années plus tard, Li [20, 21] a mis au point une généralisation complète des résultants et sous-résultants à tout anneau de Ore. Dans cette partie, nous exposons le travail de Li, tout en l'illustrant de nombreux exemples. Nous optons toutefois pour une présentation légèrement différente de ce que l'on trouve habituellement : au lieu de travailler avec des matrices, nous faisons la part belle aux applications linéaires et aux constructions algébriques, rendant ainsi notre présentation abstraite mais aussi, nous l'espérons, plus concise et efficace sur certains points.

Dans toute cette partie, on considère un anneau commutatif $\mathfrak{A}$ sur lequel on ne fait *a priori* aucune hypothèse supplémentaire. On fixe un endomorphisme d'anneaux $\theta : \mathfrak{A} \to \mathfrak{A}$ ainsi qu'une θ-dérivation $\partial : \mathfrak{A} \to \mathfrak{A}$, de manière à pouvoir considérer l'anneau de Ore $\mathfrak{A}[X, \theta, \partial]$.

On note $\mathcal{P}_n$ l'ensemble des polynômes de Ore de $\mathfrak{A}[X, \theta, \partial]$ à coefficients dans $\mathfrak{A}$ de degré strictement inférieur à n, que l'on voit comme un $\mathfrak{A}$-module à gauche. Clairement, $\mathcal{P}_n$ est libre de rang n ; on le munit de sa base canonique $(1, X, \dots, X^{n-1})$.

2.1 Définitions et propriétés du résultant

Soient $A, B \in \mathfrak{A}[X, \theta, \partial]$ deux polynômes de Ore et soient a et b deux entiers tels que $a \geq \deg A$ et $b \geq \deg B$. À ces données, on associe l'application de Sylvester :

$$\begin{array}{rccc} \mathrm{Syl}^{a,b}(A,B) : & \mathcal{P}_b \times \mathcal{P}_a & \longrightarrow & \mathcal{P}_{a+b} \\ & (U, V) & \mapsto & UA + VB. \end{array}$$

Il est immédiat de vérifier que $\mathrm{Syl}^{a,b}(A, B)$ est $\mathfrak{A}$-linéaire. On remarque également que les espaces de départ et d'arrivée de $\mathrm{Syl}^{a,b}(A, B)$ sont tous les deux des $\mathfrak{A}$-modules libres de rang $a+b$ et que ceux-ci sont munis de bases canoniques, à savoir

$$\begin{array}{rl} \big((1,0),(X,0),\dots,(X^{b-1},0),(0,1),(0,X),\dots,(0,X^{a-1})\big) & \text{pour } \mathcal{P}_b \times \mathcal{P}_a \\ \text{et} \quad \big(1, X, \dots, X^{a+b-1}\big) & \text{pour } \mathcal{P}_{a+b}. \end{array}$$

Définition 2.1.1. Le *résultant* des polynômes A et B, calculé en degrés a et b, est le déterminant de $\mathrm{Syl}^{a,b}(A, B)$ calculé dans les bases canoniques de $\mathcal{P}_b \times \mathcal{P}_a$ et $\mathcal{P}_{a+b}$.
On le note $\mathrm{Res}^{a,b}(A, B)$.

Exemple 2.1.2. Sur l'anneau de Ore $\mathbb{Z}[t][X, \frac{d}{dt}]$, considérons les deux polynômes de Ore suivants :

$$\begin{aligned} A &= X^3 + (t^2 + 3t)X^2 + (5t^3 + 4t)X + (2t^5 + 6t^2 + 2) \\ B &= X^2 + (t^2 + 4t)X + (4t^3 + 2t). \end{aligned}$$

L'application de Sylvester associée (en degrés $a = \deg A = 3$ et $b = \deg B = 2$) est, par définition, l'application linéaire :

$$\begin{array}{rccc} \mathrm{Syl}(A,B) : & \mathcal{P}_2 \times \mathcal{P}_3 & \longrightarrow & \mathcal{P}_5 \\ & (U, V) & \mapsto & UA + VB. \end{array}$$

Sa matrice dans les bases canoniques est donc une matrice carrée de taille 5 dont les colonnes successive contiennent les coefficients de A, XA, B, XB et X^2B. Un calcul aboutit ainsi à :

$$\mathrm{Syl}(A,B) = \begin{pmatrix} 2t^5+6t^2+2 & 10t^4+12t & 4t^3+2t & 12t^2+2 & 24t \\ 5t^3+4t & 2t^5+21t^2+6 & t^2+4t & 4t^3+4t+4 & 24t^2+6 \\ t^2+3t & 5t^3+6t+3 & 1 & t^2+4t & 4t^3+6t+8 \\ 1 & t^2+3t & 0 & 1 & t^2+4t \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \tag{14}$$

où, dans l'écriture précédente, dans un léger abus, on a identifié l'application $\mathrm{Syl}(A,B)$ à sa matrice. Le résultant de A et B est, par définition, le déterminant de $\mathrm{Syl}(A,B)$. Un calcul montre que celui-ci s'annule : on a $\mathrm{Res}(A,B)=0$.

Comme nous le verrons dans la suite (*cf* théorème 2.1.7), ceci indique que les polynômes A et B ne sont pas premiers entre eux, au moins lorsque ceux-ci sont vus comme des polynômes de Ore à coefficients dans le corps des fractions de $\mathbb{Z}[t]$, à savoir $\mathbb{Q}(t)$.

Dans la suite, lorsque $a=\deg A$ et $b=\deg B$, nous omettrons l'exposant a,b dans l'écriture ; autrement dit, on écrira simplement $\mathrm{Syl}(A,B)$ pour $\mathrm{Syl}^{\deg A,\deg B}(A,B)$ et, de même, $\mathrm{Res}(A,B)$ pour $\mathrm{Res}^{\deg A,\deg B}(A,B)$.

Comme dans le cas commutatif, il existe de nombreuses formules reliant les valeurs des résultants lorsque l'on fait varier les polynômes ou les entiers a et b. La proposition 2.1.3 ci-après en donne un échantillon, certainement pas exhaustif. Avant de pouvoir l'énoncer au mieux, introduisons deux notations supplémentaires. Premièrement, si $P\in\mathfrak{A}[X,\theta,\partial]$ est un polynôme de Ore et si n est un entier, désignons par $P[n]$ le coefficient de P en X^n. Deuxièmement, pour deux entiers $i\leq j$ et pour un scalaire $\lambda\in\mathfrak{A}$, définissons :

$$\Theta_i^j(\lambda) = \theta^i(\lambda)\cdot\theta^{i+1}(\lambda)\cdots\theta^{i+j-1}(\lambda) \tag{15}$$

avec la convention que $\Theta_i^j(\lambda)=1$ lorsque $i=j$.

Proposition 2.1.3. *Pour des polynômes de Ore $A,B\in\mathfrak{A}[X,\theta,\partial]$, des scalaires $\alpha,\beta\in\mathfrak{A}$ et des entiers a,b,n avec $a\geq\deg A$, $b\geq B$ et $n\geq 0$, on a :*

1. $\mathit{Res}^{b,a}(B,A) = (-1)^{ab}\cdot\mathit{Res}^{a,b}(A,B)$
2. $\mathit{Res}^{a,b}(\alpha A,\beta B) = \Theta_0^b(\alpha)\,\Theta_0^a(\beta)\cdot\mathit{Res}^{a,b}(A,B)$
3. $\mathit{Res}^{a+n,b}(A,B) = \Theta_a^{a+n}(B[b])\cdot\mathit{Res}^{a,b}(A,B)$
4. $\mathit{Res}^{a,b+n}(A,B) = (-1)^{an}\,\Theta_b^{b+n}(A[a])\cdot\mathit{Res}^{a,b}(A,B)$

Démonstration. **1.** On constate que $\mathrm{Syl}^{b,a}(B,A) = \mathrm{Syl}^{a,b}(A,B)\circ\tau$ où $\tau:\mathcal{P}_b\times\mathcal{P}_a\to\mathcal{P}_a\times\mathcal{P}_b$ est l'application qui envoie le couple (U,V) sur (V,U). En prenant les déterminants, on en déduit que $\mathrm{Res}^{b,a}(B,A) = (\det\tau)\cdot\mathrm{Res}^{a,b}(A,B)$. Or, manifestement, dans les bases canoniques $\mathcal{P}_b\times\mathcal{P}_a$ et $\mathcal{P}_a\times\mathcal{P}_b$, l'application τ est une permutation qui a exactement ab inversions. Son déterminant est donc $(-1)^{ab}$.

2. On raisonne de manière similaire en écrivant $\mathrm{Syl}^{a,b}(\alpha A,\beta B)$ comme la composée de deux applications, à savoir $\mathrm{Syl}^{a,b}(\alpha A,\beta B) = \mathrm{Syl}^{a,b}(A,B)\circ\mu$ où $\mu:\mathcal{P}_b\times\mathcal{P}_a\to\mathcal{P}_b\times\mathcal{P}_a$ envoie (U,V) sur $(U\alpha,V\beta)$. Dans la base canonique de $\mathcal{P}_b\times\mathcal{P}_a$, la matrice de μ est triangulaire supérieure et ses élements diagonaux sont, dans l'ordre α, $\theta(\alpha)$, …, $\theta^{b-1}(\alpha)$, β, $\theta(\beta)$, …, $\theta^{a-1}(\beta)$. Son déterminant vaut donc $\Theta_0^b(\alpha)\,\Theta_0^a(\beta)$ et la formule annoncée s'en déduit.

3. On dispose du diagramme commutatif suivant :

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{P}_b\times\mathcal{P}_a & \longrightarrow & \mathcal{P}_b\times\mathcal{P}_{a+n} & \longrightarrow & \mathcal{P}_{a+n}/\mathcal{P}_a & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \mathrm{Syl}^{a,b}(A,B)} & & \downarrow{\scriptstyle \mathrm{Syl}^{a+n,b}(A,B)} & & \downarrow{\scriptstyle \mu:V\mapsto VB} & & \\ 0 & \longrightarrow & \mathcal{P}_{a+b} & \longrightarrow & \mathcal{P}_{a+n+b} & \longrightarrow & \mathcal{P}_{a+n+b}/\mathcal{P}_{a+b} & \longrightarrow & 0 \end{array}$$

dont les lignes sont des suites exactes. Au niveau des matrices, ce diagramme nous apprend que la matrice $\mathrm{Syl}^{a+n,b}(A,B)$ est une matrice triangulaire par blocs dont les blocs diagonaux sont, d'une part, la matrice de $\mathrm{Syl}^{a,b}(A,B)$ et, d'autre part, la matrice de l'application μ écrite dans les bases $(X^a,\dots,X^{a+n+1})$ et $(X^{a+b},\dots,X^{a+b+n-1})$ respectivement. Ainsi trouve-t-on l'égalité :

$$\mathrm{Res}^{a+n,b}(A,B) = \det(\mu)\cdot \mathrm{Res}^{a,b}(A,B).$$

Par ailleurs, la matrice de μ est, elle-même, triangulaire supérieure avec pour coefficients diagonaux $\theta^a(B[b]),\ \dots,\ \theta^{a+n-1}(B[b]))$. Elle a donc pour déterminant $\Theta_a^{a+n}(B[b])$, ce qui conclut.
4. se déduit en combinant **1** et **3**. □

Corollaire 2.1.4. *Le résultant $\mathrm{Res}^{a,b}(A,B)$ s'annule dès lors que $a > \deg A$ et $b > \deg B$.*

Démonstration. Par la proposition 2.1.3, on a l'identité :

$$\mathrm{Res}^{a,b}(A,B) = \Theta^a_{\deg A}(B[b])\cdot \mathrm{Res}^{\deg A,b}(A,B).$$

De plus, le coefficient multiplicatif $\Theta^a_{\deg A}(B[b])$ s'annule étant donné que $B[b]=0$ et que $a>\deg A$. On en déduit que $\mathrm{Res}^{a,b}(A,B)=0$. □

Une autre propriété intéressante du résultant est sa fonctorialité qui précise son comportement par changement d'anneau. Précisément, on se donne un second anneau $\mathfrak{A}'$ muni d'un endomorphisme $\theta':\mathfrak{A}'\to\mathfrak{A}'$ et d'une θ'-dérivation ∂' et on suppose en outre que l'on dispose d'un morphisme d'anneaux $f:\mathfrak{A}\to\mathfrak{A}'$ commutant aux structures supplémentaires, c'est-à-dire satisfaisant aux axiomes $f\circ\theta=\theta'\circ f$ et $f\circ\partial=\partial'\circ f$. Sous ces hypothèses, f induit un morphisme $f_\star:\mathfrak{A}[X,\theta,\partial]\to\mathfrak{A}'[X,\theta',\partial']$ au niveau des anneaux de Ore.

Proposition 2.1.5. *Avec les notations précédentes, pour deux polynômes de Ore $A,B\in\mathfrak{A}[X,\theta,\partial]$ et deux entiers $a\geq\deg A$ et $b\geq\deg B$, on a :*

$$\mathit{Res}^{a,b}(f_\star A, f_\star B) = f\big(\mathit{Res}^{a,b}(A,B)\big).$$

Démonstration. Pour un entier n, notons $\mathcal{P}'_n$ le $\mathfrak{A}'$-module des polynômes de Ore à coefficients dans $\mathfrak{A}'$ de degré strictement inférieur à n. C'est un module libre de rang n. De plus, $\mathcal{P}'_n=\mathfrak{A}'\otimes_{\mathfrak{A}}\mathcal{P}_n$ et cette identification met en correspondance les bases canoniques de $\mathcal{P}'_n$, d'une part, et de $\mathcal{P}_n$, d'autre part. En outre, on a l'identité $\mathrm{Syl}^{a,b}(f_\star A,f_\star B)=f\otimes\mathrm{Syl}^{a,b}(A,B)$, à partir de quoi on déduit la proposition en prenant les déterminants. □

Remarque 2.1.6. Si f est injective, on déduit de la proposition 2.1.5 que $\mathrm{Res}(f_\star A,f_\star B)=f\big(\mathrm{Res}(A,B)\big)$ pour tous polynômes de Ore $A,B\in\mathfrak{A}[X,\theta,\partial]$. Ceci n'est, par contre, pas vrai en toute généralité étant donné que les coefficients dominants de A (resp. de B) peuvent être annulés par f, faisant ainsi chuter le degré de $f_\star B$ (resp. de $f_\star B$).

Enfin, lorsque l'anneau de base $\mathfrak{A}$ est un corps, le résultant permet de détecter la relative primalité de deux polynômes de Ore. Plus précisément, on a le résultat suivant.

Théorème 2.1.7. *On suppose que $\mathfrak{A}$ est un corps et on le note K. Pour $A,B\in K[X,\theta,\partial]$, on a l'équivalence suivante :* $\textsc{rgcd}(A,B)=1$ *si et seulement si* $\mathit{Res}(A,B)\neq 0$.

Démonstration. Si le résultant de A et B ne s'annule pas, l'application de Sylvester $\mathrm{Syl}(A,B)$ est un isomorphisme. Il en résulte qu'il existe des polynômes de Ore U et V tels que $UA+VB=1$. On en déduit que l'idéal à gauche de $K[X,\theta,\partial]$ engendré par A et B est $K[X,\theta,\partial]$ tout entier, c'est-à-dire que $\textsc{rgcd}(A,B)=1$.

Réciproquement, supposons que $\textsc{rgcd}(A,B)=1$. Nous allons montrer que $\mathrm{Syl}(A,B)$ est injective. Considérons pour cela un couple $(U,V)\in\mathcal{P}_{\deg B}\times\mathcal{P}_{\deg A}$ dans son noyau. On a alors

$UA = -VB$. Notons P cette quantité commune. Il s'agit d'un multiple à gauche commun à A et B. On en déduit que si $P \neq 0$, on doit avoir $\deg P \geq \deg \mathrm{LLCM}(A,B)$ et, par suite :

$$\deg U \geq \deg \mathrm{LLCM}(A,B) - \deg A \quad \text{et} \quad \deg V \geq \deg \mathrm{LLCM}(A,B) - \deg B.$$

Comme A et B sont supposés premiers entre eux, il résulte de la proposition 1.3.13 que $\deg \mathrm{LLCM}(A,B) = \deg A + \deg B$. On en déduit que $\deg U \geq \deg A$ et $\deg V \geq \deg B$. Ceci est une contradiction, de laquelle on déduit que $P = 0$. Il en résulte que $U = V = 0$ et donc que $\mathrm{Syl}(A,B)$ est injective. □

2.2 Cofacteurs et coefficients de Bézout

Nous venons de voir que le résultant permet de détecter la relative primalité des polynômes de Ore. Il se trouve, de surcroît qu'il est également possible de retrouver les coefficients de Bézout par l'intermédiaire de constructions algébriques analogues à celles que nous avons menées pour construire le résultant. L'outil essentiel permettant de mener à terme ces constructions est la notion d'adjoint d'une application linéaire (ou, au choix, d'une matrice). Rappelons ici rapidement que l'adjoint d'une matrice carrée M, noté $\mathrm{adj}(M)$ est la transposée de la comatrice de M. Il vérifie la relation fondamentale :

$$M \cdot \mathrm{adj}(M) = \mathrm{adj}(M) \cdot M = \det(M) I$$

où I est la matrice identité de taille adéquate. Si E et F sont des $\mathfrak{A}$-modules libres de même rang, chacun muni d'une base, et si $f : E \to F$ est une application linéaire, l'*adjoint* de f est, par définition, l'application $\mathrm{adj}(f) : F \to E$ dont la matrice (dans les bases choisies) est l'adjoint de la matrice de f (à nouveau dans les bases choisies). Pour de nombreux compléments sur cette notion, nous renvoyons la lectrice à l'appendice A.1.

Dans le cas de l'application de Sylvester, l'application adjointe permet de retrouver les coefficients de Bézout. Précisément, étant donnés comme précédemment des polynômes de Ore $A, B \in \mathfrak{A}[X, \theta, \partial]$ ainsi que deux entiers a, b avec $a \geq \deg A$ et $b \geq \deg B$, on définit les polynômes $U^{a,b}(A,B)$ et $V^{a,b}(A,B)$ par l'égalité :

$$\big(U^{a,b}(A,B),\, V^{a,b}(A,B)\big) = \mathrm{adj}(\mathrm{Syl}^{a,b}(A,B))(1).$$

On les appelle les *cofacteurs* de A et B. Lorsque $a = \deg A$ et $b = \deg B$, on omettra souvent l'exposant a, b dans l'écriture ; autrement dit, on écrira simplement $U(A,B)$ et $V(A,B)$ à la place de $U^{\deg A, \deg B}(A,B)$ et $V^{\deg A, \deg B}(A,B)$.

Il suit de la proposition A.1.2 que $\mathrm{Syl}^{a,b}(A,B)\big(U^{a,b}(A,B),\, V^{a,b}(A,B)\big) = \det\big(\mathrm{Syl}^{a,b}(A,B)\big)$, c'est-à-dire concrètement que :

$$U^{a,b}(A,B) \cdot A + V^{a,b}(A,B) \cdot B = \mathrm{Res}^{a,b}(A,B). \tag{16}$$

Sur un corps, lorsque $\mathrm{Res}^{a,b}(A,B)$ ne s'annule pas, on retrouve ainsi les coefficients de Bézout. L'approche précédente a l'avantage de s'étendre à un anneau de base quelconque. Par ailleurs, de même que pour les résultants, la construction des cofacteurs est fonctorielle dans le sens où celle-ci se comporte bien vis-à-vis des changements d'anneaux.

Exemple 2.2.1. Reprenons les polynômes de Ore de l'exemple (2.1.2), à savoir :

$$A = X^3 + (t^2 + 3t)X^2 + (5t^3 + 4t)X + (2t^5 + 6t^2 + 2)$$
$$B = X^2 + (t^2 + 4t)X + (4t^3 + 2t).$$

Nous avions déjà vu que $\mathrm{Res}(A,B) = 0$. Calculons maintenant les cofacteurs correspondants. Par définition ceux-ci se lisent sur la première colonne de l'adjoint de la matrice $\mathrm{Syl}(A,B)$ donnée

par la formule (14). Un calcul aboutit à :

$$\mathrm{adj}(\mathrm{Syl}(A,B)) = \begin{pmatrix} -8t^4-16t^3+6t^2+24t & \cdots & \cdots & \cdots & \cdots \\ -2t^3-4t^2+4 & \cdots & \cdots & \cdots & \cdots \\ 4t^6+8t^5-12t^3-12t^2+20 & \cdots & \cdots & \cdots & \cdots \\ 6t^4+12t^3-6t^2-20t & \cdots & \cdots & \cdots & \cdots \\ 2t^3+4t^2-4 & \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

d'où on déduit que :

$$\begin{aligned} U(A,B) &= \big(-2t^3-4t^2+4\big)X + \big(-8t^4-16t^3+6t^2+24t\big) \\ V(A,B) &= \big(2t^3+4t^2-4\big)X^2 + \big(6t^4+12t^3-6t^2-20t\big)X + \big(4t^6+8t^5-12t^3-12t^2+20\big). \end{aligned}$$

et on peut vérifier, si on le souhaite, la relation :

$$U(A,B)\cdot A + V(A,B)\cdot B = 0.$$

Remarquons que l'égalité ci-dessus indique que $L = U(A,B)\cdot A = -V(A,B)\cdot B$ est un multiple à gauche commun à A et B et est donc également d'un multiple à gauche de $\mathrm{LLCM}(A,B)$ (au moins si l'on prend soin de voir A et B comme des polynômes de Ore à coefficients dans $\mathbb{Q}(t)$ pour que la notion de LLCM soit bien définie). Sur cet exemple particulier, on peut vérifier à la main que L est, à une constante multiplicative près, égal à $\mathrm{LLCM}(A,B)$. Cette coïncidence n'en est, en réalité, pas une car on peut démontrer de manière générale que si le RGCD de deux polynômes de Ore A et B est de degré 1 alors $\mathrm{LLCM}(A,B)$ est proportionnel à $U(A,B)\cdot A = -V(A,B)\cdot B$ (*cf* théorème 2.3.5, page 26).

On dispose d'un formulaire pour les cofacteurs analogue à celui que nous avions énoncé dans la proposition 2.1.3 pour les résultants.

Proposition 2.2.2. *Pour des polynômes de Ore $A, B \in \mathfrak{A}[X,\theta,\partial]$, des scalaires $\alpha,\beta \in \mathfrak{A}$ et des entiers a, b, n avec $a \geq \deg A$, $b \geq B$ et $n \geq 0$, on a :*

1. $U^{b,a}(B,A) = (-1)^{ab}\cdot V^{a,b}(A,B)$
   $V^{b,a}(B,A) = (-1)^{ab}\cdot U^{a,b}(A,B)$
2. $U^{a,b}(\alpha A,\beta B)\cdot\alpha = \Theta_0^b(\alpha)\,\Theta_0^a(\beta)\cdot U^{a,b}(A,B)$
   $V^{a,b}(\alpha A,\beta B)\cdot\beta = \Theta_0^b(\alpha)\,\Theta_0^a(\beta)\cdot V^{a,b}(A,B)$
3. $U^{a+n,b}(A,B) = \Theta_a^{a+n}(B[b])\cdot U^{a,b}(A,B)$
   $U^{a+n,b}(A,B) = \Theta_a^{a+n}(B[b])\cdot U^{a,b}(A,B)$
4. $U^{a,b+n}(A,B) = (-1)^{an}\,\Theta_b^{b+n}(A[a])\cdot U^{a,b}(A,B)$
   $V^{a,b+n}(A,B) = (-1)^{an}\,\Theta_b^{b+n}(A[a])\cdot V^{a,b}(A,B)$

Démonstration. **1.** On rappelle qu'on a l'égalité $\mathrm{Syl}^{b,a}(B,A) = \mathrm{Syl}^{a,b}(A,B)\circ\tau$ où $\tau : \mathcal{P}_b\times\mathcal{P}_a \to \mathcal{P}_a\times\mathcal{P}_b$ est l'application qui envoie le couple (U,V) sur (V,U). On en déduit que :

$$\mathrm{adj}\big(\mathrm{Syl}^{b,a}(B,A)\big) = \mathrm{adj}(\tau)\circ\mathrm{adj}\big(\mathrm{Syl}^{a,b}(A,B)\big).$$

De plus, comme τ est bijectif on a $\mathrm{adj}(\tau) = \det(\tau)\,\tau^{-1} = (-1)^{ab}\tau$ (*cf* proposition 2.1.3 pour le calcul du déterminant de τ). Les égalités annoncées s'ensuivent.

2. De même que précédemment, on écrit $\mathrm{Syl}^{a,b}(\alpha A,\beta B)$ comme la composée $\mathrm{Syl}^{a,b}(A,B)\circ\mu$ où $\mu : \mathcal{P}_b\times\mathcal{P}_a \to \mathcal{P}_b\times\mathcal{P}_a$ envoie (U,V) sur $(U\alpha, V\beta)$. Ainsi, trouve-t-on :

$$\mathrm{adj}\big(\mathrm{Syl}^{a,b}(\alpha A,\beta B)\big) = \mathrm{adj}(\mu)\circ\mathrm{adj}\big(\mathrm{Syl}^{a,b}(A,B)\big). \tag{17}$$

D'après le calcul de la proposition 2.1.3, on sait que $\det(\mu) = \Theta_0^b(\alpha)\Theta_0^a(\beta)$. On déduit ainsi de la proposition A.1.2 que $\mu \circ \mathrm{adj}(\mu) = \Theta_0^b(\alpha)\Theta_0^a(\beta)$. En combinant avec (17), on obtient :

$$\mu \circ \mathrm{adj}\big(\mathrm{Syl}^{a,b}(\alpha A, \beta B)\big) = \Theta_0^b(\alpha)\Theta_0^a(\beta) \cdot \mathrm{adj}\big(\mathrm{Syl}^{a,b}(A, B)\big)$$

ce qui entraîne le résultat voulu en regardant la valeur en 1.

3. Reprenons le diagramme commutatif à lignes exactes introduit dans la proposition 2.1.3 :

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \mathcal{P}_b \times \mathcal{P}_a & \longrightarrow & \mathcal{P}_b \times \mathcal{P}_{a+n} & \longrightarrow & \mathcal{P}_{a+n}/\mathcal{P}_a & \longrightarrow & 0 \\
 & & \downarrow{\scriptstyle \mathrm{Syl}^{a,b}(A,B)} & & \downarrow{\scriptstyle \mathrm{Syl}^{a+n,b}(A,B)} & & \downarrow{\scriptstyle \mu:V\mapsto VB} & & \\
0 & \longrightarrow & \mathcal{P}_{a+b} & \longrightarrow & \mathcal{P}_{a+n+b} & \longrightarrow & \mathcal{P}_{a+n+b}/\mathcal{P}_{a+b} & \longrightarrow & 0.
\end{array}$$

On en déduit, par la proposition A.1.4, que :

$$\mathrm{adj}\big(\mathrm{Syl}^{a+n,b}(A, B)\big) = \det(\mu) \cdot \mathrm{adj}\big(\mathrm{Syl}^{a,b}(A, B)\big).$$

Par ailleurs, le calcul du déterminant de μ a déjà été mené dans la démonstration de la proposition 2.1.3, il vaut $\Theta_a^{a+n}(B[b])$. Le résultat annoncé s'en déduit.

4. se déduit de **1** et **3**. □

2.3 La théorie des sous-résultants

Dans le cas où les polynômes A et B ne sont pas premiers entre eux, le résultant $\mathrm{Res}(A, B)$ s'annule et l'égalité (16) ne nous permet pas de déterminer le RGCD de A et de B. Entre autres choses, la théorie des sous-résultants complète de manière satisfaisante la théorie des résultants sur ce point.

On considère comme précédemment $A, B \in \mathfrak{A}[X, \theta, \partial]$ deux polynômes de Ore ainsi que des entiers $a \geq \deg A$ et $b \geq \deg B$ mais on donne à présent, en plus, un entier $m \in \{0, 1, \ldots, \min(a, b)\}$. Lorsque $a = b$, on suppose de plus $m < a = b$. On considère l'application de Sylvester tronquée :

$$\begin{array}{rccc}
\mathrm{Syl}_m^{a,b}(A, B) : & \mathcal{P}_{b-m} \times \mathcal{P}_{a-m} & \longrightarrow & \mathcal{P}_{a+b-m}/\mathcal{P}_m \\
 & (U, V) & \mapsto & UA + VB
\end{array}$$

où le quotient $\mathcal{P}_{a+b-m}/\mathcal{P}_m$ est muni de la base $(X^m, X^{m+1}, \ldots, X^{a+b-m-1})$. On définit également $\mathrm{Syl}_m^{a,b}(A, B)$ pour $m = -1$ par :

$$\begin{array}{rccc}
\mathrm{Syl}_{-1}^{a,b}(A, B) : & \mathcal{P}_{b+1} \times \mathcal{P}_{a+1} & \longrightarrow & \mathfrak{A}X^{-1} \oplus \mathcal{P}_{a+b+1} \\
 & (U, V) & \mapsto & 0 \oplus (UA + VB)
\end{array}$$

où X^{-1} est une nouvelle variable formelle. On munit en outre le $\mathfrak{A}$-module libre $\mathfrak{A}X^{-1} \oplus \mathcal{P}_{a+b+1}$ de sa base canonique $(X^{-1}, 1, X, \ldots, X^{a+b})$.

Définition 2.3.1. Avec les notations précédentes :

– le m-ième *sous-résultant scalaire* de A et B en degrés a et b est :

$$\mathrm{sres}_m^{a,b}(A, B) = \det\big(\mathrm{Syl}_m^{a,b}(A, B)\big).$$

– les m-ièmes *cofacteurs* de A et B en degrés a et b sont :

$$\big(U_m^{a,b}(A, B),\ V_m^{a,b}(A, B)\big) = \mathrm{adj}\big(\mathrm{Syl}_m^{a,b}(A, B)\big)(X^m).$$

– le m-ième *sous-résultant* de A et B en degrés a et b est :

$$\mathrm{sRes}_m^{a,b}(A, B) = U_m^{a,b}(A, B) \cdot A + V_m^{a,b}(A, B) \cdot B.$$

Comme dans le cas des résultants, lorsque $a = \deg A$ et $b = \deg B$, on s'autorise à ne pas écrire l'exposant a, b, c'est-à-dire à écrire simplement $\mathrm{sres}_m(A,B)$, $\mathrm{sRes}_m(A,B)$, $U_m(A,B)$ et $V_m(A,B)$ à la place respectivement de $\mathrm{sres}_m^{\deg A,\deg B}(A,B)$, $\mathrm{sRes}_m^{\deg A,\deg B}(A,B)$, $U_m^{\deg A,\deg B}(A,B)$ et $V_m^{\deg A,\deg B}(A,B)$. Il résulte de la proposition A.1.2 que $\mathrm{sRes}_m^{a,b}(A,B) \equiv \mathrm{sres}_m^{a,b}(A,B) \cdot X^m \pmod{\mathcal{P}_m}$. Autrement dit, $\mathrm{sRes}_m^{a,b}(A,B)$ est un polynôme de Ore de degré au plus m dont le coefficient en X^m est $\mathrm{sres}_m^{a,b}(A,B)$. Pour $m = 0$, on observe que résultants et sous-résultants coïncident, à savoir précisément :

$$\begin{gathered}\mathrm{sres}_0^{a,b}(A,B) = \mathrm{sRes}_0^{a,b}(A,B) = \mathrm{Res}^{a,b}(A,B),\\ U_0^{a,b}(A,B) = U^{a,b}(A,B) \quad \text{et} \quad V_0^{a,b}(A,B) = V^{a,b}(A,B)\end{gathered}$$

pour tous polynômes de Ore A et B et tous entiers $a \geq \deg A$ et $b \geq \deg B$.

Par ailleurs, de même que pour les résultants et les cofacteurs, les constructions des m-ièmes sous-résultants et m-ièmes cofacteurs sont fonctorielles dans le sens où elles se comportent bien par rapport aux changements d'anneaux.

Exemple 2.3.2. Reprenons à nouveau les polynômes de Ore déjà considérés dans les exemples 2.1.2 et 2.2.1 :

$$\begin{aligned}A &= X^3 + (t^2+3t)X^2 + (5t^3+4t)X + (2t^5+6t^2+2)\\ B &= X^2 + (t^2+4t)X + (4t^3+2t)\end{aligned}$$

et calculons leurs sous-résultants. Il résulte de la définition que la matrice de Sylvester tronquée $\mathrm{Syl}_1(A,B)$ est la sous-matrice de $\mathrm{Syl}(A,B)$ obtenue en retirant la première et la dernière ligne ainsi que la deuxième et la dernière colonne. Concrètement :

$$\mathrm{Syl}_1(A,B) = \begin{pmatrix} 5t^3+4t & t^2+4t & 4t^3+4t+4\\ t^2+3t & 1 & t^2+4t\\ 1 & 0 & 1\end{pmatrix}.$$

Le premier sous-résultant scalaire de A et B est le déterminant de cette matrice qui vaut $2t^3+4t^2-4$. Les premiers cofacteurs s'obtiennent à nouveau en regardant la première colonne de la matrice adjointe $\mathrm{adj}(\mathrm{Syl}_1(A,B))$. Un calcul donne $U_1(A,B) = 1$ et $V_1(A,B) = -X + t$, à partir de quoi on déduit la valeur du premier sous-résultant :

$$\begin{aligned}\mathrm{sRes}_1(A,B) &= U_1(A,B)\cdot A + V_1(A,B)\cdot B = A + (-X+t)B\\ &= \big(2t^3+4t^2-4\big)X + \big(2t^5+4t^3-4t^2\big).\end{aligned}$$

Passons à présent au second sous-résultant. La matrice de Sylvester $\mathrm{Syl}_2(A,B)$ est la matrice carrée de taille 1 dont l'unique coefficient est 1. On obtient donc directement $\mathrm{sres}_2(A,B) = 1$, $U_2(A,B) = 0$, $V_2(A,B) = 1$ puis, par suite, $\mathrm{sRes}_2(A,B) = B$.

Plus généralement, le m-ième sous-résultant de A et B admet une formule explicite lorsque $m = \min(a,b)$ comme le précise le lemme suivant.

Lemme 2.3.3. *Pour des polynômes de Ore $A, B \in \mathfrak{A}[X,\theta,\partial]$ et des entiers a, b avec $a \geq \deg A$, $b \geq \deg B$ et $a > b$, on a :*

$$\begin{aligned}\mathit{sres}_b^{a,b}(A,B) &= \Theta_0^{a-b}(B[b])\\ U_b^{a,b}(A,B) &= 0\\ V_b^{a,b}(A,B) &= \Theta_1^{a-b}(B[b])\\ \mathit{sRes}_b^{a,b}(A,B) &= \Theta_1^{a-b}(B[b])\cdot B\end{aligned}$$

où on rappelle que $B[b]$ désigne le coefficient de B devant X^b et que $\Theta_i^j(\alpha)$ est défini par la formule (15).

Démonstration. Revenant aux définitions, on s'aperçoit que l'application de Sylvester $\mathrm{Syl}_b(A,B)$ est la fonction $\mathcal{P}_{a-b} \to \mathcal{P}_a/\mathcal{P}_b$, $V \mapsto VB$. Dans les bases canoniques, sa matrice est ainsi triangulaire supérieure et ses coefficients diagonaux sont successivement $B[b], \theta(B[b]), \dots, \theta^{a-b-1}(B[b])$. Son déterminant est donc égal à $\Theta_0^{a-b}(B[b])$, ce qui démontre la formule pour le sous-résultant scalaire.

Étant donné que $\mathrm{Syl}_b(A,B)$ est triangulaire supérieure, il en est de même de son adjoint. De plus, son coefficient en haut à gauche est égal au produit des coefficients diagonaux de $\mathrm{Syl}_b(A,B)$ excepté le premier. Il vaut donc $\Theta_1^{a-b}(B[b])$, à partir de quoi on déduit que les formules annoncées pour les cofacteurs. Enfin, la dernière formule pour le sous-résultant s'en déduit en appliquant la définition. □

À nouveau, on dispose d'un formulaire analogue à celui de la proposition 2.1.3 dans le cas des sous-résultants.

Proposition 2.3.4. *Soient $A, B \in \mathfrak{A}[X,\theta,\partial]$, $\alpha, \beta \in \mathfrak{A}$ et des entiers a, b, m, n avec $a \geq \deg A$, $b \geq B$, $-1 \leq m \leq \min(a,b)$ et $n \geq 0$. De plus, si $a = b$, on suppose $j < a$. On a alors le formulaire suivant :*

1. $\mathit{sRes}_m^{b,a}(B,A) = (-1)^{(a-m)(b-m)} \cdot \mathit{sRes}_m^{a,b}(A,B)$
   $U_m^{b,a}(B,A) = (-1)^{(a-m)b} \cdot V_m^{a,b}(A,B)$
   $V_m^{b,a}(B,A) = (-1)^{(a-m)b} \cdot U_m^{a,b}(A,B)$
2. $\mathit{sRes}_m^{a,b}(\alpha A, \beta B) = \Theta_0^{b-m}(\alpha)\,\Theta_0^{a-m}(\beta) \cdot \mathit{sRes}_m^{a,b}(A,B)$
   $U_m^{a,b}(\alpha A, \beta B) \cdot \alpha = \Theta_0^{b-m}(\alpha)\,\Theta_0^{a-m}(\beta) \cdot U_m^{a,b}(A,B)$
   $V_m^{a,b}(\alpha A, \beta B) \cdot \beta = \Theta_0^{b-m}(\alpha)\,\Theta_0^{a-m}(\beta) \cdot V_m^{a,b}(A,B)$
3. $\mathit{sRes}_m^{a+n,b}(A,B) = \Theta_{a-m}^{a-m+n}(B[b]) \cdot \mathit{sRes}_m^{a,b}(A,B)$
   $U_m^{a+n,b}(A,B) = \Theta_{a-m}^{a-m+n}(B[b]) \cdot U_m^{a,b}(A,B)$
   $V_m^{a+n,b}(A,B) = \Theta_{a-m}^{a-m+n}(B[b]) \cdot V_m^{a,b}(A,B)$
4. $\mathit{sRes}_m^{a,b+n}(A,B) = (-1)^{(a-m)n}\,\Theta_{b-m}^{b-m+n}(A[a]) \cdot \mathit{sRes}_m^{a,b}(A,B)$
   $U_m^{a,b+n}(A,B) = (-1)^{(a-m)n}\,\Theta_{b-m}^{b-m+n}(A[a]) \cdot U_m^{a,b}(A,B)$
   $V_m^{a,b+n}(A,B) = (-1)^{(a-m)n}\,\Theta_{b-m}^{b-m+n}(A[a]) \cdot V_m^{a,b}(A,B)$

Esquisse de la démonstration. La démonstration est très proche de ce que nous avons déjà fait précédemment. Pour cette raison, nous n'en donnons qu'une esquisse laissant l'exercice au lecteur de rédiger les détails s'il le souhaite.

On démontre, en premier lieu, les formules analogues pour les j-ièmes sous-résultants scalaires (obtenues en remplaçant partout sRes par sres) pour lequelles on suit pas à pas la démonstration de la proposition 2.1.3. Dans un deuxième temps, on démontre des formules de transformation pour les j-ièmes cofacteurs en suivant à nouveau pas à pas les arguments de la proposition 2.2.2. On en déduit enfin la proposition en appliquant la définition des sous-résultants. □

Comme énoncé au début de ce numéro, les sous-résultants et leurs cofacteurs permettent de retrouver les RGCD et les identités de Bézout dans le cas où les polynômes de départ A et B ne sont pas premiers entre eux.

Théorème 2.3.5. *On suppose que l'anneau de base $\mathfrak{A}$ est un corps et on le note K. Soient $A, B \in K[X,\theta,\partial]$ deux polynômes de Ore. On suppose que d est le plus petit entier tel que $\mathrm{sres}_d(A,B) \neq 0$. Alors :*

1. *il existe $c \in K$, $c \neq 0$ tel que $\mathrm{RGCD}(A,B) = c \cdot \mathit{sRes}_d(A,B)$, et*
2. *il existe $c' \in K$, $c' \neq 0$ tel que $\mathrm{LLCM}(A,B) = c' \cdot U_{d-1}(A,B) \cdot A = -c' \cdot V_{d-1}(A,B) \cdot B$.*

Démonstration. Notons $m = \deg \text{RGCD}(A,B)$ et supposons par l'absurde que $m < d$. Nous allons démontrer que l'application $\text{Syl}_m(A,B)$ est injective, ce qui contredira la nullité de $\text{sres}_m(A,B)$. Pour simplifier les écritures, posons $a = \deg A$ et $b = \deg B$. Soit $(U,V) \in \mathcal{P}_{b-m} \times \mathcal{P}_{a-m}$ un couple dans le noyau de $\text{Syl}_m(A,B)$. Par définition, cela signifie que $\deg(UA+VB) < m$. Comme $UA+VB$ doit, d'autre part, être un multiple de $\text{RGCD}(A,B)$, on en déduit que $UA+VB = 0$, *i.e.* $UA = -VB$. Le polynôme de Ore $P = UA$ est ainsi un multiple à gauche commun à A et B ; il est donc également un multiple de $\text{LLCM}(A,B)$. De plus, par la proposition 1.3.13, on sait que $\deg \text{LLCM}(A,B) = a+b-m$. On en déduit ainsi que soit U est nul, soit $\deg U \geq b-m$. La deuxième possibilité étant exclue, on trouve $U = 0$, d'où on déduit finalement que $VB = 0$ puis que $V = 0$ par intégrité de $K[X,\theta,\partial]$. On a ainsi démontré que $\text{Syl}_m(A,B)$ est injective. L'hypothèse $m < d$ est ainsi absurde ; autrement dit $m = d$, ce qui signifie que $\text{RGCD}(A,B)$ est de degré d. Par ailleurs, l'égalité :

$$U_d(A,B)\cdot A + V_d(A,B)\cdot B = \text{sRes}_d(A,B)$$

montre que $\text{RGCD}(A,B)$ est un diviseur de $\text{sRes}_d(A,B)$. On en déduit la première assertion de la proposition.

Considérons maintenant l'application $\text{Syl}_{d-1}(A,B)$ et, comme précédemment, calculons son noyau. Soit $(U,V) \in \mathcal{P}_{a-d+1} \times \mathcal{P}_{b-d+1}$ tel que $UA+VB$ soit de degré strictement inférieur à $d-1$. Étant donné que $UA+VB$ est un multiple de $\text{RGCD}(A,B)$ qui est de degré d, on déduit que $UA+VB = 0$. Il en résulte que $P = UA = -VB$ est un multiple à gauche de $\text{LLCM}(A,B)$ et, par suite, par comparaison des degrés, que $P = \lambda \cdot \text{LLCM}(A,B)$ pour un certain $\lambda \in K$. Si on définit U_0 et V_0 par les égalités $\text{LLCM}(A,B) = U_0 A = -V_0 B$, on en déduit que $U = \lambda U_0$ et $V = \lambda V_0$. Autrement dit, $\ker \text{Syl}_{d-1}(A,B)$ est le sous-espace de dimension 1 engendré par le vecteur (U_0,V_0). Par la proposition A.1.3, on déduit que l'image de $\text{adj}(\text{Syl}_{d-1}(A,B))$ est également le sous-espace engendré par (U_0,V_0). En particulier, il existe un scalaire $c' \in K$ tel que $U_{d-1}(A,B) = c'U_0$ et $V_{d-1}(A,B) = c'V_0$. En revenant à la définition de U_0 et V_0, on trouve :

$$\text{LLCM}(A,B) = c'\cdot U_{d-1}(A,B)\cdot A = -c'\cdot V_{d-1}(A,B)\cdot B.$$

Pour conclure, il ne reste plus qu'à montrer que $c' \neq 0$. Pour cela, on remarque qu'il n'existe aucun polynôme de degré $d-1$ de la forme $UA+VB$; autrement dit, le monôme X^{d-1} n'appartient pas à l'image à $\text{Syl}_{d-1}(A,B)$. Par la proposition A.1.3 à nouveau, on trouve que X^{d-1} n'est pas dans le noyau de $\text{adj}(\text{Syl}_{d-1}(A,B))$, c'est-à-dire que $(U_{d-1}(A,B), V_{d-1}(A,B)) \neq 0$. Ceci implique que $c' \neq 0$ et la démonstration est terminée. □

Exemple 2.3.6. Reprenons rapidement l'exemple 2.3.2 où nous avions calculé les sous-résultants des polynômes de Ore :

$$\begin{aligned} A &= X^3 + (t^2+3t)X^2 + (5t^3+4t)X + (2t^5+6t^2+2)\\ B &= X^2 + (t^2+4t)X + (4t^3+2t)\end{aligned}$$

et obtenu les résultats suivants :

$$\begin{aligned} \text{sRes}_0(A,B) &= 0\\ \text{sRes}_1(A,B) &= \big(2t^3+4t^2-4\big)X + \big(2t^5+4t^4-4t^2\big)\\ \text{sRes}_2(A,B) &= B = X^2 + \big(t^2+4t\big)X + \big(4t^3+2t\big).\end{aligned}$$

Ainsi, d'après le théorème 2.3.5, si l'on voit A et B comme des polynômes de Ore à coefficients dans $\mathbb{Q}(t)$, le RGCD de A et B est, à un scalaire multiplicatif près, le premier sous-résultant non nul, c'est-à-dire $\text{sRes}_1(A,B)$. Après division par le coefficient dominant, on trouve de cette manière $\text{RGCD}(A,B) = X + t^2$.

Travailler avec les sous-résultants présente plusieurs intérêts. Tout d'abord, la théorie des sous-résultants, contrairement à l'algorithme d'Euclide, ne requiert pas que l'anneau de base soit un corps. Elle peut ainsi être considérée comme une extension de l'approche classique à un anneau de base quelconque (mais néanmoins commutatif). Cependant, même dans le cas d'un corps, l'utilisation des sous-résultants présente des avantages car elle permet, à moindre frais, d'obtenir des bornes sur la taille des coefficients des RGCD et des polynômes de Bézout. Ceci peut être intéressant pour les applications algorithmiques. Ces aspects seront détaillés dans le §2.4 ci-après.

