

# Géométrie affine des espaces de matrices, et questions de décompositions quadratiques

Habilitation à diriger des recherches  
de l'Université Paris-Saclay

présentée et soutenue à Versailles, le 6 juin 2023, par

**Clément de SEGUINS PAZZIS**

## Composition du jury

**Grégory BERHUY**

Professeur des Universités,  
Université Grenoble-Alpes

Rapporteur

**Matej BREŠAR**

Professor of Mathematics,  
University of Ljubljana

Rapporteur

**Alexander GUTERMAN**

Professor of Mathematics,  
Bar-Ilan University

Rapporteur

**Pierre BAUMANN**

Chargé de Recherches,  
Université de Strasbourg

Examineur

**Philippe CALDERO**

Maître de Conférences,  
Université Claude Bernard Lyon 1

Examineur

**Bernhard KELLER**

Professeur des Universités,  
Université Paris Cité

Président

**Vincent SÉCHERRE**

Professeur des Universités,  
UVSQ

Examineur

**Dimitri ZVONKINE**

Directeur de Recherche,  
UVSQ

Examineur



# Avant-propos

*We [he and Halmos] share a philosophy about linear algebra : we think basis-free, we write basis-free, but when the chips are down we close the office door and compute with matrices like fury.*

*Nous [Paul Halmos et lui] avons la même philosophie envers l'algèbre linéaire : nous pensons sans référence aux bases, écrivons nos articles sans référence aux bases, mais lorsque les dés sont jetés nous fermons la porte de notre bureau et pratiquons le calcul matriciel frénétiquement.*

Irving Kaplansky

Le présent manuscrit a plusieurs objectifs. Le premier est bien entendu de convaincre le lecteur de ma capacité à diriger les recherches d'un potentiel étudiant en thèse, en démontrant que j'ai su être pour moi-même un guide avisé. Le deuxième, de fournir un panorama de travaux récents et significatifs en algèbre linéaire et quadratique classique. Le dernier, non des moindres, est de convaincre le lecteur que ce domaine possède un passé récent digne d'intérêt voire un avenir radieux.

Il serait étrange de ne pas souligner le caractère singulier du contexte dans lequel ont été conçues et écrites les près de 1500 pages de production mathématique que recouvre ce manuscrit, et dont seulement cinquante sont le fruit d'une collaboration avec un pair. Ces travaux sont en effet sans rapport aucun avec ma thèse de doctorat sur la K-théorie équivariante, ni même la topologie algébrique dont j'ai fait ma spécialité de DEA. Ils trouvent au contraire leurs racines dans des recherches personnelles que j'ai commencé à mener à partir de l'année 2006, alors que j'étais pleinement rentré dans ma carrière de professeur de Taupe ; et tous ces résultats ont été découverts, rédigés et publiés tout en conservant ce statut d'enseignant à plein temps. Cela ne m'a pas empêché d'avoir un pied dans le monde universitaire, via l'Université de Versailles Saint-Quentin-en-Yvelines qui m'a rapidement accueilli comme associé au Laboratoire de Mathématiques de Versailles, les nombreuses invitations à des conférences et séminaires, en France comme à l'étranger, et ma participation aux tâches d'évaluation et de recension de manuscrits.

D'aucuns me reprochent parfois de raffoler des détails et de manquer d'esprit de synthèse. Et les presque deux cents pages de ce mémoire pourraient leur donner raison si elles n'avaient pour ambition de condenser et commenter une si grande quantité de travaux. Mais je suivrai ici leurs conseils et vais donc clore immédiatement cet avant-propos par le passage obligé – quoique fort agréable – des remerciements.

Mes premiers remerciements vont à Rached Mneimné, qui m'a un peu forcé la main pour écrire mon ouvrage sur les formes quadratiques et a eu bien raison de le faire. L'expertise que j'ai acquise à cette occasion n'a cessé de m'aider lors de mes travaux ultérieurs. Lorsque j'ai commencé ma carrière, Rached était connu en France comme « Monsieur Matrices ». Aujourd'hui, il ne me déplairait pas que l'on fasse ainsi référence à moi même si ce titre me paraîtrait réducteur.

Mon prédécesseur, Saab Abou-Jaoudé, pourra aussi mesurer sa propre contribution à mon œuvre. Non seulement, quelques-uns de mes travaux sont fondés sur l'une de ses idées (voir le chapitre 2), mais les nombreux échanges que j'ai eus avec lui lors de la période 2006-2010 m'ont sans cesse poussé à progresser. C'est Saab qui a insisté pour que je publie mes découvertes, d'abord à la RMS (Revue de la Filière Mathématique) puis dans des revues internationales à comité de lecture. J'espère qu'en lisant ce manuscrit il ressentira une forme de fierté.

Du côté universitaire, je remercie Vincent Sécherre, Ariane Mézard et Yvan Martel, qui m'ont proposé cette association avec le LMV qui dure depuis 2013 (de fait sinon de droit), et surtout Peter Šemrl, qui m'a invité de nombreuses fois à Ljubljana pour y présenter mes travaux, m'a introduit dans la communauté internationale de l'Algèbre linéaire et n'a cessé de me prodiguer des conseils avisés sur la façon dont je pourrais faire évoluer mes recherches.

À titres divers, je voudrais remercier Bernard Randé, qui a supervisé mes premiers articles publiés à la RMS, Nicolas Tosel, qui m'a permis très tôt de prendre conscience de mon potentiel, François Moulin, qui fut un collaborateur impeccable depuis 2009, et enfin Jean Nougayrède qui a régulièrement relu des parties de mes articles les plus importants avant qu'ils soient proposés aux revues.

Je tiens à remercier Grégory Berhuy, qui m'a fait l'honneur de prendre un temps considérable pour prendre connaissance de mes travaux et réaliser un pré-rapport pour la commission de l'Université Paris-Saclay. Et j'associe bien évidemment à ces remerciements les autres rapporteurs et membres du jury : Matej Brešar, Alexander Guterman, Pierre Baumann, Philippe Caldero, Bernhard Keller, Vincent Sécherre et Dimitri Zvonkine. J'espère que tous auront trouvé de bien belles mathématiques en récompense du temps qu'ils ont consacré à l'examen de ce mémoire.

Mes derniers remerciements iront naturellement à mon épouse Stéphanie, qui a consenti à ce que je fasse autant de recherche mathématique sur mon temps libre.

A Versailles, décembre 2022.

# Table des matières

<b>1</b>	<b>Tour d’horizon</b>	<b>10</b>
1.1	De l’esprit . . . . .	10
1.2	Du fond . . . . .	11
1.2.1	Géométrie affine des espaces de matrices . . . . .	11
1.2.2	Décompositions en objets quadratiques . . . . .	16
1.3	De la forme . . . . .	17
1.4	Des notations et conventions . . . . .	19
<b>2</b>	<b>Espaces de matrices de rang majoré</b>	<b>21</b>
2.1	Articles concernés . . . . .	21
2.2	Bibliographie historique . . . . .	21
2.3	Problématique et historique . . . . .	22
2.4	Historique de mes travaux personnels . . . . .	25
2.5	Sous-espaces de grande dimension . . . . .	27
2.5.1	De Dieudonné à Atkinson-Lloyd . . . . .	27
2.5.2	La méthode d’Abou-Jaoudé . . . . .	29
2.5.3	La méthode itérative et ses fruits . . . . .	31
2.6	Sous-espaces de grande dimension : cas des matrices symétriques/alternées	35
2.7	Sous-espaces vectoriels primitifs de rang majoré . . . . .	38
2.8	Perspectives . . . . .	41
<b>3</b>	<b>Dépendance linéaire locale et réflexivité algébrique</b>	<b>43</b>
3.1	Articles concernés . . . . .	43
3.2	Bibliographie historique . . . . .	43
3.3	Problématique et histoire . . . . .	44
3.3.1	Principales définitions . . . . .	44
3.3.2	Motivation et histoire des principaux résultats . . . . .	45
3.3.3	Historique de mes travaux . . . . .	46
3.4	La dualité opérateur-vecteur . . . . .	48
3.5	Application de la dualité opérateur-vecteur aux problématiques de rang . . . . .	49
3.6	Application de la dualité opérateur-vecteur à la détermination des espaces localement liés minimaux . . . . .	51
3.7	Les algèbres à division à gauche bilinéarisable . . . . .	53
3.7.1	Du problème du rang minimal aux algèbres à division LDB	53
3.7.2	Les algèbres à division LDB pour elles-mêmes (début) . .	54
3.7.3	Les algèbres à division LDB pour elles-mêmes (suite) . . .	57
3.7.4	Les algèbres à division LDB pour elles-mêmes (fin) . . . .	58

3.8	Réflexivité algébrique et rang minimal sur un corps fini . . . . .	60
3.9	Perspectives . . . . .	60
<b>4</b>	<b>Morphismes Im-compatibles</b>	<b>61</b>
4.1	Articles concernés . . . . .	61
4.2	Fonctions Im-compatibles . . . . .	61
4.3	Éléments de motivation . . . . .	63
4.4	Principaux résultats . . . . .	64
4.5	Méthodes . . . . .	65
4.6	Extensions, état des lieux . . . . .	67
4.7	Bilan et perspectives . . . . .	68
<b>5</b>	<b>Préserveurs</b>	<b>69</b>
5.1	Articles concernés . . . . .	69
5.2	Bibliographie historique . . . . .	69
5.3	Introduction . . . . .	70
5.4	Application linéaires non bijectives conservant l'inversibilité [1] .	73
5.5	Bijections affines conservant l'inversibilité [2] . . . . .	76
5.6	Injectons linéaires conservant l'inversibilité à partir d'un grand espace de matrices [3] . . . . .	78
5.7	Applications linéaires conservant une fonction de Schur [6] . . . .	81
5.8	Sur le théorème fondamental de la géométrie des matrices rec- tangulaires [4] . . . . .	85
5.9	Transformations continues de l'espace-temps [5] . . . . .	89
<b>6</b>	<b>Espaces de matrices à spectre limité</b>	<b>93</b>
6.1	Articles concernés . . . . .	93
6.2	Bibliographie historique . . . . .	93
6.3	Panorama historique . . . . .	94
6.4	Historique de mes travaux . . . . .	97
6.5	Sous-espaces de matrices à spectre trivial . . . . .	99
6.5.1	Des sous-espaces affines de rang minoré aux sous-espaces vectoriels à spectre trivial . . . . .	99
6.5.2	Majoration de la dimension d'un sous-espace à spectre trivial : la méthode combinatoire . . . . .	100
6.5.3	La méthode de compatibilité diagonale . . . . .	102
6.5.4	Sous-espaces vectoriels à spectre trivial de dimension maxi- male . . . . .	105
6.5.5	Applications aux sous-espaces affines de rang minoré . . . .	108
6.5.6	Sous-espaces nilpotents : le cas d'un corps gauche . . . . .	110
6.5.7	Sous-espaces à spectre de cardinal limité . . . . .	110
6.6	Lien avec les sous-espaces primitifs de matrices de rang majoré .	115
6.7	Le problème de Gerstenhaber structuré . . . . .	117
6.7.1	Introduction et réduction au cas non dégénéré . . . . .	117
6.7.2	Les principaux résultats . . . . .	118
6.7.3	Les méthodes en caractéristique différente de 2 . . . . .	121
6.7.4	Les méthodes en caractéristique 2 . . . . .	125
6.8	Bilan et perspectives . . . . .	126

<b>7</b>	<b>Espaces de matrices diagonalisables</b>	<b>128</b>
7.1	Articles concernés . . . . .	128
7.2	Bibliographie historique . . . . .	128
7.3	Problématique et historique . . . . .	129
7.4	Le principal théorème . . . . .	129
7.5	Lien avec les sous-espaces à spectre trivial . . . . .	131
7.6	Application à la préservation de la diagonalisabilité . . . . .	133
7.7	Perspectives . . . . .	134
<b>8</b>	<b>Décompositions à l'aide d'objets quadratiques</b>	<b>135</b>
8.1	Articles concernés . . . . .	135
8.2	Bibliographie historique . . . . .	136
8.3	Problématique . . . . .	137
8.4	Historique de mes travaux . . . . .	138
8.5	Le problème à deux termes . . . . .	141
8.5.1	Énoncé du problème . . . . .	141
8.5.2	La dichotomie exceptionnel-régulier . . . . .	141
8.5.3	Intervention des algèbres de quaternions pour l'étude des ( $p, q$ )-sommes régulières . . . . .	144
8.5.4	Sur le cas des endomorphismes exceptionnels . . . . .	146
8.5.5	Le problème des produits . . . . .	150
8.6	Le problème à trois termes en dimension finie . . . . .	152
8.6.1	Décompositions en sommes ou combinaisons linéaires à trois termes . . . . .	152
8.6.2	Décompositions en produits à trois termes . . . . .	159
8.7	Sommes d'idempotents en caractéristique non nulle . . . . .	161
8.8	Décompositions en dimension infinie . . . . .	164
8.8.1	Décomposition en sommes de quatre termes . . . . .	164
8.8.2	Décompositions à trois termes . . . . .	167
8.8.3	Cas d'un endomorphisme à valeur propre dominante . . . . .	169
8.8.4	Décompositions en produits . . . . .	170
8.8.5	Perspectives . . . . .	173
8.9	Perspectives : le problème bi-quadratique . . . . .	176
<b>9</b>	<b>Travaux inclassables</b>	<b>179</b>
9.1	Une démonstration originale du lemme de Noether-Deuring [1] . . . . .	179
9.1.1	La démonstration classique . . . . .	180
9.1.2	Les grandes lignes de ma démonstration du lemme de Noether-Deuring . . . . .	181
9.1.3	Épilogue . . . . .	182
9.2	Sur les sous-algèbres de matrices carrées à commutant trivial [2] . . . . .	182
9.3	Sur l'instabilité par multiplication des grands sous-espaces de ma- trices carrées [3] . . . . .	184
9.4	Sur la classe de similitude d'une matrice d'un groupe orthogonal ou symplectique en caractéristique 2 [4] . . . . .	186
9.5	Sur le sous-monoïde engendré par la classe de similitude d'une matrice singulière [5] . . . . .	190
9.6	Sur l'exponentielle de matrice [6] . . . . .	192
9.7	Sur l'écriture d'une matrice de trace nulle comme commutateur [7] . . . . .	194

9.8 Sur la topologie de l'ensemble des racines  $p$ -ièmes d'une matrice  
nilpotente complexe [8] . . . . . 197



# Bibliographie

- [1] C. de Seguins Pazzis, The singular linear preservers of non-singular matrices. *Linear Algebra Appl.* **433-2** (2010) 483–490.
- [2] C. de Seguins Pazzis, Invariance of simultaneous similarity and equivalence of matrices under extension of the ground field. *Linear Algebra Appl.* **433-3** (2010) 618–624.
- [3] C. de Seguins Pazzis, On linear combinations of two idempotent matrices over an arbitrary field. *Linear Algebra Appl.* **433-3** (2010) 625–636.
- [4] C. de Seguins Pazzis, On decomposing any matrix as a linear combination of three idempotents. *Linear Algebra Appl.* **433-4** (2010) 843–855.
- [5] C. de Seguins Pazzis, On sums of idempotent matrices over a field of positive characteristic. *Linear Algebra Appl.* **433-4** (2010) 856–866.
- [6] C. de Seguins Pazzis, The affine preservers of non-singular matrices. *Arch. Math.* **95** (2010) 333–342.
- [7] C. de Seguins Pazzis, On the matrices of given rank in a large subspace. *Linear Algebra Appl.* **435-1** (2011) 147–151.
- [8] B. Randé, C. de Seguins Pazzis, The linear preservers of real diagonalisable matrices. *Linear Algebra Appl.* **435-6** (2011) 1257–1266.
- [9] C. de Seguins Pazzis, On small matrix subalgebras with a trivial centralizer. *Linear Multilinear Algebra* **59-7** (2011) 761–771.
- [10] C. de Seguins Pazzis, To what extent is a large space of matrices not closed under product? *Linear Algebra Appl.* **435-11** (2011) 2708–2721.
- [11] C. de Seguins Pazzis, When does a linear map belong to at least one orthogonal or symplectic group? *Linear Algebra Appl.* **436-5** (2012) 1385–1405.
- [12] C. de Seguins Pazzis, Large affine spaces of matrices with rank bounded below. *Linear Algebra Appl.* **437-2** (2012) 499–518.
- [13] C. de Seguins Pazzis, The linear preservers of non-singularity in a large space of matrices. *Linear Algebra Appl.* **436** (2012) 3507–3530.
- [14] C. de Seguins Pazzis, Spaces of matrices with a sole eigenvalue. *Linear Multilinear Algebra.* **60-10** (2012) 1165–1190.
- [15] C. de Seguins Pazzis, Sums of two triangularizable quadratic matrices over an arbitrary field. *Linear Algebra Appl.* **436-9** (2012) 3293–3302.
- [16] C. de Seguins Pazzis, The semigroup generated by the similarity class of a singular matrix. *Comm. Algebra.* **40-9** (2012) 3414–3419.
- [17] C. de Seguins Pazzis, On commuting matrices and exponentials. *Proc. Amer. Math. Soc.* **141** (2013) 763–774.

- [18] Clément de Seguins Pazzis, Large affine spaces of non-singular matrices. *Trans. Amer. Math. Soc.* **365** (2013) 2569–2596.
- [19] C. de Seguins Pazzis, On Gerstenhaber’s theorem for spaces of nilpotent matrices over a skew field. *Linear Algebra Appl.* **438-11** (2013) 4426–4438.
- [20] C. de Seguins Pazzis, On the minimal rank in non-reflexive operator spaces over finite fields. *Linear Algebra Appl.* **462** (2014) 126–132.
- [21] C. de Seguins Pazzis, Commutators from a hyperplane of matrices. *Electronic J. Linear Algebra* **27** (2014) 39–54.
- [22] C. de Seguins Pazzis, From primitive spaces of bounded rank matrices to a generalized Gerstenhaber theorem. *Quart. J. Math.* **65-2** (2014) 319–325.
- [23] C. de Seguins Pazzis, Local linear dependence seen through duality II. *Linear Algebra Appl.* **462** (2014) 133–185.
- [24] C. de Seguins Pazzis, Spaces of matrices with few eigenvalues. *Linear Algebra Appl.* **449** (2014) 210–311.
- [25] C. de Seguins Pazzis, P. Šemrl, Hua’s fundamental theorem of geometry of rectangular matrices over EAS division rings. *J. Algebra* **439** (2015) 159–187.
- [26] C. de Seguins Pazzis, Local linear dependence seen through duality I. *J. Pure Appl. Algebra* **219** (2015) 2144–2188.
- [27] C. de Seguins Pazzis, LDB division algebras. *J. Algebra* **423** (2015) 259–288.
- [28] C. de Seguins Pazzis, Range-compatible homomorphisms on matrix spaces. *Linear Algebra Appl.* **484** (2015) 237–289.
- [29] C. de Seguins Pazzis, The classification of large spaces of matrices with bounded rank. *Israel J. Math.* **208** (2015) 219–259.
- [30] C. de Seguins Pazzis, Affine spaces of symmetric or alternating matrices with bounded rank. *Linear Algebra Appl.* **504** (2016) 503–558.
- [31] C. de Seguins Pazzis, Large spaces of symmetric or alternating matrices with bounded rank. *Linear Algebra Appl.* **508** (2016) 146–189.
- [32] C. de Seguins Pazzis, Large spaces of bounded rank matrices revisited. *Linear Algebra Appl.* **504** (2016) 124–189
- [33] C. de Seguins Pazzis, Lines of full rank matrices in large subspaces. *Oper. Matrices* **10** (2016) 985–996.
- [34] C. de Seguins Pazzis, Primitive spaces of matrices with upper rank two over the field with two elements. *Linear Multilinear Algebra* **64** (2016) 1321–1353.
- [35] C. de Seguins Pazzis, Quasi-range-compatible affine maps on large operator spaces. *Linear Multilinear Algebra* **64** (2016) 1056–1085.
- [36] C. de Seguins Pazzis, Range-compatible homomorphisms on spaces of symmetric or alternating matrices. *Linear Algebra Appl.* **503** (2016) 135–163.
- [37] C. de Seguins Pazzis, The Flanders theorem over division rings. *Linear Algebra Appl.* **493** (2016) 313–322.
- [38] C. de Seguins Pazzis, A note on sums of three square zero matrices. *Linear Multilinear Algebra* **65** (2017) 787–805.
- [39] C. de Seguins Pazzis, Sums of three quadratic endomorphisms of an infinite-dimensional vector space. *Acta Sci. Math. (Szeged)* **83** (2017) 83–111.

- [40] C. de Seguins Pazzis, P. Šemrl, Continuous space-time transformations. *Adv. Geometry* **18-3** (2018) 385–393.
- [41] C. de Seguins Pazzis, From trivial spectrum subspaces to spaces of diagonalisable real matrices. *Math. Proc. R. Ir. Acad.* **118-A** (2018) 5–8.
- [42] C. de Seguins Pazzis, Sums of quadratic endomorphisms of an infinite-dimensional vector space. *Proc. Edinburgh Math. Soc.* **61** (2018) 437–447.
- [43] C. de Seguins Pazzis, Range-compatible homomorphisms over the field with two elements. *Electronic Linear Algebra* **34** (2018) 71–114.
- [44] C. de Seguins Pazzis, On the linear preservers of Schur matrix functionals. *Linear Algebra Appl.* **567** (2019) 63–117.
- [45] C. de Seguins Pazzis, The structured Gerstenhaber problem I. *Linear Algebra Appl.* **567** (2019) 263–298.
- [46] C. de Seguins Pazzis, The structured Gerstenhaber problem II. *Linear Algebra Appl.* **569** (2019) 113–145.
- [47] C. de Seguins Pazzis, The space of all  $p$ -th roots of a nilpotent complex matrix is path-connected. *Linear Algebra Appl.* **596** (2020) 106–116.
- [48] C. de Seguins Pazzis, The structured Gerstenhaber problem III. *Linear Algebra Appl.* **601** (2020) 134–169.
- [49] C. de Seguins Pazzis, Products of involutions in the stable general linear group. *J. Algebra* **530** (2019) 235–289.
- [50] C. de Seguins Pazzis, Products of involutions of an infinite-dimensional vector space. *Canadian J. Math.* **73** (2021) 195–220.
- [51] C. de Seguins Pazzis, Sums and products of pairs of quadratic endomorphisms of countable-dimensional vector spaces. *Linear Multilinear Algebra*, accepté, en attente de publication
- [52] C. de Seguins Pazzis, The sum and the product of two quadratic matrices : regular cases. *Adv. Appl. Clifford Algebras* **32** (2022).
- [53] C. de Seguins Pazzis, The sum of two quadratic matrices : Exceptional cases. *Linear Algebra Appl.* **653** (2022) 357–394
- [54] C. de Seguins Pazzis, The product of two invertible quadratic matrices : Exceptional cases. *Linear Algebra Appl.* **662** (2023) 67–109.

# Chapitre 1

## Tour d'horizon

### 1.1 De l'esprit

Avant de rentrer pleinement dans le vif du sujet, je souhaite évoquer l'esprit dans lequel j'ai initié puis mené les travaux dont il est question ici.

Lors de ma thèse, je pense avoir démontré une capacité à bâtir de grands échafaudages théoriques pour résoudre un problème ardu. Ce qui avait motivé mon retour à l'enseignement pur, c'était la difficulté de communiquer, au-delà d'un cercle très restreint, la beauté de ces constructions. J'ai ainsi pris le parfait contrepied en allant résolument vers une approche de résolution de problèmes portant sur des objets que tous les taupins connaissent. Je me suis laissé guidé par mon sens esthétique et le plaisir de démontrer des énoncés *a priori* difficiles, en essayant de trouver la parfaite combinaison entre un problème de formulation simple, un résultat élégant que l'on peut facilement conjecturer, et une résolution astucieuse nécessitant une certaine maestria technique.

Quant à l'algèbre linéaire et quadratique, je m'inscris résolument dans la tradition de Dieudonné : elle doit être faite sur un corps (commutatif) quelconque, les restrictions sur la nature du corps de base ne doivent être justifiées que par la nature profonde du problème qu'on se pose. Je me détache cependant de Dieudonné quant au style de rédaction, pour lequel j'ai cherché un meilleur équilibre entre les calculs matriciels explicites et le point de vue géométrique des applications linéaires. Dans la droite ligne de la citation de Kaplansky qui ouvre ce manuscrit, l'essentiel de ce que j'ai trouvé a été obtenu par des calculs matriciels purs et durs, et je n'ai pas systématiquement cherché à en effacer les traces par des reformulations purement géométriques. Je n'ai adopté le point de vue géométrique que lorsqu'il permettait une bien meilleure compréhension, par exemple dans mes travaux sur les morphismes Im-compatibles.

Lorsqu'on aborde un domaine aussi apparemment battu et rebattu que l'algèbre linéaire, on s'imagine bien que les questions que l'on se pose ont déjà été abordées voire entièrement résolues par d'autres. Je ne me suis donc jamais étonné de systématiquement découvrir des précurseurs à mes travaux : ceux-ci sont très souvent des améliorations et prolongements de résultats défrichés par d'autres avant moi. Quelques travaux font exception à cette règle, notamment le thème des morphismes Im-compatibles et la découverte des algèbres à division à gauche bilinéarisable. Pour le reste, je me suis efforcé d'innover

techniquement, d'enrichir les connaissances sur des questions déjà abordées par d'autres, de les pousser parfois dans des directions inattendues, et enfin de tirer consciencieusement sur les fils qui se sont présentés à moi lorsqu'ils méritaient des développements.

Le lecteur verra une évolution commune dans mes travaux, où pour chaque grand thème le domaine d'étude passe progressivement de structures faibles à celles issues de l'algèbre quadratique : par exemple, le problème des décompositions en somme ou produit d'endomorphismes quadratiques, que j'avais initialement étudié dans les algèbres d'endomorphismes, est maintenant en cours d'investigation dans le cadre d'espaces munis de formes bilinéaires symétriques ou alternées non dégénérées.

Le lecteur pourra être étonné de ne pas me voir utiliser d'outil particulièrement sophistiqué. Une raison à cela réside dans la nature des problèmes posés : le souci de généralité maximale vis-à-vis du corps de base contraint souvent à des méthodes élémentaires, par exemple dans le problème des espaces de matrices ayant au plus  $k$  valeurs propres distinctes dans le corps de base (cette propriété ne définissant pas un sous-ensemble algébrique de l'espace de matrices envisagé).

Enfin, et après une forme de papillonnage initial dont trouvera une trace dans les œuvres inclassables regroupées dans le dernier chapitre, mes recherches se sont assez vite cristallisées autour de deux grands thèmes : la géométrie affine des espaces de matrices (ou d'opérateurs) d'une part, la problématique de la décomposition en objets quadratiques dans les algèbres classiques d'autre part. Ce sont ces deux thèmes principaux que je vais évoquer à grands traits dans la suite de cette introduction.

## 1.2 Du fond

Le présent mémoire porte sur des questions en algèbre linéaire classique. On travaille donc avec des espaces de matrices sur un corps, parfois des espaces d'applications linéaires en dimension quelconque. Nous pouvons ranger l'essentiel de nos travaux dans deux grandes catégories :

- la géométrie affine des espaces de matrices ;
- les questions de décomposition en objets quadratiques dans les algèbres de matrices et d'endomorphismes.

Aucune connexion profonde n'existant entre ces deux thèmes, je les aborderai séparément dans ce qui suit. Quant à la grosse poignée de travaux qui ne se rattachent pas directement à l'un de ces thèmes, je n'en dirai rien dans cette introduction et renvoie directement au chapitre 9, où sont évoqués ces œuvres inclassables.

### 1.2.1 Géométrie affine des espaces de matrices

Il s'agit d'examiner les structures affines dans les espaces de matrices ou d'applications linéaires vis-à-vis de certaines propriétés ou invariants remarquables. Le plus souvent, la propriété sélectionnée  $\mathcal{P}$  définit un cône  $\mathcal{C}$  de sommet 0 (éventuellement époiné) dans l'espace vectoriel  $\mathcal{M}$  de matrices ou d'opérateurs envisagés. Les questions naturelles que l'on se pose sont alors :

- celle de la structure des sous-espaces vectoriels ou affines inclus dans ce cône  $\mathcal{C}$  (ou plutôt, pour les sous-espaces vectoriels, ceux inclus dans

$\mathcal{C} \cup \{0\}$  afin de tenir compte de cônes époinés); c'est le problème des **structures linéaires dans les cônes remarquables**;

- des transformations affines de l'espace  $\mathcal{M}$  qui stabilisent le cône  $\mathcal{C}$  (soit  $f(\mathcal{C}) \subset \mathcal{C}$ ), voire le laissent globalement invariant ( $f(\mathcal{C}) = \mathcal{C}$ ).

La deuxième question relève essentiellement de la problématique des **préserveurs linéaires** (linear preserve problems, en anglais). Cette problématique a fait l'objet d'une quantité considérable d'articles au cours des quarante dernières années, et l'on constate que ce filon semble maintenant largement épuisé depuis une décennie. La plupart des problèmes classiques les plus intéressants (préservation du rang, de l'inversibilité, de la nilpotente etc) ont des solutions bien établies depuis longtemps, et il n'y a probablement plus de grande avancée à attendre dans cette direction. En conséquence, j'ai consacré très peu de mes travaux à ces questions de préservation linéaire (voir néanmoins le chapitre 5), pour réserver l'essentiel de mes recherches à celle des structures linéaires dans les cônes remarquables.

Historiquement, les questions de structures linéaires sont apparues soit comme objet d'étude à part entière (Gerstenhaber, pour les sous-espaces vectoriels du cône nilpotent) soit comme outil pour étudier des problèmes de préserveurs linéaires (Dieudonné, pour la préservation de l'inversibilité). Et, avant que je n'entreprenne mes travaux, l'étude des structures linéaires dans les cônes remarquables avait été plutôt délaissée au regard de tout ce qui avait été accompli sur les préserveurs linéaires. Je me suis donc trouvé dans la situation où beaucoup de résultats de base pouvaient être améliorés de maintes façons.

Pour rentrer plus avant dans la question des structures linéaires, il faut citer les deux questions fondamentales que l'on se pose lorsque l'on dispose d'un cône époiné  $\mathcal{C}$  de sommet 0 dans un espace vectoriel :

- quelle est la dimension maximale pour un sous-espace vectoriel (ou seulement affine) inclus dans  $\mathcal{C} \cup \{0\}$  ?
- quels sont les sous-espaces vectoriels maximaux inclus dans  $\mathcal{C} \cup \{0\}$  ?

Il faut impérativement dissiper une idée fautive que la théorie des formes quadratiques pourrait faire naître chez le lecteur. Dans un espace quadratique  $(E, q)$  en caractéristique différente de 2, on sait que tous les sous-espaces vectoriels maximaux inclus dans le cône nilpotent  $\mathcal{C}(q)$  ont la même dimension : cette dimension est l'indice de Witt de la forme  $q$ . Mais pour les cônes que nous allons considérer (cône des matrices singulières, des matrices nilpotentes, par exemple) rien de tel : en général les sous-espaces vectoriels maximaux inclus dans ces cônes n'ont pas tous la même dimension ! Par exemple, dans le cône des matrices singulières de  $M_3(\mathbb{K})$ , les trois sous-espaces vectoriels suivants sont maximaux bien qu'ils soient de dimensions respectives 6, 5 et 3 :

- celui formé des matrices de la forme  $\begin{bmatrix} ? & ? & 0 \\ ? & ? & 0 \\ ? & ? & 0 \end{bmatrix}$  ;
- celui formé des matrices de la forme  $\begin{bmatrix} ? & ? & ? \\ ? & 0 & 0 \\ ? & 0 & 0 \end{bmatrix}$  ;
- celui formé des matrices alternées, i.e. antisymétriques de diagonale nulle (précision indispensable en caractéristique 2).

La question renferme donc visiblement une grande richesse, et il est illusoire d'espérer obtenir une caractérisation générale des sous-espaces vectoriels maxi-

maux. On se cantonne donc à certaines questions plus simples, la plus classique étant la détermination des sous-espaces vectoriels (ou affines) inclus dans  $\mathcal{C} \cup \{0\}$  et de *dimension maximale*. C'est exactement l'approche de Dieudonné (1948) pour le cône des matrices singulières, de Flanders (1962) pour le cône des matrices rectangulaires de rang au plus  $r$ , de Gerstenhaber (1958) pour le cône des matrices nilpotentes, et de Randé (1992) pour le cône des matrices réelles diagonalisables. Toujours dans cet ordre d'idée, on peut aller plus loin en tentant de caractériser les sous-espaces vectoriels de dimension proche de la dimension maximale (travail initié par Atkinson et Lloyd en 1980 pour les espaces de matrices carrées de rang majoré par une constante).

D'autres approches sont aussi possibles, notamment la notion de primitivité introduite par Atkinson et Lloyd et redécouverte par Eisenbud et Harris, pour les sous-espaces de matrices de rang majoré.

Le problème des structures linéaires dans les cônes remarquables de matrices ou d'opérateurs a mobilisé tous types de techniques au fil des années. Des techniques de géométrie algébrique pour le cône des matrices de rang au moins  $r$  (et l'étude des sous-espaces vectoriels), mais aussi de topologie algébrique pour le cône des matrices de rang  $r$ . En général, l'usage de ces techniques puissantes limite l'étude à des corps réels ou algébriquement clos, parfois seulement infinis. À côté de cela, de nombreux résultats sont valables indépendamment du corps de base, et leur démonstration nécessite souvent des méthodes beaucoup plus élémentaires mais subtiles. La méthode de Flanders-Atkinson-Lloyd, pour étudier les sous-espaces vectoriels de matrices de rang majoré par une constante  $r$ , est fondée sur un argument assez élémentaire de géométrie différentielle sur des courbes (cela nécessite des corps ayant suffisamment d'éléments).

La méthode la plus fine est celle de projection/extraction des blocs, aussi appelée « décorticage de la dimension » (*dimension splicing* en anglais). Par exemple, la méthode mise en place par mon collègue Saab Abou-Jaoudé pour établir la généralisation du théorème de Flanders sur les sous-espaces de matrices de rang au plus  $r$  fonctionne comme suit : on se réduit à la situation où le sous-espace vectoriel  $\mathcal{V}$  en question contient la matrice  $J_r = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ , et on décompose selon le même format chaque matrice de  $\mathcal{V}$  sous la forme

$$M = \begin{bmatrix} A(M) & C(M) \\ B(M) & D(M) \end{bmatrix}.$$

Une approche naïve consisterait à écrire

$$\dim \mathcal{V} \leq \dim A(\mathcal{V}) + \dim B(\mathcal{V}) + \dim C(\mathcal{V}) + \dim D(\mathcal{V})$$

et à examiner si la condition de rang sur les matrices de  $\mathcal{V}$  produit des contraintes sur les blocs dans la décomposition. Mais cette approche est tout à fait insuffisante en général, et il faut la raffiner. On utilise donc par exemple le théorème du rang pour écrire

$$\dim \mathcal{V} = \dim A(\mathcal{V}) + \dim \mathcal{W}$$

où  $\mathcal{W}$  est l'ensemble des matrices  $M \in \mathcal{V}$  pour lesquelles  $A(M) = 0$ . Ces matrices sont plus simples et permettent des manipulations plus efficaces. Et on peut poursuivre : dans la méthode d'Abou-Jaoudé, on poursuit le décorticage

jusqu'au bout, introduisant successivement

$$\mathcal{W}' := \{M \in \mathcal{W} : B(M) = 0\} \quad \text{et} \quad \mathcal{W}'' := \{M \in \mathcal{W}' : C(M) = 0\}.$$

Le théorème du rang donne alors

$$\dim \mathcal{V} = \dim A(\mathcal{V}) + \dim B(\mathcal{W}) + \dim C(\mathcal{W}') + \dim D(\mathcal{W}'').$$

Et plus on effeuille l'espace  $\mathcal{V}$ , plus les matrices sont simples et susceptibles d'apporter des informations importantes en les manipulant convenablement.

Enfin, ces méthodes élémentaires font presque systématiquement usage de raisonnements par récurrence sur le format des matrices envisagées.

J'ai abordé les problématiques de structures linéaires en revenant aux deux groupes de théorèmes fondateurs : ceux de Dieudonné et Flanders pour les sous-espaces de matrices de rang majoré, celui de Gerstenhaber pour les sous-espaces vectoriels de matrices carrées nilpotentes. Dans chaque situation, j'ai essayé de repenser intégralement les méthodes, sans chercher d'emblée à comprendre les démonstrations historiques, ce qui m'a permis d'avoir une approche originale. Puis, j'ai orienté mes travaux autour de cette base avec diverses formes de généralisation ou d'adaptation :

- généralisation à des corps quelconques (voire des corps gauches) lorsque les théorèmes connus requéraient des hypothèses voulant que le corps de base soit de cardinal suffisamment grand ;
- généralisation à des sous-espaces vectoriels de dimension proche de la dimension critique (uniquement pour les problématiques de rang, à ce jour) ;
- adaptation à des contextes de structures quadratiques (cônes définis par le rang dans des espaces de matrices symétriques ou antisymétriques, cônes nilpotents dans des espaces vectoriels d'endomorphismes autoadjoints ou antiautoadjoints) ;
- variantes des hypothèses sur le spectre (majoration du nombre de valeurs propres dans le corps de base, ou dans une clôture algébrique).

Les variantes et généralisations m'ont presque toujours conduit à repenser les méthodes, si bien que j'ai découvert pas moins de trois nouvelles démonstrations profondément différentes du théorème de Gerstenhaber sur les sous-espaces nilpotents (sans parler des variantes pour plusieurs d'entre elles).

Compte tenu de mon souhait d'obtenir des résultats les plus généraux possibles, mes méthodes ont le plus souvent été très élémentaires : j'ai développé une maîtrise de la méthode de projection/extraction des blocs. Dans tous les cas étudiés, j'ai aussi mis en évidence l'importance de s'intéresser aux éléments de très petit rang dans les espaces de matrices envisagés. S'il fallait donner un nom à cette gamme de méthodes, j'adopterais volontiers la désignation de « combinatoire matricielle »<sup>1</sup> Et si je devais isoler un résultat emblématique, ce serait certainement ma généralisation du théorème de Gerstenhaber à la détermination des sous-espaces formés de matrices sans valeur propre non nulle et de dimension maximale, et de son application aux sous-espaces affines de matrices inversibles et de dimension maximale. La classification de ces derniers est, de manière remarquable, directement liée à celle des formes quadratiques

---

1. Attention, il ne s'agit pas là de s'appuyer sur des méthodes combinatoires pour aborder des problèmes d'espaces de matrices, à l'instar des travaux de Roy Meshulam.



anisotropes à similitude près sur le corps de base, et la démonstration est un de mes plus beaux tours de force [18].

Mes recherches sur les généralisations des théorèmes de Flanders et Dieudonné ont permis de dégager un champ d'étude à part entière, qui est celui des morphismes Im-compatibles, version duale du problème de la réflexivité algébrique. Mes méthodes pour déchiffrer les espaces de matrices de rang majoré et de grande dimension font apparaître de façon systématique la nécessité d'avoir des théorèmes décrivant explicitement les morphismes Im-compatibles, et j'ai développé des méthodes d'étude spécifiques de ces derniers.

Sans lien apparent avec les problématiques de géométrie affine des cônes de matrices, j'ai étudié la notion de liaison locale pour les familles d'opérateurs ou les espaces vectoriels d'opérateurs. Et j'ai refondé les techniques d'étude de ces derniers sur la dualité opérateur-vecteur, qui avait été entrevue par d'autres mais n'avait jamais été utilisée de manière systématique. Cette méthode m'a permis de raccrocher la plupart des résultats connus sur la liaison locale à des techniques bien établies sur les sous-espaces vectoriels de matrices de rang majoré. En particulier, j'ai mis en évidence une correspondance entre les espaces localement liés minimaux et les espaces semi-primitifs de rang majoré, et en ai profité pour approfondir les connaissances sur ces derniers en maîtrisant la méthode des matrices génériques d'Atkinson (où l'on travaille avec des matrices dont les coefficients sont des polynômes homogènes de degré 1).

Enfin, l'étude de l'optimalité de certains résultats connus antérieurement sur les espaces localement liés m'a fait découvrir une très intéressante variante du problème des algèbres de composition de Hurwitz, appelées algèbres à division bilinéarisable à gauche. Grâce à ma connaissance des méthodes cliffordiennes, j'ai entièrement élucidé leur structure.

## Des ponts

Les problèmes que j'ai affrontés en géométrie affine des espaces de matrices pourraient paraître disparates malgré une formulation similaire. Quoi de commun en effet entre les questions de rang, de nilpotence, ou de diagonalisabilité? Au fil du temps, et par souci d'approfondir de curieuses analogies, je suis parvenu à dresser des ponts entre des questions en apparence totalement différentes. Ces ponts sont systématiquement venus de l'application de deux méthodes de dualité :

- le plus souvent, la dualité opérateur-vecteur, où l'on fait agir les vecteurs sur les opérateurs : les propriétés d'un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$  se convertissent ainsi en des propriétés de son dual-opérateur  $\widehat{\mathcal{S}}$ , constitué des fonctions d'évaluation  $s \in \mathcal{S} \mapsto s(x) \in V$  ;
- plus exceptionnellement, la dualité trace, au sens de la forme bilinéaire  $(u, v) \in \mathcal{L}(U, V) \times \mathcal{L}(V, U) \mapsto \text{tr}(u \circ v)$  ; des propriétés d'un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$  se convertissent en des propriétés de son orthogonal  $\mathcal{S}^\perp \subset \mathcal{L}(V, U)$ .