2.4 Lien avec l'algorithme d'Euclide

Tout le long de ce numéro, on suppose que $\mathfrak{A}$ est un corps et on le note K. Nous avons vu précédemment (*cf* théorème 2.3.5) que les sous-résultants sont liés de près aux RGCD. Il se trouve que ce lien est encore plus profond et se voit déjà au niveau de l'algorithme d'Euclide, les restes intermédiaires obtenus au cours de l'exécution de cet algorithme s'interprétant eux-aussi (à quelques renormalisations près) comme des sous-résultants.

Ce point de vue est intéressant à double titre. D'une part, il fournit une méthode rapide de calcul des sous-résultants et, d'autre part, il fournit les clés théoriques pour une analyse fine de la croissance de la taille des coefficients des polynômes qui apparaissent au cours de l'exécution (d'une version modifiée) de l'algorithme d'Euclide.

2.4.1 Sous-résultants et division euclidienne

Afin de faire le rapprochement entre sous-résultants et algorithme d'Euclide, une étape préliminaire consiste à comprendre comment les sous-résultants et les cofacteurs se comportent par division euclidienne. C'est l'objet de la proposition suivante.

Proposition 2.4.1. *On conserve les notations précédentes et on se donne des entiers a et b avec $a \geq \deg A$ et $b \geq \deg B$. Soit m un entier compris entre -1 et $\min(a,b)$ avec $j \neq \min(a,b)$ si $a = b$. Soient Q et R des polynômes de Ore tels que $A = QB + R$ et $\deg Q \leq \deg A - \deg B$. Alors :*

$$\begin{aligned} \mathit{sres}_m^{a,b}(A,B) &= \mathit{sres}_m^{a,b}(R,B) \\ U_m^{a,b}(A,B) &= U_m^{a,b}(R,B) \\ V_m^{a,b}(A,B) &= V_m^{a,b}(R,B) - U_m^{a,b}(R,B)\cdot Q \\ \mathit{sRes}_m^{a,b}(A,B) &= \mathit{sRes}_m^{a,b}(R,B) \end{aligned}$$

Démonstration. On considère l'application $\mathfrak{A}$-linéaire :

$$\begin{array}{rccc} \varphi : & \mathcal{P}_{b-m} \times \mathcal{P}_{a-m} & \longrightarrow & \mathcal{P}_{b-m} \times \mathcal{P}_{a-m} \\ & (U,V) & \mapsto & (U,\, V+UQ). \end{array}$$

Cette application est bien définie grâce à l'hypothèse faite sur le degré de Q. La relation $UA + VB = U(QB+R) + VB = UR + (V+UQ)B$ indique que $\mathrm{Syl}_m^{a,b}(A,B) = \mathrm{Syl}_m^{a,b}(R,B) \circ \varphi$. En passant aux déterminants, on en déduit $\mathrm{sres}_m^{a,b}(A,B) = \mathrm{sres}_m^{a,b}(R,B)\cdot \det(\varphi)$. D'autre part, la matrice de φ dans les bases canoniques est manifestement triangulaire inférieure et n'a que des 1 sur la diagonale. On en déduit que $\det(\varphi) = 1$ et, finalement, que $\mathrm{sres}_m^{a,b}(A,B) = \mathrm{sres}_m^{a,b}(R,B)$ comme annoncé.

Pour les cofacteurs, on écrit :

$$\mathrm{adj}\big(\mathrm{Syl}_m^{a,b}(A,B)\big) = \mathrm{adj}(\varphi) \circ \mathrm{adj}\big(\mathrm{Syl}_m^{a,b}(R,B)\big).$$

Comme φ est de déterminant 1, on a $\varphi \circ \mathrm{adj}(\varphi) = \mathrm{id}$, *i.e.* $\mathrm{adj}(\varphi) = \varphi^{-1}$. On en déduit que $\mathrm{adj}(\varphi)$ est l'application qui envoie un couple (U', V') sur $(U',\, V'-U'Q)$, à partir de quoi les égalités énoncées dans la proposition s'ensuivent.

Enfin, pour les sous-résultants, on revient à la définition et on écrit :

$$\begin{aligned}\mathrm{sRes}_m^{a,b}(A,B) &= U_m^{a,b}(A,B)\cdot A + V_m^{a,b}(A,B)\cdot B\\ &= U_m^{a,b}(R,B)\cdot A + \big(V_m^{a,b}(R,B) - U_m^{a,b}(R,B)\cdot Q\big)\cdot B\\ &= U_m^{a,b}(R,B)\cdot (A - QB) + V_m^{a,b}(R,B)\cdot B\\ &= U_m^{a,b}(R,B)\cdot R + V_m^{a,b}(R,B)\cdot B = \mathrm{sRes}_m^{a,b}(R,B)\end{aligned}$$

ce qui est bien la formule annoncée. □

Corollaire 2.4.2. *Soient Q et R des polynômes de Ore tels que $A = QB + R$ et $\deg Q \leq \deg A - \deg B$. On pose $a = \deg A$, $b = \deg B$ et $r = \deg R$. Alors :*

$$\begin{aligned}\mathit{sRes}_m(A,B) &= (-1)^{(r-j)(b-j)}\,\Theta_{r-j}^{a-j}(B[b])\cdot \mathit{sRes}_m(B,R)\\ U_m(A,B) &= (-1)^{(r-j)(b-j)}\,\Theta_{r-j}^{a-j}(B[b])\cdot U_m(B,R)\\ V_m(A,B) &= (-1)^{(r-j)(b-j)}\,\Theta_{r-j}^{a-j}(B[b])\cdot \big(V_m(B,R) - U_m(B,R)\cdot Q\big).\end{aligned}$$

Démonstration. Les formules du corollaire s'obtiennent en combinant celles de la proposition 2.4.1 avec le troisième jeu de formules de la proposition 2.3.4. □

Remarque 2.4.3. Le corollaire 2.4.2 vaut en particulier lorsque Q et R sont respectivement le quotient et le reste de la division euclidienne à droite de A par B.

2.4.2 Cas de non-annulation des sous-résultants scalaires

Dans toute la suite, nous fixons deux polynômes de Ore $A, B \in K[X,\theta,\partial]$ que l'on suppose non nuls. Quitte à permuter A et B, on suppose également que $\deg A \geq \deg B$ et on pose $d = \deg B$. Dans ce numéro, en guise d'échauffement avant d'étudier le cas général (qui sera traité au §2.4.3), nous supposons en outre que les sous-résultants scalaires $\mathrm{sres}_m(A,B)$ $(0 \leq j < d)$ sont tous non nuls.

On définit trois suites finies $(R_i)_{0\leq i\leq d+1}$, $(U_i)_{0\leq i\leq d+1}$ et $(V_i)_{0\leq i\leq d+1}$ de la manière suivante :

$$\begin{aligned}&\text{pour } i = 0 : \quad R_0 = A,\ U_0 = 1,\ V_0 = 0\\ &\text{pour } i = 1 : \quad R_1 = B,\ U_1 = 0,\ V_0 = 1\\ &\text{pour } i \geq 2 : \quad R_i = \mathrm{sRes}_{d+1-i}(A,B),\ U_i = U_{d+1-1}(A,B),\ V_i = V_{d+1-i}(A,B).\end{aligned}$$

Notre objectif est de démontrer que les R_i, U_i et V_i satisfont des relations de récurrence semblables à celles qui apparaissent dans l'algorithme d'Euclide et qui, en particulier, permettent de calculer les sous-résultants et les cofacteurs sans avoir à revenir à la définition. Remarquons pour commencer que, grâce à l'hypothèse qui a été faite, on sait que R_i est de degré $d+1-i$ pour tout $i \in \{1,\dots,d+1\}$. Pour ces mêmes indices i, notons r_i le coefficient dominant de R_i. Convenons également que $r_0 = 1$. Enfin, si A et B sont deux polynômes de Ore, notons $A \,\%\, B$ (resp. $A \mathbin{/\!/} B$) le reste (resp. le quotient) de la division euclidienne à droite de A par B.

Théorème 2.4.4. *Avec les notations précédentes, on a :*

$$\begin{aligned}R_{i+1} &= \lambda_i\cdot\big(R_{i-1} \,\%\, R_i\big)\\ U_{i+1} &= \lambda_i\cdot\big(U_{i-1} - (R_{i-1} \mathbin{/\!/} R_i)\cdot U_i\big)\\ V_{i+1} &= \lambda_i\cdot\big(V_{i-1} - (R_{i-1} \mathbin{/\!/} R_i)\cdot V_i\big)\end{aligned}$$

avec $\lambda_i = \Theta_0^2\left(\dfrac{r_i}{r_{i-1}}\right) = \dfrac{r_i\cdot\theta(r_i)}{r_{i-1}\cdot\theta(r_{i-1})}$.

Démonstration. Nous allons démontrer la proposition par récurrence sur i. Fixons donc i et supposons donc que les formules de la proposition soient vraies pour tout indice $j < i$. Posons $m = d-1-i$. Le corollaire 2.4.2 implique alors l'égalité :

$$\mathrm{sRes}_m(R_{j-1}, R_j) = \Theta_{i-j-1}^{i-j+1}(r_j) \cdot \Theta_0^{i-j}(\lambda_j^{-1}) \cdot \mathrm{sRes}_m(R_j, R_{j-1})$$

pour tout $j < i$. En mettant ensemble toutes ces identités, on déduit :

$$\mathrm{sRes}_m(A, B) = \mathrm{sRes}_m(R_{i-1}, R_i) \cdot \prod_{j=1}^{i-1} \Theta_{i-j-1}^{i-j+1}(r_j) \cdot \Theta_0^{i-j}(\lambda_j^{-1}).$$

Étant donné que $m = \deg R_i$, le lemme 2.3.3 nous apprend que le m-ième sous-résultant de R_{i-1} et R_i vaut R_i. Afin de démontrer la première formule de la proposition, il ne reste donc plus qu'à vérifier que le facteur $\prod_{j=1}^{i-1} \Theta_{i-j-1}^{i-j+1}(r_j) \cdot \Theta_0^{i-j}(\lambda_j^{-1})$ est égal à 1. Or, par définition :

$$\prod_{j=1}^{i-1} \Theta_0^{i-j}(\lambda_j) = \prod_{j=1}^{i-1} \frac{\Theta_0^{i-j}(r_j)\, \Theta_1^{i-j+1}(r_j)}{\Theta_0^{i-j}(r_{j-1})\, \Theta_1^{i-j+1}(r_{j-1})}.$$

On remarque que l'expression du dénominateur vaut 1 lorsque $j = 0$ (car $r_0 = 1$) et lorsque $j = i$. En faisant un changement d'indice dans le produit au dénominateur, on trouve ainsi :

$$\begin{aligned}\prod_{j=1}^{i-1} \Theta_0^{i-j}(\lambda_j) &= \prod_{j=1}^{i-1} \frac{\Theta_0^{i-j}(r_j)\, \Theta_1^{i-j+1}(r_j)}{\Theta_0^{i-j-1}(r_j)\, \Theta_1^{i-j}(r_j)} \\ &= \prod_{j=1}^{i-1} \theta^{i-j-1}(r_j)\, \theta^{i-j}(r_j) = \prod_{j=1}^{i-1} \Theta_{i-j-1}^{i-j+1}(r_j)\end{aligned}$$

ce qui conclut. Les deux autres formules se démontrent de manière entièrement analogue. □

Le théorème 2.4.4 a un intérêt calculatoire évident : il permet de calculer de proche en proche les sous-résultants pendant un algorithme d'Euclide « renormalisé » sans avoir à évaluer des déterminants de grande taille. Par ailleurs, il se trouve que l'algorithme d'Euclide « renormalisé » est, lui-même, en général, meilleur que l'algorithme d'Euclide standard car les coefficients intermédiaires qu'il fait intervenir sont plus petits et peuvent, en outre, être bornés *a priori*. Également, si les coefficients des polynômes initiaux A et B appartient tous à un sous-anneau $\mathcal{O}_K$ de K stable par θ et ∂, la fonctorialité des sous-résultants entraîne que $\mathrm{Res}_m(A, B), U_m(A, B), V_m(A, B) \in \mathcal{O}_K[X, \theta, \partial]$ pour tout m. On en déduit que l'algorithme d'Euclide renormalisée ne produit, dans ce cas, que des polynômes de Ore à coefficients dans $\mathcal{O}_K$, ceci malgré les divisions qui apparaissent. À titre de comparaison, cette propriété n'est pas valable pour l'algorithme d'Euclide usuel. L'exemple ci-après illustre ces phénomènes.

Exemple 2.4.5. Dans l'anneau de Ore $\mathbb{Q}(t)[X, \frac{d}{dt}]$, considérons les deux polynômes de Ore $A = X^3 - 2tX^2 + 5t^2X + 3$ et $B = X^3 - t^3X^2 + tX + t^2$. La formule récurrente du théorème 2.4.4 permet de calculer les sous-résultants de A et B de proche en proche. On obtient, ce faisant, les résultats suivants :

$$\begin{aligned}R_2 &= \big(t^3 - 2t\big)X^2 + \big(5t^2 - t\big)X - \big(t^2 - 3\big) \\ R_3 &= \big(5t^8 - 10t^6 - t^5 + 30t^4 - 13t^3 + 11t^2 + 6t\big)X + \big(t^6 - 8t^4 + t^3 + 22t^2 - 3t - 6\big) \\ R_4 &= -25t^{12} + 60t^{10} + 16t^9 - 290t^8 + 82t^7 + 275t^6 + 44t^5 - 255t^4 - 93t^3 + 555t^2 + 216t + 66\end{aligned}$$

À titre de comparaison, voici les restes successifs que l'on obtient avec l'algorithme d'Euclide non modifié :

$$R'_2 = (t^3-2t)X^2 + (5t^2-t)X - (t^2-3)$$
$$R'_3 = \frac{5t^8-10t^6-t^5+30t^4-13t^3+11t^2+6t}{t^6-4t^4+4t^2}\,X + \frac{t^6-8t^4+t^3+22t^2-3t-6}{t^6-4t^4+4t^2}$$
$$R'_4 = \frac{(t^2-2)^2\cdot(25t^{12}-60t^{10}-16t^9+290t^8-82t^7-275t^6-44t^5+255t^4+93t^3-555t^2-216t-66)}{(5t^7-10t^5-t^4+30t^3-13t^2+11t+6)^2}$$

On constate que ceux-ci sont proportionnels aux sous-résultants mais que des facteurs « parasites » apparaissent, à la fois au numérateur et au dénominateur. Dans le cas des polynômes de cet exemple (où A et B sont des polynômes de Ore unitaires de degré 3), ces facteurs parasites restent néanmoins de taille modérée. Ce n'est toutefois rapidement plus le cas lorsque le degré des polynômes d'entrée augmente.

2.4.3 Le cas général

Nous en venons à présent au cas général où certains sous-résultants scalaires peuvent s'annuler. En fait, la démarche est identique, seuls les calculs sont un peu plus délicats à mener. Nous changeons néanmoins légèrement de point de vue et définissons à présent la suite (R_i) par les relations de récurrence :

$$R_0 = A, \quad R_1 = B, \quad R_{i+1} = \lambda_i \cdot \big(R_{i-1} \,\%\, R_i\big)$$

où les $\lambda_i \in K$ sont des constantes non nulles que nous préciserons par la suite. Bien entendu, comme habituellement, la construction s'arrête dès lors que l'on a atteint un reste R_i égal à 0. Posons $d_i = \deg R_i$ et notons $r_i = R_i[d_i]$ le coefficient dominant de R_i. On remarque que les d_i ne dépendent pas des choix des λ_i.

Nous allons chercher à ajuster les λ_i de sorte que les R_i s'interprètent comme des sous-résultants. On part de la formule $R_{i+1} = \mathrm{sRes}_{d_i-1}^{d_i,d_i-1}(R_i, R_{i+1})$ qui provient directement du lemme 2.3.3. L'idée consiste à présent à exprimer $\mathrm{sRes}_{d_i-1}^{d_i,d_i-1}(R_i, R_{i+1})$ en fonction d'un sous-résultant de A et B en utilisant les formules de la proposition 2.3.4 et du corollaire 2.4.2. Pour un entier m dans l'intervalle $[d_{i+1},\, d_i-1]$ et un entier $j < i$, la combinaison des deux énoncés précédents permet d'écrire :

$$\begin{aligned}
\mathrm{sRes}_m(R_{j-1}, R_j) &= (-1)^{(d_{j+1}-m)(d_j-m)} \cdot \Theta_{d_{j+1}-m}^{d_{j-1}-m}(r_j) \cdot \Theta_0^{d_j-m}(\lambda_j^{-1}) \cdot \mathrm{sRes}_m(R_j, R_{j+1}) \\
\mathrm{sRes}_m(R_{i-1}, R_i) &= \Theta_0^{d_{i-1}-m}(r_i) \cdot \Theta_0^{d_i-m}(\lambda_i^{-1}) \cdot \mathrm{sRes}_m^{d_i,m}(R_i, R_{i+1}) \\
&= \Theta_0^{d_{i-1}-m}(r_i) \cdot \Theta_0^{d_i-m}(\lambda_i^{-1}) \cdot \Theta_1^{d_i-m}(R_{i+1}[m]) \cdot R_{i+1}.
\end{aligned}$$

où la dernière égalité provient du lemme 2.3.3. On en déduit que $\mathrm{sRes}_m(A, B) = c_{i,m} R_{i+1}$ où le coefficient $c_{i,m}$ est donné par la formule :

$$c_{i,m} = \Theta_0^{d_{i-1}-m}(r_i) \cdot \Theta_0^{d_i-m}(\lambda_i^{-1}) \cdot \Theta_1^{d_i-m}(R_{i+1}[m]) \cdot \prod_{j=1}^{i-1} (-1)^{(d_{j+1}-m)(d_j-m)} \cdot \Theta_{d_{j+1}-m}^{d_{j-1}-m}(r_j) \cdot \Theta_0^{d_j-m}(\lambda_j^{-1}).$$

Pour $m = d_i-1$, une légère simplification se produit et conduit à :

$$c_{i,d_i-1} = \Theta_0^{d_{i-1}-d_i+1}(r_i) \cdot \lambda_i^{-1} \cdot \prod_{j=1}^{i-1} (-1)^{(d_{j+1}-d_i+1)(d_j-d_i+1)} \cdot \Theta_{d_{j+1}-d_i+1}^{d_{j-1}-d_i+1}(r_j) \cdot \Theta_0^{d_j-d_i+1}(\lambda_j^{-1}).$$

Nous choisissons λ_i de sorte à rendre ce coefficient égal à 1. Autrement dit, nous posons :

$$\lambda_i = \Theta_0^{d_{i-1}-d_i+1}(r_i) \cdot \prod_{j=1}^{i-1} (-1)^{(d_{j+1}-d_i+1)(d_j-d_i+1)} \cdot \Theta_{d_{j+1}-d_i+1}^{d_{j-1}-d_i+1}(r_j) \cdot \Theta_0^{d_j-d_i+1}(\lambda_j^{-1}). \tag{18}$$

Nous simplifierons cette formule par la suite (*cf* Eq. (19) ci-après). En attendant, remarquons qu'avec ce choix, nous avons la proposition suivante qui donne l'expression de tous les sous-résultants en fonction des R_i.

Proposition 2.4.6. *Avec les notations précédentes, pour tout indice $i \geq 1$ pour lequel R_{i+1} est défini, on a :*

- *si* $m = d_i - 1$, $\quad \mathit{sRes}_m(A,B) = R_{i+1}$,
- *si* $d_{i+1} < m < d_i - 1$, $\quad \mathit{sRes}_m(A,B) = 0$,
- *si* $m = d_{i+1}$, $\quad \mathit{sRes}_m(A,B) = \pm\, \Theta_0^{d_i - d_{i+1} - 1}\left(\dfrac{r_i \cdot \theta(r_{i+1})}{\lambda_1 \cdot \lambda_2 \cdots \lambda_i}\right) \cdot R_{i+1}$.

Démonstration. Lorsque $m = d_i - 1$, l'égalité annoncée provient du choix de λ_i. D'autre part, pour $m \in [d_i - 1, d_{i+1}]$, on vérifie l'égalité :

$$c_{i,m} = \pm\, \theta^{d_i - 1 - m}(c_{i,d_i - 1}) \cdot \Theta_0^{d_i - 1 - m}\left(\frac{r_i \cdot \theta(R_{i+1}[m])}{\lambda_1 \cdot \lambda_2 \cdots \lambda_i}\right)$$

Or, on rappelle que les λ_i ont été choisis de sorte que $c_{i,d_i - 1} = 1$. On en déduit que

$$\mathrm{sRes}_m(A,B) = \pm\, \Theta_0^{d_i - 1 - m}\left(\frac{r_i \cdot \theta(R_{i+1}[m])}{\lambda_1 \cdot \lambda_2 \cdots \lambda_i}\right) \cdot \mathrm{sRes}_{d_i - 1}(A,B).$$

Ceci donne directement la relation annoncé pour $m = d_{i+1}$. Lorsque $d_{i+1} < m < d_i - 1$, on constate que le coefficient $R_{i+1}[m]$ s'annule. Il en est donc de même que $\Theta_1^{d_i - m}(R_{i+1}[m])$ étant donné que $d_i - 1 - m > 0$. Par suite $\mathrm{sRes}_m(A,B) = 0$, lui aussi. Enfin, l □

Remarque 2.4.7. Bien que cela soit un peu fastidieux, il est bien sûr tout à fait possible de déterminer complètement le signe dans la troisième égalité de la proposition 2.4.6. Après calcul, on trouve qu'il vaut $(-1)^{(d_i - d_{i+1} - 1)(d_1 - d_i - 1 + i)}$.

Remarque 2.4.8. Lorsque $d_{i+1} = d_i - 1$, la proposition 2.4.6 donne deux expressions différentes pour $\mathrm{sRes}_{d_i - 1}(A,B) = \mathrm{sRes}_{d_{i+1}}(A,B)$. Bien sûr, celles-ci conduisent à la même valeur étant donné, d'une part, que le facteur $\Theta_0^{d_i - d_{i+1} - 1}(\cdots)$ vaut 1 puisque l'exposant $d_i - d_{i+1} - 1$ vaut 0 et, d'autre part, que l'on peut montrer que le signe est positif grâce à la formule donnée dans la remarque 2.4.7 ci-dessus.

Simplification de l'expression des λ_i. Pour alléger les formules à suivre, commençons par noter ε_i le signe qui apparaît dans l'expression de λ_i, à savoir

$$\varepsilon_i = (-1)^{\sum_{j=1}^{i-1}(d_{j+1} - d_i + 1)(d_j - d_i + 1)}.$$

Soit $i \geq 2$. En appliquant $\theta^{d_{i-1} - d_i}$ à la formule (18) prise en $i-1$, on trouve la relation :

$$\begin{aligned}\theta^{d_{i-1} - d_i}(\lambda_{i-1}) &= \varepsilon_{i-1} \cdot \Theta_{d_{i-1} - d_i}^{d_{i-2} - d_i + 1}(r_{i-1}) \cdot \prod_{j=1}^{i-2} \Theta_{d_{j+1} - d_i + 1}^{d_{j-1} - d_i + 1}(r_j) \cdot \Theta_{d_{i-1} - d_i}^{d_j - d_i + 1}(\lambda_j^{-1}) \\ &= \frac{\varepsilon_{i-1}}{\varepsilon_i} \cdot \frac{\Theta_0^{d_{i-1} - d_i}(\lambda_1 \cdots \lambda_{i-2}) \cdot \Theta_0^{d_{i-1} - d_i + 1}(\lambda_{i-1})}{\Theta_1^{d_{i-1} - d_i}(r_{i-1}) \cdot \Theta_0^{d_{i-1} - d_i + 1}(r_i)} \cdot \lambda_i\end{aligned}$$

ce qui donne, en simplifiant par $\theta^{d_{i-1} - d_i}(\lambda_{i-1})$ et en réorganisant les facteurs :

$$\lambda_i = \frac{\varepsilon_i}{\varepsilon_{i-1}} \cdot \frac{\Theta_1^{d_{i-1} - d_i}(r_{i-1}) \cdot \Theta_0^{d_{i-1} - d_i + 1}(r_i)}{\Theta_0^{d_{i-1} - d_i}(\lambda_1 \cdots \lambda_{i-1})}.$$

Pour ce qui concerne le signe, on remarque que l'exposant sur (-1) qui apparaît dans l'expression de $\frac{\varepsilon_i}{\varepsilon_{i-1}}$ est $e = \sum_{j=1}^{i-1}(d_{j+1} - d_i + 1)(d_j - d_i + 1) - \sum_{j=1}^{i-2}(d_{j+1} - d_{i-1} + 1)(d_j - d_{i-1} + 1)$. Or, en développant et en réorganisant les termes, on remarque que :

$$\begin{aligned}(d_{j+1} - d_i + 1)(d_j - d_i + 1) - (d_{j+1} - d_{i-1} + 1)(d_j - d_{i-1} + 1)& \\ \equiv (d_{i-1} - d_i)(d_j - d_{j+1} + 1) \pmod 2.&\end{aligned}$$

Ainsi en sommant, on obtient $e \equiv (d_{i-1} - d_i)(d_1 - d_i + i) + 1 \pmod 2$ et donc, finalement :

$$\lambda_i = (-1)^{(d_{i-1}-d_i)(d_1-d_i+i)+1} \cdot \frac{\Theta_1^{d_{i-1}-d_i}(r_{i-1}) \cdot \Theta_0^{d_{i-1}-d_i+1}(r_i)}{\Theta_0^{d_{i-1}-d_i}(\lambda_1 \cdots \lambda_{i-1})}. \tag{19}$$

En multiplication l'identité précédente par $\lambda_1 \cdots \lambda_{i-1}$, on s'aperçoit que la quantité $\mu_i = \lambda_2 \cdots \lambda_i$ vérifie une relation de récurrence simple d'ordre 1 qui s'écrit :

$$\begin{aligned}\mu_1 &= \Theta_0^{d_0-d_1+1}(r_1) \\ \mu_i &= (-1)^{(d_{i-1}-d_i)(d_1-d_i+i)+1} \cdot \Theta_1^{d_{i-1}-d_i}\left(\frac{r_{i-1}}{\mu_{i-1}}\right) \cdot \Theta_0^{d_{i-1}-d_i+1}(r_i) \quad \text{pour } i \geq 2. \end{aligned} \tag{20}$$

En intégrant ces nouvelles relations à la récurrence définissant les R_i, on peut calculer les μ_i en même temps que les R_i pour un coût raisonnable. À partir de là, on retrouve également facilement les λ_i soit en utilisant (19), soit en utilisant directement la relation $\lambda_i = \frac{\mu_i}{\mu_{i-1}}$. Soulignons enfin que, dans le cas particulier (mais générique) où $d_i = d_{i-1} - 1$, le facteur $\Theta_1^{d_{i-1}-d_i}\left(\frac{r_{i-1}}{\mu_{i-1}}\right)$ disparaît et on a simplement $\mu_i = \pm\, \Theta_0^2(r_i) = \pm\, r_i \cdot \theta(r_i)$ (où le signe est explicite). En particulier, on constate que cette dernière formule ne fait plus intervenir μ_{i-1}. On retrouve, de cette manière, les résultats du §2.4.2.

Remarque 2.4.9. La formule (20) est également valable pour $i = 1$ si l'on prend soin de poser $\mu_0 = -1$ et $r_0 = 1$.

Remarque 2.4.10. La relation de récurrence définissant les μ_i se réécrit de manière encore légèrement plus concise si l'on introduit la variable $\eta_i = \frac{\mu_i}{r_i}$ puisque l'on a alors :

$$\eta_i = (-1)^{(d_{i-1}-d_i)(d_1-d_i+i)+1} \cdot \frac{\Theta_1^{d_{i-1}-d_i+1}(r_i)}{\Theta_1^{d_{i-1}-d_i}(\eta_{i-1})} \quad \text{pour } i \geq 1$$

où, conformément à la remarque 2.4.9 ci-dessus, on convient d'initialiser la récurrence par $\eta_0 = -1$ (la valeur de r_0 n'est pas utilisée ici, il n'est donc pas nécessaire de la spécifier). Les λ_i se retrouvent à partir des η_i soit à partir de la formule $\lambda_i = \frac{\eta_i r_i}{\eta_{i-1} r_{i-1}}$, soit à partir de l'égalité (19) qui, dans notre contexte, se réécrit comme ceci :

$$\lambda_i = (-1)^{(d_{i-1}-d_i)(d_1-d_i+i)+1} \cdot \frac{\Theta_0^{d_{i-1}-d_i+1}(r_i)}{r_{i-1} \cdot \Theta_0^{d_{i-1}-d_i}(\eta_{i-1})}.$$

À propos des cofacteurs. Pour conclure, signalons rapidement qu'il est également possible de calculer les cofacteurs par une méthode analogue. Il suffit, pour ce faire, de définir les suites (U_i) et (V_i) à l'aide des formules de récurrence suivantes :

$$\begin{aligned} U_0 = 1,\ V_0 = 0 \\ U_1 = 0,\ V_1 = 1 \\ U_{i+1} = \lambda_i \cdot \big(U_{i-1} - (R_{i-1} /\!\!/ R_i) \cdot U_i\big) \\ V_{i+1} = \lambda_i \cdot \big(V_{i-1} - (R_{i-1} /\!\!/ R_i) \cdot V_i\big). \end{aligned}$$

On peut alors montrer un analogue de la proposition 2.4.6 en appliquant exactement les mêmes techniques que précédemment, à savoir précisément :

- si $m = d_i - 1$, $\quad U_m(A,B) = U_{i+1} \quad$ et $\quad V_m(A,B) = V_{i+1}$,
- si $d_{i+1} < m < d_i - 1$, $\quad U_m(A,B) = V_m(A,B) = 0$,
- si $m = d_{i+1}$, $\quad U_m(A,B) = \pm\, \Theta_0^{d_i - d_{i+1} - 1}\left(\frac{r_i \cdot \theta(r_{i+1})}{\lambda_1 \cdot \lambda_2 \cdots \lambda_i}\right) \cdot U_{i+1}$,

$$V_m(A,B) = \pm\, \Theta_0^{d_i - d_{i+1} - 1}\left(\frac{r_i \cdot \theta(r_{i+1})}{\lambda_1 \cdot \lambda_2 \cdots \lambda_i}\right) \cdot V_{i+1}.$$

où, à nouveau, le signe dans les deux dernières équations peut être rendu explicite conformément à la remarque 2.4.7.

3 Algèbres simples centrales

Dans certains cas, les anneaux de polynômes de Ore qui ont été introduits au §1 peuvent être comparés à des anneaux de matrices, ce qui fournit de nouveaux outils puissants pour les étudier. Voici un exemple typique de ce phénomène. Dans le cas de l'anneau de Ore $\mathbb{C}[X, \mathrm{conj}]$ (où on rappelle que conj est le morphisme de conjugaison complexe), on rappelle (*cf* remarque 1.3.6) qu'on dispose d'un morphisme dit d'*évaluation* :

$$\begin{array}{rcl} \mathbb{C}[X,\mathrm{conj}] & \longrightarrow & \mathrm{End}_{\mathbb{Z}}(\mathbb{C}) \\ P(X) = a_0 + a_1 X + \cdots + a_d X^d & \mapsto & P(\mathrm{conj}) = a_0 \cdot \mathrm{id}_{\mathbb{C}} + a_1 \cdot \mathrm{conj} + a_2 \cdot \mathrm{conj}^2 + \cdots + a_d \cdot \mathrm{conj}^d \\ & & (z \mapsto a_0 z + a_1 \bar{z} + a_2 z + a_3 \bar{z} + \cdots) \end{array}.$$

Celui-ci prend ses valeurs dans $\mathrm{End}_{\mathbb{R}}(\mathbb{C})$, l'anneau des endomorphismes $\mathbb{R}$-linéaires de $\mathbb{C}$ et il se trouve qu'il induit un isomorphisme entre $\mathbb{C}[X, \mathrm{conj}]/(X^2 - 1)$ et $\mathrm{End}_{\mathbb{R}}(\mathbb{C})$. À partir de là, on peut déduire que les diviseurs non triviaux de $X^2 - 1$ dans $\mathbb{C}[X, \mathrm{conj}]$ sont paramétrés par les diviseurs de zéro de $\mathrm{End}_{\mathbb{R}}(\mathbb{C})$, c'est-à-dire par les endomorphismes de rang 1. En général, la situation n'est pas toujours aussi transparente mais la philosophie demeure la même : on compare les quotients des anneaux de Ore à certaines algèbres de matrices et, de cette comparaison, on réinterprète dans le langage de l'algèbre linéaire les propriétés de factorisation des polynômes de Ore.

Le cadre pertinent dans lequel s'exprime et s'étudie cette correspondance est celui des algèbres d'Azumaya dont le cœur est la théorie des algèbres simples centrales. Le but de cette partie est de mettre en place les éléments de ces théories qui nous seront utiles dans la suite. Nous insisterons particulièrement sur la notion de *norme réduite* (qui doit être pensé comme l'équivalent de la notion de déterminant d'une matrice) qui jouera un rôle fondamental pour les applications à la factorisation dans les anneaux de Ore.

En aucun cas, ce chapitre ne doit être considéré comme un traité complet sur les algèbres simples centrales et les algèbres d'Azumaya. Pour de nombreux compléments sur le sujet, nous renvoyons la lectrice et le lecteur aux livres [10, 4, 24].

Dans toute cette partie, la lettre K designe un corps fixé.

3.1 Définition et exemples

Si A est un anneau, on rappelle qu'un idéal bilatère $\mathcal{I}$ de A est un sous-groupe additif de A vérifiant la condition supplémentaire suivante : pour tout $a \in A$ et tout $x \in \mathcal{I}$, les produits ax et xa appartiennent à $\mathcal{I}$. Autrement dit, un idéal bilatère de A est un sous-ensemble de A qui est, à la fois, un idéal à gauche et un idéal à droite. On rappelle également que le centre de A, parfois noté $Z(A)$, est défini comme l'ensemble des éléments de A qui commutent avec tous les élements de A :

$$Z(A) = \big\{\, x \in A \quad \text{tels que} \quad \forall y \in A,\ xy = yx \,\big\}.$$

Définition 3.1.1. Soit A une K-algèbre (non nécessairement commutative).
On dit que A est *simple* si ses seuls idéaux bilatères sont 0 et A.
On dit que A est *centrale* si le centre de A est K.

Les deux exemples d'algèbres simples centrales que nous allons présenter ci-dessous sont fondamentaux ; nous énoncerons et démontrerons dans la suite un théorème de structure qui présente toute algèbre simple centrale de dimension finie comme une combinaison de ces deux modèles.

3.1.1 Premier exemple : algèbre de matrices

Étant donné un espace vectoriel V de dimension finie sur K, l'algèbre $\mathrm{End}(V)$ des endomorphismes de K-linéaires de V est une K-algèbre simple centrale. En particulier, pour $V = K^n$, on trouve que l'algèbre de matrices $M_n(K)$ est une algèbre simple centrale sur K.

Démontrons ce résultat. Le fait que $\mathrm{End}(V)$ soit centrale est un exercice classique. Considérons une application K-linéaire $f : V \to V$ et supposons que f commute avec tous les éléments de $\mathrm{End}(V)$. S'il existe un élément $x \in V$ pour lequel x et $f(x)$ ne sont pas liés, on pourrait construire une application $g : V \to V$ qui s'annule sur x mais pas sur $f(x)$. Mais alors, on aurait $f \circ g(x) = 0$ mais $g \circ f(x) \neq 0$, contredisant ainsi l'hypothèse faite sur f. On en déduit qu'il existe une fonction $\lambda : V\backslash\{0\} \to K$ telle que $f(x) = \lambda_x \cdot x$ pour tout $x \neq 0$. Il s'agit, à présent, de montrer que la fonction $x \mapsto \lambda_x$ est constante. Pour cela, on considère $x, y \in V$. Si la famille (x, y) est liée, le résultat suit de la linéarité de f. Sinon, il suit de l'égalité $f(x + y) = f(x) + f(y)$ que $\lambda_{x+y}x + \lambda_{x+y}y = \lambda_x x + \lambda_y y$. En identifiant les coefficients, on obtient $\lambda_x = \lambda_{x+y} = \lambda_y$, ce qui conclut.

Il reste à démontrer que $\mathrm{End}(V)$ est une algèbre simple. On utilise pour cela le lemme suivant.

Lemme 3.1.2. *Soient $f, g : V \to V$ deux applications K-linéaires. Alors, il existe $\varphi, \psi : V \to V$ telles que* $\mathrm{im}(f \circ \varphi + g \circ \psi) = \mathrm{im}f + \mathrm{im}g$.

Démonstration. On considère $(e_i)_{i\in I}$ une base de V ainsi qu'un sous-ensemble $J \subset I$ tel que la famille $(e_i)_{i\in J}$ soit une base de $\mathrm{im}f + \mathrm{im}g$. Pour $i \in J$, on choisit un couple $(x_i, y_i) \in V^2$ tel que $e_i = f(x_i)+g(y_i)$. On considère à présent les applications $\varphi, \psi : V \to V$ définies ainsi : pour $i \in J$, on pose $\varphi(e_i) = x_i$ et $\psi(e_i) = y_i$ tandis que pour $i \notin J$, on pose $\varphi(e_i) = \psi(e_i) = 0$. Clairement l'image de $f\circ\varphi+g\circ\psi$ contient tous les e_i pour $i \in J$, de sorte que l'on a $\mathrm{im}(f\circ\varphi+g\circ\psi) = \mathrm{im}f+\mathrm{im}g$ comme voulu. □

Considérons à présent un idéal bilatère non nul I de $\mathrm{End}(V)$. On souhaite démontrer que $I = \mathrm{End}(V)$. Soit $f \in \mathrm{End}(V)$, $f \neq 0$. Soit $x \in \mathrm{im}f$. Par hypothèse I contient la composée $\varphi \circ f$ pour tout endomorphisme φ de V. De plus l'image de $\varphi \circ f$ contient manifestement le vecteur $\varphi(x)$. Une application répétée du lemme 3.1.2 assure que I contient un endomorphisme surjectif, et donc bijectif. Autrement dit I contient un élément inversible et, par conséquent, $I = \mathrm{End}(V)$. Il en résulte que $\mathrm{End}(V)$ est une algèbre simple.

Le lemme 3.1.2 permet en réalité d'être plus précis et de décrire complètement les idéaux à gauche et à droite de $\mathrm{End}(V)$. En effet, on a le résultat suivant.

Proposition 3.1.3. *Les idéaux à droite de $End(V)$ sont principaux, c'est-à-dire de la forme $f \cdot End(V)$ pour un certain $f \in End(V)$. De plus, on a la description concrète suivante :*

$$f \cdot End(V) = \big\{\, g \in End(V) \quad tel\ que \quad \mathrm{im}f \subset \mathrm{im}g \,\big\}. \tag{21}$$

Démonstration. Soit $\mathcal{I}$ un idéal à droite de $\mathrm{End}(V)$. Soit $f \in \mathcal{I}$ un endomorphisme pour lequel $\dim_K \mathrm{im}f$ est maximale. Considérons $g \in \mathcal{I}$. D'après le lemme 3.1.2, il existe un élement de $\mathcal{I}$ dont l'image est $\mathrm{im}f + \mathrm{im}g$. Par maximalité, on doit avoir $\mathrm{im}g \subset \mathrm{im}f$.

Choisissons à présent une base $(e_i)_{i\in I}$ de V. Pour tout $i \in J$, on choisit également un antécédent $x_i \in V$ de $g(e_i)$ par f ; cela est possible en vertu de l'inclusion $\mathrm{im} g \subset \mathrm{im} f$. Si $\varphi : V \to V$ est l'application K-linéaire définie par $\varphi(e_i) = x_i$ pour $i \in I$, on vérifie directement que $g = f \circ \varphi$. Ainsi $g \in f \cdot \mathrm{End}(V)$ et la première assertion est démontrée.

L'égalité (21) se démontre de manière similaire. □

Pareillement (ou, au choix, en utilisant la dualité), on démontre que les idéaux à gauche de $\mathrm{End}(V)$ sont également principaux, c'est-à-dire de la forme $\mathrm{End}(V) \cdot f$ et que :

$$\mathrm{End}(V) \cdot f = \big\{\, g \in \mathrm{End}(V) \quad \text{tel que} \quad \ker g \subset \ker f \,\big\}.$$

On constate sur ces descriptions que les idéaux à droite (resp. à gauche) de $\mathrm{End}(V)$ sont paramétrés par les sous-espaces vectoriels de V, en l'occurrence $\mathrm{im} f$ (resp. $\ker f$) où f est un générateur. Nous verrons dans la suite, lorsque nous étudierons l'équivalence de Morita (*cf* §3.2.3) comment ces résultats s'étendent aux modules (à gauche et à droite) sur $\mathrm{End}(V)$.

Deuxième exemple : algèbre de quaternions

On suppose ici (pour simplifier) que K est le corps des nombres réels $\mathbb{R}$. L'algèbre des quaternions $\mathbb{H}$ est définie comme l'espace vectoriel réel $\mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k$ muni de la table de multiplication suivante :

$\times$	**1**	**i**	**j**	**k**
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

(22)

Attention, cette table se lit en mettant à gauche l'élément indiquant la ligne (ce sont donc ceux qui apparaissent sur la première colonne) et à droite celui indiquant la colonne. Par exemple, on a $ij = k$ et $ji = -k$. On résumé souvent cette table de mulitiplication par la relation mnémotechnique $i^2 = j^2 = k^2 = ijk = -1$. En effet, en supposant que $\mathbb{H}$ est intègre, cette relation permet de reconstruire la table. Par exemple, en simplifiant par k à droite, on déduit de l'égalité $k^2 = ijk$ que $ij = k$.

Un calcul facile conduit à la relation $(a + bi + cj + dk) \cdot (a - bi - cj - dk) = a^2 + b^2 + c^2 + d^2$. On en déduit directement que tout quaternion non nul $a + ib + jc + kd$ est inversible, d'inverse :

$$\frac{a}{a^2+b^2+c^2+d^2} - \frac{b}{a^2+b^2+c^2+d^2}\, i - \frac{c}{a^2+b^2+c^2+d^2}\, j - \frac{d}{a^2+b^2+c^2+d^2}\, k \qquad (23)$$

le dénominateur ne s'annulant pas puisque a, b, c, d sont réels. On dit que $\mathbb{H}$ est une *algèbre à divisions*. Il est clair, par ailleurs, que cette propriété implique que $\mathbb{H}$ est une algèbre simple puisqu'un idéal bilatère contenant un élément inversible est nécessairement trivial.

Le fait que le centre de $\mathbb{H}$ est $\mathbb{R}$ se vérifie aisément à la main. Soit, en effet, $x = a+bi+cj+dk$ un élement central de $\mathbb{H}$. En écrivant que x commute avec i, on obtient $ai-b-ck+dj = ai-b+ck-dj$, soit $c = d = 0$. À présent, le fait que x commute avec j implique $b = 0$. Il reste donc $x = a \in \mathbb{R}$.

D'autres remarques intéressantes peuvent être faites sur les quaternions. La première est que le corps de nombres complexes $\mathbb{C}$ apparaît naturellement comme un sous-corps de $\mathbb{H}$, par exemple en voyant $\sqrt{-1} \in \mathbb{C}$ comme l'élement $i \in \mathbb{H}$. L'écriture $a + bi + cj + dk = (a + bi) + (c + di)j$ (pour $a, b, c, d \in \mathbb{R}$) fait en outre apparaître $\mathbb{H}$ comme un espace vectoriel complexe de dimension 2 dont une base est $(1, j)$. Dans cette base, la multiplication à droite par $x = a + bi + cj + dk$ (qui est bien $\mathbb{C}$-linéaire) a pour matrice :

$$M_x = \begin{pmatrix} a + b\sqrt{-1} & -c + d\sqrt{-1} \\ c + d\sqrt{-1} & a - b\sqrt{-1} \end{pmatrix}.$$

de sorte que l'application $\iota : \mathbb{H} \to M_2(\mathbb{C})$, $x \mapsto M_x$ est un morphisme d'anneaux injectif. Notons $\mathbb{H}_{\mathbb{C}} = \mathbb{C} \otimes_{\mathbb{R}} \mathbb{H}$ la $\mathbb{C}$-algèbre obtenue à partir de $\mathbb{H}$ en étendant les scalaires à $\mathbb{C}$: comme $\mathbb{C}$-espace vectoriel, $\mathbb{H}_{\mathbb{C}} = \mathbb{C} \oplus \mathbb{C}i \oplus \mathbb{C}j \oplus \mathbb{C}k$ tandis que la loi de multiplication sur $\mathbb{H}_{\mathbb{C}}$ est encore donnée par la table (22). Le morphisme ι induit un morphisme de $\mathbb{C}$-algèbres $\iota_{\mathbb{C}} : \mathbb{H}_{\mathbb{C}} \to M_2(\mathbb{C})$, qui est un isomorphisme. Autrement dit, après extension des scalaires à $\mathbb{C}$, l'algèbre de quaternions $\mathbb{H}$ devient isomorphe à une algèbre de matrices.

Cette description a une vertu supplémentaire : elle permet de donner une nouvelle interprétation, plus intrinsèque, à la quantité $a^2 + b^2 + c^2 + d^2$ qui est apparue précédemment (voir par exemple Eq. (23)). En effet, un calcul immédiat montre que $a^2 + b^2 + c^2 + d^2$ n'est autre que le déterminant de M_x (pour $x = a + bi + cj + dk$). L'application $x \mapsto a^2 + b^2 + c^2 + d^2$ s'appelle la *norme réduite*. Elle aura un sens, grâce à l'interprétation « déterminantale », dans le cadre d'une algèbre simple centrale quelconque et jouera un rôle essentiel dans l'étude des polynômes de Ore au §4.

3.2 Théorèmes de structure

Bien que la définition des algèbres simples centrales paraisse *a priori* relativement souple et semble laisser la porte ouverte à des nombreux exemples variés, il se trouve qu'il n'en est rien et, qu'au contraire, la théorie qui en résulte est très rigide, comme nous allons le voir tout au long de cette partie.

3.2.1 Le théorème d'Artin–Wedderburn

Une *K-algèbre à divisions* est une K-algèbre (non nécessairement commutative) dans laquelle tout élément non nul est inversible. Il est clair que toute algèbre à divisions est simple puisqu'un idéal contenant un élément inversible est nécessairement trivial.

Théorème 3.2.1 (Artin–Wedderburn). *Toute K-algèbre simple centrale A est isomorphe à $M_n(D)$ pour une certaine K-algèbre à divisions centrale D.*

Démonstration. Soit A une algèbre simple centrale sur k. On considère un A-module irréductible S. On définit $D' = \mathrm{End}_A(S)$ comme l'ensemble des endomorphismes A-linéaires de S. Clairement D est une K-algèbre. Soit $f \in D'$, $f \neq 0$. Le noyau de f est un sous-A-module strict de S. Par simplicité, on en déduit qu'il est nul, c'est-à-dire que f est injective. De même, l'image de f est un sous-A-module non nul de S et est donc égale à S tout entier. Ainsi f est surjective. Il s'ensuit que f est bijective. En résumé, nous venons de démontrer que tout élément non nul de D' est inversible, d'où on déduit que D' est une algèbre à divisions.

On munit S d'une structure de D'-espace vectoriel à gauche *via* la loi de multiplication externe $f \cdot s = f(s)$ pour $f \in D$ et $s \in S$. Il se trouve que la théorie des espaces vectoriels sur les algèbres à divisions est similaire à la théorie classique sur les corps (*cf* appendice A.2 pour plus de détails). En particulier, S admet une base sur D' et le choix d'une telle base identifie $\mathrm{End}_{D'}(S)$ à une algèbre de matrices sur l'algèbre à divisions opposée $D = (D')^{\mathrm{op}}$. Par ailleurs, pour tout $a \in A$, l'application $m_a : S \to S, s \mapsto as$ est dans $\mathrm{End}_{D'}(S)$. On obtient de cette manière un morphisme d'anneaux $\varphi : A \to \mathrm{End}_{D'}(S)$. Son noyau est un idéal bilatère de A. Par simplicité, on en déduit que φ est injective. Nous allons montrer que c'est en fait un isomorphisme.

Soit $f \in \mathrm{End}_{D'}(S)$. On veut montrer que f est la multiplication par un certain élément $a \in A$. Pour cela, on fixe une base $(v_1, \ldots, v_n)$ de S sur k. On note $\mathrm{End}_A(S^n)$ l'ensemble des endomorphismes A-linéaires de S^n. Concrètement, sachant que $\mathrm{End}_A(S) = D'$, on déduit que tout morphisme dans $\mathrm{End}_A(S^n)$ est de la forme :

$$(s_1, \ldots, s_n) \mapsto \left(\textstyle\sum_{i=1}^n d_{i,1}(s_i), \ldots, \sum_{i=1}^n d_{i,n}(s_i) \right)$$

pour des $d_{i,j} \in D'$. Autrement dit $\mathrm{End}_A(S^n)$ s'identifie canoniquement à $M_n(D')$. On considère l'application $f_V : S^n \to S^n$, $(s_1, \ldots, s_n) \mapsto (f(s_1), \ldots, f(s_n))$. On déduit de la description

précédente que f_V commute à tout morphisme de $\mathrm{End}_A(S^n)$ puisque, par définition, f est D'-linéaire et donc commute à tous les $d_{i,j}$.

Posons $v = (v_1, \ldots, v_n) \in S^n$ et notons Av le sous-A-module de S^n engendré par v. Étant donné un sous-ensemble I de $\{1, \ldots, n\}$, notons S_I le sous-ensemble de S^n formé des n-uplets $(s_1, \ldots, s_n)$ pour lesquels $s_i = 0$ dès que $i \in I$. Soit I un sous-ensemble de cardinal maximal de $\{1, \ldots, n\}$ tel que $Av \cap S_I = 0$. Montrons que $Av + S_I = S^n$. Pour cela, il suffit de vérifier que $S_i \subset Av + S_I$ pour tout i. Or par simplicité de S_i, l'intersection $S_i \cap (Av + S_I)$ est soit égale à S_i, soit égale à 0. Dans le premier cas, l'inclusion annoncée est prouvée tandis que, dans le second cas, on déduit $Av \cap S_{I\cup\{i\}} = 0$, ce qui contredit la maximalité de I. En conclusion, on obtient $Av \oplus S_I = S^n$, la somme étant directe puisque S_I est choisi de façon à ce que $Av \cap S_I = 0$. Soulignons que Av et S_I sont tous les deux des A-modules et que la décomposition ci-dessus est compatible à la structure de A-module.

Soit $\pi : S^n \to S^n$ la projection sur Av correspondant à la décomposition $S^n = Av \oplus S_I$ que l'on vient d'établir. Du fait que $\pi \in \mathrm{End}_A(S^n)$, on déduit que π commute à f_V et donc, en particulier, que $f_V(v_1, \ldots, v_n) = \pi \circ f_V(v_1, \ldots, v_n) \in Av$. Autrement dit, il existe $a \in A$ tel que $f(v_i) = av_i$ pour tout i. Comme la famille $(v_1, \ldots, v_n)$ engendre S sur k, on en déduit que f est la multiplication par a. Autrement dit, $f = \varphi(a)$ et on a ainsi trouvé un antécédent à f par φ.

Il résulte de ce qui précède que $A \simeq M_n(D)$. Il ne reste donc plus qu'à démontrer que le centre de D est K. Soit donc $d \in D$ un élément central. Du fait que d commute avec tous les éléments de D, on déduit plus généralement qu'il commute avec tous les matrices de $M_n(D)$, c'est-à-dire avec tous les éléments de A. Étant donné que A est centrale par hypothèse, on obtient finalement $d \in K$. □

Il est en général possible d'avoir, à peu de frais, des renseignements supplémentaires sur l'algèbre à divisions D promise par le théorème d'Artin–Wedderburn. Tout d'abord, il se trouve qu'elle est entièrement déterminée (à isomorphisme près) par A. C'est une conséquence de l'équivalence de Morita que nous exposerons au §3.2.3 (voir proposition 3.2.9).

Par ailleurs, si K est un corps fini, on a nécessairement $D = K$; c'est une conséquence directe du théorème de Wedderburn qui affirme que toute algèbre à divisions finie est commutative [4, Théorème III.5]. Lorsque $K = \mathbb{R}$, on peut montrer que $D = \mathbb{R}$ ou $D = \mathbb{H}$ (l'algèbre des quaternions introduites précédemment). Dans le cas général, les algèbres à divisions de centre K sont classifiées par ce que l'on appelle le *groupe de Brauer* de K. Il existe des outils puissants (de nature cohomologique) pour étudier ce groupe. On renvoie le lecteur à [10, 24] pour de nombreux compléments à ce sujet.

3.2.2 Extension des scalaires

Soit A une K-algèbre simple centrale. Si L est une extension de K, il est possible de former le produit tensoriel $L \otimes_K A$, qui est une L-algèbre. Concrètement, si $(\lambda_i)_{i\in I}$ est une base de L sur K, tout élément de $L \otimes_K A$ s'écrit de manière unique sous la forme $x = \sum_{i\in I} \lambda_i \otimes a_i$ pour une famille $(a_i)_{i\in I}$ d'éléments de A presque tous nuls. Le produit, quant à lui, se calcule « coordonnée par coordonnée » par la formule :

$$(\lambda \otimes a) \cdot (\mu \otimes b) = (\lambda\mu) \otimes (ab)$$

pour $\lambda, \mu \in L$ et $a, b \in A$.

Proposition 3.2.2. *Soit A une K-algèbre simple centrale. Alors $L \otimes_K A$ est une algèbre simple centrale sur L.*

Démonstration. Le fait que $L \otimes_K A$ soit centrale résulte aisément de la description explicite que nous avons donnée ci-dessus. Montrons à présent que $L \otimes_K A$ est simple. Soit $\mathcal{I}$ un idéal bilatère non nul de $L \otimes_K A$. Considérons un élément $x \in \mathcal{I}$, $x \neq 0$ que l'on écrit sous la forme

$x = \sum_{i\in I} \lambda_i \otimes x_i$ pour une famille $(x_i)_{i\in I}$ d'éléments de A presque tous nuls. Choisissons x de sorte que le nombre de x_i non nuls soit minimal. Fixons un indice i_0 pour lequel $x_{i_0} \neq 0$. L'idéal bilatère de A engendré par x_{i_0} est égal à A par simplicité. Ainsi il existe $a_1, \dots, a_m, a'_1, \dots, a'_m \in A$ tels que $\sum_{j=1}^m a_j x_{i_0} a'_j = 1$. L'élément $\sum_{j=1}^m (1 \otimes a_j) \cdot x \cdot (1 \otimes a'_j)$ appartient encore à $\mathcal{I}$ et a une composante en i_0 qui est égale à 1. Quitte à remplacer x par ce nouvel élément, on peut supposer que $x_{i_0} = 1$. Pour $a \in A$, examinons l'élément de $\mathcal{I}$ suivant :

$$(1 \otimes a) \cdot x - x \cdot (1 \otimes a) = \sum_{i\in I} \lambda_i \otimes (ax_i - x_i a).$$

Sa composante en i_0 s'annule puisqu'on a supposé $x_{i_0} = 1$. Par minimalité, les autres composantes doivent donc s'annuler aussi, *i.e* $x_i a = a x_i$ pour tout $i \in I$. Comme ceci vaut pour tout $a \in A$, on déduit que les x_i sont tous centraux. Étant donné que A est une K-algèbre centrale, cela signifie que $x_i \in K$ pour tout $i \in I$. Il résulte de cela que $x = \sum_{i\in I} \lambda_i \otimes x_i$ est dans L. Ainsi $L \otimes_K A$ est, elle aussi, centrale. □

Théorème 3.2.3. *Soit A une K-algèbre simple centrale. Il existe une extension finie L/K et un entier n tels que $L \otimes_K A \simeq M_n(L)$ (comme K-algèbres).*

Démonstration. Considérons une clôture algébrique $\bar{K}$ de K et formons le produit tensoriel $\bar{A} = \bar{K} \otimes_K A$. D'après la proposition 3.2.2, $\bar{A}$ est une $\bar{K}$-algèbre simple centrale et, par le théorème d'Artin–Wedderburn, est donc de la forme $M_n(D)$ pour une $\bar{K}$-algèbre à divisions D centrale.

Considérons $x \in D$ et notons $\bar{K}[x]$ la sous-algèbre de D engendré par x. Clairement $\bar{K}[x]$ est commutative et de dimension finie sur $\bar{K}$. De plus, elle est intègre car D est une algèbre à divisions. Ainsi $\bar{K}[x]$ est une extension de $\bar{K}$ et, comme $\bar{K}$ est algébriquement clos, on a nécessairement $\bar{K}[x] = \bar{K}$, *i.e.* $x \in \bar{K}$. En conclusion, on a démontré que $D = \bar{K}$ et donc que $\bar{A} \simeq M_n(\bar{K})$.

Fixons un isomorphisme $\bar{\varphi} : \bar{A} \simeq M_n(\bar{K})$ et notons M sa matrice écrite dans des K-bases de A et de $M_n(K)$ respectivement. Si L désigne l'extension finie de K engendrée par les coefficients de M, l'isomorphisme $\bar{\varphi}$ provient d'un isomorphisme $\varphi_L : L \otimes_K A \to M_n(L)$. Le théorème est démontré. □

Corollaire 3.2.4. *Soit A une K-algèbre simple centrale. La dimension de A sur K est un carré.*

Démonstration. Soit $\bar{A} = \bar{K} \otimes_K A$. Clairement $\dim_K A = \dim_{\bar{K}} \bar{A}$. Or, d'après le théorème 3.2.3, $\bar{A}$ est isomorphe à une algèbre de matrices sur $\bar{K}$, d'où on déduit que sa dimension est un carré. □

Afin de pouvoir utiliser toute la puissance de la théorie de Galois, il est important d'avoir une version plus précise du théorème 3.2.3 permettant de choisir l'extension L galoisienne sur K. Ce raffinement est l'objet du théorème de Wedderburn–Koethe, énoncé ci-dessous.

Théorème 3.2.5 (Wedderburn–Koethe)**.** *Soit A une K-algèbre simple centrale. Il existe une extension finie* séparable *L/K et un entier n tels que $L \otimes_K A \simeq M_n(L)$ (comme K-algèbres).*

Démonstration. Si K est de caractéristique nulle, le théorème a déjà été démontré puisque toute extension finie de K est automatiquement séparable. Nous supposons donc que K est de caractéristique p avec $p > 0$.

De même que dans la démonstration du théorème 3.2.3, il suffit de démontrer que si K est séparablement clos, il n'existe pas de K-algèbre à divisions non triviale. Soient donc K un corps séparablement clos et D une K-algèbre à divisions centrale que l'on suppose non triviale. Soit $x \in D$, $x \notin K$. Soit $K[x]$ le sous-anneau de D engendré par x. Il s'agit d'une extension finie de K. Comme, par hypothèse, K est séparablement clos, l'extension $K[x]/K$ est nécessairement purement inséparable. Autrement dit, il existe un entier m tel que $x^{p^m} \in K$.

Soit $\sigma : D \to D, d \mapsto xdx^{-1}$ l'application de conjugaison par x. D'après ce qui précède, on a $\sigma^{p^m} = 1$, soit encore $(\sigma - 1)^{p^m}$. Soit r l'indice de nilpotence de $(\sigma - 1)$; c'est un entier strictement supérieur à 1 puisque $x \notin K$. Soit $a \in K$ tel que $(\sigma - 1)^{r-1}(a) \neq 0$. Posons $u = (\sigma - 1)^{r-1}(a)$ et $v = (\sigma - 1)^{r-2}(a)$ de sorte que l'on ait $\sigma(u) = u$ et $\sigma(v) = u + v$. En posant $t = u^{-1}v \in D$, on trouve $\sigma(t) = t + 1$. Ainsi l'endomorphisme de K-algèbres $\sigma : D \to D$ stabilise le sous-corps commutatif $K[t]$ et agit non trivialement sur celui-ci. Ceci est une contradiction car K est supposé séparablement clos. □

3.2.3 La théorie de Morita

L'équivalence de Morita est un énoncé algébrique très général qui, étant donné un anneau R quelconque, non nécessairement commutatif, établit un lien étroit entre les modules sur $M_n(R)$ et ceux sur R.

Avant d'énoncer le résultat de Morita de manière précise, on a besoin d'introduire quelques objets. Dans tout ce numéro, on fixe un anneau R (non nécessairement commutatif) ainsi qu'un entier n strictement positif. On introduit le R-module $P = R^n$ que l'on identifie à l'espace des vecteurs colonne de taille n à coefficients dans R. On voit P à la fois comme un module à droite sur R et un module à gauche sur $M_n(R)$ *via* l'action naturelle $M \cdot x = Mx$ (multiplication matricielle). On note $P^\star$ le dual algébrique de P. Par définition, $P^\star = \mathrm{Hom}_R(P, R)$; c'est l'ensemble des applications R-linéaires de P dans R. Les éléments de $P^\star$ peuvent être écrits comme des vecteurs ligne, le vecteur ligne L étant en correspondance avec l'application linéaire $\ell : R^n \to R, X \mapsto LX$ (X étant vu ici comme un vecteur colonne). On muni $P^\star$ d'une structure de module à gauche sur R et de module à droite sur $M_n(R)$.

On rappelle par ailleurs qu'étant donné un anneau A, un module à droite M sur A et un module à gauche N sur A, le produit tensoriel $M \otimes_A N$ peut être défini : c'est l'ensemble des sommes formelles de tenseurs purs $x \otimes y$ (avec $x \in M$ et $y \in N$) modulo les relations $xa \otimes y = x \otimes ay$ pour $a \in A$, $x \in M$ et $y \in N$ ainsi que les relations usuelles d'additivité. Si de plus M (resp. N) est muni d'une structure supplémentaire de module à gauche (resp. à droite) sur un anneau auxiliaire B qui commute avec l'action de A, alors le produit tensoriel $M \otimes_A N$ hérite d'une structure de B-module à gauche (resp. à droite).

Lemme 3.2.6. *1. L'application :*

$$\begin{array}{rccc} \alpha : & P \otimes_R P^\star & \longrightarrow & M_n(R) \\ & C \otimes L & \mapsto & CL \end{array}$$

est un isomorphisme de modules à gauche et à droite sur $M_n(R)$.

2. L'application :

$$\begin{array}{rccc} \beta : & P^\star \otimes_{M_n(R)} P & \longrightarrow & R \\ & L \otimes C & \mapsto & LC \end{array}$$

est un isomorphisme de modules à gauche et à droite sur R.

Démonstration. Il est immédiat de vérifier la linéarité annoncée des applications α et β.

Soient $(E_1, \ldots, E_n)$ la base canonique de P et $(E_1^\star, \ldots, E_n^\star)$ sa base duale. La famille $(E_i \otimes E_j^\star)_{1 \leq i,j \leq n}$ est une base de $P \otimes_R P^\star$. Par ailleurs, un calcul direct montre que $E_i \otimes E_j^\star$ s'envoie par α sur la matrice E_{ij} dont tous les coefficients sont nuls à l'exception de celui en position (i, j) qui vaut 1. Le fait que α soit un isomorphisme en résulte sachant que la famille $(E_{ij})_{1 \leq i,j \leq n}$ est une base de $M_n(R)$.

Notons $L_0 = (1\,0\,0\,\cdots\,0) \in P^\star$ et $C_0 = L_0^T \in P$, le vecteur colonne transposé de L_0. Il est clair que $\beta(L_0 \otimes C_0) = 1$; ainsi β est surjectif. Soit à présent $X \in P^\star \otimes_{M_n(R)} P$ que l'on écrit comme une somme formelle $X = L_1 \otimes C_1 + L_2 \otimes C_2 + \cdots + L_m \otimes C_m$ (pour $m \in \mathbb{N}$, $L_i \in P^\star$ et $C_i \in P$). On suppose que $\beta(X) = 0$. Pour $i \in \{1, \ldots, m\}$, posons $u_i = L_i C_i \in R$ et considérons la

matrice M_i dont la première ligne est L_i et les autres lignes sont nulles. Un calcul direct aboutit à $L_0 M_i = L_i$ et $M_i C_i = C_0 u_i$, d'où on déduit :

$$L_i \otimes C_i = (L_0 M_i) \otimes C_i = L_0 \otimes (M_i C_i) = L_0 \otimes (C_0 u_i) = \big(L_0 \otimes C_0\big) \cdot u_i.$$

En sommant sur i, on obtient finalement $X = (L_0 \otimes C_0) \cdot (u_1 + \cdots + u_m) = (L_0 \otimes C_0) \cdot \beta(X) = 0$. On en déduit l'injectivité de β et le lemme est démontré. □

Corollaire 3.2.7 (Équivalence de Morita)**.** *Soient* $\mathrm{Mod}_R^{\mathrm{g}}$ *(resp.* $\mathrm{Mod}_{M_n(R)}^{\mathrm{g}}$*) la catégorie des modules à gauche sur* R *(resp. sur* $M_n(R)$*). Les foncteurs :*

$$\begin{pmatrix} \mathrm{Mod}_R^{\mathrm{g}} & \longrightarrow & \mathrm{Mod}_{M_n(R)}^{\mathrm{g}} \\ X & \mapsto & P \otimes_R X \end{pmatrix} \quad \textit{et} \quad \begin{pmatrix} \mathrm{Mod}_{M_n(R)}^{\mathrm{g}} & \longrightarrow & \mathrm{Mod}_R^{\mathrm{g}} \\ Y & \mapsto & P^\star \otimes_{M_n(R)} Y \end{pmatrix}$$

réalisent des équivalences de catégories inverses l'une de l'autre entre les catégories $\mathrm{Mod}_R^{\mathrm{g}}$ *et* $\mathrm{Mod}_{M_n(R)}^{\mathrm{g}}$.

Démonstration. Le lemme 3.2.6 permet de vérifier directement que la composée des foncteurs dans les deux sens est isomorphe à l'identité. □

Le cas particulier où R est un corps est particulièrement intéressant. En effet, dans ce cas, il est bien connu que deux K-espaces vectoriels sont isomorphes si et seulement s'ils ont le même dimension. On en déduit que deux $M_n(K)$-modules à gauche sont isomorphes si et seulement s'ils ont la même dimension en tant que K-espace vectoriel (noter que $\dim_K(P \otimes_K X) = n \cdot \dim_K X$). Ce résultat vaut encore si K est remplacé par une algèbre à divisions D, d'après les résultats de l'appendice A.2. Du théorème d'Artin–Wedderburn (*cf* théorème 3.2.1), on déduit le résultat encore plus général suivant :

Proposition 3.2.8. *Soit* A *une* K*-algèbre simple centrale. Deux* A*-modules à gauche sont isomorphes si et seulement s'ils ont la même dimension en tant que* K*-espaces vectoriels.*

Comme autre corollaire, l'équivalence de Morita implique l'unicité de l'écriture du théorème d'Artin–Wedderburn. Précisément :

Proposition 3.2.9. *Soient* D, D' *deux algèbres à divisions et* n, n' *deux entiers tels que* $M_n(D) \simeq M_{n'}(D')$*. Alors* $D \simeq D'$ *et* $n = n'$.

Démonstration. Posons $A = M_n(D)$. D'après la démonstration du théorème 3.2.1, il suffit de démontrer que tous les A-modules simples à gauche sont isomorphes. D'après le corollaire 3.2.7, ceci revient à démontrer que tous les D-espaces vectoriels de dimension 1 sont isomorphes, ce qui est bien le cas. □

Remarque 3.2.10. Bien entendu, l'équivalence de Morita est encore valable si l'on remplace partout « gauche » par « droite » : il existe une équivalence de catégories entre la catégorie des $M_n(R)$-modules à droite et celle des R-modules à droite. Il est possible également d'obtenir une version mixte qui stipule que la catégorie des $M_n(R)$-bimodules [1] est équivalente à celle des R-bimodules. Les foncteurs réalisant cette équivalence sont :

$$\begin{pmatrix} \mathrm{biMod}_R & \longrightarrow & \mathrm{biMod}_{M_n(R)} \\ X & \mapsto & P \otimes_R X \otimes_R P^\star \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} \mathrm{biMod}_{M_n(R)} & \longrightarrow & \mathrm{biMod}_R \\ Y & \mapsto & P^\star \otimes_{M_n(R)} Y \otimes_{M_n(R)} P \end{pmatrix}.$$

En particulier, on retrouve comme ceci que, si K est un corps, $M_n(K)$ ne possède pas d'idéaux bilatères non triviaux.

1. Si A est un anneau, un A-bimodule est un groupe additif muni d'une multiplication externe à gauche et d'une multiplication externe à droite par les éléments de A, de manière à ce que les multiplications à gauche commutent avec celles à droite.

3.2.4 Le théorème de Noether–Skolem

Le théorème de Noether–Skolem est un résultat de rigidité sur les endomorphismes d'une algèbre simple centrale. Il nous sera utile à plusieurs reprises dans la suite pour montrer que les constructions que nous allons élaborer ne dépendent d'aucun choix.

Théorème 3.2.11 (Noether–Skolem)**.** *Soit A une K-algèbre simple centrale. Tout endomorphisme de K-algèbres de A est de la forme $x \mapsto gxg^{-1}$ pour un certain élément inversible g de A.*

Le théorème de Noether–Skolem possède une reformulation dans le langage de la théorie des modules qui en donne un nouvel éclairage et qui nous sera utile dans la démonstration. Cette reformulation s'énonce comme suit.

Corollaire 3.2.12. *Soit D une algébre à divisions centrale sur K.*
Soit V un bimodule sur D. On suppose que tout pour $\lambda \in K$, les multiplications à gauche et à droite par λ coïncident. Si V est de dimension finie sur K, alors $V \simeq D^n$ comme D-bimodule avec $n = \frac{\dim_K V}{\dim_K D}$.

Commençons par démontrer que le théorème 3.2.11 pour $A = M_n(D)$ implique le corollaire 3.2.12 pour l'algèbre à divisions D et l'entier n. Considérons donc un D-bimodule V vérifiant les hypothèses de corollaire. On suppose que $\dim_K V = n \cdot \dim_K D$. Du fait que D est une algèbre à divisions, on déduit que $V \simeq D^n$ en tant que D-espace vectoriel à gauche. Soit $(e_1, \dots, e_n)$ une base de V sur D. Pour toute matrice $M = (m_{ij})_{1 \leq i,j \leq n} \in M_n(D)$, l'application $d_M : V \to V$ de multiplication à droite par M :

$$\sum_{i=1}^{n} d_i e_i \mapsto \sum_{i=1}^{n} \sum_{j=1}^{n} d_i e_j m_{ij}$$

est D-linéaire à gauche. Notons $\sigma(M)$ sa matrice dans la base $(e_1, \dots, e_n)$. Il est évident que l'application $\sigma : M_n(D) \to M_n(D)$ ainsi définie est un morphisme d'anneaux. Par le théorème 3.2.11, il existe une matrice inversible $G \in M_n(D)$ telle que $\sigma(M) = GMG^{-1}$ pour tout $M \in M_n(D)$. Autrement dit, l'application d_M s'identifie à la multiplication à *gauche* par M dans une certaine base $(g_1, \dots, g_n)$. Ainsi l'application D-linéaire à gauche $D^n \to V$ qui envoie la base canonique sur la base de (g_i) est un morphisme de bimodules et le corollaire est démontré.

Venons-en maintenant à la démonstration du théorème de Noether–Skolem. D'après le théorème d'Artin–Wedderburn, A est isomorphe à $M_n(D)$ pour un certain entier n et une certaine algèbre à divisions D. On peut donc supposer que A est de cette forme.

Montrons tout d'abord que, si le théorème de Noether–Skolem est vrai pour D, alors il l'est aussi pour $M_n(D)$ pour tout entier n. Pour ce faire, considérons un endomorphisme de K-algèbres $\sigma : M_n(D) \to M_n(D)$. Posons $P = D^n$ et munissons-le de deux structures de $M_n(D)$-modules :

- d'une part, de sa structure naturelle en faisant agir la matrice $M \in M_n(D)$ agit sur P par multiplication à gauche et
- d'autre part, de la structure naturelle twistée par σ obtenue en faisant agir $M \in M_n(D)$ agit par multiplication par $\sigma(M)$.

Pour distinguer ces deux modules, notons-les respectivement P_1 et P_2. Munissons également P_1 et P_2 de leur structure naturelle de D-module à droite. Soient V_1 et V_2 les D-espaces vectoriels à gauche associés respectivement à P_1 et P_2 par l'équivalence de Morita (*cf* corollaire 3.2.7). Ils héritent en outre d'une structure de D-espace vectoriel à droite, faisant d'eux des bimodules sur D. Par ailleurs, ils sont de dimension 1 sur D. Par le corollaire 3.2.12 appliqué avec $n = 1$, il existe un isomorphisme de bimodules $f : V_1 \to V_2$. *Via* l'équivalence de Morita, f correspond à un automorphisme D-linéaire à droite $\varphi : P \to P$ vérifiant :

$$\forall M \in M_n(D), \quad \forall x \in P, \quad \varphi(Mx) = \sigma(M) \cdot \varphi(x). \tag{24}$$

Comme φ est D-linéaire à droite, il est donné par la multiplication à gauche par une matrice $G \in M_n(D)$; de plus, G est inversible puisque φ est un automorphisme. Ainsi la relation (24) s'écrit encore $GMx = \sigma(M)Gx$. Comme ceci est vrai pour tout x, on trouve $GM = \sigma(M)G$, soit encore $\sigma(M) = GMG^{-1}$.

Il reste à démontrer le théorème de Noether–Skolem lorsque A est une algèbre à divisions D. Soit $\sigma : D \to D$ un morphisme de K-algèbres. D'après le théorème 3.2.3, il existe une extension finie L/K telle que $D_L = L \otimes_K D$ soit isomorphe à une algèbre de matrices $M_n(L)$. De plus $\sigma_L = \mathrm{id} \otimes \sigma$ définit un endomorphisme de L-algèbres de D_L. D'après ce que nous avons déjà démontré, il existe un élément $G_L \in D_L$ tel que $\sigma_L(X) = G_L X G_L^{-1}$ pour tout $X \in D_L$. En particulier, pour $X \in D$, on a la relation $\sigma(X)G_L = G_L X$. Fixons $(\lambda_1, \dots, \lambda_s)$ une base de L sur K et écrivons G_L sous la forme $G_L = \lambda_1 G_1 + \cdots + \lambda_s G_s$ avec $G_i \in D$. La relation $\sigma(X)G_L = G_L X$ entraîne que $\sigma(X)G_i = G_i X$ pour tout $i \in \{1, \dots, s\}$ et tout $X \in D$. Or, il existe au moins un indice i pour lequel G_i est non nul. Comme D est une algèbre à divisions, ce G_i est inversible et le théorème de Noether–Skolem est démontré.

3.3 La norme réduite

Les théorèmes de rigidité que nous avons présentés précédemment permettent d'étendre certaines constructions classiques dans le cas des matrices au contexte des algèbres simples centrales. Dans ce numéro, nous utilisons ce yoga pour étendre la notion de déterminant.

3.3.1 Construction par descente galoisienne

Fixons une algèbre simple centrale A. Par le théorème de Wedderburn–Koethe, il existe une extension finie séparable L de K et un isomorphisme de L-algèbres $f : L \otimes_K A \xrightarrow{\sim} M_n(L)$ pour un certain entier n. Quitte à étendre L, on peut supposer que l'extension L/K est galoisienne, ce que l'on fait à partir de maintenant. Intéressons-nous à l'application composée :

$$N_{\mathrm{rd}} : A \hookrightarrow L \otimes_K A \xrightarrow{f} M_n(L) \xrightarrow{\det} L.$$

Lemme 3.3.1. *L'application N_{rd} prend ses valeurs dans K.*

Démonstration. Posons $A_L = L \otimes_K A$ et considérons $\sigma \in \mathrm{Gal}(L/K)$. Clairement σ induit des automorphismes de A_L et de $M_n(L)$ que, par abus de notations, nous continuons à appeler σ. Une vérification immédiate montre que le commutateur $\sigma f \sigma^{-1} f^{-1} : M_n(L) \to M_n(L)$ est un isomorphisme de L-algèbres. Par le théorème de Noether–Skolem, il existe une matrice $G \in \mathrm{GL}_n(L)$ telle que $\sigma f \sigma^{-1} f^{-1}(M) = GMG^{-1}$ pour tout $M \in M_n(L)$. En prenant le déterminant, on trouve :

$$\forall M \in M_n(L), \quad \det(\sigma f \sigma^{-1} f^{-1}(M)) = \det(M).$$

En appliquant ceci avec $M = f(x)$ pour $x \in A$, on obtient $\sigma N_{\mathrm{rd}}(x) = N_{\mathrm{rd}}(x)$ (noter que σ commute avec det). Comme ceci est vrai pour tout $\sigma \in \mathrm{Gal}(L/K)$, on conclut que $N_{\mathrm{rd}}(x) \in K$. □

Lemme 3.3.2. *L'application N_{rd} ne dépend pas de la trivialisation f (ni de l'extension L).*

Démonstration. Considérons deux trivialisations $f : L \otimes_K A \to M_n(L)$ et $f' : L' \otimes_K A \to M_n(L')$. Il s'agit de montrer que les deux applications $\det \circ f$ et $\det \circ f'$ coïncident sur A. Choisissons une troisième extension F de K contenant à la fois L et L' et notons encore $f, f' : M \otimes_K A \to M_n(F)$ les applications déduites respectivement de f et f' après extension des scalaires à F. La composée $f' \circ f^{-1} : M_n(F) \to M_n(F)$ est alors un morphisme de K-algèbres ; d'après le théorème de Skolem–Noether, il est donc de la forme $X \mapsto GXG^{-1}$ pour une certaine matrice $G \in GL_n(F)$. En particulier, pour tout $x \in A$, on a $f'(x) = Gf(x)G^{-1}$, ce qui conduit directement à $\det f'(x) = \det f(x)$ en prenant le déterminant. □

Il résulte des lemmes 3.3.1 et 3.3.2 que N_{rd} définit une application canonique de A dans K ; on l'appelle la *norme réduite* de A. Elle vérifie les propriétés suivantes, héritées directement des propriétés classiques du déterminant :

– pour $x \in K$, on a $N_{\text{rd}}(x) = x^n$ (où n est défini par $[A : K] = n^2$),
– pour $x, y \in A$, on a $N_{\text{rd}}(xy) = N_{\text{rd}}(x)N_{\text{rd}}(y)$.

Dans le cas matriciel, une matrice M et son déterminant sont liées par la relation

$$M \cdot \text{adj}(M) = \det(M)$$

où $\text{adj}(M)$ désigne la matrice adjointe de M définie comme la transposée de la comatrice de M (*cf* appendice A.1 pour de nombreux compléments sur la notion d'adjoint). Il se trouve que cette relation s'étend au cas des algèbres simples centrales. Plus précisément, étant données une algèbre simple centrale A et une trivialisation $f : L \otimes_K A \to M_n(L)$ (pour une extension L/K que l'on peut supposer galoisienne), on définit l'application adj sur A comme la composée :

$$\text{adj} : A \hookrightarrow L \otimes_K A \xrightarrow{f} M_n(L) \xrightarrow{\text{adj}} M_n(L) \xrightarrow{f^{-1}} L \otimes_K A.$$

De même que pour la norme réduite, on démontre que adj prend ses valeurs dans A et ne dépend pas de la trivialisation choisie, définissant ainsi une application canonique $\text{adj} : A \to A$ qui vérifie la propriété $x \cdot \text{adj}(x) = \text{adj}(x) \cdot x = N_{\text{rd}}(x)$ pour tout $x \in A$. En particulier, on observe que x est toujours un diviseur à gauche et à droite de sa norme réduite $N_{\text{rd}}(x)$.

Exemple 3.3.3. Dans le cas particulier de l'algèbre de quaternions $\mathbb{H}$, la trivialisation

$$\begin{array}{rcl} \mathbb{C} \otimes_{\mathbb{R}} \mathbb{H} & \longrightarrow & M_2(\mathbb{C}) \\ a + bi + cj + dk & \mapsto & \begin{pmatrix} a + b\sqrt{-1} & -c + d\sqrt{-1} \\ c + d\sqrt{-1} & a - b\sqrt{-1} \end{pmatrix} \end{array}$$

montre que la norme réduite du quaternion $a + bi + cj + dk$ ($a, b, c, d \in \mathbb{R}$) est $a^2 + b^2 + c^2 + d^2$. Par ailleurs, son adjoint est le quaternion associé à la matrice adjointe qui vaut

$$\begin{pmatrix} a - b\sqrt{-1} & c - d\sqrt{-1} \\ -c - d\sqrt{-1} & a + b\sqrt{-1} \end{pmatrix}.$$

On trouve ainsi $\text{adj}(a + bi + cj + dk) = a - bi - cj - dk$.

Remarque 3.3.4. De manière analogue, on peut encore définir la trace réduite d'un élément de A ou, plus généralement, son polynôme caractéristique. Nous n'insistons pas davantage car cela ne nous sera pas utile pour la suite. Cependant, nous encouragerons la lectrice et le lecteur interessés à faire l'exercice.

3.3.2 Une définition alternative

Il existe une seconde construction de la norme réduite qui ne fait pas apparaître aussi clairement le lien avec le déterminant matriciel mais qui aura l'avantage de se généraliser plus simplement par la suite.

On considère toujours une K-algèbre simple centrale. Si C est une sous-algèbre de A, on peut évidemment voir A comme une C-algèbre et donc, en particulier, comme un C-module à gauche. On se donne une sous-K-algèbre *commutative* C de A pour laquelle A est un C-module libre. Sous ces hypothèses, si $x \in A$, on peut définir sa norme sur C comme suit. On considère l'application $m_x : A \to A, a \mapsto ax$ de multiplication à droite par x. Elle est C-linéaire et cela a donc un sens de considérer son déterminant. On définit $N_{A/C}(x) = \det m_x$. Nous allons démontrer dans la suite que, sous certaines hypothèses, l'application $N_{A/C}$ coïncide avec la norme réduite. Cependant, avant d'énoncer un résultat précis, nous nous proposons d'examiner quelques exemples qui, en plus de faciliter la compréhension, nous seront utiles pour la preuve.

Exemple 3.3.5. Commençons par l'algèbre des quaternions qui est plus simple. Clairement $\mathbb{H}$ possède une sous-algèbre commutative isomorphe à $\mathbb{C}$, c'est celle engendrée par 1 et i. De plus $\mathbb{H}$ est de dimension 2 sur $\mathbb{C}$. On est donc dans les conditions d'applications des propositions 3.3.9 et 3.3.10. Une base de $\mathbb{H}$ sur $\mathbb{C}$ est, par exemple, la famille $(1, j)$ puisque tout quaternion $a+bi+cj+dk$ s'écrit encore $(a+bi)+(c+di)j$ (on prendra garde à bien mettre les « scalaires » à gauche et les éléments de la base à droite). Soit $x = a + bi + cj + dk \in \mathbb{H}$. La matrice de la multiplication à droite par x dans la base $(1, j)$ s'écrit :

$$M_x = \begin{pmatrix} a+bi & -c+di \\ c+di & a-bi \end{pmatrix}.$$

Ainsi $N_{\mathbb{H}/\mathbb{C}}(M_x) = a^2+b^2+c^2+d^2$ et on retrouve bien, comme ceci, la norme réduite de l'élément x. (La lectrice assidue aura remarqué que le calcul que nous venons de faire est exactement le même que celui que nous avions déjà fait au §3.1 lorsque nous avions introduit $\mathbb{H}$.) D'autre part, pour $x \neq 0$, la matrice adjointe de M_x est :

$$\mathrm{adj}(M_x) = \begin{pmatrix} a-bi & c-di \\ -c-di & a+bi \end{pmatrix}$$

et on reconnaît la multiplication par l'élément $a - bi - cj - dk$ qui est exactement l'adjoint de x.

Il s'avère que $\mathbb{H}$ possède de nombreuses autre sous-algèbres commutatives de dimension 2. Précisément, si $x_0 \in \mathbb{H}\backslash\mathbb{R}$, la sous-algèbre $\mathbb{R}[x_0] \subset \mathbb{H}$ est de dimension 2 et isomorphe à $\mathbb{C}$. Par exemple, pour $x_0 = 1 + i + j$, on trouve $x_0^2 = 2x_0 - 3$, ce qui implique que $\mathbb{R}[x_0] \simeq \mathbb{C}$ *via* le morphisme envoyant x_0 sur $\frac{1+\sqrt{-2}}{2} \in \mathbb{C}$. Regardons brièvement ce qui se serait passé si nous avions choisi $C = \mathbb{R}[x_0]$. La famille $(1, i)$ est une base de $\mathbb{H}$ sur C et un quaternion quelconque se décompose sur cette base comme suit :

$$a + bi + cj + dk = (a - c - d + cx_0) + (b - c + d - dx_0)i.$$

Dans la base $(1, i)$, la multiplication à droite par $x = a + bi + cj + dk$ a pour matrice :

$$\begin{pmatrix} a-c-d+cx_0 & -b+d-c-dx_0 \\ b-c+d-dx_0 & a+d+c-cx_0 \end{pmatrix}$$

dont le déterminant est $a^2 + b^2 + c^2 + d^2$. On retrouve ainsi à nouveau la norme réduite de x.

Exemple 3.3.6. Soit $\mathcal{D}$ la sous-algèbre de $M_n(K)$ formée des matrices diagonales. Clairement $\mathcal{D}$ est commutative. Pour $j \in \{1, \ldots, n\}$, définissons la matrice $E_j \in M_n(K)$ dont tous les coefficients sont nuls, exceptés ceux de la j-ième colonne qui sont égaux à 1. Pour $M = (m_{i,j})_{1\leq i,j\leq n} \in M_n(K)$, on a la décomposition :

$$M = \sum_{j=1}^{n} \mathrm{Diag}(m_{1,j}, m_{2,j}, \ldots, m_{n,j}) \cdot E_j$$

où $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ désigne la matrice diagonale dont le coefficient en position (i, i) est λ_i. On en déduit que la famille des E_j forme une base de $M_n(K)$ sur $\mathcal{D}$. Soit à présent $M = (m_{i,j})_{1\leq i,j\leq n} \in M_n(K)$. Un calcul simple montre que la matrice de la multiplication à droite par M sur $M_n(K)$ dans la base $(E_1, \ldots, E_n)$ est $\big(\mathrm{Diag}(m_{j,i}, \ldots, m_{j,i})\big)_{1\leq i,j\leq n}$, c'est-à-dire la transposée de M. Son déterminant est donc égal à celui de M, *i.e.* à $N_{\mathrm{rd}}(M)$.

Exemple 3.3.7. Toujours dans le cas des algèbres de matrices, il est possible de choisir une autre algèbre, qui sera notée $\mathcal{T}$, pour lequel la coïncidence remarquable $N_{A/\mathcal{T}} = N_{\mathrm{rd}}$ est encore valable.

Il s'agit de l'algèbre consitituée des matrices triangulaires inférieures de la forme :

$$T_n(a_1, a_2, \ldots, a_n) = \begin{pmatrix} a_1 & & & & \\ a_2 & a_1 & & & \\ a_3 & a_2 & a_1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ a_n & \cdots & a_3 & a_2 & a_1 \end{pmatrix} \quad \text{avec} \quad a_1, \ldots, a_n \in K.$$

Il est facile de vérifier que $\mathcal{T}$ est la sous-algèbre de $M_n(K)$ engendrée par la matrice nilpotente $T_n(0, 1, 0, \ldots, 0)$. En particulier, il s'agit d'une algèbre commutative. Soit $\iota : K \to \mathcal{T}$, $a \mapsto T_n(a, 0, \ldots, 0)$ le morphisme définissant la structure de K-algèbre sur $\mathcal{T}$. Pour $i \in \{1, \ldots, n\}$, considérons la matrice $B_i \in M_n(K)$ dont tous les coefficients sont nulles à l'exception de celui à la position $(1, i)$ qui est égale à 1. Pour $M \in M_n(K)$, un calcul immédiat conduit à :

$$M = T_n(C_1)B_1 + T_n(C_2)B_2 + \cdots + T_n(C_n)B_n$$

où C_i désigne le i-ième vecteur colonne de M. De plus, une telle décomposition est unique. Il en résulte que la famille $(B_1, \ldots, B_n)$ est une base de $M_n(K)$ sur T. Par ailleurs, étant donnée une matrice $M = (m_{i,j})_{1 \leq i,j \leq n} \in M_n(K)$, on a $B_i M = \iota(m_{i,1})B_1 + \iota(m_{i,2})B_2 + \cdots + \iota(m_{i,n})B_n$. Autrement dit, la matrice (dans la base $(B_1, \ldots, B_n)$) de l'application de multiplication à droite par M n'est autre que la transposée de $\iota(M)$. En particulier, on en déduit que $N_{M_n(K)/\mathcal{T}}(M) = N_{\mathrm{rd}}(M)$.

Revenons-nous à présent au cas général. On rappelle qu'on a fixé une K-algèbre simple centrale A. D'après le corollaire 3.2.4, la dimension sur A sur K est un carré ; écrivons donc $[A : K] = n^2$. À partir de maintenant, nous fixons une sous-K-algèbre commutative C de A sur laquelle nous faisons l'hypothèse suivante.

Hypothèse 3.3.8.

(1) En tant que C-module, A est libre de rang n.

(2) En tant que $\bar{K}$-algèbre (où $\bar{K}$ désigne une clôture algébrique de K), $\bar{K} \otimes_K C$ est engendrée par un unique élément.

Proposition 3.3.9. *Si C vérifie l'hypothèse 3.3.8, alors $N_{A/C}(x) = N_{rd}(x)$ pour tout $x \in A$.*

Démonstration. Soit $\bar{K}$ une clôture algébrique de K et soient $\bar{A} = \bar{K} \otimes_K A$ et $\bar{C} = \bar{K} \otimes_K C$. D'après la proposition 3.2.2, $\bar{A}$ est une $\bar{K}$-algèbre simple centrale. De plus, si C vérifie l'hypothèse 3.3.8 alors il en est de même pour la sous-$\bar{K}$-algèbre $\bar{C}$ de $\bar{A}$. En outre, $N_{A/C}$ et $N_{\bar{A}/\bar{C}}$ coïncident sur A et, de même, la norme réduite de $\bar{A}$ coïncide avec celle de A en restriction sur A. Ainsi, pour démontrer la proposition, on peut supposer que K est algébriquement clos.

Sous cette hypothèse supplémentaire, le théorème 3.2.3 entraîne que $A \simeq M_n(K)$. Dans le reste de la démonstration, on identifie sans commentaire supplémentaire ces deux anneaux. Soit c un générateur de $M_n(K)$. Comme K est algébriquement clos, le théorème de décomposition de Jordan indique que c est semblable à une matrice diagonale par blocs de la forme :

$$\begin{pmatrix} T_{n_1}(a_1, 1, 0, \ldots, 0) & & & \\ & T_{n_2}(a_2, 1, 0, \ldots, 0) & & \\ & & \ddots & \\ & & & T_{n_k}(a_k, 1, 0, \ldots, 0) \end{pmatrix}.$$

où la notation T_{n_i} est empruntée à l'exemple 3.3.6 et où les n_i sont des entiers strictement positifs dont la somme vaut n. Bien sûr, dans le cas où $n_i = 1$, on convient que la matrice $T_1(a_i, 1, 0, \ldots, 0)$ est la matrice 1×1 dont l'unique coefficient est a_i. On en déduit que, quitte à

modifier l'isomorphisme entre A et $M_n(K)$, C est ainsi incluse dans la sous-algèbre $\mathcal{T}$ de $M_n(K)$ formée des matrices de la forme :

$$T(\alpha_1,\dots,\alpha_n)=\begin{pmatrix} T_{n_1}(\alpha_1,\alpha_2,\dots,\alpha_{s_1}) & & & \\ & T_{n_2}(\alpha_{s_1+1},\alpha_{s_1+2},\dots,\alpha_{s_2}) & & \\ & & \ddots & \\ & & & T_{n_k}(\alpha_{s_{k-1}+1},\alpha_{s_{k-1}+2},\dots,\alpha_n) \end{pmatrix}.$$

où les α_i parcourent K et où $s_j=n_1+\cdots+n_j$ (pour $1\leq j\leq k$) par définition. De plus, le (1) de l'hypothèse 3.3.8 implique que C est de dimension n sur K. Comme il en est de même de $\mathcal{T}$, on obtient $C=\mathcal{T}$.