Grâce à l'une ou l'autre de ces dualités, des connexions surprenantes se sont faites jour entre mes différents thèmes d'étude, et j'ai pu ainsi :

- traduire toutes les questions de liaison locale sur les espaces d'opérateurs en problèmes d'espaces d'opérateurs de rang majoré ; et par ce biais simplifier considérablement l'approche des premiers et obtenir de nombreux

résultats inédits ;

- mettre en évidence un lien inattendu entre la structure des sous-espaces vectoriels à spectre trivial et celle des espaces semi-primitifs de matrices de rang majoré, débouchant sur une démonstration extrêmement efficace du théorème principal de classification des premiers ; et ce lien est encore porteur d'un grand potentiel pour des études similaires dans le futur ;
- obtenir la structure des espaces vectoriels de matrices réelles diagonalisables de dimension maximale comme application quasi-immédiate, par dualité-trace, du théorème de structure des sous-espaces vectoriels à spectre trivial de dimension maximale ;
- lier la notion de morphisme Im-compatible à celle de réflexivité algébrique, obtenant ainsi de nombreux résultats inédits sur cette dernière ;
- mettre en évidence, pour les théorèmes les plus fins sur les morphismes Im-compatibles, un lien avec la structure des sous-espaces de matrices de rang très petit (via la dualité trace).

J'avais abordé les questions précédentes avec une approche de pure résolution de problèmes, et j'ai le sentiment qu'au fil des années tout cela a mué en une forme étrange de théorie dans laquelle les espaces semi-primitifs de matrices de rang majoré tiennent une place centrale.

### 1.2.2 Décompositions en objets quadratiques

Le deuxième champ que j'ai investi pleinement est celui des décompositions en objets quadratiques dans les algèbres classiques. Pour donner une motivation à cette étude, au-delà de la volonté de généraliser quelques résultats anciens, on peut revenir sur la problématique des parties génératrices dans les groupes classiques.

S'il fallait résumer le choix des parties génératrices "classiques" dans les groupes classiques, on pourrait présenter la chose en disant que les éléments de ces parties génératrices sont ceux du groupe qui sont, dans un certain sens, les plus proches possibles de l'élément neutre sans y être égaux pour autant. Ainsi :

- dans les groupes de permutation, les transpositions sont les permutations différentes de l'identité qui perturbent le plus faible nombre de points ;
- dans les groupes linéaires, les transvections et dilatations sont les automorphismes différents de l'identité dont l'espace des vecteurs fixes est de plus grande dimension possible ;
- dans les groupes orthogonaux, les réflexions sont les automorphismes orthogonaux différents de l'identité dont l'espace des vecteurs fixes est de plus grande dimension possible.

La philosophie des décompositions en somme ou produit d'objets quadratiques est proche de cet esprit. Un élément  $x$  d'une  $\mathbb{K}$ -algèbre  $\mathcal{A}$  est dit quadratique lorsque  $x^2$  est combinaison linéaire de  $1_{\mathcal{A}}$  et  $x$ , autrement dit lorsque  $x$  admet un polynôme annulateur de degré 2. L'idée est que les éléments scalaires (ceux de  $\mathbb{K}1_{\mathcal{A}}$ ) sont les éléments extrêmement simples, et que les éléments quadratiques sont ceux qui sont les plus proches de cette extrême simplicité au sens algébrique. Et, à partir de là, on se pose le même genre de questions que pour les parties génératrices classiques des groupes classiques, mais en imposant le type d'élément quadratique, au sens d'imposer un polynôme annulateur de degré 2. Par exemple, on cherche à savoir quels éléments sont sommes d'idempotents

(sans limitation sur le nombre de termes), produits d'idempotents, sommes d'un idempotent et d'un élément de carré nul, produits d'éléments de carré nul etc. Ces problèmes ont commencé à être étudiés à la fin des années 1960, avec une nette accélération au début des années 1990, puis un relatif oubli jusqu'à ce que je reprenne le flambeau.

En général, ces problèmes n'ont pas de solution intelligible sans rentrer dans les détails de l'algèbre considérée, et j'ai choisi de les examiner dans les algèbres d'endomorphismes en dimension finie mais aussi en dimension infinie. En dimension finie, les méthodes font la part belle à la réduction sous toutes ses formes (réduites de Frobenius, de Jordan) et vont jusqu'à utiliser des structures algébriques plus riches. En dimension infinie, ce sont des constructions fondées sur les ordinaux qui offrent les méthodes les plus efficaces.

L'essentiel de mon travail sur la question a porté sur les décompositions avec un tout petit nombre de termes : en dimension infinie, mes travaux portent sur les décompositions à trois ou quatre termes, exceptionnellement à deux. En dimension finie, mon grand œuvre est la résolution complète du problème à deux termes, pour les sommes comme pour les produits : j'ai unifié et largement prolongé des travaux disparates sur la question, et mis en évidence les structures quaternioniques cachées dans les résultats connus. Et je prolonge actuellement ces recherches sur le problème à deux termes dans le cadre d'algèbres à involution, ce qui fait appel à des invariants de conjugaison plus riches développés dans les années 1970 et donne enfin un véritable champ d'application pratique à ces derniers. Ce domaine est en plein essor, et l'exposé de mes résultats de la dernière décennie donne un bon panorama de sa vitalité.

Je me permettrai enfin de clore ce tour d'horizon par l'énoncé de quelques résultats emblématiques que j'ai obtenus :

- tout endomorphisme d'un espace vectoriel (de dimension finie ou non) est combinaison linéaire de trois projecteurs ;
- pour toute matrice carrée  $A$  de taille  $n$  et de déterminant  $\pm 1$ , la matrice augmentée  $A \oplus I_n$  est produit de trois matrices involutives ;
- sur un corps de caractéristique  $p > 0$ , toute matrice carrée de format assez grand et à trace dans le sous-corps premier est somme de cinq matrices idempotentes, et même de quatre si le corps de base est premier.

### 1.3 De la forme

Le reste du mémoire est organisé en huit chapitres, dont on peut dégager quatre sous-ensembles suffisamment distincts et d'inégal poids.

- (i) *Sous-espaces de matrices sous contrainte de rang.* Ce thème est abordé dans les chapitres 2 à 5. Le chapitre initial, consacré aux sous-espaces vectoriels de matrices de rang majoré, aborde la problématique de front. Le chapitre 3 évoque la notion de familles d'opérateurs localement liées, qui se raccroche au chapitre 2 par une astuce de dualité opérateur-vecteur. Le chapitre 4 concerne les morphismes Im-compatibles, notion apparue dans mes travaux sur la généralisation du théorème d'Atkinson-Lloyd, qui se rattache à la notion de réflexivité algébrique (évoquée aussi dans le chapitre 3) et s'est ensuite développée comme un objet d'étude à part entière depuis 2013. Cette partie du manuscrit se clot par l'évocation, dans le chapitre 5,

de six travaux sur les préserveurs, dont quatre sont en lien direct avec les problématiques développées dans le chapitre 2.

- (ii) *Sous-espaces de matrices carrées sous contrainte de réduction.* Les contraintes sur la nature et le nombre de valeurs propres font l'objet du long chapitre 6, tandis que les espaces vectoriels de matrices diagonalisables sont évoqués dans le chapitre 7.
- (iii) *Décompositions en objets quadratiques.* Le chapitre 8 y est entièrement consacré.
- (iv) *Travaux inclassables.* Le chapitre 9 regroupe huit articles, presque tous écrits dans mes toutes premières années de travail de chercheur, et qui ne s'inscrivent pas bien l'une des trois autres catégories.

Chaque chapitre est en principe lisible de manière autonome. Pour en faciliter la lecture, j'ai choisi d'ouvrir chaque chapitre par la liste de mes articles qu'il recouvre, puis par une bibliographie historique de travaux d'auteurs antérieurs ou contemporains sur le même thème. Le référencement bibliographique est toujours strictement interne au chapitre considéré : la bibliographie de mes œuvres complètes, reproduite avant la présente introduction, ne sert donc de référence que pour celle-ci.

Les références bibliographiques sont données par ordre chronologique et non par ordre alphabétique des auteurs, et en ce qui concerne mes publications c'est non la chronologie stricte de publication qui est prise en compte – elle souffre parfois de la variabilité inhérente au système d'évaluation par les pairs – mais celle de soumission des manuscrits, qui est bien plus pertinente pour considérer l'évolution d'un travail. Exceptionnellement, comme le chapitre 9 a trait à des thèmes extrêmement variés, je n'y ai pas inclus de bibliographie historique mais ai usé de notes de bas de page pour citer les articles d'autres auteurs. Dans les autres chapitres, les notes de bas de page sont aussi utilisées pour citer des références qui, si elles ne sont pas directement dans le thème abordé, fournissent des lemmes techniques indispensables.

Dans chaque chapitre, les bibliographies précèdent un aperçu historique indiquant où en était la recherche avant que je n'explore le thème en question. Cet exposé est généralement suivi d'explications sur mes motivations et la chronologie de mes recherches.

Après ces grands traits, on rentre enfin pleinement dans le vif du sujet en présentant mes résultats et en discutant si possible des méthodes. J'ai essayé de respecter une division par article lorsque je le jugeais pertinent.

Dans quelques rares cas, je donne des démonstrations complètes de certains lemmes critiques. Dans d'autres, je fournis un éclairage géométrique à des travaux présentés sous un angle très matriciel, éclairage qui n'est pas toujours donné dans les articles mentionnés. Et à une occasion je corrige une erreur vénielle dans une démonstration.

En définitive, tous ces chapitres ont été conçus de telle sorte qu'un mathématicien non spécialiste des thèmes abordés puisse y lire un bel état de l'art, et en particulier du mien.

## 1.4 Des notations et conventions

Par souci d'économie, inférieur sera employé pour dire inférieur ou égal, supérieur pour supérieur ou égal.

Dans l'ensemble, j'utilise tout le système de notation classique français. On note donc  $\mathbb{N}$  l'ensemble des entiers naturels,  $\mathbb{N}^*$  celui des entiers naturels non nuls. La notation  $\llbracket i, j \rrbracket$  recouvre l'ensemble des entiers compris au sens large entre  $i$  et  $j$ .

Le corps de base est généralement noté  $\mathbb{K}$  : le groupe de ses éléments non nuls est alors noté  $\mathbb{K}^*$ , sa caractéristique  $\chi(\mathbb{K})$ , et son cardinal  $|\mathbb{K}|$ . Tout du long, on fixe une clôture algébrique  $\overline{\mathbb{K}}$  de  $\mathbb{K}$ .

Pour des espaces vectoriels  $U$  et  $V$ , on note  $\mathcal{L}(U, V)$  l'espace vectoriel des applications linéaires de  $U$  vers  $V$  ; on note  $\mathcal{L}(U)$  l'algèbre des endomorphismes de  $U$ , et  $\text{GL}(U)$  le groupe de ses automorphismes. On note  $U^* = \mathcal{L}(U, \mathbb{K})$  l'espace dual de  $U$ . Pour  $u \in \mathcal{L}(U, V)$ , on note  $u^t$  l'application linéaire transposée, définie comme suit :

$$u^t : \varphi \in V^* \mapsto \varphi \circ u \in U^*.$$

Étant donné des sous-espaces vectoriels  $E$  et  $F$  respectivement de  $U$  et  $U^*$ , on note

$$E^\circ := \{\varphi \in U^* : \forall x \in E, \varphi(x) = 0\} \quad \text{et} \quad {}^\circ F := \{x \in U : \forall \varphi \in F, \varphi(x) = 0\},$$

appelés respectivement dual-orthogonal de  $E$  et pré-dual-orthogonal de  $F$ .

La notation  $M_{n,p}(\mathbb{K})$  recouvre l'ensemble des matrices à  $n$  lignes,  $p$  colonnes et à coefficients dans  $\mathbb{K}$ , et  $M_n(\mathbb{K})$  celui des matrices carrées à  $n$  lignes. Pour une matrice rectangulaire  $M$ , on note  $m_{i,j}$  (ou  $M_{i,j}$  si nécessaire) son coefficient situé sur la ligne  $i$  et la colonne  $j$  ; on note aussi  $M^T$  la transposée de  $M$ . Les lignes de  $M$  sont notées  $L_1(M), \dots, L_n(M)$ , ses colonnes  $C_1(M), \dots, C_p(M)$ . Le rang de  $M$  est noté  $\text{rg } M$  ; lorsque  $M$  est carrée sa trace et son déterminant sont notés respectivement  $\text{tr } M$  et  $\det M$ , et son polynôme caractéristique  $\chi_M$ .

On note  $\text{GL}_n(\mathbb{K})$  le groupe des matrices inversibles de  $M_n(\mathbb{K})$ , et  $I_n$  la matrice identité d'ordre  $n$ . On note  $S_n(\mathbb{K})$  le sous-espace vectoriel de  $M_n(\mathbb{K})$  formé des matrices symétriques, et  $A_n(\mathbb{K})$  celui formé des matrices alternées : on rappelle qu'une matrice alternée est une matrice antisymétrique de coefficients diagonaux nuls (cette dernière précision est essentielle pour les corps de caractéristique 2). L'ensemble des matrices hermitiennes complexes de taille  $n$  est noté  $\mathcal{H}_n(\mathbb{C})$ .

On note  $\mathfrak{sl}_n(\mathbb{K})$  l'algèbre de Lie des matrices de  $M_n(\mathbb{K})$  de trace nulle. On note  $T_n^+(\mathbb{K})$ ,  $T_n^-(\mathbb{K})$ ,  $T_n^{++}(\mathbb{K})$  et  $T_n^{--}(\mathbb{K})$  les sous-ensembles de  $M_n(\mathbb{K})$  formés respectivement des matrices triangulaires supérieures, triangulaires inférieures, triangulaires supérieures strictes et triangulaires inférieures strictes.

Pour deux matrices carrées  $A \in M_n(\mathbb{K})$  et  $B \in M_p(\mathbb{K})$ , on note

$$A \oplus B := \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \in M_{n+p}(\mathbb{K}).$$

Pour  $i \in \llbracket 1, n \rrbracket$  et  $j \in \llbracket 1, p \rrbracket$ , on note  $E_{i,j}$  la matrice élémentaire de  $M_{n,p}(\mathbb{K})$  dont tous les coefficients sont nuls à l'exception du coefficient en position  $(i, j)$ , lequel vaut 1.

Enfin, on note le plus souvent  $X$  l'indéterminée des polynômes considérés. Ainsi,  $\mathbb{K}[X]$  désigne l'algèbre des polynômes en l'indéterminée  $X$  et à coefficients

dans le corps  $\mathbb{K}$ . Pour un polynôme unitaire  $P := X^n - \sum_{k=0}^{n-1} a_k X^k$  de degré  $n > 0$  à coefficients dans  $\mathbb{K}$ , on note sa **matrice compagnon**

$$C(P) := \begin{bmatrix} 0 & & & (0) & a_0 \\ 1 & 0 & & & a_1 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & & 0 & a_{n-2} \\ (0) & \cdots & 0 & 1 & a_{n-1} \end{bmatrix} \in M_n(\mathbb{K}).$$

## Chapitre 2

# Espaces de matrices de rang majoré

### 2.1 Articles concernés

- [1] C. de Seguins Pazzis, The affine preservers of non-singular matrices. *Arch. Math.* **95** (2010) 333–342.
- [2] C. de Seguins Pazzis, The classification of large spaces of matrices with bounded rank. *Israel J. Math.* **208** (2015) 219–259.
- [3] C. de Seguins Pazzis, Local linear dependence seen through duality II. *Linear Algebra Appl.* **462** (2014) 133–185.
- [4] C. de Seguins Pazzis, Affine spaces of symmetric or alternating matrices with bounded rank. *Linear Algebra Appl.* **504** (2016) 503–558.
- [5] C. de Seguins Pazzis, Large spaces of symmetric or alternating matrices with bounded rank. *Linear Algebra Appl.* **508** (2016) 146–189.
- [6] C. de Seguins Pazzis, The Flanders theorem over division rings. *Linear Algebra Appl.* **493** (2016) 313–322.
- [7] C. de Seguins Pazzis, Lines of full rank matrices in large subspaces. *Oper. Matrices* **10** (2016) 985–996.
- [8] C. de Seguins Pazzis, Large spaces of bounded rank matrices revisited. *Linear Algebra Appl.* **504** (2016) 124–189.
- [9] C. de Seguins Pazzis, Primitive spaces of matrices with upper rank two over the field with two elements. *Linear Multilinear Algebra* **64** (2016) 1321–1353.

### 2.2 Bibliographie historique

- [Die48] J. Dieudonné, Sur une généralisation du groupe orthogonal à quatre variables. *Arch. Math.* **1** (1948) 282–287.
- [Fla62] H. Flanders, On spaces of linear transformations with bounded rank. *J. Lond. Math. Soc.* **37** (1962) 10–16.
- [Atk80a] M.D. Atkinson, S. Lloyd, Large spaces of matrices of bounded rank. *Quart. J. Math. Oxford* (2) **31** (1980) 253–262.

- [Atk80b] M.D. Atkinson, S. Lloyd, Primitive spaces of matrices of bounded rank. *J. Austral. Math. Soc. (Ser. A)* **30** (1980) 473–482.
- [Atk83] M.D. Atkinson, Primitive spaces of matrices of bounded rank II. *J. Austral. Math. Soc. (Ser. A)* **34** (1983) 306–315.
- [Mes85] R. Meshulam, On the maximal rank in a subspace of matrices. *Quart. J. Math. Oxford (2)* **36** (1985) 225–229.
- [Bea87] L.B. Beasley, Null spaces of spaces of matrices of bounded rank. *Current Trends in Matrix Theory*, Elsevier, 1987, 45–50.
- [Eis88] D. Eisenbud, J. Harris, Vector spaces of matrices of low rank. *Adv. Math.* **37** (1988) 135–155.
- [Mes89] R. Meshulam, On two extremal matrix problems. *Linear Algebra Appl.* **114-115** (1989) 261–271.
- [Loe94] R. Loewy, N. Radwan, Spaces of symmetric matrices of bounded rank. *Linear Algebra Appl.* **197-198** (1994) 189–215.
- [Loe01] R. Loewy, Large spaces of symmetric matrices of bounded rank are decomposable. *Linear Multilinear Algebra* **48** (2001) 355–382.
- [Mes17] R. Meshulam, Maximal rank in matrix spaces via graph matchings. *Linear Algebra Appl.* **529** (2017) 1–11.

## 2.3 Problématique et historique

Le thème général abordé dans ce chapitre est celui des sous-espaces vectoriels de matrices rectangulaires sur un corps – ou encore, d’applications linéaires entre deux espaces vectoriels *a priori* distincts de dimension finie – vérifiant toutes (à l’exception éventuelle de la matrice nulle) des conditions restrictives sur leur rang.

Historiquement, la première observation substantielle sur ce thème est le théorème suivant, généralement attribué à Issai Schur, et que nous donnons sous sa formulation opérateurs et sans restriction de dimension.

**Théorème 2.1** (Théorème de Schur). *Soit  $\mathcal{V}$  un sous-espace vectoriel d’applications linéaires. Si tous les éléments de  $\mathcal{V}$  sont de rang au plus 1, alors l’une des situations suivantes se produit :*

- (i) *Tous les éléments non nuls de  $\mathcal{V}$  ont la même image.*
- (ii) *Tous les éléments non nuls de  $\mathcal{V}$  ont le même noyau.*

Alternativement, en dimension finie le résultat peut se reformuler en termes de produit tensoriel : dans le produit tensoriel  $E \otimes F$  de deux espaces vectoriels  $E$  et  $F$ , tout sous-espace vectoriel formé de tenseurs purs est de la forme  $x \otimes F'$  pour un sous-espace vectoriel  $F'$  de  $F$  et un vecteur  $x$  de  $E$ , ou  $E \otimes y$  pour un sous-espace vectoriel  $E'$  de  $E$  et un vecteur  $y$  de  $F$ .

À partir de la fin des années 1940, l’intérêt des chercheurs pour la question a été croissant. Et les recherches se sont déployées selon deux axes :

- Étude des sous-espaces vectoriels de matrices dont le rang est majoré par une constante  $r$  donnée. Les premiers travaux sont dus à Dieudonné [Die48] et Flanders [Fla62].



- Étude des sous-espaces vectoriels de matrices dans lesquels toute matrice *non nulle* est de rang supérieur ou égal à une constante  $r$  donnée ; ou de rang égal à  $r$ .

Dans le deuxième axe de recherche, la plupart des résultats dépendent très étroitement de l'arithmétique du corps de base. Par exemple, si l'on considère un sous-espace vectoriel de  $M_n(\mathbb{K})$  dans lequel toute matrice non nulle est de rang  $n$  (i.e. inversible), un tel espace est assez évidemment de dimension au plus  $n$  (car l'annulation de la première colonne force l'annulation de la matrice), mais la dimension maximale est 1 si  $\mathbb{K} = \mathbb{C}$ , si  $\mathbb{K} = \mathbb{R}$  elle ne vaut  $n$  que si  $n \in \{1, 2, 4, 8\}$  (conséquence du célèbre théorème de Bott, Milnor et Kervaire sur la dimension possible pour une algèbre à division sur  $\mathbb{R}$ ), et c'est systématiquement  $n$  si  $\mathbb{K} = \mathbb{Q}$  (par existence d'une extension algébrique de degré  $n$  de  $\mathbb{Q}$ ). Ce genre de problème fait souvent intervenir des résultats de théorie de Galois (pour les corps finis notamment), des techniques de géométrie algébrique pour les corps algébriquement clos, et même de la topologie algébrique pour le corps des réels. Nous nous sommes tenus largement à l'écart de cet axe particulier de recherche, principalement par préférence pour les énoncés très généraux et indépendants du corps de base envisagé. En conséquence, dans ce qui suit nous évoquerons uniquement l'histoire des travaux portant sur le premier axe de recherche sur les sous-espaces de matrices avec contraintes de rang.

Précisons d'abord quelques notations et concepts importants. On se donne trois entiers naturels  $n, p, r$  tels que  $r < \min(n, p)$ , et l'on rappelle que l'on travaille sur un corps  $\mathbb{K}$ . Pour simplifier les énoncés, on note

$$\text{rgmax } \mathcal{V} := \max\{\text{rg } M \mid M \in \mathcal{V}\}$$

le rang maximal pour un sous-ensemble non vide  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$ . On adopte bien sûr la même notation pour un sous-ensemble d'applications linéaires (de rang fini) d'un espace vectoriel vers un autre.

Bien entendu, la relation pertinente pour étudier les ensembles de matrices vis-à-vis du rang est celle de l'**équivalence** : deux sous-ensembles  $\mathcal{A}$  et  $\mathcal{B}$  de  $M_{n,p}(\mathbb{K})$  sont équivalents lorsqu'il existe deux matrices inversibles  $P \in GL_n(\mathbb{K})$  et  $Q \in GL_p(\mathbb{K})$  telles que  $\mathcal{B} = P \mathcal{A} Q$ , autrement dit lorsque  $\mathcal{A}$  et  $\mathcal{B}$  représentent le même ensemble d'applications linéaires dans des bases *a priori* différentes des espaces de départ et d'arrivée. Deux matrices équivalentes ayant même rang, le rang maximal est le même dans  $\mathcal{A}$  et  $\mathcal{B}$ . En outre, le rang maximal est invariant par transposition.

Voici maintenant deux exemples standards de sous-espaces de rang maximal  $r$  :

- l'espace  $\mathcal{R}(0, r)$  des matrices dont les  $p-r$  dernières colonnes sont nulles ;
- l'espace  $\mathcal{R}(r, 0)$  des matrices dont les  $n-r$  dernières lignes sont nulles.

Le théorème de Schur se reformule par exemple en disant que tout sous-espace vectoriel  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  de rang maximal 1 est équivalent à un sous-espace vectoriel de  $\mathcal{R}(0, 1)$  ou de  $\mathcal{R}(1, 0)$ . Et un mathématicien non averti pourrait penser que, dans la droite ligne de ce résultat, tout sous-espace vectoriel de rang maximal inférieur ou égal à  $r$  doit être équivalent à un sous-espace vectoriel de  $\mathcal{R}(0, r)$  ou  $\mathcal{R}(r, 0)$ .

Pour dissiper ce type d'idée fautive, rien de tel que de construire une gamme encore plus vaste de sous-espaces intéressants, traditionnellement appelés **espaces de compression** : pour tout  $(s, t) \in \llbracket 0, n \rrbracket \times \llbracket 0, p \rrbracket$  tel que  $s + t = r$ ,

on note  $\mathcal{R}(s, t)$  (ou plus exactement  $\mathcal{R}(s, t)_{n,p}$  lorsque le contexte n'indique pas précisément le format des matrices envisagées) l'espace formé des matrices de la forme

$$\begin{bmatrix} [?]_{s \times t} & [?]_{s \times (p-t)} \\ [?]_{(n-s) \times t} & 0 \end{bmatrix}.$$

Il est facile de vérifier que  $\mathcal{R}(s, t)$  est un sous-espace vectoriel de rang maximal  $s + t$ , et qu'il est maximal parmi ces sous-espaces vectoriels! En outre, il est facile de démontrer que les sous-espaces  $\mathcal{R}(s, t)$  sont deux à deux inéquivalents. En combinant ces deux principes, il vient qu'aucun espace de compression n'est équivalent à un sous-espace vectoriel d'un autre espace de compression. En particulier, dans l'espace  $M_3(\mathbb{K})$ , l'espace vectoriel des matrices de la forme

$$\begin{bmatrix} * & * & * \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix}$$

est de rang maximal 2, mais n'est équivalent à aucun sous-espace vectoriel de  $\mathcal{R}(2, 0)$  ou de  $\mathcal{R}(0, 2)$ . Nous dirons qu'une partie  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  est  **$r$ -décomposable** lorsqu'elle est équivalente à une partie d'un espace de compression de rang maximal  $r$ .

Bien entendu, pour connaître les sous-espaces vectoriels de rang maximal  $r$ , il suffit de connaître ceux qui sont maximaux. Mais à nouveau les espaces de compression ne suffisent pas à tout décrire, même à équivalence près. Par exemple, pour tout  $n \in \mathbb{N}^*$ , on peut facilement démontrer que  $A_{2n+1}(\mathbb{K})$  est un sous-espace de matrices de rang maximal  $2n$ , et qu'il est maximal dans  $M_{2n+1}(\mathbb{K})$  parmi les sous-espaces de matrices singulières. Et il est facile de voir que  $A_{2n+1}(\mathbb{K})$  n'est pas équivalent à un espace de compression.

Le programme général de détermination des sous-espaces maximaux, parmi ceux de rang maximal  $r$ , est nettement trop ambitieux. Et historiquement, les résultats les plus intéressants sont venus des deux approches suivantes :

- (i) L'étude des sous-espaces *de grande dimension*. On commence par déterminer la dimension maximale pour un sous-espace vectoriel de rang maximal  $r$ , puis on étudie la structure des sous-espaces vectoriels de dimension maximale. Enfin, et c'est nettement plus délicat, on tente d'élucider la structure de tous les sous-espaces maximaux de dimension suffisamment proche de la dimension maximale. C'est l'approche de Dieudonné [Die48] et Flanders [Fla62] (calcul de la dimension maximale et détermination des sous-espaces la réalisant), et d'Atkinson et Lloyd [Atk80a] (étude des sous-espaces de dimension proche de la dimension maximale).
- (ii) La limitation de l'étude aux sous-espaces ne pouvant se déduire, à partir de techniques d'extension rudimentaires, d'autres sous-espaces de *format* plus petit : c'est l'étude des espaces **primitifs** initiée par Atkinson et Lloyd [Atk80b], prolongée par Atkinson seul [Atk83], et redécouverte ultérieurement par Eisenbud et Harris [Eis88]. Cette approche est bien adaptée aux matrices de petit rang (voir le titre de l'article d'Eisenbud et Harris, qui fait explicitement mention d'espaces vectoriels de faible rang).

Le problème a été ici posé dans le cadre de matrices rectangulaires. Mais il admet également des variantes dans à peu près tous les cas de matrices spéciales envisagées, par exemple les matrices carrées symétriques/alternées/hermitiennes.

Ces variantes du problème, suivant toujours le schéma mis en place par Dieudonné et Flanders, ont été étudiées initialement par Meshulam, Radwan et Loewy [Mes89, Loe94, Loe01] autour de la décennie 1990-2000.

## 2.4 Historique de mes travaux personnels

Je me suis intéressé au problème de Flanders pour la première fois à l'été 2006. Il s'agissait de recherches indépendantes sans ambition de publication aucune. De fil en aiguille, j'ai trouvé au printemps 2007 de nouvelles idées qui ont permis de généraliser les résultats de Flanders, datant de 1962, à des matrices sur un corps quelconque. Les résultats furent à l'époque publiés sous la forme d'un article pour la RMS<sup>1</sup>, y compris l'application aux préserveurs linéaires de l'inversibilité. Au cours de mes recherches, que j'avais orientées vers des méthodes itératives (récurrence sur le format des matrices envisagées) où l'on regardait finement les matrices de rang 1 figurant dans l'espace considéré, mon collègue Saab Abou-Jaoudé avait trouvé un raffinement de la méthode de Flanders et me l'avait communiqué. Ce raffinement, consistant en une application de la technique de projection/extraction des blocs, permettait d'éviter toute récurrence et se révéla redoutablement efficace.

Je me remis à la tâche au cours de l'hiver 2010, alors que je découvrais l'existence des travaux d'Atkinson et Lloyd. Je pris très vite conscience que les techniques qu'Abou-Jaoudé et moi-même avions développées permettraient très probablement de généraliser les résultats d'Atkinson et Lloyd, notamment en se débarrassant des hypothèses de cardinalité. Et c'est ce qui fut fait en quelques jours. J'avais pour cela mis le doigt sur la nécessité de renforcer la conclusion du théorème de Dieudonné sur les sous-espaces vectoriels de matrices singulières. En effet, la partie la plus simple du théorème indique, par contraposition, que tout sous-espace vectoriel  $\mathcal{V}$  de  $M_n(\mathbb{K})$  de codimension strictement inférieure à  $n$  contient une matrice inversible. Pour les besoins de mon étude, il me fallait en plus savoir que  $\mathcal{V}$  était *engendré* par ses matrices inversibles. À l'époque, j'avais rédigé tout un article pour démontrer précisément cela, mais tout s'éclaira lorsque je pris conscience que ce résultat était contenu de manière déguisée dans l'article de Dieudonné. En effet, si  $\mathcal{V}$  n'est pas engendré par ses matrices inversibles, ces dernières sont piégées dans un hyperplan vectoriel  $H$  de  $\mathcal{V}$ , et tout hyperplan affine  $\mathcal{H}$  de  $\mathcal{V}$  strictement parallèle à  $H$  est alors constitué de matrices singulières, et de codimension au plus  $n$  dans  $M_n(\mathbb{K})$ . Mais le théorème de Dieudonné (théorème 2.2) assure qu'un tel espace affine  $\mathcal{H}$  ne peut exister sauf dans un cas très particulier pour les corps à deux éléments. J'ai donc jeté ce projet d'article à la corbeille, mais comme j'avais besoin d'une généralisation de ce résultat aux sous-espaces affines de matrices rectangulaires et avec un rang maximal  $r$  quelconque (ce qui n'a rien d'évident à partir des résultats de Dieudonné), j'ai inclus la démonstration – largement simplifiée grâce au point de vue des sous-espaces affines – dans un petit article sur les préserveurs affines de l'inversibilité [1]. Ensuite, ma généralisation du théorème d'Atkinson et Lloyd, sur les sous-espaces de rang majoré dont la dimension est proche de la dimension maximale, fut soumise pour être publiée seulement cinq années plus tard [2]. Et, en parallèle, j'appliquai ces nouveaux résultats à la détermination des

---

1. Clément de Seguins Pazzis, Sous-espaces vectoriels de matrices de rang au plus  $p$ . RMS **118-4** (2008) 33–61.

préserveurs linéaires de l'inversibilité définis sur des sous-espaces de matrices carrées de petite codimension (voir la section 5.6).

Je découvris l'existence de la théorie d'Atkinson et Lloyd sur les sous-espaces primitifs en 2012 lors de mes travaux sur les espaces d'opérateurs localement liés. J'avais moi-même déjà trouvé une partie de leur résultat principal, mais ai tout de même dû m'incliner devant la subtilité des méthodes déployées par Atkinson dans son article final [Atk83]. Et, bien que j'aie substantiellement amélioré les résultats d'Atkinson dans [3], je ne peux pas prétendre avoir révolutionné le sujet. Mes découvertes majeures sur la question résident dans l'observation du lien entre sous-espaces primitifs et espaces d'opérateurs localement liés minimaux – lien qui avait totalement échappé aux autres chercheurs à l'époque – mais aussi d'un lien beaucoup moins évident entre le théorème d'Atkinson sur les sous-espaces primitifs et le problème de Gerstenhaber sur les sous-espaces vectoriels de matrices nilpotentes (voir la section 6.6). Et j'ai contribué à la zoologie des cas exotiques en déterminant complètement les sous-espaces primitifs de matrices de rang au plus 2 sur  $\mathbb{F}_2$ , avec diverses applications [9].

Je laissai ensuite temporairement ces questions derrière moi, et n'y revint qu'en 2014 lorsque je me décidai enfin à attaquer la question des sous-espaces vectoriels de matrices symétriques de rang majoré. En 2010, j'avais déjà essayé de l'attaquer, sans succès. La méthode d'Abou-Jaoudé se révèle en effet totalement inopérante dans ce cadre. Il me fallut donc réinventer les méthodes, ce qui passa par une forme retour à des idées plus anciennes, que j'ai considérablement dépoussiérées. La méthode que je trouvai alors est fondée sur une récurrence sur le format des matrices, une prise en compte systématique des sous-espaces affines plutôt que des sous-espaces vectoriels, un examen des matrices de petit rang (rang au plus 1 pour les matrices rectangulaires, rang au plus 2 pour les matrices symétriques/alternées) dans le sous-espace envisagé, et enfin le développement parallèle de résultats sur les morphismes Im-compatibles, pour lesquels j'avais enfin trouvé en 2013 le cadre conceptuel adéquat et les outils appropriés (voir le chapitre 4).

La période 2014-2016 connut alors une très forte concentration de mes travaux sur ces questions. J'écrivis le court article [6] en guise d'introduction à ces nouvelles méthodes : le théorème de Flanders y est élargi au cadre d'un corps gauche de degré fini sur son centre. Ensuite, je publiai une série de deux articles sur les sous-espaces vectoriels de matrices symétriques/alternées de rang majoré, améliorant nettement tous les résultats connus sur la question [4, 5] et enfin je publiai une sorte d'article définitif sur le théorème de Flanders [8], utilisant la nouvelle méthode pour aller encore plus loin dans la généralisation des résultats d'Atkinson et Lloyd. Parallèlement, je dus étoffer les résultats sur les morphismes Im-compatibles dans toutes sortes de directions, mais je n'aborderai cela que dans le chapitre 4. Et enfin, parce que j'avais besoin de cet outil pour [8], j'étudiai aussi des généralisations de la contraposée du théorème de Dieudonné : si un sous-espace de  $M_n(\mathbb{K})$  est de codimension assez petite, alors il contient nécessairement une matrice inversible ; mon problème était de trouver la meilleure contrainte de codimension garantissant qu'il y ait non seulement des matrices inversibles, mais qu'il y ait des droites affines entières de matrices inversibles selon toute direction raisonnable : ce fut fait dans [7].

## 2.5 Sous-espaces de grande dimension

### 2.5.1 De Dieudonné à Atkinson-Lloyd

J'ai déjà cité le résultat de Schur sur les sous-espaces de matrices de rang au plus 1 (théorème 2.1). Après Schur, il a fallu attendre Dieudonné pour voir apparaître un nouveau résultat important sur les sous-espaces de rang majoré :

**Théorème 2.2** (Dieudonné (1948)). *Soit  $\mathcal{V}$  un sous-espace affine de  $M_n(\mathbb{K})$  ne contenant aucune matrice inversible. Alors,  $\dim \mathcal{V} \leq n(n-1)$ , et en cas d'égalité :*

- (i) ou bien  $\mathcal{V}$  est équivalent à  $\mathcal{R}(0, n-1)$  ou à  $\mathcal{R}(n-1, 0)$  ;
- (ii) ou bien  $n = |\mathbb{K}| = 2$  et  $\mathcal{V}$  est équivalent à

$$\left\{ \begin{bmatrix} x & y \\ 0 & x+1 \end{bmatrix} \mid (x, y) \in \mathbb{K}^2 \right\}.$$

Dieudonné s'était penché sur les sous-espaces affines de matrices singulières pour accéder à une description des automorphismes de l'espace vectoriel  $M_n(\mathbb{K})$  qui conservent l'inversibilité au sens fort. Nous reviendrons sur cet aspect dans le chapitre 5.

Quelques années plus tard, Flanders, visiblement ignorant des résultats de Dieudonné, s'intéressa aux sous-espaces de matrices rectangulaires et démontra le résultat suivant :

**Théorème 2.3** (Flanders (1962)). *Soit  $1 \leq r \leq p \leq n$  des entiers naturels, et  $\mathcal{V}$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$ . On suppose  $\chi(\mathbb{K}) \neq 2$  et  $|\mathbb{K}| > r$ . Si  $\text{rgmax } \mathcal{V} \leq r$ , alors  $\dim \mathcal{V} \leq nr$ , et en cas d'égalité ou bien  $\mathcal{V}$  est équivalent à  $\mathcal{R}(0, r)$ , ou bien  $n = p$  et  $\mathcal{V}$  est équivalent à  $\mathcal{R}(r, 0)$ .*

Les techniques mises en œuvre par Flanders sont profondément différentes de celles de Dieudonné. Alors que ce dernier s'appuie sur une récurrence assez technique et très élémentaire, Flanders utilise des techniques rudimentaires de géométrie algébrique : prenant une matrice  $A \in \mathcal{V}$  de rang  $r$  (pour peu qu'elle existe) ainsi qu'une matrice  $M \in \mathcal{V}$  non nulle, le fait que toute matrice de la droite affine  $A + \mathbb{K}M$  soit de rang au plus  $r$  apporte des informations cruciales sur la matrice  $M$ . Si par exemple  $\mathbb{K}$  est infini, on sait que l'espace tangent au point  $A$  à la variété déterminantale  $\mathcal{G}_r$  des matrices de  $M_{n,p}(\mathbb{K})$  de rang exactement  $r$  est constitué des matrices  $M$  envoyant  $\text{Ker } A$  dans  $\text{Im } A$ . Autrement dit, si  $A$  est de la forme

$$A = \begin{bmatrix} P & [0]_{r \times (p-r)} \\ [0]_{(n-r) \times r} & [0]_{(n-r) \times (p-r)} \end{bmatrix}$$

avec  $P \in \text{GL}_r(\mathbb{K})$ , alors toute matrice de l'espace tangent à  $\mathcal{G}_r$  en  $A$  est de la forme

$$M = \begin{bmatrix} [?]_{r \times r} & C(M) \\ L(M) & [0]_{(n-r) \times (p-r)} \end{bmatrix}.$$

Mieux, si  $A + \text{Vect}(M)$  est incluse dans  $\mathcal{G}_r$ , une dérivation à l'ordre 2 des mineurs extraits d'ordre  $r+1$  donne la relation

$$L(M)P^{-1}C(M) = 0. \tag{2.1}$$

Cette observation permet ensuite à Flanders d'obtenir son résultat, les contraintes obtenues à l'ordre 2 limitant la dimension grâce à l'orthogonalité pour une certaine forme quadratique. Cette considération explique la contrainte de caractéristique par Flanders ; néanmoins, une modification mineure de ses arguments permet de s'en affranchir.

Un argument très simple, dont ne trouvons pas trace dans la littérature de l'époque, aurait permis à Flanders d'obtenir la majoration de la dimension sur un corps *fini* quelconque dans le cas de matrices carrées. Il consiste, sur un tel corps, à montrer qu'il existe, pour tous  $n \geq p \geq r \geq 1$ , un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$  de dimension  $nr$  dont tous les éléments non nuls sont de rang strictement supérieur à  $n - r$ . Ce point s'obtient en considérant des applications polynomiales construites en combinant linéairement des puissances de l'endomorphisme de Frobenius de  $\overline{\mathbb{K}}$  sur une extension finie de  $\mathbb{K}$  : voir mon article RMS (op. cit. note en bas de la page 25) à ce sujet. Ensuite, un simple argument d'intersection permet de conclure.

Il fallut attendre plus de vingt ans pour voir Roy Meshulam [Mes85] démontrer enfin le théorème de Flanders sans condition de cardinalité ni de caractéristique. Meshulam établit un lien entre le problème de Flanders et des questions purement combinatoires, et il s'appuya sur un théorème de König consacré aux matrices à coefficients dans  $\{0, 1\}$ .

Entretemps, Atkinson et Lloyd [Atk80a] avaient étudié la question suivante : peut-on plus généralement donner la structure de sous-espaces vectoriels de  $M_{n,p}(\mathbb{K})$  de grande dimension et de rang majoré ?

Leur réponse fut la suivante, limitée aux matrices carrées :

**Théorème 2.4** (Atkinson-Lloyd (1980)). *Soit  $1 \leq r \leq n$  deux entiers naturels, et  $\mathcal{V}$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  de rang maximal  $r$ . Si  $\dim \mathcal{V} \geq nr - n + 1$  et  $|\mathbb{K}| > r$ , alors  $\mathcal{V}$  est équivalent à une partie de l'un des espaces  $\mathcal{R}(0, r)$ ,  $\mathcal{R}(1, r - 1)$ ,  $\mathcal{R}(r - 1, 1)$  ou  $\mathcal{R}(r, 0)$ .*

Le résultat a ensuite été facilement prolongé par Beasley à des matrices rectangulaires [Bea87].

La démonstration d'Atkinson et Lloyd reprend les idées de Flanders et les pousse plus loin, mais ne contient pas d'innovation technique majeure.

Vingt ans plus tard, aucune avancée substantielle n'avait été faite sur ces théorèmes. En particulier, la question en suspens était la généralisation du résultat d'Atkinson et Lloyd à un corps arbitraire, et éventuellement l'attaque de dimensions encore inférieures. Le contre-exemple suivant sur  $\mathbb{F}_2$ , publié par Meshulam [Mes85], semble avoir suscité bien des craintes : les matrices de  $M_3(\mathbb{F}_2)$  de la forme

$$\begin{bmatrix} a & ? & ? \\ 0 & b & ? \\ 0 & 0 & a + b \end{bmatrix}$$

sont toutes singulières. Elles forment portant un sous-espace vectoriel de dimension 5 qui n'est équivalent à aucun sous-espace vectoriel de l'un des espaces  $\mathcal{R}(2, 0)$ ,  $\mathcal{R}(1, 1)$  et  $\mathcal{R}(0, 2)$ . Cet exemple est intimement lié au sous-espace affine exceptionnel mis en évidence par Dieudonné et cité dans le théorème 2.2.

Et voilà où en était la recherche en 2010.

## 2.5.2 La méthode d'Abou-Jaoudé

L'idée centrale d'Abou-Jaoudé est un retour à la situation canonique étudiée par Flanders. Prenons un sous-espace affine  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  de rang maximal  $r$ , et notons  $V$  sa direction. On suppose  $n = p$  (le cas rectangulaire nécessite une légère adaptation que nous n'évoquerons pas ici). Écrivons toute matrice  $A$  de  $M_n(\mathbb{K})$  sous la forme

$$A = \begin{bmatrix} K(A) & C(A) \\ L(A) & D(A) \end{bmatrix}$$

où  $K(A) \in M_r(\mathbb{K})$  etc. Supposons que  $\mathcal{V}$  contienne une matrice  $A$  telle que  $K(A)$  soit inversible. Dans ce cas, même si  $C(A)$ ,  $L(A)$  et  $D(A)$  sont tous nuls, il n'est plus vrai en général que toute matrice de la direction  $V$  de  $\mathcal{V}$  ait son bloc inférieur droit nul (penser au plan sur  $\mathbb{F}_2$  engendré par deux matrices diagonales). La relation (2.1) mise en évidence par Flanders sur les deux blocs hors-diagonaux est tout aussi fautive à ce degré de généralité. L'innovation d'Abou-Jaoudé a été de ne pas prendre une matrice arbitraire dans  $V$ , mais seulement une matrice dont le bloc supérieur gauche est nul, autrement dit

$$M = \begin{bmatrix} 0_r & C(M) \\ L(M) & D(M) \end{bmatrix}.$$

En cela, il utilise judicieusement la technique de projection/extraction des blocs. Le théorème du rang lui permet de voir que

$$\dim \mathcal{V} \leq \dim K(\mathcal{V}) + \dim W,$$

où  $W$  est le sous-espace vectoriel de  $V$  formé des matrices de la forme particulière indiquée. Et il s'agira alors de majorer judicieusement la dimension de  $W$ , celle de  $K(\mathcal{V})$  étant trivialement majorée par  $r^2$ .

Pour  $M$  dans  $W$  et  $A$  dans  $\mathcal{V}$ , un simple pivot fournit la relation polynomiale

$$D(M) + D(A) = L(M + A) K(A)^{-1} C(M + A)$$

(on peut aussi penser au complément de Schur). En soustrayant à celle-ci l'égalité obtenue pour  $M = 0$ , on trouve

$$D(M) = L(M) K(A)^{-1} C(M) + L(M) K(A)^{-1} C(A) + L(A) K(A)^{-1} C(M).$$

En particulier,  $D(M) = 0$  dès que  $L(M) = 0$  et  $C(M) = 0$ . En outre, une polarisation de l'identité précédente donne

$$\forall (M, N) \in W^2, L(M)K(A)^{-1}C(N) + L(N)K(A)^{-1}C(M) = 0.$$

En particulier, pour le sous-espace vectoriel

$$W' := \{M \in W : L(M) = 0\},$$

on trouve

$$\forall (M, N) \in W \times W', L(M)K(A)^{-1}C(N) = 0. \quad (2.2)$$

Or, le théorème du rang fournit l'identité

$$\dim \mathcal{V} = \dim K(\mathcal{V}) + \dim W + \dim W'.$$

En exploitant finement la relation (2.2), on peut obtenir

$$\dim W + \dim W' \leq (n - r)r.$$

La majoration brutale  $\dim K(\mathcal{V}) \leq r^2$  permet alors d'obtenir la majoration de la dimension dans le théorème de Flanders.

C'est à partir de ce point précis que mes idées prennent le relais de celles d'Abou-Jaoudé. D'abord, en analysant la situation précédente dans le cas de la dimension maximale, on voit que les différentes inégalités de dimension deviennent des égalités, et en raisonnant finement on parvient à les exploiter pour reconstituer la structure de  $\mathcal{V}$  : nous renvoyons à [1] pour les détails d'une telle démonstration et nous nous contentons de citer le résultat central qui y est établi :

**Théorème 2.5** (Théorème 4 de [1]). *Soit  $n \geq p > r \geq 1$  des entiers, et  $\mathcal{V}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  dans lequel toute matrice est de rang inférieur ou égal à  $r$ . Alors  $\dim \mathcal{V} \leq nr$ . Si en outre  $\dim \mathcal{V} = nr$  alors l'une des trois situations suivantes est vérifiée :*

- (i)  $\mathcal{V}$  est équivalent à  $\mathcal{R}(0, r)$  ;
- (ii)  $n = p$  et  $\mathcal{V}$  est équivalent à  $\mathcal{R}(r, 0)$  ;
- (iii)  $n = p = 2$ ,  $r = 1$ ,  $|\mathbb{K}| = 2$  et  $\mathcal{V}$  est équivalent à

$$\left\{ \begin{bmatrix} x & y \\ 0 & x+1 \end{bmatrix} \mid (x, y) \in \mathbb{K}^2 \right\}.$$

J'ai ensuite prolongé l'argumentation pour généraliser le théorème d'Atkinson et Lloyd à un corps quelconque : supposons dans ce cadre que  $\mathcal{V}$  soit un sous-espace *vectoriel* et que  $\dim V > nr - r + 1$  (le cas d'égalité est un peu plus délicat mais repose sur les mêmes idées fondamentales). D'abord, on peut améliorer marginalement les identités précédentes en prenant un bloc  $K(A)$  plus général : on trouve alors

$$\forall (M, N) \in W \times W', \forall A \in V, L(M)K(A)^{\text{ad}}C(N) = 0, \quad (2.3)$$

où  $K(A)^{\text{ad}}$  est l'adjoint principal de  $K(A)$ , autrement dit la transposée de sa comatrice.

Nous supposons dans un premier temps que  $C(W') \neq \{0\}$  et  $L(W) \neq \{0\}$ . Prenons une colonne non nulle  $X$  d'une matrice de  $C(W')$ , et une ligne non nulle  $Y$  d'une matrice de  $L(W)$ . Ainsi,

$$\forall A \in V, YK(A)^{\text{ad}}X = 0.$$

On ne perd bien sûr pas de généralité particulière à supposer que  $X^T = Y = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$ . Dans ce cas, on voit que pour tout  $A \in V$ , son bloc  $(r-1) \times (r-1)$  supérieur gauche, noté  $Z(A)$ , est singulier ! Autrement dit,  $Z(V)$  est un sous-espace vectoriel de  $M_{r-1}(\mathbb{K})$  de rang maximal au plus  $r-2$  ! Le théorème de Dieudonné montre alors que  $\text{codim}_{M_{r-1}(\mathbb{K})} Z(V) \geq r-1$ . En combinant cela avec les techniques précédentes, on parvient alors à  $\dim V \leq nr - r + 1$ . Si  $\dim V = nr - r + 1$ , alors le cas critique dans le théorème de Dieudonné permet



d'obtenir la structure de  $Z(V)$ , puis de reconstituer celle de  $V$  (ce n'est pas immédiat!).

Il reste alors à examiner la situation où  $C(W') = \{0\}$  ou  $L(W) = \{0\}$ , qui est la plus délicate. Dans ce cas, quitte à transposer  $V$  on peut supposer que  $C(W') = \{0\}$ . Comme vu précédemment, cela indique que toute matrice de  $V$  ayant toutes ses  $r$  premières colonnes nulles est totalement nulle! En d'autres termes, on peut mettre  $V$  sous la forme de l'espace des matrices de la forme

$$[N \quad \varphi(N)]$$

où  $N$  vit dans un sous-espace vectoriel  $\overline{V}$  de  $M_{n,r}(\mathbb{K})$ , de même dimension que  $V$ , et  $\varphi : \overline{V} \rightarrow M_{n,p-r}(\mathbb{K})$  est une fonction linéaire.

La suite du travail fait apparaître la notion d'application Im-compatible, qui fait l'objet du chapitre 4 et a été découverte à cette occasion. Nous allons démontrer que chaque colonne de  $\varphi(N)$  est dans l'image de la matrice  $N$  pour tout  $N \in \overline{V}$ . Notons  $V_0$  le sous-espace vectoriel de  $\overline{V}$  formé de ses matrices de dernière ligne nulle. On écrit ces éléments

$$N = \begin{bmatrix} U(N) \\ [0]_{1 \times r} \end{bmatrix} \quad \text{où } U(N) \in M_{n-1,r}(\mathbb{K}).$$

Supposons que le coefficient de  $\varphi(N)$  en position  $(n, 1)$  ne soit pas systématiquement nul lorsque  $N \in V_0$ . Les éléments  $N$  de  $V_0$  tels que  $\varphi(N)_{n,1} = 1$  forment donc un hyperplan affine  $\mathcal{H}$  de  $V_0$ , et  $\text{rgmax} U(\mathcal{H}) < r$ . Le théorème de Flanders généralisé par nos soins aux sous-espaces affines (voir le théorème 2.5) garantit donc que

$$\dim U(\mathcal{H}) \leq (n-1)(r-1),$$

alors que

$$\dim U(\mathcal{H}) \geq \dim \overline{V} - r - 1 > nr - 2r \geq nr - n - r + 1 = (n-1)(r-1).$$

C'est contradictoire. Ainsi  $\varphi(N)_{n,1} = 0$  pour tout  $N \in V_0$ , et plus généralement  $\varphi(N)_{n,j} = 0$  pour tout  $N \in V_0$  et tout  $j \in \llbracket 1, p-r \rrbracket$ . En généralisant encore, on en déduit que pour tout  $N \in \overline{V}$ , tout hyperplan vectoriel incluant  $\text{Im } N$  inclut aussi  $\text{Im } \varphi(N)$ , ce qui par dualité donne  $\text{Im } \varphi(N) \subset \text{Im } N$ . Ainsi, chaque fonction  $N \mapsto C_j(\varphi(N))$  est Im-compatible au sens du chapitre 4.

Par une étude des morphismes Im-compatibles définis sur un sous-espace de matrices rectangulaires de petite codimension, qui se fait de manière totalement autonome à la démonstration précédente, on parvient alors à montrer qu'après transvections sur les colonnes l'espace  $V$  se réduit à un sous-espace vectoriel de  $\mathcal{R}(0, r)$ .

### 2.5.3 La méthode itérative et ses fruits

Malgré le succès de la méthode précédente pour généraliser le théorème d'Atkinson-Lloyd, elle a montré ses limites. En particulier, je n'ai pas réussi à l'adapter au cadre des sous-espaces de matrices symétriques/alternées. À mon sens, la méthode d'Abou-Jaoudé souffre d'une approche trop globale.

Pour obtenir des résultats plus poussés, il s'est avéré indispensable de revenir à des arguments encore plus élémentaires que ceux de la méthode précédente : la clef réside dans l'analyse fine des matrices de rang 1 qui appartiennent à la

direction de  $\mathcal{V}$ . Les idées sont en fait proches de celles utilisées par Dieudonné en [Die48], mais différent en quelques étapes cruciales. Je vais expliquer ci-après comment on peut envisager une démonstration, par simple récurrence sur le triplet  $(n, p, r)$  (ou, si l'on préfère, sur la valeur de  $n+p+r$ ), de théorèmes de type Flanders et Atkinson-Lloyd. Les lignes qui suivent reprennent des explications figurant dans l'introduction de [8].

Donnons-donc un sous-espace affine  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  de rang maximal au plus  $r$ , et notons  $V$  sa direction. On va s'intéresser de très près à l'intersection de  $V$  avec les sous-espaces maximaux de matrices de rang au plus 1.

La méthode se déploie en trois sous-méthodes. La première sous-méthode est baptisée « **Effacer une ligne et une colonne** » (en abrégé, **ELC**) : on y suppose que  $V$  contient une matrice de rang 1, mettons  $E_{1,1}$ . On écrit alors par blocs toute matrice  $M$  de  $\mathcal{V}$  sous la forme

$$M = \begin{bmatrix} ? & [?]_{1 \times (p-1)} \\ [?]_{(n-1) \times 1} & K(M) \end{bmatrix} \quad \text{avec } K(M) \in M_{n-1,p-1}(\mathbb{K}).$$

Il est alors facile de constater que  $K(\mathcal{V})$  est un sous-espace affine de  $M_{n-1,p-1}(\mathbb{K})$  de rang maximal au plus  $r-1$ . Par ailleurs, le théorème du rang fournit l'inégalité

$$\dim K(\mathcal{V}) \geq \dim \mathcal{V} - (p-1) - \dim U,$$

où  $U$  désigne le sous-espace vectoriel de  $V$  formé des matrices dont les  $p-1$  dernières colonnes sont nulles. Dès que la dimension de  $U$  est suffisamment petite, une hypothèse de récurrence peut s'appliquer à  $K(\mathcal{V})$ .

Si  $K(\mathcal{V})$  est inclus dans l'espace de compression  $\mathcal{R}(i, r-1-i)$ , alors  $\mathcal{V}$  est inclus dans  $\mathcal{R}(i+1, r-i)$ , qui n'est pas bien sûr un espace de rang maximal  $r$  en général. Cependant, avec un travail supplémentaire (éventuellement assez technique), il est possible de montrer sous des hypothèses spécifiques que  $\mathcal{V}$  se retrouve équivalent à un sous-espace affine de  $\mathcal{R}(i, r-i)$  ou de  $\mathcal{R}(i+1, r-i-1)$ . Un point clef de ce procédé de relèvement réside dans l'usage de théorèmes antérieurement établis pour les sous-espaces affines : plus précisément, pour démontrer la généralisation du théorème d'Atkinson-Lloyd pour les sous-espaces affines, on s'appuie sur le théorème de Flanders pour les sous-espaces affines ; pour généraliser encore le théorème d'Atkinson-Lloyd à une gamme de dimensions supérieures pour les sous-espaces vectoriels, on s'appuie sur le théorème d'Atkinson-Lloyd généralisé aux espaces affines. Précisons que ces extensions aux sous-espaces affines sont absolument cruciales pour les petits corps finis, mais qu'elles sont gratuites dès que le corps de base est de cardinal suffisamment grand : en effet, en exploitant la nullité des mineurs, on voit que si le corps de base possède au moins  $r+2$  éléments alors tout sous-espace affine de rang maximal majoré par  $r$  a son sous-espace vectoriel engendré de rang maximal tout aussi majoré par  $r$ .

La notation suivante sera utile pour aborder ces idées plus clairement :

**Notation 1.** Pour un sous-ensemble  $S$  de  $M_{n,p}(\mathbb{K})$ , une droite vectorielle  $D$  de  $\mathbb{K}^n$ , ainsi qu'un hyperplan vectoriel  $H$  de  $\mathbb{K}^p$ , convenons de noter

$$S_H := \{M \in S : H \subset \text{Ker } M\} \quad \text{et} \quad S^D := \{M \in S : \text{Im } M \subset D\}.$$

Pour que la méthode ELC soit fructueuse, on est conduit à chercher un hyperplan vectoriel  $H$  de  $\mathbb{K}^p$  tel que  $V_H$  soit non nul et de petite dimension (dans

la situation particulière mentionnée plus haut, on prend  $H = \{0\} \times \mathbb{K}^{p-1}$  ; mais par transposition il est aussi bon de disposer d'une droite vectorielle  $D$  de  $\mathbb{K}^n$  telle que  $V^D$  soit non nulle et de petite dimension. Cela conduit naturellement à s'intéresser aux éléments de rang 1 dans  $V$ . Nous y reviendrons ultérieurement.

La deuxième sous-méthode est la méthode « **Effacer une colonne** » (en abrégé : EC) : elle s'applique lorsque l'on dispose d'un hyperplan vectoriel  $H$  de  $\mathbb{K}^p$  pour lequel  $V_H = \{0\}$ , mettons  $H = \mathbb{K}^{p-1} \times \{0\}$  pour fixer les idées. Écrivons dans ce cas toute matrice  $M$  de  $\mathcal{V}$  sous la forme

$$M = \begin{bmatrix} H(M) & [?]_{n \times 1} \end{bmatrix} \quad \text{avec } H(M) \in M_{n,p-1}(\mathbb{K}).$$

Bien sûr,  $H(\mathcal{V})$  est alors un sous-espace affine de  $M_{n,p-1}(\mathbb{K})$  de même dimension que  $\mathcal{V}$  et de rang maximal au plus  $r$ . Il est alors possible de récupérer des informations sur la structure de  $H(\mathcal{V})$  en appliquant une hypothèse de récurrence ; on est ensuite confronté à un problème de relèvement assez voisin de celui qui se pose dans la méthode ELC. Toutefois, se pose dans ce cas un problème spécifique lié à la situation où  $r = p - 1$  : c'est le problème du **relèvement spécial**. On dispose dans ce cas critique d'une fonction affine  $F : H(\mathcal{V}) \rightarrow \mathbb{K}^n$  vérifiant

$$\forall M \in \mathcal{V}, \quad M = \begin{bmatrix} H(M) & F(H(M)) \end{bmatrix}.$$

On reconnaît là une situation très voisine à celle évoquée à la fin du paragraphe 2.5.2. Dans la plupart des cas, il est alors possible de montrer que  $F$  est Im-compatible (ou pire, quasi-Im-compatible, ce qui est une version affaiblie de l'Im-compatibilité, voir le chapitre 4). En utilisant des résultats de classification des fonctions Im-compatibles, il est alors possible – sauf dans quelques cas très spécifiques où la structure de  $H(\mathcal{V})$  est en fait déjà assez simple – de trouver un vecteur non nul annulé par toutes les matrices de  $\mathcal{V}$ , et d'en déduire qu'alors  $\mathcal{V}$  est équivalent à un sous-espace de  $\mathcal{R}(0, r)$ .

Bien sûr, en transposant la méthode EC, on obtient la méthode EL (« effacer une ligne »), qu'il est inutile de décrire en détail, sinon qu'elle est pertinente lorsqu'on a trouvé une droite vectorielle  $D$  de  $\mathbb{K}^n$  telle que  $V^D = \{0\}$ .

En général, on essaie d'appliquer l'une des méthodes ELC, EC ou EL. Pour cela, il est besoin de trouver soit un hyperplan vectoriel  $H$  pour lequel  $V_H$  est de petite dimension (éventuellement nulle), soit une droite vectorielle  $D$  pour laquelle  $V^D$  est de petite dimension (éventuellement nulle). Dans [8], l'existence de tels sous-espaces vectoriels est établie à partir d'un raffinement de la méthode d'Abou-Jaoudé (couplée à quelques idées supplémentaires) ; on montre plus spécifiquement le résultat suivant :

**Lemme 2.6** (Lemme 2.5 de [8]). *Soit  $n, p$  des entiers naturels non nuls. Soit  $\mathcal{S}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  de rang maximal  $r < \min(n, p)$ , et de direction notée  $S$ . L'une des trois situations suivantes se produit alors :*

- (i) *Il existe un hyperplan vectoriel  $H$  de  $\mathbb{K}^p$  tel que  $\dim S_H \leq \frac{r-1}{2}$ .*
- (ii) *Il existe une droite vectorielle  $D$  de  $\mathbb{K}^n$  telle que  $\dim S^D \leq \frac{r-1}{2}$ .*
- (iii)  *$r$  est pair et  $\mathcal{S}$  est équivalent à l'espace de compression  $\mathcal{R}(r/2, r/2)$ .*

Nous ne dirons rien de plus sur cette méthode, sinon qu'elle nécessite à un moment le raffinement suivant du théorème de Dieudonné, qui a donné lieu à une publication séparée [7] :

**Théorème 2.7** (Théorème 3 de [7]). *Soit  $n \geq p \geq 2$  deux entiers naturels, et  $\mathcal{V}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  tel que  $\text{codim } \mathcal{V} < n - 1$ . Soit  $N$  dans la direction de  $\mathcal{V}$ , telle que  $\text{rg } N < p$ . Il existe alors  $A$  dans  $\mathcal{V}$  telle que tout élément de  $A + \mathbb{K}N$  soit de rang  $p$ .*

Dans ce résultat, la borne  $n-1$  sur la codimension est optimale, et la nécessité de prendre  $N$  de rang strictement inférieur à  $p$  est imposée par le cas éventuel d'un corps algébriquement clos (où toute droite affine dirigée par une matrice inversible contient une matrice singulière).

La méthode par récurrence évoquée plus haut nécessite d'être très méticuleux, mais elle a permis des avancées assez substantielles du point de vue des résultats : ces résultats, qui sont les meilleurs connus à ce jour, ont été démontrés dans [8] et nous les citons ci-après. Le premier généralise le théorème d'Atkinson et Lloyd aux sous-espaces affines (à l'exception du cas de la dimension critique) :

**Théorème 2.8** (Théorème 1.5 de [8]). *Soit  $\mathcal{V}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  de rang maximal au plus  $r \geq 1$ , où  $n \geq p > r$ . Supposons  $\dim \mathcal{V} > nr - (n - p + r) + 1$ . Alors,  $\mathcal{V}$  est  $r$ -décomposable.*

Le théorème précédent peut être raffiné en prenant en compte deux dimensions supplémentaires (ce qui nécessite un énorme surcroît de travail!). Pour énoncer le résultat, il faut citer un cas particulier sur les corps à 3 éléments : si  $|\mathbb{K}| = 3$ , le sous-espace affine

$$\mathcal{U}_3(\mathbb{K}) := \left\{ \begin{bmatrix} x & a & b \\ 0 & x+1 & c \\ 0 & 0 & x-1 \end{bmatrix} \mid (x, a, b, c) \in \mathbb{K}^4 \right\}$$

est en effet de rang maximal 2 et de codimension  $nr - (n - p + r)$  (ici,  $n = p = 3$  et  $r = 2$ ), mais il n'est pas 2-décomposable.

**Théorème 2.9** (Théorème 1.6 de [8]). *Soit  $n, p, r$  trois entiers naturels tels que  $n \geq p > r > 0$ . Soit  $\mathcal{V}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  de rang maximal au plus  $r$ . Supposons que  $\dim \mathcal{V} \geq nr - (n - p + r)$  et  $|\mathbb{K}| > 2$ . Alors ou bien  $\mathcal{V}$  est  $r$ -décomposable, ou bien  $|\mathbb{K}| = 3$  et  $\mathcal{V}$  est équivalent à  $\mathcal{U}_3(\mathbb{K})$ .*

Le théorème ultime porte uniquement sur les sous-espaces vectoriels. Voici son énoncé :

**Théorème 2.10** (Théorème 1.7 de [8]). *Soit  $n, p, r$  trois entiers naturels tels que  $n \geq p > r > 0$ . Soit  $V$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$  de rang maximal au plus  $r$ . Supposons que  $\dim V \geq nr - 2(n - p + r) + 2$  si  $|\mathbb{K}| > 2$ , et  $\dim V \geq nr - 2(n - p + r) + 4$  sinon. Alors :*

- ou bien  $V$  est  $r$ -décomposable ;
- ou bien  $(n, p, r) = (4, 4, 3)$ ,  $|\mathbb{K}| = 3$  et  $V$  est équivalent à l'espace

$$\mathcal{U}_4(\mathbb{K}) := \left\{ \begin{bmatrix} x & a & b & c \\ 0 & y & d & e \\ 0 & 0 & x+y & f \\ 0 & 0 & 0 & x-y \end{bmatrix} \mid (a, b, c, d, e, f, x, y) \in \mathbb{K}^8 \right\}.$$

Ces résultats sont optimaux pour le cas  $n \geq p = r + 1$ . En effet, dans ce cas et si  $r \geq 2$ , alors l'espace formé des matrices de la forme

$$\begin{bmatrix} A & [?]_{3 \times (p-3)} \\ [0]_{(n-3) \times 3} & [?]_{(n-3) \times (p-3)} \end{bmatrix}, \text{ où } A \in A_3(\mathbb{K}),$$

est de rang maximal  $r$  et de dimension  $nr - 2(n - p + r) + 1$ . Pourtant, il n'est pas  $r$ -décomposable. De même, on peut construire sur  $\mathbb{F}_2$  un sous-espace vectoriel de rang maximal  $r$  et de dimension  $nr - 2(n - p + r) + 3$  qui n'est pas  $r$ -décomposable (remplacer le bloc alterné par une matrice triangulaire supérieure de trace nulle et de taille 3).

Pour formuler des conjectures sur des dimensions encore inférieures, il faut probablement faire appel à la notion de sous-espace primitif (voir le paragraphe 2.7). En tout état de cause, il est fort probable que mes techniques aient atteint leurs limites : des idées radicalement nouvelles sont sans doute nécessaires pour aller plus loin.

## 2.6 Sous-espaces de grande dimension : cas des matrices symétriques/alternées

L'étude des sous-espaces de matrices de rang majoré a été prolongée à des sous-espaces de matrices de forme particulière. Meshulam [Mes89] a étudié le cas des sous-espaces vectoriels de matrices symétriques ou alternées. Il a obtenu une majoration optimale de la dimension d'un tel sous-espace, mais toujours sous les restrictions de cardinalité du théorème de Flanders (ni caractéristique 2, ni petits cardinaux ne sont pris en compte). Sa méthode était fondée sur l'observation d'un lien avec le problème du mariage dans un graphe<sup>2</sup> (« graph matching »). Les idées de Meshulam ont ensuite été prolongées par Radwan et Loewy [Loe94] pour obtenir la structure des sous-espaces vectoriels de dimension maximale (toujours avec les mêmes hypothèses de cardinalité), puis Loewy seul [Loe01] a élucidé en 2001 la structure des sous-espaces de dimension proche de la dimension maximale, sous les mêmes hypothèses et en se restreignant aux petits rangs.

Pour comprendre un peu mieux de quoi il retourne, il faut introduire l'équivalent des espaces de compression pour les matrices symétriques/alternées.

Soit  $n, s, t$  des entiers naturels tels que  $2s + t \leq n$ . On définit  $WS_{n,s,t}(\mathbb{K})$  (respectivement,  $WA_{n,s,t}(\mathbb{K})$ ) comme le sous-espace vectoriel constitué des matrices de  $S_n(\mathbb{K})$  (respectivement, de  $A_n(\mathbb{K})$ ) de la forme suivante :

$$\begin{bmatrix} [?]_{s \times s} & [?]_{s \times t} & [?]_{s \times (n-s-t)} \\ [?]_{t \times s} & [?]_{t \times t} & [0]_{t \times (n-s-t)} \\ [?]_{(n-s-t) \times s} & [0]_{(n-s-t) \times t} & [0]_{(n-s-t) \times (n-s-t)} \end{bmatrix}.$$

On vérifie que

$$\text{rgmax } WS_{n,s,t}(\mathbb{K}) = 2s + t$$

et, si  $t$  est *impair*, que

$$\text{rgmax } WA_{n,s,t}(\mathbb{K}) = 2s + t - 1.$$

---

2. Après mes travaux sur la question, Meshulam a amélioré sa propre technique pour éliminer toute hypothèse sur le cardinal du corps de base, voir [Mes17].

On pose

$$s_{n,s,t} := \dim \text{WS}_{n,s,t}(\mathbb{K}) \quad \text{et} \quad a_{n,s,t} := \dim \text{WA}_{n,s,t}(\mathbb{K}),$$

et l'on vérifie que

$$s_{n,s,t} = \binom{s+1}{2} + \binom{t+1}{2} + s(n-s) \quad \text{et} \quad a_{n,s,t} = \binom{s}{2} + \binom{t}{2} + s(n-s).$$

Fixons deux entiers  $n, r$  tels que  $0 \leq r \leq n$ . Les suites  $(s_{n,s,r-2s})_{0 \leq s \leq \lfloor r/2 \rfloor}$  et  $(a_{n,s,r+1-2s})_{0 \leq s \leq \lfloor (r+1)/2 \rfloor}$  sont alors strictement convexes. Lorsque  $r$  est petit, elles prennent leur valeur maximale au dernier indice, lorsque  $r$  est proche de  $n$  elles prennent leur valeur maximale à l'indice 0.

La relation pertinente pour étudier les sous-espaces de matrices symétriques/alternées est celle de congruence : deux sous-ensembles  $\mathcal{A}$  et  $\mathcal{B}$  de  $M_n(\mathbb{K})$  sont congrus lorsqu'il existe  $P \in \text{GL}_n(\mathbb{K})$  telle que  $P^T \mathcal{A} P = \mathcal{B}$ , autrement dit lorsque  $\mathcal{A}$  et  $\mathcal{B}$  représentent le même ensemble de formes bilinéaires, dans des bases *a priori* différentes.

Dans [4], nous avons commencé par améliorer les résultats de Meshulam, Loewy et Radwan en les généralisant à n'importe quel corps. Y furent démontrés les deux résultats principaux qui suivent :

**Théorème 2.11** (Théorème 1.3 de [4]). *Soit  $n$  et  $s$  deux entiers naturels tels que  $2s < n$ , et  $S$  un sous-espace vectoriel de  $A_n(\mathbb{K})$  tel que  $\text{rgmax } S \leq 2s$ . Alors,*

$$\dim S \leq \max(a_{n,0,2s+1}, a_{n,s,1}),$$

*et en cas d'égalité  $S$  est congru à  $\text{WA}_{n,s,1}(\mathbb{K})$  ou à  $\text{WA}_{n,0,2s+1}(\mathbb{K})$ .*

**Théorème 2.12** (Théorème 1.4 de [4]). *Soit  $n$  et  $r$  deux entiers naturels tels que  $r < n$ . Soit  $S$  un sous-espace vectoriel de  $S_n(\mathbb{K})$  tel que  $\text{rgmax } S \leq r$ .*

(a) *Si  $r$  est pair alors pour  $s := r/2$  on a*

$$\dim S \leq \max(s_{n,0,r}, s_{n,s,0}),$$

*et en cas d'égalité ou bien  $S$  est congru à  $\text{WS}_{n,s,0}(\mathbb{K})$  ou à  $\text{WS}_{n,0,r}(\mathbb{K})$ , ou bien  $\mathbb{K}$  est de caractéristique 2 et  $S$  est congru à  $\text{WA}_{n,0,r+1}(\mathbb{K})$ .*

(b) *Si  $r$  est impair alors pour  $s := (r-1)/2$  on a*

$$\dim S \leq \max(s_{n,0,r}, s_{n,s,1}),$$

*et en cas d'égalité  $S$  est congru à  $\text{WS}_{n,s,1}(\mathbb{K})$  ou à  $\text{WS}_{n,0,r}(\mathbb{K})$ .*

Encore mieux, sous les mêmes hypothèses nous avons classifié les sous-espaces affines de rang majoré (toujours dans [4]). De nombreux exemples spéciaux apparaissent lorsque  $\mathbb{K}$  est de cardinal 2 : nous ne les exposerons pas ici. Dans [5], nous sommes allés encore plus loin et avons abouti aux deux théorèmes suivants, qui sont l'équivalent, pour le problème des sous-espaces de matrices symétriques/alternées de rang majoré, du théorème d'Atkinson et Lloyd sur les sous-espaces de matrices carrées de rang majoré.

**Théorème 2.13** (Théorème 1.4 de [5]). *Soit  $S$  un sous-espace vectoriel de  $A_n(\mathbb{K})$ , et  $r = 2s$  un entier pair tel que  $0 < r < n$ . Supposons que*

$$\dim S > \max(a_{n,1,r-1}, a_{n,s-1,3}) \quad \text{et} \quad \text{rgmax } S \leq r.$$

*Alors,  $S$  est congru à un sous-espace de  $WA_{n,0,r+1}(\mathbb{K})$  ou de  $WA_{n,s,1}(\mathbb{K})$ .*

**Théorème 2.14** (Théorème 1.3 de [5]). *Supposons  $|\mathbb{K}| > 2$ . Soit  $S$  un sous-espace vectoriel de  $S_n(\mathbb{K})$ , et  $r$  un entier tel que  $1 < r < n$ . Écrivons  $r = 2s + \varepsilon$  pour un entier  $s > 0$  et un  $\varepsilon \in \{0, 1\}$ . Supposons que*

$$\dim S > \max(s_{n,1,r-2}, s_{n,s-1,2+\varepsilon}) \quad \text{et} \quad \text{rgmax } S \leq r.$$

*Alors, l'une des situations suivantes se présente :*

- (i)  *$S$  est congru à un sous-espace de  $WS_{n,0,r}(\mathbb{K})$  ;*
- (ii)  *$S$  est congru à un sous-espace de  $WS_{n,s,\varepsilon}(\mathbb{K})$  ;*
- (iii)  *$\mathbb{K}$  est de caractéristique 2,  $r$  est pair et  $S$  est congru à un sous-espace de  $WA_{n,0,r+1}(\mathbb{K})$ .*

Le cas du corps à deux éléments est beaucoup plus compliqué, en particulier à cause de l'échec d'un lemme critique. Il semble nécessaire, pour le traiter, d'élargir l'énoncé à des sous-espaces affines, ce qui promet d'être nettement plus coûteux : j'avoue ne pas avoir eu le courage de me lancer dans cette entreprise jusqu'à maintenant.

Par comparaison, Loewy se limitait dans [Loe01] au cas des matrices symétriques, aux corps de caractéristique différente de 2, et imposait d'avoir  $2n > 5s + 1$  et  $|\mathbb{K}| \geq 2s + 3$  !

Les techniques qui m'ont permis d'obtenir les résultats indiqués plus haut sont fondées sur une adaptation de ma méthode itérative déjà expliquée pour le cas des matrices rectangulaires (section 2.5.3), et en réalité c'est l'étude du cas symétrique/alterné qui m'a obligé à inventer cette méthode. Ses caractéristiques essentielles sont :

- une récurrence sur le format des matrices ;
- la nécessité, pour démontrer les théorèmes 2.13 et 2.14, de disposer d'une version des théorèmes 2.11 et 2.12 pour les sous-espaces *affines* de matrices ;
- le lien avec les morphismes Im-compatibles sur des espaces de matrices symétriques/alternées (voir le chapitre 4 et en particulier la section 4.6) pour montrer qu'un sous-espace est congru à une partie de  $WA_{n,0,r+1}(\mathbb{K})$  ou de  $WS_{n,0,r}(\mathbb{K})$ .

Au-delà du simple lien thématique, il y a une dépendance profonde entre le théorème d'Atkinson-Lloyd pour les matrices rectangulaires et les problèmes considérés plus haut : ce lien surgit lorsque l'on cherche à établir qu'un sous-espace  $S$  est congru à un sous-espace de  $WA_{n,s,1}(\mathbb{K})$  ou de  $WA_{n,s,\varepsilon}(\mathbb{K})$ . Mettons que l'on dispose d'un sous-espace  $S$  de  $A_n(\mathbb{K})$  tel que  $\text{rgmax } S \leq r$ , et que l'on ait réussi à ramener après congruence la situation à celle où toute matrice  $M$  de  $S$  s'écrit par blocs sous la forme suivante :

$$M = \begin{bmatrix} [?]_{s \times s} & -B(M)^T & [?]_{s \times 1} \\ B(M) & [0]_{(n-s-1) \times (n-s-1)} & C(M) \\ [?]_{1 \times s} & -C(M)^T & 0 \end{bmatrix}$$

où  $B(M) \in M_{n-s-1,s}(\mathbb{K})$  et  $C(M) \in \mathbb{K}^{n-s-1}$ . Alors, pour n'importe quelle matrice  $M$  dans  $S$ , on trouve

$$\text{rg } M \geq \text{rg } A(M) + \text{rg } (-A(M)^T)$$

où

$$A(M) := \begin{bmatrix} B(M) & C(M) \end{bmatrix} \in M_{n-s-1,s+1}(\mathbb{K}).$$

Le rang maximal dans  $A(S)$  est donc au plus  $s$ . Comme  $n - s - 1 \geq s + 1$ , si la dimension de  $A(S)$  est assez grande, alors le théorème d'Atkinson-Lloyd permet de trouver un vecteur  $X \in \mathbb{K}^s$  tel que  $C(M) = B(M)X$  pour tout  $M \in S$ , et à partir de là une transvection symétrique permet de conclure que  $S$  est congru à un sous-espace vectoriel de  $\text{WA}_{n,s,1}(\mathbb{K})$ .

## 2.7 Sous-espaces vectoriels primitifs de rang majoré

Développée en premier lieu par Atkinson et Lloyd [Atk80a], et étudiée plus tard par Eisenbud et Harris avec des méthodes de géométrie algébrique [Eis88], la notion de sous-espace primitif de rang majoré est un point de vue sur la question qui se détache nettement de celui considéré dans les paragraphes précédents.

L'idée générale est qu'un espace de rang majoré est primitif lorsqu'il ne peut se déduire, par un procédé de complétion élémentaire, d'espaces de matrices de rang majoré ayant moins de lignes ou de colonnes. Plus précisément, un espace  $\mathcal{V}$  de matrices de rang maximal  $r$  est dit **primitif** lorsque toutes les conditions suivantes sont réunies :

- (i)  $\mathcal{V}$  est inéquivalent à un sous-espace de  $\mathcal{R}(0, p - 1)$  ;
- (ii)  $\mathcal{V}$  est inéquivalent à un sous-espace de  $\mathcal{R}(n - 1, 0)$  ;
- (iii)  $\mathcal{V}$  est inéquivalent à un sous-espace  $\mathcal{T}$  dans lequel toute matrice s'écrit  $M = \begin{bmatrix} H(M) & [?]_{n \times 1} \end{bmatrix}$  et le rang maximal dans  $H(\mathcal{T})$  est  $r - 1$  ;
- (iv)  $\mathcal{V}$  est inéquivalent à un sous-espace  $\mathcal{T}$  dans lequel toute matrice s'écrit  $M = \begin{bmatrix} H(M) \\ [?]_{1 \times p} \end{bmatrix}$  et le rang maximal dans  $H(\mathcal{T})$  est  $r - 1$ .

On dit aussi que  $\mathcal{V}$  est :

- **réduit en colonne** lorsqu'il vérifie la condition (i) ;
- **réduit en ligne** lorsqu'il vérifie la condition (ii) ;
- **réduit** lorsqu'il vérifie les conditions (i) et (ii) ;
- **semi-primitif** lorsqu'il vérifie les conditions (i), (ii) et (iii).

On dit enfin que  $\mathcal{V}$  est **défectueux** lorsque  $\text{rgmax}(\mathcal{V}) < p$ . Et l'on vérifie facilement que tout espace semi-primitif est défectueux.

L'exemple fondamental suivant a été introduit par Atkinson [Atk83] (dans sa version transposée) : donnons-nous deux espaces vectoriels  $U$  et  $V$  de dimensions finies respectivement notées  $p$  et  $n$ , et prenons une surjection linéaire

$$f : U \wedge U \rightarrow V.$$

On choisit des bases respectives  $\mathbf{B}$  et  $\mathbf{C}$  de  $U$  et  $V$ . On introduit

$$\mathcal{S}_f := \{y \mapsto f(x \wedge y) \mid x \in U\},$$



sous-espace vectoriel de  $\mathcal{L}(U, V)$ , et l'on note  $\mathcal{M}$  l'espace de matrices représentant  $\mathcal{S}_f$  dans les bases  $\mathbf{B}$  et  $\mathbf{C}$ .

Tous les éléments de  $\mathcal{S}_f$  sont non injectifs, donc  $\mathcal{M}$  est défectueux. De plus, la surjectivité de  $f$  garantit que  $\mathcal{M}$  vérifie la condition (ii). Enfin, si  $n > 1 + \frac{(p-1)(p-2)}{2}$  alors on démontre (voir le corollaire 4.10 de [3]) que  $\mathcal{M}$  vérifie les conditions (i) et (iii), et même que  $\mathcal{M}$  est primitif (proposition 4.14 de [3]). En particulier, le cas  $f = \text{id}_{U \wedge U}$  montre l'existence d'un sous-espace primitif de  $M_{\binom{p}{2}, p}(\mathbb{K})$  de rang maximal  $p - 1$ .

Nous dirons qu'un espace vectoriel de matrices  $\mathcal{M}$  est **de type alterné** lorsqu'il représente un espace de type  $\mathcal{S}_f$  dans certaines bases.

Atkinson a établi une forme de réciproque, qui est le théorème majeur sur les sous-espaces semi-primitifs :

**Théorème 2.15** (Atkinson (1983)). *Soit  $\mathcal{M}$  un sous-espace semi-primitif de  $M_{n,p}(\mathbb{K})$ , de rang maximal  $r$ . On suppose  $|\mathbb{K}| > r$ . Alors :*

(a) *On a  $n \leq \binom{r+1}{2}$ .*

(b) *Si en outre  $n > 1 + \binom{r}{2}$ , alors  $r = p - 1$  et  $\mathcal{M}$  est de type alterné.*

Signalons qu'Atkinson démontre ce résultat pour les sous-espaces primitifs mais que la démonstration s'adapte sans problème aux sous-espaces semi-primitifs ; signalons aussi que c'est la version transposée qui est établie par Atkinson (où l'on inverse les nombres de lignes et de colonnes ; les calculs effectués par Atkinson se déroulent plus naturellement dans la version que nous exposons ici).