La fin de la démonstration reprend désormais l'argumentation développée dans les exemples 3.3.5 et 3.3.6. Pour $i\in\{1,\dots,n\}$, on note $B_i\in M_n(K)$ la matrice dont tous les coefficients sont nuls sauf ceux en position $(1+s_{j-1},\ i)$ (pour $1\leq j\leq k$ et où, par convention, $s_0=1$) qui valent 1. Un calcul sans difficultés aboutit, pour tout matrice $M\in M_n(K)$, à :

$$M=T(C_1)B_1+\cdots+T(C_n)B_n$$

où C_i désigne la i-ième colonne de M. La famille $(B_1,\dots,B_n)$ est ainsi une base de $M_n(K)$ sur C. De plus, étant donnée une matrice $M\in M_n(K)$, on vérifie que, dans cette base, la matrice de la multiplication à droite par M s'identifie à la transposée de M. La proposition en découle. □

De la même manière, voyant toujours m_x comme un endomorphisme C-linéaire de A, on peut considérer son adjoint : il s'agit d'une application C-linéaire $\mathrm{adj}(m_x):A\to A$ vérifiant l'égalité $m_x\circ\mathrm{adj}(m_x)=\mathrm{adj}(m_x)\circ m_x=N_{\mathrm{rd}}(x)\mathrm{id}_A$.

Proposition 3.3.10. *Avec les notations et les hypothèses précédentes, l'application $\mathrm{adj}(m_x)$ est la multiplication à droite par $\mathrm{adj}(x)$. En particulier $\mathrm{adj}(x)=\mathrm{adj}(m_x)(1)$.*

Démonstration. Lorsque x est inversible dans A, on peut écrire $\mathrm{adj}(x)=N_{\mathrm{rd}}(x)\cdot x^{-1}$ et $\mathrm{adj}(m_x)=\det(m_x)\cdot m_x^{-1}$. Étant donné que $\det(m_x)=N_{\mathrm{rd}}(x)$ (par la proposition 3.3.9), on en déduit que $\mathrm{adj}(m_x)$ et la multiplication à droite par $\mathrm{adj}(x)$ coïncident tous deux avec $\det(m_x)\cdot m_x^{-1}$ et donc sont égaux. Le cas général s'en déduit par un argument de continuité pour la topologie de Zariski. □

3.4 Les algèbres d'Azumaya

La notion d'algèbre d'Azumaya est l'extension de celle d'algèbre simple centrale où cas où la base n'est plus nécessairement un corps mais un anneau commutatif quelconque.

Définition 3.4.1. Soit Z un anneau commutatif. Une Z-algèbre A est une *algèbre d'Azumaya* si pour tout idéal premier $\mathfrak{p}$ de Z, le produit tensoriel $\mathrm{Frac}(Z/\mathfrak{p})\otimes_Z A$ est une algèbre simple centrale sur $\mathrm{Frac}(Z/\mathfrak{p})$.

Remarque 3.4.2. Si Z est un corps, l'unique idéal premier de Z est l'idéal nul ; les notions d'algèbre d'Azumaya et d'algèbre simple centrale coïncident donc dans ce cas.

Dans le langage de la géométrie algébrique, la défintion d'une algèbre d'Azumaya est de nature locale ; en effet, il s'avère qu'une algèbre d'Azumaya est une algèbre qui est localement une algèbre simple centrale pour la topologie de Zariski. Mieux encore, on peut montrer qu'une algèbre d'Azumaya est une algèbre qui est localement isomorphe à une algèbre de matrices pour la topologie étale [2]. Pour cette raison, étudier les algèbres d'Azumaya dans le cadre conceptuel de la topologie étale s'avère particulièrement puissant. C'est, par exemple, le chemin qui est

2. La raison en est que les extensions séparables sont des recouvrements pour la topologie étale ; le théorème de Wedderburn–Koethe nous enseigne donc que toute algèbre simple centrale se trivialise sur un recouvrement étale.

suivi dans [11]. Toutefois, pour ce cours, nous avons préféré suivre une voie plus pédestre qui n'utilise pas la topologie étale mais, en contrepartie, ne fonctionne que sous certaines hypothèses restrictives (qui seront toutefois vérifiées dans les applications que nous avons en vue).

Dans toute la suite, on fixe un anneau commutatif Z et une algèbre d'Azumaya A sur Z. On suppose qu'il existe une sous-Z-algèbre C de A de façon à ce que les hypothèses suivantes sont vérifiées :

Hypothèse 3.4.3.

(1) L'anneau Z n'a pas d'éléments nilpotents,

(2) Pour tout idéal premier $\mathfrak{p}$ de A, la $\mathrm{Frac}(Z/\mathfrak{p})$-sous-algèbre $\mathrm{Frac}(Z/\mathfrak{p}) \otimes_Z C$ de $\mathrm{Frac}(Z/\mathfrak{p}) \otimes_Z A$ vérifie l'hypothèse 3.3.8,

(3) C est libre comme Z-module et admet une base de la forme $(1, c_2, c_3, \ldots, c_m)$.

Ces hypothèses ne sont pas automatiques mais elles seront vérifiées dans les cas que nous considèrerons dans la suite de ce cours. Notons également qu'il peut exister — et, qu'en pratique, il existera — plusieurs sous-algèbres C satisfaisant les conditions requises. À ce propos, être capable de jongler avec différents C s'avèrera fructueux à plusieurs reprises.

Sous l'hypothèse 3.4.3, on peut définir une application de norme réduite en copiant la seconde définition que nous avons vue dans le cas des algèbres simples centrales. Étant donné une sous-algèbre commutative C de A vérifiant l'hypothèse 3.4.3, on définit $N_{A/C}(x)$ comme le déterminant de l'application C-linéaire $m_x : A \to A, a \mapsto ax$.

Proposition 3.4.4. *Pour un élément $x \in A$ donné, $N_{A/C}(x)$ appartient à Z et ne dépend pas du choix de C vérifiant l'hypothèse 3.4.3.*

Démonstration. Soit C une sous-algèbre de A vérifiant les conditions de l'hypothèse 3.4.3. Soit $(1, c_2, \ldots, c_m)$ une base de C sur Z. Soit $\mathfrak{p}$ un idéal premier de Z. Posons $Z_\mathfrak{p} = \mathrm{Frac}(Z/\mathfrak{p})$, $A_\mathfrak{p} = Z_\mathfrak{p} \otimes_Z A$ et $C_\mathfrak{p} = Z_\mathfrak{p} \otimes_Z C$. Par hypothèse, la proposition 3.3.9 s'applique à la $Z_\mathfrak{p}$-algèbre simple centrale $A_\mathfrak{p}$ munie de la sous-algèbre $C_\mathfrak{p}$, donnant ainsi la congruence :

$$N_{A/C}(x) \equiv N_{\mathrm{rd}}(x \bmod \mathfrak{p}) \pmod{\mathfrak{p}} \tag{25}$$

où $x \bmod \mathfrak{p}$ désigne l'image de x dans $A_\mathfrak{p}$. Par ailleurs, de part sa définition, $N_{A/C}(x)$ est un élement de C. Il s'écrit donc de manière unique sous la forme $N_{A/C}(x) = \lambda_1 + \lambda_2 c_2 + \cdots + \lambda_m c_m$ avec $\lambda_i \in Z$ pour tout i. La congruence (25) implique que $\lambda_i \equiv 0 \pmod{p}$ pour tout $i \geq 2$. Comme ceci est vrai pour tout idéal premier $\mathfrak{p}$, tous les λ_i pour $i \geq 2$ appartiennent à l'intersection de tous les idéaux premiers de Z et sont donc nuls puisqu'on a supposé que Z n'a pas d'éléments nilpotents. On en déduit que $N_{A/C}(x) \in Z$.

Montrons maintenant que $N_{A/C}(x)$ ne dépend pas du choix du C. Soient donc C_1 et C_2 deux sous-algèbres de A vérifiant les conditions de l'hypothèse 3.4.3. La congruence (25) est alors valable pour C_1 et C_2, d'où on déduit que $N_{A/C_1}(x) \equiv N_{A/C_2}(x) \pmod{\mathfrak{p}}$ pour tout idéal premier $\mathfrak{p}$ de Z. La différence $N_{A/C_1}(x) - N_{A/C_2}(x)$ appartient donc à tous les idéaux premiers de Z et est donc nulle. Autrement dit $N_{A/C_1}(x) = N_{A/C_2}(x)$. □

Grâce à la proposition 3.4.4, on dispose d'une application bien définie $N_{\mathrm{rd}} : A \to Z$ (construite comme l'application norme $N_{A/C}$ pour un certain C vérifiant l'hypothèse 3.4.3). De manière similaire, on construit une application $\mathrm{adj}_C : A \to A$ en faisant correspondre à $x \in A$ l'image de 1 par l'application adjointe $\mathrm{adj}(m_x)$. En reprenant les arguments de la démonstration de la proposition 3.4.4, on montre que

$$\mathrm{adj}_C(x) \equiv \mathrm{adj}(x \bmod \mathfrak{p}) \pmod{\mathfrak{p}} \tag{26}$$

où l'application adj qui apparaît dans le membre de droite est l'adjoint dans l'algèbre simple centrale $\mathrm{Frac}(Z/\mathfrak{p}) \otimes_Z A$. On en déduit que $\mathrm{adj}_C(x)$ ne dépend pas du choix de C. À partir de maintenant, on la notera simplement adj. Les applications N_{rd} et adj vérifient les propriétés suivantes.

Proposition 3.4.5. *Pour deux éléments $x, y \in A$ et un idéal premier $\mathfrak{p}$ de Z, on a :*

(i) *si $x \in Z$, alors $N_{rd}(x) = x^n$ et $adj(x) = x^{n-1}$,*

(ii) *$N_{rd}(xy) = N_{rd}(x)N_{rd}(y)$ et $adj(xy) = adj(y)adj(x)$,*

(iii) *$x \cdot adj(x) = adj(x) \cdot x = N_{rd}(x)$,*

(iv) *si $x \equiv y \pmod{\mathfrak{p}A}$, alors $N_{rd}(x) \equiv N_{rd}(y) \pmod{\mathfrak{p}}$, et $adj(x) \equiv adj(y) \pmod{\mathfrak{p}A}$.*

En particulier x apparaît comme un diviseur à gauche et à droite de sa norme réduite.

Démonstration. Tout découle directement des congruences (25)–(26), du fait que les propriétés (i)–(iii) sont vérifiées pour les algèbres simples centrales et du fait que si deux éléments de A (resp. de Z) sont égaux si et seulement s'ils sont congrus modulo $\mathfrak{p}$ (resp. modulo $\mathfrak{p}A$) pour tout idéal premier $\mathfrak{p}$ de Z. □

4 Polynômes de Ore et algèbres d'Azumaya

Dans cette partie, nous faisons le lien entre les anneaux de polynômes de Ore que nous avions introduits au §1 et les notions d'algèbres simples centrales et d'algèbres d'Azumaya étudiées au §3. Précisément, étant donné un corps K, nous démontrons, sous certaines hypothèses de finitude et à quelques modifications éventuelles près, que les anneaux de Ore $K[X, \theta, \partial]$ sont des algèbres d'Azumaya sur leur centre. Nous déduisons ensuite de cette observation une nouvelle panoplie d'outils puissants pour étudier de façon détaillée les propriétés de factorisation dans $K[X, \theta, \partial]$.

Dans toute cette partie, on continue de travailler avec un anneau de base K qui est un corps. Quitte à appliquer un twist d'Hilbert (voir §1.2), on peut supposer que soit $\partial = 0$, soit $\theta = \mathrm{id}$. C'est ce que nous ferons par la suite.

4.1 Description du centre

Une étape préliminaire nécessaire, avant de mettre en œuvre la stratégie très brièvement esquissée dans l'introduction de cette partie, ci-dessus, est de décrire le centre des anneaux de Ore $K[X, \theta]$ et $K[X, \partial]$. C'est ce à quoi nous nous attelons dans ce numéro.

4.1.1 Le cas de $K[X, \theta]$

Soit $\theta : K \to K$ un endomorphisme d'anneaux fixé. On désigne par F le sous-ensemble de K formé des $x \in K$ tels que $\theta(x) = x$. On vérifie sans peine que F est un sous-corps de K. De plus, étant donné que θ est l'identité sur F, le sous-anneau de $K[X, \theta]$ formé des polynômes de Ore à coefficients dans F est commutatif ; plus précisément, il est isomorphe à $F[X]$.

Proposition 4.1.1. *Le centre de $K[X, \theta]$ est F si θ est d'ordre infini et $F[X^r]$ si θ est d'ordre $r \in \mathbb{N}$.*

Démonstration. Soit $P = a_0 + a_1 X + \cdots + a_n X^n$ un élément central de $K[X, \theta]$. Un calcul donne :

$$P \cdot X - X \cdot P = \sum_{i=1}^{n} \big(\theta(a_i) - a_i\big) X^i.$$

Ainsi, comme P est supposé central, le polynôme ci-dessus s'annule, d'où on déduit que $\theta(a_i) = a_i$ pour tout i. Autrement dit $P \in F[X]$. D'autre part, pour $a \in K$, on a :

$$P \cdot a - a \cdot P = \sum_{i=1}^{n} a_i \big(a - \theta^i(a)\big) X^i.$$

Comme précédemment, on en déduit que, dès lors que a_i ne s'annule pas, on doit avoir $\theta^i(a) = a$ pour tout $a \in K$, *i.e.* $\theta^i = \mathrm{id}$. La proposition en résulte. □

Exemple 4.1.2. La proposition 4.1.1 s'applique directement aux cas des anneaux de Ore $\mathbb{C}[X, \mathrm{conj}]$ et $\mathbb{F}_{p^r}[X, \mathrm{Frob}]$ et indique que les centres de ces anneaux sont respectivement $\mathbb{R}[X^2]$ et $\mathbb{F}_p[X^r]$.

4.1.2 Le cas de $K[X, \partial]$

Soit $\partial : K \to K$ une dérivation. On note F le sous-corps des constantes de K, c'est-à-dire le sous-corps de F formé des éléments $x \in K$ tels que $\partial(x) = 0$. Le sous-anneau de $K[X, \theta]$ formé des polynômes de Ore à coefficients dans F est commutatif et isomorphe à un anneau de polynômes commutatif ; on le notera $F[X]$.

Lorsque K est de caractéristique $p > 0$, suivant les constructions de l'appendice A.3, on introduit l'anneau de Ore $F[Y, \mathrm{Frob}]$ où $\mathrm{Frob} : F \to F, x \mapsto x^p$ est le morphisme de Frobenius absolu. Cet anneau agit de manière naturelle sur l'ensemble des dérivations de K *via* la formule $P \cdot d = P_{\mathrm{lin}}(d)$ où P_{lin} est le polynôme linéarisé associé à P défini dans l'exemple 1.1.3 (voir aussi appendice A.3). Soit $\mathrm{Ann}(\partial)$ un générateur de l'idéal des polynômes $P \in F[Y, \mathrm{Frob}]$ tels que $P \cdot \partial = 0$. Contrairement au cas considéré dans l'appendice A.3, il est possible ici d'avoir $\mathrm{Ann}(\partial) = 0$. Lorsque $\mathrm{Ann}(\partial)$ est non nul, on suppose en outre qu'il est normalisé de manière à être unitaire.

Proposition 4.1.3. *Si K est de caractéristique 0, le centre de $K[X, \partial]$ est F dès lors que $\partial \neq 0$. Si K est de caractéristique $p > 0$, le centre de $K[X, \partial]$ est $F[\mathrm{Ann}(\partial)_{lin}(X)]$.*

Démonstration. Supposons, dans un premier temps, que K soit de caractéristique nulle. Soit $P = a_0 + a_1 X + \cdots + a_n X^n \in K[X, \partial]$ un polynôme central. On suppose que n est choisi de sorte que $a_n \neq 0$. De la relation $XP - PX = \sum_{i=1}^{n} \partial(a_i) X^i$, on déduit que $\partial(a_i) = 0$ pour tout indice i. Autrement dit $P \in F[X]$. Par ailleurs, étant donné $a \in K$, un calcul montre que le coefficient en X^{n-1} de $Pa - aP$ est $n a_n \partial(a)$. Ainsi, en supposant qu'il existe $a \in K$ tel que $\partial(a) \neq 0$, on obtient $n = 0$, d'où $P \in F$ comme voulu.

Venons-en maintenant au cas où la caractéristique de K est $p > 0$. Soit $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X) \in K[X, \partial]$. Si $\mathrm{Ann}(\partial)$ s'écrit $\mathrm{Ann}(\partial) = c_0 + c_1 Y + \cdots + c_m Y^m$, on a :

$$Z(X) = c_0 + c_1 X^p + \cdots + c_m X^{p^m}.$$

Commençons par montrer que $Z(X)$ est un polynôme central. Il suffit pour cela de vérifier que $Z(X)$ commute avec X ainsi qu'avec tous les scalaires $a \in K$. La commutation avec X résulte du fait que $Z(X)$ est à coefficients dans F. Pour ce qui concerne la commutation avec $a \in K$, on écrit :

$$aZ(X) - Z(X)a = \sum_{i=0}^{m} c_i \left(aX^{p^i} - X^{p^i} a\right) = \sum_{i=0}^{m} c_i \, \partial^{p^i}(a) = 0.$$

Posons $\mathcal{Z} = F[Z(X)]$. D'après ce qui précéde, $\mathcal{Z}$ est inclus dans le centre de $K[X, \partial]$. Par ailleurs, si $C \neq 0$, choisissons m de sorte que $c_m \neq 0$. L'anneau de Ore $K[X, \partial]$ est alors un $\mathcal{Z}$-module libre de base $(1, X, X^2, \ldots, X^{p^m - 1})$. Soit P un élement central de $K[X, \theta]$ que l'on écrit sous la forme $P = z_0 + z_1 X + z_2 X^2 + \cdots + z_{p^m - 1} X^{p^m - 1}$. Si $P = 0$, on a clairement $P \in \mathcal{Z}$ et il n'y a rien à démontrer. Sinon, appelons s le plus grand indice pour lequel $z_s \neq 0$. Bien sûr $s < p^m$. Comme dans le cas de la caractéristique nulle, un calcul montre que :

$$aP - Pa = z'_0 + z'_1 X + \cdots + z'_{s-2} X^{s-2} + s z_s \partial(a) X^{s-1}$$

pour certains coefficients $z'_i \in \mathcal{Z}$. On en déduit que $s z_s \partial(a) = 0$. Comme ceci doit être vrai pour tout a, on aboutit à $s \equiv 0 \pmod{p}$. Écrivons $s = ps'$ et posons $Q = P - z_s X^s = z_0 + z_1 X + z_2 X^2 + \cdots + z_{s-1} X^{s-1}$. Soit t le plus grand entier strictement inférieur à s tel que $z_t \neq 0$.

$$(z_s X^s)a - a(z_s X^s) = s' z_s \partial^p(a) X^{s-p} + \cdots$$
$$Qa - aQ = t z_t \partial(a) X^{t-1} + \cdots$$

où les points de suspension représentent une somme de termes de « degrés en X » inférieurs. Ainsi si $t > s - p + 1$, on doit avoir $tz_t\partial(a) = 0$ pour tout a. Ceci entraîne que p divise t, ce qui est incompatible avec la condition $t > s - p + 1$. Si $t = s - p + 1$, la condition que l'on obtient s'écrit $s'z_s\partial^p(a) + tz_t\partial(a) = 0$. Dès lors que $m > 1$, ceci ne peut se produire que si $s'z_s = tz_t = 0$. On déduit ainsi que s' doit être un multiple de p, c'est-à-dire que s doit être un multiple de p^2. Enfin, si $t < s - p + 1$, on trouve la condition $s'z_s = 0$, de laquelle on déduit pareillement que s est un multiple de p^2.

En répétant l'argument, on démontre par récurrence sur i que s est un multiple de p^i pour tout entier $i \leq m$. Ainsi p^m divise s et la condition $s < p^m$ ne laisse plus que la possibilité $s = 0$. On en déduit que $P = z_0 \in \mathcal{Z}$. Le centre de $K[X, \partial]$ est donc bien réduit à $\mathcal{Z}$.

Enfin, le cas où $Z(X) = 0$ se traite de la manière totalement similaire. □

Dans le cas d'un endomorphisme θ d'ordre fini, le degré de l'extension K/F était égal à l'ordre de θ et donc également au plus petit degré d'un polynôme central non constant. La proposition A.3.2 de l'appendice A.3 montre que ce résultat demeure dans le cas différentiel lorsque le centre de $K[X, \partial]$ est d'indice fini. Il sera amené à jouer un rôle important dans la suite. Soulignons également que, dans le cas différentiel, l'extension K/F est purement inséparable alors qu'elle est séparable dans le cas d'un endomorphisme.

Exemple 4.1.4. Attardons-nous quelques instants sur plusieurs exemples en caractéristique p.

a- On prend $K = \mathbb{F}_p(t)$ et $\partial = \frac{d}{dt}$. Le noyau de ∂ est le sous-corps $F = \mathbb{F}_p(t^p)$. D'autre part, on vérifie que $\partial^p = 0$ (alors que $\partial \neq 0$). Ainsi $\mathrm{Ann}(\partial) = Y$. D'après la proposition 4.1.3, le centre de $\mathbb{F}_p(t)[X, \frac{d}{dt}]$ est $\mathbb{F}_p(t^p)[X^p]$.

b- On prend encore $K = \mathbb{F}_p(t)$ mais on le munit à présent de la dérivation $\partial = t\frac{d}{dt}$. À nouveau, le sous-corps des constantes F est $\mathbb{F}_p(t^p)$. Le calcul de ∂^p est, par contre, différent. En effet, on vérifie immédiatement que $\partial(t) = t$, de sorte que $\partial^p(t) = t$ également. Comme ∂ et ∂^p sont des dérivations qui coïncident sur le générateur t, on obtient $\partial^p = \partial$. Le polynôme $\mathrm{Ann}(\partial)$ est donc, cette fois-ci, $\mathrm{Ann}(\partial) = Y - 1$. En conclusion, on trouve que le centre de $\mathbb{F}_p(t)[X, t\frac{d}{dt}]$ est $\mathbb{F}_p(t^p)[X^p - X]$.

c- Soient $K = \mathbb{F}_p(x, y)$ et $\partial = \frac{d}{dx} - x^{p-1}\frac{d}{dy}$. Un calcul donne $\partial(x) = 1$ et $\partial(y) = -x^{p-1}$. Ainsi trouve-t-on $\partial^p(x) = 0$ et $\partial^p(y) = 1$, d'où on déduit que la dérivation ∂^p s'identifie à $\frac{d}{dy}$. Elle est donc linéairement indépendante de ∂. Le calcul de ∂^{p^2} aboutit aisément à $\partial^{p^2} = 0$. Il en résulte que $\mathrm{Ann}(\partial) = Y^2$. Reste à déterminer le sous-corps des constantes. Soit $f \in K$ tel que $\partial(f) = 0$. Cela implique évidemment que $\partial^p(f) = 0$ et donc, comme $\partial^p = \frac{d}{dy}$, on obtient $\partial(f) = \frac{df}{dx} = 0$. Ainsi $\frac{df}{dx} = \frac{df}{dy} = 0$, ce qui implique que $f \in \mathbb{F}_p(x^p, y^p)$. Réciproquement, il est clair que tous les éléments de $\mathbb{F}_p(x^p, y^p)$ sont annulés par la dérivation ∂. Ces calculs étant faits, on déduit de la proposition 4.1.3, que le centre de $\mathbb{F}_p(x, y)[X, \frac{d}{dx} - x^{p-1}\frac{d}{dy}]$ est $\mathbb{F}_p(x^p, y^p)[X^{p^2}]$.

Dans tous les cas, on remarque que le sous-corps des puissances p-ièmes de K est d'indice fini dans K. Ceci implique *a fortiori* que l'extension K/F est fini et donc que l'espace des dérivations $\mathrm{Der}(K)$ est de dimension finie sur F (puisqu'il est inclus dans $\mathrm{End}_F(K)$). Ainsi, le fait que le centre de $K[X, \partial]$ ne soit pas réduit aux constantes mais soit un authentique anneau de polynômes est attendu. En réalité, dans les trois exemples présentés ci-dessus, le sous-corps F s'identifiait au sous-corps des puissances p-ièmes. Ce n'est, bien sûr, en général, pas toujours le cas. Par exemple, si le sous-corps de constantes de $K = \mathbb{F}_p(x, y)$ pour la dérivation $\frac{d}{dx}$ est $\mathbb{F}_p(x^p, y)$ alors que le sous-corps des puissances p-ièmes de K est $\mathbb{F}_p(x^p, y^p)$.

4.2 Une algèbre d'Azumaya

D'après le calcul que nous venons de faire, il ne peut être vrai que les anneaux de Ore $\mathcal{A} = K[X, \theta, \partial]$ sont toujours des algèbres d'Azumaya sur leur centre $\mathcal{Z}$. En effet, une condition

nécessaire pour cela est que $\mathcal{Z}$ soit d'indice fini dans $\mathcal{A}$. Or dans le cas d'un endomorphisme (resp. d'une dérivation), ceci ne se produit que lorsque θ est d'ordre fini (resp. lorsque K est de caractéristique positive et $\mathrm{Ann}(\partial) \neq 0$). Toutefois, nous allons démontrer dans ce numéro que lorsque ces conditions sont remplies, la propriété d'Azumaya est bel et bien vérifiée, à quelques modifications mineures près. Comme d'habitude, nous traitons séparément le cas d'un endomorphisme et celui d'une dérivation.

4.2.1 Le cas de $K[X, \theta]$

On considère $\theta : K \to K$ un endomorphisme d'anneaux d'ordre r. Comme précédemment, on note F le sous-corps de K fixé par θ. D'après la proposition 4.1.1, le centre de l'anneau de Ore $K[X, \theta]$ est F^r.

Il n'est pas vrai que $K[X, \theta]$ est une algèbre d'Azumaya sur $F[X^r]$. En effet, pour l'idéal premier $\mathfrak{p} = (X^r)$, l'algèbre $K[X, \theta]/(X^r)$ n'est manifestement pas simple : elle possède un idéal bilatère qui est celui engendré par X. Toutefois, si l'on excepte ce cas, les conditions d'algèbres d'Azumaya sont bien remplies. Pour exprimer cela rigoureusement, on inverse formellement la variable X. Concrètement, on introduit l'anneau $K[X, \theta][\frac{1}{X}]$ dont les éléments sont les expressions de la forme :

$$a_v X^v + a_{v+1} X^{v+1} + \cdots a_{n-1} X^{n-1} + a_n X^n$$

pour un certain entier *relatif* v et des élément $a_i \in K$. Ces expressions s'additionnent comme des polynômes de Laurent usuels et se multiplient grâce à la loi $X^i \cdot a = \theta^i(a) X^i$, l'exposant i pouvant être positif ou négatif. On notera que cela ne pose pas de difficultés étant que θ est bijectif puisqu'il est supposé d'ordre fini. L'anneau commutatif des polynômes de Laurent traditionnels $F[X^r][\frac{1}{X^r}]$ s'injecte, de manière naturelle, dans $K[X, \theta][\frac{1}{X}]$. De même que dans la proposition 4.1.1, on démontre que $F[X^r][\frac{1}{X^r}]$ est le centre de $K[X, \theta][\frac{1}{X}]$.

Le théorème 4.2.1 ci-après est le résultat principal de ce numéro. Il apparaît, textuellement sous cette forme, dans un article d'Ikehata [12, 13]. Néanmoins, il transparait déjà de manière extrêmement forte dans l'un des premiers articles de Jacobson sur le sujet [14].

Théorème 4.2.1. *Soit $\theta : K \to K$ un automorphisme d'ordre fini r. Alors $K[X, \theta][\frac{1}{X}]$ est une algèbre d'Azumaya sur $F[X^r][\frac{1}{X^r}]$.*

Démonstration. Écrivons $\mathcal{Z} = F[X^r][\frac{1}{X^r}]$. Soit $\mathfrak{p}$ un idéal premier de $\mathcal{Z}$. Concrètement $\mathfrak{p}$ est soit l'idéal nul, soit l'idéal principal engendré par un polynôme unitaire irréductible sur F en la variable X^r. Supposons pour commencer que $\mathfrak{p} = (N(X^r))$ pour un certain polynôme $N(X^r) \in F[X^r]$. Étant donné qu'on a inversé formellement X, on a $N(X^r) \neq X^r$. Notons d le degré de P, vu comme un polynôme en X^r. Comme $N(X^r)$ est premier avec X et X^r, les quotients $K[X, \theta][\frac{1}{X}]/N(X^r)$ et $F[X^r][\frac{1}{X^r}]/N(X^r)$ s'identifient respectivement à $K[X, \theta]/N(X^r)$ et $F[X^r]/N(X^r)$; il s'agit donc de démontrer que $K[X, \theta]/N(X^r)$ est une algèbre simple centrale sur $F[X^r]/N(X^r)$.

Commençons par montrer qu'elle est centrale. On procède de même que dans la démonstration de la proposition 4.1.1. Soit $P \in K[X, \theta]/N(X^r)$ un élément central. Il s'écrit de manière unique sous la forme $P(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_{dr-1} X^{dr-1}$ avec $a_i \in K$. Pour $a \in K$, $a \neq 0$, un calcul donne :

$$X \cdot P(X) \cdot X^{-1} = \theta(a_0) + \theta(a_1) X + \theta(a_2) X^2 + \cdots + \theta(a_{dr-1}) X^{dr-1} \tag{27}$$

$$\text{et} \quad a^{-1} \cdot P(X) \cdot a = a_0 + a_1 \frac{\theta(a)}{a} X + a_2 \frac{\theta^2(a)}{a} X^2 + \cdots + a_{dr-1} \frac{\theta^{dr-1}(a)}{a} X^{dr-1} \tag{28}$$

d'où on déduit, d'une part, que $a_i = \theta(a_i) = a_i \frac{\theta^i(a)}{a}$ pour tout i. Cela implique que $a_i \in F$ pour tout i et que, si i n'est pas multiple de r, alors $a_i = 0$. Ainsi $P(X) \in F[X^r]/N(X^r)$ comme souhaité.

Montrons maintenant que $K[X,\theta]/N(X^r)$ est une algèbre simple. Soit $\mathcal{I}$ un idéal bilatère de $K[X,\theta]/N(X^r)$ que l'on suppose non nul. On considère $P(X) = a_0 + a_1X + \cdots + a_{dr-1}X^{dr-1}$ un élément non nul de I, choisi de façon à ce que le nombre de a_i non nuls soit minimal. Quitte à diviser P par une puissance de X (ce qui est possible puisque X est inversible), on peut supposer que $a_0 \neq 0$. Quitte à diviser par a_0, on peut supposer, mieux encore, que $a_0 = 1$. D'après la formule (27), le polynôme $P(X) - X \cdot P(X) \cdot X^{-1}$ a, au moins, un coefficient non nul de plus de P. Ainsi, par minimalité, on trouve $P = X \cdot P(X) \cdot X^{-1}$. Autrement dit $P(X)$ commute avec X. De la même manière, à partir de la formule (28), on trouve que $P(X)$ commute avec tous les éléments $a \in K$. Ainsi $P(X)$ est central et, d'après ce que nous avons déjà vu, $P(X)$ est dans $F[X^r]/N(X^r)$. Ce dernier anneau étant un corps, on en déduit que $P(X)$ est inversible. L'idéal $\mathcal{I}$ contient ainsi un élément inversible ; il est donc égal à $K[X,\theta]/N(X^r)$ tout entier.

Il reste à traiter le cas de l'idéal nul, c'est-à-dire à démontrer que $F(X^r) \otimes_{F[X^r]} K[X,\theta]$ est une algèbre simple centrale sur $F(X^r)$. Ceci se fait en reprenant ligne à ligne la démonstration précédente. □

Le théorème 4.2.1 implique que les algèbres quotient $K[X,\theta]/N(X^r)$ (pour un polynôme irréductible $N(X^r) \in F[X^r][\frac{1}{X}]$) deviennent isomorphes à des algèbres de matrices sur des algèbres à divisions. Il y a certains cas (principalement lorsque l'algèbre à divisions est triviale) où cet isomorphisme peut être rendu complètement explicite. Par exemple, dans le cas du polynôme $N(X^r) = X^r - 1$, cet isomorphisme est donné par le morphisme d'évaluation que nous avons déjà rencontré dans la remarque 1.3.6. Plus précisément, ce morphisme induit un isomorphisme :

$$\begin{array}{rccc} \text{ev} : & K[X,\theta]/(X^r-1) & \xrightarrow{\sim} & \mathrm{End}_F(K) \simeq M_r(F) \\ & a_0 + a_1X + \cdots + a_{r-1}X^{r-1} & \mapsto & a_0 + a_1\theta + \cdots + a_{r-1}\theta^{r-1} \end{array}$$

Plus généralement, étant donnés $\lambda \in K$ et $a = N_{K/F}(\lambda) = \lambda \cdot \theta(\lambda) \cdots \theta^{r-1}(\lambda)$, le morphisme d'évaluation en $\lambda\theta$ identifie de la même manière $K[X,\theta]/(X^r - a)$ avec $\mathrm{End}_F(K)$.

Bien entendu, il se peut que $K[X,\theta]/N(X^r)$ ne soit pas directement isomorphe à une algèbre de matrices sur $F[X^r]/N(X^r)$ mais, bel et bien, sur une algèbre à divisions non triviale. L'exemple simple suivant illustre ce fait.

Exemple 4.2.2. Plaçons-nous dans l'anneau de Ore $\mathbb{C}[X, \mathrm{conj}]$ et considérons le polynôme central $N(X^2) = X^2 - c$ avec $c \in \mathbb{R}$. Lorsque $c > 0$, il résulte que ce nous avons fait précédemment que $\mathbb{C}[X, \mathrm{conj}]/(X^2 - c)$ est isomorphe à $M_2(\mathbb{R})$. Si $a \in \mathbb{C}$ est un nombre complexe de module $\sqrt{c}$, un tel isomorphisme est donné par l'application qui envoie $u + vX$ sur la matrice de l'application $\mathbb{R}$-linéaire $\mathbb{C} \to \mathbb{C}$, $z \mapsto uz + av\bar{z}$. Concrètement, pour $a = \sqrt{c}$, on obtient l'isomorphisme :

$$\begin{array}{rcl} \mathbb{C}[X, \mathrm{conj}]/(X^2-c) & \xrightarrow{\sim} & M_2(\mathbb{R}) \\ u + vX & \mapsto & \begin{pmatrix} \Re(u) + \Re(v)\sqrt{c} & -\Im(u) + \Im(v)\sqrt{c} \\ \Im(u) + \Im(v)\sqrt{c} & \Re(u) - \Re(v)\sqrt{c} \end{pmatrix}. \end{array}$$

Par contre, lorsque $c < 0$, il se trouve que $\mathbb{C}[X,\theta]/(X^2 - c)$ est une algèbre à divisions. Ceci se voit directement grâce à la formule explicite suivante $(u + vX) \cdot (\bar{u} - vX) = |u|^2 - c|v|^2$ (pour $u, v \in \mathbb{C}$) : comme la quantité $|u|^2 - c|v|^2$ ne s'annule que lorsque $u = v = 0$, on en déduit que $u + vX$ est inversible dès qu'il est non nul. Au passage, on notera que $|u|^2 - c|v|^2$ est la norme réduite de $u + vX$ et que $\bar{u} - vX$ est son adjoint (dans la $\mathbb{R}$-algèbre simple centrale $\mathbb{C}[X, \mathrm{conj}]/(X^2 - c)$).

De manière générale, il est également possible de décrire une extension (séparable) explicite sur laquelle l'algèbre simple centrale $K[X,\theta]/N(X^r)$ devient triviale (*i.e.* isomorphe à une algèbre de matrices). Précisément, notons E l'extension de F définie par le polynôme $N(X^r)$ (*i.e.* $E = F[X^r]/N(X^r)$) et plongeons E dans une clôture algébrique de K. Soit EK l'extension

composée de E et K et soit $[X^r] \in E$ la classe du polynôme X^r. Avec ces notations, on dispose toujours d'un isomorphisme entre $EK \otimes_E K[X,\theta]/C(X^r)$ et $M_r(EK)$ défini ainsi :

$$\lambda \otimes 1 \mapsto \lambda I_r \quad ; \quad 1 \otimes a \mapsto \begin{pmatrix} a & & & \\ & \theta(a) & & \\ & & \ddots & \\ & & & \theta^{r-1}(a) \end{pmatrix} \quad ; \quad 1 \otimes X \mapsto \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ [X^r] & & & \end{pmatrix}$$

pour $\lambda \in EK$ et $a \in K$. Dans le cas particulier où $K \subset E$, on a $EK = E$ et l'isomorphisme précédent se réduit à un isomorphisme entre $K[X,\theta]/N(X^r)$ et $M_r(E)$. Autrement dit, l'algèbre simple centrale $K[X,\theta]/N(X^r)$ est triviale.

Plus généralement, les formules ci-dessus définissent un isomorphisme $K \otimes_F K[X,\theta] \simeq M_r(K[X^r])$ qui se spécialise, modulo chaque polynôme $N(X^r)$, en l'isomorphisme discuté à l'alinéa précédent.

4.2.2 Le cas de $K[X,\partial]$

Dans le cas d'une dérivation, on dispose de résultats similaires. À vrai dire, ils sont même meilleurs car il n'est plus nécessaire d'inverser formellement la variable X. On pose également $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X)$.

Le théorème suivant est l'analogue du théorème 4.2.1. Dans le cas d'un corps de la forme $k(t)$ (où k est un corps de caractéristique p) muni de la dérivation $\frac{d}{dt}$, le théorème suivant est dû à Revoy [23].

Théorème 4.2.3. *Soit $\partial : K \to K$ une dérivation pour laquelle $\mathrm{Ann}(\partial) \neq 0$. Alors $K[X,\partial]$ est une algèbre d'Azumaya sur $F[Z(X)]$.*

Démonstration. La démonstration est très similaire à celle du théorème 4.2.1. Posons, pour simplifier les écritures, $\mathcal{Z} = F[Z(X)]$. Soit $N(T)$ un polynôme irréductible unitaire à coefficients dans F. Nous devons démontrer que $K[X,\partial]/N(Z(X))$ est une algèbre simple centrale sur $F[A(X)]/N(Z(X))$. Le calcul du centre de $K[X,\partial]/N(Z(X))$ se fait de même que dans la démonstration de la proposition 4.1.3 ; nous ne le reproduisons pas ici.

Montrons que $K[X,\partial]/N(Z(X))$ est une algèbre simple. Soit $\mathcal{I}$ un idéal bilatère non nul de $K[X,\partial]/N(Z(X))$. Soit $P(X) \in \mathcal{I}$ un élément non nul. Si $d = \deg_X N(Z(X))$, on peut écrire de manière unique $P(X)$ sous la forme :

$$P(X) = a_0 + a_1 X + a_2 X^2 + \cdots + a_{d-1} X^{d-1}$$

pour des $a_i \in K$. Notons s le plus grand entier pour lequel $a_s \neq 0$ et choisissons $P(X) \in \mathcal{I}$ de sorte que s soit le plus petit possible. Quitte à diviser par a_s, on peut normaliser $P(X)$ de sorte que $a_s = 1$. Le polynôme de Ore

$$X \cdot P(X) - P(X) \cdot X = \partial(a_0) + \partial(a_1)X + \cdots + \partial(a_{s-1})X^{s-1}$$

est ainsi de « degré » strictement inférieur à s. Par minimalité, on trouve $X \cdot P(X) = P(X) \cdot X$, ce qui veut dire que $P(X)$ commute avec X. De même, un calcul montre que, pour $a \in K$, le « degré » de $a \cdot P(X) - P(X) \cdot a$ est stictement inférieur à s et donc, par minimalité, que $P(X)$ commute avec a. En conclusion, $P(X)$ est un élément central dans $K[X,\partial]/N(Z(X))$; il appartient donc à $F[A(X)]/N(Z(X))$. Comme ce dernier est un corps, on trouve que $P(X)$ est inversible et, par conséquent, que l'idéal qu'il engendre est égal à l'anneau $K[X,\partial]/N(Z(X))$ tout entier. □

De la même manière que dans le cas d'un automorphisme, le morphisme d'évaluation induit un isomorphisme

$$\begin{array}{rccc} \text{ev} : & K[X,\partial]/Z(X) & \xrightarrow{\sim} & \mathrm{End}_F(K) \simeq M_{p^m}(F) \\ a_0 + a_1 X + \cdots + a_{p^m-1} X^{p^m-1} & & \mapsto & a_0 + a_1\partial + \cdots + a_{p^m-1}\partial^{p^m-1} \end{array}$$

où m est le degré de $\mathrm{Ann}(\partial) \in F[Y, \mathrm{Frob}]$. Ceci permet de réaliser le quotient $K[X,\partial]/Z(X)$ comme une algèbre de matrices.

Exemple 4.2.4. Dans le cas particulier de l'anneau de Ore $\mathbb{F}_p(t)[X, \frac{d}{dt}]$, on peut expliciter un isomorphisme entre $\mathbb{F}_p(t)[X, \frac{d}{dt}]$ et une algèbre de matrices, après une certaine extension des scalaires. Précisément, notons comme habituellement $F = \mathbb{F}_p(t^p)$ et posons $\mathcal{Z} = F[X^p]$; c'est le centre de $\mathbb{F}_p(t)[X, \frac{d}{dt}]$. Soit $\mathcal{Z}'$ l'extension de $\mathcal{Z}$ obtenue en ajoutant l'élement T solution de l'équation $T^p - T = t^p X^p$, *i.e.* $Z' = Z[T]/(T^p - T - t^p X^p)$. Avec ces notations, on a un isomorphisme $\mathcal{Z}' \otimes_{\mathcal{Z}} \mathbb{F}_p(t)[X, \frac{d}{dt}] \to M_p(\mathcal{Z}')$ défini par :

$$\lambda \otimes 1 \mapsto \lambda I_p \quad ; \quad 1 \otimes t \mapsto \begin{pmatrix} & & T \\ 1 & & \\ & \ddots & \\ & 1 & \end{pmatrix} \quad ; \quad 1 \otimes X \mapsto \begin{pmatrix} & 1 & \\ & & \ddots \\ & & 1 \\ [X^r] & & \end{pmatrix}$$

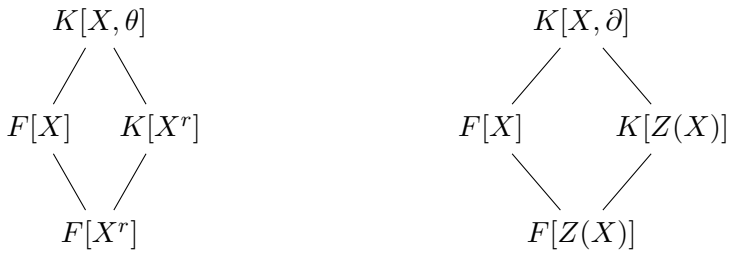
avec $\lambda \in \mathcal{Z}'$. Après réduction modulo un polynôme central de la forme $N(Z(X))$, cet isomorphisme fournit une trivialisation explicite de $\mathbb{F}_p(t)[X, \frac{d}{dt}]/N(Z(X))$.

4.3 La norme réduite sur les anneaux de Ore

Maintenant que nous savons que $K[X,\theta][\frac{1}{X}]$ et $K[X,\partial]$ sont des algèbres d'Azumaya, la prochaine étape consiste à exhiber des sous-algèbres C qui satisfont à l'hypothèse 3.4.3 dans le but de construire les applications de norme réduite et d'adjoint.

4.3.1 Construction de la norme réduite

Les diagrammes suivants font apparaître des candidats naturels pour être la sous-algèbre C recherchée :



Ci-dessus, on a repris les notations du paragraphe précédent : dans le cas d'un automorphisme θ (resp. d'une dérivation ∂), F désigne le sous-corps de K formé des points fixes de θ (resp. des constantes, *i.e.* des éléments $x \in K$ tels que $\partial(x) = 0$). La lettre r, quant à elle, désigne l'ordre de θ, tandis que $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X)$.

Proposition 4.3.1. *On conserve les notations précédentes.*

- (θ) *Dans le cas d'un automorphisme, les sous-algèbres $K[X^r][\frac{1}{X^r}]$ et $F[X][\frac{1}{X}]$ de $K[X,\theta][\frac{1}{X}]$ vérifient l'hypothèse 3.4.3 ;*
- (∂) *Dans le cas d'une dérivation, la sous-algèbre $F[X]$ de $K[X,\partial]$ vérifient l'hypothèse 3.4.3.*

Démonstration. Commençons par le cas d'un automorphisme. Clairement le centre $F[X^r]$ n'a pas d'éléments nilpotents. Pour terminer la démonstration, il suffit donc de vérifier que, pour $C = K[X^r][\frac{1}{X^r}]$ ou $C = F[X][\frac{1}{X}]$, les conditions suivantes sont satisfaites :

(i) C est libre de rang r sur $F[X^r][\frac{1}{X^r}]$ et possède une base de la forme $(1, c_2, \dots, c_r)$;

(ii) $K[X, \theta][\frac{1}{X}]$ est libre de rang r sur C ;

(iii) en tant que $F[X^r][\frac{1}{X^r}]$-algèbre, C est engendrée par un unique élément.