En outre, Atkinson a mis à profit les mêmes techniques dans [Atk83] pour élucider la structure des sous-espaces primitifs de rang maximal 3, mais uniquement sur un corps algébriquement clos.

Le théorème d'Atkinson sur les sous-espaces primitifs fait indubitablement penser au théorème des mêmes Atkinson et Lloyd sur les sous-espaces de rang majoré de grande dimension (voir le théorème 2.4), mais ici ce n'est pas la dimension de l'espace  $\mathcal{M}$  qui est en cause, c'est le format des matrices de  $\mathcal{M}$  ! On verra dans la section 6.6 que le théorème d'Atkinson a une application spectaculaire à la structure des sous-espaces vectoriels de matrices à spectre trivial (avec application immédiate aux sous-espaces vectoriels de matrices nilpotentes).

Notre motivation première pour étudier les sous-espaces semi-primitifs résidait dans leur lien avec la problématique de la dépendance linéaire locale. Nous repoussons les explications à ce sujet au chapitre 3, et nous nous contenterons ici de signaler que le théorème d'Atkinson a une version duale en termes d'espaces d'opérateurs localement liés.

Dans la même optique que notre travail sur les sous-espaces de grande dimension, nous nous sommes posés deux questions :

- (i) Peut-on se dispenser de la condition de cardinalité sur le corps  $\mathbb{K}$  ?
- (ii) Dans quelle mesure peut-on prolonger le résultat d'Atkinson à des sous-espaces de matrices dont le nombre de lignes est encore inférieur à la borne du théorème 2.15 ?

L'étude des matrices de petite taille montre que les sous-espaces semi-primitifs de format  $\frac{p(p-1)}{2} \times p$  ne sont pas tous, sur un corps arbitraire, de la forme indiquée

dans le théorème d'Atkinson et Lloyd. En particulier, nous avons complètement classifié, sur le corps  $\mathbb{F}_2$ , les sous-espaces vectoriels primitifs de matrices de rang au plus 2 en usant de méthodes élémentaires d'extraction [9]. Nous avons notamment mis en évidence plusieurs sous-espaces primitifs de  $M_3(\mathbb{F}_2)$  deux à deux inéquivalents et ne représentant pas  $\text{id}_{\mathbb{K}^3 \wedge \mathbb{K}^3}$  (certains étaient déjà connus). En revanche, il est remarquable que la majoration  $n \leq \frac{r(r+1)}{2}$  reste valable dans le cas particulier  $r = 2$  sur  $\mathbb{F}_2$  ! Nous conjecturons que la partie « majoration » du théorème reste valable sur un corps quelconque ; cette conjecture semble robuste au vu des applications du théorème – voir la section 6.6 – mais à ce jour nous n'avons trouvé aucune piste pour l'attaquer.

Il est temps de préciser quelques idées-clefs pour le théorème d'Atkinson. D'abord, étant donné  $(s, t) \in \llbracket 0, n \rrbracket \times \llbracket 0, p \rrbracket$ , on dit que  $\mathcal{M}$  est  $(s, t)$ -décomposé lorsque toute matrice dans  $\mathcal{M}$  est de la forme

$$M = \begin{bmatrix} [?] & C(M) \\ B(M) & [0] \end{bmatrix}$$

où  $B(M)$  est de format  $(n - s) \times t$  et  $C(M)$  de format  $s \times (p - t)$ . Dans ce cas,

$$\text{rgmax}(B(\mathcal{M})) + \text{rgmax}(C(\mathcal{M})) \leq \text{rgmax } \mathcal{M},$$

et cette propriété utilise de manière cruciale l'hypothèse  $|\mathbb{K}| > \text{rgmax } \mathcal{M}$ . Atkinson fait un usage systématique des matrices *génériques* : considérons un anneau  $A$  de polynômes à coefficients dans  $\mathbb{K}$  (et à nombre fini d'indéterminées). Une **matrice générique** de  $\mathcal{M}$  est une matrice de  $M_{n,p}(A)$  dont les coefficients sont des éléments homogènes de degré 1 de  $A$  et telle que  $\mathcal{M}$  soit l'espace des matrices obtenues en spécialisant dans  $\mathbb{K}$  toutes les indéterminées. En général, le rang d'une telle matrice générique est au moins  $\text{rgmax } \mathcal{M}$ , et il y a égalité si  $|\mathbb{K}| > \text{rgmax } \mathcal{M}$ . Une partie de la démonstration d'Atkinson repose sur des calculs directs sur les matrices génériques. En particulier, le lemme suivant, qui prolonge un résultat de Flanders (voir le paragraphe 2.5.1) est très utile :

**Lemme 2.16** (Flanders-Atkinson). *Soit  $\mathcal{S}$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$ .*

*On suppose que  $\mathcal{S}$  contient  $J_r := \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$  et que  $\text{rgmax } \mathcal{S} = r < |\mathbb{K}|$ . Alors, pour tout élément  $M \in \mathcal{S}$ , que l'on décompose*

$$M = \begin{bmatrix} A(M) & C(M) \\ B(M) & D(M) \end{bmatrix} \quad \text{où } A(M) \in M_r(\mathbb{K}), \text{ etc,}$$

*les identités suivantes sont vraies :*

$$D(M) = 0 \quad \text{et} \quad \forall k \in \mathbb{N}, B(M)A(M)^k C(M) = 0.$$

La démonstration percutante qui suit figure dans notre premier article sur les espaces d'opérateurs localement liés<sup>3</sup>.

*Démonstration.* Puisque  $r < |\mathbb{K}|$ , la contrainte de rang survit à toute extension des scalaires. Prenons donc le corps des fractions  $\mathbb{L}$  de l'anneau des séries formelles  $\mathbb{K}[[s]]$  en une indéterminée  $s$ . Soit  $M \in \mathcal{S}$ . On observe que

$$J_r - sM = \begin{bmatrix} I_r - sA(M) & -sC(M) \\ -sB(M) & -sD(M) \end{bmatrix}.$$

<sup>3</sup>. C. de Seguins Pazzis, Local linear dependence seen through duality I. J. Pure Appl. Algebra **219** (2015) 2144–2188.

Or  $I_r - sA(M)$  est inversible, d'inverse  $\sum_{k=0}^{+\infty} s^k A(M)^k$ . Puisqu'en outre  $\text{rg}(J_r - sM) \leq r$ , un pivot donne  $-sD(M) = -sB(M)(I_r - sA(M))^{-1}(-sC(M))$ . Les identités annoncées s'obtiennent alors par unicité des coefficients du développement en série formelle.  $\square$

Du lemme précédent découle que sous l'hypothèse habituelle de cardinalité,  $\mathcal{M}$  est toujours équivalent à un sous-espace  $(r, r)$ -décomposé. Atkinson applique ensuite son lemme, sous ses hypothèses, à une matrice générique de  $\mathcal{M}$  et obtient des identités exploitables pour étudier les sous-espaces semi-primitifs ayant un grand nombre de lignes.

En poussant plus loin les techniques de matrices génériques qu'Atkinson avait développées, je suis parvenu à l'amélioration suivante de son résultat :

**Théorème 2.17** (Théorème 5.13 de [3]). *Soit  $\mathcal{M}$  un sous-espace semi-primitif de  $M_{n,p}(\mathbb{K})$ , de rang maximal  $r > 0$ . On suppose  $|\mathbb{K}| > r$  et  $n > 3 + \binom{r-1}{2}$ . Alors, ou bien  $\mathcal{M}$  est de type alterné, ou bien il est équivalent à un espace  $\mathcal{M}'$  où toute matrice est de la forme*

$$M = \begin{bmatrix} [?]_{1 \times r} & [?]_{1 \times (p-r)} \\ H(M) & [0]_{(n-1) \times (p-r)} \end{bmatrix}$$

et  $H(\mathcal{M}')$  est de type alterné et de rang maximal  $r - 1$ .

J'ai enfin classifié, sur n'importe quel corps ayant plus de trois éléments (c'est la contrainte incontournable pour appliquer les techniques précédentes) les sous-espaces semi-primitifs de rang maximal 3 (voir la section 5.3 de [3]), ce qu'Atkinson, Eisenbud et Harris n'avaient fait que pour les corps algébriquement clos. En s'appuyant sur les théorèmes précédents, ainsi que sur une classification rigoureuse des espaces de type alterné ayant un grand nombre de lignes, on se réduit à la classification à congruence près des sous-espaces vectoriels de dimension 4, 5 ou 6 de  $A_4(\mathbb{K})$ , ce qui s'obtient grâce à la classification à congruence près des sous-formes du Pfaffien. Ce faisant, nous en avons profité pour corriger une erreur dans un article de Gauger<sup>4</sup> cité par Atkinson : cette erreur est liée à la classification à congruence près des sous-formes de rang 3 du Pfaffien, et doit sans aucun doute être mise sur le compte du fait que Gauger n'avait pas conscience d'un tel lien.

## 2.8 Perspectives

À mon sens, les méthodes évoquées dans ce chapitre ont atteint leurs limites. La longueur des articles [8] et [5], la nécessité de développements de moins en moins fluides sur les morphismes Im-compatibles (notamment celui sur les morphismes quasi-Im-compatibles) doit convaincre qu'une nouvelle révolution des méthodes est nécessaire pour aller plus loin.

Sur l'approche de Dieudonné-Flanders etc, un point mineur a été laissé en suspens et je m'en occuperai certainement à terme : les espaces de matrices

---

4. M.A. Gauger, On the classification of metabelian Lie algebras. Trans. Amer. Math. Soc. **179** (1973) 293–329.

hermitiennes. En effet, dans une approche classique cette étude doit accompagner celle de sous-espaces de matrices symétriques ou alternées. Il faudra donc étudier les morphismes Im-compatibles correspondants, seulement considérer des espaces vectoriels sur le sous-corps des scalaires autoadjoints (semi-sous-espaces vectoriels). Mais on ne verra là que le prolongement de stratégies bien rodées, avec peut-être quelques surprises ici ou là, comme des contre-exemples inattendus.

Quant aux sous-espaces primitifs, il y a probablement encore de la marge pour aller plus loin sans technicité supplémentaire excessive. Dans des travaux non encore publiés, j'ai démontré qu'avec les techniques d'Atkinson, on pouvait encore décrire les sous-espaces semi-primitifs de  $M_{n,p}(\mathbb{K})$  de rang au plus  $r$  lorsque  $n > 6 + \binom{r-2}{2}$  et  $r > 2$ . Mais la question la plus intéressante qui reste en suspens sur ces espaces n'est pas de cet ordre : c'est celle de l'extension à un corps quelconque de l'énoncé sur le nombre maximal de lignes dans un sous-espace semi-primitif. La résolution de cette question, que j'appelle **conjecture d'Atkinson**, serait probablement porteuse d'applications fructueuses au vu des liens entre les sous-espaces primitifs et les espaces de matrices à spectre trivial. Hélas, à ce jour je n'ai strictement aucune piste pour m'y attaquer.

## Chapitre 3

# Dépendance linéaire locale et réflexivité algébrique

### 3.1 Articles concernés

- [1] C. de Seguins Pazzis, Local linear dependence seen through duality I. *J. Pure Appl. Algebra* **219** (2015) 2144–2188.
- [2] C. de Seguins Pazzis, Local linear dependence seen through duality II. *Linear Algebra Appl.* **462** (2014) 133–185.
- [3] C. de Seguins Pazzis, LDB division algebras. *J. Algebra* **423** (2015) 259–288.
- [4] C. de Seguins Pazzis, On the minimal rank in non-reflexive operator spaces over finite fields. *Linear Algebra Appl.* **462** (2014) 126–132.
- [5] C. de Seguins Pazzis, Primitive spaces of matrices with upper rank two over the field with two elements. *Linear Multilinear Algebra* **64** (2016) 1321–1353.

### 3.2 Bibliographie historique

- [Ami65] S.A. Amitsur, Generalized polynomial identities and pivotal monomials. *Trans. Amer. Math. Soc.* **114** (1965) 210–226.
- [Lar88] D.A. Larson, Reflexivity, algebraic reflexivity and linear interpolation. *Amer. J. Math.* **283** (1988) 283–299.
- [Din96] L. Ding, On a pattern of reflexivity of operator spaces. *Proc. Amer. Math. Soc.* **124** (1996) 3101–3108.
- [Bre99] M. Brešar, P. Šemrl, On locally linearly dependent operators and derivations. *Trans. Amer. Math. Soc.* **351** (1999) 1257–1275.
- [Mes02] R. Meshulam, P. Šemrl, Locally linearly dependent operators. *Pacific J. Math.* **203-2** (2002) 441–459.
- [Mes04] R. Meshulam, P. Šemrl, Locally linearly dependent operators and reflexivity of operator spaces. *Linear Algebra Appl.* **383** (2004) 143–150.
- [Mes07] R. Meshulam, P. Šemrl, Minimal rank and reflexivity of operator spaces. *Proc. Amer. Math. Soc.* **135** (2007) 1839–1842.
- [Che08] M. Chebotar, P. Šemrl, Minimal locally linearly dependent spaces of operators. *Linear Algebra Appl.* **429** (2008) 887–900.

## 3.3 Problématique et histoire

### 3.3.1 Principales définitions

Soit  $U$  et  $V$  deux espaces vectoriels, et  $f_1, \dots, f_n$  des applications linéaires de  $U$  vers  $V$ . On dit que la famille  $(f_1, \dots, f_n)$  est **localement liée** (en anglais : *locally linearly dependent*, abrégé en LLD) lorsque la famille de vecteurs  $(f_1(x), \dots, f_n(x))$  est liée pour tout  $x \in U$ .

Il y a deux situations évidentes, dites triviales, où la famille  $(f_1, \dots, f_n)$  est localement liée :

- $(f_1, \dots, f_n)$  est liée dans l'espace vectoriel  $\mathcal{L}(U, V)$  ;
- l'espace  $V$  est de dimension strictement inférieure à  $n$ .

Un exemple d'apparition de la liaison locale figure dans la caractérisation classique des homothéties d'un espace vectoriel  $U$  : cette dernière indique en effet qu'est une homothétie tout endomorphisme  $u$  de  $U$  tel que  $(\text{id}_U, u)$  soit localement liée. Classiquement, ce résultat se généralise comme suit :

**Théorème 3.1.** *Soit  $(f, g)$  un couple localement lié d'applications linéaires de  $U$  vers  $V$ . Ou bien  $(f, g)$  est liée, ou bien  $f$  et  $g$  sont de même image et de rang 1.*

En revanche, un exemple simple suffit à se convaincre qu'il existe des familles d'opérateurs localement liées qui ne vérifient aucune des conditions de trivialité. Raisonnons matriciellement et prenons une base  $(A_1, A_2, A_3)$  de l'espace  $A_3(\mathbb{K})$  des matrices antisymétriques de taille 3. Pour tout  $X \in \mathbb{K}^3 \setminus \{0\}$ , la famille  $(A_1X, A_2X, A_3X)$  ne peut engendrer  $\mathbb{K}^3$  car tous ses vecteurs sont dans l'orthogonal de  $X$  pour la forme bilinéaire symétrique standard  $(Y, Z) \mapsto Y^T Z$ . Ainsi,  $(A_1, A_2, A_3)$  est localement liée. Et clairement  $\text{Im } A_1 + \text{Im } A_2 + \text{Im } A_3 = \mathbb{K}^3$ , si bien que  $(A_1, A_2, A_3)$  n'est pas un triplet localement lié trivial.

Le bon point de vue pour aborder la question est en réalité celui des espaces d'opérateurs. Reprenons une famille  $(f_1, \dots, f_n) \in \mathcal{L}(U, V)^n$ . Si elle est libre, la liaison locale se traduit par le fait que tout  $x \in U$  est annulé par un élément non nul de  $\text{Vect}(f_1, \dots, f_n)$ , autrement dit que ce dernier espace est localement lié au sens de la définition suivante :

**Définition 2.** Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$ . On dit que  $\mathcal{S}$  est **localement lié** lorsque

$$\forall x \in U, \exists f \in \mathcal{S} \setminus \{0\} : f(x) = 0.$$

Le problème est ainsi reporté sur la compréhension des espaces d'opérateurs localement liés, ce qui s'avère être le point de vue le plus efficace.

Avant de mieux expliquer les motivations qui ont conduit à s'intéresser à ces objets, il est utile d'aborder une notion assez voisine, qui est celle de réflexivité algébrique.

**Définition 3.** Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$ . La clôture réflexive de  $\mathcal{S}$ , notée  $\mathcal{R}(\mathcal{S})$ , est l'ensemble des applications linéaires  $g : U \rightarrow V$  telles que

$$\forall x \in U, \exists f \in \mathcal{S} : g(x) = f(x).$$

La clôture réflexive de  $\mathcal{R}(\mathcal{S})$  est systématiquement un sous-espace vectoriel de  $\mathcal{L}(U, V)$  incluant  $\mathcal{S}$ , et l'on démontre facilement que  $\mathcal{R}(\mathcal{R}(\mathcal{S})) = \mathcal{R}(\mathcal{S})$ . On dit que  $\mathcal{S}$  est **(algébriquement) réflexif** lorsqu'il est égal à sa clôture réflexive, autrement dit lorsque, pour une application linéaire  $g : U \rightarrow V$ , la propriété  $\forall x \in U, \exists f \in \mathcal{S} : g(x) = f(x)$  implique  $\exists f \in \mathcal{S} : \forall x \in U, g(x) = f(x)$  (cela peut donc être vu comme un problème d'interversion de quantificateurs).

Un lien essentiel existe entre les deux notions : supposons que  $\mathcal{S}$  ne soit *pas* réflexif, et prenons un élément  $g \in \mathcal{R}(\mathcal{S}) \setminus \mathcal{S}$ . L'espace  $\mathcal{S} \oplus \mathbb{K}g$  est alors localement lié : en effet, pour tout  $x \in U$ , on peut trouver  $f_x \in \mathcal{S}$  tel que  $g(x) = f_x(x)$ , et alors  $x$  est annulé par l'opérateur non nul  $g - f_x$ . Ainsi,  $\mathcal{S}$  apparaît comme un hyperplan vectoriel d'un espace d'opérateurs localement lié. Attention, il n'est pas vrai réciproquement que tout hyperplan d'un espace d'opérateurs localement lié soit non réflexif. Prenons par exemple l'espace  $\mathcal{S} = \mathcal{L}(\mathbb{K}^2, \mathbb{K})$ , et la première projection canonique  $\pi_1$  de  $\mathbb{K}^2$  dans  $\mathbb{K}$ . On montre facilement que  $\mathbb{K}\pi_1$  est réflexif (c'est plus généralement le cas pour n'importe quelle droite vectorielle d'opérateurs), en revanche  $\mathcal{S}$  est évidemment localement lié (il vérifie la seconde condition de trivialité).

### 3.3.2 Motivation et histoire des principaux résultats

On a déjà fait observer que la notion de liaison locale est une généralisation naturelle de la caractérisation des homothéties d'un espace vectoriel (comme endomorphismes  $u$  tels que  $(x, u(x))$  soit liée pour tout  $x$ ). Kaplansky remarqua aussi, grâce à la réduction de Frobenius, que pour un endomorphisme  $u$  d'un espace vectoriel de dimension finie, la liaison locale de la famille  $(\text{id}, u, u^2, \dots, u^n)$  implique sa liaison.

Il semble qu'il faille remonter à Amitsur pour trouver une trace d'étude un peu plus systématique des familles localement liées. Dans [Ami65], il montra qu'obtenir des éléments non nuls de petit rang dans un espace d'opérateurs localement lié permettait d'obtenir des résultats sur les anneaux vérifiant des identités polynomiales généralisées (précisément, les opérateurs en question sont des morphismes de groupes abéliens, le groupe d'arrivée étant enrichi en un espace vectoriel à gauche sur un corps gauche). Amitsur montra en particulier que, dans ce contexte, un espace d'opérateurs localement lié contient nécessairement un élément non nul de rang au plus  $\binom{n+1}{2} - 1$ .

Deux décennies plus tard, Larson étudia la propriété de réflexivité algébrique [Lar88], démontra notamment que tout espace d'opérateurs localement lié de dimension finie possède un élément non nul de rang fini, et en déduisit que tout espace d'opérateurs localement lié minimal de dimension finie est constitué d'éléments de rang fini (ce qui permet de réduire entièrement la détermination de ces espaces à celle d'opérateurs entre espaces vectoriels de dimension finie).

Encore plus récemment, Brešar et Šemrl [Bre99] furent confrontés, pour classer les couples commutants  $(d, g)$  de dérivations continues d'une algèbre de Banach  $\mathcal{A}$  telles que  $(dg)(x)$  soit quasi-nilpotent pour tout vecteur  $x$  de  $\mathcal{A}$ , de déterminer les triplets  $(f_1, f_2, f_3)$  localement liés. Ils améliorèrent en outre la borne trouvée par Amitsur, en montrant que sur un corps infini toute  $n$ -liste localement liée  $(f_1, \dots, f_n)$  possède une combinaison linéaire non triviale de rang strictement inférieur à  $n$ . Meshulam et Šemrl montrèrent ultérieurement [Mes07] que ce dernier résultat vaut aussi pour les corps finis, en usant d'un argument

de dénombrement. Étudiant l’optimalité du résultat, ils démontrèrent même le théorème suivant :

**Théorème 3.2** (Meshulam et Šemrl (2002)). *Soit  $\mathcal{S}$  un sous-espace localement lié de  $\mathcal{L}(U, V)$ , de dimension  $n > 0$ . On suppose  $|\mathbb{K}| \geq n + 2$ . Alors, ou bien  $\mathcal{S}$  contient un élément non nul de rang strictement inférieur à  $n - 1$ , ou bien tous les éléments non nuls de  $\mathcal{S}$  sont de rang  $n - 1$ .*

Des problèmes similaires de rang furent étudiés pour les espaces d’opérateurs non réflexifs. L’idée est, en trouvant des contraintes sur ces derniers, d’obtenir des conditions suffisantes simples pour qu’un espace d’opérateur soit réflexif. On peut faire remonter à Ding [Din96] le premier résultat quantitatif significatif : il démontra que tout espace d’opérateurs non réflexif de dimension  $n$  possède un élément non nul de rang inférieur à  $n^2$ . Pour les corps de cardinal suffisamment grand, cette borne fut nettement améliorée par Meshulam et Šemrl dans [Mes04], en prolongement de leurs méthodes sur les espaces localement liés :

**Théorème 3.3** (Meshulam et Šemrl (2004)). *Soit  $\mathcal{S}$  un sous-espace non réflexif de  $\mathcal{L}(U, V)$ , de dimension  $n > 0$ . On suppose  $|\mathbb{K}| > n + 2$ . Alors,  $\mathcal{S}$  contient un élément non nul de rang inférieur à  $2n - 2$ .*

Par suite, et sous l’hypothèse de cardinalité de ce théorème, est réflexif tout espace vectoriel d’opérateurs  $\mathcal{S}$  de dimension  $n$  dont tous les éléments non nuls sont de rang strictement supérieur à  $2n - 2$ .

Meshulam et Šemrl améliorèrent même le résultat précédent sur un corps algébriquement clos, démontrant que le majorant  $2n - 2$  peut y être remplacé par  $n$ , et ils postulèrent que le résultat valait sur n’importe quel corps infini.

Les espaces localement liés minimaux de dimension finie furent étudiés par Chebotar et Šemrl dans [Che08], tentant ainsi de prolonger les résultats obtenus par Brešar et Šemrl dans [Bre99] sur les triplets. Ils obtinrent que pour un tel espace minimal de dimension  $n$ , et sous contrainte que le corps de base soit de cardinal assez grand, la somme des images des opérateurs est de dimension au plus  $\frac{n(n-1)}{2}$ , et dans le cas d’égalité ils parvinrent à trouver une forme explicite pour ces espaces (forme faisant intervenir des matrices alternées).

### 3.3.3 Historique de mes travaux

Mon attention a été attirée sur la question des familles d’opérateurs localement liées par une discussion que j’ai eue avec Peter Šemrl lors de ma première venue à Ljubljana au début du printemps 2012. Au bord du lac de Bled, je soumis à Peter quelques idées disparates, et je lui fis notamment part de mes recherches sur la possibilité d’écrire toute matrice de trace nulle comme un crochet de Lie à partir d’un hyperplan vectoriel fixé<sup>1</sup>. Ce travail, que j’avais réalisé plusieurs années auparavant, nécessitait la connaissance des couples  $(f, g)$  d’endomorphismes d’un espace vectoriel de dimension finie tels que  $(\text{id}, f, g)$  soit localement liée. J’étais à l’époque venu à bout de cette question par de très lourds calculs, et demandai à Peter s’il avait déjà rencontré cette situation. Tout sourire, il m’indiqua qu’il avait précisément beaucoup travaillé sur les familles d’opérateurs localement liées. De fait, c’était avec Meshulam le spécialiste de

1. Voir C. de Seguins Pazzis, Commutators from a hyperplane of matrices. Electronic J. Linear Algebra **27** (2014) 39–54.



ce sujet. Il m'indiqua les principaux résultats qu'il avait trouvés, et je repartis en France avec un paquet de tirés à part sous le bras, intrigué par le sujet et désireux de l'approfondir.

Quelques jours plus tard, alors que j'essayais de retrouver moi-même le premier résultat de Brešar et Šemrl sur le rang [Bre99], je pris conscience que les espaces d'opérateurs localement liés n'étaient qu'une forme déguisée d'espaces d'opérateurs de rang majoré. Dans les articles publiés sur la question jusqu'à présent, ce point de vue émergait parfois, mais il n'avait clairement pas été utilisé de manière systématique. Par exemple, le résultat de base de Brešar et Šemrl n'était qu'une traduction du résultat le plus simple dans le lemme de Flanders-Atkinson (lemme 2.16 du chapitre 2). Les autres résultats de Meshulam et Šemrl étaient en revanche plus originaux et n'avaient pas encore eu leur contrepartie dans la théorie des espaces vectoriels de rang majoré. Avançant rapidement, je découvris l'article de Chebotar et Šemrl et fus convaincu à sa lecture que le point de vue des espaces de rang majoré permettrait de donner une démonstration beaucoup plus courte de leur résultat. Et non seulement je réalisai ce projet très vite, mais j'obtins des résultats encore plus généraux. Mais je constatai par hasard quelques semaines plus tard que je n'avais fait que redécouvrir la théorie d'Atkinson et Lloyd des espaces primitifs de matrices de rang majoré (voir le paragraphe 2.7). Je m'aperçus qu'Atkinson avait des techniques plus performantes que les miennes et que son usage des matrices génériques était redoutablement efficace. J'appris donc à maîtriser sa méthode, et parvins grâce à cela à aller sensiblement plus loin que lui dans l'analyse des sous-espaces primitifs minimaux à grand nombre de lignes par rapport au nombre de colonnes.

Je passai ensuite à l'étude des espaces non réflexifs. J'améliorai à la marge les contraintes de cardinalité pour le théorème 3.3, et étudiâi ensuite l'optimalité de la borne  $2n - 2$ . Je parvins d'abord facilement à fabriquer des corps pour lesquels la borne  $n$ , obtenue par Meshulam et Šemrl dans [Mes04] sur les corps algébriquement clos, ne convenait pas. Ensuite, je parvins par une étude minutieuse à utiliser les techniques sur les espaces de rang majoré pour comprendre que tous les espaces où il n'existe pas d'élément non nul de rang strictement inférieur à  $2n - 2$  étaient produits à partir d'une structure algébrique curieuse que j'ai baptisée les algèbres à division à gauche bilinéarisable (algèbres à division LDB<sup>2</sup>). En ce mois de mai 2012, je consultai Bruno Kahn, spécialiste des formes quadratiques, pour savoir si les algèbres à division LDB avaient été étudiées par le passé, et il apparut que rien n'était connu sur la question. J'entrepris alors d'analyser ces objets, mais il fallut attendre de nombreux mois avant que je ne puisse progresser. Une percée fut finalement faite à l'automne 2012, où je trouvai enfin un angle d'attaque pour élucider le mystère de ces objets. Pris par d'autres préoccupations, et notamment la publication de l'intégralité de mes résultats sur les espaces d'opérateurs localement liés [1, 2], je laissai la chose en l'état jusqu'à l'été 2013, où je clos définitivement la question de la structure des algèbres à division LDB en les mettant en correspondance avec les algèbres de composition de Hurwitz.

Une chose restait en suspens, et c'était la validité du théorème 3.3 sans hypothèse sur le cardinal du corps de base. À l'hiver 2014, je trouvai comment obtenir ce résultat par un argument de dénombrement sensiblement plus fin que celui utilisé par Meshulam et Šemrl pour le résultat correspondant sur les

---

2. Left division bilinearizable algebras.

espaces d'opérateurs localement liés.

Depuis, je ne suis plus revenu sur la question des espaces localement liés, sinon pour tirer des conséquences sur leur structure à partir de résultats sur les sous-espaces vectoriels de matrices de rang majoré [5].

### 3.4 La dualité opérateur-vecteur

Commençons par quelques définitions utiles. Des sous-espaces vectoriels respectifs  $\mathcal{S} \subset \mathcal{L}(U, V)$  et  $\mathcal{S}' \subset \mathcal{L}(U', V')$  sont dits **équivalents** lorsqu'il existe des isomorphismes  $\varphi : U \xrightarrow{\sim} U'$  et  $\psi : V \xrightarrow{\sim} V'$  tels que  $\mathcal{S}' = \{\psi \circ f \circ \varphi^{-1} \mid f \in \mathcal{S}\}$ , et on parlera de **similitude** lorsque  $U = U'$ ,  $V = V'$  et que l'on peut prendre  $\varphi = \psi$ .

Pour un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$ , on appelle **noyau** de  $\mathcal{S}$  l'intersection des noyaux de ses éléments, et **image essentielle** de  $\mathcal{S}$  la somme des images de ses éléments. On dit que  $\mathcal{S}$  est :

- **réduit au départ** lorsque son noyau est  $\{0\}$  ;
- **réduit à l'arrivée** lorsque son image essentielle est  $V$  ;
- **réduit** lorsqu'il est à la fois réduit au départ et à l'arrivée.

Si l'on note  $U_0$  le noyau de  $\mathcal{S}$  et  $V_0$  son image essentielle, tout opérateur  $f \in \mathcal{S}$  induit un opérateur linéaire  $\bar{f} : U/U_0 \rightarrow V_0$ . L'ensemble  $\bar{\mathcal{S}}$  des opérateurs linéaires ainsi formés est un sous-espace vectoriel réduit de  $\mathcal{L}(U/U_0, V_0)$ , et l'application  $f \in \mathcal{S} \mapsto \bar{f} \in \bar{\mathcal{S}}$  est un isomorphisme d'espaces vectoriels conservant le rang. Un espace d'opérateurs  $\mathcal{S}$  est localement lié si et seulement si son réduit  $\bar{\mathcal{S}}$  est localement lié. On peut donc entièrement limiter l'étude à des espaces réduits d'opérateurs.

Le principe de la dualité opérateur-vecteur est de faire agir les vecteurs de départ sur les opérateurs du sous-espace  $\mathcal{S}$ . Pour un vecteur  $x \in U$ , on introduit l'opérateur linéaire

$$\hat{x} : f \in \mathcal{S} \mapsto f(x) \in V.$$

On obtient ainsi un sous-espace d'opérateurs  $\hat{\mathcal{S}}$  de  $\mathcal{L}(\mathcal{S}, V)$ , appelé **dual-opérateur** de  $\mathcal{S}$ . Le dual-opérateur est toujours réduit au départ, car si un opérateur de  $\mathcal{S}$  s'annule en tout vecteur de  $U$ , alors il est nul. Lorsque  $\mathcal{S}$  est réduit au départ, l'application  $x \in U \mapsto \hat{x} \in \hat{\mathcal{S}}$  est un isomorphisme d'espaces vectoriels et on dispose d'un isomorphisme canonique de  $\mathcal{S}$  sur  $\hat{\mathcal{S}}$  qui à  $f$  associe  $[\hat{x} \mapsto f(x)]$ .

La dualité opérateur-vecteur permet de convertir des propriétés de l'espace  $\mathcal{S}$  en des propriétés de son dual-opérateur. Pour ce qui nous concerne dans ce chapitre, le résultat essentiel est le suivant, et il est parfaitement évident :

**Proposition 3.4.** *Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$ . Pour que  $\mathcal{S}$  soit localement lié, il est nécessaire et suffisant qu'aucun élément non nul de  $\hat{\mathcal{S}}$  ne soit injectif.*

Une application immédiate et spectaculaire de cette observation réside dans la classification des couples  $(f, g)$  d'opérateurs localement liés. Supposons un tel couple libre et considérons le dual-opérateur de  $\mathcal{S} := \text{Vect}(f, g)$ . Ainsi tout élément non nul de  $\hat{\mathcal{S}}$  est de rang 1. Par le théorème de Schur (théorème 2.1), ou bien il existe un hyperplan de  $\mathcal{S}$  sur lequel tout élément de  $\hat{\mathcal{S}}$  est nul, ou bien il existe une droite vectorielle  $D$  incluant l'image de tout élément de  $\hat{\mathcal{S}}$ . La première option est impossible puisque  $\hat{\mathcal{S}}$  est réduit au départ. Cela valide la

deuxième option, et une droite  $D$  associée inclut nécessairement  $\text{Im } f$  et  $\text{Im } g$ . Ainsi,  $f$  et  $g$  sont de rang 1 et de même image.

Dans toute la suite, nous supposons  $U$  et  $V$  de dimension finie pour simplifier le discours. Donnons-nous un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$ . Si l'on se donne des bases respectives  $\mathbf{B}$  et  $\mathbf{C}$ , de tailles respectives  $n$  et  $m$ , l'espace  $\mathcal{M}$  des matrices représentant le dual-opérateur  $\widehat{\mathcal{S}}$  dans ces bases est un sous-espace vectoriel de  $M_{m,n}(\mathbb{K})$  dans lequel tout élément est de rang strictement inférieur à  $n$ .

Les questions de rang dans  $\mathcal{S}$  se traduisent par des questions de transitivité dans le dual-opérateur  $\widehat{\mathcal{S}}$ . Ainsi, montrer que  $\mathcal{S}$  contient un élément non nul de rang  $r$  revient à montrer que dans l'espace de départ du dual-opérateur, il existe un vecteur non nul  $x$  tel que  $\dim(\widehat{\mathcal{S}}x) = r$ . Ou encore qu'il existe  $X \in \mathbb{K}^n \setminus \{0\}$  tel que  $\dim MX = r$ . On va pouvoir alors systématiquement se raccrocher à des techniques connues (ou des dérivées de techniques connues) sur les espaces de matrices de rang majoré pour analyser ces questions, et le plus souvent on se raccrochera à celles découvertes et développées par Flanders, Atkinson et Lloyd.

Avant de voir cela, il convient donc de rappeler les deux résultats techniques principaux déjà cités dans le chapitre 2 :

**Lemme 3.5** (Lemme de Flanders-Atkinson). *Soit  $\mathcal{M}$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$ , de rang maximal  $r$ . On suppose que  $|\mathbb{K}| > r$  et que  $\mathcal{M}$  contient la matrice  $\begin{bmatrix} I_r & [0] \\ [0] & [0] \end{bmatrix}$ . Toute matrice de  $\mathcal{M}$  s'écrit alors sous la forme*

$$M = \begin{bmatrix} A(M) & C(M) \\ B(M) & [0] \end{bmatrix}$$

où  $A(M) \in M_r(\mathbb{K})$ ,  $B(M) \in M_{n-r,r}(\mathbb{K})$  et  $C(M) \in M_{r,p-r}(\mathbb{K})$ , et est vérifiée la relation  $B(M)A(M)^k C(M) = 0$  pour tout  $k \in \mathbb{N}$ .

**Lemme 3.6** (Lemme des blocs (Atkinson-Lloyd)). *Soit  $\mathcal{M}$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$ , de rang maximal  $r$ . On suppose que  $|\mathbb{K}| > r$ . On suppose trouvé un couple  $(n', p') \in \llbracket 1, n-1 \rrbracket \times \llbracket 1, p-1 \rrbracket$  tel que toute matrice de  $\mathcal{M}$  s'écrive sous la forme*

$$\begin{bmatrix} A(M) & C(M) \\ B(M) & [0] \end{bmatrix}$$

où  $A(M) \in M_{n',p'}(\mathbb{K})$ ,  $B(M) \in M_{n-n',p'}(\mathbb{K})$  et  $C(M) \in M_{n',p'-p}(\mathbb{K})$ . Alors

$$\text{rgmax } B(\mathcal{M}) + \text{rgmax } C(\mathcal{M}) \leq \text{rgmax } \mathcal{M}.$$

### 3.5 Application de la dualité opérateur-vecteur aux problématiques de rang

Pour illustrer notre propos, prenons le tout premier résultat significatif sur le rang (non nul) minimal dans un espace localement lié. C'est le lemme de Brešar et Šemrl :

**Théorème 3.7** (Brešar et Šemrl (1999)). *On suppose  $\mathbb{K}$  infini. Soit  $\mathcal{S}$  un espace d'opérateurs localement lié de dimension  $n > 0$ . Il existe alors  $f \in \mathcal{S} \setminus \{0\}$  tel que  $\text{rg } f < n$ .*

Par le prisme de la dualité, ce résultat se déduit directement de la partie la plus simple du lemme de Flanders-Atkinson, celle qui indique la nullité du bloc inférieur droit ! En effet, cette partie du lemme se traduit par le fait que, dans tout espace d'opérateurs linéaires  $\mathcal{T}$  entre espaces vectoriels  $U$  et  $V$  de dimension finie, si l'on prend un opérateur  $f_0$  de rang maximal  $r$  et  $|\mathbb{K}| > r$ , alors tout opérateur dans  $\mathcal{T}$  envoie  $\text{Ker } f_0$  dans  $\text{Im } f_0$  (c'est naturel si  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ , car on reconnaît dans cette condition celle qui définit l'espace vectoriel tangent en  $f_0$  à la variété des opérateurs de rang  $r$ ).

Et particulier si  $r < \dim U$ , il suffira de choisir un vecteur non nul  $x$  de  $\text{Ker } f_0$  pour observer que  $\mathcal{T}x \subset \text{Im } f_0$  et ainsi  $\mathcal{T}x$  est de dimension au plus  $r$  ! Cette propriété, appliquée à  $\mathcal{T} = \widehat{\mathcal{S}}$ , donne *immédiatement* le théorème plus haut, et permet même de voir que l'infinitude de  $\mathbb{K}$  peut être remplacée par l'hypothèse plus faible  $|\mathbb{K}| \geq n$ .

La même technique a permis de redémontrer le théorème de Meshulam et Šemrl qui interroge l'optimalité de la borne  $n$  dans ce théorème (voir le théorème 3.2). L'argument est à nouveau limpide : supposons donc disposer d'un sous-espace vectoriel  $\mathcal{M}$  de  $M_{m,n}(\mathbb{K})$  de rang maximal strictement inférieur à  $n$  dans lequel  $\dim \mathcal{M}X \geq n$  pour tout  $X \in \mathbb{K}^n \setminus \{0\}$ , et supposons  $|\mathbb{K}| \geq n$ . D'abord, ce qui précède montre que le rang maximal dans  $\mathcal{M}$  est nécessairement  $n - 1$ . Ensuite, on remarque que  $\dim \mathcal{M}X \leq n - 1$  dès que  $X$  est un vecteur dans le noyau d'un élément de rang  $n - 1$  de  $\mathcal{M}$  (encore une application du lemme de Flanders). L'astuce – qu'il faut formaliser un peu plus sur les corps finis – est de montrer que les vecteurs ainsi considérés forment une partie algébriquement dense de  $\mathbb{K}^n$ . Et cela s'obtient assez facilement : après réduction, on peut supposer que  $\mathcal{M}$  contient la matrice blocs  $J_{n-1} = \begin{bmatrix} I_{n-1} & [0] \\ [0] & [0] \end{bmatrix}$ , on saura aussi par les contraintes supposées que  $\mathcal{M}$  contient, pour tout  $X \in \mathbb{K}^{n-1}$ , une matrice de la forme  $M_X = \begin{bmatrix} ? & X \\ [0] & [0] \end{bmatrix}$ . En examinant le noyau de  $J_{n-1} + tM_X$  et en faisant varier  $t$  et  $X$ , on obtient la densité algébrique espérée, ce qui permet de conclure.

Terminons en voyant comment la technique permet d'éclairer la démonstration du théorème 3.3 tout en relâchant légèrement la contrainte de cardinalité sur le corps de base. Donnons-nous donc un espace d'opérateur  $\mathcal{T}$  non réflexif et de dimension  $n \geq 2$ , entre espaces vectoriels de dimension finie (on peut toujours se ramener à ce cas grâce à un théorème de Larson [Lar88]). Supposons en outre  $|\mathbb{K}| > n$ . On peut voir  $\mathcal{T}$  comme un hyperplan d'un espace localement lié  $\mathcal{S}$ . Considérons le dual-opérateur  $\widehat{\mathcal{S}}$  ainsi qu'un élément  $\varphi$  de rang maximal dans celui-ci. Si  $\text{Ker } \varphi \cap \mathcal{T} \neq \{0\}$  alors le lemme de Flanders-Atkinson montre que tout élément non nul  $f$  de  $\text{Ker } \varphi \cap \mathcal{T}$  vérifie  $\text{rg } f \leq \text{rg } \varphi \leq n$ .

Supposons  $\text{Ker } \varphi \cap \mathcal{T} = \{0\}$  dans toute la suite, ce qui donne en particulier  $\text{rg } \varphi = n$ . Prenons  $f_0 \in \text{Ker } \varphi$ . Le lemme de Flanders-Atkinson donne  $\text{rg } f_0 \leq n$ . En outre, le théorème 3.2 (avec l'hypothèse de cardinalité rectifiée sur  $\mathbb{K}$ ) assure que si  $\text{rg } f_0 = n$  alors  $\text{rg } f \leq n$  pour tout  $f \in \mathcal{T}$ .

Supposons enfin  $\text{rg } f_0 < n$ . Représentons  $\widehat{\mathcal{S}}$  dans une base adaptée à la décomposition  $\mathcal{S} = \mathcal{T} \oplus \mathbb{K}f_0$  et dans une base adaptée au sous-espace  $\text{Im } \varphi$ . Les matrices de l'espace  $\mathcal{M}$  associé s'écrivent donc toutes sous la forme

$$M = \begin{bmatrix} [?]_{(n-1) \times n} & C(M) \\ B(M) & [0] \end{bmatrix}.$$

Il suffit alors que  $|\mathbb{K}| > n$  pour que le lemme des blocs donne  $\text{rgmax} B(\mathcal{M}) + \text{rgmax} C(\mathcal{M}) \leq n$ , si bien que  $\text{rgmax} B(\mathcal{M}) \leq n - 1$ . Ainsi, une nouvelle application du lemme de Flanders-Atkinson fournit un vecteur  $X \in \mathbb{K}^n \setminus \{0\}$  tel que  $\dim B(\mathcal{M})X \leq n - 1$ . Cela se traduit bien par l'existence d'un élément  $h$  de  $\mathcal{T}$  tel que  $\text{rg} h \leq (n - 1) + (n - 1) = 2n - 2$ .

Meshulam et Šemrl ont montré [Mes04] que la borne supérieure  $2n - 2$  du théorème 3.3 peut être remplacée par  $n$  sur un corps algébriquement clos. La méthode de dualité opérateur-vecteur permet d'éclairer largement leur démonstration, mais nous n'en dirons rien ici. Meshulam et Šemrl pensaient que cette borne  $n$  valait pour tout corps de cardinal assez grand, mais nous avons démontré qu'il n'en est rien : par exemple, dès que l'on peut trouver un hyperplan vectoriel  $H$  de  $A_n(\mathbb{K})$  dans lequel tout élément non nul est inversible, il est facile de voir que le sous-espace  $H$  (identifié à un espace d'opérateurs) a pour clôture réflexive  $A_n(\mathbb{K})$ , alors que d'une part tout élément non nul de  $H$  est de rang  $n$ , et d'autre part  $\dim H = n - 1$ . L'existence d'un tel hyperplan est garantie par exemple sur  $\mathbb{R}$  en dimensions 2, 4 et 8 (cela s'obtient en utilisant une algèbre de Hurwitz).

Nous laissons temporairement de côté les interrogations sur l'optimalité de la borne  $2n - 2$  et y reviendrons dans la section 3.7. Terminons cette partie en donnant une belle application du théorème 3.7, ou plutôt de sa généralisation à n'importe quel corps (dont on rappelle que le traitement des corps finis se fait par un argument de dénombrement). Cette application nous a été soufflée par Jean-Pierre Barani.

**Proposition 3.8.** *Soit  $\mathbb{K} - \mathbb{L}$  une extension finie de corps, de degré  $d$ . Soit  $E$  un  $\mathbb{L}$ -espace vectoriel de dimension finie, et  $F$  un sous- $\mathbb{K}$ -espace vectoriel de  $E$  dont la dimension est multiple de  $d$ . Il existe alors un sous- $\mathbb{L}$ -espace vectoriel  $G$  de  $E$  tel que  $F \oplus G = E$ .*

*Démonstration.* On peut évidemment supposer  $F \subsetneq E$ , et il suffira tout aussi évidemment, par récurrence sur la dimension de  $F$ , de montrer l'existence d'un vecteur  $x$  de  $E \setminus \{0\}$  tel que  $F \cap \mathbb{L}x = \{0\}$ . Pour cela, on considère

$$\mathcal{S} := \{\varphi_\lambda : x \in E \mapsto [\lambda x] \in E/F \mid \lambda \in \mathbb{L}\}.$$

Pour tout  $\lambda \in \mathbb{L}^*$ , l'opérateur  $\varphi_\lambda$  est surjectif, donc de rang au moins  $d$ . Comme  $\mathcal{S}$  est de dimension  $d$ , le théorème de Brešar-Šemrl-Meshulam montre qu'il n'est pas localement lié, d'où un vecteur  $x \in E \setminus \{0\}$  tel que  $[\lambda x] \neq 0$  pour tout  $\lambda \in \mathbb{L}^*$ . Autrement dit,  $\mathbb{L}x \cap F = \{0\}$ , ce qui achève la démonstration.  $\square$

### 3.6 Application de la dualité opérateur-vecteur à la détermination des espaces localement liés minimaux

La dualité opérateur-vecteur permet de relier la détermination des espaces localement liés minimaux à celle des sous-espaces semi-primitifs d'Atkinson (voir le paragraphe 2.7). Prenons en effet un espace localement lié minimal  $\mathcal{S} \subset \mathcal{L}(U, V)$ , que l'on peut supposer réduit sans perdre de généralité. Le dual-opérateur  $\widehat{\mathcal{S}}$  est alors réduit. Enfin, il est semi-primitif car s'il violait la condition (iii) des espaces semi-primitifs, il existerait un hyperplan vectoriel  $H$  de  $\mathcal{S}$  tel

que tout élément de  $\{\widehat{x}|_H \mid x \in U\}$  soit non injectif, autrement dit  $H$  serait localement lié et on contredirait ainsi la minimalité de  $\mathcal{S}$ .

Enfin, et sous réserve que  $\dim \mathcal{S} \geq 2$ , le rang maximal dans  $\widehat{\mathcal{S}}$  doit être  $\dim \mathcal{S} - 1$  car s'il est strictement inférieur alors tout hyperplan vectoriel de  $\mathcal{S}$  est localement lié.

Il est facile enfin de voir réciproquement que, pour tout espace d'opérateurs  $\mathcal{S} \subset \mathcal{L}(U, V)$  de dimension  $n \geq 2$ , si  $\widehat{\mathcal{S}}$  est semi-primitif et de rang maximal  $n - 1$  alors  $\mathcal{S}$  est localement lié minimal.

L'étude des espaces d'opérateurs localement liés minimaux est donc entièrement réduite à celle des espaces semi-primitifs d'opérateurs (ou de matrices). Cette dernière est bien balisée et l'on y dispose de théorèmes puissants (voir le paragraphe 2.7) démontrés avec des outils efficaces comme les matrices génériques. *A contrario*, les méthodes directes fondées sur le point de vue « localement lié » rendent les études extrêmement lourdes.

Dans [Che08], Chebotar et Šemrl étudièrent la dimension d'arrivée maximale pour un espace localement lié minimal réduit de dimension donnée, et ils étudièrent le cas où cette dimension maximale est réalisée. Ils ne firent, et au prix d'une hypothèse de cardinalité non naturelle et légèrement trop forte ( $|\mathbb{K}| \geq n+2$  au lieu de  $|\mathbb{K}| \geq n$ ), que retrouver le cas limite dans le théorème d'Atkinson sur les sous-espaces semi-primitifs. Leur démonstration extrêmement technique est battue à plate couture par l'approche dualité opérateur-vecteur. Chebotar et Šemrl étaient tout simplement inconscient de la théorie d'Atkinson, tout comme ils étaient insuffisamment convaincus de la pertinence de l'approche par dualité. De mon côté, en cherchant à améliorer au printemps 2012 leur résultat, j'ai redécouvert seul la notion d'espace semi-primitif et ai pu redémontrer le théorème d'Atkinson de manière indépendante, mais avec des méthodes moins performantes à l'époque que celles d'Atkinson. C'est par hasard que j'ai fait la découverte des travaux d'Atkinson et Lloyd sur la question et que j'ai pu ensuite aller plus loin.

À partir de là, il n'y a plus grand-chose à dire sur le sujet : si l'on veut étudier les sous-espaces localement liés minimaux, la seule méthode efficace connue est de se ramener, par dualité opérateur-vecteur, aux sous-espaces semi-primitifs de rang majoré. Tout le travail technique se fait alors dans ce dernier cadre, avec les méthodes appropriées. Nous terminerons tout de même en citant et en commentant le théorème principal de [Bre99] sur les triplets localement liés :

**Théorème 3.9** (Brešar, Šemrl (1999)). *Soit  $(f, g, h)$  un triplet d'opérateurs localement lié d'un espace vectoriel  $U$  vers un espace vectoriel  $V$ . On suppose  $|\mathbb{K}| > 2$ . L'une des trois conditions suivantes est alors vérifiée :*

- (i) *La famille  $(f, g, h)$  est liée dans  $\mathcal{L}(U, V)$ .*
- (ii) *Il existe un sous-espace vectoriel  $P$  de  $V$  tel que  $\dim P \leq 2$  et  $f, g, h$  aient tous leur image incluse dans  $P$ .*
- (iii) *Il existe deux triplets  $(\alpha, \beta, \gamma) \in \mathbb{K}^3$  et  $(\alpha', \beta', \gamma') \in \mathbb{K}^3$  non colinéaires tels que  $\alpha f + \beta g + \gamma h$  et  $\alpha' f + \beta' g + \gamma' h$  aient tous leur image incluse dans une même droite vectorielle.*
- (iv) *La somme  $V_0 := \text{Im } f + \text{Im } g + \text{Im } h$  est de dimension 3, l'espace  $\text{Ker } f \cap \text{Ker } g \cap \text{Ker } h$  est de codimension 3, et pour n'importe quel supplémentaire  $U_0$  de ce dernier dans  $U$ , il existe des bases respectives de  $U_0$  et  $V_0$  telles que les éléments de  $\text{Vect}(f|_{U_0}, g|_{U_0}, h|_{U_0})$  soient représentés dans ces bases par les matrices de  $A_3(\mathbb{K})$ .*

Pour démontrer cela, on peut déjà supposer  $(f, g, h)$  libre et  $\mathcal{S} := \text{Vect}(f, g, h)$  réduit. Si  $\mathcal{S}$  n'est pas minimal, on en extrait un plan vectoriel localement lié, puis la caractérisation des couples localement liés donne que (iii) est vérifiée. Si  $\mathcal{S}$  est minimal et  $\dim V \leq 2$ , alors (ii) est vérifiée. Si enfin  $\mathcal{S}$  est minimal et  $\dim V \geq 3$ , alors le théorème d'Atkinson (théorème 2.15) donne que  $\widehat{\mathcal{S}}$  est de type alterné et associé à la fonction  $f = \text{id} : \mathbb{K}^3 \wedge \mathbb{K}^3 \rightarrow \mathbb{K}^3 \wedge \mathbb{K}^3$ . Il est alors facile de remonter au fait que  $\mathcal{S}$  est lui-même représenté par l'espace de matrices  $A_3(\mathbb{K})$ , ce qui donne (iv).

## 3.7 Les algèbres à division à gauche bilinéarisable

### 3.7.1 Du problème du rang minimal aux algèbres à division LDB

Revenons au théorème 3.3 portant sur le rang minimal (non nul) dans un espace non réflexif. La question de l'optimalité du majorant  $2n - 2$  avait été posée par Meshulam et Šemrl. Rappelons qu'ils avaient nettement amélioré ce majorant pour les corps algébriquement clos ( $n$  au lieu de  $2n - 2$ ) et que nous avons trouvé des exemples où la borne  $n$  ne convenait pas (voir la fin de la section 3.5).

Nous avons alors entrepris une analyse minutieuse des cas où la borne  $2n - 2$  est atteinte : généralisons immédiatement l'étude à un hyperplan vectoriel  $\mathcal{T}$  de dimension  $n$  d'un espace localement lié  $\mathcal{S}$ , supposant que tout élément non nul de  $\mathcal{T}$  est de rang au moins  $2n - 2$ .

On ne considèrera évidemment que le cas réduit. Dans ce cas, en reprenant les arguments évoqués dans la section 3.5 pour trouver l'existence d'un élément de  $\mathcal{T} \setminus \{0\}$  de rang au plus  $2n - 2$ , et en travaillant finement sur les matrices blocs représentant les éléments du dual-opérateur  $\widehat{\mathcal{S}}$ , nous sommes parvenus à la suite d'une longue analyse (sept pages) à la conclusion suivante : ce dual-opérateur est de dimension  $2n - 2$  et représentable, dans des bases bien choisies, par un espace formé des matrices de la forme suivante :

$$\begin{bmatrix} F(Y) & X & [0]_{(n-1) \times 1} \\ E(X) & [0]_{(n-1) \times 1} & Y \end{bmatrix} \quad \text{où } (X, Y) \in (\mathbb{K}^{n-1})^2,$$

les fonctions  $F : \mathbb{K}^{n-1} \rightarrow M_{n-1}(\mathbb{K})$  et  $E : \mathbb{K}^{n-1} \rightarrow M_{n-1}(\mathbb{K})$  sont linéaires et telles que  $F(X)$  et  $E(X)$  soient inversibles pour tout  $X \in \mathbb{K}^{n-1} \setminus \{0\}$ , et enfin elles vérifient la contrainte de colinéarité

$$\forall X \in \mathbb{K}^{n-1} \setminus \{0\}, \forall Y \in \mathbb{K}^{n-1}, F(Y)E(X)^{-1}Y \in \mathbb{K}X. \quad (3.1)$$

Les fonctions  $F$  et  $G$  sont évidemment liées aux structures d'algèbres à division. Rappelons qu'une algèbre à division sur un corps  $\mathbb{K}$  est simplement un couple  $(A, \star)$  formé d'un  $\mathbb{K}$ -espace vectoriel  $A$  (non nul et de dimension finie) et d'une fonction bilinéaire  $\star : A^2 \rightarrow A$  qui soit régulière, autrement dit  $x \star y \neq 0$  pour tous  $x, y$  non nuls dans  $A$ . Par curryfication, se donner une telle algèbre à division revient à se donner une application linéaire  $d : A \rightarrow \mathcal{L}(A)$  envoyant tout élément non nul de  $A$  sur un automorphisme du  $\mathbb{K}$ -espace vectoriel  $A$ .

Dans la conclusion plus haut, les fonctions  $F$  et  $E$  permettent de créer deux lois  $\star$  et  $\bullet$  sur  $\mathbb{K}^{n-1}$  par

$$X \star Y := F(Y)X \quad \text{et} \quad X \bullet Y := E(Y)X.$$

La contrainte de colinéarité (3.1), appliquée au couple  $(Y, E(Y)X)$ , indique que  $X \star (X \bullet Y)$  est colinéaire à  $Y$  pour tous  $X, Y$  dans  $\mathbb{K}^{n-1}$  ( $y$  compris lorsque  $X = 0$ , évidemment).

J'ai ainsi mis le doigt sur une structure nouvelle définie comme suit :

**Définition 4.** Une **algèbre à division à gauche bilinéarisable**, en abrégé algèbre à division LDB<sup>3</sup>, est un triplet  $(A, \star, \bullet)$  dans lequel  $(A, \star)$  et  $(A, \bullet)$  sont deux algèbres à division et  $x \star (x \bullet y)$  est colinéaire à  $y$  pour tous  $x, y$  dans  $A$ .

La terminologie s'explique bien : dans une algèbre à division  $(A, \star)$ , pour  $x$  dans  $A$  et  $y$  dans  $A \setminus \{0\}$ , l'équation  $y \star z = x$  d'inconnue  $z$  possède une unique solution dans  $A$ , et cette solution est naturellement appelée «  $x$  divisé à gauche par  $y$  ». Dans une algèbre à division à gauche bilinéarisable, cette solution est presque égale à  $y \bullet x$ , « presque » au sens de la colinéarité. La fonction de division à gauche est ponctuellement égale à la deuxième loi à colinéarité près, d'où la terminologie « division à gauche bilinéarisable ».

Nous nous inquiéterons ultérieurement de l'existence de telles structures, mais concluons d'abord sur le lien avec les espaces non réflexifs, ou plutôt avec les hyperplans d'espaces d'opérateurs localement liés. On introduit l'application bilinéaire

$$\Gamma_A : \begin{cases} (A \times \mathbb{K}^2) \times A^2 & \longrightarrow A^2 \\ ((x, (\lambda, \mu)), (y, z)) & \longmapsto (x \star z + \lambda y, x \bullet y + \mu z). \end{cases}$$

L'espace d'opérateurs tordus<sup>4</sup>  $\mathcal{T}_A$  est défini comme l'image de la curryfiée  $X \in A \times \mathbb{K}^2 \mapsto \Gamma(X, -) \in \mathcal{L}(A^2)$ , et c'est d'ailleurs un espace d'endomorphismes ; on a une bijection linéaire naturelle de  $A \times \mathbb{K}^2$  sur  $\mathcal{T}_A$ , et ce dernier est de dimension  $\dim A + 2$ . L'étude citée précédemment, et qui constitue la fin de la toute dernière partie de [1], se conclut par le théorème suivant :

**Théorème 3.10** (Théorème 6.1 de [1]). *Soit  $\mathcal{S}$  un espace d'opérateurs localement lié réduit et de dimension  $n + 1$ . Supposons que  $|\mathbb{K}| > n$  et qu'il existe un hyperplan vectoriel de  $\mathcal{S}$  dont tout élément non nul est de rang supérieur à  $2n - 2$ . Il existe alors une algèbre à division LDB  $(A, \star, \bullet)$  telle que  $\mathcal{S}$  soit équivalent à  $\mathcal{T}_A$ .*

Inversement, il n'est pas difficile de démontrer que pour une algèbre à division LDB  $(A, \star, \bullet)$ , l'espace  $\mathcal{T}_A$  est toujours localement lié.

### 3.7.2 Les algèbres à division LDB pour elles-mêmes (début)

Le théorème 3.10 étant acquis, il me restait à comprendre la structure de ces mystérieuses algèbres à division LDB. Il est d'abord facile d'en construire à l'aide des algèbres de composition classique. Par exemple, si l'on prend une algèbre de quaternions  $\mathcal{Q}$  sur  $\mathbb{K}$  qui est un corps gauche, sa conjugaison  $x \mapsto \bar{x}$  et sa norme  $N$ , on dispose de l'identité  $x(\bar{xy}) = N(x)y$  pour tout  $(x, y) \in \mathcal{Q}^2$ . Ainsi,  $(\mathcal{Q}, (x, y) \mapsto xy, (x, y) \mapsto \bar{xy})$  est une algèbre à division LDB. Ce principe fonctionne aussi pour les octonions généralisés (même sans avoir l'associativité de la multiplication).

3. Left division bilinearizable division algebra.

4. Par opposition à l'espace d'opérateurs classique  $\{a \star - \mid a \in A\} \subset \mathcal{L}(A)$  associé à  $\star$ .



Plus généralement, les algèbres de Hurwitz fournissent des algèbres à division LDB. Rappelons qu'une algèbre de Hurwitz est une algèbre à division  $(A, \star)$  possédant un élément neutre (bilatère) et munie d'une forme quadratique non dégénérée  $N$  telle que

$$\forall(x, y) \in A^2, N(x \star y) = N(x)N(y).$$

Dans ce cas, il est classique que pour la conjugaison  $x \mapsto \bar{x}$  associée, qui est la symétrie  $N$ -orthogonale par rapport à la droite dirigée par le neutre, on ait l'identité  $x \star (\bar{x} \star y) = N(x)y$  pour tous  $x, y$  dans  $A$ . Le triplet  $(A, \star, (x, y) \mapsto \bar{x} \star y)$  est donc une algèbre à division LDB de forme quadratique associée  $N$ . Quant aux algèbres de Hurwitz, leur classification est connue : elles sont toutes de dimension 1, 2, 4 ou 8 ; celles de dimension 4 sont rattachées aux algèbres de quaternions, et celles de dimension 8 aux octonions généralisés.

En caractéristique 2, on a en outre des exemples remarquables fournis par les phénomènes de radicalité. Nous dirons qu'une extension  $\mathbb{L}$  de  $\mathbb{K}$  est **hyper-radicielle** lorsqu'elle est de degré fini et  $x^2 \in \mathbb{K}$  pour tout  $x \in \mathbb{L}$ . Le triplet  $(\mathbb{L}, \pi, \pi)$  est alors une algèbre à division LDB pour l'application produit  $\pi : (x, y) \mapsto xy$ .

Et il n'est pas difficile de fournir à partir de là de nombreux exemples montrant l'optimalité de la borne  $2n - 2$  dans le théorème de Meshulam et Šemrl, mais seulement pour des valeurs spécifiques de  $n$ , et bien sûr seulement aussi pour des corps spécifiques.

Nous avons voulu aller plus loin et élucider complètement la structure des algèbres à division LDB. Pendant de nombreux mois, nous avons buté sur cette question, ne trouvant aucun angle d'attaque pour avancer au-delà de considérations très simples. Et, de manière absolument remarquable, c'est par une analyse poussée de la structure de l'espace d'opérateurs tordus  $\mathcal{T}_A$  qu'une véritable brèche s'est ouverte, brèche dans laquelle nous nous sommes alors engouffrés avant de tout emporter.

Avant d'expliquer cette brèche, il faut revenir sur quelques considérations très élémentaires sur les algèbres à division LDB. Donnons-nous donc une algèbre à division LDB  $(A, \star, \bullet)$ . On peut préciser l'hypothèse de colinéarité en repérant par des arguments d'homogénéité de polynômes qu'il existe une (unique) forme quadratique  $q$  *anisotrope* sur  $A$  vérifiant l'identité

$$\forall(x, y) \in A^2, x \star (x \bullet y) = q(x)y.$$

C'est déjà une contrainte assez importante, et en particulier sur un corps fini il n'existe pas d'algèbre à division LDB dont la dimension excède 2. Il faut aussi remarquer qu'en caractéristique 2 la forme  $q$  est certes anisotrope mais elle peut être dégénérée, au sens où la forme polaire  $(x, y) \mapsto q(x+y) - q(x) - q(y)$  peut être dégénérée. Par exemple, dans les algèbres à division LDB citées plus haut et associées à des extensions hyper-radicielles, la fonction  $q$  est en fait  $\mathbb{F}_2$ -linéaire, si bien que sa forme polaire est nulle ! Nous dirons que  $q$  est la forme quadratique associée à notre algèbre à division LDB, et que l'algèbre à division LDB est dégénérée lorsque  $q$  est dégénérée.

Ensuite, on obtient facilement le principe d'inversion : on dispose de l'identité inverse

$$\forall(x, y) \in A^2, x \bullet (x \star y) = q(x)y,$$

si bien que  $(A, \bullet, \star)$  est une algèbre à division LDB de même forme quadratique que  $(A, \star, \bullet)$ . En parlant d'unicité, toute loi (bilinéaire régulière)  $\bullet'$  telle que  $x \star (x \bullet' y)$  soit systématiquement colinéaire à  $y$  est un multiple de  $\bullet$  par un scalaire. La structure du triplet  $(A, \star, \bullet)$  est donc presque entièrement déterminée par celle de la simple algèbre à division  $(A, \star)$ .

L'espace d'opérateurs tordus  $\mathcal{T}_A$  a aussi une forme quadratique naturellement associée : celle déduite de la forme

$$\widetilde{q}_A : \begin{cases} A \times \mathbb{K}^2 & \longrightarrow \mathbb{K} \\ (x, (\lambda, \mu)) & \longmapsto q(x) - \lambda\mu \end{cases}$$

par le transport de structure  $X \in A \times \mathbb{K}^2 \mapsto \Gamma(X, -) \in \mathcal{T}_A$ . On démontre facilement que l'endomorphisme  $\Gamma((x, (\lambda, \mu)), -)$  est bijectif si et seulement si  $q(x) - \lambda\mu \neq 0$ , auquel cas il est de rang  $2 \dim A$  (sinon il est de rang 0 ou  $\dim A$ ). La forme  $\widetilde{q}_A$  est évidemment isotrope contrairement à  $q_A$  ; les hyperplans de  $\mathcal{T}_A$  dans lesquels tout élément non nul est de rang  $2 \dim A$  sont ceux sur lesquels la forme correspondant à  $\widetilde{q}_A$  est anisotrope. Il peut bien sûr ne pas en exister (par exemple si  $\mathbb{K}$  est fini et  $A$  de dimension 2!). Et enfin on peut démontrer que  $\mathcal{T}_A$  est inclus dans la clôture réflexive de chacun d'entre eux, et même qu'il est égal à cette clôture réflexive si  $\dim A > 2$ .

En dehors de ces considérations élémentaires, il est difficile d'aller beaucoup plus loin, on peut cependant citer et démontrer un résultat sur la dimension d'une algèbre à division LDB :

**Lemme 3.11.** *Soit  $(A, \star, \bullet)$  une algèbre à division LDB telle que  $\dim A > 1$ . Alors  $\dim A$  est paire.*

*Démonstration.* On ne perd pas de généralité à supposer  $A = \mathbb{K}^n$  où  $n \geq 2$ . Pour  $X$  dans  $A$ , on note alors  $B(X)$  et  $C(X)$  les matrices canoniquement associées respectivement à  $X \star -$  et  $X \bullet -$ , si bien que  $B(X)C(X) = q(X)I_n$ . La première ligne de  $B(X)$  s'écrit  $X^T M$ , celle de  $C(X)$  s'écrit  $NX$ , pour des matrices fixes  $M, N$  de  $M_n(\mathbb{K})$ . Les matrices  $M$  et  $N$  sont inversibles car  $B(X)$  et  $C(X)$  sont inversibles pour tout  $X \in \mathbb{K}^n \setminus \{0\}$ . L'identité  $\forall X \in \mathbb{K}^n, B(X)C(X) = q(X)I_n$  donne alors, par extraction du coefficient en position  $(1, 2)$ , que  $MN$  est alternée. Comme cette matrice est inversible, on conclut que  $n$  est pair.  $\square$

Avant de poursuivre, il faut introduire des notions d'isomorphie adéquates entre algèbres à division LDB. Pour les simples algèbres à division, la notion pertinente est celle d'*isotypie* : deux algèbres à division  $(A, \star)$  et  $(B, \star')$  sont **isotypiques** lorsqu'il existe trois isomorphismes d'espaces vectoriels  $f, g, h$ , les deux premiers de  $B$  sur  $A$  et le troisième de  $A$  sur  $B$ , tels que

$$\forall (x, y) \in B^2, x \star' y = h(f(x) \star g(y)).$$

**Définition 5.** Deux algèbres à division LDB  $(A, \star, \bullet)$  et  $(B, \star', \bullet')$  sont dites :

- **faiblement équivalentes** lorsque  $(A, \star)$  et  $(B, \star')$  sont isotypiques ;
- **équivalentes** lorsqu'il existe des isomorphismes d'espaces vectoriels  $f$  et  $g$  de  $B$  sur  $A$ , et un isomorphisme d'espaces vectoriels  $h$  de  $A$  sur  $B$  tels que, pour tout  $(x, y) \in B^2$ ,

$$x \star' y = h(f(x) \star g(y)) \quad \text{et} \quad x \bullet' y = g^{-1}(f(x) \bullet h^{-1}(y)).$$

Par le principe d'unicité de la deuxième loi, on obtient facilement que si  $(A, \star, \bullet)$  et  $(B, \star', \bullet')$  sont faiblement équivalentes alors  $(A, \star, \bullet)$  et  $(B, \star', \lambda \bullet')$  sont équivalentes pour un scalaire  $\lambda$  non nul. Deux algèbres à division LDB faiblement équivalentes ont des formes quadratiques associées semblables (c'est-à-dire équivalentes à multiplication près par un scalaire), alors que deux algèbres à division équivalentes ont des formes quadratiques associées équivalentes.

Lorsque deux algèbres à division LDB  $A$  et  $B$  sont faiblement équivalentes (respectivement, équivalentes), il est facile de constater que les espaces d'opérateurs tordus  $\mathcal{T}_A$  et  $\mathcal{T}_B$  sont équivalents (respectivement, semblables), mais la réciproque n'a rien d'évident.

### 3.7.3 Les algèbres à division LDB pour elles-mêmes (suite)

Plusieurs mois après avoir fait les observations qui précèdent, je n'avais pas avancé d'un pouce. Tout s'est débloquent brutalement lorsque j'ai entrepris quelques manipulations en apparence anodines. L'espace d'opérateurs tordus  $\mathcal{T}_A$  est muni de la forme quadratique  $\varphi_A$  correspondant à  $\widetilde{q}_A$ , et j'ai choisi de m'intéresser à l'effet d'une transformation  $\varphi_A$ -orthogonale sur ses éléments vus comme des opérateurs. Une observation critique a alors été faite en étudiant le cas des réflexions parallèlement à des droites dirigées par des éléments de la forme  $\Gamma(X, -)$  où  $X \in (A \setminus \{0\}) \times \mathbb{K}^2$ .

**Lemme 3.12** (Lemme de rectification). *Notons  $\pi$  la projection canonique de  $A \times \mathbb{K}^2$  sur  $A \times \{0\}$ . Pour  $a \in A \times \mathbb{K}^2$  anisotrope pour  $\widetilde{q}_A$ , notons  $s_a$  la symétrie  $\widetilde{q}_A$ -orthogonale de  $A \times \mathbb{K}^2$  parallèlement à  $\mathbb{K}a$ .*

*Soit  $a \in A \times \mathbb{K}^2$  anisotrope tel que  $\pi(a) \neq 0$ . Il existe alors deux automorphismes  $F$  et  $G$  de l'espace vectoriel  $A^2$  tels que*

$$\forall X \in A \times \mathbb{K}^2, \quad \Gamma(s_a(X), -) = G \circ \Gamma(s_{\pi(a)}(X), -) \circ F.$$

Ce lemme apparemment anodin a permis de tout débloquent. D'abord, par une récurrence assez subtile où l'on applique le lemme dans d'autres algèbres à division LDB que  $A$ , on obtient une généralisation à une composée quelconque de symétries de type particulier :

**Lemme 3.13** (Lemme de rectification généralisé). *Soit  $a_1, \dots, a_p$  des vecteurs anisotropes de  $A \times \mathbb{K}^2$  pour  $\widetilde{q}_A$  tels que  $\pi(a_1), \dots, \pi(a_p)$  soient tous non nuls. Posons  $u := s_{a_1} \circ \dots \circ s_{a_p}$  et  $v := s_{\pi(a_1)} \circ \dots \circ s_{\pi(a_p)}$ . Il existe alors deux automorphismes  $F$  et  $G$  de l'espace vectoriel  $A^2$  tels que*

$$\forall X \in A \times \mathbb{K}^2, \quad \Gamma(u(X), -) = G \circ \Gamma(v(X), -) \circ F.$$

Grâce à un lemme indiquant que le groupe orthogonal de  $(A \times \mathbb{K}^2, \widetilde{q}_A)$  est engendré par les réflexions associées à des vecteurs de  $(A \setminus \{0\}) \times \mathbb{K}^2$ , on obtient un résultat final synthétique mais qui ne conserve pas toute la précision du lemme précédent :

**Proposition 3.14** (Principe de rectification). *Soit  $u$  une isométrie linéaire de  $(A \times \mathbb{K}^2, \widetilde{q}_A)$ . Il existe alors une isométrie linéaire  $v$  de  $(A \times \mathbb{K}^2, \widetilde{q}_A)$  fixant tous les éléments de  $\{0\} \times \mathbb{K}^2$ , ainsi que deux automorphismes  $F$  et  $G$  de l'espace vectoriel  $A^2$  tels que*

$$\forall X \in A \times \mathbb{K}^2, \quad \Gamma(u(X), -) = G \circ \Gamma(v(X), -) \circ F.$$

L'utilisation de ce résultat n'est pas tout à fait évidente, et voici la première application que nous en avons tirée dans [1]. Supposons disposer de deux algèbres à division LDB  $(A, \star, \bullet)$  et  $(B, \star', \bullet')$  dont les espaces d'opérateurs tordus  $\mathcal{T}_A$  et  $\mathcal{T}_B$  sont équivalents. On peut alors trouver un isomorphisme d'espaces vectoriels  $K : A \times \mathbb{K}^2 \xrightarrow{\sim} B \times \mathbb{K}^2$  ainsi que des automorphismes  $F$  et  $G$  de l'espace vectoriel  $B^2$  tels que

$$\forall X \in A \times \mathbb{K}^2, \quad \Gamma(K(X), -) = G \circ \Gamma(X, -) \circ F.$$

Compte tenu du contrôle d'inversibilité par les formes  $\widetilde{q}_A$  et  $\widetilde{q}_B$ , on obtient par le Nullstellensatz quadratique que  $K$  est une similitude de  $(A \times \mathbb{K}^2, \widetilde{q}_A)$  vers  $(B \times \mathbb{K}^2, \widetilde{q}_B)$ . On peut alors (théorème de prolongement de Witt) trouver une isométrie  $u$  de  $(B \times \mathbb{K}^2, \widetilde{q}_B)$  telle que  $u \circ K$  envoie  $(0_A, (1, 0))$  sur  $(0_B, (1, 0))$  et  $(0_A, (0, 1))$  sur  $(0_B, (0, \lambda))$  pour un certain  $\lambda$  dans  $\mathbb{K}^*$ . En utilisant le principe de rectification, on se réduit donc, dans la situation initiale, au cas où  $K(0_A, (1, 0)) = (0_B, (1, 0))$  et  $K(0_A, (0, 1)) = (0_B, (0, \lambda))$ . Et à partir de là, par orthogonalité  $K$  envoie  $A \times \{0\}$  sur  $B \times \{0\}$ , puis un travail direct sur la forme des opérateurs dans  $\mathcal{T}_A$  et  $\mathcal{T}_B$  permet de dégager que  $(A, \star, \bullet)$  et  $(B, \star', \bullet')$  sont faiblement équivalentes (et même équivalente si  $\lambda = 1$ , ce que l'on peut obtenir si l'équivalence initiale était une similitude entre espaces d'endomorphismes).

On a en particulier établi le théorème suivant :

**Théorème 3.15** (Théorème 5.10 de [1]). *Soit  $A$  et  $B$  deux algèbres à division LDB. Les conditions suivantes sont équivalentes :*

- (i) *Les espaces d'endomorphismes  $\mathcal{T}_A$  et  $\mathcal{T}_B$  sont semblables (respectivement, équivalents).*
- (ii) *Les algèbres à division LDB  $A$  et  $B$  sont équivalentes (respectivement, faiblement équivalentes).*

À l'époque de [1], l'application la plus immédiate que nous avons vu de ce résultat est l'observation qu'une algèbre à division LDB  $(A, \star, \bullet)$  est toujours équivalente à son algèbre opposée  $(B, \bullet, \star)$  où  $B = A$ . Tout simplement, les espaces d'opérateurs tordus  $\mathcal{T}_A$  et  $\mathcal{T}_B$  se déduisent l'un de l'autre par conjugaison par l'isomorphisme d'échange  $(x, y) \in A^2 \mapsto (y, x) \in A^2$ .

### 3.7.4 Les algèbres à division LDB pour elles-mêmes (fin)

Nous étions à ce stade encore loin de bien comprendre la situation, mais une utilisation plus fine des lemmes de rectification permet enfin d'emporter le morceau à l'été 2013.

Convenons que, pour un vecteur  $e$  de  $A$ , l'algèbre à division LDB  $(A, \star, \bullet)$  est *e-standard* lorsque  $e \star - = \text{id}_A$  (i.e.  $e$  est neutre à gauche pour  $\star$ ) et

$$\forall (x, y) \in A^2, \quad x \bullet y = \bar{x} \star y$$

où  $x \mapsto \bar{x}$  est l'opposée de la symétrie  $q$ -orthogonale parallèlement à  $\mathbb{K}e$ . Les algèbres à division LDB associées aux algèbres de Hurwitz sont évidemment *e-standard* lorsque  $e$  est leur neutre bilatère. Et une observation élémentaire est que si  $A$  est *e-standard* alors  $q(e) = 1$ .

J'ai alors démontré le lemme critique suivant grâce au lemme de rectification généralisé :

**Lemme 3.16** (Lemme de standardisation). *Soit  $(A, \star, \bullet)$  une algèbre à division LDB, de forme quadratique associée notée  $q$ , et  $e$  un vecteur de  $A$  tel que  $q(e) = 1$ . Il existe alors des lois  $\star'$  et  $\bullet'$  sur  $A$  telles que  $(A, \star', \bullet')$  soit une algèbre à division LDB  $e$ -standard, équivalente à  $(A, \star, \bullet)$  et de même forme quadratique associée.*

Autrement dit, on peut décréter que  $A$  est  $e$ -standard sans fondamentalement rien changer au problème, tant que l'on dispose d'un bon candidat  $e$ .

La démonstration se fait en deux étapes. D'abord, on réussit à réaliser une décomposition judicieuse de la réflexion de  $A \times \mathbb{K}^2$  associée au vecteur  $(0_A, (1, -1))$  en la composée de trois réflexions par rapport à des vecteurs de  $A \times \mathbb{K}^2$  qui ont tous pour projeté sur  $A$  un multiple non nul de  $e$ , et on tire du lemme de rectification généralisé qu'il existe deux automorphismes  $g$  et  $h$  de l'espace vectoriel  $A$  tels que

$$\forall(x, y) \in A^2, \quad x \bullet y = h(s_e(x) \star g(y)),$$

où l'on a noté  $s_e$  la réflexion de  $(A, q_A)$  parallèlement à  $\mathbb{K}e$ .

Et à partir de là, c'est un jeu assez facile de modifier les lois  $\star$  et  $\bullet$  pour se ramener à la situation d'une algèbre à division LDB  $e$ -standard (voir les pages 271 et 272 de [3]).

On n'a plus alors qu'à travailler sur des algèbres à division LDB  $(A, \star, \bullet)$  qui sont  $e$ -standard pour un certain vecteur  $e$ . Dans cette situation, on se focalise sur la première loi  $\star$  et l'on se réduit facilement à des méthodes d'algèbres de Clifford. Par exemple, on note que  $(x \star -)^2 = -q(x)\text{id}_A$  pour tout  $x$  dans  $\{e\}^{\perp q}$ . Et si l'on prend donc un sous-espace vectoriel  $V$  de  $\{e\}^{\perp q}$ , on obtient naturellement un morphisme de l'algèbre de Clifford  $C(-q|_V)$  vers  $\mathcal{L}(A)$ . En outre, pour peu que  $q|_V$  soit non dégénérée et  $V$  de dimension paire, un tel morphisme est nécessairement injectif (simplicité des algèbres de Clifford) et on peut alors en tirer des conclusions sur la dimension de  $A$ . Par exemple, supposons  $\chi(\mathbb{K}) \neq 2$ . En notant  $n := \dim A$  et en supposant  $n \geq 3$ , on sait déjà que  $n$  est pair (lemme 3.11) et on peut prendre  $V$  de codimension 2 dans  $A$ , ce qui donne un morphisme injectif de  $C(-q|_V)$  dans  $\mathcal{L}(A)$ . La comparaison des dimensions donne alors  $2^{n-2} \leq n^2$ , et on en tire facilement  $n \leq 8$ . Le cas  $n = 6$  peut ensuite être écarté en utilisant plus finement encore les théorèmes de structure sur les algèbres de Clifford.

On pressent à ce stade que les méthodes d'algèbres de Clifford doivent permettre d'élucider totalement la structure recherchée, et c'est effectivement ce qui s'est produit. Et comme il n'y a pas de grosse surprise dans les méthodes employées, je me contenterai d'achever cet exposé en citant les principaux résultats obtenus à l'issue de cette étude. Ils apportent une réponse définitive aux questions posées initialement.

**Théorème 3.17** (Théorème 1.2 de [3]). *Toute algèbre à division LDB dégénérée est faiblement équivalente à une algèbre à division LDB associée à une extension hyper-radicielle.*

*Toute algèbre à division LDB non dégénérée est faiblement équivalente à une algèbre à division LDB associée à une algèbre de Hurwitz.*

**Théorème 3.18** (Théorème 1.5 de [3]). *Deux algèbres à division LDB sont équivalentes si et seulement si les formes quadratiques associées sont équivalentes.*

*Deux algèbres à division LDB sont faiblement équivalentes si et seulement si les formes quadratiques associées sont semblables.*

### 3.8 Réflexivité algébrique et rang minimal sur un corps fini

Une dernière question était celle de la validité de la conclusion du théorème 3.3 pour n'importe quel corps.

Dans [4], j'ai montré qu'elle tenait effectivement pour n'importe quel corps fini, ce qui aboutit au résultat définitif suivant :

**Théorème 3.19.** *Soit  $\mathcal{T}$  un sous-espace vectoriel non réflexif de  $\mathcal{L}(U, V)$ , de dimension  $n \geq 2$ . L'espace  $\mathcal{T}$  possède alors un élément  $f \neq 0$  tel que  $\text{rg } f \leq 2n - 2$ .*

Donnons rapidement quelques idées de démonstration : on part de  $f \in \mathcal{R}(\mathcal{T}) \setminus \mathcal{T}$  et on construit l'espace localement lié  $\mathcal{S} := \mathcal{T} \oplus \mathbb{K}f$ . Pour une fois, nous n'allons pas nous limiter à l'observation que  $\mathcal{T}$  est un hyperplan vectoriel d'un espace localement lié, nous allons utiliser plus spécifiquement l'observation que  $f$  appartient à la clôture réflexive de  $\mathcal{S}$ , ce qui se traduit par le fait que tout vecteur de  $U$  est annulé par un élément de l'hyperplan affine  $\mathcal{H} := f + \mathcal{T}$ , et pas seulement par un élément non nul de  $\mathcal{S}$ !

La démonstration se fait alors par l'absurde, en supposant  $\mathcal{T}$  dénué d'élément non nul de rang inférieur à  $2n - 2$ . On montre successivement :

- que  $\mathcal{H}$  contient un élément de rang inférieur à  $n - 1$  ;
- puis que  $\mathcal{H}$  contient au plus un élément de rang minimal ;
- que le rang minimal dans  $\mathcal{H}$  est  $n - 1$  ;

et on conclut enfin à une contradiction. En dehors de l'étape d'unicité d'un élément de  $\mathcal{H}$  de rang strictement inférieur à  $n$  (qui repose sur l'hypothèse de base du raisonnement par l'absurde), tous les arguments reposent sur un dénombrement des vecteurs des noyaux des différents opérateurs de  $\mathcal{H}$ .