Ces trois propriétés se vérifient sans difficultés. Pour $C = K[X^r][\frac{1}{X^r}]$, on constate que tout base de K/F contenant l'élément 1 fournit une base convenable de C sur $F[X^r][\frac{1}{X^r}]$, que la famille $(1, X, \dots, X^{r-1})$ est une base de $K[X, \theta][\frac{1}{X}]$ sur C et, finalement, que comme $F[X^r][\frac{1}{X^r}]$-algèbre, C est engendrée par l'élément X. Pour $C = F[X][\frac{1}{X}]$, une base possible de C sur $F[X^r][\frac{1}{X^r}]$ est $(1, X, \dots, X^{r-1})$, tout base de K/F fournit une base de $K[X, \theta][\frac{1}{X}]$ sur C et, enfin, le théorème de l'élément primitif appliqué à l'extension K/F montre que C est une $F[X^r][\frac{1}{X^r}]$-algèbre engendrée par un unique élément.

L'argumentation est similaire dans le cas d'une dérivation. □

Remarque 4.3.2. Malheureusement, il n'est pas vrai en général que la sous-algèbre $K[Z(X)]$ de $K[X, \partial]$ vérifie l'hypothèse 3.4.3, le problème venant du fait que l'extension purement inséparable K/F n'est pas monogène (même après extension des scalaires à une clôture algébrique). Précisément, la structure de F-algèbre de K est donnée par l'isomorphisme :

$$K \simeq F[X_1, X_2, \dots, X_m]/(X_1^p - a_1,\ X_2^p - a_2, \dots, X_m^p - a_m) \tag{29}$$

où m est un entier et les a_i sont des éléments de F que l'on peut supposer deux à deux distincts. Pour établir cet isomorphisme, on peut procéder ainsi. On considère $x_1 \in K$, $x_1 \notin F$. Étant donné que $\partial(x_1^p) = px_1^{p-1}\partial(x_1) = 0$, on obtient $x_1^p \in F$. Ainsi x_1 engendre une extension de F de degré p et le polynôme minimal de x_1 est $X^p - a_1$ où $a_1 = x_1^p$ par définition. Autrement dit, l'extension $F(x_1)$ est isomorphe au corps $F[X_1]/(X_1^p - a_1)$. On choisit maintenant, s'il existe, un élément $x_2 \in K$, $x_2 \notin F(x_1)$. De même que précédemment, on trouve que le polynôme minimal de x_2 sur $F(x_1)$ est $x_2^p - a_2$ avec $a_2 = x_2^p \in F$. Ainsi l'extension $F(x_1, x_2)$ est isomorphe à $F[X_1, X_2]/(X_1^p - a_1, X_2^p - a_2)$. Continuant ce processus jusqu'à ce qu'il ne soit plus possible de choisir un nouvel x_i, on obtient l'isomorphisme (29).

À partir de la description précédente, il est possible de reprendre l'argumentation de la démonstration de la proposition 3.3.9 et d'aboutir à la congruence de normes :

$$N_{K[X,\partial]/K[\mathrm{Ann}(\partial)_{\mathrm{lin}}(X)]} \equiv N_{\mathrm{rd}} \pmod{\mathfrak{p}}$$

pour tout idéal premier $\mathfrak{p}$ de $F[Z(X)]$. Ainsi bien que l'hypothèse 3.4.3 soit mise en défaut, la conclusion de la proposition 3.3.9 vaut pour $C = K[Z(X)]$. Il en est de même de la conclusion de la proposition 3.3.10.

Il résulte de la proposition 4.3.1 que l'algèbre d'Azumaya $K[X, \theta][\frac{1}{X}]$ (resp. $K[X, \partial]$) est équipée d'applications :

$$\begin{aligned} & N_{\mathrm{rd}} : K[X, \theta][\tfrac{1}{X}] \to F[X^r][\tfrac{1}{X^r}] && (\text{resp. } N_{\mathrm{rd}} : K[X, \partial] \to F[\mathrm{Ann}(\partial)_{\mathrm{lin}}(X)]) \\ \text{et} \quad & \mathrm{adj} : K[X, \theta][\tfrac{1}{X}] \to K[X, \theta][\tfrac{1}{X}] && (\text{resp. adj} : K[X, \partial] \to K[X, \partial]). \end{aligned}$$

Dans le cas d'un endomorphisme, on peut en outre s'affranchir de l'inversion formelle de X comme le garantit le lemme suivant.

Lemme 4.3.3. *Dans le cas d'un endomorphisme, l'application N_{rd} (resp. adj) envoie $K[X, \theta]$ sur $F[X^r]$ (resp. $K[X, \theta]$).*

Démonstration. Cela suit directement du fait que la norme réduite (resp. l'adjoint) d'un polynôme de Ore P s'interprète comme le déterminant (resp. la valeur en 1 de l'adjoint) de l'application $K[X,\theta] \to K[X,\theta], R \mapsto RP$ vue comme application $K[X^r]$-linéaire. □

Exemple 4.3.4. Soit $a \in \mathbb{C}$. Calculons la norme réduite du polynôme de Ore $X-a \in \mathbb{C}[X, \mathrm{conj}]$. Pour cela, nous disposons *a priori* de deux méthodes différentes correspondant aux deux choix possibles de la sous-algèbre commutative, soit $\mathbb{C}[X^2]$, soit $\mathbb{R}[X]$. Faisons le calcul successivement avec ces deux méthodes. Pour commencer, supposons que l'on ait choisi $C = \mathbb{C}[X^2]$. Une base de $\mathbb{C}[X, \mathrm{conj}]$ sur $\mathbb{C}[X^2]$ est $(1, X)$. Dans cette base, la multiplication à droite par $X-a$ est :

$$M_1 = \begin{pmatrix} -a & X^2 \\ 1 & \bar{a} \end{pmatrix}.$$

La norme réduite est le déterminant de cette matrice et est donc $|a|^2 - X^2$. On constate, en particulier, qu'il s'agit bien d'un polynôme à coefficients réels (alors que les coefficients de la matrice M_1 font intervenir des nombres complexes).

Utilisons maintenant plutôt la sous-algèbre $C = \mathbb{R}[X]$. Une base de $\mathbb{C}[X, \mathrm{conj}]$ sur $\mathbb{R}[X]$ est $(1, i)$. Écrivons $a = \alpha + i\beta$. Les égalités :

$$\begin{aligned} 1 \times (X-a) &= X - \alpha - i\beta = (X-\alpha) - \beta i \\ i \times (X-a) &= iX - i\alpha + \beta = -Xi - i\alpha + \beta = \beta - (X+\alpha)i \end{aligned}$$

(on prendra garde à bien positionner les éléments de la base à droite) montrent que la matrice dans la base $(1, i)$ de la multiplication à droite par $X-a$ est :

$$M_2 = \begin{pmatrix} X-\alpha & \beta \\ -\beta & -X-\alpha \end{pmatrix}.$$

Son déterminant vaut $\alpha^2 + \beta^2 - X^2 = |a|^2 - X^2$. On retrouve bien le même résultat qu'avec la première méthode. À nouveau, on peut être surpris par le fait que le déterminant de M_2 ne fait pas intervenir de terme en X (alors qu'il en apparaissait dans ses coefficients).

4.3.2 Premières propriétés de la norme réduite

Pour uniformiser les notations, posons à partir de maintenant $\mathcal{A} = K[X,\theta]$ dans le cas d'un endomorphisme et $\mathcal{A} = K[X,\partial]$ dans le cas d'une dérivation. Notons également $\mathcal{Z}$ le centre de $\mathcal{A}$, *i.e.* $\mathcal{Z} = F[Z(X)]$ où $Z(X) = X^r$ dans le cas d'un endomorphisme et $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X)$ dans le cas d'une dérivation. Dans le cas d'une dérivation, convenons également que $r = \deg Z(X)$, soit encore $r = p^m$ avec $m = \deg_Y \mathrm{Ann}(\partial)$. Dans tous les cas, les applications de norme réduite et d'adjoint que nous venons de construire envoient respectivement $\mathcal{A}$ dans $\mathcal{Z}$ et $\mathcal{A}$ dans $\mathcal{A}$. La proposition 3.4.5 s'instancie, dans ce contexte, en l'énoncé suivant :

Proposition 4.3.5. *Pour deux polynômes de Ore $P, Q \in \mathcal{A}$ et un idéal premier $\mathfrak{p}$ de $\mathcal{Z}$, on a :*

(i) *si $P \in \mathcal{Z}$, alors $N_{rd}(P) = P^r$ et $adj(P) = P^{r-1}$,*

(ii) *$N_{rd}(PQ) = N_{rd}(P)N_{rd}(Q)$ et $adj(PQ) = adj(Q)adj(P)$,*

(iii) *$P \cdot adj(P) = adj(P) \cdot P = N_{rd}(P)$,*

(iv) *si $P \equiv Q \pmod{\mathfrak{p}\mathcal{A}}$, alors $N_{rd}(P) \equiv N_{rd}(Q) \pmod{\mathfrak{p}}$ et $adj(P) \equiv adj(Q) \pmod{\mathfrak{p}\mathcal{A}}$.*

Démonstration. Dans le cas de $K[X,\partial]$, il s'agit exactement de la proposition 3.4.5. Dans le cas de $K[X,\theta]$, les trois premières propriétés résultent également formellement de la proposition 3.4.5. Ce n'est, par contre, pas directement le cas pour le dernier énoncé ; en effet, la proposition 3.4.5 ne fournit *a priori* que l'implication :

(iv') si $P \equiv Q \pmod{\mathfrak{p}\mathcal{A}[\frac{1}{X}]}$,
alors $N_{\text{rd}}(P) \equiv N_{\text{rd}}(Q) \pmod{\mathfrak{p}\mathcal{Z}[\frac{1}{X^r}]}$ et $\text{adj}(P) \equiv \text{adj}(Q) \pmod{\mathfrak{p}\mathcal{A}[\frac{1}{X}]}$.

Cependant, si $\mathfrak{p}$ n'est pas l'idéal nul ni l'idéal principal (X^r), on a $\mathcal{A}[\frac{1}{X}]/\mathfrak{p}\mathcal{A}[\frac{1}{X}] \simeq \mathcal{A}/\mathfrak{p}\mathcal{A}$ et $\mathcal{Z}[\frac{1}{X^r}]/\mathfrak{p}\mathcal{Z}[\frac{1}{X^r}] \simeq \mathcal{Z}/\mathfrak{p}$. Ainsi, dans ce cas, (iv') est équivalente à (iv). De plus, lorsque $\mathfrak{p} = 0$, il est évident que (iv) est vraie. Il ne reste donc plus que le cas où $\mathfrak{p} = (X^r)$, qui se vérifie simplement à l'aide de la définition de la norme réduite. □

Dans le cadre particulier des anneaux de Ore, des propriétés supplémentaires peuvent être dégagées.

Proposition 4.3.6. *Pour $P \in \mathcal{A}$, on a :*

(i) $\deg N_{rd}(P) = r \cdot \deg P$,

(ii) $lc(N_{rd}(P)) = (-1)^{(r-1)\cdot \deg P} N_{K/F}(lc(P))$

où $lc(\cdot)$ désigne le coefficient dominant d'un polynôme de Ore.

Remarque 4.3.7. Dans la proposition ci-dessus, la notation $\deg$ fait toujours référence au degré en tant que polynôme de Ore, c'est-à-dire au degré calculé dans l'anneau $\mathcal{A}$. Une confusion pourrait être possible car $N_{\text{rd}}(P)$ vit dans l'anneau $\mathcal{Z}$ qui est, lui-même, un anneau de polynômes en une variable. Or, le polynôme $N_{\text{rd}}(P)$ n'a évidemment pas le même degré lorsqu'il est vu comme un élément de $\mathcal{Z}$ ou comme un élément de $\mathcal{A}$, étant donné que la variable sur $\mathcal{Z}$, à savoir $Z(X)$, est elle-même un polynôme de Ore de degré r. Précisément, on a la relation $\deg_{\mathcal{Z}} N = r \cdot \deg N$ pour $N \in \mathcal{Z}$. En particulier, le premier énoncé de la proposition 4.3.6 peut se réécrire, si on le souhaite, $\deg_{\mathcal{Z}} N_{\text{rd}}(P) = \deg P$.

Démonstration de la proposition 4.3.6. Écrivons P sous la forme $a_0 + a_1 X + a_2 X^2 \cdots + a_d X^d$ avec $a_d \neq 0$ de sorte que $\deg P = d$. On fixe une base de K sur F que l'on note $\mathcal{B}$. Dans le cas d'un endomorphisme, l'extension K/F est galoisienne et on suppose que $\mathcal{B}$ est une base normale, *i.e.* de la forme $(a, \theta(a), \ldots, \theta^{r-1}(a))$. Clairement $\mathcal{B}$ est aussi une base de $\mathcal{A}$ vu comme module sur $F[X]$. Soit M_P la matrice, dans cette base, de la multiplication à droite par P, c'est-à-dire de l'application $F[X]$-linéaire $\mathcal{A} \to \mathcal{A}$, $R \mapsto RP$. De même, pour $a \in K$, on note M_a (resp. M_X) la matrice de la multiplication à droite par a (resp. par X). Un calcul immédiat montre que, dans le cas d'un endomorphisme, $M_X = M_\theta^{-1} X$ tandis que, dans le cas d'une dérivation, $M_X = X - M_\partial$ où M_θ (resp. M_∂) désigne la matrice de θ (resp. de ∂) dans la base $\mathcal{B}$. Par ailleurs, on a l'identité :

$$M_P = M_{a_0} + M_{a_1} M_X + M_{a_2} M_X^2 + \cdots + M_{a_d} M_X^d$$

et, par définition, $N_{\text{rd}}(P)$ est le déterminant de M_P. Du fait que M_P est une matrice de taille $r = [K : F]$ et que les coefficients de M_X sont tous des polynômes de degré au plus 1, on déduit que $\deg N_{\text{rd}}(P) \leq rd$. De plus, le coefficient en X^{rd} est égal au déterminant de $M_{a_d} M_\theta^{-d}$ (resp. de M_{a_d}) dans le cas d'un endomorphisme (resp. d'une dérivation). Par définition, le déterminant de M_d est égal à $N_{K/F}(a_d)$. Ceci conclut la démonstration dans le cas d'une dérivation car on a toujours $(-1)^{r-1} = (-1)^{p^m - 1} = 1$ dans K. Dans le cas d'un endomorphisme, on se souvient que $\mathcal{B}$ est une base normale. Ainsi, la matrice M_θ est :

$$\begin{pmatrix} & & & 1 \\ 1 & & & \\ & \ddots & & \\ & & 1 & \end{pmatrix}$$

(les coefficients non renseignés étant tous nuls) et a donc $(-1)^{r-1}$ pour déterminant. □

Souvenons-nous (*cf* §1.3) que les anneaux de Ore $K[X,\theta]$ et $K[X,\partial]$ possèdent des propriétés arithmétiques remarquables puisque ce sont des anneaux euclidiens. En particulier, nous avons défini et étudié la notion de RGCD (plus grand commun diviseur à droite) sur ces anneaux. Les énoncés que nous allons dégager ci-après indiquent quelques compatibilités entre RGCD et norme réduite.

Avant d'entrer dans le vif du sujet, faisons une remarque importante. L'anneau $\mathcal{Z}$ étant un anneau de polynômes (commutatif) en une variable, la notion de PGCD a un sens dans $\mathcal{Z}$. Ainsi si N_1 et N_2 sont deux polynômes centraux, on peut, d'une part, les considérer comme deux polynômes de Ore et calculer leur RGCD dans $\mathcal{A}$ mais on peut aussi, d'autre part, les considérer comme des éléments de $\mathcal{Z}$ et calculer leur PGCD (commutatif) dans $\mathcal{Z}$. Il s'avère que cela ne pose aucune ambiguïté car ces deux opérations conduisent au même résultat. En effet, considérons deux polynômes $A(T)$ et $B(T)$ à coefficients dans F, avec $B(T) \neq 0$. Notons $A(Z(X))$ et $B(Z(X))$ les éléments de $\mathcal{Z}$ qu'ils définissent. La division euclidienne dans $\mathcal{Z}$ de A par B s'écrit sous la forme $A(T) = Q(T) \cdot B(T) + R(T)$ où $\deg_T R < \deg_T B$. Remplaçant T par $Z(X)$, on obtient l'écriture :

$$A(Z(X)) = Q(Z(X)) \cdot B(Z(X)) + R(Z(X)) \tag{30}$$

que l'on peut voir comme une égalité de polynômes de Ore dans $\mathcal{A}$. De plus $\deg R(Z(X)) < \deg B(Z(X))$. Ainsi l'égalité (30) n'est autre que la division eucidienne à droite de $A(Z(X))$ par $B(Z(X))$ dans l'anneau $\mathcal{A}$. On déduit de cela que, s'il débute avec deux polynômes centraux, l'algorithme d'Euclide produit le même résultat qu'il soit effectué dans $\mathcal{A}$ ou dans $\mathcal{Z}$. Ainsi les notions de RGCD et PGCD coïncident pour les polynômes centraux.

Lemme 4.3.8. *Pour $P \in \mathcal{A}$ et $N \in \mathcal{Z}$, on a l'équivalence suivante :* $\text{RGCD}(P,N) = 1$ *si et seulement si* $\text{PGCD}(N_{rd}(P), N) = 1$.

Démonstration. Quitte à décomposer N en produit d'irréductibles (dans l'anneau de polynômes $\mathcal{Z}$), on peut supposer que N est irréductible. Supposons que P et N soient premiers entre eux. Alors, par le théorème de Bézout, il existe Q tel que $PQ \equiv 1 \pmod N$. D'après la proposition 4.3.5, cela implique $N_{\text{rd}}(P)N_{\text{rd}}(Q) \equiv 1 \pmod N$, d'où on déduit que $N_{\text{rd}}(P)$ et N sont premiers entre eux.

Réciproquement, supposons que P et N ne soient pas premiers entre eux. On peut alors écrire $P = P'D$ et $N = N'D$ pour un certain polynôme de Ore D de degré strictement positif. En prenant la norme réduite, on déduit des écritures précédentes que $N_{\text{rd}}(D)$ divise $N_{\text{rd}}(P)$ et $N_{\text{rd}}(N)$. Or, d'après la proposition 4.3.6, $N_{\text{rd}}(D)$ est un polynôme de degré strictement positif. Ainsi les polynômes $N_{\text{rd}}(P)$ et $N_{\text{rd}}(N)$ ne sont pas premiers entre eux. Par ailleurs, en vertu de la proposition 4.3.5, on sait que $N_{\text{rd}}(N) = N^r$. Comme N est irréductible (dans $\mathcal{Z}$), on en déduit que $N_{\text{rd}}(P)$ et N ne sont pas premiers entre eux, comme voulu. □

Corollaire 4.3.9. *Soit $P \in \mathcal{A}$. Si P est irréductible dans $\mathcal{A}$, alors $N_{rd}(P)$ est, à une constante multiplicative près, une puissance d'un polynôme irréductible de $\mathcal{Z}$.*

Démonstration. Supposons par l'absurde que $N_{\text{rd}}(P)$ s'écrive sous la forme $N_{\text{rd}}(P) = N_1N_2$ où $N_1, N_2 \in \mathcal{Z}$ sont premiers entre eux. Soit $D = \text{RGCD}(P, N_1)$. Par le lemme 4.3.8, on sait que le degré de D est strictement positif. Écrivons $N_1 = QD$ et $P = RD$ avec $Q, R \in \mathcal{A}$. Pour conclure, il suffit de montrer que R n'est pas constant. Or, en prenant les normes réduites, on trouve $N_1^r = N_{\text{rd}}(N_1) = N_{\text{rd}}(Q)N_{\text{rd}}(D)$ et $N_1N_2 = N_{\text{rd}}(P) = N_{\text{rd}}(R)N_{\text{rd}}(D)$. De la première égalité, il résulte que $N_{\text{rd}}(D)$ est premier avec N_2. Sachant cela, on déduit de la seconde égalité que N_2 divise $N_{\text{rd}}(R)$, ce qui ne peut se produire si R est constant. □

Comme nous le verrons en détails dans la suite, l'exposant e pour lequel $N_{\text{rd}}(P) = c\,N^e$ avec $c \in F$ et N irréductible dans $\mathcal{Z}$ joue un rôle important dans la compréhension des propriétés fines de factorisation dans $\mathcal{A}$. Pour cette raison, nous en ferons une étude détaillée aux §4.4–4.6

et montrerons notamment qu'il s'interprète de manière naturelle dans le langage des algèbres simples centrales.

Enfin, dans le cas des polynômes *irréductibles*, la norme réduite permet également de reconnaître les paires de polynômes associés au sens de la définition 1.4.8.

Proposition 4.3.10. *Pour deux polynômes de Ore irréductibles $P, P' \in \mathcal{A}$, on a l'équivalence suivante : P et P' sont associés si et seulement si $N_{rd}(P) = c\, N_{rd}(P')$ pour une constante non nulle $c \in F$.*

Démonstration. Supposons, tout d'abord, que P et P' soient associées, c'est-à-dire que les modules M_P et $M_{P'}$ soient isomorphes. Soit $\mathcal{I}$ l'idéal de $\mathcal{Z}$ formé des polynômes centraux N tels que $NM_P = 0$. Puisque $N_{\mathrm{rd}}(P)$ est un multiple de P, on a $N_{\mathrm{rd}}(P) \in \mathcal{I}$. Étant donné que $M_P \simeq M_{P'}$, on obtient de même $N_{\mathrm{rd}}(P') \in \mathcal{I}$. Ainsi $\mathcal{I}$ contient $\mathrm{PGCD}(N_{\mathrm{rd}}(P), N_{\mathrm{rd}}(P'))$. Or $\mathcal{I}$ ne contient manifestement pas 1 (puisque $M_P \neq 0$) ; on en déduit que $N_{\mathrm{rd}}(P)$ et $N_{\mathrm{rd}}(P')$ possèdent un diviseur commun non trivial. Or, en vertu du corollaire 4.3.9, on sait par ailleurs que $N_{\mathrm{rd}}(P)$ et $N_{\mathrm{rd}}(P')$ sont, tous les deux, des puissances de polynômes irréductibles de $\mathcal{Z}$. On en déduit que $N_{\mathrm{rd}}(P) = \lambda\, N_0^e$ et $N_{\mathrm{rd}}(P') = \lambda'\, N_0^{e'}$ pour des entiers e, e', des constantes non nulles $\lambda, \lambda' \in F$ et le *même* polynome irréductible $N_0 \in \mathcal{Z}$. De plus, de l'égalité des degrés de P et P', on déduit, avec la proposition 4.3.6, que $\deg N_{\mathrm{rd}}(P) = \deg N_{\mathrm{rd}}(P')$ et donc, par suite, que $e = e'$ et que $N_{\mathrm{rd}}(P') = \frac{\lambda'}{\lambda}\, N_{\mathrm{rd}}(P)$.

Réciproquement, supposons que $N_{\mathrm{rd}}(P) = c\, N_{\mathrm{rd}}(P')$ avec $c \in F$, $c \neq 0$. Puisque P est irréductible, on peut écrire comme précédemment, $N_{\mathrm{rd}}(P) = \lambda\, N_0^e$ pour un certain polynôme irréductible N_0 de $\mathcal{Z}$. De plus, par le lemme 4.3.8, on a $\mathrm{RGCD}(P, N_0) \neq 1$. Par irréductibilité, on en déduit que P divise N_0 à droite. De la même manière, P' est, lui aussi, un diviseur à droite de N_0. Les $\mathcal{A}$-modules M_P et $M_{P'}$ apparaissent donc comme des modules sur l'anneau quotient $\mathcal{A}/N_0\mathcal{A}$ qui est une algèbre simple centrale. D'autre part, l'égalité des degrés $\deg N_{\mathrm{rd}}(P) = \deg N_{\mathrm{rd}}(P')$ implique celle des degrés de P et P' par la proposition 4.3.6. On déduit ainsi de la proposition 3.2.8 que les $(\mathcal{A}/N_0\mathcal{A})$-modules M_P et $M_{P'}$ sont isomorphes. Autrement dit, P et P' sont associés. □

Nous insistons sur le fait que l'hypothèse d'irréductibilité dans la proposition 4.3.10 ne peut être évitée. Un contre-exemple simple, dans le cas où elle n'est pas vérifiée est donné par les polynômes de Ore $P = X^2 - 1$ et $P' = (X-i)(X+i) = X^2 - 2iX + 1$ dans $\mathbb{C}[X, \mathrm{conj}]$. Étant donné que P est central, on a $N_{\mathrm{rd}}(P) = P^2 = (X^2-1)^2$. D'autre part, $N_{\mathrm{rd}}(P') = N_{\mathrm{rd}}(X-i)N_{\mathrm{rd}}(X+i) = (1 - X^2)^2$ d'après le calcul de l'exemple 4.3.4. Ainsi $N_{\mathrm{rd}}(P) = N_{\mathrm{rd}}(P')$. Par contre, il se trouve que P et P' ne sont pas associés. En effet, supposons par l'absurde que ce soit le cas, c'est-à-dire que l'on puisse trouver quatre nombre complexes a, b, c et d tels que $P \cdot (aX + b) = (cX + d) \cdot P'$ avec $\mathrm{RGCD}(aX + b, P') = 1$. En développant et en identifiant les coefficients, on obtient les quatre équations

$$a = c \quad ; \quad b = -2ic + d \quad ; \quad -a = c - 2id \quad ; \quad -b = d$$

desquelles on déduit $b = ia$. Ainsi $aX + b = a(X + i)$ et ce dernier polynôme n'est pas premier avec P' comme il devrait l'être pour que P et P' soient associés.

4.3.3 Une autre définition de la norme réduite

Précédemment, nous avons déjà défini la norme réduite d'un polynôme de Ore comme une norme (au sens classique) relativement à une certaine extension. Il existe une autre interprétation de la norme réduite comme polynôme caractéristique d'une certaine application. Dans certains cas, cette définition alternative peut s'avérer plus facile à calculer ou fait apparaître plus clairement certaines propriétés.

Si V est un espace vectoriel de dimension finie sur un corps et si $f : V \to V$ est une application linéaire, nous notons $\chi_f(T) = \det(T - f)$ le polynôme caractéristique de f. Remarquons, en

particulier, que notre choix de normalisation est tel que $\chi_f(T)$ est toujours un polynôme unitaire, quelle que soit la dimension de V.

Avant d'énoncer le résultat principal de cette partie, rappelons que nous avons avons posé $Z(X) = X^r$ dans le cas d'un endomorphisme et $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X)$ dans le cas d'une dérivation, de manière à ce que le centre $\mathcal{Z}$ de $\mathcal{A}$ soit égal à $F[Z(X)]$.

Proposition 4.3.11. *Soit $P \in \mathcal{A}$. Soit $z_P : \mathcal{A}/\mathcal{A}P \to \mathcal{A}/\mathcal{A}P$ l'application de multiplication par $Z(X)$. Alors*

$$N_{rd}(P) = (-1)^{(r-1)\cdot \deg P} \cdot N_{K/F}(lc(P)) \cdot \chi_{z_P}(Z(X)).$$

Démonstration. Si P s'écrit sous la forme $P = BA$ avec $A, B \in \mathcal{A}$, on sait par la proposition 1.4.4 que les deux lignes du diagramme commutatif suivant sont exactes :

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \mathcal{A}/\mathcal{A}B & \longrightarrow & \mathcal{A}/\mathcal{A}P & \longrightarrow & \mathcal{A}/\mathcal{A}A & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle z_B} & & \downarrow{\scriptstyle z_P} & & \downarrow{\scriptstyle z_A} & & \\ 0 & \longrightarrow & \mathcal{A}/\mathcal{A}B & \longrightarrow & \mathcal{A}/\mathcal{A}P & \longrightarrow & \mathcal{A}/\mathcal{A}A & \longrightarrow & 0 \end{array}$$

On en déduit que $\chi_{z_P}(Z(X)) = \chi_{z_B}(Z(X)) \cdot \chi_{z_A}(Z(X))$. Étant donné que les quantités $N_{\mathrm{rd}}(P)$, $(-1)^{(r-1)\cdot \deg P}$ et $N_{K/F}(\mathrm{lc}(P))$ sont également multiplicatives par rapport à P, il suffit de démontrer la proposition dans le cas où P est un polynôme de Ore irréductible.

Supposons donc que P est irréductible. Écrivons $N_{\mathrm{rd}}(P) = N(Z(X))$ pour un certain polynôme $N \in F[T]$ (la lettre T désignant une nouvelle variable auxiliaire). Du fait que $N_{\mathrm{rd}}(P)$ est un multiple de P (*cf* proposition 4.3.5), on déduit que $N(z_P) = 0$. Autrement dit N est un polynôme annulateur de l'endomorphisme z_P. De plus, on sait par le corollaire 4.3.9 que N s'écrit sous la forme N_0^e pour un certain polynôme irréductible $N_0 \in F[T]$ et un certain entier $e > 0$. On en déduit que χ_{z_P} est, lui aussi, nécessairement une puissance de N_0. Par ailleurs, son degré est égal à celui de P, de même que le degré de N d'après la proposition 4.3.6 (voir aussi la remarque 4.3.7 juste après). Il en résulte que $N = c \cdot \chi_{z_P}$ pour un certain scalaire non nul $c \in F$. En évaluant pour $T = Z(X)$, on obtient $N_{\mathrm{rd}}(P) = c \cdot \chi_{z_P}(Z(X))$. Il ne reste plus qu'à déterminer la constance c ce qui se fait en identifiant les coefficients dominants que l'on connaît grâce à la proposition 4.3.6. □

Exemple 4.3.12. Dans la continuité de l'exemple 4.3.4, calculons la norme de réduite de $X-a \in \mathbb{C}[X, \mathrm{conj}]$ (pour $a \in \mathbb{C}$) à l'aide de la proposition 4.3.11 ci-dessus. Le quotient

$$V = \mathbb{C}[X, \mathrm{conj}]/\mathbb{C}[X, \mathrm{conj}](X - a)$$

est un $\mathbb{C}$-espace vectoriel de dimension 1 qui s'identifie à $\mathbb{C}$ *via* l'application $\mathbb{C}[X, \mathrm{conj}] \to \mathbb{C}$ qui fait à correspondre à un polynôme de Ore P le reste dans la division euclidienne à droite de P par $X-a$. La matrice de z_{X-a} (définie comme la multiplication par $Z(X) = X^2$ dans V) est ainsi la matrice 1×1 dont l'unique coefficient est le reste de la division euclidienne de X^2 par $X-a$, noté $X^2 \,\%\, (X - a)$ dans la suite. Or, un calcul simple donne :

$$X^2 = (X + \bar{a})(X - a) + |a|^2$$

d'où on obtient $X^2 \,\%\, (X-a) = |a|^2$. Le polynôme caractéristique de z_{X-a} est ainsi $T - |a|^2$. En appliquant la proposition 4.3.11, on obtient $N_{\mathrm{rd}}(P) = -X^2 + |a|^2$, ce qui confirme une fois de plus le résultat que nous avions déjà obtenu.

La proposition 4.3.11 est particulièrement intéressante dans le cas d'une dérivation (*i.e.* $\mathcal{A} = K[X, \partial]$) car elle permet de relier la norme réduite à d'autres invariants classiques. Précisément, étant donné un $K[X, \partial]$-module M, il est usuel de définir la *p-courbure* de M comme l'application

$c : M \to M$ donné par la multiplication par X^p. Cette application vérifie la loi de Leibniz pour la dérivation ∂^p :

$$\forall f \in K, \quad \forall x \in M, \quad c(fx) = \partial^p(f)\, x + f\, c(x).$$

En particulier, lorsque $\partial^p = 0$ — ce qui est par exemple pour $\partial = \frac{d}{dt}$ agissant sur $\mathbb{F}_p(t)$ — la p-courbure est une application linéaire. Sous cette même hypothèse, et en supposant de plus que $\partial \neq 0$, la proposition 4.3.11 nous apprend que le polynôme caractéristique de la p-courbure du module $\mathcal{A}/\mathcal{A}P$ est formé relié à la norme réduite de P puisque son évaluation en X^p s'identifie à $N_{\text{rd}}(P)$ à une constante multiplicative près.

4.4 Application à la factorisation

On conserve les notations et les hypothèses de la partie précédente. On pose $\mathcal{A} = K[X, \theta]$ (resp. $\mathcal{A} = K[X, \partial]$) dans le cas d'un endomorphisme (resp. d'une dérivation), on appelle F le sous-corps de K fixé par θ (resp. annulé par ∂) et on note $\mathcal{Z}$ le centre de $\mathcal{A}$. On suppose que $\mathcal{Z}$ n'est pas réduit à F. D'après la proposition 4.1.1 (resp. la proposition 4.1.3), $\mathcal{Z} = F[Z(X)]$ où $Z(X)$ est une puissance de X (resp. $Z(X) = \text{Ann}(\partial)_{\text{lin}}(X)$). On note $r = \deg Z(X)$.

4.4.1 De la factorisation dans $\mathcal{Z}$ à la factorisation dans $\mathcal{A}$

Les propriétés de la norme réduite $N_{\text{rd}} : \mathcal{A} \to \mathcal{Z}$ (construite et étudiée au §4.3) permettent de transférer une partie des questions de factorisations de l'anneau $\mathcal{A}$ à $\mathcal{Z}$. Ce phénomène de transfert est particulièrement intéressant puisque $\mathcal{Z}$ est un anneau de polynômes commutatif dans lequel les propriétés de factorisation sont très bien comprises. Le lemme clé qui permet ce transfert est le suivant.

Lemme 4.4.1. *Soit $P \in \mathcal{A}$ et soit $N \in \mathcal{Z}$ un diviseur de $N_{rd}(P)$ de degré strictement positif. Alors* $\text{RGCD}(N, P)$ *est un diviseur à droite de P de degré strictement positif.*

Démonstration. Par définition, il est clair que $\text{RGCD}(N, P)$ est un diviseur à droite de P. Il est de degré strictement positif grâce au lemme 4.3.8. □

Le lemme 4.4.1 peut s'utiliser comme suit. Soit $P \in \mathcal{A}$. On considère un facteur irréductible $N_1 \in \mathcal{Z}$ de $N_{\text{rd}}(P)$. D'après le lemme 4.4.1, le polynôme de Ore $D_1 = \text{RGCD}(N_1, P)$ est un diviseur à droite de P de degré strictement positif. Par ailleurs, du fait que D_1 divise N_1, on déduit que $N_{\text{rd}}(D_1)$ divise $N_{\text{rd}}(N_1) = N_1^r$. On a ainsi isolé un facteur de P qui ne contribue à la norme réduite qu'au niveau de la place N_1. À présent, on écrit P comme un produit $P = P_1 D_1$ (avec $P_1 \in \mathcal{A}$) et on applique à nouveau la même méthode avec le facteur restant P_1. *In fine,* on aboutit à une factorisation de la forme :

$$P = D_m \cdot D_{m-1} \cdots D_2 \cdot D_1$$

où chaque facteur D_i a la propriété de diviser un polynôme central N_i qui est irréductible dans $\mathcal{Z}$. L'étude des propriétés de factorisation de ces polynômes particuliers fera l'objet du §4.4.2 ; en attendant, pour nous échauffer, examinons un cas particulièrement favorable pour lequel on peut démontrer à moindre frais que les facteurs D_i sont tous irréductibles.

Ce cas particulier est celui où la norme réduite $N_{\text{rd}}(P)$ est un polynôme séparable, c'est-à-dire se factorise sous la forme :

$$N_{\text{rd}}(P) = \lambda \cdot N_m \cdot N_{m-1} \cdots N_2 \cdot N_1$$

où λ est une constante non nulle et où les N_i sont unitaires, irréductibles dans $\mathcal{Z}$ et deux à deux distincts. Comme précédemment, posons $D_1 = \text{RGCD}(P, N_1)$; c'est un diviseur à droite de P de degré strictement positif. De plus, sa norme réduite $N_{\text{rd}}(D_1)$ divise simultanément

$N_{\text{rd}}(P)$ et $N_{\text{rd}}(N_1) = N_1^r$. Comme, en outre, $N_{\text{rd}}(D_1)$ n'est pas constant, on a nécessairement $N_{\text{rd}}(D_1) = \pm N_1$ (où, de surcroît, le signe est connu grâce à la proposition 4.3.6). Puisque N_{rd} est une fonction multiplicative, il résuite du fait que $N_{\text{rd}}(D_1)$ est irréductible que le polynôme de Ore D_1 est, lui-même, irréductible. Écrivons maintenant $P = P_1 D_1$ avec $P_1 \in \mathcal{A}$. On a :

$$N_{\text{rd}}(P_1) = \frac{N_{\text{rd}}(P)}{N_{\text{rd}}(D_1)} = \pm\, \lambda \cdot N_m \cdot N_{m-1} \cdots N_2$$

et on peut réappliquer le même raisonnement avec P_1. Au bout de m étapes, on obtient une factorisation *complète* de P qui s'écrit $P = \text{lc}(P) \cdot D_m \cdot D_{m-1} \cdots D_2 \cdot D_1$ avec $N_{\text{rd}}(D_i) = \pm N_i$ pour tout i.

Bien sûr, si σ est une permutation de $\{1, 2, \ldots, m\}$, la norme $N_{\text{rd}}(P)$ se factorise également sous la forme $N_{\text{rd}}(P) = c \cdot N_{\sigma(m)} \cdot N_{\sigma(m-1)} \cdots N_{\sigma(2)} \cdot N_{\sigma(1)}$, factorisation à partir de laquelle on déduit, exactement de la même manière que précédemment, une factorisation de P que l'on note :

$$P = \text{lc}(P) \cdot D_m^{(\sigma)} \cdot D_{m-1}^{(\sigma)} \cdots D_2^{(\sigma)} \cdot D_1^{(\sigma)}$$

avec $N_{\text{rd}}(D_i^{(\sigma)}) = \pm N_{\sigma(i)}$. Mieux encore, *toutes* les factorisations de P en produits d'irréductibles unitaires sont de cette forme. En effet, si $P = \text{lc}(P) \cdot P_m \cdots P_2 \cdot P_1$ est une telle factorisation, la norme réduite de P_1 est un facteur de $N_{\text{rd}}(P)$ qui, d'après le corollaire 4.3.9, est une puissance d'un polynôme irréductible. Nécessairement, on a donc $N_{\text{rd}}(P_1) = \pm N_i$ pour un certain entier i et, par suite, $P_1 = \text{RGCD}(P, N_i)$ pour cet indice i particulier. Appliquant à nouveau plusieurs fois le même argument, on construit de proche en proche une permutation σ de $\{1, 2, \ldots, m\}$ pour laquelle $P_i = D_i^{(\sigma)}$ pour tout i.

Remarque 4.4.2. En accord avec le théorème de factorisation de Ore (*cf* théorème 1.5.6), on observe qu'étant données deux permutations σ et σ' de $\{1, \ldots, m\}$, les polynômes de Ore $D_i^{(\sigma)}$ ($1 \leq i \leq m$) sont bien associés (au sens de la définition 1.4.8), à réordonnement près, aux polynômes de Ore $D_i^{(\sigma')}$ ($1 \leq i \leq m$). En effet, on a vu que l'ensemble des $\pm N_{\text{rd}}(D_i^{(\sigma)})$ (le signe étant choisi de sorte que le polynôme obtenu soit unitaire), d'une part, et celui des $\pm N_{\text{rd}}(D_i^{(\sigma')})$ (même convention pour le choix du signe), d'autre part, coïncident tous les deux avec l'ensemble des N_i. La proposition 4.3.10 s'applique et assure bien que les $D_i^{(\sigma)}$ et $D_i^{(\sigma')}$ sont associés par paires.

4.4.2 Factorisation des diviseurs d'irréductibles centraux

Grâce à la réduction que nous venons de faire, il ne nous reste plus qu'à étudier la factorisation des polynômes de Ore P de degré strictement positif qui sont des diviseurs d'un polynôme central irréductible dans $\mathcal{Z}$.

Soit $N \in \mathcal{Z}$ un polynôme central irréductible. Dans le cas d'un endomorphisme (*i.e.* $\mathcal{A} = K[X, \theta]$), on suppose en outre que N n'est pas proportionnel à X^r. On note $\deg_{\mathcal{Z}} N$ le degré de N, considéré comme un polynôme en la variable $Z(X)$, *i.e.* si on écrit $N = N_0(Z(X))$ (où on rappelle que $Z(X)$ est la variable sur $\mathcal{Z}$), alors $\deg_{\mathcal{Z}} N = \deg N_0$. L'algèbre quotient $\mathcal{A}/N\mathcal{A}$ est une algèbre simple centrale sur le corps $\mathcal{Z}/N\mathcal{Z}$. D'après le théorème d'Artin–Wedderburn (*cf* théorème 3.2.1), on a un isomorphisme de $\mathcal{Z}/N\mathcal{Z}$-algèbres $\mathcal{A}/N\mathcal{A} \simeq M_{d(N)}(D_N)$ pour un entier $d(N)$ et une algèbre à divisions D_N de centre $\mathcal{Z}/N\mathcal{Z}$. Un calcul de dimension montre que $[D_N : \mathcal{Z}/N\mathcal{Z}] = e(N)^2$ avec $e(N) = \frac{r}{d(N)}$.

Dans le cas où $\mathcal{A} = K[X, \theta]$ et où $N = \lambda X^r$ avec $\lambda \in F$, $\lambda \neq 0$, on convient que $e(N) = 1$. Dans un léger abus, convenons également que, dans ce cas, $\mathbb{P}^{n-1}(D_N)$ soit un ensemble de cardinal 1 quelque soit l'entier n (sans, pour autant, que D_N ne soit défini).

Théorème 4.4.3. *Soit N un polynôme irréductible dans $\mathcal{Z}$.*
Pour un polynôme de Ore $P \in \mathcal{A}$ divisant N :

(i) *le degré de P est divisible par $e(N) \cdot \deg_{\mathcal{Z}} N$;*

(ii) *le polynôme P est irréductible dans $\mathcal{A}$ si et seulement si $\deg P = e(N) \cdot \deg_{\mathcal{Z}} N$;*

(iii) *l'ensemble des diviseurs à droite de P qui sont irréductibles et unitaires est en bijection naturelle avec l'espace projectif*[3] *$\mathbb{P}^{n-1}(D_N)$ pour $n = \frac{\deg P}{e(N)\cdot \deg_{\mathcal{Z}} N}$;*

(iv) *$N_{rd}(P) = \lambda N^m$ pour $\lambda \in F$, $\lambda \neq 0$ et $m = \frac{\deg P}{\deg_{\mathcal{Z}} N} = n \cdot e(N)$.*

Démonstration. On vérifie à la main sans difficulté que la proposition est valide lorsque $X = \lambda X^r$ ($\lambda \in F$, $\lambda \neq 0$) dans le cas où $\mathcal{A} = K[X, \theta]$.

Pour simplifier les notations, posons $d = \deg_{\mathcal{Z}} N$ pour le reste de la démonstration. Le théorème d'équivalence de Morita (*cf* corollaire 3.2.7) assure qu'il y a une équivalence de catégories entre les catégories $\mathrm{Mod}^{\mathrm{g}}_{D_N}$ et $\mathrm{Mod}^{\mathrm{g}}_{\mathcal{A}/N\mathcal{A}}$ (où on rappelle que la notation $\mathrm{Mod}^{\mathrm{g}}_R$ désigne la catégorie des modules à gauche sur R) qui est compatible à la dimension dans le sens où, si $X \in \mathrm{Mod}^{\mathrm{g}}_{D_N}$ et $Y \in \mathrm{Mod}^{\mathrm{g}}_{\mathcal{A}/N\mathcal{A}}$ se correspondent *via* cette équivalence, on a $\dim_{\mathcal{Z}/N\mathcal{Z}} Y = r \cdot \dim_{\mathcal{Z}/N\mathcal{Z}} X = r \cdot e(N) \cdot \dim_{D_N} X$. Du fait $[K : F] = r$ et $[\mathcal{Z}/N\mathcal{Z} : F] = d$, on déduit de l'égalité précédente que :

$$\dim_K Y = d \cdot e(N) \cdot \dim_{D_N} X. \tag{31}$$

Considérons à présent $P \in \mathcal{A}$, un diviseur de N. Le $\mathcal{A}$-module quotient $M_P = \mathcal{A}/\mathcal{A}\mathcal{P}$ est annulé par N et hérite ainsi naturellement d'une structure de module sur $\mathcal{A}/N\mathcal{A}$. Soit V le D_N-espace vectoriel à gauche correspondant à M_P par l'équivalence de Morita. L'équation aux dimensions (31) s'écrit ici :

$$\deg P = d \cdot e(N) \cdot \dim_{D_N} V. \tag{32}$$

Ceci démontre (i). L'énoncé (ii) en découle également après avoir remarqué que l'équivalence de Morita met en correspondre les $(\mathcal{A}/N\mathcal{A})$-modules simples avec les D_n-espaces vectoriels de dimension 1.

L'équivalence de Morita induit une bijection entre les quotients de M_P et ceux de V et donc, également, entre les quotients simples de M_P et les quotients de V de dimension 1. Or, d'après la proposition 1.4.4, les quotients de M_P sont en bijection avec les diviseurs à droite de P qui sont unitaires ; les quotients simples de M_P sont ainsi en bijection avec les diviseurs à droite de P qui sont à la fois unitaires et irréductibles. Par ailleurs, par définition, les quotients de V de dimension 1 sont paramétrés par l'espace projectif $\mathbb{P}(V)$. Comme V est de dimension $n = \frac{\deg P}{d\cdot e(N)}$ d'après la formule (32), on a $\mathbb{P}(V) \simeq \mathbb{P}^{n-1}(D_N)$ et l'assertion (iii) est démontrée.