### 3.9 Perspectives

Mes publications sur les espaces localement liés ont totalement refondé le thème, mettant au centre la méthode de dualité opérateur-vecteur, mais elles ont aussi eu pour effet d'assécher le sujet. Les pistes ouvertes par Meshulam et Šemrl ont été largement refermées, les études d'optimalité ont révélé des structures très intéressantes que j'ai complètement élucidées, et j'ai du mal à concevoir des prolongements substantiels sur ce thème. Tout juste pourrait-on examiner ce qu'il advient de la borne  $2n - 2$  du théorème 3.3 pour les corps finis lorsqu'on remplace la propriété de non-réflexivité par le fait d'être un hyperplan vectoriel d'un espace localement lié, mais cela ne devrait pas mener bien loin.

Il est en tout cas significatif que plus aucune publication sur le sujet n'ait vu le jour depuis 2014.

## Chapitre 4

# Morphismes Im-compatibles

### 4.1 Articles concernés

- [1] C. de Seguins Pazzis, The linear preservers of non-singularity in a large space of matrices. *Linear Algebra Appl.* **436** (2012) 3507–3530.
- [2] C. de Seguins Pazzis, The classification of large spaces of matrices with bounded rank. *Israel J. Math.* **208** (2015) 219–259.
- [3] C. de Seguins Pazzis, Range-compatible homomorphisms on matrix spaces. *Linear Algebra Appl.* **484** (2015) 237–289.
- [4] C. de Seguins Pazzis, Large spaces of bounded rank matrices revisited. *Linear Algebra Appl.* **504** (2016) 124–189.
- [5] C. de Seguins Pazzis, Quasi-range-compatible affine maps on large operator spaces. *Linear Multilinear Algebra* **64** (2016) 1056–1085.
- [6] C. de Seguins Pazzis, Range-compatible homomorphisms on spaces of symmetric or alternating matrices. *Linear Algebra Appl.* **503** (2016) 135–163.
- [7] C. de Seguins Pazzis, Large spaces of symmetric or alternating matrices with bounded rank. *Linear Algebra Appl.* **508** (2016) 146–189.
- [8] C. de Seguins Pazzis, Primitive spaces of matrices with upper rank two over the field with two elements. *Linear Multilinear Algebra* **64** (2016) 1321–1353.
- [9] C. de Seguins Pazzis, Range-compatible homomorphisms over the field with two elements. *Electron. J. Linear Algebra* **34** (2018) 71–114.

### 4.2 Fonctions Im-compatibles

**Définition 6.** Soit  $U, V$  deux espaces vectoriels de dimension finie et  $\mathcal{S}$  une partie de  $\mathcal{L}(U, V)$ . Une fonction  $F : \mathcal{S} \rightarrow V$  est dite **Im-compatible**<sup>1 2</sup> lorsque

$$\forall s \in \mathcal{S}, F(s) \in \text{Im } s,$$

autrement dit lorsque  $\forall s \in \mathcal{S}, \exists x \in U : F(s) = s(x)$ .

---

1. La terminologie nous est propre.  
2. En anglais, « range-compatible ».

En version matricielle, la définition prend la forme suivante :

**Définition 7.** Soit  $\mathcal{M}$  une partie de  $M_{n,p}(\mathbb{K})$ . Une fonction  $F : \mathcal{M} \rightarrow \mathbb{K}^n$  est dite **Im-compatible** lorsqu'elle associe à toute matrice de  $\mathcal{M}$  une combinaison linéaire de ses colonnes.

Exemple essentiel de telles fonctions : les **évaluations**, autrement dit les fonctions de la forme

$$\hat{x} : s \in \mathcal{S} \mapsto s(x),$$

pour un vecteur  $x \in U$  fixé. On dit que  $F$  est **locale** lorsque c'est une évaluation. En version matricielle, la fonction  $F : \mathcal{M} \rightarrow \mathbb{K}^n$  est locale lorsqu'elle s'écrit  $M \mapsto MX$  pour un vecteur  $X$  fixé de  $\mathbb{K}^p$ , autrement dit lorsque  $F(M)$  est combinaison linéaire des colonnes de  $M$  à coefficients indépendants de  $M$ .

Lorsque  $U$  ou  $V$  est de dimension 1, on démontre facilement que toute fonction linéaire Im-compatible sur  $\mathcal{S}$  est locale (dans le cas où  $U$  est de dimension 1, ce résultat est une reformulation de la caractérisation classique des homothéties d'un espace vectoriel). Il est alors naturel de se poser la question suivante :

Étant donné un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$ , est-ce que toute fonction *linéaire* Im-compatible  $F : \mathcal{S} \rightarrow V$  est locale ? Et est-ce que tout morphisme (de groupes) Im-compatible  $F : \mathcal{S} \rightarrow V$  est local ?

Bien sûr, il n'y a rien à attendre d'intéressant sur les fonctions Im-compatibles sans hypothèse de nature algébrique sur la fonction  $F$  et l'ensemble  $\mathcal{S}$ .

La réponse aux questions précédentes est – on pouvait s'en douter – négative en toute généralité. Voici deux exemples matriciels simples :

*Exemple 4.1.* On considère le sous-espace vectoriel

$$\mathcal{C} := \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mid (a, b) \in \mathbb{R}^2 \right\} \in M_2(\mathbb{R})$$

et la fonction linéaire

$$F : \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \mapsto \begin{bmatrix} a \\ 0 \end{bmatrix}.$$

On vérifie sans peine que  $F$  est Im-compatible (remarquer que tout élément non nul de  $\mathcal{C}$  est inversible) mais non locale.

*Exemple 4.2.* On considère le sous-espace vectoriel

$$\mathcal{U}_2(\mathbb{K}) := \left\{ \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \mid (a, b) \in \mathbb{K}^2 \right\} \subset M_2(\mathbb{K})$$

et la fonction linéaire

$$G : M(a, b) := \begin{bmatrix} a & b \\ 0 & a \end{bmatrix} \mapsto \begin{bmatrix} 0 \\ a \end{bmatrix}.$$

Si  $a \neq 0$  alors  $M(a, b)$  est inversible et automatiquement  $G(M(a, b))$  est dans son image, sinon  $G(M(a, b))$  est nulle donc aussi dans l'image de  $M(a, b)$ . En revanche, un calcul élémentaire permet de voir que  $G$  n'est pas locale.

Enfin l'exemple suivant est spectaculaire mais tout à fait spécifique aux corps à deux éléments.



*Exemple 4.3.* L'application  $\Delta$  qui à une matrice de  $S_n(\mathbb{F}_2)$  (symétrique de taille  $n$  à coefficients dans  $\mathbb{F}_2$ ) associe son vecteur diagonal  $\Delta(M)$  est Im-compatible. En effet, pour tout  $M \in S_n(\mathbb{F}_2)$ , il suffit de montrer que  $\Delta(M)$  est orthogonal au noyau de  $M$ , or pour un vecteur  $X$  de  $\text{Ker } M$  on dispose de l'égalité  $X^T M X = 0$ , qui se réécrit  $(X^T \Delta(M))^2 = 0$ .

En revanche, il est facile de voir que  $\Delta$  n'est pas locale dès que  $n \geq 2$ .

Il y a néanmoins un cas spécifique tout à fait intéressant :

**Théorème 4.1.** *Soit  $U$  et  $V$  deux espaces vectoriels de dimension finie. Toute fonction linéaire Im-compatible de  $\mathcal{L}(U, V)$  dans  $V$  est alors locale.*

Ce résultat n'est pas difficile (c'est un bon exercice pour taupins : il généralise la caractérisation classique des homothéties d'un espace vectoriel), et on peut suspecter qu'il ait été découvert par divers mathématiciens de manière indépendante à diverses époques. Dans le paragraphe suivant, nous allons donner quelques motivations pour étudier la question des morphismes Im-compatibles sur des sous-espaces plus variés.

### 4.3 Éléments de motivation

Nous avons déjà vu à la fin du paragraphe 2.5.2 comment les morphismes Im-compatibles apparaissaient naturellement pour étudier la structure des sous-espaces vectoriels de petite codimension à rang majoré. Rappelons rapidement l'idée : dans une situation réduite, on se donnait un sous-espace vectoriel  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  (où  $n > p$ ) et une fonction linéaire  $\varphi : \mathcal{V} \rightarrow \mathbb{K}^n$  telle que

$$\forall N \in \mathcal{V}, \text{rg} \begin{bmatrix} N & \varphi(N) \end{bmatrix} \leq p.$$

Sous réserve que la codimension de  $\mathcal{V}$  soit assez petite, la version affine du théorème de Flanders permettait d'établir que  $\varphi$  est Im-compatible. Si l'on peut obtenir que  $\varphi$  est locale, alors elle s'écrit  $N \mapsto NX$  pour un  $X \in \mathbb{K}^p$ , et le vecteur non nul  $Y := \begin{bmatrix} X \\ -1 \end{bmatrix}$  est alors dans le noyau de chaque matrice de la forme  $\begin{bmatrix} N & \varphi(N) \end{bmatrix}$ , ce qui permet de voir que le sous-espace formé par ces matrices est équivalent à un sous-espace vectoriel de  $\mathcal{R}(0, p-1)$  (voir le chapitre 2 pour les notations).

Les morphismes Im-compatibles interviennent également dans l'étude des fonctions linéaires conservant l'inversibilité (voir la section 5.6). Dans [1], nous étudions, pour un sous-espace vectoriel  $\mathcal{V}$  de  $M_n(\mathbb{K})$  de codimension assez petite (concrètement, au plus  $n-2$ ), les fonctions linéaires  $\Phi : \mathcal{V} \rightarrow M_n(\mathbb{K})$  telles que  $\Phi(M)$  soit inversible si et seulement si  $M$  est inversible (conservation de l'inversibilité au sens fort). Après un travail substantiel, on démontre que, quitte à composer  $\varphi$  à gauche par la transposition et/ou par une fonction de la forme  $M \mapsto PM$  avec  $P \in \text{GL}_n(\mathbb{K})$ , la situation se réduit à celle où  $\text{Im } \Phi(M) \subset \text{Im } M$  pour tout  $M \in M_n(\mathbb{K})$ . En d'autres termes, dans cette situation chacune des fonctions

$$\Phi_k : M \in \mathcal{V} \mapsto C_k(\Phi(M)), \quad \text{où } 1 \leq k \leq n$$

est linéaire et Im-compatible. Si toutes ces fonctions sont locales, on en déduit que  $\Phi : M \mapsto MQ$  pour une matrice  $Q \in M_n(\mathbb{K})$ . On peut alors établir que

$Q$  est nécessairement inversible (typiquement, car  $\Phi(\mathcal{V})$  contient une matrice inversible, toujours par le théorème de Dieudonné). On obtiendrait alors que  $\Phi$  se prolonge en un morphisme de Frobenius (voir le chapitre 5).

Les morphismes Im-compatibles sont enfin liés, par dualité opérateur-vecteur, à la notion de réflexivité algébrique que nous avons vue dans le chapitre 3. Rappelons que la clôture réflexive d'un sous-espace vectoriel  $\mathcal{S}$  de  $\mathcal{L}(U, V)$  est définie comme

$$\mathcal{R}(\mathcal{S}) := \{g \in \mathcal{L}(U, V) : \forall x \in U, \exists f \in \mathcal{S} : g(x) = f(x)\},$$

et que c'est un sous-espace vectoriel de  $\mathcal{L}(U, V)$  incluant  $\mathcal{S}$ . On dit que  $\mathcal{S}$  est (algébriquement) réflexif lorsque  $\mathcal{R}(\mathcal{S}) = \mathcal{S}$ .

Rappelons aussi la définition de l'espace dual-opérateur : étant donné  $x \in U$ , on pose  $\hat{x} : \varphi \in \mathcal{S} \mapsto \varphi(x) \in V$ , et

$$\hat{\mathcal{S}} := \{\hat{x} \mid x \in U\} \subset \mathcal{L}(\mathcal{S}, V).$$

Bien sûr,  $x \in U \mapsto \hat{x} \in \hat{\mathcal{S}}$  est linéaire et surjective, et même bijective si  $\mathcal{S}$  est réduit au départ (voir le paragraphe 3.4).

Donnons-nous alors  $g \in \mathcal{R}(\mathcal{S})$ . Étant donné  $x \in U$ , la condition  $\mathcal{S}x = \{0\}$  implique  $g(x) = 0$ . La fonction  $g$  induit donc une fonction linéaire  $G : \hat{\mathcal{S}} \rightarrow V$  telle que  $G(\hat{x}) = g(x)$  pour tout  $x \in U$ . L'hypothèse  $g \in \mathcal{R}(\mathcal{S})$  traduit alors l'Im-compatibilité de  $G$ . En outre, la condition  $g \in \mathcal{S}$  est équivalente au fait que  $G$  soit locale. On met en évidence, par ce biais, un isomorphisme

$$\mathcal{R}(\mathcal{S})/\mathcal{S} \xrightarrow{\cong} \mathcal{L}_{\text{IC}}(\hat{\mathcal{S}}, V)/\mathcal{L}_{\text{loc}}(\hat{\mathcal{S}}, V),$$

où  $\mathcal{L}_{\text{IC}}(\hat{\mathcal{S}}, V)$  et  $\mathcal{L}_{\text{loc}}(\hat{\mathcal{S}}, V)$  désignent respectivement le sous-espace vectoriel de  $\mathcal{L}(\hat{\mathcal{S}}, V)$  formé des fonctions Im-compatibles, et celui formé des fonctions locales.

On peut aussi mettre en évidence un isomorphisme

$$\mathcal{L}_{\text{IC}}(\mathcal{S}, V)/\mathcal{L}_{\text{loc}}(\mathcal{S}, V) \xrightarrow{\cong} \mathcal{R}(\hat{\mathcal{S}})/\hat{\mathcal{S}}.$$

En particulier, toute application linéaire Im-compatible sur  $\mathcal{S}$  est locale si et seulement si  $\hat{\mathcal{S}}$  est réflexif. Et, réciproquement,  $\mathcal{S}$  est réflexif si et seulement si toute application linéaire Im-compatible sur  $\hat{\mathcal{S}}$  est locale.

## 4.4 Principaux résultats

Les deux premiers éléments de motivation du paragraphe précédent m'ont poussé à m'intéresser prioritairement aux morphismes Im-compatibles définis sur des espaces d'opérateurs de *petite codimension*. Le cas de la codimension nulle est déjà connu (c'est le théorème 4.1). Dans toute la suite,  $U$  et  $V$  désignent deux espaces vectoriels de dimension finie sur  $\mathbb{K}$ . Notre premier résultat substantiel fut le suivant, dans l'ordre chronologique :

**Théorème 4.2** (Lemme 8 de [2]). *Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$  tel que  $\text{codim } \mathcal{S} \leq \dim V - 2$ . Toute fonction linéaire Im-compatible sur  $\mathcal{S}$  est alors locale.*

La borne  $\dim V - 2$  n'est cependant pas optimale ! On peut sensiblement l'améliorer, comme nous le découvrièmes plus tard. En fait, c'est dans le cadre des morphismes de groupes qu'elle est optimale :

**Théorème 4.3** (Théorème 1.3 de [3]). *Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$  tel que  $\text{codim } \mathcal{S} \leq \dim V - 2$ . Tout morphisme (de groupes additifs) Im-compatible sur  $\mathcal{S}$  est alors local.*

L'optimalité est simple à voir dans le cas matriciel. Pour cela (et en vue d'exemples ultérieurs), il est utile d'introduire une notation pour les *prolongements naturels* d'un morphisme Im-compatible matriciel : soit  $\mathcal{M}$  un sous-espace vectoriel de  $M_{n,p}(\mathbb{K})$ , et soit  $n' \geq n$  et  $p' \geq p$  des entiers. Nous notons  $\widetilde{\mathcal{M}}^{(n',p')}$  l'ensemble des matrices de la forme

$$\begin{bmatrix} A & [?]_{n \times (p'-p)} \\ [0]_{(n'-n) \times p} & [?]_{(n'-n) \times (p'-p)} \end{bmatrix} \quad \text{où } A \in \mathcal{M}.$$

C'est évidemment un sous-espace vectoriel de  $M_{n',p'}(\mathbb{K})$  de codimension  $(n' - n)p + \text{codim } \mathcal{M}$ . En outre, étant donné un morphisme Im-compatible  $F : \mathcal{M} \rightarrow \mathbb{K}^n$ , il est bien clair que

$$\widetilde{F}^{(n',p')} : \begin{bmatrix} A & [?]_{n \times (p'-p)} \\ [0]_{(n'-n) \times p} & [?]_{(n'-n) \times (p'-p)} \end{bmatrix} \mapsto \begin{bmatrix} F(A) \\ [0]_{(n'-n) \times 1} \end{bmatrix}$$

est encore un morphisme Im-compatible, et il est local si et seulement si  $F$  est local.

Supposons maintenant que le corps  $\mathbb{K}$  ne soit pas premier. Il existe alors un endomorphisme  $\varphi : \mathbb{K} \rightarrow \mathbb{K}$  du groupe  $(\mathbb{K}, +)$  qui n'est pas  $\mathbb{K}$ -linéaire. Comme  $\varphi(0) = 0$ , on voit que  $\varphi$  est un morphisme Im-compatible non linéaire de  $M_1(\mathbb{K})$  dans  $\mathbb{K}^1$ , donc son extension  $\widetilde{\varphi}^{(n,p)}$  donne un morphisme Im-compatible non linéaire de  $\widetilde{M_1(\mathbb{K})}^{(n,p)}$  dans  $\mathbb{K}^n$  dont l'espace de départ est de codimension  $n - 1$  dans  $M_{n,p}(\mathbb{K})$ . En particulier,  $\widetilde{\varphi}^{(n,p)}$  n'est pas local.

L'extension aux fonctions seulement additives est motivée par celle aux fonctions quasi-linéaires (la somme de deux fonctions quasi-linéaires associées à des automorphismes distincts du corps  $\mathbb{K}$  n'est pas quasi-linéaire en général, mais c'est au moins un morphisme de groupes).

Revenons au théorème 4.2 : comme annoncé, la borne  $\dim V - 2$  n'est pas optimale. Le théorème suivant donne la meilleure borne possible sur un corps arbitraire :

**Théorème 4.4.** *Soit  $\mathcal{S}$  un sous-espace vectoriel de  $\mathcal{L}(U, V)$ , où  $\dim U \geq 2$  et  $\dim V \geq 2$ , tel que  $\text{codim } \mathcal{S} \leq 2 \dim V - 3$  si  $|\mathbb{K}| > 2$ , et  $\text{codim } \mathcal{S} \leq 2 \dim V - 4$  sinon. Alors, toute application linéaire Im-compatible sur  $\mathcal{S}$  est locale.*

L'optimalité de la borne est illustrée, pour un corps ayant plus de deux éléments, par une extension du morphisme présenté dans l'exemple 4.2.

Le cas du corps  $\mathbb{F}_2$  est plus délicat. Il est lié à l'exemple 4.3, dont le cas particulier  $n = 2$  montre, après extension, l'optimalité de la borne  $2 \dim V - 4$  dans le théorème 4.4.

## 4.5 Méthodes

La première démonstration publiée du théorème 4.2 (voir [2]) reposait sur une récurrence sur la dimension de l'espace de départ  $U$ . Il est apparu que cette méthode avait atteint ses limites et ne permettait pas d'attaquer un énoncé comme le théorème 4.4.

La méthode qui s'est révélée la plus fructueuse consiste à procéder plutôt par récurrence sur la dimension de l'espace d'arrivée  $V$ . Elle est fondée sur l'observation élémentaire que la propriété d'Im-compatibilité est préservée par passage au quotient dans l'espace d'arrivée : c'est la technique dite de **projection**. Donnons-nous en effet un sous-espace vectoriel  $V_0$  de  $V$ , et notons  $\pi : V \rightarrow V/V_0$  la projection canonique associée. Par composition, chaque élément  $s$  de  $\mathcal{S}$  donne naissance à l'opérateur  $\pi \circ s \in \mathcal{L}(U, V/V_0)$ , et on dispose donc d'un espace d'opérateurs

$$\mathcal{S} \bmod V_0 := \{\pi \circ s \mid s \in \mathcal{S}\} \subset \mathcal{L}(U, V/V_0).$$

De manière remarquable, tout morphisme Im-compatible  $F : \mathcal{S} \rightarrow V$  induit un morphisme Im-compatible  $F \bmod V_0$  de  $\mathcal{S} \bmod V_0$  dans  $V/V_0$ , comme suit : le morphisme de groupes  $s \in \mathcal{S} \mapsto [F(s)] \in V/V_0$  est nul sur le noyau de  $s \mapsto \pi \circ s$ , car pour tout  $s \in \mathcal{S}$ , si  $\pi \circ s = 0$  alors  $\text{Im } s \subset V_0$  et donc  $F(s) \in V_0$  par Im-compatibilité de  $F$ . Ainsi, ce morphisme se factorise via la projection canonique de  $\mathcal{S}$  sur  $\mathcal{S} \bmod V_0$ , d'où le diagramme commutatif suivant :

$$\begin{array}{ccc} \mathcal{S} & \xrightarrow{F} & V \\ s \mapsto \pi \circ s \downarrow & & \downarrow \pi \\ \mathcal{S} \bmod V_0 & \xrightarrow{F \bmod V_0} & V/V_0. \end{array}$$

Bien sûr,  $F \bmod V_0$  renferme une partie de l'information sur la structure de  $F$ , et en considérant  $F \bmod V_1$  et  $F \bmod V_2$  pour deux sous-espaces  $V_1$  et  $V_2$  de  $V$  d'intersection nulle, on récupère *toute* l'information sur  $F$ .

Dans la pratique, on quotiente le plus souvent par des *droites* : on se donne ainsi un vecteur  $y \in V \setminus \{0\}$  et on examine  $F \bmod \mathbb{K}y$  (plus simplement notée  $F \bmod y$ , tandis que  $\mathcal{S} \bmod \mathbb{K}y$  est aussi notée  $\mathcal{S} \bmod y$ ). Et pour faire tourner les récurrences, il est essentiel de remarquer que le théorème du rang fournit l'égalité

$$\text{codim}(\mathcal{S} \bmod y) = \text{codim } \mathcal{S} - \text{codim}_{\mathcal{L}(U, \mathbb{K}y)}(\mathcal{S} \cap \mathcal{L}(U, \mathbb{K}y)).$$

À titre d'illustration, je vais maintenant donner une démonstration presque intégralement détaillée du théorème 4.3.

*Démonstration du théorème 4.3.* On procède par récurrence sur la dimension de  $V$ . Le cas où  $\dim V < 2$  est trivial. Supposons donc  $\dim V \geq 2$ . Un vecteur  $y \in V$  sera dit **adapté à  $\mathcal{S}$**  lorsqu'il est non nul et  $\mathcal{L}(U, \mathbb{K}y) \not\subset \mathcal{S}$ . Dans ce cas, on a bien  $\text{codim}(\mathcal{S} \bmod y) \leq \dim(V/\mathbb{K}y) - 2$  et l'hypothèse de récurrence peut donc s'appliquer.

Si l'on prend une base de  $V$ , et si aucun de ses vecteurs n'est adapté à  $\mathcal{S}$ , alors  $\mathcal{S} = \mathcal{L}(U, V)$  et on n'a pas de mal à conclure grâce à une adaptation du théorème 4.1 aux morphismes de groupes (ce qui utilise  $\dim V \geq 2$ ).

Supposons désormais que  $\mathcal{S} \neq \mathcal{L}(U, V)$ . Alors dans toute base de  $V$  il y a automatiquement un vecteur adapté à  $\mathcal{S}$ . L'ensemble des vecteurs inadaptés à  $\mathcal{S}$  n'engendre donc pas  $V$ , ainsi il est piégé dans un hyperplan vectoriel de  $V$ . Le complémentaire de ce dernier contient au moins deux vecteurs non colinéaires  $y_1$  et  $y_2$ .

En appliquant l'hypothèse de récurrence à  $F \bmod y_1$  et  $F \bmod y_2$ , on récupère donc deux vecteurs  $x_1$  et  $x_2$  de  $U$  tels que  $F(s) = s(x_1) \bmod \mathbb{K}y_1$  et  $F(s) =$

$s(x_2) \bmod \mathbb{K}y_2$  pour tout  $s \in \mathcal{S}$ . Si  $x_1 = x_2$ , on conclut que  $F$  est l'évaluation en  $x_1$ .

Supposons enfin  $x_1 \neq x_2$ . Quitte à retrancher à  $F$  la fonction locale  $s \mapsto s(x_2)$ , on peut supposer  $x_2 = 0$  (et donc  $x_1 \neq 0$ ). Notons qu'alors  $s(x_1) \in \text{Vect}(y_1, y_2)$  pour tout  $s \in \mathcal{S}$ . En se souvenant de l'hypothèse  $\text{codim } \mathcal{S} \leq \dim V - 2$ , on obtient que  $\mathcal{S}$  est exactement l'ensemble des  $s \in \mathcal{L}(U, V)$  tels que  $s(x_1) \in \text{Vect}(y_1, y_2)$ . Prenons donc  $s \in \mathcal{S}$  de rang 1 telle que  $s(x_1) = y_1 + y_2$ . Alors  $F(s) \in \mathbb{K}(y_1 + y_2)$  par Im-compatibilité, mais  $F(s) \in \mathbb{K}y_2$  d'après ce qui précède. Donc  $F(s) = 0$ . Et à nouveau  $s(x_1) \in \mathbb{K}y_1$  par ce qui précède, ce qui est faux.  $\square$

La situation du théorème 4.4 conduit même à un autre lien remarquable avec la théorie des espaces vectoriels de rang majoré. Dans ce cadre-là, les vecteurs non nuls  $y$  inadaptés à  $\mathcal{S}$  sont en effet ceux pour lesquels  $\mathcal{S} \cap \mathcal{L}(U, \mathbb{K}y)$  est de codimension au plus 2 dans  $\mathcal{L}(U, \mathbb{K}y)$ . Or, par dualité-trace<sup>3</sup>, cette propriété revient à dire que  $t \in \mathcal{S}^\perp \mapsto t(y)$  est de rang au plus 2. Si le corps de base est infini, on montre facilement qu'il existe une base de  $V$  formée de vecteurs adaptés à  $\mathcal{S}$  pourvu qu'il y ait au moins un tel vecteur (non nul, bien sûr). Par suite, s'il n'existe pas de telle base alors le dual-opérateur  $\widehat{\mathcal{S}^\perp}$  de  $\mathcal{S}^\perp$  a tous ses éléments de rang au plus 2. Cela conduit à un lien fructueux avec la classification des espaces vectoriels de matrices de rang au plus 2 (voir en particulier [8, 9])!

## 4.6 Extensions, état des lieux

J'ai présenté plus haut les deux résultats les plus emblématiques de mes travaux sur les morphismes Im-compatibles. Le théorème 4.3 a été baptisé « Premier Théorème de Classification ». Dans [3], j'ai complètement décrit la structure des morphismes Im-compatibles (non supposés linéaires!) sous la condition de codimension du théorème 4.4 : c'est ce que j'ai baptisé le « Second Théorème de Classification ». L'étude a été complétée dans [9] par une zoologie du cas critique  $\text{codim } \mathcal{S} = 2 \dim V - 3$  sur  $\mathbb{F}_2$ , ce qui fait intervenir ma classification [8] des sous-espaces vectoriels de matrices de rang au plus 2 sur  $\mathbb{F}_2$ .

Dans [5], le théorème 4.2 a été généralisé aux applications affines Im-compatibles définies sur des sous-espaces affines de petite codimension, et j'ai également établi une version du théorème 4.4 lorsque la propriété d'Im-compatibilité est remplacée par une propriété légèrement plus faible, que j'ai baptisé **quasi-Im-compatibilité**. En bref, un morphisme  $F : \mathcal{S} \rightarrow V$  est **quasi-Im-compatibilité** lorsque  $F(s) \in \text{Im } s$  pourvu que  $\text{Im } s$  n'inclue pas une certaine droite vectorielle de  $V$ . Ces raffinements sont apparus indispensables pour attaquer l'extension du théorème d'Atkinson et Lloyd (voir le chapitre 2) à des dimensions plus basses, et ils interviennent de manière critique dans [4].

Enfin, l'équivalent du théorème 4.3 a été établi en [6] pour des espaces de matrices symétriques ou alternées. Encore une fois, l'étude est motivée par celle des espaces vectoriels de matrices symétriques/alternées de grande dimension et de rang majoré, et elle a permis de généraliser très largement des résultats de Raphael Loewy sur ces espaces (voir [7]). Une difficulté dans ce cadre a été de réaliser des démonstrations par récurrence sur la dimension. En effet, la matrice résultant de l'effacement de la dernière ligne d'une matrice

3. L'orthogonal de  $\mathcal{S}$  pour la dualité-trace, noté  $\mathcal{S}^\perp$ , est l'orthogonal à gauche de  $\mathcal{S}$  pour la forme bilinéaire non dégénérée  $(u, v) \mapsto \text{tr}(u \circ v)$  sur  $\mathcal{L}(V, U) \times \mathcal{L}(U, V)$ .

symétrique ou alternée n'est même pas carrée! La solution à cette difficulté consiste à considérer la situation plus générale des sous-espaces vectoriels de  $\mathcal{L}(U, U_0^*)$ , où  $U_0$  est un sous-espace vectoriel de  $U$  : on décrète qu'un élément  $f$  de  $\mathcal{L}(U, U_0^*)$  est symétrique (respectivement, alternée) lorsque la forme bilinéaire  $(x, y) \mapsto f(y)[x]$  est symétrique (respectivement, alternée) sur  $U_0^2$ , et on s'intéresse exclusivement aux sous-espaces vectoriels de  $\mathcal{L}(U, U_0^*)$  formés d'applications linéaires symétriques (respectivement, alternées). Le cas des matrices symétriques (respectivement, alternées) correspond à la situation où  $U = U_0$ . Enfin, la technique de projection fonctionne encore à condition, pour un sous-espace vectoriel  $H$  de  $U_0^*$ , d'identifier naturellement le quotient  $U_0^*/H$  à l'espace dual de  ${}^oH$ , prédual-orthogonal de  $H$ .

## 4.7 Bilan et perspectives

Au vu de ce qui précède, je considère que le gros du travail sur les morphismes Im-compatibles est derrière nous. Les techniques introduites en [3] sont maintenant bien établies et ont montré leur efficacité. Un élément remarquable, dans les techniques utilisées, est l'utilisation systématique de récurrence sur la dimension de l'espace d'arrivée, alors que les problèmes sur les espaces d'opérateurs ayant motivé l'étude de l'Im-compatibilité ne peuvent pas raisonnablement être traités de la sorte. L'Im-compatibilité, motivée par ces problèmes, s'est développée comme un champ quasi-autonome de résultats et de techniques, qui utilise parfois des théorèmes fins sur les sous-espaces de matrices de petit rang (voir la fin du paragraphe 4.5).

Pour la poursuite de recherches sur ce thème, je me bornerai à indiquer deux pistes qui me semblent dignes d'intérêt :

- Trouver l'équivalent du théorème 4.4 pour les espaces de matrices symétriques ou alternées. Une conjecture est faite dans [6], et il est probable qu'avec un certaine dose de travail elle puisse être confirmée avec les outils actuellement disponibles.
- Adapter le travail précédent au cadre hermitien : on munit  $\mathbb{K}$  d'une involution non triviale, on note  $\mathbb{F}$  le sous-corps des points fixes de cette involution, et on s'intéresse aux morphismes Im-compatibles définis sur un sous- $\mathbb{F}$ -espace vectoriel de celui des matrices  $\mathbb{F}$ -hermitiennes à coefficients dans  $\mathbb{K}$ , de petite codimension dans celui-ci. Bien sûr, de tels résultats devraient s'appliquer à l'obtention d'un théorème de type Atkinson-Lloyd pour les espaces de matrices hermitiennes de rang majoré.

# Chapitre 5

## Préserveurs

### 5.1 Articles concernés

- [1] C. de Seguins Pazzis, The singular linear preservers of non-singular matrices. *Linear Algebra Appl.* **433-2** (2010) 483–490.
- [2] C. de Seguins Pazzis, The affine preservers of non-singular matrices. *Arch. Math.* **95** (2010) 333–342.
- [3] C. de Seguins Pazzis, The linear preservers of non-singularity in a large space of matrices. *Linear Algebra Appl.* **436** (2012) 3507–3530.
- [4] C. de Seguins Pazzis, P. Šemrl, Hua’s fundamental theorem of geometry of rectangular matrices over EAS division rings. *J. Algebra* **439** (2015) 159–187.
- [5] C. de Seguins Pazzis, P. Šemrl, Continuous space-time transformations. *Adv. Geometry* **18-3** (2018) 385–393.
- [6] C. de Seguins Pazzis, On the linear preservers of Schur matrix functionals. *Linear Algebra Appl.* **567** (2019) 63–117.

### 5.2 Bibliographie historique

- [Fro97] G. Frobenius, Über die Darstellung der endlichen Gruppen durch Linear Substitutionen. *Sitzungsber. Deutsch. Akad. Wiss. Berlin.* (1897) 994–1015.
- [Hua45a] L.K. Hua, Geometries of matrices I. Generalizations of von Staudt’s theorem. *Trans. Amer. Math. Soc.* **57** (1945) 441–481.
- [Hua45b] L.K. Hua, Geometries of matrices  $I_1$ . Arithmetical construction. *Trans. Amer. Math. Soc.* **57** (1945) 482–490.
- [Hua46a] L.K. Hua, Geometries of symmetric matrices over the real field I. *Dokl. Akad. Nauk. SSSR* **53** (1946) 95–97.
- [Hua46b] L.K. Hua, Geometries of symmetric matrices over the real field II. *Dokl. Akad. Nauk. SSSR* **53** (1946) 195–196.
- [Hua47a] L.K. Hua, Geometries of matrices II. Study of involutions in the geometry of symmetric matrices. *Trans. Amer. Math. Soc.* **61** (1947) 193–228.

- [Hua47b] L.K. Hua, Geometries of matrices III. Fundamental theorems in the geometries of symmetric matrices. *Trans. Amer. Math. Soc.* **61** (1947) 229–255.
- [Die48] J. Dieudonné, Sur une généralisation du groupe orthogonal à quatre variables. *Arch. Math.* **1** (1948) 282–287.
- [Hua49] L.K. Hua, Geometry of symmetric matrices over any field with characteristic other than two. *Ann. Math.* **50** (1949) 8–31.
- [Al50] A.D. Alexandrov, Seminar Report. *Uspehi Mat. Nauk* **5** (1950) 187.
- [Hua51] L.K. Hua, A theorem on matrices over a sfield and its applications. *Acta Math. Sinica* **1** (1951) 109–163.
- [Al67] A.D. Alexandrov, A contribution to chronogeometry. *Canad. J. Math.* **19** (1967) 1119–1128.
- [Bot67] E.P. Botta, Linear transformations that preserve the permanent. *Proc. Amer. Math. Soc.* **18** (1967) 566–569.
- [Al75] A.D. Alexandrov, Mappings of spaces with families of cones and space-time transformations. *Annali di Matematica* **103** (1975) 229–257.
- [Bot78] E.P. Botta, Linear maps that preserve singular and nonsingular matrices. *Linear Algebra Appl.* **20** (1978) 45–49.
- [Duf94] M.A. Duffner, Linear transformations that preserve immanants. *Linear Algebra Appl.* **197-198** (1994) 567–588.
- [Lus06] F. Lussier, F. Moulin, Endomorphismes laissant stable le groupe linéaire. *RMS* **116-3** (2006) 38–49.
- [Hua08] W.-l. Huang, P. Šemrl, Adjacency preserving maps on Hermitian matrices. *Canad. J. Math.* **60** (2008) 1050–1066.
- [Rod10] L. Rodman, P. Šemrl, A localization technique for linear preserver problems. *Linear Algebra Appl.* **433** (2010) 2257–2268.
- [Duf13] M.A. Duffner, H.F. da Cruz, A relation between the determinant and the permanent on singular matrices. *Linear Algebra Appl.* **438** (2013) 3654–3660.
- [Sem14] P. Šemrl, The optimal version of Hua’s fundamental theorem of geometry of rectangular matrices. *Mem. Amer. Math. Soc.* **232** (1089) (2014), vi+74 pp.

### 5.3 Introduction

Les questions de préservation linéaire ont été très à la mode dans la période 1970-2010. L’essentiel des grands résultats ayant été largement traité avant moi, je n’ai consacré qu’une proportion modeste de mes recherches à ces questions même si mon intérêt pour le sujet reste vif.

Je me suis principalement attaché à revenir aux problèmes classiques en leur donnant des généralisations, les poussant parfois dans des directions inattendues. On pourrait résumer mon intérêt à la question en revenant aux trois grands travaux classiques qui ont lancé le sujet. Le précurseur fut Frobenius lui-même [Fro97], qui étudia les endomorphismes de  $M_n(\mathbb{K})$  qui conservent le déterminant (sur un corps infini). Les solutions sont les fonctions de la forme

$$M \mapsto PMQ \quad \text{ou} \quad M \mapsto PM^TQ$$



où  $(P, Q) \in \mathrm{GL}_n(\mathbb{K})^2$  est tel que  $\det(PQ) = 1$ .

Dieudonné a amélioré le résultat de Frobenius en 1948 [Die48], en s'intéressant aux automorphismes  $f$  de l'espace vectoriel  $M_n(\mathbb{K})$  qui conservent l'inversibilité au sens fort (c'est-à-dire que  $f(\mathrm{GL}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$ ). Il faut mentionner que le style très littéraire de Dieudonné rend parfois difficile la compréhension des hypothèses précises : par exemple, il faut savoir qu'une « transformation » est nécessairement une bijection pour Dieudonné, et que « laisser invariant » l'ensemble  $X$  signifie pour lui que la fonction considérée  $f$  vérifie  $f(X) = X$  et non seulement  $f(X) \subset X$ . Toujours est-il que Dieudonné donna une formule explicite pour ces automorphismes, très proche de celle obtenue par Frobenius : il s'agit des applications d'une des formes

$$M \mapsto PMQ \quad \text{ou} \quad M \mapsto PM^TQ, \quad \text{où } (P, Q) \in \mathrm{GL}_n(\mathbb{K})^2.$$

Nous dirons qu'une application de ce type est un **morphisme de Frobenius** (même si on devrait parler de morphisme de Dieudonné).

Une chose remarquable dans l'article de Dieudonné est la méthode de démonstration. Pour obtenir son résultat, Dieudonné étudie les sous-espaces vectoriels (et même les sous-espaces affines !) de dimension maximale inclus dans le cône des matrices singulières, ce qui nous renvoie directement au chapitre 2. Ensuite, pour un automorphisme linéaire  $f$  de  $M_n(\mathbb{K})$  conservant l'inversibilité, Dieudonné examine l'action de  $f$  sur l'ensemble de ces sous-espaces vectoriels de codimension minimale  $n$ . Par des considérations de dimension d'intersection, il obtient que  $f$  conserve ou échange les deux orbites sous l'action naturelle de  $\mathrm{GL}_n(\mathbb{K})^2$  (par  $(P, Q).V = PVQ^{-1}$ ), met en correspondance ces orbites avec l'espace projectif  $\mathbb{P}(\mathbb{K}^n)$ , et conclut (en dimension au moins 3) grâce à une utilisation judicieuse du théorème fondamental de la géométrie projective. Dieudonné intitula son article *Sur une généralisation du groupe orthogonal à quatre variables* et il vaut la peine d'élucider ce titre un peu mystérieux. Le groupe orthogonal dont il est question est celui de la forme quadratique  $M \mapsto \det M$  sur  $M_2(\mathbb{K})$ . Les automorphismes de  $M_2(\mathbb{K})$  qui conserve la singularité sont ceux qui conservent l'isotropie pour cette forme. Classiquement, et parce que cette forme quadratique est régulière et isotrope, ces automorphismes sont précisément les similitudes pour la forme  $\det$  (autrement dit les automorphismes  $f$  pour lesquels il existe un  $\lambda \in \mathbb{K}^*$  tel que  $\forall M \in M_2(\mathbb{K}), \det(f(M)) = \lambda \det M$ ). Dieudonné aurait donc dû parler de groupe des similitudes et non de groupe orthogonal (le groupe orthogonal pour  $\det$  étant précisément obtenu comme cas particulier du théorème de Frobenius cité plus haut). On peut aussi souligner que Dieudonné ne restreint absolument pas le corps de base sinon par une hypothèse de commutativité.

Enfin, à l'origine de nos travaux les plus récents se trouve la singulière série de Lo-Keng Hua [Hua45a, Hua45b, Hua46a, Hua46b, Hua47a, Hua47b, Hua49, Hua51], qui établit les théorèmes fondamentaux de la géométrie des espaces de matrices. La géométrie considérée par Hua est la géométrie discrète définie par la distance arithmétique : la distance arithmétique de deux matrices  $A$  et  $B$  rectangulaires de même format est définie comme le rang de  $A - B$  (elle vérifie bien les axiomes d'une distance). En géométrie des matrices, deux matrices de même format sont dites adjacentes lorsqu'elle sont à distance non nulle minimale possible. Et cette géométrie se décline aussi pour des matrices spéciales : on a ainsi une géométrie des matrices rectangulaires, des matrices symétriques, des matrices hermitiennes, et des matrices alternées (où la distance minimale non

nulle possible est 2 et non pas 1, cas exceptionnel dans ces théorèmes). Et pour chacun de ces types d'espaces de matrices, Hua est parvenu à décrire les permutations  $f$  qui conservent l'adjacence pour les paires (au sens que deux éléments  $A$  et  $B$  sont adjacents si et seulement si  $f(A)$  et  $f(B)$  sont adjacents). Un des intérêts de l'approche de Hua est de ne faire strictement aucune hypothèse de linéarité, ni même d'additivité, sur les bijections qu'il manipule. Citons son principal résultat : pour un corps  $\mathbb{K}$ , toute bijection de  $M_n(\mathbb{K})$  dans lui-même qui conserve l'adjacence est de l'une des deux formes suivantes :

- ou bien  $M \mapsto A_0 + PM^\sigma Q$ , où  $A_0 \in M_n(\mathbb{K})$ ,  $P, Q$  sont dans  $GL_n(\mathbb{K})$ , et  $\sigma$  est un automorphisme du corps  $\mathbb{K}$  ;
- ou bien  $M \mapsto A_0 + P(M^\sigma)^T Q$ , où  $A_0 \in M_n(\mathbb{K})$ ,  $P, Q$  sont dans  $GL_n(\mathbb{K})$ , et  $\sigma$  est un automorphisme du corps  $\mathbb{K}$ .

La méthode de Hua ressemble sensiblement à celle de Dieudonné, en ce qu'il s'intéresse à la structure des ensembles maximaux de matrices deux à deux adjacents. Il observe qu'il y a essentiellement deux formes possibles pour ces ensembles lorsqu'ils contiennent la matrice nulle : soit l'ensemble des matrices s'annulant sur un hyperplan donné de  $\mathbb{K}^n$ , soit l'ensemble des matrices dont l'image est incluse dans une droite vectorielle donnée de  $\mathbb{K}^n$ . La démonstration procède ensuite en un examen de l'action de la transformation donnée sur ces ensembles adjacents maximaux, en exploitant le théorème fondamental de la géométrie projective (ou sa version affine, souvent plus directement manipulable).

Les travaux dont je vais rendre compte ici sont tous liés de manière directe à ces grands classiques. J'ai principalement étendu le théorème de Dieudonné dans quatre directions différentes :

- dans [1], j'ai relâché l'hypothèse de bijectivité et la conservation de l'inversibilité au sens fort, en m'intéressant plus généralement aux endomorphismes de l'espace vectoriel  $M_n(\mathbb{K})$  qui envoient toute matrice inversible sur une matrice inversible ;
- dans [2], j'ai prolongé le résultat de Dieudonné aux bijections affines, et j'ai étudié le cas très particulier de  $M_2(\mathbb{F}_2)$  et son lien avec le groupe symétrique  $\mathfrak{S}_6$  et le groupe symplectique  $Sp_4(\mathbb{F}_2)$ .
- dans [3], j'ai étudié la préservation de l'inversibilité par une injection linéaire  $\varphi : V \hookrightarrow M_n(\mathbb{K})$  partant d'un sous-espace  $V$  de petite codimension dans  $M_n(\mathbb{K})$  ;
- enfin, dans [6] j'ai prolongé le théorème de Dieudonné à la conservation du cône de nullité d'une *fonctionnelle de Schur*, notion qui englobe à la fois le déterminant et le permanent.

Dans chaque cas, je n'ai pas dérogré à l'idée fondamentale de Dieudonné (utilisation des sous-espaces de grande dimension inclus dans le cône des matrices singulières), mais l'utilisation de cette idée est déclinée très différemment selon la situation.

Mes deux autres travaux évoqués dans ce chapitre se rattachent quant à eux à ceux de Hua. Dans [4], qui a été initié par Peter Šemrl, nous avons démontré une version du théorème de Hua sur les matrices rectangulaires avec des hypothèses très affaiblies pour des corps (gauche) particuliers (qui englobent tout de même les corps finis, le corps des réels et le corps gauche des quaternions). Enfin, dans [5], Šemrl et moi avons établi une variante du théorème fondamental de la géométrie des matrices hermitiennes complexes, dans le cas  $2 \times 2$ , avec des

hypothèses affaiblies du point de vue de la préservation, mais en adoptant une hypothèse de nature topologique ; ce travail est aussi à mettre en relation avec le théorème d’Alexandrov sur les transformations de l’espace-temps qui conservent les cônes de lumière [Al50, Al67, Al75].

J’ai volontairement exclu de ce chapitre mon article sur les applications linéaires préservant la diagonalisabilité. Cet article est abordé dans le chapitre 7, et la partie la plus intéressante de celui-ci ne réside pas dans la recherche des préserveurs à proprement parler.

## 5.4 Application linéaires non bijectives conservant l’inversibilité [1]

Dans le théorème de Dieudonné [Die48], on considère les bijections linéaires  $f : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$  vérifiant  $f(\mathrm{GL}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$  (conservation de l’inversibilité au sens fort), et on trouve que ces fonctions sont les morphismes de Frobenius décrits dans l’introduction. Il est naturel de chercher à affaiblir l’une ou l’autre des hypothèses, soit en retirant l’hypothèse d’injectivité sur  $f$ , soit en affaiblissant la conservation de l’inversibilité pour la remplacer par la simple stabilisation  $f(\mathrm{GL}_n(\mathbb{K})) \subset \mathrm{GL}_n(\mathbb{K})$ . Et d’autres étaient passés avant moi pour étudier ces questions :

- Dans [Bot78], Botta démontre que si l’on remplace  $f(\mathrm{GL}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$  par  $f^{-1}(\mathrm{GL}_n(\mathbb{K})) = \mathrm{GL}_n(\mathbb{K})$ , alors la bijectivité de  $f$  est automatique. Il remarque en outre que si  $\mathbb{K}$  est algébriquement clos, alors on peut relâcher simultanément les deux hypothèses tout en conservant la même conclusion.
- Par ailleurs, une analyse fine des arguments de Dieudonné permet sans trop de peine de constater que la bijectivité et la stabilisation faible  $f(\mathrm{GL}_n(\mathbb{K})) \subset \mathrm{GL}_n(\mathbb{K})$  suffisent à faire fonctionner l’essentiel de la démonstration.

Mon attention a été attirée sur ce problème par un article de mes actuels collègues François Lussier et François Moulin [Lus06]. Dans celui-ci, ils abordent un certain nombre de thèmes classiques, et ils présentent une classe de solutions *singulières*, c’est-à-dire non bijectives, au problème de la préservation au sens faible.

Supposons disposer d’un sous-espace vectoriel  $V$  de  $M_n(\mathbb{K})$ , de dimension  $n$  et dans lequel tout élément non nul est inversible. En fixant un isomorphisme d’espaces vectoriels  $\alpha : \mathbb{K}^n \xrightarrow{\sim} V$  ainsi qu’un vecteur non nul  $X \in \mathbb{K}^n$ , on peut remarquer que chacune des applications  $M \in M_n(\mathbb{K}) \mapsto \alpha(MX) \in M_n(\mathbb{K})$  et  $M \in M_n(\mathbb{K}) \mapsto \alpha(M^T X) \in M_n(\mathbb{K})$  est linéaire et envoie toute matrice inversible sur une matrice inversible (en effet, lorsque  $M$  est inversible,  $MX$  est non nul donc  $\alpha(MX)$  aussi). Par exemple l’endomorphisme

$$\begin{bmatrix} a & c \\ b & d \end{bmatrix} \mapsto \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

de  $M_2(\mathbb{R})$  envoie toute matrice inversible sur une matrice inversible. Ici, on a pris pour  $V$  le plan vectoriel engendré par  $I_2$  et  $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ , pour  $\alpha$  l’isomorphisme naturel qui est la réciproque de l’extraction de la première colonne, et on a considéré le premier type d’application. Les espaces  $V$  envisagés sont profondément liés aux **algèbres à division** sur  $\mathbb{K}$  au sens le plus faible qui soit.

Une telle algèbre à division est simplement la donnée d'un  $\mathbb{K}$ -espace vectoriel  $D$  de dimension finie non nulle muni d'une application  $\mathbb{K}$ -bilinéaire  $\star : D^2 \rightarrow D$  qui soit non dégénérée, autrement dit  $a \star b \neq 0$  pour tous  $a, b$  non nuls dans  $D$ . Pour une telle algèbre à division, l'application  $a \mapsto a \star -$  définit un isomorphisme de  $D$  sur un sous-espace vectoriel de  $\mathcal{L}_{\mathbb{K}}(D)$  de même dimension que  $D$  et dont tout élément non nul est inversible. Et inversement, pour tout sous-espace  $V$  du type précédent, en se donnant un isomorphisme arbitraire  $\alpha : \mathbb{K}^n \xrightarrow{\sim} D$ , on récupère une structure d'algèbre à division sur  $\mathbb{K}^n$  en prenant le produit défini par  $X \star Y := \alpha(X)Y$ .

La question posée par mes collègues, et qui clot leur article, est celle de l'existence de solutions autres que les morphismes de Frobenius d'une part, et les solutions dégénérées liées aux algèbres à division d'autre part. Et mon article [1] répond entièrement à cette question par la négative : les solutions dégénérées trouvées par Moulin et Lussier sont les seules possibles. En particulier, on retrouve les résultats de Botta sur le cas algébriquement clos (dans celui-ci, toute algèbre à division est de dimension 1, et il n'y a alors pas de cas réellement dégénéré); de plus, sur le corps des réels il n'existe de solution dégénérée que pour les dimensions 2, 4 et 8, conséquence du célèbre théorème de Bott-Milnor-Kervaire sur les dimensions possibles pour une algèbre à division réelle<sup>1</sup>. En outre, la plupart des variantes du théorème de Dieudonné s'obtiennent à très peu de frais à partir de notre théorème général, que nous citons maintenant :

**Théorème 5.1** (Théorème 2 de [1]). *Soit  $\mathbb{K}$  un corps et  $n$  un entier naturel non nul. Les endomorphismes  $f$  de  $M_n(\mathbb{K})$  tels que  $f(\mathrm{GL}_n(\mathbb{K})) \subset \mathrm{GL}_n(\mathbb{K})$  sont les applications d'un des types suivants :*

- (i) *les morphismes de Frobenius ;*
- (ii) *les applications de la forme  $X \mapsto \alpha(MX)$  ou  $X \mapsto \alpha(M^T X)$  où  $X \in \mathbb{K}^n \setminus \{0\}$  et  $\alpha$  est un isomorphisme de  $\mathbb{K}^n$  sur un sous-espace vectoriel de dimension  $n$  de  $M_n(\mathbb{K})$  dont tout élément non nul est inversible.*

La méthode reprend l'idée de base de Dieudonné. Il sera utile pour en discuter d'introduire les notations suivantes :

- pour une droite vectorielle  $D$  de  $\mathbb{K}^n$ , on note  $\mathcal{L}_D$  le sous-espace vectoriel des matrices de  $M_n(\mathbb{K})$  qui s'annulent en tout vecteur de  $D$  (on parle de sous-espace vectoriel **de type noyau**) ;
- pour un hyperplan vectoriel  $H$  de  $\mathbb{K}^n$ , on note  $\mathcal{L}^H$  le sous-espace vectoriel des matrices de  $M_n(\mathbb{K})$  dont l'image est incluse dans  $H$  (on parle de sous-espace vectoriel **de type image**).

Le théorème de Dieudonné pour les sous-espaces vectoriels singuliers de  $M_n(\mathbb{K})$  indique que tout sous-espace vectoriel de matrices singulières de  $M_n(\mathbb{K})$  est de codimension au moins  $n$ , et que les espaces cités précédemment sont exactement ceux de codimension  $n$ . Prenons donc  $f : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$  linéaire et stabilisant  $\mathrm{GL}_n(\mathbb{K})$ . L'idée de base est que l'image *reciproque* par  $f$  d'un sous-espace vectoriel  $W$  de  $M_n(\mathbb{K})$  constituée de matrices singulières est aussi constitué de matrices singulières, et sa codimension dans  $M_n(\mathbb{K})$  est *au plus* celle de  $W$ . Il vient en particulier que  $f^{-1}(\mathcal{L}_D)$ , pour  $D$  droite vectorielle de

1. R. Bott, J. Milnor, On the parallelizability of the spheres. Bull. Amer. Math. Soc. **64** (1958) 87–89.

M. Kervaire, Non-parallelizability of the  $n$ -sphere for  $n > 7$ . Proc. Nat. Acad. Sci. **44** (1958) 280–283.

$\mathbb{K}^n$ , est systématiquement de type noyau ou de type image. Avec les hypothèses fortes de Dieudonné, il est facile, par des calculs de dimensions d'intersection, de montrer que la fonction ainsi créée conserve ou renverse les types (on suppose  $n \geq 2$  tout du long, puisque le cas  $n = 1$  est trivial). Ici, il est beaucoup plus subtil de montrer un résultat équivalent, mais cela fonctionne tout de même (voir la proposition 6 de [1]). On montre plus précisément que lorsque  $D, D'$  sont deux droites vectorielles de  $\mathbb{K}^n$ , les sous-espaces  $f^{-1}(\mathcal{L}_D)$  et  $f^{-1}(\mathcal{L}_{D'})$  sont tous deux de type noyau ou tous deux de type image. Et plus généralement, soit l'image réciproque de tout sous-espace de type noyau est de type noyau, soit l'image réciproque de tout sous-espace de type image est de type image. En précomposant  $f$  par la transposition, on se ramène à la première situation, que nous supposons désormais vérifiée.

Le reste de la démonstration est profondément différent de l'approche de Dieudonné, et c'est là que réside notre principale innovation. Au lieu de chercher à examiner aussi les images réciproques des sous-espaces de type image, on reste sur ceux de type noyau et on observe la signification des résultats obtenus en termes de colonnes. Si par exemple (cas auquel on se ramène facilement par composition à droite et à gauche par des morphismes de Frobenius) on a  $f^{-1}(\mathcal{L}_D) = \mathcal{L}_D$  lorsque  $D$  est engendré par le premier vecteur de la base canonique, nous savons que l'image par  $f$  d'une matrice  $M$  a sa première colonne nulle si et seulement si la première colonne de  $M$  est nulle. Le lemme de factorisation des applications linéaires indique alors que la première colonne de  $f(M)$  est une fonction linéaire de la première colonne de  $M$ , et mieux cette fonction est bijective. Autrement dit, on a l'identité  $\forall M \in M_n(\mathbb{K}), C_1(f(M)) = \varphi_1(C_1(M))$  pour un automorphisme  $\varphi_1$  de  $\mathbb{K}^n$ . Plus généralement, on dispose du résultat suivant : pour tout  $X \in \mathbb{K}^n \setminus \{0\}$ , il existe un vecteur  $X' \in \mathbb{K}^n \setminus \{0\}$  (unique à colinéarité près) et un automorphisme  $\varphi_X$  de  $\mathbb{K}^n$  tel que  $f(M)X = \varphi_X(MX')$  pour tout  $M \in M_n(\mathbb{K})$ . Ce *principe de localisation* avait été découvert parallèlement par Rodman et Šemrl sans que je le sache, et je les ai devancés dans la publication de l'application la plus remarquable de celui-ci. Ils ont publié leur article quelques semaines après le mien, avec d'autres exemples intéressants d'utilisation de cette technique [Rod10].

La fin du travail consiste à examiner comment les vecteurs  $X'$  se répartissent dans  $\mathbb{K}^n$  : on regarde le sous-espace engendré, on en extrait une base etc, et pour des vecteurs  $X_1, \dots, X_p$  correspondant à une telle base, on parvient alors à trouver une relation entre les automorphismes  $\varphi_{X_1}, \dots, \varphi_{X_p}$  (ils sont tous colinéaires!) puis à montrer que la situation  $1 < p < n$  conduirait à une contradiction. Le cas  $p = n$  donne ensuite que  $f$  est un morphisme de Frobenius (du premier type, autrement dit sans transposition), alors que le cas  $p = 1$  donne la situation dégénérée observée par Lussier et Moulin.

Quant aux prolongements possibles de cette méthode, ils ont déjà été largement explorés par Rodman et Šemrl (conservation du rang maximal pour des matrices rectangulaires non carrées, conservation du rang 1, conservation des groupes unitaires ou orthogonaux), mais on peut fort bien imaginer d'autres utilisations dans des contextes de matrices particulières.

## 5.5 Bijections affines conservant l'inversibilité [2]

Il y a une forme de paradoxe dans le travail de Dieudonné [Die48] sur les automorphismes de l'espace vectoriel  $M_n(\mathbb{K})$  qui conservent l'inversibilité. En effet, Dieudonné s'intéresse dans la première partie aux sous-espaces *affines* de  $M_n(\mathbb{K})$  constitués de matrices singulières et de dimension maximale, mais il n'applique en définitive son résultat qu'aux sous-espaces vectoriels. Il était donc naturel de s'interroger sur une description potentielle des automorphismes *affines* de  $M_n(\mathbb{K})$  qui stabilisent  $GL_n(\mathbb{K})$ . Une simple adaptation de la méthode de Dieudonné permet tout simplement de montrer le résultat suivant :

**Théorème 5.2** (Théorème 3, point (a), de [2]). *À l'exception du cas où  $n = |\mathbb{K}| = 2$ , les automorphismes affines de  $M_n(\mathbb{K})$  qui stabilisent  $GL_n(\mathbb{K})$  sont les morphismes de Frobenius.*

Concrètement, on démontre que dans le cas où  $(n, |\mathbb{K}|) \neq (2, 2)$ , tout automorphisme affine  $f$  de  $M_n(\mathbb{K})$  qui stabilise  $GL_n(\mathbb{K})$  s'avère linéaire, puis on utilise le théorème de Dieudonné, ou plutôt sa version renforcée où l'on considère la stabilisation du groupe linéaire et non son invariance globale. Pour cela, et en reprenant les notations du paragraphe précédent, on se donne une liste  $(D_1, \dots, D_n)$  de droites vectorielles de  $\mathbb{K}^n$  dont la somme vaut  $\mathbb{K}^n$ , et on remarque que  $f^{-1}\{0\} = \bigcap_{k=1}^n f^{-1}(\mathcal{L}_{D_k})$ . Par ailleurs chaque  $f^{-1}(\mathcal{L}_{D_k})$  est un sous-espace affine de codimension  $n$  constitué de matrices singulières, donc c'est un sous-espace vectoriel, et finalement  $0 \in f^{-1}\{0\}$ , d'où l'on tire la linéarité de  $f$ . En revanche, dans le cas particulier où  $n = |\mathbb{K}| = 2$ , Dieudonné avait bien remarqué l'existence d'un plan affine de matrices singulières ne passant pas par 0 : celui des matrices triangulaires supérieures de trace 1.

Laissons pour un court moment ce cas particulier de côté, et donnons sans démonstration quelques corollaires faciles du théorème précédent.

**Théorème 5.3.** *À l'exception du cas où  $n = |\mathbb{K}| = 2$  :*

- (a) *les endomorphismes affines  $f$  de  $M_n(\mathbb{K})$  tels que  $f^{-1}(GL_n(\mathbb{K})) = GL_n(\mathbb{K})$  sont les morphismes de Frobenius ;*
- (b) *les endomorphismes affines  $f$  de  $M_n(\mathbb{K})$  tels que  $f(GL_n(\mathbb{K})) = GL_n(\mathbb{K})$  sont les morphismes de Frobenius.*

À l'instar de [1], on pourrait aller encore plus loin et s'interroger sur les endomorphismes affines de  $M_n(\mathbb{K})$  qui stabilisent  $GL_n(\mathbb{K})$ . Là, une très vaste classe de solutions dégénérées existe, puisque dès que l'on se donne un sous-espace affine  $\mathcal{F}$  inclus dans  $GL_n(\mathbb{K})$ , ce qui nous renvoie au chapitre 6, n'importe quelle fonction affine de  $M_n(\mathbb{K})$  dans  $\mathcal{F}$  stabilise  $GL_n(\mathbb{K})$ . Une question ouverte naturelle est celle de l'existence de fonctions affines non bijectives de  $M_n(\mathbb{K})$  dans lui-même qui stabilisent  $GL_n(\mathbb{K})$  mais dont l'image n'est pas incluse dans  $GL_n(\mathbb{K})$ . À ce jour, nous n'en connaissons pas et n'avons pas réellement travaillé activement sur ce problème.

Il nous reste à examiner le cas particulier où  $n = 2$  et  $\mathbb{K} = \mathbb{F}_2$ . Là réside un deuxième paradoxe de l'article de Dieudonné, car ce cas met en évidence un isomorphisme entre d'une part le groupe  $\mathcal{AG}_2(\mathbb{F}_2)$  des automorphismes affines de  $M_2(\mathbb{F}_2)$  qui stabilisent (et donc laissent globalement invariant, finitude oblige) le groupe  $GL_2(\mathbb{F}_2)$ , d'autre part le groupe symplectique  $Sp_4(\mathbb{F}_2)$  et le groupe

symétrique  $\mathfrak{S}_6$ . Il est donc particulièrement surprenant que cela ait échappé à Dieudonné, qui était pourtant friand de la mise en évidence d'isomorphismes exceptionnels. Ici, la première observation est celle de l'existence d'éléments non linéaires dans  $\mathcal{AG}_2(\mathbb{F}_2)$ . Un exemple donné dans mon article est l'application

$$M \mapsto (1 + \theta(M)) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + M$$

où  $\theta$  est la forme linéaire attribuant à toute matrice  $M \in M_2(\mathbb{F}_2)$  la somme de ses quatre coefficients. On peut vérifier facilement que cette application  $u$ , qui est visiblement affine et involutive, stabilise  $GL_2(\mathbb{F}_2)$  : observer pour cela que les seules matrices inversibles  $M$  pour lesquelles  $\theta(M) = 0$  ont exactement deux coefficients non nuls.

Revenons à la situation générale en rappelant que  $\mathcal{AG}_2(\mathbb{F}_2)$  désigne le groupe des automorphismes affines de  $M_2(\mathbb{F}_2)$  stabilisant  $GL_2(\mathbb{F}_2)$ . Nous avons immédiatement un homomorphisme naturel

$$\psi : \mathcal{AG}_2(\mathbb{F}_2) \longrightarrow \mathfrak{S}(GL_2(\mathbb{F}_2)),$$

et notre résultat majeur (point (c) du théorème 3 de [2]) stipule qu'il s'agit d'un isomorphisme de groupes. Il y a six matrices dans  $GL_2(\mathbb{F}_2)$ , si bien que  $\mathcal{AG}_2(\mathbb{F}_2)$  est isomorphe à  $\mathfrak{S}_6$ .

Nous terminerons en donnant les grandes lignes de la démonstration. D'abord il est immédiat que le morphisme mis en évidence est injectif, en remarquant que  $GL_2(\mathbb{F}_2)$  est une partie génératrice de l'espace affine  $M_2(\mathbb{K})$ . Ensuite, nous avons mis en évidence un lien entre le groupe symplectique d'une forme symplectique  $B$  de rang 4 sur  $\mathbb{F}_2$  d'une part, et  $\mathcal{AG}_2(\mathbb{F}_2)$  d'autre part. Soit donc  $B$  une telle forme, d'espace de définition  $V$ . Les formes quadratiques sur  $V$  de forme polaire  $B$  forment un espace affine  $\mathcal{Q}(B)$  de dimension 4, contenant deux classes d'équivalence (au sens de l'équivalence entre formes quadratiques). Ces deux classes d'équivalence sont repérées par l'invariant d'Arf des formes en question. Et l'invariant d'Arf s'identifie à un déterminant. Plus concrètement, si l'on se place dans une base symplectique  $(e_1, e_2, f_1, f_2)$  de  $V$ , si bien que la matrice de  $B$  dans cette base s'écrive  $\begin{bmatrix} 0 & I_2 \\ I_2 & 0 \end{bmatrix}$ , l'ensemble  $\mathcal{Q}(B)$  est mis en correspondance (affine) avec l'ensemble des matrices de la forme

$$M(a, b, c, d) := \begin{bmatrix} a & 0 & 1 & 0 \\ 0 & b & 0 & 1 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & d \end{bmatrix},$$

et l'invariant d'Arf de la forme correspondant à  $M(a, b, c, d)$  est  $ac + bd$ . On dispose donc d'un isomorphisme d'espaces affines de  $\mathcal{Q}(B)$  sur  $M_2(\mathbb{F})$  attribuant à la forme correspondant à  $M(a, b, c, d)$  la matrice  $\begin{bmatrix} a & d \\ b & c \end{bmatrix}$ , isomorphisme qui transforme invariant d'Arf en déterminant.

Par ailleurs, en attribuant à toute transformation symplectique  $u \in Sp(B)$  l'automorphisme affine  $q \in \mathcal{Q}(B) \mapsto q \circ u^{-1} \in \mathcal{Q}(B)$ , on obtient un homomorphisme de  $Sp(B)$  sur le groupe des transformations affines de  $\mathcal{Q}(B)$  conservant l'invariant d'Arf, et il est facile d'observer que cet homomorphisme est

injectif, et mieux encore qu'il induit un isomorphisme de  $\mathrm{Sp}(B)$  sur le groupe des permutations de l'ensemble  $\mathcal{Q}_1(B)$  formé des éléments de  $\mathcal{Q}(B)$  d'invariant d'Arf 1. Ce point s'obtient en constatant d'une part que les deux groupes ont le même cardinal, et d'autre part que pour toute transformation symplectique  $u \in \mathrm{Sp}(B) \setminus \{\mathrm{id}\}$  on peut expliciter une forme quadratique  $q \in \mathcal{Q}_1(B)$  telle que  $q \circ u \neq q$ . Cette observation était déjà au cœur du chapitre XXXIII de mon ouvrage sur les formes quadratiques<sup>2</sup>.

Via l'identification précédente, on en déduit alors un homomorphisme  $\varphi : \mathrm{Sp}(B) \rightarrow \mathcal{AG}(\mathbb{F}_2)$  dont la composée par  $\psi$  se révèle être bijective. Il vient que  $\varphi$  et  $\psi$  sont eux-mêmes des isomorphismes, ce qui achève l'étude.

Enfin, en remarquant que la forme polaire  $B$  du déterminant sur  $M_2(\mathbb{F}_2)$  est symplectique, on peut facilement créer, par passage à la partie linéaire, un homomorphisme naturel  $u \in \mathcal{AG}_2(\mathbb{F}_2) \mapsto \vec{u} \in \mathrm{Sp}(B)$ . Son injectivité est alors facile, puis l'égalité des cardinaux déjà établie permet de conclure que l'on a mis en évidence un nouvel isomorphisme de groupes, plus canonique que les précédents à vrai dire.

## 5.6 Injections linéaires conservant l'inversibilité à partir d'un grand espace de matrices [3]

Les travaux précédents parlaient d'applications définies sur des espaces de matrices tout à fait précis (systématiquement  $M_n(\mathbb{K})$ ). À cette époque, je m'intéressais très généralement aux propriétés des sous-espaces de matrices de petite codimension. En particulier, un corollaire du théorème de Dieudonné [Die48] sur les sous-espaces affines de matrices singulières de  $M_n(\mathbb{K})$  est qu'à l'exception du cas où  $n = |\mathbb{K}| = 2$ , tout sous-espace vectoriel  $V$  de  $M_n(\mathbb{K})$  de codimension strictement inférieure à  $n$  contient non seulement au moins une matrice inversible, mais en possède toute une base. Pour le voir, on remarque que s'il n'existait pas de base de  $V$  formée de matrices inversibles, les matrices inversibles de  $V$  seraient piégées dans un hyperplan vectoriel  $H$  de  $V$ , et tout hyperplan affine strictement parallèle à  $H$  serait constitué de matrices singulières, violant le théorème de Dieudonné pour les sous-espaces affines. Cette observation a été une des clefs de mon extension à un corps quelconque d'un théorème d'Atkinson et Lloyd (voir le chapitre 2). À partir de là, il était donc naturel d'examiner les automorphismes d'un tel espace vectoriel  $V$  qui conservent l'inversibilité, et d'examiner s'ils possèdent encore la forme de Frobenius.

Une difficulté réside dans le contre-exemple suivant : si l'on prend, pour  $n \geq 3$ , le sous-espace  $V$  des matrices de la forme  $\begin{bmatrix} N & C \\ 0 & a \end{bmatrix}$  avec  $N \in M_{n-1}(\mathbb{K})$ ,  $C \in \mathbb{K}^{n-1}$  et  $a \in \mathbb{K}$ , l'application linéaire

$$u : \begin{bmatrix} N & C \\ 0 & a \end{bmatrix} \mapsto \begin{bmatrix} N & C + n_{2,2} E_{1,1} \\ 0 & a \end{bmatrix}$$

s'avère être injective, et elle conserve l'inversibilité (car l'inversibilité s'observe sur les seules propriétés des blocs diagonaux). Et il est pourtant aisé de vérifier que  $u$  n'a pas la forme d'un morphisme de Frobenius. Ici,  $V$  a pour codimension

<sup>2</sup>. C. de Seguins Pazzis, Invitation aux formes quadratiques. Mathématiques en devenir, Calvage et Mounet, Paris, 2011.



$n - 1$ , qui est pourtant la codimension critique en deçà de laquelle on est certain que  $V$  est engendré par ses matrices inversibles! Ainsi, on doit se limiter aux sous-espaces  $V$  de codimension au plus  $n - 2$  (et non  $n - 1$ ) pour espérer généraliser le théorème de Dieudonné.

Voici donc notre théorème principal :

**Théorème 5.4** (Théorème 1 de [3]). *Soit  $V$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  tel que  $\text{codim } V \leq n - 2$ , et  $f : V \hookrightarrow M_n(\mathbb{K})$  une injection linéaire telle que  $f(M) \in \text{GL}_n(\mathbb{K}) \Leftrightarrow M \in \text{GL}_n(\mathbb{K})$  pour tout  $M \in V$ . Alors, hormis dans le cas où  $(n, |\mathbb{K}|, \text{codim } V) = (3, 2, 1)$ , la fonction  $f$  est la restriction à  $V$  d'un morphisme de Frobenius.*

On notera qu'ici on fait une hypothèse de conservation d'inversibilité au sens fort. Nous discuterons ultérieurement du cas particulier  $(n, |\mathbb{K}|, \text{codim } V) = (3, 2, 1)$ .

Examinons les grandes lignes de la démonstration. Nous supposons  $n > 2$  (sinon on est directement ramené au théorème de Dieudonné). Le début de la démarche est sensiblement le même que dans le théorème de Dieudonné : en conservant les notations de la section 5.4, on observe que, pour toute droite vectorielle  $D$  de  $\mathbb{K}^n$ , et en considérant l'orthogonalité pour la forme bilinéaire  $(X, Y) \mapsto X^T Y$  sur  $\mathbb{K}^n$ , les sous-espaces  $f^{-1}(\mathcal{L}_D)$  et  $f^{-1}(\mathcal{L}^{D^\perp})$  sont formés de matrices toutes singulières, de codimension au plus  $n$  dans  $V$ , donc de codimension au plus  $2n - 2$  dans  $M_n(\mathbb{K})$ . On est donc précisément dans le cas couvert par le théorème d'Atkinson et Lloyd généralisé par nos soins (voir le chapitre 2), l'exclusion du cas  $(n, |\mathbb{K}|, \text{codim } V) = (3, 2, 1)$  étant fort heureuse. Il reste la situation épineuse où  $f^{-1}(\mathcal{L}_D)$  ou  $f^{-1}(\mathcal{L}^{D^\perp})$  pourrait être équivalent à l'un des espaces de compression  $\mathcal{R}(1, n - 2)$  ou  $\mathcal{R}(n - 2, 1)$  (voir à nouveau le chapitre 2). Et la plus grosse partie de la démonstration consiste à démontrer qu'une telle éventualité est impossible, par une analyse plus poussée de la situation, analyse nécessitant l'examen d'autres images réciproques de sous-espaces singuliers de grande dimension : nous ne dirons rien de plus sur ce nœud particulièrement délicat mais incontournable (voir la partie 3 de [3]).

Une fois écartée la situation bloquante, on se retrouve en terrain presque familier, chaque  $f^{-1}(\mathcal{L}_D)$  ou  $f^{-1}(\mathcal{L}^{D^\perp})$  apparaissant comme l'intersection avec  $V$  d'un sous-espace singulier de type noyau ou image (là, on utilise de manière critique le fait que  $f$  conserve la singularité aussi bien que l'inversibilité). Il est alors relativement aisé, par des calculs de dimensions d'intersection, de montrer la *cohérence* de ces types, à savoir : ou bien  $f^{-1}(\mathcal{L}^{D^\perp})$  est systématiquement de type image lorsque  $D$  parcourt l'espace projectif  $\mathbb{P}(\mathbb{K}^n)$ , et  $f^{-1}(\mathcal{L}_D)$  est systématiquement de type noyau lorsque  $D$  parcourt  $\mathbb{P}(\mathbb{K}^n)$ . Comme on ne perd aucune généralité à post-composer  $f$  par la transposition, on peut se réduire au premier cas. Et alors, en appliquant les résultats obtenus à  $f^{-1} : f(V) \hookrightarrow M_n(\mathbb{K})$ , on obtient une bijection  $\varphi : \mathbb{P}(\mathbb{K}^n) \xrightarrow{\sim} \mathbb{P}(\mathbb{K}^n)$  telle que  $f(V \cap \mathcal{L}^{D^\perp}) = f(V) \cap \mathcal{L}^{\varphi(D)^\perp}$  pour tout  $D \in \mathbb{P}(\mathbb{K}^n)$ . Ensuite, en utilisant le théorème fondamental de la géométrie projective ainsi que l'effet de  $f$  sur les sous-espaces de la forme  $V \cap \mathcal{L}_D$ , on peut démontrer que  $\varphi$  est une homographie. Composant  $f$  à gauche par un morphisme de Frobenius, on se réduit alors à la situation à  $f$  envoie  $V \cap \mathcal{L}^{D^\perp}$  dans  $\mathcal{L}^{D^\perp}$  pour toute droite vectorielle  $D$  de  $\mathbb{K}^n$ .

Ainsi, après composition à gauche de  $f$  par un morphisme de Frobenius, on s'est ramené à la situation où  $f(M) \subset \text{Im } M$  pour toute matrice  $M \in M_n(\mathbb{K})$ .

On reconnaît alors le problème des applications linéaires Im-compatibles (voir le chapitre 4) en travaillant sur les colonnes de  $f(M)$  ! Comme  $\text{codim } V \leq n - 2$ , le théorème 4.2 montre qu'il existe pour tout  $j \in \llbracket 1, n \rrbracket$  un vecteur  $X_j \in \mathbb{K}^n$  tel que  $C_j(f(M)) = MX_j$  pour tout  $M \in V$ . La matrice  $Q := [X_1 \ \cdots \ X_n]$  vérifie donc  $f(M) = MQ$  pour tout  $M \in M_n(\mathbb{K})$ . Et on trouve alors facilement que  $Q$  est inversible. On a ainsi reconnu un morphisme de Frobenius, ce qui achève la démonstration.

Notre travail [3] ne se limite pas au théorème 5.4. Nous avons aussi effectué une étude quasi-exhaustive du cas particulier  $(n, |\mathbb{K}|, \text{codim } V) = (3, 2, 1)$ , autrement dit celui des hyperplans vectoriels de  $M_3(\mathbb{F}_2)$ . Le cas particulier dans le théorème d'Atkinson-Lloyd est un frein majeur pour appliquer la stratégie précédente, et de fait il existe des contre-exemples au théorème dans ce cas. Pour les obtenir, nous avons procédé à une analyse au cas par cas en isolant les trois classes d'équivalence d'hyperplans de  $M_3(\mathbb{F}_2)$  (déterminées par le rang des matrices non nulles orthogonales à ces hyperplans pour la forme bilinéaire  $(A, B) \mapsto \text{tr}(AB)$ ). D'abord, par un argument de dénombrement de matrices inversibles, on démontre qu'une injection linéaire conservant l'inversibilité au sens fort et définie à partir d'un tel hyperplan doit avoir son image dans la même classe d'équivalence, et on se ramène alors à considérer des *endomorphismes* de représentants particuliers pour chacune de ces classes d'équivalence. Pour les hyperplans  $\{I_3\}^\perp = \mathfrak{sl}_3(\mathbb{F}_2)$  et  $\{I_2 \oplus 0\}^\perp$ , une étude explicite, qui reprend d'ailleurs la stratégie mise en place pour le cas général évoqué plus haut, montre que tout automorphisme linéaire de l'un d'entre eux qui conserve l'inversibilité (que ce soit au sens faible ou fort, c'est équivalent par un argument de finitude) est restriction d'un morphisme de Frobenius. En revanche, pour l'orthogonal de la matrice  $E_{3,3}$  dans  $M_3(\mathbb{F}_2)$ , nous sommes parvenus à construire un automorphisme conservant l'inversibilité et qui ne se prolonge pas en un morphisme de Frobenius. Explicitement, nous avons démontré que

$$\begin{bmatrix} a & c & e \\ b & d & f \\ g & h & 0 \end{bmatrix} \longmapsto \begin{bmatrix} a & c & e \\ b+h+e & d & f \\ g & h & 0 \end{bmatrix}$$

est un tel contre-exemple (on se référera à la page 3523 de [3] pour des explications détaillées à ce sujet).

Nous avons enfin examiné la possibilité d'affaiblir des hypothèses dans le théorème 5.4. Principalement, nous avons examiné s'il était possible de remplacer la conservation de l'inversibilité au sens fort par la simple stabilisation  $f(V \cap \text{GL}_n(\mathbb{K})) \subset \text{GL}_n(\mathbb{K})$ . En nous limitant aux corps infinis, nous avons observé que la méthode utilisée par Lussier et Moulin [Lus06] pouvait être adaptée à notre situation. La démonstration passe notamment par la constatation du fait que la fonction polynomiale  $\det|_V$  est irréductible (et à nouveau la contrainte de codimension  $\text{codim } V \leq n - 2$  est optimale pour avoir ce résultat : le même espace cité pour produire un contre-exemple au début de cette partie fournit une situation où la restriction du déterminant à  $V$  n'est pas irréductible). On renvoie à la proposition 11 de [3], que l'on démontre par récurrence sur  $n$ , suivie de la proposition 12 où l'on démontre que toute fonction polynomiale sur  $V$  qui s'annule en tout point de  $V \setminus \text{GL}_n(\mathbb{K})$  est un multiple du déterminant.

## Perspectives

Le travail de [3] ouvre des perspectives intéressantes. L'une d'entre elles consisterait à examiner ce que devient le théorème 5.4 sur les corps finis en remplaçant la conservation forte de l'inversibilité par une conservation faible. Une autre direction possible, plus immédiatement accessible, consiste à s'intéresser à des espaces de matrices rectangulaires. Typiquement, on voudrait trouver, pour des entiers  $n > p$  fixés, la plus grande dimension  $d(n, p)$  telle que, pour tout sous-espace vectoriel  $V$  de  $M_{n,p}(\mathbb{K})$  de codimension au plus  $d(n, p)$  et toute injection linéaire  $f : V \hookrightarrow M_{n,p}(\mathbb{K})$ , la conservation du rang  $p$  au sens fort par l'application  $f$  la contraint à être la restriction d'un morphisme de Frobenius  $M \mapsto PMQ$  (où  $P \in GL_n(\mathbb{K})$  et  $Q \in GL_p(\mathbb{K})$ ). Et ici la dimension  $d(n, p)$  ne sera pas toujours  $n - 2$ , car il y a une autre contrainte fonction de  $p$  qui est liée à la classification des morphismes Im-compatibles. C'est ce problème qui a motivé une étude plus approfondie des morphismes Im-compatibles, étude qui a ensuite rejailli sur mes travaux sur les sous-espaces de rang majoré. Il ne reste donc plus qu'à enfin rédiger et publier ces travaux ...

## 5.7 Applications linéaires conservant une fonction de Schur [6]

À la suite d'une présentation en 2011 de mes travaux au séminaire de l'UVSQ sur le théorème d'Atkinson et Lloyd et son application à la préservation de l'inversibilité, Ariane Mézard m'a demandé si mes méthodes pouvaient être adaptées à la préservation des *immanents*. Les immanents sont une généralisation, introduite par Littlewood et Richardson, de la notion de déterminant. Il est bien connu que les caractères sur  $\mathfrak{S}_n$  sont à valeurs dans  $\mathbb{Z}$ . Pour un tel caractère irréductible  $\chi$ , on peut considérer l'immanent associé

$$d_\chi : A \mapsto \sum_{\sigma \in \mathfrak{S}_n} \chi(\sigma) \prod_{j=1}^n a_{\sigma(j),j}.$$

On reconnaît comme cas particulier le déterminant (pour le caractère signature) et le permanent (pour le caractère constant). Plus généralement, on associe à toute application  $f : \mathfrak{S}_n \rightarrow \mathbb{K}$  la **fonction de Schur**

$$\tilde{f} : A \in M_n(\mathbb{K}) \mapsto \sum_{\sigma \in \mathfrak{S}_n} f(\sigma) \prod_{j=1}^n a_{\sigma(j),j}.$$

Duffner [Duf94] a étudié les endomorphismes de  $M_n(\mathbb{C})$  qui conservent un immanent donné, autrement dit qui vérifient  $\forall A \in M_n(\mathbb{C}), d_\chi(f(A)) = d_\chi(A)$ , le caractère irréductible  $\chi$  étant fixé. La question qui m'était posée était celle de l'adaptation à un corps quelconque.

Il est bon de revenir brièvement sur le cas des permanents. Dans celui-ci, les préserveurs sont beaucoup plus rares que pour le déterminant, et sont de nature essentiellement combinatoire. Ainsi, pour peu que  $n > 2$  et  $\chi(\mathbb{K}) \neq 2$ , les endomorphismes de  $M_n(\mathbb{K})$  qui conservent le permanent sont les fonctions de l'une des formes

$$A \mapsto K \star (PAQ) \quad \text{ou} \quad A \mapsto K \star (PA^TQ),$$

où  $P$  et  $Q$  sont des matrices *de permutation* de  $M_n(\mathbb{K})$ , la notation  $\star$  recouvre le produit de Hadamard (produit coefficient par coefficient), et enfin  $K$  est une matrice de rang 1 dont le produit des termes diagonaux vaut 1 (ce qui permet d'avoir  $\prod_{j=1}^n K_{\sigma(j),j} = 1$  pour tout  $\sigma \in \mathfrak{S}_n$ ). Ce résultat est dû à Botta [Bot67] pour un corps de cardinal suffisamment grand, et il semble que j'aie été le premier à l'obtenir sans cette hypothèse.