Enfin, pour démontrer (iv), il suffit de remarquer que, puisque P divise N, on trouve que $N_{\mathrm{rd}}(P)$ divise $N_{\mathrm{rd}}(N) = N^r$ et donc $N_{\mathrm{rd}}(P) = \lambda N^m$ pour un certain $\lambda \in F$ et un certain entier m. On trouve finalement la valeur de m en comparant les degrés. □

Il résulte également de la démonstration du théorème 4.4.3 que si $P \in \mathcal{A}$ est un polynôme unitaire de Ore divisant N de degré $n \cdot d \cdot e(N)$, toute factorisation complète de P en produits de facteurs irréductibles unitaires est de la forme $P = P_n \cdot P_{n-1} \cdots P_2 \cdot P_1$ où les P_i sont tous de degré $d \cdot e(N)$ et de norme réduite $N^{e(N)}$. De plus, ces factorisations sont naturellement paramétrées par l'espace des drapeaux complets de D_N^n.

Exemple 4.4.4. Plaçons-nous dans $\mathcal{A} = \mathbb{C}[X, \mathrm{conj}]$ dont le centre $\mathcal{Z}$ est égal à $\mathbb{R}[X^2]$. Soit $N = X^2 - c \in \mathcal{Z}$ avec $c \in \mathbb{R}$, $c \neq 0$. D'après le calcul qui a été présenté dans l'exemple 4.2.2, l'algèbre quotient $\mathcal{Z}/N\mathcal{Z}$ est isomorphe à $M_2(\mathbb{R})$ (resp. à $\mathbb{H}$) lorsque $c > 0$ (resp. lorsque $c < 0$). On en déduit que $e(N) = 1$ si $c > 0$ et que $e(N) = 2$ sinon.

Dans le cas où $c < 0$, le théorème 4.4.3 nous assure donc que N est irréductible dans $\mathcal{A}$. Par contre, lorsque $c > 0$, elle nous dit que N se factorise comme le produit de deux facteurs

3. On renvoie à l'appendice A.2 pour un topo rapide sur les espaces projectifs sur des algèbres à divisions non nécessairement commutatives.

irréductibles unitaires de degré 1. En outre, ces factorisations sont paramétrées par l'espace projectif $\mathbb{P}^1(\mathbb{R})$. Ceci confirme le calcul que nous avions fait dans l'exemple 1.1.4 où nous avions vu que les factorisations de N s'écrivent sous la forme :

$$X^2 - c = (X - a) \cdot (X + \bar{a})$$

où a est un nombre complexe de module $\sqrt{c}$.

Le théorème 4.4.3 nous permet également de généraliser l'exemple de « factorisation séparable » que nous avions étudié au §4.4.1. Précisément, considérons un polynôme de Ore $P \in \mathcal{A}$ dont la norme réduite s'écrit :

$$N_{\text{rd}}(P) = \lambda \cdot N_m^{e(N_m)} \cdot N_{m-1}^{e(N_{m-1})} \cdots N_2^{e(N_2)} \cdot N_1^{e(N_1)}$$

où $\lambda \in F$, $\lambda \neq 0$ et les N_i sont des polynômes centraux deux et deux distincts et irréductibles dans $\mathcal{Z}$. Comme en §4.4.1, posons $D_1 = \text{RGCD}(P, N_1)$. Il s'agit d'un diviseur de P de degré strictement positif. De plus, d'après le théorème 4.4.3, sa norme réduite $N_{\text{rd}}(D_1)$ doit être égal à $\pm N_1^{n \cdot e(N_1)}$ pour un certain entier n. Mais comme $N_{\text{rd}}(D_1)$ divise $N_{\text{rd}}(P)$, on a nécessairement $n = 1$. En appliquant à nouveau le théorème 4.4.3, on en déduit que D_1 est irréductible. On écrit à présent $P = P_1 D_1$ avec $P_1 \in \mathcal{A}$. La multiplicativité de la norme réduite implique :

$$N_{\text{rd}}(P_1) = \pm\lambda \cdot N_m^{e(N_m)} \cdot N_{m-1}^{e(N_{m-1})} \cdots N_2^{e(N_2)}$$

et on peut donc appliquer la même construction à partir de P_1 : on pose $D_2 = \text{RGCD}(P_1, N_2)$ et, comme précédemment, on démontre que D_2 est irréductible dans $\mathcal{A}$. De proche en proche, on parvient à une factorisation complète de P qui prend la forme $P = \text{lc}(P) \cdot D_m \cdot D_{m-1} \cdots D_1$ avec $N_{\text{rd}}(D_i) = \pm N_i^{e(N_i)}$ pour tout i. De plus, comme cela a déjà été discuté au §4.4.1, il est associé à chaque permutation σ de $\{1, \ldots, m\}$ une nouvelle factorisation complète de P qui s'écrit :

$$P = \text{lc}(P) \cdot D_m^{(\sigma)} \cdot D_{m-1}^{(\sigma)} \cdots D_2^{(\sigma)} \cdot D_1^{(\sigma)}.$$

Ces factorisations sont deux à deux différentes et toutes les factorisations complètes de P s'obtiennent de cette manière ; il y en a ainsi exactement $m!$.

4.4.3 L'image de la norme réduite

Le théorème 4.4.3 permet enfin de décrire complètement l'image de la norme réduite $N_{\text{rd}} : \mathcal{A} \to \mathcal{Z}$. Pour un polynôme irréductible N de $\mathcal{Z}$, notons $v_N : \mathcal{Z} \to \mathbb{N} \cup \{+\infty\}$ la valuation N-adique ; autrement dit, pour $S \in \mathcal{Z}$, par définition, $v_N(S)$ est le plus grand entier v pour lequel N^v divise P. Commençons par isoler un lemme qui est un raffinement du corollaire 4.3.9 que nous avions vu au §4.3.2 où nous avons regroupé quelques propriétés fondamentales de la norme réduite.

Lemme 4.4.5. *Soit $P \in \mathcal{A}$ un polynôme de Ore irréductible. Alors $N_{rd}(P) = \lambda\, N^{e(N)}$ pour $\lambda \in F$ et pour un polynôme irréductible N de $\mathcal{Z}$.*

Démonstration. Soit $N \in \mathcal{Z}$ un diviseur irréductible de $N_{\text{rd}}(P)$. Le lemme 4.3.8 assure que $\text{PGCD}(P, N) \neq 1$. Comme P est irréductible, on en déduit que P divise N. Le théorème 4.4.3 permet alors de conclure. □

Proposition 4.4.6. *Un polynôme $S \in \mathcal{Z}$ est dans l'image de la norme réduite $N_{rd} : \mathcal{A} \to \mathcal{Z}$ si et seulement si les deux conditions suivantes sont satisfaites :*

(1) $(-1)^{(r-1)\cdot \deg_{\mathcal{Z}} S} \text{lc}(S)$ *est une norme dans l'extension K/F, et*

(2) $e(N)$ *divise $v_N(S)$ pour tout polynôme irréductible N de $\mathcal{Z}$.*

Démonstration. Considérons $P \in \mathcal{A}$. Le fait que $(-1)^{(r-1)\cdot \deg_{\mathcal{Z}} N_{\mathrm{rd}}(P)} \mathrm{lc}(N_{\mathrm{rd}}(P))$ soit une norme dans l'extension K/F résulte directement de la proposition 4.3.6. Écrivons, à présent, la décomposition de P en produits de facteurs irréductibles : $P = P_1 \cdots P_2 \cdots P_m$. Par le lemme 4.4.5, on a $N_{\mathrm{rd}}(P_i) = \lambda_i N_i^{e(N_i)}$ pour un $\lambda_i \in F$ et un polynôme irréductible N_i de $\mathcal{Z}$. Il en résulte que :

$$N_{\mathrm{rd}}(P) = \lambda N_1^{e(N_1)} N_2^{e(N_2)} \cdots N_m^{e(N_m)}$$

pour un certain $\lambda \in F$ et, par suite, que $e(N)$ divise $v_N(N_{\mathrm{rd}}(P))$ pour tout polynôme irréductible N de $\mathcal{Z}$. Réciproquement, soit $S \in \mathcal{Z}$ dont la décomposition en produit de facteurs irréductibles dans $\mathcal{Z}$ s'écrit :

$$S = (-1)^{(r-1)\cdot \deg_{\mathcal{Z}} S} N_{K/F}(\lambda) \cdot N_1^{d_1 e(N_1)} N_2^{d_2 e(N_2)} \cdots N_m^{d_m e(N_m)}$$

pour $\lambda \in K$ et des polynômes irréductibles unitaires $N_i \in \mathcal{Z}$ deux à deux distincts. Pour $i \in \{1, \ldots, m\}$, soit P_i un diviseur irréductible unitaire de N_i. Du théorème 4.4.3 et de la proposition 4.3.6, on déduit que $N_{\mathrm{rd}}(P_i) = (-1)^{(r-1)\cdot \deg P_i} N_i^{e(N_i)}$. Par multiplicativité, $N_{\mathrm{rd}}(\lambda P_1 P_2 \cdots P_m) = S$, ce qui démontre que S est dans l'image de la norme réduite. □

Corollaire 4.4.7. *Si K est un corps fini*[4]*, alors l'application $N_{rd} : \mathcal{A} \to \mathcal{Z}$ est surjective.*

Démonstration. D'après la proposition 4.4.6, il suffit de démontrer que, sous l'hypothèse que K est un corps fini, d'une part, la norme $N_{K/F} : K^\star \to F^\star$ est surjective et, d'autre part, que pour tout polynôme irréductible $N \in \mathcal{Z}$, on a $e(N) = 1$. La première assertion est un résultat classique de la théorie des corps finis. Quant à la seconde, elle résulte du fait que toute algèbre à divisions finie est commutative ; c'est le théorème de Wedderburn. □

4.5 Calcul de la norme réduite d'un polynôme de degré 1

On conserve les notations et hypothèses de la partie précédente : $\mathcal{A}$ désigne un anneau de Ore qui peut être soit $K[X, \theta]$ pour un endomorphisme $\theta : K \to K$ soit $K[X, \partial]$ pour une dérivation $\partial : K \to K$. On note $\mathcal{Z}$ le centre de $\mathcal{A}$ et on suppose que $\mathcal{Z}$ est d'indice fini dans $\mathcal{A}$. Les propositions 4.1.1 et 4.1.3 donnent une description explicite de $\mathcal{Z}$ qui s'identifie dans tous les cas à $F[Z(X)]$ où F est un sous-corps de K et où $Z(X) \in \mathcal{A}$. Dans ces conditions, nous avons construit une application de norme réduite $N_{\mathrm{rd}} : \mathcal{A} \to \mathcal{Z}$ dont nous avons longuement étudié les propriétés dans les paragraphes précédents.

Le but de ce numéro est de calculer la norme réduite d'un polynôme de Ore unitaire de degré 1, c'est-à-dire de la forme $X - a$ pour un certain $a \in K$. Hormis l'intérêt intrinsèque de l'exercice, ce calcul aura des applications dans la suite lorsque nous chercherons à estimer les invariants $e(N)$ qui sont apparus au §4.4.2.

Le cas d'un endomorphisme. On commence par traiter le cas (nettement plus simple) d'un anneau de Ore associé à un automorphisme $\theta : K \to K$ d'ordre fini, noté r. Dans ce cas, le centre de $\mathcal{A} = K[X, \theta]$ est $\mathcal{Z} = F[X^r]$.

Proposition 4.5.1. *Dans le cas où $\mathcal{A} = K[X, \theta]$, pour tout $a \in K$, on a :*

$$N_{rd}(X - a) = (-1)^{r-1} \cdot \left(X^r - N_{K/F}(a)\right)$$

où $N_{K/F}(a) = a \cdot \theta(a) \cdots \theta^{r-1}(a)$ désigne la norme de K à F de a.

4. Ceci ne peut pas se produire dans le cas où $\mathcal{A} = K[X, \partial]$ avec $\partial \neq 0$.

Démonstration. Par la proposition 4.3.6, on sait déjà que $N_{\mathrm{rd}}(X-a)$ est un polynôme en X^r de degré 1 et que son coefficient dominant est $(-1)^{r-1}$. Il ne reste donc qu'à déterminer le coefficient constant. Pour cela, on considère la sous-algèbre commutative $C = F[X]$ de $\mathcal{A}$. Par définition, la norme réduite de $X-a$ est égale au déterminant de l'application C-linéaire $m_{X-a} : \mathcal{A} \to \mathcal{A}$, $P \mapsto P \cdot (X-a)$. Clairement m_{X-a} se réduit modulo X sur l'application $m_{-a} : K \to K$ de multiplication par $(-a)$. On en déduit que :

$$\det m_{X-a} \equiv N_{K/F}(-a) = (-1)^r \, N_{K/F}(a) \pmod{X}.$$

La proposition s'ensuit. □

Le cas d'une dérivation. Venons-en maintenant au cas différentiel, c'est-à-dire au cas où $\mathcal{A} = K[X, \partial]$ pour une dérivation $\partial : K \to K$. Il s'avère que le cas où $\partial^p = 0$ est plus simple et nous servira comme calcul intermédiaire pour aborder le cas général. Concentrons-nous donc, dans un premier temps, sur celui-ci. Si l'on suppose en outre que $\partial \neq 0$, le centre de $\mathcal{A}$ est $\mathcal{Z} = F[X^p]$ où F est le sous-corps des constantes de K.

Lemme 4.5.2. *Dans le cas où $\mathcal{A} = K[X, \partial]$ avec $\partial \neq 0$ et $\partial^p = 0$, pour tout $f \in K$, on a :*

$$N_{rd}(X - f) = X^p - \big(f^p + \partial^{p-1}(f)\big).$$

Démonstration. Commençons par supposer que f admet une primitive, c'est-à-dire qu'il existe $g \in K$ tel que $\partial(g) = f$. On a alors $\partial^{p-1}(f) = \partial^p(g) = 0$ et il s'agit donc de montrer que $N_{\mathrm{rd}}(X-f) = X^p - f^p$. Lorsque $f = 0$, cela résulte directement de la définition de la norme réduite comme norme par rapport à l'extension $\mathcal{A}/K[X^p]$. Sinon $\partial(g) \neq 0$, ce qui assure que $g \notin F$. Comme l'extension K/F est de degré premier p, on en déduit que $K = F(g)$ et donc que la famille $(1, g, \ldots, g^{p-1})$ forme une base de K sur F. Introduisons la sous-algèbre commutative $C = F[X]$ de $\mathcal{A}$. La famille $(1, g, \ldots, g^{p-1})$ est également une base de $\mathcal{A}$ sur C. De plus, les relations :

$$g^i \cdot (X-f) = (X-f) \cdot g^i + if \cdot g^{i-1} \quad (0 \leq i < p)$$

montrent que, dans cette base, la matrice de la multiplication à droite par $X-f$ est une matrice triangulaire supérieure dont tous les coefficients de la diagonale sont égaux à $X-f$. On en déduit que son déterminant, qui est la norme réduite de $X-f$, vaut $(X-f)^p = X^p - f^p$.

Passons maintenant au cas général. On considère, cette fois-ci, la sous-algèbre commutative $C = K[X^p]$ et on voit $\mathcal{A}$ comme un C-module libre de base $(1, X, \ldots, X^{p-1})$. Pour $i \in \{0, 1, \ldots, p-1\}$, on a :

$$X^i \cdot (X-f) = X^{i+1} - \sum_{j=0}^{i} \binom{i}{j} \partial^{i-j}(f) X^j.$$

En prenant le déterminant, on en déduit que :

$$N_{\mathrm{rd}}(X-f) = X^p - D(f, \partial(f), \ldots, \partial^{p-1}(f))$$

où $D \in \mathbb{F}_p[X_0, \ldots, X_{p-1}]$ est le polynôme en p variables défini par :

$$D(X_0, X_1, \ldots, X_{p-1}) = \begin{vmatrix} X_0 & X_1 & X_2 & X_3 & \cdots & X_{p-2} & X_{p-1} \\ 1 & X_0 & 2X_1 & 3X_2 & \cdots & \binom{p-2}{1} X_{p-3} & \binom{p-1}{1} X_{p-2} \\ & 1 & X_0 & 2X_1 & \cdots & \binom{p-2}{2} X_{p-4} & \binom{p-1}{2} X_{p-3} \\ & & 1 & X_0 & \cdots & \binom{p-2}{3} X_{p-5} & \binom{p-1}{3} X_{p-4} \\ & & & \ddots & \ddots & \vdots & \vdots \\ & & & & \ddots & X_0 & \binom{p-1}{p-2} X_1 \\ & & & & & 1 & X_0 \end{vmatrix}.$$

Écrivons $D = X_0^p + X_{p-1} + Q(X_0, \dots, X_{p-1})$. On veut démontrer que Q est le polynôme nul. En développant le déterminant ci-dessus selon la dernière colonne, on s'aperçoit que Q ne dépend pas de sa dernière variable X_{p-1}. Plaçons-nous à présent dans l'anneau de Ore $\bar{\mathbb{F}}_p(t)[X, \frac{d}{dt}]$ et considérons l'élément $f = a_0 + a_1 t + \cdots + a_{p-2} t^{p-2} \in \bar{\mathbb{F}}_p(t)$ pour des constantes $a_i \in \mathbb{F}_p$. Clairement f admet une primitive. On en déduit, d'après la première partie de la démonstration, que $N_{\mathrm{rd}}(X - f) = X^p - f^p$ pour ce f particulier. Ainsi trouve-t-on $Q(f, f', \dots, f^{(p-2)}) = 0$ et, par suite, en évaluation en $t = 0$, que :

$$Q(a_0,\, a_1,\, 2a_2,\, 6a_3,\, \dots,\, (p-2)!\, a_{p-2}) = 0.$$

Comme ceci est vrai pour tout choix de $a_0, \dots, a_{p-2} \in \bar{\mathbb{F}}_p$, on obtient $Q = 0$ dans $\mathbb{F}_p[X_0, \dots, X_{p-2}]$ comme voulu. □

Le lemme précédent est intéressant car il peut se reformuler en terme de p-courbure, ce qui nous donnera la clé pour ôter l'hypothèse que nous avons faite sur ∂. Précisément, on a l'énoncé suivant.

Proposition 4.5.3. *Pour une dérivation $\partial : K \to K$ et $f \in K$, on a l'identité :*

$$(\partial + f\, \mathrm{id}_K)^p = \partial^p + \big(f^p + \partial^{p-1}(f)\big)\, \mathrm{id}_K$$

où $(\partial + f\, \mathrm{id}_K)^p = (\partial + f\, \mathrm{id}_K) \circ \cdots \circ (\partial + f\, \mathrm{id}_K)$ (p fois)

Démonstration. En développant la composée $(\partial + f\, \mathrm{id}_K)^p$, on voit que :

$$(\partial + f\, \mathrm{id}_K)^p = \partial^p + \sum_{i=0}^{p-1} Q_i(f, \partial(f), \dots, \partial^{p-1}(f)) \cdot \partial^i$$

où les $Q_i \in \mathbb{F}_p[X_0, \dots, X_{p-1}]$ sont des polynômes *universels* (c'est-à-dire ne dépendant pas de f) en p variables. À partir de là, on déduit que, pour démontrer le résultat escompté, on peut supposer que $K = \bar{\mathbb{F}}_p(t)$ et $\partial = \frac{d}{dt}$. On a alors $\partial^p = 0$ et le lemme 4.5.2 s'applique, donnant la relation :

$$N_{\mathrm{rd}}(X - f) = P(X^p) \quad \text{pour} \quad P(T) = T - \big(f^p + \partial^{p-1}(f)\big). \tag{33}$$

Par ailleurs, on sait grâce à la proposition 4.3.11 que $P(T)$ est le polynôme caractéristique de l'application $\nabla : M \to M$, $R \mapsto RX^p$ avec $M = K[X, \partial]/K[X, \partial]\,(X - f)$. De plus, on observe que M est K-espace vectoriel de dimension 1 engendré par la classe de 1, notée m. Un calcul immédiat conduit à $\nabla(ym) = (\partial + f\, \mathrm{id}_K)(y) \cdot m$, de sorte que $\nabla^p(ym) = (\partial + f\, \mathrm{id}_K)^p(y) \cdot m$. On déduit de cette dernière formule, d'une part, que l'application $(\partial + f\, \mathrm{id}_K)^p$ est K-linéaire et, d'autre part, que son polynôme caractéristique (ou, ce qui revient au même, celui de ∇^p) est $T - (\partial + f\mathrm{id}_K)^p(1)$. En comparant avec (33), on en déduit que $(\partial + f\mathrm{id}_K)^p(1) = f^p + \partial^{p-1}(f)$ et, par suite, que :

$$(\partial + f\, \mathrm{id}_K)^p = \big(f^p + \partial^{p-1}(f)\big)\, \mathrm{id}_K$$

puisque deux endomorphismes K-linéaires de K qui coïncident en 1 sont nécessairement égaux. La proposition est ainsi démontrée. □

Un des intérêts de la formulation utilisée dans la proposition 4.5.3 est qu'elle admet une extension naturelle à toute expression de la forme $P_{\mathrm{lin}}(\partial + f\, \mathrm{id}_K)$ où P_{lin} est un polynôme linéarisé (*cf* exemple 1.1.3). Précisément, rappelons que nous avions introduit l'anneau de Ore $F[Y, \mathrm{Frob}]$ et que nous avions vu que celui-ci agit sur l'ensemble des dérivations de K : si $P \in F[Y, \mathrm{Frob}]$ et si $\partial : K \to K$ est une dérivation, on peut former la dérivation $P \cdot \partial = P_{\mathrm{lin}}(\partial)$ où P_{lin} est le polynôme linéarisé associé à P (*cf* exemple 1.1.3 à nouveau).

Définition 4.5.4. Étant donnés un polynôme de Ore $P \in F[Y, \mathrm{Frob}]$, une dérivation $\partial : K \to K$ et une fonction $f \in K$, on définit :

$$\nu_\partial(P, f) = \sum_{i=0}^{n} \sum_{j=0}^{i} c_i \cdot (\partial^{p^j - 1}(f))^{p^{i-j}}$$

si P s'écrit $P = c_0 + c_1 X + \cdots + c_n X^n$.

La construction ci-dessus est F-linéaire vis-à-vis de P (pour la structure de F-espace vectoriel à gauche sur $F[Y, \mathrm{Frob}]$) et additive vis-à-vis de f. De plus, elle vérifie la formule de composition suivante :

$$\forall P, Q \in F[Y, \mathrm{Frob}],\ \forall \partial \in \mathrm{Der}(K),\ \forall f \in K, \quad \nu_\partial(PQ, f) = \nu_{Q\cdot\partial}(P, \nu_d(Q, f)) \tag{34}$$

où on rappelle que la notation $Q \cdot \partial$ fait référence à la dérivation $Q_{\mathrm{lin}}(\partial)$ et où, bien entendu, le produit PQ est calculé dans $F[Y, \mathrm{Frob}]$. Pour démontrer cette formule, on se ramène, grâce à l'additivité, au cas où P et Q sont des monômes où elle résulte alors d'un calcul facile à mener. Avec ces définitions, la généralisation annoncée de la proposition 4.5.3 se formule comme suit.

Proposition 4.5.5. *Soit $\partial : K \to K$ une dérivation. Pour $P \in F[Y, Frob]$ et $f \in K$, on a l'identité :*

$$P_{lin}(\partial + f\ \mathrm{id}_K) = P\cdot\partial + \nu_\partial(P, f)\ \mathrm{id}_K.$$

Démonstration. Par F-linéairité vis-à-vis de P, il suffit de démontrer la proposition lorsque P est de la forme Y^i pour $i \in \mathbb{N}$. On procède par récurrence sur i. Lorsque $i = 1$, l'énoncé que l'on souhaite démontrer est exactement celui de la proposition 4.5.3, il n'y a donc rien de plus à faire. Pour passer de i à $i + 1$, on écrit :

$$\begin{aligned}
(\partial + f\ \mathrm{id}_K)^{p^{i+1}} &= \left((\partial + f\ \mathrm{id}_K)^{p^i}\right)^p \\
&= \left(\partial^{p^i} + \nu_\partial(Y^i, f)\ \mathrm{id}_K\right)^p && \text{par hypothèse de recurrence} \\
&= \partial^{p^{i+1}} + \left(\nu_\partial(Y^i, f)^p + \partial^{p^i(p-1)}\nu_\partial(Y^i, f)\right)\mathrm{id}_K && \text{par la proposition 4.5.3} \\
&= \partial^{p^{i+1}} + \nu_{\partial^{p^i}}(Y, \nu_\partial(Y^i, f))\ \mathrm{id}_K \\
&= \partial^{p^{i+1}} + \nu_\partial(Y^{i+1}, f)\ \mathrm{id}_K && \text{par la formule (34)}
\end{aligned}$$

ce qui conclut. □

Nous sommes enfin prêts à mener à bien le calcul de la norme réduite de $X - f \in K[X, \partial]$ dans le cas général. Rappelons rapidement les notations : $\mathrm{Ann}(\partial) \in F[Y, \mathrm{Frob}]$ désigne le polynôme unitaire de plus petit degré qui annule ∂ (que l'on suppose exister afin que $K[X, \partial]$ soit une algèbre d'Azumaya) et $Z(X) = \mathrm{Ann}(\partial)_{\mathrm{lin}}(X)$. Le centre de $\mathcal{A} = K[X, \partial]$ est $\mathcal{Z} = F[Z(X)]$ où F est le sous-corps des constantes de K.

Définition 4.5.6. On définit la fonction $\nu_{(K,\partial)/F}$ par la formule

$$\nu_{(K,\partial)/F}(f) = \nu_\partial(\mathrm{Ann}(\partial), f)$$

pour tout $f \in K$ (où on rappelle la fonction ν_∂ a été introduite dans la définition 4.5.4).

Dans la suite, on notera généralement $\nu_{K/F}$ à la place de $\nu_{(K,\partial)/F}$, sachant que la dérivation ∂ sur K est fixée une fois pour toutes et que cet abus ne prêtera donc pas à confusion.

Proposition 4.5.7. *Dans le cas où $\mathcal{A} = K[X, \partial]$, on a :*

$$N_{rd}(X - f) = Z(X) - \nu_{K/F}(f).$$

Démonstration. D'après la proposition 4.3.11, la norme réduite de $X-f$ s'écrit comme $\chi(Z(X))$ où $\chi(T)$ est le polynôme caractéristique de l'application de multiplication par $Z(X)$ agissant sur le module quotient $M = K[X,\partial]/K[X,\partial]\,(X-f)$. Or, on a vu dans la démonstration de la proposition 4.5.3 que la multiplication à gauche par X agit sur M *via* la fonction $(\partial + f\,\mathrm{id}_K)$. On en déduit que la multiplication à gauche par $Z(X)$ sur M est donnée par l'endomorphisme de K :

$$Z(\partial + f\,\mathrm{id}_K) = \mathrm{Ann}(\partial)\cdot\partial + \nu_\partial(\mathrm{Ann}(\partial), f)\,\mathrm{id}_K = \nu_{K/F}(f)\,\mathrm{id}_K.$$

Ci-dessus, la première égalité est conséquence de la proposition 4.5.5 tandis que la seconde est vraie car $\mathrm{Ann}(\partial)\cdot\partial$ s'annule par construction. Le polynôme caractéristique de $Z(X)$ agissant sur M est ainsi $\chi(T) = T - \nu_{K/F}(f)$, ce qui conclut la démonstration. □

Remarque 4.5.8. Il résulte de la proposition 4.5.7 que la fonction $\nu_{K/F}$ prend ses valeurs dans le sous-corps des constantes F. Cette propriété peut se démontrant directement à l'aide d'un calcul en vérifiant que $\partial(\nu_{K/F}(f))$ s'annule pour tout $f \in K$.

La proposition 4.5.7 peut être mise à profit pour déduire des propriétés supplémentaires de la fonction $\nu_{K/F}$. C'est le cas, par exemple, de la proposition 4.5.9 ci-après qui peut être vue comme un analogue différentiel du théorème de Hilbert 90.

Proposition 4.5.9. *Pour $f \in K$, on a l'équivalence suivante : $\nu_{K/F}(f) = 0$ si et seulement si f est de la forme $f = \frac{\partial(g)}{g}$ pour $g \in K$, $g \neq 0$.*

Démonstration. Soit $g \in K$, $g \neq 0$. On pose $f = \frac{\partial(g)}{g}$. Par le lemme 1.3.5, on sait que le reste de la division euclidienne de $Z(X)$ par $X-f$ est $g^{-1}\cdot Z(\partial)(g)$ et, donc, s'annule. On en déduit que $X-f$ est un diviseur à droite de $Z(X)$. Il en résulte que $N_{\mathrm{rd}}(X-f)$ n'est pas premier avec $Z(X)$ puis, comme ces deux polynômes centraux sont de degré 1 dans $\mathcal{Z}$, on en déduit que $N_{\mathrm{rd}}(X-f) = Z(X)$. Il découle à présent de la proposition 4.5.7 que $\nu_{K/F}(f) = 0$.

Réciproquement, considérons un élément $f \in K$ tel que $\nu_{K/F}(f) = 0$. De la proposition 4.5.7, on déduit que $N_{\mathrm{rd}}(X-f) = Z(X)$ et, par suite, que $X-f$ divise $Z(X)$. Définissons $\mathcal{A} = K[X,\partial]$ et $E = \mathrm{End}_F(K)$. Le morphisme d'évaluation $P(X) \mapsto P(\partial)$ induit un isomorphisme $\mathcal{A}/Z(X)\mathcal{A} \simeq E$. De plus, le quotient $\mathcal{A}/\mathcal{A}(X-f)$ est un E-module qui est de dimension p^m sur F. On déduit ainsi de la proposition 3.2.8 que $\mathcal{A}/\mathcal{A}(X-f) \simeq K$ comme E-module (où E agit sur K de manière naturelle). Considérons le morphisme composé :

$$\varepsilon : E \xrightarrow{\sim} \mathcal{A}/Z(X)\mathcal{A} \longrightarrow \mathcal{A}/\mathcal{A}(X-f) \xrightarrow{\sim} K.$$

De la E-linéarité, on déduit que ε est nécessairement de la forme $\varphi \mapsto \varphi(g)$ pour un certain $g \in K$. Clairement $g \neq 0$ puisque ε est surjectif. À partir de la description explciite de ε que nous venons d'établir, on obtient $\varepsilon(\partial - h\,\mathrm{id}_K) = 0$ pour $h = \frac{\partial(g)}{g}$. En remontant les isomorphismes, on en déduit que le polynôme de Ore $X-h$ s'annule dans le quotient $\mathcal{A}/\mathcal{A}(X-f)$. Autrement dit $X-f$ est un diviseur à droite de $X-h$. Par égalité des degrés, ceci implique $X-f = X-h$ et donc finalement $f = h = \frac{\partial(g)}{g}$. □

On déduit de la proposition 4.5.9 que les diviseurs irréductibles unitaires de $Z(X)$ sont exactement les polynômes de Ore de la forme $X - \frac{\partial(g)}{g}$ pour $g \in K$, $g \neq 0$. De plus, pour deux éléments non nuls $g, h \in K$, la condition $\frac{\partial(g)}{g} = \frac{\partial(h)}{h}$ est équivalente à $\partial(\frac{g}{h}) = 0$, soit encore à $\frac{g}{h} \in F$. On en déduit que les diviseurs irréductibles unitaires de $Z(X)$ sont canoniquement paramétrés par l'espace projectif $\mathbb{P}(K)$ où, ici, K comme un espace vectoriel sur F. Puisque K est de dimension p^m sur F, ce résultat confirme et précise l'assertion (iii) du théorème 4.4.3.

4.6 Une étude de l'exposant $e(N)$

Rappelons qu'étant donné un polynôme central N qui est irréductible dans $\mathcal{Z}$, on a défini au §4.4.2 l'entier $e(N) = \sqrt{\dim_{\mathcal{Z}/N\mathcal{Z}} D_N}$ où D_N est l'algèbre à divisions (déterminée à isomorphisme près) pour laquelle $\mathcal{A}/N\mathcal{A}$ est isomorphe à une algèbre de matrices sur D_N. Cet entier $e(N)$ joue un rôle important pour ce qui concerne la factorisation. En effet, le théorème 4.4.3 nous apprend qu'il détermine les degrés des diviseurs irréductibles de N. Dans ce numéro, nous donnons quelques propriétés supplémentaire de cet invariant $e(N)$ permettant de mieux le cerner.

4.6.1 Le cas d'un polynôme central de degré 1

Afin d'uniformiser les notations, posons $\nu_{K/F} = N_{K/F}$ et $Z(X) = X^r$ dans le cas d'un automorphisme θ d'ordre r. Avec ces conventions, la formule

$$N_{\text{rd}}(X - a) = Z(X) - \nu_{K/F}(a) \tag{35}$$

est vraie de manière uniforme pour tout $a \in K$.

Proposition 4.6.1. *Pour $c \in F$ et $N = Z(X) - c \in \mathcal{Z}$, on a l'équivalence suivante : $e(N) = 1$ si et seulement si c est dans l'image de l'application $\nu_{K/F} : K \to F$.*

Démonstration. Supposons que $e(N) = 1$. Soit P un diviseur irréductible unitaire de N. D'après le théorème 4.4.3, P est de degré 1 et sa norme réduite est proportionnelle à N. En écrivant $P = X - a$ (avec $a \in K$), on aboutit en comparant avec la formule (35) à $c = \nu_{K/F}(a)$. Ainsi c est dans l'image de $\nu_{K/F}$.

Réciproquement, supposons que c s'écrive $c = \nu_{K/F}(a)$ avec $a \in K$. Alors $N_{\text{rd}}(X - a) = \pm N$, ce qui démontre que $X - a$ est un diviseur de N. Ainsi N admet un diviseur de degré 1 et le théorème 4.4.3 assure que $e(N) = 1$. □

Exemple 4.6.2. Lorsque $\mathcal{A} = \mathbb{C}[X, \text{conj}]$, la proposition 4.6.1 nous apprend que $e(X^2 - c) = 1$ si $c > 0$ alors que $e(X^2 - c) > 1$ si $c < 0$. Dans le dernier cas, on a nécessairement $e(X^2 - c) = 2$ puisqu'il est toujours vrai que $e(N) \leq r$ et que $r = 2$ dans notre exemple. Ces résultats confirment (par une méthode légèrement plus rapide) ce que nous avions déjà trouvé dans l'exemple 4.4.4.

Il est possible de mettre en évidence de nouvelles contraintes en faisant intervenir les extensions intermédiaires de K/F. Dans le cas d'un endomorphisme, la théorie de Galois nous dit que toute extension L comprise entre F et K est de la forme

$$L = \left\{ x \in K \quad \text{t.q.} \quad \theta^s(x) = x \right\} \tag{36}$$

pour un certain entier s. De plus, si s est choisi minimal, il s'agit d'un diviseur de r et on a $[K{:}L] = \frac{r}{s}$ ainsi que $[L{:}F] = s$. Dans cette situation, on pose $\nu_{K/L} = N_{K/L}$ et $\nu_{L/F} = N_{L/F}$. Très concrètement, les fonctions $\nu_{K/L}$ et $\nu_{L/F}$ sont données par les formules suivantes :

$$\forall a \in K, \quad \nu_{K/L}(a) = a \cdot \theta^s(a) \cdot \theta^{2s}(a) \cdots \theta^{(r-1)s}(a)$$
$$\forall a \in L, \quad \nu_{L/F}(a) = a \cdot \theta(a) \cdot \theta^2(a) \cdots \theta^{s-1}(a)$$

et la formule de composition des normes s'écrit $\nu_{L/F} \circ \nu_{K/L} = \nu_{K/F}$.

Dans le cas différentiel, nous nous restreignons aux extensions intermédiaires L qui sont stables par la dérivation ∂. D'après la correspondance de Galois explicitée dans l'appendice A.3 (*cf* théorème A.3.4), l'extension L est de la forme $\ker(D{\cdot}\partial)$ pour un certain diviseur unitaire D de $\text{Ann}(\partial)$, uniquement déterminé. On écrit $\text{Ann}(\partial) = D'D$ avec $D' \in \mathbb{F}[Y, \text{Frob}]$. Les degrés

des extensions K/L et L/F sont alors respectivement $p^{\deg D'}$ et $p^{\deg D}$. Dans cette situation, les fonctions $\nu_{K/L} = \nu_{(K,D\cdot\partial)/L} : K \to L$ et $\nu_{L/F} = \nu_{(L,\partial)/F} : L \to F$ sont définies par :

$$\forall f \in K, \quad \nu_{K/L}(f) = \nu_{D\cdot\partial}(D', f)$$
$$\forall f \in L, \quad \nu_{L/F}(f) = \nu_{\partial}(D, f).$$

La formule de composition (34) implique que la relation $\nu_{L/F} \circ \nu_{K/L} = \nu_{K/F}$ vaut aussi dans le cadre différentiel.

Proposition 4.6.3. *Soient $c \in F$ et $N = Z(X) - c \in \mathcal{Z}$.*
Soit L une extension intermédiaire entre F et K que l'on suppose stable par ∂ dans le cadre différentiel. Si c est dans l'image de $\nu_{K/L}$, alors $e(N)$ divise $[L{:}F]$.

Démonstration. On introduit une sous-algèbre $\mathcal{A}_L$ de $\mathcal{A}$. Dans le cas où $\mathcal{A} = K[X, \theta]$, on définit s comme l'unique diviseur de r pour lequel l'égalité (36) est vraie et on pose $X_L = X^s$ et $\mathcal{A}_L = K[X_L, \theta^s]$. Le fait que $X_L \cdot a = \theta^s(s)X_L$ ($a \in K$) assure que $\mathcal{A}_L$ apparaît bien comme une sous-algèbre de $\mathcal{A}$. Étant donné que l'ordre de θ^s est $\frac{r}{s}$, on déduit de la proposition 4.1.1 que le centre de $\mathcal{A}_L$ est $\mathcal{Z}_L = L[X_L^{r/s}] = L[X^r]$.

Dans le cas où $\mathcal{A} = K[X, \partial]$, une construction analogue est possible. On appelle D l'unique diviseur unitaire de $\mathrm{Ann}(\partial)$ pour lequel $L = \ker(D\cdot\partial)$. On pose $X_L = D_{\mathrm{lin}}(X)$ et $\mathcal{A}_L = K[X_L, D\cdot\partial]$. De même que précédemment, la relation $X_L \cdot f = f \cdot X_L + (D\cdot\partial)(f)$ ($f \in K$) suffit à assurer que $\mathcal{A}_L$ est naturellement une sous-algèbre de $\mathcal{A}$. Par ailleurs, il résulte de la démonstration du théorème A.3.4 que $\mathrm{Ann}(D\cdot\partial)$ est égal au polynôme de Ore $D' \in F[Y, \mathrm{Frob}]$ défini par $D'D = \mathrm{Ann}(\partial)$. La proposition 4.1.3 implique donc que le centre de $\mathcal{A}_L$ est $\mathcal{Z}_L = L[D'_{\mathrm{lin}}(X_L)] = L[Z(X)]$.

En résumé, dans tous les cas, on a construit une sous-algèbre $\mathcal{A}_L$ de $\mathcal{A}$ de centre $\mathcal{Z}_L = L[Z(X)]$, qui est en outre un anneau de Ore en la variable X_L. Remarquons également que, dans tous les cas, $\deg_{\mathcal{A}} X_L = [L{:}F]$. Soit $a \in K$ un antécédent de c par l'application $\nu_{K/L}$. On forme le polynôme de Ore $P = X_L - a$ que l'on voit comme élément de $\mathcal{A}_L$. Sa norme réduite *pour l'anneau de Ore* $\mathcal{A}_L$ est égale à $\pm N$ d'après la formule (35). On en déduit que P est un diviseur de N dans $\mathcal{A}_L$. Il en est donc *a fortiori* également un dans $\mathcal{A}$. Ainsi, N admet un diviseur de degré $[L{:}F]$ dans $\mathcal{A}$, ce qui, d'après le théorème 4.4.3, implique que $e(N)$ divise $[L{:}F]$. □

Nous attirons l'attention de la lectrice et du lecteur sur le fait que, contrairement à ce qui se passait pour $L = F$, la réciproque de la proposition 4.6.3 est fausse, dans le sens où la divisibilité $e(N) \mid [L{:}F]$ n'implique généralement *pas* que c est dans l'image de $\nu_{K/L}$. Cette implication ne vaut ni dans le cas d'une dérivation, ni dans celui d'un endomorphisme qui, pourtant, paraît *a priori* plus favorable étant donné qu'il n'existe qu'une seule extension intermédiaire de degré fixé. Nous donnons ci-dessous un contre-exemple.

Exemple 4.6.4. Considérons le corps cyclotomique $K = \mathbb{Q}(\zeta_5)$ où ζ_5 désigne une racine primitive 5-ième de l'unité. Il est bien connu que l'extension $K/\mathbb{Q}$ est galoisienne et que son groupe de Galois $\mathrm{Gal}(K/\mathbb{Q})$ s'identifie canoniquement au groupe multiplicatif $(\mathbb{Z}/5\mathbb{Z})^\times$: à un élement $n \in (\mathbb{Z}/5\mathbb{Z})^\times$, il correspond l'automorphisme de K qui envoie ζ_5 sur ζ_5^n. De plus $(\mathbb{Z}/5\mathbb{Z})^\times$ est cyclique (il est engendré par exemple par la classe de 2), il en est donc de même de $\mathrm{Gal}(K/\mathbb{Q})$. Fixons un générateur θ de $\mathrm{Gal}(K/\mathbb{Q})$. L'automorphisme θ^2 engendre l'unique sous-groupe d'ordre 2 de $\mathrm{Gal}(K/\mathbb{Q})$. Soit L le sous-corps fixé par θ^2, il est engendré par l'élément $\eta = \zeta_5 + \zeta_5^{-1}$. Un calcul simple donne :

$$\eta^2 + \eta = \zeta_5^2 + \zeta_5^{-2} + 2 + \zeta_5 + \zeta_5^{-1} = 1.$$

Ainsi, en résolvant l'équation, on trouve $\eta = \frac{-1+\sqrt{5}}{2}$. Le sous-corps L s'identifie ainsi à $\mathbb{Q}(\sqrt{5})$.

Travaillons à présent dans l'anneau de Ore $\mathcal{A} = K[X, \theta]$. Clairement, l'automorphisme θ est d'ordre 4 et le sous-corps de ses points fixes est $\mathbb{Q}$. Le centre de $\mathcal{A}$ est ainsi $\mathcal{Z} = \mathbb{Q}[X^4]$.

Considérons le polynôme central $N = X^4 - c$ pour $c = -4$. Celui-ci admet des factorisations non triviales dans $\mathcal{A}$, par exemple :

$$N = X^4 + 4 = \left(X^2 + 2X + 2\right) \cdot \left(X^2 - 2X + 2\right). \tag{37}$$

Le théorème 4.4.3 permet d'en déduire que $e(N) \in \{1, 2\}$. Montrons cependant que c n'est pas une norme dans l'extension K/L. Pour cela, considérons un élément générique de K écrit sous la forme $x + y\zeta_5$, avec $x, y \in L$. Sa norme vaut :

$$N_{K/L}(x + y\zeta_5) = (x + y\zeta_5) \cdot \theta^2(x + y\zeta_5) = (x + y\zeta_5)(x + y\zeta_5^{-1}) = (x^2 + y^2) + xy\eta$$

et ne peut donc pas être égale à $c = -4$ puisque $x^2 + y^2$ est toujours un réel positif ou nul étant donné que $L = \mathbb{Q}(\sqrt{5}) \subset \mathbb{R}$. Ceci fournit un contre-exemple à une éventuelle réciproque de la proposition 4.6.3.

Pour conclure, remarquons qu'étant donné que c n'est pas une norme dans l'extension K/L, il n'en est pas une non plus dans l'extension $K/\mathbb{Q}$. En effet, de $c = N_{K/\mathbb{Q}}(a)$ on déduirait $c = N_{K/L}(a\theta(a))$. Il résulte ainsi de la proposition 4.6.1 que $e(N) > 1$. *In fine*, on trouve $e(N) = 2$ d'où on déduit, par le théorème 4.4.3, que la factorisation (37) est une factorisation de N en produit de polynômes de Ore irréductibles dans $\mathcal{A}$.

Exemple 4.6.5. Considérons le corps $K = \mathbb{F}_p(t)$ muni de sa dérivation canonique $\partial = \frac{d}{dt}$ et formons l'anneau de Ore $\mathcal{A} = K[X, \partial] = \mathbb{F}_p(t)[X, \frac{d}{dt}]$. Le sous-corps des constantes de ∂ est $F = \mathbb{F}_p(t^p)$. De plus $\partial^p = 0$. On déduit de cela que le centre de $\mathcal{A}$ est $\mathcal{Z} = F[X^p]$. Revenant aux définitions, on trouve la formule explicite suivante $\nu_{K/F}(f) = f^p + \frac{d^{p-1}f}{dt^{p-1}}$ pour toute fonction $f \in K$. Pour un élément $c \in F$ fixé, considérons le polynôme central $N = X^p - c \in \mathcal{Z}$. D'après la proposition 4.6.1, l'invariant $e(N)$ vaut 1 si et seulement si l'équation différentielle

$$\frac{d^{p-1}y(t)}{dt^{p-1}} + y(t)^p = c \tag{38}$$

d'inconnue $y(t)$ a une solution. De plus, le cas échéant, toute solution $y(t)$ de (38) fournit un facteur irréductible de N, à savoir $X - y(t)$. À noter que, connaissant ces facteurs, il est possible de construire une factorisation complète de N en prenant de LLCM. Si, au contraire, l'équation différentielle (38) n'a pas de solutions, la proposition 4.6.1 nous enseigne que $e(N) > 1$. Comme, de plus, $e(N)$ doit être un diviseur de p, on a nécessairement $e(N) = p$ dans ce cas. On déduit ainsi du théorème (4.4.3) que N est irréductible lorsque (38) n'a pas de solutions.

Concluons cet exemple en expliquant brièvement comment l'équation différentielle (38) peut être résolue. Premièrement, on observe que cette équation différentielle est $\mathbb{F}_p$-linéaire puisqu'il en est ainsi de l'élévation à la puissance p en caractéristique p. Ainsi cela a un sens de se concentrer tout d'abord sur l'équation différentielle homogène :

$$\frac{d^{p-1}y(t)}{dt^{p-1}} + y(t)^p = 0. \tag{39}$$

Par la proposition 4.5.9, on sait que les solutions de cette dernière équation sont exactement les fonctions $y(t)$ de la forme $y(t) = \frac{f'(t)}{f(t)}$ pour $f(t) \in \mathbb{F}_p(t)$. Dans le cas particulier considéré ici, ce résultat peut se redémontrer directement. En effet, un calcul montre que, pour tout $\alpha \in \bar{\mathbb{F}}_p$, la fonction $y(t) = \frac{1}{t-\alpha}$ est solution de (39). Par additivité, on en déduit que toute fonction $y(t)$ de la forme $y(t) = \frac{f'(t)}{f(t)}$ avec $f(t) \in \mathbb{F}_p(t)$ est également solution de (39). Réciproquement, considérons maintenant une solution $y(t)$ de (39). Écrivons le décomposition en éléments simples de $y(t)$ dans $\bar{\mathbb{F}}_p(t)$:

$$y(t) = P(t) + \frac{P_1(t)}{(t-\alpha_1)^{n_1}} + \frac{P_2(t)}{(t-\alpha_2)^{n_2}} + \cdots + \frac{P_s(t)}{(t-\alpha_1)^{n_s}}$$

où les α_i sont des éléments deux à deux distincts de $\bar{\mathbb{F}}_p$, les n_i sont des entiers strictement positifs et $P(t)$ et les $P_i(t)$ sont des polynômes à coefficients dans $\bar{\mathbb{F}}_p$ avec $\deg P_i(t) < n_i$ pour tout i. Quitte à retirer une quantité de la forme $\frac{f'(t)}{f(t)}$, on peut supposer en outre que $P_i(t) \notin \mathbb{F}_p$ dès que $n_i = 1$. Avec cette hypothèse, un calcul montre que les dénominateurs $(t-\alpha_1)^{pn_i}$ subsistent tous dans $\frac{d^{p-1}y(t)}{dt^{p-1}} + y(t)^p$. On en déduit que $y(t) = P(t)$ puis, avec des considérations de degré, que $y(t) = 0$.

Pour ce qui concerne l'équation inhomogène (38), la méthode suivie précédemment (qui consiste à décomposer une solution éventuelle en éléments simples) permet de construire une solution particulière, s'il en existe, ou au contraire de démontrer qu'il n'y en a pas. Dans le cas où l'équation (38), la linéarité montre que celles-ci sont également paramétrées par l'espace projectif $\mathbb{P}^{p-1}(\mathbb{F}_p(t^p))$. Cela confirme, dans le cas de l'anneau de Ore $\mathcal{A} = \mathbb{F}_p(t)[X, \frac{d}{dt}]$, la troisième assertion du théorème 4.4.3.

4.6.2 Le cas général

On considère à présent un polynôme central N de degré quelconque que l'on écrit sous la forme $N = N_0(Z(X))$ où N_0 est un polynôme à coefficients dans F. On suppose que N_0 est irréductible. Soit $F' = \mathcal{Z}/N\mathcal{Z} = F[T]/N_0(T)$; c'est une extension finie de F qui se plonge naturellement dans $\mathcal{A}/N\mathcal{A}$. Désignons par c' la classe de $Z(X) \in \mathcal{Z}$ dans F' (ou, ce qui revient au même, la classe de $T \in F[T]$ dans F'). Posons encore :

$$\mathcal{A}' = F' \otimes_F \mathcal{A} = (F' \otimes_F K)[X, \mathrm{id} \otimes \bullet]$$

où $\bullet = \theta$ dans le cas d'un endomorphisme et $\bullet = \partial$ dans le cas d'une dérivation. En particulier $\mathcal{A}'$ est encore un anneau de Ore. D'autre part, son centre $\mathcal{Z}'$ s'identifie canoniquement à $F' \otimes_F \mathcal{Z}$. On dispose d'un morphisme d'anneaux surjectif $\mathcal{A}' \to \mathcal{A}/N\mathcal{A}$ dont le noyau contient manifestement l'élément $N' = Z(X) - c'$. Il induit ainsi un morphisme surjectif :

$$\psi : \mathcal{A}'/N'\mathcal{A}' \longrightarrow \mathcal{A}/N\mathcal{A}. \tag{40}$$

Ce dernier est, en fait, un isomorphisme car les ensembles de départ et d'arrivée sont tous les deux des espaces vectoriels de dimension $\deg N$ sur K. De cette manière, on ramène le problème du calcul de $e(N)$ à celui du calcul de $e(N')$. L'avantage à procéder ainsi est que, désormais, le polynôme N', vu comme élément de $\mathcal{Z}'$, est de degré 1, ce qui nous place en situation pour appliquer les résultats du §4.6.1. Il y a toutefois un petit écueil car $\mathcal{A}'$ est un anneau de polynômes de Ore sur l'anneau $F' \otimes_F K$ qui n'est pas un corps — mais un produit de corps — lorsque les extensions F'/F et K/F ne sont pas linéairement disjointes. Au prix d'un effort minimal, il est cependant possible d'étendre les résultats du §4.6.1 pour couvrir ce cas. Néanmoins, dans de nombreuses situations, il est possible d'éviter ce désagrément et de se ramener, *via* une construction un peu plus complexe, à un anneau de polynômes de Ore définis sur un corps. C'est ce que nous nous proposons d'expliquer ci-dessous (en traitant à part le cas d'un endomorphisme et celui d'une dérivation).

Le cas d'un endomorphisme. Plaçons-nous pour commencer dans le cas d'un endomorphisme, *i.e.* $\mathcal{A} = K[X, \theta]$. Comme nous l'avons mentionné précédemment, le produit tensoriel $F' \otimes_F K$ n'est généralement pas un corps, mais un produit de corps. Précisément, soit C le plus grand sous-corps de K qui se plonge de F' et soit $d = [C{:}F]$. La correspondance de Galois nous apprend que C est le sous-corps des points fixes de θ^d. Soit $K' = F' \otimes_C K$. Le fait que les extensions F'/C et K/C soient linéairement disjointes assure que K' est un corps. Dans le suite, nous identifierons sans commentaire F' et K à des sous-corps de K'. Le produit tensoriel $F' \otimes_F K$ se décompose alors comme suit :

$$\begin{array}{ccc} F' \otimes_F K & \xrightarrow{\sim} & (K')^d \\ a \otimes b & \mapsto & \big(ab,\, a\theta(b),\, a\theta^2(b), \ldots,\, a\theta^{d-1}(b)\big). \end{array} \tag{41}$$

De plus id $\otimes\, \theta^d$ définit un automorphisme de K' (que l'on notera simplement θ^d dans la suite) et l'action de θ sur $F' \otimes_F K$ se transporte sur $(K')^d$ comme suit :

$$\theta(x_0, x_1, \ldots, x_{d-1}) = (x_1, x_2, \ldots, x_{d-1}, \theta^d(x_0))$$

pour $x_0, x_1, \ldots, x_{d-1} \in K'$.

Proposition 4.6.6. *On conserve les notations précédentes. Alors :*

(i) $e(N) = 1$ *si et seulement si* c' *est une norme dans l'extension* K'/F',

(ii) si L' *est une extension intermédiaire entre* F' *et* K' *pour laquelle* c' *est une norme dans l'extension* K'/L', *alors* $e(N)$ *divise* $[L':F']$.

Démonstration. L'isomorphisme (41) induit un isomorphisme entre anneaux de Ore

$$\mathcal{A}' = (K' \otimes_F K)[X, \theta] \simeq (K')^d[X, \theta]. \tag{42}$$

De plus, on peut construire explicitement un isomorphisme

$$(K')^d[X, \theta] \simeq M_d(K'[X^d, \theta^d]) \tag{43}$$

en envoyant un d-uplet $(x_0, \ldots, x_{d-1}) \in (K')^d$ sur la matrice diagonale $\mathrm{Diag}(x_0, \ldots, x_{d-1})$ et en envoyant la variable X sur la matrice :

$$M_X = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ X^d & & & \end{pmatrix}.$$

Le fait que ces formules définissent bien un morphisme d'anneaux provient de l'égalité matricielle $M_X \cdot \mathrm{Diag}(x_0, \ldots, x_{d-1}) = \mathrm{Diag}(x_1, \ldots, x_{d-1}, \theta^d(x_0)) \cdot M_X$, qui est valable dans $M_d(K'[X^d, \theta^d])$. En combinant les isomorphismes (42) et (43), on obtient une identification entre $\mathcal{A}'/N'\mathcal{A}'$ et l'algèbre de matrices $M_d(K'[X^d, \theta^d]/(X^r - c'))$. À partir de là, la proposition qu'on souhaite démontrer est une conséquence directe des propositions 4.6.1 et 4.6.3. □

Corollaire 4.6.7. *Avec les notations précédentes,* $e(N)$ *divise* $\frac{r}{d}$.
En particulier, $e(N) = 1$ *dès lors que* K *se plonge dans* F'.

Démonstration. Il suffit d'appliquer le (ii) de la proposition 4.6.6 avec $L' = K'$. □

Le cas d'une dérivation. Venons-en maintenant au cas d'une dérivation. On rappelle que l'extension K/F est purement inséparable. Ainsi, si l'extension F'/F est, au contraire, séparable, les extensions K/F et F'/F sont linéairement disjointes, de sorte que le produit tensoriel $F' \otimes_F K$ est un corps. L'isomorphisme (40) permet ainsi de ramener le calcul de $e(N)$ à celui de $e(N')$ où N' est un polynôme central de degré 1 dans un anneau de Ore défini sur un corps et ainsi d'utiliser les résultats du §4.6.1.

Toutefois, dans le cas général, l'extension F'/F peut ne pas être séparable. Dans ce cas, on fait intervenir le plus grand sous-corps F'_{sep} de F' qui est séparable sur F. Concrètement, si on écrit $N(X)$ sous la forme $N(X) = N_{\mathrm{sep}}(Z(X)^{p^n})$ où $N_{\mathrm{sep}} \in F[T]$ est un polynôme séparable et n est un entier, on a $F'_{\mathrm{sep}} = F'[U]/N_{\mathrm{sep}}(U)$ et $F' = F'_{\mathrm{sep}}[V]/(V^{p^n} - U)$. De plus, $F'_{\mathrm{sep}} \otimes_F K$ est un corps. Soit $\mathcal{A}'_{\mathrm{sep}} = F'_{\mathrm{sep}} \otimes_F K = (F'_{\mathrm{sep}} \otimes_F K)[X, \partial]$. Il s'agit d'un anneau de polynôme de Ore sur un corps dont le centre $\mathcal{Z}'_{\mathrm{sep}}$ s'identifie à $F'_{\mathrm{sep}}[Z(X)]$. On dispose de plus d'un isomorphisme :

$$\mathcal{A}'_{\mathrm{sep}}/\big(Z(X)^{p^n} - [U]\big)\mathcal{A}'_{\mathrm{sep}} \simeq \mathcal{A}/N\mathcal{A}$$

où $[U]$ désigne la classe de la variable U dans F'_{sep}.

Quitte à remplacer $\mathcal{A}$ par $\mathcal{A}'_{\text{sep}}$, l'argumentation ci-dessus permet de se ramener au cas où le polynôme N est de la forme $N(X) = Z(X)^{p^n} - a$ pour un certain entier $n \geq 1$ et un certain élément $a \in F$. C'est ce que nous supposerons à partir de maintenant. Le corps F' est alors obtenu à partir de F en ajoutant une racine p^n-ième de a. Comme F contient les puissances p-ièmes des éléments de K, on est dans l'alternative suivante :

- soit a n'admet pas de racine p-ième dans K et alors les extensions F'/F et K/F sont linéairement disjointes,
- soit a admet une racine p-ième b dans K et alors le corps $C = F[b]$ se plonge à la fois dans F' et K et les extensions F'/C et K/C sont linéairement disjointes.

Dans le premier cas, le produit tensoriel $F' \otimes_F K$ est un corps, et on peut appliquer directement les résultats du §4.6.1. Concentrons-nous donc à présent sur le second cas. Nous faisons l'hypothèse supplémentaire suivante [5] :

Hypothèse 4.6.8. Le sous-corps C de K est stable par la dérivation ∂.

Sous l'hypothèse précédente, il suit de la correspondance de Galois dans le cadre différentiel (*cf* théorème A.3.4 de l'appendice A.3) que C est le sous-corps des constantes de K pour la dérivation $\delta = \partial^p + c\partial$ pour un certain $c \in F$. Soit $K' = F' \otimes_C K$. Puisque δ agit trivialement sur C, elle induit une dérivation $\mathrm{id} \otimes \delta$ sur K' que, dans un léger abus, nous continuerons à noter δ dans la suite. On a les descriptions explicites suivantes :

$$K' = K[U]/(U^{p^{n-1}} - b) \quad \text{et} \quad F' \otimes_F K = K[U]/(U^{p^n} - a) = K[U]/(U^{p^{n-1}} - b)^p$$

et la dérivation ∂ (resp. δ) agit sur $F' \otimes_F K$ (resp. sur K') coordonnée par coordonnée. Soit $u \in K'$ l'image de la variable U. Pour $f \in K[U]/(U^{p^n} - a)$, on note $f(u)$ son image dans K'. Le « développement de Taylor selon ∂ [6] » au voisinage de u permet de définir un isomorphisme d'anneaux comme suit :

$$\begin{array}{ccc} K[U]/(U^{p^n} - a) & \xrightarrow{\sim} & K'[T]/T^p \\ f & \mapsto & f(u) + \partial(f(u))T + \partial^2(f(u))\frac{T^2}{2} + \cdots + \partial^{p-1}(f(u))\frac{T^{p-1}}{(p-1)!}. \end{array}$$

Ceci peut se voir comme une application de la proposition A.3.5 de l'appendice A.3. En outre la dérivation $\mathrm{id} \otimes \partial$ induit sur $K'[T]/T^p$ une dérivation qui s'écrit explicitement de la manière suivante :

$$a_0 + a_1T + \cdots + a_{p-1}T^{p-1} \mapsto a_1 + 2a_2T + \cdots + (p-1)\,a_{p-1}T^{p-2} + \big(\delta(a_0) - ca_1\big)T^{p-1}.$$

Au vu de ces descriptions, on démontre par le calcul qu'on a un isomorphisme explicite

$$(K'[T]/T^p)[X, \partial] \xrightarrow{\sim} M_p\big(K'[Y, \delta]\big) \tag{44}$$

obtenue en envoyant les éléments de K' sur les matrices scalaires correspondantes et T et X respectivement sur les matrices :

$$M_T = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \quad \text{et} \quad M_X = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & 0 & p-1 \\ Y & -c & 0 & \cdots & 0 \end{pmatrix}.$$

De même que dans le cas d'un endomorphisme, l'isomorphisme (44) permet de se ramener à un anneau de polynômes de Ore définis sur un corps et d'utiliser ainsi les résultats du §4.6.1.

5. Cette hypothèse n'est pas du tout automatique, ni même générique. Toutefois, sans trop s'éloigner du cadre fixé dans ce cours, il semble difficile de progresser sans celle-ci. Dans le cas où elle n'est pas vérifiée, il semble nécessaire de devoir travailler avec des polynômes de Ore définis sur des anneaux qui ne sont pas des corps.

6. Pour avoir l'isomorphisme, on aurait pu également utiliser la dérivation classique $\frac{d}{dU}$; cependant, celle-ci ne permet pas d'avoir les autres descriptions nécessaires à la mise en place complète du raisonnement.

Remarque 4.6.9. Plus généralement, en utilisant les résultats de l'appendice A.3 (*cf* §A.3.3), on peut démontrer que si F' est une extension de F *non nécessairement monogène* pour laquelle $C = K \subset F'$ (l'intersection étant prise dans une clôture algébrique) est stable par ∂, alors :

$$(F' \otimes_F K)[X, \partial] \simeq M_{[C:F]}\big((F' \otimes_C K)[D_{\text{lin}}(X), D{\cdot}\partial]\big)$$

où $D \in F[Y, \text{Frob}]$ est le polynôme de Ore correspondant à la sous-extension C de K/F *via* la bijection du théorème A.3.4.

Exemple 4.6.10. Poursuivons l'exploration entamée dans l'exemple 4.6.5 concernant l'anneau de Ore $\mathcal{A} = \mathbb{F}_p(t)[X, \frac{d}{dt}]$. Le centre de $\mathcal{A}$ est $\mathcal{Z} = F[X^p]$ avec $F = \mathbb{F}_p(t^p)$. Soit $N(X) = N_0(X^p)$ un polynôme central avec N_0 irréductible dans $F[T]$. Soit F' l'extension de F définie par le polynôme N_0. Soit $c(t)$ la classe de T dans $F' = F[T]/N_0(T)$.

Supposons dans un premier temps que F'/F ne soit pas séparable. Alors $N_0(T)$ est de la forme $N_0(T) = N_1(T^p)$ pour un certain polynôme N_1 à coefficients dans F. Écrivons $N_1(U) = c_0 + c_1U + \cdots + c_dU^d$ avec $c_i \in F$. Les c_i ont une racine p-ième dans F. Autrement dit, on peut écrire $c_i = f_i^p$ pour des $f_i \in \mathbb{F}_p(t)$, ce qui montre que le polynôme $N(X)$ se factorise sous la forme :

$$N(X) = \big(f_0 + f_1X^p + \cdots + f_dX^{pd}\big)^p \tag{45}$$

dans l'anneau commutatif $\mathbb{F}_p(t)[X^p]$ et donc *a fortiori* dans $\mathcal{A}$. On en déduit que N n'est pas irréductible et donc, par le théorème 4.4.3, que $e(N) > 1$. On a ainsi nécessairement $e(N) = p$, ce qui démontre grâce à une nouvelle application du théorème 4.4.3 que l'égalité (45) est une factorisation de $N(X)$ en produit de facteurs irréductibles dans $\mathcal{A}$.

Supposons à présent que l'extension F'/F soit séparable. Elle est alors linéairement disjointe de K/F où $K = \mathbb{F}_p(t)$. Le produit tensoriel $K' = F' \otimes_F K$ est alors un corps. Précisément, si $N_0(T) = c_0 + c_1T + \cdots + c_dT^d$ avec $c_i \in F$, le corps K' s'obtient à partir de K en ajoutant une racine du polynôme irréductible $f_0 + f_1X + \cdots + f_dX^d$ où $f_i \in K$ est une racine p-ième de c_i. En vertu du théorème 4.4.3 combiné à l'isomorphisme (40), on trouve que $e(N) = 1$ si et seulement si l'équation différentielle

$$\frac{d^{p-1}y(t)}{dt^{p-1}} + y(t)^p = c(t)$$

a une solution dans K'. On est ainsi amené à résoudre une équation différentielle dans une extension finie de K, ce qui géométriquement correspond à résoudre une équation différentielle sur une courbe algébrique qui apparaît comme un revêtement fini (possiblement ramifié) de la droite projective. La théorie des courbes algébriques permet d'étendre les méthodes détaillées dans l'exemple 4.6.10 à cette nouvelle situation.

A Appendices

A.1 Adjoint d'une matrice et d'une application linéaire

L'adjoint d'une application linéaire a été utilisée à plusieurs reprises dans ce cours : elle a été utile tout d'abord dans le §2 pour la définition des cofacteurs sous-résultants puis, dans un second temps, elle a joué un rôle central dans le §4 car c'est bel et bien elle qui a permis de faire le lien entre un polynôme de Ore et sa norme réduite.

Le but de cet appendice est de faire le point sur la notion d'adjoint. Dans le cas d'une matrice, cette notion est classique : l'adjoint est la transposée de la matrice de cofacteurs. Toutefois, étant donné le rôle joué par l'adjoint dans ce cours, nous avons trouvé souhaitable de prendre le temps d'approcher cette notion selon plusieurs points de vue différents et d'en établir, de manière complète, toutes les propriétés que nous avons eu à utiliser.

Dans toute la suite de l'appendice, on fixe un anneau commutatif unitaire que l'on désigne par la lettre $\mathfrak{A}$.

A.1.1 Rappels sur les algèbres extérieures

Avant d'entrer dans le vif du sujet, nous commençons par quelques rappels utiles sur les algèbres extérieures (sans démonstration). Soient E un $\mathfrak{A}$-module. Pour tout entier $m \geq 0$, la m-ième *puissance extérieure* de E, notée $\bigwedge^m E$, est définie comme le quotient de $E \otimes E \otimes \cdots \otimes E$ (m fois) par le sous-module engendré par les éléments $x_1 \otimes \cdots \otimes x_m$ pour lesquels il existe deux indices $i \neq j$ tels que $x_i = x_j$. Traditionnellement, on note $x_1 \wedge \cdots \wedge x_m$ l'image de $x_1 \otimes \cdots \otimes x_m$ dans $\bigwedge^m E$. On a ainsi $x_1 \wedge \cdots \wedge x_m = 0$ dès lors que le même vecteur apparaît au moins deux fois parmi les x_i. En développant le produit extérieur $\cdots \wedge (x+y) \wedge \cdots (x+y) \wedge \cdots$, on trouve que

$$\cdots \wedge x \wedge \cdots y \wedge \cdots = \cdots \wedge y \wedge \cdots x \wedge \cdots .$$

À partir de là, on déduit qu'étant donnés des vecteurs $x_1, \ldots, x_m \in E$ et une permutation σ de $\{1, \ldots, m\}$ de signature $\varepsilon(\sigma)$, la relation

$$x_{\sigma(1)} \wedge \cdots \wedge x_{\sigma(m)} = \varepsilon(\sigma) \cdot x_1 \wedge \cdots \wedge x_m$$

est vérifiée dans $\bigwedge^m E$.

La construction des puissances extérieures est fonctorielle : si E et F sont deux $\mathfrak{A}$-modules, toute application $\mathfrak{A}$-linéaire $f : E \to F$ induit, pour tout entier m, une application $\mathfrak{A}$-linéaire $\bigwedge^m f : E \to F$ définie par :

$$(\textstyle\bigwedge^m f)(x_1 \wedge \cdots \wedge x_m) = f(x_1) \wedge \cdots \wedge f(x_m)$$

pour $x_1, \ldots, x_m \in E$.

Lorsque l'entier 2 est inversible dans $\mathfrak{A}$, on peut voir, de manière alternative, $\bigwedge^m E$ comme un sous-module du produit tensoriel $E \otimes E \otimes \cdots \otimes E$ (m fois) *via* le morphisme $\mathfrak{A}$-linéaire :

$$x_1 \wedge \cdots \wedge x_m \mapsto \sum_{\sigma \in \mathfrak{S}_m} \varepsilon(\sigma) \cdot x_1 \otimes x_2 \otimes \cdots \otimes x_m$$

où $\mathfrak{S}_m$ désigne l'ensemble des permutations de $\{1, \ldots, m\}$.

Dualité. Si E est un $\mathfrak{A}$-module, on note $E^\star$ son dual algébrique défini par $E^\star = \mathrm{Hom}_{\mathfrak{A}}(E, \mathfrak{A})$. Les puissances extérieures de $E^\star$ ont une interprétation simple en termes de formes multilinéaires alternées. Précisément, si m est un entier, on définit une *forme m-linéaire alternée* sur E comme une application multilinéaire $f : E^m \to \mathfrak{A}$ vérifiant l'axiome supplémentaire suivante : $f(x_1, \ldots, x_m) = 0$ s'il existe deux indices distincts i et j tels que $x_i = x_j$. Comme précédemment, on démontre que cela implique que :

$$f(x_{\sigma(1)}, \ldots, x_{\sigma(m)}) = \varepsilon(\sigma) \cdot f(x_1, \ldots, x_m)$$

pour tous $x_1, \ldots, x_m \in E$ et toute permutation σ de $\{1, \ldots, m\}$. On note $\mathrm{Alt}_m(E)$ le $\mathfrak{A}$-module des formes m-linéaires alternées sur E. Lorsque $m = 1$, une forme 1-linéaire alternée n'est autre qu'une forme linéaire. Autrement dit $\mathrm{Alt}_1(E) = E^\star$. Plus généralement, pour $m \geq 1$, on dispose d'une construction qui fabrique une forme m-linéaire alternée à partir d'un élément de $\bigwedge^m E^\star$: si $\ell_1, \ldots, \ell_m$ sont des formes linéaires sur E, on fait correspondre à $\ell_1 \wedge \cdots \wedge \ell_m \in \bigwedge^m E^\star$, la forme m-linéaire alternée définie par :

$$(x_1, \ldots, x_m) \mapsto \sum_{\sigma \in \mathfrak{S}_m} \varepsilon(\sigma) \cdot \ell_1(x_{\sigma(1)})\, \ell_2(x_{\sigma(2)}) \cdots \ell_m(x_{\sigma(m)}).$$

On vérifie par le calcul que la somme du membre de droite s'annule dès lors que deux ℓ_i ou deux x_i sont égaux. On en déduit que la définition précédente a bien un sens. Cette association définit un isomorphisme $\bigwedge^m E^\star \simeq \mathrm{Alt}_m(E)$.

Par ailleurs, on dispose d'un morphisme bilinéaire $B : \mathrm{Alt}_m(E) \times \bigwedge^m E \to \mathfrak{A}$ obtenu simplement en évaluant une forme m-linéaire en un m-uplet de vecteurs, ce qui s'écrit avec des formules

$$B\big(f,\, x_1 \wedge \cdots \wedge x_m\big) = f(x_1, \ldots, x_m).$$

Le fait que f soit une forme m-linéaire alternée assure que la définition précédente a bien un sens, c'est-à-dire qu'elle ne dépend pas de la manière d'écrire un élément de $\bigwedge^m E$ comme une combinaison linéaire formelles d'éléments de la forme $x_1 \wedge \cdots \wedge x_m$. La forme bilinéaire B est une dualité parfaite et identifie ainsi $\mathrm{Alt}_m(E)$ au dual de $\bigwedge^m E$. En mettant ensemble ce résultat et celui de l'alinéa précédent, on obtient une identification canonique :

$$\bigwedge^m E^\star \simeq \big(\bigwedge^m E\big)^\star.$$

Cette identification est, en outre, fonctorielle : si $f : E \to F$ est une application $\mathfrak{A}$-linéaire entre $\mathfrak{A}$-modules, la transposée[7] de $\bigwedge^m f$ est la m-ième puissance extérieure de la transposée de f.

Cas des modules libres. On suppose à partir de maintenant que le $\mathfrak{A}$-module E est libre de rang n et on en fixe une base $(e_1, \ldots, e_n)$. Le dual de E est alors également libre de rang n et on note $(e_1^\star, \ldots, e_n^\star)$ la base duale de $(e_1, \ldots, e_n)$.

Pour toute partie I de $\{1, \ldots, n\}$ de cardinal m, on note $e_I = e_{i_1} \wedge \cdots \wedge e_{i_m}$ où $i_1 < \cdots < i_m$ sont les éléments de I triés par ordre croissant. On démontre aisément que $\bigwedge^m E$ est un module libre de rang $\binom{n}{m}$ dont une base est donnée par les e_I pour I décrivant les parties de $\{1, \ldots, n\}$ de cardinal m. En particulier, lorsque $m = n$, la puissance extérieure $\bigwedge^n E$ est de rang 1 et est engendré par le vecteur $e_1 \wedge \cdots \wedge e_n$. Pour $m > n$, on a $\bigwedge^m E = 0$.

Ce qui précède s'applique évidemment aussi à $E^\star$. En particulier, la n-ième puissance extérieure $\bigwedge^n E^\star$ est un module libre de rang 1 engendré par le vecteur $e_1^\star \wedge \cdots \wedge e_n^\star$. *Via* l'identification avec l'espace $\mathrm{Alt}_n(E)$ des formes n-linéaires alternées sur E, l'élément $e_1^\star \wedge \cdots \wedge e_n^\star$ correspond à l'application « déterminant dans la base $(e_1, \ldots, e_n)$ » que l'on note $\det_E$. Étant donné un entier $m \in \{0, \ldots, n\}$, l'application $\det_E$ peut également être vue comme une forme bilinéaire sur $\bigwedge^m E \times \bigwedge^{m-n} E$ associant à un couple $(x_1 \wedge \cdots \wedge x_m,\, x_{m+1} \wedge \cdots \wedge x_n)$ le scalaire $\det_E(x_1, \ldots, x_n)$. Cette forme bilinéaire est une dualité parfaite et identifie ainsi $\bigwedge^m E$ au dual de $\bigwedge^{m-n} E$, c'est-à-dire à $\bigwedge^{m-n} E^\star \simeq \mathrm{Alt}_{m-n}(E)$. En particulier, pour $m = n-1$, on obtient un isomorphisme entre $\bigwedge^{n-1} E$ et $E^\star$.

A.1.2 Définition et propriétés de l'adjoint

Soient E et F deux $\mathfrak{A}$-modules libres. On suppose que E et F sont tous les deux de rang n et qu'ils sont munis de bases privilégiées, notées $(e_1, \ldots, e_n)$ et $(f_1, \ldots, f_n)$ respectivement. Comme nous l'avons expliqué précédemment, le choix de ces bases définit, *via* l'application déterminant, des isomorphismes

$$\bigwedge^{n-1} E \simeq E^\star \quad \text{et} \quad \bigwedge^{n-1} F \simeq F^\star$$

où $E^\star$ et $F^\star$ sont les duaux algébriques respectifs de E et F.

Soit $f : E \to F$ une application linéaire. Sa $(n-1)$-ième puissance extérieure définit une application $\bigwedge^{n-1} E \to \bigwedge^{n-1} F$ que l'on peut voir, grâce aux identifications précédentes, comme une application linéaire de $E^\star$ dans $F^\star$. On définit l'adjoint de f, noté $\mathrm{adj}(f)$, comme la transposée de $\bigwedge^{n-1} f$. Il s'agit d'une application de F dans E. Il résulte formellement des définitions que la formation de l'adjoint commute à la transposition : si $g : F^\star \to E^\star$ désigne la transposée de f, alors $\mathrm{adj}(g) : E^\star \to F^\star$ est la transposée de $\mathrm{adj}(f) : F \to E$

En remontant les définitions, on peut donner des descriptions plus concrètes de l'adjoint. Par exemple, $\mathrm{adj}(f)$ est l'application linéaire $F \to E$ caractérisée par les égalités

$$\det_E\big(x_1, \ldots, x_{n-1}, \mathrm{adj}(f)(y)\big) = \det_F\big(f(x_1), \ldots, f(x_{n-1}), y\big). \tag{46}$$

7. C'est-à-dire l'application duale

pour $x_1, \dots, x_{n-1} \in E$ et $y \in F$. En réalité, afin de définir sans ambiguïté $\mathrm{adj}(f)(y)$, il suffit de tester l'identité (46) sur un sous-ensemble de E^{n-1} dont l'image par l'application $(x_1, \dots, x_{n-1}) \mapsto x_1 \wedge \cdots \wedge x_{n-1}$ engendre $\bigwedge^{n-1} E$.

Encore plus concrètement, du point de vue matriciel, si M désigne la matrice de f dans les bases choisies, la matrice de $\bigwedge^{n-1} f$ est la matrice des cofacteurs de M et la matrice de $\mathrm{adj}(f)$ est sa transposée. C'est la définition à laquelle notre lectrice et notre lecteur sont peut-être plus habitués.

Proposition A.1.1. *Soient E, F et G trois $\mathfrak{A}$-modules libres de même rang n munis de bases priivilégiées. Pour $f : E \to F$, $g : F \to G$ deux applications linéaires, on a $\mathrm{adj}(f \circ g) = \mathrm{adj}(g) \circ \mathrm{adj}(f)$.*

Démonstration. Cela résulte du fait que $\bigwedge^{n-1}(f \circ g) = \bigwedge^{n-1} f \circ \bigwedge^{n-1} g$ tandis que la dualité renverse le sens des flèches. □

Proposition A.1.2. *Avec les notations précédentes, on a*

$$\mathrm{adj}(f) \circ f = \det(f)\, \mathrm{id}_E \quad \text{et} \quad f \circ \mathrm{adj}(f) = \det(f)\, \mathrm{id}_F.$$

Démonstration. Soit $x \in E$. En prenant $y = f(x)$ dans (46), on trouve :

$$\begin{aligned} \det{}_E\big(x_1, \dots, x_{n-1}, \mathrm{adj}(f) \circ f(x)\big) &= \det{}_F\big(f(x_1), \dots, f(x_{n-1}), f(x)\big) \\ &= \det(f) \cdot \det{}_E(x_1, \dots, x_{n-1}, x) \\ &= \det{}_E(x_1, \dots, x_{n-1}, \det(f)\, x) \end{aligned}$$

ce qui implique que $\mathrm{adj}(f) \circ f(x) = \det(f)\, x$.

Pour démontrer la seconde égalité, on introduit g la transposée de f. Par ce qui a déjà été fait, on sait que $\mathrm{adj}(g) \circ g = \det(g)\, \mathrm{id}_{F^\star}$. Ceci donne $f \circ \mathrm{adj}(f) = \det(f)\, \mathrm{id}_F$ en tranposant puisque l'on sait, d'une part, que l'adjoint commute à la transposition et, d'autre part, que une application linéaire et sa transposée ont même déterminant. □

Lorsque f est inversible, la proposition A.1.2 nous apprend que $\mathrm{adj}(f) = \det(f)\, f^{-1}$. Lorsque f n'est pas inversible, cet formule ne vaut plus évidemment. Cependant, certaines informations sur $\mathrm{adj}(f)$ sont facilement accessibles. La proposition suivante en donne quelques unes dans le corps des corps.

Proposition A.1.3. *On suppose que $\mathfrak{A}$ est un corps.*

1. *Si le noyau de f est de dimension 1, alors $\ker \mathrm{adj}(f) = \mathrm{im} f$ et $\mathrm{im}\, \mathrm{adj}(f) = \ker f$*
2. *Si le noyau de f est de dimension au moins 2, alors $\mathrm{adj}(f) = 0$.*

Démonstration. Supposons d'abord que $\ker f$ soit de dimension 1. Sous cette hypothèse, le déterminant de f s'annule et la relation $\mathrm{adj}(f) \circ f = 0$ implique alors que $\ker \mathrm{adj}(f) \subset \mathrm{im} f$. Pour montrer l'égalité, il suffit de montrer que $\mathrm{adj}(f)$ ne s'annule pas en dehors de $\mathrm{im} f$. Soit $y \notin \mathrm{im} f$. Considérons également $y_1, \dots, y_{n-1}$ une base de $\mathrm{im} f$ (qui est bien de dimension $n-1$ par le théorème du rang). La famille $(y_1, \dots, y_{n-1}, y)$ est alors une base de F de sorte que $\det_F(y_1, \dots, y_{n-1}, y) \neq 0$. Si $x_i \in E$ désigne un antécédent de y_i, la relation (46) appliquée avec $x_1, \dots, x_{n-1}, y$ entraîne que $\mathrm{adj}(f)(y)$ est non nul.

L'égalité $f \circ \mathrm{adj}(f) = 0$ implique l'inclusion $\ker f \subset \mathrm{im}\, \mathrm{adj}(f)$. L'égalité suit à présent de l'égalité des dimensions.

Supposons à présent que $\dim \ker f \geq 2$. Soit $y \in F$. Pour $x_1, \dots, x_{n-1} \in E$, l'espace engendré par $f(x_1), \dots, f(x_{n-1})$ est de dimension au plus $n-2$. Ainsi la famille $f(x_1), \dots, f(x_{n-1}), y$ est nécessairement liée. La formule (46) entraîne alors que $\det_E(x_1, \dots, x_{n-1}, \mathrm{adj}(f)(y)) = 0$ pour tous $x_1, \dots, x_{n-1} \in E$. Ceci n'est possible que si $\mathrm{adj}(f)(y) = 0$. □

Si M est une matrice triangulaire par blocs, son adjoint $\mathrm{adj}(M)$ est encore triangulaire par blocs et ses blocs diagonaux s'expriment en fonction des blocs diagonaux de M. Ci-après, nous établissons une version de ce résultat exprimée dans le langage des applications linéaires. Étant donné un entier $m \leq n$, on note E_m (resp. F_m) le sous-module de E (resp. de F) engendré par les m premiers vecteurs de base. Le quotient E/E_m (resp. F/F_m) est un module libre de rang $n-m$ que l'on munit de la base image de $(e_{m+1}, \dots, e_n)$ (resp. $(f_{m+1}, \dots, f_n)$). Si $x_1, \dots, x_n$ sont des vecteurs de E avec $x_1, \dots, x_m \in E_m$, on a :

$$\det_E(x_1, \dots, x_n) = \det_{E_m}(x_1, \dots, x_m) \cdot \det_{E/E_m}(\bar{x}_{m+1}, \dots, \bar{x}_n)$$

où $\bar{x}_i$ désigne la classe de x_i dans le quotient E/E_m. Bien entendu, une formule analogue vaut pour F.

Proposition A.1.4. *Soit $f : E \to F$ une application linéaire et soit $m \in \{1, \dots, m\}$. On suppose que f envoie E_m sur F_m. On note $f_m : E_m \to F_m$ l'application induite par restriction et $g_m : E/E_m \to F/F_m$ l'application induite par passage au quotient.*
Alors $\mathrm{adj}(f)$ envoie F_m sur E_m et, de plus :
- *l'application $F_m \to E_m$ qu'elle induit par restriction est $\det(g_m)\,\mathrm{adj}(f_m)$, et*
- *l'application $F/F_m \to E/E_m$ qu'elle induit par passage au quotient est $\det(f_m)\,\mathrm{adj}(g_m)$.*

Démonstration. On considère des vecteurs $x_1, \dots, x_{m-1} \in E_m$, $x_{m+1}, \dots, x_n \in E$ et $y \in F_m$. D'après la formule (46) appliquée à f_m, on a :

$$\det_{E_m}\big(x_1, \dots, x_{m-1}, \mathrm{adj}(f_m)(y)\big) = \det_{F_m}\big(f(x_1), \dots, f(x_{m-1}), y\big).$$

On en déduit que :

$$\begin{aligned}
&\det_E\big(x_1, \dots, x_{m-1}, \mathrm{adj}(f_m)(y), x_{m+1}, \dots, x_n\big) \\
&\qquad = \det_{E_m}\big(x_1, \dots, x_{m-1}, \mathrm{adj}(f_m)(y)\big) \cdot \det_{E/E_m}\big(\bar{x}_{m+1}, \dots, \bar{x}_n\big) \\
&\qquad = \det_{F_m}\big(f(x_1), \dots, f(x_{m-1}), y\big) \cdot \det_{E/E_m}\big(\bar{x}_{m+1}, \dots, \bar{x}_n\big).
\end{aligned}$$

Par ailleurs :

$$\det_{F/F_m}\big(g_m(\bar{x}_{m+1}), \dots, g_m(\bar{x}_n)\big) = \det(g_m) \cdot \det_{E/E_m}\big(\bar{x}_{m+1}, \dots, \bar{x}_n\big).$$

En mettant les deux formules précédentes ensemble, on déduit :

$$\begin{aligned}
&\det_E\big(x_1, \dots, x_{m-1}, \det(g_m)\mathrm{adj}(f_m)(y), x_{m+1}, \dots, x_n\big) \\
&\qquad = \det_{F_m}\big(f(x_1), \dots, f(x_{m-1}), y\big) \cdot \det_{F/F_m}\big(g_m(\bar{x}_{m+1}), \dots, g_m(\bar{x}_n)\big) \\
&\qquad = \det_F\big(f(x_1), \dots, f(x_{m-1}), y, f(x_{m+1}), \dots, f(x_n)\big) \\
&\qquad = \det_E\big(x_1, \dots, x_{m-1}, \mathrm{adj}(f)(y), x_{m+1}, \dots, x_n\big)
\end{aligned}$$

à partir de quoi il résulte que $\mathrm{adj}(f)(y) = \det(g_m)\mathrm{adj}(f_m)(y)$. On en déduit que $\mathrm{adj}(f)$ envoie F_m sur E_m et que les applications $\mathrm{adj}(f)$ et $\det(g_m)\mathrm{adj}(f_m)$ coïncident sur F_m.

La dernière assertion se démontre de manière analogue. □

A.2 Espaces vectoriels sur les algèbres à divisions

Une *algèbre à divisions* est un anneau (non nécessairement commutatif) dans lequel tout élément non nul est inversible. Autrement dit, les axiomes d'algèbres à divisions sont exactement les axiomes de corps sauf que l'on impose pas la commutativité de la multiplication. Pour cette raison, une algèbre à divisions est parfois également appelée un *corps gauche*.

Le but de ce court appendice est d'énoncer et de démontrer les théorèmes de structure sur les espaces vectoriels (de dimension finie) sur les algèbres à divisions. En réalité, nous verrons que la théorie, tant du point de vue des résultats que des démonstrations, ressemble pratiquement en tout point à celles des espaces vectoriels usuels sur les corps commutatifs.

Bases d'un espace vectoriel. Soient D une algèbre à divisions et soit V un espace vectoriel sur D. La terminologie suivante est très classique.

Définition A.2.1. On dit que la famille $(v_1, \dots, v_n)$ d'éléments de V est *libre* s'il n'existe pas de relation de dépendance linéaire non triviale entre les v_i, c'est-à-dire concrètement si une égalité de la forme $d_1v_1 + d_2v_2 + \cdots + d_nv_n = 0$ (avec $d_i \in D$) implique $d_1 = d_2 = \cdots = d_n = 0$.
On dit que la famille $(v_1, \dots, v_n)$ d'éléments de V est *génératrice* si tout vecteur de $v \in V$ s'écrit sous la forme $v = d_1v_1 + d_2v_2 + \cdots + d_nv_n = 0$ pour des scalaires $d_1, \dots, d_n \in D$.
On dit que la famille $(v_1, \dots, v_n)$ d'éléments de V est une *base* de V si elle est à la fois libre et génératrice.

La donnée d'une famille d'éléments de V de cardinal n est équivalente à celle d'une application D-linéaire $D^n \to V$: à la famille $\mathcal{V} = (v_1, \dots, v_n)$, on fait correspondre l'application linéaire

$$\begin{array}{rccc} \varphi_{\mathcal{V}} : & D^n & \longrightarrow & V \\ & (d_1, \dots, d_n) & \mapsto & d_1v_1 + d_2v_2 + \cdots + d_nv_n. \end{array}$$

Avec ce langage, dire que la famille $\mathcal{V}$ est libre (resp. génératrice, resp. une base) revient à dire que l'endomorphisme $\varphi_{\mathcal{V}}$ est injectif (resp. surjectif, resp. bijectif). En particulier, la donnée d'une base de V équivaut à la donnée d'une identification de V avec D^n pour un certain entier n.

Théorème A.2.2. *Soit V un espace vectoriel sur D qui admet une famille génératrice finie. Alors D admet une base.*

Démonstration. Soit $\mathcal{V} = (v_1, \dots, v_n)$ une famille génératrice de V de cardinal minimal. Nous allons démontrer que $\mathcal{V}$ est une base de D. Pour cela, il suffit de montrer qu'elle est libre. Soient $d_1, \dots, d_n$ des éléments de D pour lesquels on a la relatrion de dépendance linéaire $d_1v_1 + \cdots + d_nv_n = 0$. Supposons, par l'absurde, qu'il existe un indice i tel que $d_i = 0$. Alors, en multipliant par d_i^{-1} à gauche, on peut écrire :

$$v_i = d_i^{-1}d_1v_1 + \cdots + d_i^{-1}d_{i-1}v_{i-1} + d_i^{-1}d_{i+1}v_{i+1} + \cdots + d_i^{-1}d_nv_n.$$

On en déduit que la famille $\mathcal{V}\backslash\{v_i\}$ est également une famille génératrice de V, ce qui contredit la minimalité du cardinal de $\mathcal{V}$. □

Remarque A.2.3. La démonstration que nous venons de faire assure que toute famille génératrice de cardinal minimal est une base de V. On notera qu'il est vrai, également, que toute famille libre de cardinal maximal est une base de V. En effet, considérons une telle famille $\mathcal{V} = (v_1, \dots, v_n)$ et supposons par l'absurde qu'il existe un vecteur v qui ne s'écrit pas comme combinaison linéaire des v_i. J'affirme alors que la famille $(v_1, \dots, v_n, v)$ est encore libre, contredisant ainsi la maximalité de $\mathcal{V}$. En effet, supposons donnée une relation de dépendance linéaire de la forme :

$$d_1v_1 + \cdots + d_nv_n + dv = 0 \tag{47}$$

avec $d_1, \dots, d_n, d \in D$. Si on avait $d \neq 0$, on pourrait écrire comme précédemment d comme combinaison linéaire des d_i, contredisant ainsi notre hypothèse. Ainsi $d = 0$ et la relation (47) se réduit à $d_1v_1 + \cdots + d_nv_n = 0$. On en déduit que $d_1 = \cdots = d_n = 0$ grâce à la liberté de $\mathcal{V}$. La liberté de $(v_1, \dots, v_n, v)$ s'ensuit.

Notion de dimension. Il est bien connu que toutes les bases d'un K-espace vectoriel donné sont de même cardinal, lorsque K est un corps commutatif. Ce résultat s'étend *de facto* aux espaces vectoriels sur des corps gauches. Plus précisément, on a la proposition suivante.

Proposition A.2.4. *Toute famille libre (resp. génératrice) de D^n est de cardinal $\leq n$ (resp. de cardinal $\geq n$).*

Démonstration. Soit $(v_1, \dots, v_m)$ une famille libre de D^n. Le vecteur v_1 est évidemment non nul. Quitte à permuter l'ordre de coordonnées sur D_n, on peut supposer que sa première coordonnée, notée d_1 est non nulle. Posons $w_1 = d_1^{-1} v_1$. La famille $(w_1, v_2, \dots, v_m)$ est clairement encore libre. On écrit à présent $v_2 = (\delta_1, \delta_2, \dots, \delta_n)$ et on pose $w_2' = v_2 - \delta_1 w_1$ de manière à annuler la première coordonnée de w_2'. La liberté de la famille $(w_1, v_2, \dots, v_m)$ assure que le vecteur w_2' n'est pas nul. Il a ainsi une coordonnée non nulle et, quitte à faire une nouvelle permutation des coordonnées sur D^n, on peut supposer que c'est la deuxième. Notons d_2 la valeur de cette coordonnée et posons $w_2 = d_2^{-1} w_2$. On vérifie sans peine que la famille $(w_1, w_2, v_3, \dots, v_n)$ est libre. De proche en proche, on construit une nouvelle famille *libre* $(w_1, \dots, w_m)$ ayant la propriété suivante : pour tout i, les $(i-1)$ premières coordonnées de w_i sont nulles. Si on avait $m > n$, ceci impliquerait que le vecteur w_{n+1} s'annule, ce qui ne peut se produire dans une famille libre. On en déduit que $m \leq n$ comme voulu.

Soit maintenant $\mathcal{V} = (v_1, \dots, v_m)$ une famille génératrice de D^n. À partir de $\mathcal{V}$, on construit une famille *génératrice* $(w_1, \dots, w_{m'})$ de D^n de cardinal $m' \leq m$ ayant en outre la propriété suivante : pour tout i, les $(i-1)$ premières coordonnées de w_i sont nulles et la i-ième coordonnée de w_i vaut 1. La construction est similaire à ce qui a été fait à l'alinéa précédent sauf que si l'un des vecteurs w_i' rencontrés en cours de route s'annule, on le supprime simplement de la famille. Supposons par l'absurde que $m' < n$. J'affirme alors que le vecteur $v = (0, \dots, 0, 1)$ ne saurait s'écrire comme combinaison linéaire des w_i. En effet, si une écriture

$$v = d_1 w_1 + d_2 w_2 + \cdots + d_{m'} w_{m'}$$

existait, on obtiendrait $d_1 = 0$ en examinant la première coordonnée, puis $d_2 = 0$ en examinant la deuxième et, ainsi de suite, $d_i = 0$ pour tout $i \in \{1, \dots, m'\}$. Ceci impliquerait donc en particulier que $v = 0$, ce qui n'est pourtant pas le cas. On en déduit que $m' \geq n$ et donc *a fortiori* que $m \geq n$. □

Corollaire A.2.5. *Soit V un espace vectoriel sur D qui admet une famille génératrice finie. Alors toutes les bases de V ont le même cardinal.*

Démonstration. Soit $\mathcal{V}$ une base de D dont on note n le cardinal. L'application $\varphi_{\mathcal{V}}$ définit alors une identification entre V et D^n. Le corollaire résulte ainsi directement de la proposition A.2.4. □

Comme dans le cas commutatif, le cardinal commun des bases de V est appelé la *dimension* de V sur D et se note $\dim_D V$ ou, plus simplement, $\dim V$ lorsque cela ne prête pas à confusion. On dit que V est de *dimension finie* lorsqu'il admet une base finie, c'est-à-dire d'après le théorème A.2.2 lorsqu'il admet une famille génératrice finie.

Une autre corollaire de la proposition A.2.4 est le théorème dit de la base incomplète qui s'énonce comme suit.

Proposition A.2.6 (Théorème de la base incomplète)**.** *Soit V un D-espace vectoriel de dimension finie n. Soit $(v_1, \dots, v_m)$ une famille libre d'éléments de V. Alors il existe des vecteurs $v_{m+1}, \dots, v_n \in V$ tels que la famille $(v_1, \dots, v_n)$ soit une base de V.*

Démonstration. Choisissons une famille libre $\mathcal{V}$ de la forme $(v_1, \dots, v_m, w_1, \dots, w_s)$ de cardinal maximal. Une telle famille existe car on sait, d'une part, que $(v_1, \dots, v_m)$ est une famille libre et, d'autre part, que le cardinal de toute famille libre est majoré par n. En reprenant la démonstration faite dans la remarque A.2.3, on voit que la famille $\mathcal{V}$ est une base de V et est donc également de cardinal n. □

Les propriétés classiques de la dimension demeurent également. La proposition suivante en donne une liste, bien entendu non exhaustive.

Proposition A.2.7. *Soit V et W deux D-espaces vectoriels de dimension finie.*

(i) *si* $\mathcal{V} = (v_1, \dots, v_{\dim V})$ *est une famille de cardinal* $\dim V$*, alors* $\mathcal{V}$ *est libre si et seulement si* $\mathcal{V}$ *est génératrice si et seulement si* $\mathcal{V}$ *est une base ;*

(ii) *si* $V \subset W$*, alors* $\dim V \leq \dim W$ *et l'égalité des dimensions a lieu si et seulement si* $V = W$ *;*

(iii) $\dim(V \oplus W) = \dim V + \dim W$ *;*

(iv) *si* $W \subset V$*, alors* $\dim(V/W) = \dim V - \dim W$ *;*

(v) *pour toute application* D*-linéaire* $f : V \to W$*, on a* $\dim \ker f + \dim \mathrm{im} f = \dim V$.

Esquisse de démonstration. Le (i) résulte du fait qu'une base est une famille libre de cardinal maximal ou, au choix, une famille génératrice de cardinal minimal (*cf* remarque A.2.3). Pour le (ii), on remarque qu'une base de V est une famille libre de W, et qu'elle est en outre de cardinal maximal en cas d'égalité des dimensions. Le (iii) est immédiat. Le (iv) résulte du théorème de la base incomplète et le (v), enfin, résulte de l'isomorphisme standard $V/\ker f \simeq \mathrm{im} f$. □

Applications linéaires et matrices. Soit D^{op} l'anneau opposé à D. Par définition, D^{op} et D ont le même ensemble sous-jacent et la même loi d'addition mais la loi de multiplication sur D^{op}, notée $\&$, est défini par $a\&b = ba$ pour $a, b \in D^{\mathrm{op}} = D$. Bien sûr, D^{op} est aussi une algèbre à divisions.

Soient m et n deux entiers. La donnée d'une application linéaire $f : D^n \to D^m$ est équivalente à la donnée de la double famille $(d_{i,j})_{1 \leq i \leq n, 1 \leq j \leq m}$ d'éléments de D ainsi définie :

$$f(0, \dots, 0, 1, 0, \dots, 0) = (d_{i,1}, d_{i,2}, \dots, d_{i,m})$$

où dans le vecteur du membre du gauche, la coordonnée égale à 1 est en i-ième position. Il est commode de considérer la famille des $d_{i,j}$ non pas comme une matrice à coefficients dans D, mais comme une matrice à coefficients dans D^{op}. On la note alors $\mathrm{Mat}(f)$. Avec cette convention, on dispose de la relation $\mathrm{Mat}(f \circ g) = \mathrm{Mat}(f) \times \mathrm{Mat}(g)$ où $\times$ désigne le produit matriciel usuel. Cette construction définit ainsi une bijection D-linéaire $\mathrm{Hom}_D(D^n, D^m) \to M_{n,m}(D^{\mathrm{op}})$ qui, lorsque $n = m$, est un isomorphisme d'anneaux $\mathrm{End}(D^n) \simeq M_n(D^{\mathrm{op}})$.

Si maintenant V et W sont deux espaces vectoriels de dimension finie équipés respectivement des bases $\mathcal{V} = (v_1, \dots, v_n)$ et $\mathcal{W} = (w_1, \dots, w_m)$ et si $f : V \to W$ est une application linéaire, on définit la matrice de f dans les bases $\mathcal{V}$ et $\mathcal{W}$ comme la matrice de l'application composée :

$$D^n \xrightarrow{\varphi_{\mathcal{V}}} V \xrightarrow{f} W \xrightarrow{\varphi_{\mathcal{W}}^{-1}} D^m.$$

Une fois que les bases $\mathcal{V}$ et $\mathcal{W}$ ont été fixées, la construction précédente établit un isomorphisme D-linéaire $\mathrm{Hom}_D(V, W) \to M_{n,m}(D^{\mathrm{op}})$. Lorsque $V = W$ et $\mathcal{V} = \mathcal{W}$, cet isomorphisme est, de plus, compatible à la structure d'anneau.

Espaces projectifs. Soit V un espace vectoriel de dimension finie. Par définition, un *hyperplan* de V est le noyau d'une forme linéaire surjective $\varphi : V \to D$. On définit l'espace projectif $\mathbb{P}(V)$ comme l'ensemble des hyperplans de V.

Si on fixe une base de V, la donnée d'une forme linéaire surjective $\varphi : V \to D$ est équivalente à la donnée de sa matrice (dans la base fixée) $\mathrm{Mat}(\varphi)$. En notant $n = \dim V$, ceci équivaut encore à la donnée d'un n-uplet $(d_1, \dots, d_n)$ d'éléments de D. De plus, en utilisant le théorème de la base incomplète, on démontre que deux formes linéaires φ et ψ ont même noyau si et seulement si $\varphi = d\psi$ pour un certain $d \in D$, $d \neq 0$. On en déduit que le choix d'une base de V détermine une identification

$$\mathbb{P}(V) \simeq (D^n \backslash \{0\}) \,/\, (D \backslash \{0\})$$

comme dans le cas classique. On fera attention toutefois à ce que, dans le quotient précédent, D agit par multiplication à *gauche* sur D^n. Autrement dit, on identifie les n-uplets $(d_1, \dots, d_n)$ et $(d'_1, \dots, d'_n)$ lorsqu'il existe $d \in D$, $d \neq 0$ tel que $d'_i = dd_i$ pour tout i.

On pose $\mathbb{P}^{n-1}(D) = (D^n\backslash\{0\})/(D\backslash\{0\})$ comme d'habitude et on utilise la notation usuelle $[d_1 : \ldots : d_n]$ pour représenter un élément de $\mathbb{P}^{n-1}(D)$. L'espace affine D^{n-1} s'injecte naturellement dans $\mathbb{P}^{n-1}(D)$ par l'application $(d_1, \ldots, d_{n-1}) \mapsto [d_1 : \ldots : d_{n-1} : 1]$. De plus, les éléments qui ne sont pas dans l'image de cette injection sont ceux de la forme $[d_1 : \ldots : d_{n-1} : 0]$ et sont donc en bijection naturelle avec $\mathbb{P}^{n-2}(D)$. Autrement dit :

$$\mathbb{P}^{n-1}(D) \simeq D^{n-1} \sqcup \mathbb{P}^{n-1}(D) \simeq D^{n-1} \sqcup D^{n-2} \sqcup \mathbb{P}^{n-2}(D) \simeq D^{n-1} \sqcup D^{n-2} \sqcup \cdots \sqcup D^0.$$

En particulier, pour $n = 2$, on voit que $\mathbb{P}^1(D)$ est l'union de D et d'un point supplémentaire, souvent appelé le *point à l'infini*.

A.3 Extensions différentielles finies en caractéristique positive

Dans ce deuxième appendice, nous nous plaçons dans le contexte suivant. Soit K un corps de caractéristique $p > 0$ muni d'une dérivation $\partial : K \to K$. On note F le sous-corps des constantes de K, c'est-à-dire le sous-corps de F formé des éléments $x \in K$ tels que $\partial(x) = 0$. Remarquons d'ores et déjà que F contient les puissances p-ièmes de K. En effet, si $x \in K$, on a $\partial(x^p) = px^{p-1}\partial(x) = 0$ puisque K est de caractéristique p. On en déduit en particulier que l'extension K/F est purement inséparable. Dans toute cet appendice, *on suppose que l'extension K/F est finie*. D'après ce qui précède, ceci est automatique dès lors que le sous-corps de K des puissances p-ièmes est d'indice fini dans K. Soit $\mathrm{Der}(K)$ le K-espace vectoriel des dérivations de K. Le lemme suivant est classique et facile, mais bien utile.

Lemme A.3.1. *Lorsque K est de caractéristique positive p, $\mathrm{Der}(K)$ est stable par l'opération d'élévation à la puissance p.*

Démonstration. Étant donné $d \in \mathrm{Der}(K)$, la formule de Leibniz entraîne, pour $a, b \in K$, la relation suivante :

$$d^p(ab) = \sum_{i=0}^{p} \binom{p}{i} d^i(a)d^{p-i}(b) = ad^p(b) + d^p(a)b$$

qui montre que d^p est encore une dérivation de K. □

A.3.1 Le polynôme de Ore $\mathrm{Ann}(\partial)$

Considérons l'anneau de Ore $F[Y, \mathrm{Frob}]$ où $\mathrm{Frob} : F \to F, x \mapsto x^p$ est le morphisme de Frobenius usuel. Rappelons (*cf* exemple 1.1.3) qu'à $P(X) = a_0 + a_1Y + \cdots + a_nY^n \in F[Y, \mathrm{Frob}]$, on peut associer un polynôme commutatif P_{lin} défini par

$$P_{\mathrm{lin}}(T) = a_0T + a_1T^p + \cdots + a_nT^{p^n} \in F[T]$$

de manière à ce que l'on ait $(PQ)_{\mathrm{lin}} = P_{\mathrm{lin}} \circ Q_{\mathrm{lin}}$ pour tous $P, Q \in F[Y, \mathrm{Frob}]$. L'anneau $F[Y, \mathrm{Frob}]$ agit alors sur $\mathrm{Der}(K)$ comme suit : pour $P \in F[X, \mathrm{Frob}]$ et $d \in \mathrm{Der}(K)$, on pose $P \cdot d = P_{\mathrm{lin}}(d)$. Cette action est compatible à la structure d'anneau dans le sens où

$$(P + Q) \cdot d = P \cdot d + Q \cdot d$$
$$(PQ) \cdot d = P \cdot (Q \cdot d)$$

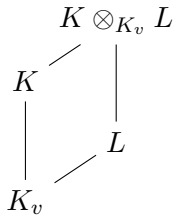
pour $P, Q \in F[Y, \mathrm{Frob}]$ et $d \in \mathrm{Der}(K)$. On en déduit que l'ensemble $\mathcal{I}$ de polynômes de Ore $P \in F[Y, \mathrm{Frob}]$ tels que $P \cdot \partial = 0$ est un idéal à gauche de $F[Y, \mathrm{Frob}]$. De plus, l'hypothèse selon laquelle K/F est une extension finie implique que $\mathcal{I} \neq 0$. En effet, on a une inclusion de K-espaces vectoriels $\mathrm{Der}(K) \subset \mathrm{End}_F(K)$ de laquelle il résulte de $\mathrm{Der}(K)$ est de dimension finie sur K. Ainsi, il existe nécessairement une relation de dépendance linéaire entre les ∂^{p^i} (pour i variant dans $\mathbb{N}$) et donc, par suite, un polynôme de Ore non nul dans l'idéal $\mathcal{I}$. Par ailleurs, d'après le corollaire 1.3.7, on sait que l'idéal $\mathcal{I}$ est principal. Il existe donc un unique polynôme de Ore unitaire $\mathrm{Ann}(\partial) \in F[Y, \mathrm{Frob}]$ pour lequel $\mathcal{I} = \big(\mathrm{Ann}(\partial)\big)$.

Proposition A.3.2. *Avec les notations précédentes, si $\mathrm{Ann}(\partial)$ est un polyôme de Ore de degré m, alors l'extension K/F est de degré p^m.*

Démonstration. Conformément à l'énoncé de la proposition, notons m le degré du polynôme de Ore $\mathrm{Ann}(\partial)$. Écrivons $\mathrm{Ann}(\partial)(Y) = P(Y) \cdot Y^v$ avec $P \in F[Y, \mathrm{Frob}]$ de coefficient constant non nul. Ainsi :

$$\mathrm{Ann}(\partial)_{\mathrm{lin}}(T) = P_{\mathrm{lin}}(T^{p^v})$$

et le polynôme $P_{\mathrm{lin}}(T)$ est séparable. En effet, son polynôme dérivé $P'_{\mathrm{lin}}(T)$ est constant, non nul. Soient $\partial_v = \partial^{p^v}$ et K_v le corps des constantes pour ∂_v. Soit L une extension séparable de K_v dans laquelle le polynôme $P_{\mathrm{lin}}(T)$ se scinde. Formons le diagramme suivant :



Les extensions K/K_v et L/K_v sont linéairement disjointes étant donné que la première est purement inséparable alors que la seconde est séparable. Il s'ensuit que $K \otimes_{K_v} L$ est un corps. De plus $[K : K_v] = [K \otimes_{K_v} L : L]$. Munissons $K \otimes_{K_v} L$ de la dérivation $\delta = \partial_v \otimes \mathrm{id}$ et voyons celle-ci comme un endomorphisme L-linéaire de $K \otimes_{K_v} L$. Clairement $P_{\mathrm{lin}}(\delta) = 0$. Du fait que P_{lin} est scindé à racines simples dans L, on déduit que δ est diagonalisable. Autrement dit, si $V \subset L$ désigne l'ensemble des racines de P_{lin} et si, pour $\lambda \in V$, on pose $E_\lambda = \ker(\delta - \lambda)$, on a la décomposition :

$$K \otimes_{K_v} L = \bigoplus_{\lambda \in V} E_\lambda. \tag{48}$$

Remarquons en outre que $E_0 = \ker \delta = L$.

Par ailleurs, la structure particulière du polynôme P_{lin} montre que V est un sous-$\mathbb{F}_p$-espace vectoriel de L. Soit $\lambda \in V$. Fixons des éléments non nuls $x \in E_\lambda$ et $y \in E_{-\lambda}$. Il est facile de vérifier que la multiplication par x définit une application L-linéaire $m_x : L \to E_\lambda$ et que, de la même manière, la multiplication par y définit une application L-linéaire $m_y : E_\lambda \to L$. La composée $m_y \circ m_x : L \to L$ est, bien sûr, la multiplication par xy. Étant donné que $K \otimes_{K_v} L$ est un corps, on a $xy \neq 0$, ce qui entraîne que $m_y \circ m_x$ est bijective. Ainsi m_x et m_y le sont également. Ceci démontre que E_λ est un espace vectoriel de dimension 1. Comme ceci est vrai pour tout λ, on déduit de la décomposition (48) que :

$$[K \otimes_{K_v} L : L] = \mathrm{Card}\, V = \deg_T P_{\mathrm{lin}}(T) = p^{\deg_Y P(Y)} = p^{m-v}. \tag{49}$$

Montrons à présent que l'extension K_v/F est de degré p^v. Pour cela, introduisons les corps $K_i = \ker(\partial^{p^i})$ pour $0 \leq i \leq v$. On a $F = K_0 \subset K_1 \subset \cdots \subset K_v$. De plus, pour tout indice i, la dérivation ∂^{p^i} définit un endomorphisme K_i-linéaire de K_{i+1}. Celui-ci est nilpotent d'indice p et son noyau est exactement K_i par définition. On en déduit que K_{i+1} est un espace vectoriel de dimension p sur K_i, c'est-à-dire que l'extension K_{i+1}/K_i est de degré p. En mettant ensemble ces résultats pour $i \in \{0, \ldots, v-1\}$, on trouve que K_v/F est de degré p^v puis finalement, en combinant avec Eq. (49), que $[K : F] = p^m$. □

Remarque A.3.3. Un examen attentif de la démonstration précédente montre que $\mathrm{Ann}(\partial)_{\mathrm{lin}}$ est le polynôme minimal de la dérivation ∂, vue comme endomorphisme F-linéaire de K. En particulier, si $m = \deg \mathrm{Ann}(\partial)$, la famille $(\mathrm{id}_K, \partial, \partial^2, \ldots, \partial^{p^m-1})$ est libre sur F. Ce résultat peut être vu comme un équivalent différentiel du théorème d'Artin d'indépendance linéaire des caractères.

A.3.2 Une correspondance de type Galois

La correspondance de Galois classique établit une bijection entre les extensions intermédiaires d'une extension galoisienne et les sous-groupes de son groupe de Galois. Dans le cas des extensions purement inséparables, un analogue de la correspondance de Galois a été mis au point par Jacobson [16, 17, 18]. Dans ce numéro, nous présentons cette correspondance dans le cas particulier des extensions K/F étudiées dans cet appendice. Comme précédemment donc, on fixe un corps K de caractéristique p muni d'une dérivation $\partial : K \to K$ et on suppose que le sous-corps F des constantes est d'indice fini dans K.

Soit L un sous-corps de K contenant F. On suppose que L est stable par la dérivation ∂ et on note $\partial_{|L}$ la restriction de ∂ à L que l'on considère comme une dérivation de L. Clairement, le corps des constantes pour la dérivation $\partial_{|L}$ est encore F. Les constructions que nous avons faites précédemment s'appliquent au corps différentiel $(L, \partial_{|L})$ et permettent de définir un polynôme de Ore $\mathrm{Ann}(\partial_{|L}) \in F[Y, \mathrm{Frob}]$; c'est le polynôme de Ore de plus petit degré dont le polynôme linéarisé associé annule la dérivation $\partial_{|L}$. Du fait que $\mathrm{Ann}(\partial)$ annule également $\partial_{|L}$ (puisqu'il annule ∂ par construction), on déduit que $\mathrm{Ann}(\partial_{|L})$ est un diviseur à droite de $\mathrm{Ann}(\partial)$ dans l'anneau $F[Y, \mathrm{Frob}]$. Réciproquement, à un diviseur à droite D de $\mathrm{Ann}(\partial)$, on peut associer le sous-corps de K formé des constantes pour la dérivation $D{\cdot}\partial$ (qui est définie, rappelons-le, comme $D_{\mathrm{lin}}(\partial)$), noté $\ker(D{\cdot}\partial)$ dans la suite.

Théorème A.3.4. *Les applications :*

$$\left\{ \begin{array}{c} \textit{sous-corps de } K \\ \textit{contenant } F\textit{, stables par } \partial \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{diviseurs à droite} \\ \textit{unitaires de } \mathrm{Ann}(\partial) \end{array} \right\}$$
$$L \mapsto \mathrm{Ann}(\partial_{|L})$$
$$\ker(D{\cdot}\partial) \hookleftarrow D$$

sont des bijections inverses l'une de l'autre.
De plus, ces bijections sont croissantes dans le sens où si L_1 et L_2 sont deux sous-extensions de K/F stables par ∂ et si D_1 et D_2 désignent les diviseurs de $\mathrm{Ann}(\partial)$ qui leur correspondent, alors $L_1 \subset L_2$ si et seulement si D_1 divise D_2 à droite.

Démonstration. Soit D un diviseur à droite unitaire de $\mathrm{Ann}(\partial)$ et soit $L = \ker(D{\cdot}\partial)$. Par définition, la dérivation $D{\cdot}\partial$ s'annule sur L. On en déduit que $\mathrm{Ann}(\partial_{|L})$ divise D à droite. De plus, la proposition A.3.2 appliquée au corps différentiel (L, ∂) nous apprend que :

$$[L{:}F] = p^{\deg \mathrm{Ann}(\partial_{|L})} \leq p^{\deg D}. \tag{50}$$

Écrivons à présent $\mathrm{Ann}(\partial) = D'D$ avec $D' \in F[Y, \mathrm{Frob}]$. La dérivation $D{\cdot}\partial$ est alors annulée par D'. Il en résulte que D' est un diviseur à droite de $\mathrm{Ann}(D{\cdot}\partial)$. La proposition A.3.2 appliquée au corps différentiel $(K, D{\cdot}\partial)$ entraîne alors :

$$[K{:}L] = p^{\deg \mathrm{Ann}(D\cdot\partial)} \leq p^{\deg D'}. \tag{51}$$

Étant donné que $[K{:}L] \cdot [L{:}F] = [K{:}F] = p^{\deg \mathrm{Ann}(\partial)} = p^{\deg D + \deg D'}$, la seule possibilité est que les inégalités (50) et (51) soient des égalités. On en déduit que les polynômes de Ore D et $\mathrm{Ann}(\partial_{|L})$ et, par suite, qu'ils sont égaux étant donné que l'on sait déjà que le second divise le premier.

Pour l'autre sens, soit L un sous-corps de K contenant F et stable par ∂. On pose $\delta = \mathrm{Ann}(\partial_{|L}){\cdot}\partial$. L'inclusion $L \subset \ker \delta$ est évidente au vu des définitions. Par ailleurs, il est clair également que $\mathrm{Ann}(\partial_{|L})$ annule la dérivation $\partial_{|\ker\delta}$. Autrement dit, $\mathrm{Ann}(\partial_{|L})$ est un multiple à droite de $\mathrm{Ann}(\partial_{|\ker\delta})$. On déduit ainsi de la proposition A.3.2 appliquée aux corps différentiels $(\ker\delta, \partial_{|\ker\delta})$ et $(L, \partial_{|L})$ que :

$$[\ker\delta : F] = p^{\deg \mathrm{Ann}(\partial_{|\ker\delta})} \leq p^{\deg \mathrm{Ann}(\partial_{|L})} = [L : F].$$

En combinant cela avec l'inclusion $L \subset \ker \delta$, on obtient $\ker \delta = L$ comme voulu.

On déduit de l'argumentation précédente que les deux applications définies dans l'énoncé de la proposition sont des bijections inverses l'une de l'autre. La propriété de croissance se vérifie facilement. En effet, si $L_1 \subset L_2$, le polynôme de Ore $\mathrm{Ann}(\partial_{|L_2})$ annule $\partial_{|L_2}$ par définition et il annule donc *a fortiori* $\partial_{|L_1}$. Ainsi $\mathrm{Ann}(\partial_{|L_1})$ divise $\mathrm{Ann}(\partial_{|L_2})$. Réciproquement si D_1 divise D_2, on peut écrire $D_2 = DD_1$, ce qui implique que $D_2 \cdot \partial = D \cdot (D_1 \cdot \partial)$ et par suite que $\ker(D_1 \cdot \partial) \subset \ker(D_2 \cdot \partial)$. □

A.3.3 Extension du corps des constantes

Dans le contexte galoisien, si on se donne une extension galoisienne K/F de degré n, il est bien connu que le produit tensoriel $K \otimes_F K$ se décompose comme un produit de n copies de K. Plus généralement, si F'/F est une seconde extension, le produit tensoriel $F' \otimes_F K$ s'écrit comme un produit d'un certain nombre de copies de l'extension composée $F'K$ (vue, par exemple, dans une clôture algébrique). Dans ce paragraphe, nous étudions l'équivalent de ces phénomènes dans le cadre différentiel en caractéristique p. Comme précédemment, on fixe un corps différentiel (K, ∂) et on suppose que le sous-corps F des constantes est d'indice fini dans K. On se donne, en outre, une deuxième extension finie F'/F (que l'on munit, si on le souhaite, de la dérivation nulle). Le produit tensoriel $F' \otimes_F K$ hérite de la dérivation $\mathrm{id} \otimes \partial$.

Soit C le plus grand sous-corps de K qui se plonge dans F'. Dans toute la suite, *on suppose C est stable par ∂*. D'après le théorème A.3.4, C est de la forme $C = \ker(D \cdot \partial)$ pour un certain polynôme de Ore unitaire $D \in F[Y, \mathrm{Frob}]$ qui est un diviseur à droite de $\mathrm{Ann}(\partial)$. De plus, les extensions K/C et F'/C sont linéairement disjointes, de sorte que le produit tensoriel $K' = F' \otimes_C K$ est un corps. Dans la suite, on identifiera sans commentaire supplémentaire F' et K à des sous-corps de K'. Le diagramme suivant résume la situation.

$$\begin{array}{ccc} & & K' = F' \otimes_C K \\ & \nearrow & | \\ K & & F' \\ | & \nearrow & \\ C & & \\ | & & \\ F & & \end{array}$$

On pose $d = \deg D$ et $m = \deg \mathrm{Ann}(\partial)$. Les extensions C/F et K/C sont respectivement de degré p^d et p^{m-d}. Considérons l'anneau commutatif de polynômes à puissances divisées tronqués $K'[T]^{\mathrm{pd}}_{<p^d}$ dont les éléments sont les sommes formelles $a_0 T^{[0]} + a_1 T^{[1]} + \cdots + a_{p^d-1} T^{[p^d-1]}$ pour des coefficients a_i dans K'. Dans l'écriture précédente, les $T^{[i]}$ désignent des symbôles qui se multiplient selon la règle suivante

$$\begin{aligned} T^{[i]} \cdot T^{[j]} &= \tbinom{i+j}{i} \cdot T^{[i+j]} && \text{si } i + j < p^d \\ &= 0 && \text{sinon} \end{aligned}$$

pour $i, j \in \{0, 1, \ldots, p^d - 1\}$. Intuitivement $T^{[i]}$ doit être pensé comme $\frac{T^i}{i!}$ sauf que cette division n'a pas de sens en caractéristique p lorsque $i \geq p$. Dans la suite, on notera souvent simplement 1 pour $T^{[0]}$ et T pour $T^{[1]}$.

Proposition A.3.5. *L'application*

$$\begin{array}{rccc} \varphi : & F' \otimes_F K & \longrightarrow & K'[T]^{\mathrm{pd}}_{<p^d} \\ & \lambda \otimes f & \mapsto & \lambda \cdot \displaystyle\sum_{0 \leq i < p^d} \partial^i(f)\, T^{[i]} \end{array}$$

est un isomorphisme d'anneaux.

Démonstration. Commençons par démontrer que φ est un morphisme d'anneaux. La compatibilité à l'addition est immédiate. Pour ce qui concerne la multiplication, il suffit de démontrer que $\varphi(\lambda\mu \otimes fg) = \varphi(\lambda \otimes f) \cdot \varphi(\mu \otimes g)$ pour $\lambda, \mu \in F'$ et $f, g \in K$. Calculons :

$$\begin{aligned}\varphi(\lambda \otimes f) \cdot \varphi(\mu \otimes g) &= \lambda\mu \cdot \left(\sum_{0 \leq i < p^d} \partial^i(f)\, T^{[i]} \right) \cdot \left(\sum_{0 \leq j < p^d} \partial^j(g)\, T^{[j]} \right) \\ &= \lambda\mu \cdot \sum_{0 \leq i,j < p^d} c_{i,j} \partial^i(f) \partial^j(g) T^{[i+j]}\end{aligned}$$

où $c_{i,j} = \binom{i+j}{i}$ si $i + j < p^d$ et $c_{i,j} = 0$ sinon. En regroupant les termes de la somme double selon la valeur de $s = i + j$, on obtient :

$$\begin{aligned}\varphi(\lambda \otimes f) \cdot \varphi(\mu \otimes g) &= \lambda\mu \cdot \sum_{0 \leq s < p^d} \sum_{i=0}^{s} \binom{s}{i} \partial^i(f) \partial^{s-i}(g)\, T^{[s]} \\ &= \lambda\mu \cdot \sum_{0 \leq s < p^d} \partial^s(fg)\, T^{[s]} \;=\; \varphi(\lambda\mu \otimes fg).\end{aligned}$$

La multiplicativité est ainsi démontrée. Il reste à démontrer que φ est bijectif. Pour cela, il suffit de vérifier que l'application :

$$\begin{array}{rccc} \psi : & C \otimes_F K & \longrightarrow & K[T]^{\mathrm{pd}}_{<p^d} \\ & \lambda \otimes f & \mapsto & \lambda \cdot \sum_{i=0}^{p^d-1} \partial^i(f)\, T^{[i]} \end{array}$$

est un isomorphisme puisque φ se déduit de ψ par extension des scalaires de C à F'. Étant donné que la source et le but de ψ sont tous les deux des K-espaces vectoriels de dimension p^d et que ψ est K-linéaire (puisque c'est un morphisme d'anneaux), il suffit de démontrer que ψ est surjectif. Raisonnons par l'absurde en supposons que l'image de ψ soit incluse dans un hyperplan. Ceci impliquerait l'existence de scalaires $\lambda_0, \ldots, \lambda_{p^d-1} \in K$ non tous nuls tels que :

$$\lambda_0 f + \lambda_1 \partial(f) + \lambda_2 \partial^2(f) + \cdots + \lambda_{p^d-1} \partial^{p^d-1}(f) = 0 \tag{52}$$

pour tout $f \in K$. On suppose que la relation précédente est choisie de manière à faire intervenir un nombre minimal de λ_i non nuls. Soit i_0 un indice pour lequel $\lambda_{i_0} \neq 0$. Quitte à diviser par λ_{i_0}, on peut supposer que $\lambda_{i_0} = 1$. En appliquant ∂ à (52), on aboutit à la nouvelle relation :

$$\partial(\lambda_0) f + \partial(\lambda_1) \partial(f) + \partial(\lambda_2) \partial^2(f) + \cdots + \partial(\lambda_{p^d-1}) \partial^{p^d-1}(f) = 0$$

dans laquelle le coefficient $\partial(\lambda_{i_0}) = \partial(1)$ est désormais nul. Par minimalité, on en déduit que $\partial(\lambda_i) = 0$ pour tout i. Autrement dit, la relation de dépendance linéaire (52) est à coefficients dans le corps des constantes F. On en déduit que la famille $(\mathrm{id}, \partial, \ldots, \partial^{p^d-1})$ est liée sur F, ce qui entre en contradiction avec la remarque A.3.3. La surjectivité de ψ est ainsi démontrée. Il en résulte, comme nous l'avons déjà expliqué, que φ est un isomorphisme. □

Le produit tensoriel $F' \otimes_F K$ est naturellement muni de la dérivation $\mathrm{id} \otimes \partial$ que, dans un léger abus de notations, on notera simplement ∂ dans la suite. Afin de compléter la proposition A.3.5, nous nous proposons de décrire la dérivation de $K'[T]^{\mathrm{pd}}_{<p^d}$ induite par ∂ *via* l'isomorphisme φ. Pour cela, munissons K' de la dérivation $\mathrm{id} \otimes D{\cdot}\partial$ que, par abus, on notera simplement $D{\cdot}\partial$ dans la suite [8]. L'anneau $K'[T]^{\mathrm{pd}}_{<p^d}$ est muni de plusieurs dérivations intéressantes. Il y a, dans un

8. On prendra garde au fait que la dérivation ∂, elle, ne s'étend pas à K' car ∂ ne s'annule pas sur C.

premier temps, la dérivation naturelle $\partial_T = \frac{d}{dT}$ qui envoie $T^{[i]}$ sur $T^{[i-1]}$ pour tout i. Pour $j < d$, on introduit la dérivation $\partial_j = T^{[p^d-1]} \cdot \partial_T^{p^j}$. Concrètement, on a :

$$\partial_j\big(a_0 + a_1T + \cdots + a_{p^d-1}T^{[p^d-1]}\big) = a_{p^j}\, T^{[p^d-1]}.$$

Enfin, on définit la dérivation $\partial' : K'[T]^{\mathrm{pd}}_{<p^d} \to K'[T]^{\mathrm{pd}}_{<p^d}$ par la formule :

$$\partial'\big(a_0 + a_1T + \cdots + a_{p^d-1}T^{[p^d-1]}\big) = (D{\cdot}\partial)(a_0)\, T^{[p^d-1]}.$$

À partir de ces dérivations, fabriquons la dérivation ∂_D définie par :

$$\partial_D = \partial_T - (c_0\partial_0 + c_1\partial_1 + \cdots + c_{d-1}\partial_{d-1}) + \partial'$$

où les c_i sont les coefficients de $D \in F[Y, \mathrm{Frob}]$, *i.e.* $D = c_0 + c_1Y + \cdots + c_{d-1}Y^{d-1} + Y^d$.

Proposition A.3.6. *Pour tout $f \in E \otimes_F L$, on a la relation de compatibilité $\varphi(\partial(f)) = \partial_D(\varphi(f))$. Autrement dit, la dérivation induite par ∂ sur $K'[T]^{\mathrm{pd}}_{<p^d}$ est ∂_D.*

Démonstration. Il suffit de vérifier que $\varphi(\lambda\otimes\partial(f)) = \partial_D(\varphi(\lambda\otimes f))$ pour tout $\lambda \in F'$ et tout $f \in K$. Posons $a_i = \lambda\partial^i(f)$ $(0 \le i \le p^d)$ et $\Phi = \varphi(\lambda\otimes f)$. Par définition, $\Phi = a_0 + a_1T + \cdots + a_{p^d-1}T^{[p^d-1]}$. D'autre part, un calcul donne :

$$\varphi(\lambda \otimes \partial(f)) = \lambda \cdot \sum_{0\le i<p^d} a_{i+1}T^{[i]} = \lambda \cdot \sum_{1\le i\le p^d} a_iT^{[i-1]} = \partial_T(\Phi) + \partial^{p^d}(f)T^{[p^d-1]}. \tag{53}$$

Or, par définition de $D{\cdot}\partial$, on a $\partial^{p^d} = D{\cdot}\partial - \big(c_0\partial + c_1\partial^p + \cdots + c_{d-1}\partial^{p^{d-1}}\big)$, d'où on déduit :

$$\partial^{p^d}(f) = (D{\cdot}\partial)(a_0) - \big(c_0a_1 + c_1a_p + c_2a_{p^2} + \cdots + c_{d-1}a_{p^{d-1}}\big). \tag{54}$$

La proposition suit en mettant ensemble (53) et (54). □

L'isomorphisme φ vérifie une relation supplémentaire de compatibilité à la réduction modulo T. Précisément, remarquons que l'application $F' \otimes_F K \to K'$, $\lambda \otimes f \mapsto \lambda f$ est bien définie. On dispose, par ailleurs, d'un morphisme naturel $\pi : K'[T]^{\mathrm{pd}}_{<p^d} \to K'$ qui consiste à envoyer un polynôme à puissances divisées sur son terme constant. On vérifie alors sans mal que le diagramme suivant est commutatif :

$$\begin{array}{ccc} F' \otimes_F K & \xrightarrow{\varphi} & K'[T]^{\mathrm{pd}}_{<p^d} \\ {\scriptstyle \lambda\otimes f\mapsto\lambda f}\downarrow & & \downarrow{\scriptstyle \pi} \\ K' & = & K' \end{array}$$

On dispose en outre d'une section évidente $\sigma : K' \to K'[T]^{\mathrm{pd}}_{<p^d}$. Rappelons-nous également que K' est muni de la dérivation $D{\cdot}\partial$. On s'attend à ce que celle-ci se compare à la dérivation $D{\cdot}\partial_D$ agissant sur $K'[T]^{\mathrm{pd}}_{<p^d}$. C'est en effet le cas comme le précise la proposition suivante.

Proposition A.3.7. *Les deux diagrammes suivants :*

$$\begin{array}{ccc} K'[T]^{\mathrm{pd}}_{<p^d} & \xrightarrow{D\cdot\partial_D} & K'[T]^{\mathrm{pd}}_{<p^d} \\ {\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\ K' & \xrightarrow{D\cdot\partial} & K' \end{array} \qquad\qquad \begin{array}{ccc} K'[T]^{\mathrm{pd}}_{<p^d} & \xrightarrow{D\cdot\partial_D} & K'[T]^{\mathrm{pd}}_{<p^d} \\ {\scriptstyle \sigma}\uparrow & & \uparrow{\scriptstyle \sigma} \\ K' & \xrightarrow{D\cdot\partial} & K' \end{array}$$

sont commutatifs.

Remarque A.3.8. On peut reformuler la proposition ci-dessus en disant que π et σ sont des morphismes différentiels de $(K'[T]^{\mathrm{pd}}_{<p^d}, D{\cdot}\partial_D)$ dans $(K', D{\cdot}\partial)$ et de $(K', D{\cdot}\partial)$ dans $(K'[T]^{\mathrm{pd}}_{<p^d}, D{\cdot}\partial_D)$ respectivement.

Démonstration de la proposition A.3.7. Définissons :

$$\partial_{D,0} = \partial_T - (c_0\partial_0 + c_1\partial_1 + \cdots + c_{d-1}\partial_{d-1})$$

de manière à ce que $\partial_D = \partial_{D,0} + \partial'$. La dérivation $\partial_{D,0}$ est K'-linéaire et sa matrice dans la base $(1, T, T^{[2]}, \ldots, T^{[p^d-1]})$ est la transposée de la matrice compagnon associée au polynôme D_{lin}. On en déduit que $D{\cdot}\partial_{D,0} = 0$. Par ailleurs, remarquons que

$$\forall k \in \{0, 1, \ldots, p^d{-}2\}, \quad \partial' \circ \partial^k_{D,0} \circ \partial' = 0. \tag{55}$$

En effet, pour $\Phi \in K'[T]^{\mathrm{pd}}_{<p^d}$, il suit de la définition que $\partial'(\Phi)$ est un monôme en $T^{[p^d-1]}$. Du fait que $\partial_{D,0}$ diminue la « valuation T-adique » d'au plus 1, on déduit que le coefficient constant de $\partial^k_{D,0} \circ \partial'(\Phi)$ est nul. Ainsi son image par ∂' s'annule également. À partir de la relation (55), on déduit que, pour tout entier $n \leq p^d$:

$$\partial^n_D = \partial^n_{D,0} + \sum_{k=0}^{n-1} \partial^k_{D,0} \circ \partial' \circ \partial^{n-1-k}_{D,0}$$

puisque les termes faisant intervenir au moins deux occurrences de ∂' s'annulent tous. De plus, lorsque $k < p^d$, on a $\partial^k_{D,0} \circ \partial' \circ \partial^{n-1-k}_{D,0}(\Phi) = (D{\cdot}\partial)(a_{n-1-k})\, T^{[p^d-1-k]}$ si les a_i sont les coefficients de Φ. On en déduit que :

$$\partial^n_D(\Phi) = \partial^n_{D,0}(\Phi) + \sum_{k=0}^{n-1} (D{\cdot}\partial)(a_{n-1-k})\, T^{[p^d-1-k]}$$

et enfin, puisque $D{\cdot}\partial_{D,0}$ s'annule, on obtient en sommant :

$$(D{\cdot}\partial_D)(\Phi) = \sum_{i=0}^{d} \sum_{k=0}^{p^i-1} c_i \cdot (D{\cdot}\partial)(a_{p^i-1-k})\, T^{[p^d-1-k]} \tag{56}$$

où les c_i sont les coefficients de D, *i.e.* $D = c_0 + c_1 Y + \cdots + c_d Y^d$. Sur la formule (56), on observe en particulier que la seule contribution au coefficient constant de $(D{\cdot}\partial_D)(\Phi)$ est obtenue pour $i = d$ et $k = p^d - 1$. Ainsi le coefficient constant de $(D{\cdot}\partial_D)(\Phi)$ est $(D{\cdot}\partial)(a_0)$. On en déduit la commutativité du premier diagramme.

Pour ce qui concerne le second diagramme, il s'agit de démontrer que $(D{\cdot}\partial_D)(\Phi) = (D{\cdot}\partial)(a_0)$ lorsque $\Phi = a_0 \in K'$, *i.e.* lorsque $a_1 = a_2 = \cdots = a_{p^d-1} = 0$. Ceci résulte à nouveau directement de la formule (56). □

Références

[1] Gorô Azumaya, *On maximally central algebras*, Nagoya Math. J. **2** (1951), 119–150.

[2] Saugata Basu, Richard Pollack, Marie-Françoise Roy *Algorithms in Real Algebraic Geometry*, Springer-Verlag (2008), second edition

[3] M. Berkovich, G. Tsirulik, *Differential resultants and some of their applications*, Differential'nye Uravneniya **22** (1986), 750–757

[4] André Blanchard, *Les corps non commutatifs*, Presses Universitaires de France (PUF)

[5] , Marc Chardin, *Differential resultants and subresultants*, Proceddings of Fundamentals of Computation Theory, Lecture Notes in Computer Science **529** (1991), 180–189

[6] Paul Cohn, *Skew field constructions*, London Mathematical Society Lecture Notes Series **27**, Cambridge University Press

[7] G. Collins, *Subresultant and Reduced Polynomial Remainder Sequences*, Journal of ACM **16** (1967), 708–712

[8] James Milne, *Class field theory*, disponible à http://www.jmilne.org/math/CourseNotes/CFT.pdf

[9] Ernst Gabidulin, *Theory of codes with maximum rank distance*, Problemy Peredachi Informatsii **21** (1985), no. 1, 3–16.

[10] Philippe Gille, Tamás Szamuely, *Central simple algebras and Galois cohomology*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[11] Alexander Grothendieck, *Le groupe de Brauer. I. Algèbres d'Azumaya et interprétations diverses*, Séminaire Bourbaki **9**, Soc. Math. France, Paris, 1995, pp. Exp. No. 290, 199–219.

[12] Shûichi Ikehata, *Azumaya algebras and skew polynomial rings*, Math. J. Okayama Univ. **23** (1981), no. 1, 19–32.

[13] Shûichi Ikehata, *Azumaya algebras and skew polynomial rings. II*, Math. J. Okayama Univ. **26** (1984), 49–57.

[14] Nathan Jacobson, *Non commutative polynomials and cyclic algebras*, Ann. of Maths **35** (1934), 197–208

[15] Nathan Jacobson, *Pseudo linear transformations*, Ann. of Math. **38** (1937), 484–506

[16] Nathan Jacobson, *Abstract derivation and Lie algebras*, Trans. AMS **43** (1937), 206–224

[17] Nathan Jacobson, *An extension of Galois theory to non-normal and non-separable fields*, Amer. J. of Math. **66** (1944), 1–29

[18] Nathan Jacobson, *Galois theory of purely inseparable fields of exponent one*, Amer. J. of Math. **66** (1944), 645–648

[19] Nathan Jacobson, *Finite-dimensional division algebras over fields*, Springer-Verlag, Berlin, 1996.

[20] Ziming Li, *A subresultant theory for linear differential, linear difference and Ore polynomials with applications*, PhD Thesis (1996)

[21] Ziming Li, *A subresultant theory for Ore polynomials and applications*, proceedings ISSAC'98

[22] Øystein Ore, *Theory of non-commutative polynomials*, Ann. of Math. **34** (1933), no. 3, 480–508.

[23] Philippe Revoy, *Algèbres de Weyl en caractéristique* p, C. R. Acad. Sci. Paris Sér. A-B **276** (1973), A225–A228.

[24] André Weil, *Basic number theory*, Classics in Mathematics, Springer-Verlag, Berlin, 1995.

[25] F. Winkler, *Polynomial Algorithms in Computer Algebra*, Springer Wien New Work (1996)