Duffner montra quant à elle qu'à l'exception de certains caractères irréductibles sur  $\mathfrak{S}_3$  et à l'exception des déterminants, tous les immanents ont leurs préserveurs linéaires de forme très voisine de celle obtenue pour le permanent (on renvoie au théorème 2 de [Duf94]).

Pour m'attaquer au problème, j'ai conservé la trame globale de la méthode de Dieudonné et ai essayé de conserver le maximum de généralité en traitant de fonctions de Schur. Le plan est globalement toujours le même :

- Étude des sous-espaces de dimension maximale inclus dans le cône  $\mathcal{C}(f) := \tilde{f}^{-1}\{0\}$  des zéros de la fonction de Schur ;
- Examen de l'action, sur ces sous-espaces, d'un préserveur linéaire de  $\tilde{f}$ .

Une difficulté immédiate à laquelle on est confronté – on le voit bien en comparant les situations du déterminant et du permanent – est que les solutions de la première étape doivent dépendre intimement du choix de la fonction  $f$ . Une autre difficulté à laquelle se heurte quiconque voudrait adapter les méthodes du chapitre 2 est l'impossibilité d'appliquer autre chose que des transformations strictement combinatoires sur les matrices envisagées (permutations de lignes et de colonnes).

Bien sûr, lorsque l'on considère une fonction de Schur et ses préserveurs, on fait en général l'hypothèse que la fonction  $f$  n'est pas identiquement nulle. Cependant, pour adapter mes méthodes, il a été indispensable d'aller plus loin et de supposer que  $f$  ne s'annule pas du tout. Une justification heuristique est la suivante. Partant d'un sous-espace affine de  $M_n(\mathbb{K})$  inclus dans  $\mathcal{C}(f)$  et dans lequel toute matrice est de la forme

$$M = \begin{bmatrix} N & [0]_{(n-1) \times 1} \\ [?]_{1 \times (n-1)} & 1 \end{bmatrix},$$

on peut écrire  $\tilde{f}(M) = \tilde{g}(N)$  pour une certaine fonction  $g : \mathfrak{S}_{n-1} \rightarrow \mathbb{K}$  indépendante de  $M$  ; il est alors crucial de pouvoir garantir que  $g$  n'est pas identiquement nulle, ce qui revient à dire que  $f$  est non nulle en au moins une permutation fixant  $n$ . On voit assez vite qu'il est bien pratique de supposer d'emblée que  $f$  ne s'annule pas (ce qui ne sera plus vérifié si l'on passe modulo  $p$  un caractère irréductible de  $\mathfrak{S}_n$ , typiquement).

Avec cette hypothèse de non-annulation, on parvient notamment à obtenir le théorème suivant :

**Théorème 5.5** (Théorème 4.2 de [6]). *Soit  $f : \mathfrak{S}_n \rightarrow \mathbb{K}^*$ . Tout sous-espace affine de  $M_n(\mathbb{K})$  inclus dans le cône  $\mathcal{C}(f)$  est alors de codimension au moins  $n$ .*

L'extension aux sous-espaces affines est, comme dans le chapitre 2, très utile pour autoriser des raisonnements par récurrence sur la taille des matrices envisagées.

Ensuite, la structure des sous-espaces vectoriels inclus dans  $\mathcal{C}(f)$  et de codimension minimale a été élucidée en fonction d'objets liés à la structure combinatoire de  $f$ , objets qu'il nous faut maintenant évoquer. D'abord, on peut faire subir à  $f$  diverses transformations sans altérer profondément la structure de la fonction de Schur associée. On a ainsi une notion de H-équivalence (qui revient à transformer  $\tilde{f}$  en  $A \mapsto \tilde{f}(K \star A)$  avec  $K$  à coefficients tous non nuls) puis de PH-équivalence, dans laquelle on s'autorise en plus des permutations de lignes et de colonnes sur  $A$ . Ensuite, à  $f$  on associe deux relations sur  $\llbracket 1, n \rrbracket$ , appelées  $f$ -équivalence en ligne et  $f$ -équivalence en colonne. Deux indices distincts  $i$  et  $j$  de  $\llbracket 1, n \rrbracket$  sont  **$f$ -équivalents en colonne** lorsqu'il existe un vecteur  $Z \in \mathbb{K}^n$  à coefficients *tous non nuls* tel que les conditions équivalentes suivantes soient réalisées :

- (i) la fonction  $\tilde{f}$  s'annule en toute matrice  $M$  telle que  $C_j(M) = Z \star C_i(M)$  ;
- (ii) on dispose de l'identité  $f(\sigma\tau_{i,j}) = -\frac{z_{\sigma(j)}}{z_{\sigma(i)}} f(\sigma)$  pour tout  $\sigma \in \mathfrak{S}_n$ .

En imposant la réflexivité, on démontre qu'il s'agit là d'une relation d'équivalence. On dit en outre que  $f$  est normalisée en colonne lorsqu'on peut systématiquement prendre  $Z = (1)_{1 \leq i \leq n}$  dans la condition précédente, autrement dit lorsque  $\tilde{f}$  s'annule systématiquement en une matrice ayant deux colonnes colinéaires en des positions  $f$ -équivalentes en colonne. On adapte la définition pour définir la  $f$ -équivalence en ligne, et on dit que  $f$  est normalisée lorsqu'elle est normalisée en ligne et en colonne, et totalement normalisée lorsqu'en outre les classes de  $f$ -équivalence (en ligne ou en colonne) sont des intervalles d'entiers, par ordre décroissant de cardinal.

Aux deux extrémités du spectre, on a le cas des permanents lorsque  $\chi(\mathbb{K}) \neq 2$  et  $n \geq 3$ , où la  $f$ -équivalence en ligne est triviale, tout comme la  $f$ -équivalence en colonne (les classes d'équivalence sont des singletons), et le cas des déterminants, où la  $f$ -équivalence en ligne est grossière tout comme la  $f$ -équivalence en colonne. Plusieurs exemples traités dans [6] montrent que d'autres situations sont possibles, mais qu'il y a des restrictions sur les relations d'équivalence sur  $\llbracket 1, n \rrbracket$  ainsi obtenues (par exemple, il n'est pas possible qu'il y ait exactement deux classes d'équivalence dont une est un singleton), et il y a des relations entre la  $f$ -équivalence en ligne et la  $f$ -équivalence en colonne (par exemple, l'une ne peut pas être grossière sans que l'autre le soit aussi). Ces questions combinatoires méritent très certainement un approfondissement.

Une fois ces bases posées, il est possible de démontrer que toute fonction  $f$  est PH-équivalente à une fonction totalement normalisée. Dans le cas où  $f$  est totalement normalisée, on peut alors facilement décrire les sous-espaces vectoriels de codimension minimale inclus dans le cône de nullité, ce qui généralise la situation rencontrée pour les permanents et les déterminants :

**Théorème 5.6.** *Soit  $f : \mathfrak{S}_n \rightarrow \mathbb{K}^*$  totalement normalisée. Les sous-espaces vectoriels de codimension  $n$  inclus dans  $\mathcal{C}(f)$  sont alors ceux d'un des deux types suivants :*

- ceux constitués des matrices  $M \in M_n(\mathbb{K})$  telles que  $MX = 0$ , pour un vecteur non nul  $X$  de  $\mathbb{K}^n$  dont les indices correspondant aux termes non nuls sont deux à deux  $f$ -équivalents en colonne ;
- ceux constitués des matrices  $M \in M_n(\mathbb{K})$  telles que  $M^T X = 0$ , pour un vecteur non nul  $X$  de  $\mathbb{K}^n$  dont les indices correspondant aux termes non nuls sont deux à deux  $f$ -équivalents en ligne.

À partir de là, une adaptation de la méthode de Dieudonné permet d'étudier non seulement les préserveurs linéaires de  $\tilde{f}$ , mais plus généralement encore les *convertisseurs* entre deux fonctions de Schur *a priori* différentes :

**Théorème 5.7.** *Soit  $f : \mathfrak{S}_n \rightarrow \mathbb{K}^*$  et  $g : \mathfrak{S}_n \rightarrow \mathbb{K}^*$ . Soit  $u : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$  linéaire telle que  $\forall M \in M_n(\mathbb{K}), \tilde{g}(u(M)) = 0 \Leftrightarrow \tilde{f}(M) = 0$ . Il existe alors un scalaire  $\lambda \in \mathbb{K}^*$  tel que :*

- (i)  $\forall M \in M_n(\mathbb{K}), \tilde{g}(u(M)) = \lambda \tilde{f}(M)$ .
- (ii) *La fonction  $f$  est PH-équivalente à  $g$  ou à sa transposée  $g^T : \sigma \mapsto g(\sigma^{-1})$ .*

En particulier, lorsque  $f$  et  $g$  sont totalement normalisées, un convertisseur linéaire de  $f$  vers  $g$  n'existe que si l'une des deux conditions suivantes est vérifiée :

- la  $f$ -équivalence en ligne est la  $g$ -équivalence en ligne, et la  $f$ -équivalence en colonne est la  $g$ -équivalence en colonne ;
- la  $f$ -équivalence en colonne est la  $g$ -équivalence en ligne, et la  $f$ -équivalence en ligne est la  $g$ -équivalence en colonne.

Armé de ce résultat très général, il est facile de contredire l'existence d'un convertisseur du permanent vers le déterminant ou du déterminant vers le permanent, tant que  $n \geq 3$  et  $\chi(\mathbb{K}) \neq 2$  et sans contrainte de cardinalité sur  $\mathbb{K}$ , ce qui généralise des résultats antérieurement connus [Duf13].

Il reste à expliciter les convertisseurs linéaires d'une fonction de Schur  $\tilde{f}$  associée à une fonction totalement normalisée  $f : \mathfrak{S}_n \rightarrow \mathbb{K}^*$  vers une autre  $\tilde{g}$ . Là, il n'est pas possible de donner une description explicite sans rentrer plus avant encore dans les entrailles de  $f$  et  $g$ . Je me suis donc limité à une description partielle de ces convertisseurs, que je vais maintenant expliciter. Nous dirons qu'une permutation  $\sigma$  est adaptée à  $f$  en colonne (respectivement, en ligne) lorsqu'elle conserve la  $f$ -équivalence en colonne (respectivement, en ligne). Une matrice  $K \in M_n(\mathbb{K})$  est dite  $f$ -adaptée lorsqu'elle a les mêmes coefficients aux positions  $(i, j)$  et  $(i', j')$  dès que  $i$  et  $i'$  sont  $f$ -équivalents en ligne, et  $j$  et  $j'$  sont  $f$ -équivalents en colonne. Enfin, on appelle  $f$ -similitude standard toute fonction de la forme  $M \mapsto PMQ$ , où  $P, Q$  sont des matrices diagonales par blocs inversibles, où les tailles des blocs diagonaux de  $P$  (respectivement, de  $Q$ ) sont les tailles des classes de  $f$ -équivalence en ligne (respectivement, en colonne) dans le même ordre.

On démontre alors le résultat suivant (qui est en fait à la source de l'énoncé (b) du théorème précédent) :

**Théorème 5.8** (Théorème 5.2 de [6]). *Soit  $n \geq 2$ ,  $f : \mathfrak{S}_n \rightarrow \mathbb{K}^*$  et  $g : \mathfrak{S}_n \rightarrow \mathbb{K}^*$  deux fonctions totalement normalisées, et  $u : M_n(\mathbb{K}) \rightarrow M_n(\mathbb{K})$  une application linéaire telle que  $\forall M \in M_n(\mathbb{K}), \tilde{g}(u(M)) = \tilde{f}(M)$ . L'une des situations suivantes a alors lieu :*

- (a) *Il existe un quadruplet  $(K, \sigma, \tau, V)$  dans lequel  $K \in M_n(\mathbb{K})$  est une matrice  $g$ -adaptée à coefficients tous non nuls,  $\sigma$  et  $\tau$  sont des éléments de  $\mathfrak{S}_n$  respectivement  $f$ -adaptés en ligne et en colonne,  $V$  est une  $f$ -similitude standard et*

$$u : M \mapsto K \star (P_\sigma V(M) P_\tau)$$

- (b) *Il existe un quadruplet  $(K, \sigma, \tau, V)$  dans lequel  $K \in M_n(\mathbb{K})$  est une matrice  $g$ -adaptée à coefficients tous non nuls,  $\sigma$  et  $\tau$  sont des éléments de  $\mathfrak{S}_n$*

respectivement  $f$ -adaptés en colonne et en ligne,  $V$  est une  $f^T$ -similitude standard et

$$u : M \mapsto K \star (P_\sigma V(M^T)P_\tau).$$

On a ici noté  $P_\sigma = (1_{i=\sigma(j)})_{1 \leq i, j \leq n}$  la matrice associée à la permutation  $\sigma \in \mathfrak{S}_n$ .

Ce théorème semble clore complètement la question, mais tel n'est pas le cas. Par le point (b) du théorème 5.7, la structure des convertisseurs se ramène à celle des préserveurs linéaires du cône de nullité de  $\tilde{f}$ . Mais dans ce cas, on ne sait par exemple pas pour quels quadruplets  $(K, \sigma, \tau, V)$  la fonction  $M \mapsto K \star (P_\sigma V(M)P_\tau)$  est réellement un préserveur de  $\tilde{f}$ . Dans le cas du déterminant ou du permanent, on retrouve certes facilement les résultats classiques, mais sans hypothèse supplémentaire sur  $f$  il n'est pas possible d'aller véritablement plus loin.

La fin de l'étude de [6] a été inspirée par la question initiale des immanents. Certes, ils ne vérifient pas tous les conditions de non-annulation, mais ils ont en retour la propriété remarquable d'être centraux. J'ai donc étudié de manière tout à fait systématique les préserveurs pour les fonctions de Schur associées aux fonctions centrales  $f$  qui ne s'annulent pas. Dans ce cas, une observation fondamentale est que les relations de  $f$ -équivalence en ligne ou en colonne sont grossières ou triviales. Le cas grossier se ramène à l'étude des déterminants. Dans le cas contraire, les deux relations sont triviales (on parle de fonction *rigide*), comme dans le cas des permanents sous les hypothèses habituelles. Il reste alors pour conclure à étudier les permutations  $\tau$  de  $\llbracket 1, n \rrbracket$  pour lesquelles il existe  $A \in M_n(\mathbb{K}^*)$  vérifiant

$$\forall M \in M_n(\mathbb{K}), \tilde{f}(MP_\tau) = \tilde{f}(A \star M).$$

Le résultat-clef est que l'ensemble  $G_f$  de ces permutations est un sous-groupe distingué de  $\mathfrak{S}_n$  ! En utilisant les résultats classiques sur la classification des sous-groupes distingués de  $\mathfrak{S}_n$ , je suis alors parvenu à complètement expliciter les fonctions  $f$  pour lesquels  $G_f$  inclut un sous-groupe distingué non trivial donné de  $\mathfrak{S}_n$  (par exemple,  $\mathfrak{A}_n$ , ou encore le groupe de Klein lorsque  $n = 4$ ).

En définitive, bien que n'ayant pas complètement résolu la question que je m'étais posée initialement, mon étude m'a conduit dans des directions inattendues, avec de nombreuses questions encore en suspens (notamment sur la structure possible pour les relations de  $f$ -équivalence en ligne ou en colonne).

## 5.8 Sur le théorème fondamental de la géométrie des matrices rectangulaires [4]

L'essentiel des idées de ce projet est venu de Peter Šemrl. Je ne suis intervenu dessus qu'alors qu'il était très avancé, et ma contribution a été de trois ordres : d'abord, retravailler la démonstration d'un résultat critique afin qu'il soit valable pour tous les corps finis (la démonstration de Peter nécessitait un corps ayant au moins quatre éléments), ensuite reformuler une bonne partie des arguments de manière géométrique afin de fluidifier considérablement la démonstration, et enfin donner un résultat beaucoup plus satisfaisant pour les corps finis (la dernière partie de l'article).

Cet article est le point quasi final d'un plan de recherche lancé plusieurs années plus tôt par Peter Šemrl pour étudier des versions des théorèmes de Hua sur la géométrie des espaces de matrices avec des hypothèses affaiblies. On rappelle que deux matrices rectangulaires de même format sont adjacentes lorsque leur différence est de rang 1. Dans tout ce qui va suivre,  $\mathbb{K}$  désigne un corps gauche (exceptionnellement, nous ne faisons donc aucune hypothèse de commutativité).

La version forte du théorème de Hua décrit les permutations  $\phi$  de  $M_{n,p}(\mathbb{K})$  telles que l'adjacence de  $A$  et  $B$  soit équivalente à celle de  $\phi(A)$  et  $\phi(B)$  pour toutes matrices  $A$  et  $B$  de  $M_{n,p}(\mathbb{K})$ . Šemrl a cherché à affaiblir les hypothèses de trois façons différentes :

- retrait de la bijectivité ;
- préservation de l'adjacence dans le sens direct uniquement (si  $A$  et  $B$  sont adjacentes alors  $\phi(A)$  et  $\phi(B)$  sont adjacentes) ;
- application entre espaces de matrices rectangulaires de formats différents (on étudie des fonctions  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$ ).

Dans son œuvre la plus aboutie sur la question [Sem14], Šemrl a montré que, même pour des matrices carrées, sur un corps général il est possible de construire des solutions très dégénérées lorsque l'on retire l'hypothèse de bijectivité et de conservation de l'adjacence dans les deux sens. Ces solutions sont liées à l'existence d'endomorphismes non surjectifs du corps de base. Afin d'obtenir des résultats significatifs sur un corps gauche quelconque pour les matrices carrées, Šemrl est obligé de rajouter une hypothèse de conservation du diamètre, à savoir qu'il existe deux matrices  $A$  et  $B$  telles que  $A - B$  et  $\phi(A) - \phi(B)$  soient inversibles (en général, la simple conservation de l'adjacence au sens faible assure que  $\text{rg}(\phi(A) - \phi(B)) \leq \text{rg}(A - B)$  quelles que soient les matrices  $A$  et  $B$ ). Et même sous ces hypothèses apparaissent certaines solutions dégénérées difficiles à décrire explicitement, ainsi que des solutions de la forme suivante :

$$A \mapsto T(I_n + A^\tau L)^{-1} A^\tau S$$

où  $T$  et  $S$  sont des matrices de  $\text{GL}_n(\mathbb{K})$ ,  $\tau$  un endomorphisme du corps gauche  $\mathbb{K}$ , et  $L \in M_n(\mathbb{K})$  une matrice telle que  $I_n + A^\tau L$  soit inversible pour tout  $A \in M_n(\mathbb{K})$  (une telle matrice  $L$  n'existe pas si  $\tau$  est bijectif). Il y a aussi des solutions de type similaire définies à partir d'antiendomorphismes de  $\mathbb{K}$ , avec une transposition en sus.

Šemrl a remarqué que ses méthodes fonctionnaient beaucoup mieux lorsque le corps (gauche) de base  $\mathbb{K}$  était EAS<sup>3</sup>, autrement dit que tout endomorphisme de  $\mathbb{K}$  soit bijectif. Parmi les corps gauches EAS importants, il y a bien sûr  $\mathbb{Q}$  et plus généralement n'importe quel corps de nombres, mais aussi les corps finis,  $\mathbb{R}$  et  $\mathbb{H}$ . En revanche, il est classique que  $\mathbb{C}$  n'est pas EAS (ce qu'on peut montrer grâce à des extensions transcendantes).

Avant de poursuivre, indiquons que ce que l'on appelle un préserveur dégénéré. Soit  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  conservant l'adjacence (au sens faible). On munit les ensembles de départ et d'arrivée de la distance arithmétique ; l'application  $\phi$  est 1-lipschitzienne. On appelle **ensemble adjacent** toute partie de  $M_{m,n}(\mathbb{K})$  dont les éléments sont deux à deux adjacents. Pour des droites vectorielles respectives  $d$  et  $e$  de  $\mathbb{K}^n$  et  $\mathbb{K}^m$ , on note  $R_d$  l'ensemble des matrices de  $M_{m,n}(\mathbb{K})$  dont le noyau inclut l'orthogonal de  $d$ , et  $L_e$  l'ensemble des matrices de  $M_{m,n}(\mathbb{K})$

---

3. Endomorphisms Are Surjective.



dont l'image est incluse dans  $e$ . Tout ensemble adjacent maximal est alors soit de la forme  $A + R_d$  pour une matrice  $A \in M_{m,n}(\mathbb{K})$  et une droite vectorielle  $d$  de  $\mathbb{K}^n$ , soit de la forme  $A + L_e$  pour une matrice  $A \in M_{m,n}(\mathbb{K})$  et une droite vectorielle  $e$  de  $\mathbb{K}^m$ . On dit enfin que  $\phi$  est dégénérée en  $A$  lorsque  $\phi(\mathcal{B}(A, 1))$  est inclus dans un ensemble de la forme  $A + (R_d \cup L_e)$ . Et on dit que  $\phi$  est **dégénérée** lorsqu'elle est dégénérée en tout point.

Sur un corps gauche infini, on peut toujours construire des préserveurs d'adjacence dégénérés. L'exemple, imaginé par Peter, est le suivant : on suppose  $m \geq n \geq 4$  et  $n$  pair, et on se donne une suite  $(\varphi_k)_k$  d'injections vers  $\mathbb{K}$ , toutes d'images disjointes et définies sur  $M_{m,n}(\mathbb{K})$  (il est facile d'obtenir l'existence d'une telle suite par considérations sur les cardinaux infinis). Ensuite, on envoie systématiquement  $A \in M_{m,n}(\mathbb{K})$ , de rang  $k$  :

- sur la matrice nulle si  $k = 0$  ;
- sur la matrice  $\sum_{i=1}^k E_{i,i} + \varphi_k(A)E_{k,k+1}$  si  $k$  est pair et  $0 < k < n$  ;
- sur la matrice  $\sum_{i=1}^k E_{i,i} + \varphi_k(A)E_{k+1,k}$  si  $k$  est impair ;
- sur la matrice  $\sum_{i=1}^k E_{i,i} + \varphi_n(A)E_{n-1,n}$  si  $k = n$ .

La construction est faite de telle sorte que  $\phi(A)$  et  $\phi(B)$  soient adjacentes dès que  $|\text{rg}(A) - \text{rg}(B)| \leq 1$ . Comme cette dernière condition est valable dès que  $A$  et  $B$  sont adjacentes,  $\phi$  conserve bien l'adjacence (au sens faible). Et  $\phi$  est même injective !

En revanche, dans la dernière partie de [4], dont je suis entièrement responsable, il est établi que les préserveurs d'adjacence dégénérés sont extrêmement simples sur les corps finis :

**Théorème 5.9** (Théorème 1.2 de [4]). *Soit  $\mathbb{K}$  un corps fini,  $m, n, p, q$  des entiers naturels non nuls, et  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  une fonction dégénérée conservant l'adjacence. Alors l'image de  $\phi$  est un ensemble adjacent.*

La démonstration de ce résultat, totalement autonome du reste de [4], est fondée sur des techniques de dénombrement : par exemple, on dénombre les matrices simultanément adjacentes à deux matrices données suivant leur distance arithmétique.

Nous ne dirons rien de plus sur les préserveurs dégénérés, et allons pour finir décrire la stratégie de démonstration du théorème principal de [4]. Nous dirons qu'une application de  $M_{m,n}(\mathbb{K})$  dans  $M_{p,q}(\mathbb{K})$  est **standard** lorsqu'elle est de l'une des deux formes suivantes :

- ou bien  $M \mapsto A_0 + P \begin{bmatrix} M^\sigma & 0 \\ 0 & 0 \end{bmatrix} Q$ , où  $A_0 \in M_{p,q}(\mathbb{K})$ ,  $P \in \text{GL}_p(\mathbb{K})$ ,  $Q \in \text{GL}_q(\mathbb{K})$ , et  $\sigma$  est un automorphisme de  $\mathbb{K}$  (ce qui nécessite  $m \leq p$  et  $n \leq q$ ) ;
- ou bien  $M \mapsto A_0 + P \begin{bmatrix} (M^T)^\tau & 0 \\ 0 & 0 \end{bmatrix} Q$ , où  $A_0 \in M_{p,q}(\mathbb{K})$ ,  $P \in \text{GL}_p(\mathbb{K})$ ,  $Q \in \text{GL}_q(\mathbb{K})$ , et  $\tau$  est un antiautomorphisme de  $\mathbb{K}$  (ce qui nécessite  $m \leq q$  et  $n \leq p$ ).

**Théorème 5.10.** *Soit  $m, n, p, q$  quatre entiers naturels non nuls, et  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  une fonction conservant l'adjacence (au sens faible). Alors,  $\phi$  est standard ou dégénérée.*

L'idée de base, assez inhabituelle dans ce genre d'étude, consiste à démarrer par le cas  $m = n = 2$ , puis à procéder par recollements/extensions successives. Bien sûr, le théorème fondamental de la géométrie affine intervient de manière critique dans l'étude, comme dans tous les travaux sur le sujet (c'est par le biais de ce théorème qu'on fait apparaître l'automorphisme ou l'antiautomorphisme de  $\mathbb{K}$  qui apparaît dans la forme standard recherchée).

Voici maintenant les principales étapes du raisonnement, déclinées en lemmes :

**Lemme 5.11.** *Soit  $m, n, p, q$  des entiers naturels non nuls tels que  $m \geq 2$  et  $n \geq 2$ . Soit  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  une application conservant l'adjacence, envoyant 0 sur 0, et non dégénérée par rapport à 0. Il existe alors des matrices inversibles  $P \in GL_m(\mathbb{K})$  et  $Q \in GL_n(\mathbb{K})$  telles que la fonction*

$$A \in M_2(\mathbb{K}) \mapsto \phi \left( P \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} Q \right)$$

*soit non dégénérée par rapport à 0 et envoie au moins une matrice de rang 2 sur une matrice de rang 2.*

On note que ce lemme est valable sans supposer que  $\mathbb{K}$  est EAS. Il pourrait donc être utile pour des études ultérieures éventuelles. Le lemme critique, dont la démonstration est dense et délicate, concerne la situation obtenue à l'issue du premier lemme. C'est là, et uniquement là, que l'on fait usage du théorème fondamental de la géométrie affine (signalons que ce théorème exclut le cas des corps à deux éléments, mais que ce dernier se règle très facilement lorsqu'une difficulté apparaît).

**Lemme 5.12.** *Soit  $p, q$  des entiers supérieurs à 2, et  $\mathbb{K}$  un corps gauche EAS. Soit  $\phi : M_2(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  conservant l'adjacence et telle que  $\phi(0) = 0$ . On suppose enfin que  $\phi$  n'envoie pas l'ensemble des matrices de rang 1 dans un ensemble adjacent et que  $\phi$  envoie au moins une matrice de rang 2 sur une matrice de rang 2. Alors  $\phi$  est standard.*

Le dernier grand résultat intermédiaire, qui ne nécessite d'ailleurs pas l'hypothèse que le corps gauche de base soit EAS, est un résultat d'extension à partir de la situation locale :

**Lemme 5.13.** *Soit  $\mathbb{K}$  un corps gauche, ainsi que des entiers  $m, n, p, q$  tous strictement supérieurs à 1. Soit  $\phi : M_{m,n}(\mathbb{K}) \rightarrow M_{p,q}(\mathbb{K})$  conservant l'adjacence. On suppose que*

$$\phi \left( \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \right) = \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \quad \text{pour tout } A \in M_2(\mathbb{K}).$$

*Alors  $m \leq p$ ,  $n \leq q$  et  $\phi$  est standard.*

La démonstration consiste d'abord en une extension de la forme de  $\phi$  aux matrices ayant toutes leurs colonnes (ou lignes) nulles à partir de la deuxième, puis une extension aux matrices de rang supérieur.

À partir de là, la démonstration du théorème 5.10 est très facile. En supposant  $\phi$  non dégénérée, on se ramène après translation au départ et à l'arrivée au cas où  $\phi$  envoie 0 sur 0 et n'est pas dégénérée par rapport à 0. Et il suffit essentiellement d'enchaîner les résultats des trois lemmes précédents pour conclure que  $\phi$  est standard.

## Perspectives

De l'avis de Peter Šemrl, les résultats de [4] combinés à ceux de [Sem14] sont probablement ce qu'on peut envisager de plus définitif dans l'étude du théorème de Hua avec hypothèses affaiblies. Un affaiblissement supplémentaire pourrait être de remplacer la conservation de l'adjacence par celle de la cohérence (deux matrices sont cohérentes lorsque leur distance est inférieure à 1), mais il est très difficile de tirer quoi que ce soit de cette simple conservation. Une observation des arguments de [4] indique qu'il est très important de savoir qu'un préserveur d'adjacence est injectif sur tout ensemble adjacent maximal, et la perte de cette injectivité est extrêmement handicapante. Rien de substantiel n'a été trouvé dans cette direction.

Les nombreux exemples de préserveurs dégénérés indiquent aussi qu'il n'y a probablement pas grand-chose non plus à attendre dans la direction d'une description explicite de ceux-ci : le meilleur résultat concerne les corps finis et c'est nous qui l'avons obtenu.

En revanche, une direction d'étude potentielle serait la transposition de ce type de recherche aux autres géométries, notamment celle des matrices alternées, qui est la plus délicate de toutes et où il reste des résultats à établir. Nous avons choisi d'en rester là provisoirement pour le moment, mais envisageons de nous remettre à la tâche un jour.

## 5.9 Transformations continues de l'espace-temps [5]

Il s'agit de ma seconde collaboration entre Peter Šemrl. Et contrairement à la première, c'est moi qui ai essentiellement trouvé tous les résultats et méthodes de cette article, Peter ayant surtout contribué à poser le problème et à me souffler d'ajouter une hypothèse de *continuité*, inhabituelle dans ce genre d'étude.

Le cadre est celui des espaces quadratiques réels. On prend un tel espace  $(E, q)$ , de dimension  $n \geq 4$  et de signature  $(1, n - 1)$ . Deux points  $a$  et  $b$  de  $E$  sont dits **cohérents** lorsque  $q(a - b) = 0$ , et ils sont dits **adjacents** lorsqu'ils sont cohérents et distincts. L'ensemble des points de  $E$  cohérents avec un point  $a$  donné est  $a + \mathcal{C}$ , où  $\mathcal{C}$  est le cône isotrope de  $(E, q)$ . Le théorème d'Alexandrov décrit les permutations de  $E$  qui conservent l'adjacence :

**Théorème 5.14** (Alexandrov). *Soit  $(E, q)$  un espace quadratique réel de dimension  $n \geq 4$  et de signature  $(1, n - 1)$ . Les permutations de  $E$  qui conservent l'adjacence sont alors les similitudes affines de  $(E, q)$ , autrement dit les fonctions de la forme  $x \mapsto a + \varphi(x - a)$ , où  $\varphi$  est une similitude vectorielle de  $(E, q)$ .*

L'étude que nous avons faite était motivée par la remarque suivante : un cas particulier du théorème d'Alexandrov est celui de l'espace  $E = \mathcal{H}_2(\mathbb{C})$  des matrices hermitiennes  $2 \times 2$  muni de la forme quadratique déterminant. Ce cas précis est aussi un cas particulier des résultats de Hua sur la géométrie des espaces de matrices hermitiennes. Et pour cette dernière Huang et Šemrl [Hua08] ont caractérisé les préserveurs de l'adjacence avec les hypothèses minimales. Les fonctions  $f : \mathcal{H}_n(\mathbb{C}) \rightarrow \mathcal{H}_n(\mathbb{C})$  préservant l'adjacence sont :

- les fonctions de la forme  $M \mapsto A_0 + \lambda PMP^*$  où  $A_0 \in \mathcal{H}_n(\mathbb{C})$ ,  $\lambda \in \mathbb{R}^*$  et  $P \in \text{GL}_n(\mathbb{C})$  ;
- les fonctions de la forme  $M \mapsto A_0 + \lambda PM^T P^*$  où  $A_0 \in \mathcal{H}_n(\mathbb{C})$ ,  $\lambda \in \mathbb{R}^*$  et  $P \in \text{GL}_n(\mathbb{C})$  ;
- les fonctions ayant leur image incluse dans une droite affine (nécessairement dirigée par une matrice de rang 1)

Peter voulait savoir, dans cette étude précise, ce qu'on pouvait obtenir si la préservation de l'adjacence était encore affaiblie en la seule préservation de la cohérence. Il fallait donc au moins commencer par étudier le premier cas non trivial, à savoir celui où  $n = 2$ , ce qui nous ramenait à la situation étudiée par Alexandrov.

Nous ne sommes parvenus à rien de bien intéressant avec des hypothèses aussi faibles, et c'est Peter qui a eu l'idée de rajouter une hypothèse de continuité, somme toute assez naturelle même si en général les études ne formulent que des hypothèses géométriques.

Nous avons alors réussi à donner une forme à peu près explicite pour les fonctions continues  $\phi : E \rightarrow E$  qui conservent la cohérence. On pourrait s'attendre à ce que la situation dégénérée se réduise à celle où  $\phi$  a son image incluse dans une **droite cohérente**, c'est-à-dire une droite affine dirigée par un vecteur  $q$ -isotrope. Cependant, je suis parvenu à construire des exemples plus complexes, où certes l'image de  $\phi$  est incluse dans un translaté du cône isotrope de  $q$ , mais où cette image n'est pas incluse dans une droite. On peut fabriquer un tel exemple comme suit : admettons qu'il soit possible de construire :

- une famille  $(x_k)_{k \in I}$  de points de  $E$  indexée par un ensemble fini ou dénombrable  $I$  ;
- une famille  $(B_k)_{k \in I}$  telle que  $B_k$  soit un voisinage ouvert de  $x_k$  pour tout  $k \in I$ , et pour tous  $k, l$  distincts aucun point de  $B_k$  ne soit cohérent avec un point de  $B_l$  ;
- et une fonction  $\lambda : E \rightarrow \mathbb{R}_+$  continue, dont l'ensemble des zéros est  $E \setminus \bigcup_{k \in I} B_k$ .

Donnons-nous alors une suite  $(z_k)_{k \in I}$  de vecteurs  $q$ -isotropes deux à deux non colinéaires, tous unitaires pour une certaine norme euclidienne sur  $E$ . La fonction

$$\phi : x \in E \mapsto \begin{cases} \lambda(x).z_k & \text{si } x \in B_k \text{ pour un } k \in I \\ 0 & \text{si } x \notin \bigcup_{k \in I} B_k \end{cases}$$

est alors continue et envoie systématiquement deux points cohérents sur deux points cohérents (si l'un des deux points est envoyé sur 0, c'est évident, sinon ils doivent appartenir au même  $B_k$  et donc être envoyés dans la droite  $\mathbb{R}z_k$ ).

Convenons d'appeler **dégénérée** toute fonction  $\phi : E \rightarrow E$  continue, préservant la cohérence et dont l'image est incluse dans un translaté du cône isotrope de  $q$ . On peut alors démontrer qu'une telle fonction est nécessairement produite par le procédé que nous venons de décrire. Enfin il est effectivement possible de prendre des familles  $(x_k)_{k \in I}$  et  $(B_k)_{k \in I}$  comme annoncé avec  $I = \mathbb{N}$ . Par exemple, on se place dans  $E = \mathbb{R}^n$  muni de la forme quadratique  $q : (t_1, \dots, t_n) \mapsto t_1^2 - \sum_{k=2}^n t_k^2$ , et on prend  $x_k := (k, 0, \dots, 0)$  et  $B_k$  la boule ouverte (pour la norme euclidienne standard) de centre  $x_k$  et de rayon  $\varepsilon \in ]0, \frac{1}{2}[$  indépendant de  $k$  et choisi assez petit pour que la contrainte de non-cohérence d'une boule à l'autre soit vérifiée.

Le reste des données est facilement choisi.

En dehors d'une description constructive des préserveurs continus dégénérés de la cohérence, le principal résultat de [5] est le théorème suivant :

**Théorème 5.15.** *Soit  $(E, q)$  un espace quadratique réel de dimension  $n \geq 4$  et de signature  $(1, n - 1)$ . Soit  $\phi : E \rightarrow E$  continue et préservant la cohérence. Alors  $\phi$  est dégénérée ou elle conserve l'adjacence.*

Ensuite, comme on sait que les préserveurs non dégénérés de l'adjacence sont les similitudes affines, le problème est entièrement résolu.

Voyons maintenant les grandes étapes de la démonstration du théorème 5.15. L'idée, naturelle, est d'examiner l'action de  $\phi$  sur les droites cohérentes et en particulier celles incluses dans un même translaté du cône isotrope.

On suppose  $\phi$  non dégénérée. On note  $\mathcal{Q}$  la quadrique projective associée au cône isotrope de  $q$ . D'abord, on montre facilement que l'image par  $\phi$  d'une droite cohérente (i.e. dirigée par un vecteur  $q$ -isotrope) ne peut être un singleton (sinon  $\phi$  aurait son image dans un cône ayant pour sommet le point de ce singleton). Par suite, pour tout  $a \in E$ , on obtient une fonction continue  $\varphi_a : \mathcal{Q} \rightarrow \mathcal{Q}$  telle que, pour tout  $D \in \mathcal{Q}$ , l'élément  $\varphi_a(D)$  soit la direction de l'unique droite cohérente incluant  $\phi(a + D)$ . En outre, la fonction  $a \in E \mapsto \varphi_a \in \mathcal{C}(\mathcal{Q}, \mathcal{Q})$  est continue, l'espace d'arrivée étant muni de la topologie compact-ouvert.

Plaçons-nous pour simplifier dans le cas où  $\phi(0) = 0$  et  $\varphi_0$  n'est pas constante. On examine alors la situation où il existe un  $a \in \mathcal{C} \setminus \{0\}$  tel que  $\phi(a) = 0$ . L'observation critique est que la valeur 0 est prise par  $\phi$  dans tout un ouvert non vide. Pour cela, l'idée est d'introduire deux vecteurs  $b, c$  de  $\mathcal{C}$  tels que  $a, b, c$  soient deux à deux non colinéaires, et  $\phi(b)$  et  $\phi(c)$  soient des éléments non cohérents de  $\mathcal{C} \setminus \{0\}$ . Puis, grâce à un lemme élémentaire sur les formes quadratiques, on trouve un  $d \in \mathcal{C} \setminus \{0\}$  qui soit simultanément cohérent avec  $a, b$  et  $c$ . On montre alors facilement que  $\phi(d) = 0$ , puis de proche en proche on parvient à montrer que  $\phi$  est nulle sur tout un voisinage de  $d$  (en examinant d'abord la situation d'un point cohérent avec  $d$  et proche de  $d$ , puis en reperturbant ce point).

Dans la situation considérée, le point 0 est non générique selon la définition qui suit : un point  $m$  de  $E$  est dit **générique** lorsque  $\phi^{-1}(\phi(m))$  est d'intérieur vide dans  $E$ . Par séparabilité, les valeurs de  $\phi$  aux points non génériques forment un ensemble au plus dénombrable.

Après translations, on peut conclure au résultat intermédiaire suivant :

**Lemme 5.16.** *Pour tout point générique  $a$  de  $E$ , si  $\varphi_a$  n'est pas constante alors  $\phi(m)$  est adjacent à  $\phi(a)$  pour tout point  $m$  adjacent à  $a$ .*

Cette observation cruciale en poche, l'essentiel de la démonstration consiste d'abord à montrer l'existence d'un point générique  $a$  pour lequel  $\varphi_a$  est non constante. Cela se démontre par l'absurde en trois étapes : supposant que  $\varphi_a$  est constante pour tout point générique  $a$ , on montre qu'elle est constante en tout point  $a$ , puis que  $a \mapsto \varphi_a$  est constante, et finalement que  $\phi$  a son image incluse dans une droite cohérente, contredisant l'hypothèse de non-dégénérescence.

Ensuite, on montre progressivement que dès que  $a$  est générique, la fonction  $\varphi_a$  est injective, puis que c'est un homéomorphisme (par le théorème d'invariance du domaine); ensuite la théorie du degré montre, par continuité de  $a \mapsto \varphi_a$ , qu'aucune des  $\varphi_a$  n'est constante. On montre alors que tout point est générique, et on conclut que  $\phi$  conserve l'adjacence grâce au lemme 5.16.

## Bilan et perspectives

Au regard de ce qui précède, la singularité de [5] au sein de mes travaux réside dans l'usage systématique d'arguments de topologie, et bien peu d'algèbre et géométrie en apparence même si quelques lemmes sur la géométrie des cônes sont d'usage constant.

Le travail précédent mériterait certainement d'être prolongé en direction de formes quadratiques de signature plus générale, mais on s'attend à une difficulté nettement accrue du fait de la forme moins contrainte des sous-ensembles cohérents maximaux. On pourrait aussi orienter les recherches vers la détermination des applications continues de  $\mathcal{H}_n(\mathbb{C})$  dans  $\mathcal{H}_n(\mathbb{C})$  qui conservent la cohérence. Autant de pistes intéressantes mais que j'ai provisoirement laissées en suspens.

## Chapitre 6

# Espaces de matrices à spectre limité

### 6.1 Articles concernés

- [1] C. de Seguins Pazzis, On the matrices of given rank in a large subspace. *Linear Algebra Appl.* **435-1** (2011) 147–151.
- [2] C. de Seguins Pazzis, Large affine spaces of matrices with rank bounded below. *Linear Algebra Appl.* **437-2** (2012) 499–518.
- [3] C. de Seguins Pazzis, Spaces of matrices with a sole eigenvalue. *Linear Multilinear Algebra.* **60-10** (2012) 1165–1190.
- [4] C. de Seguins Pazzis, Large affine spaces of non-singular matrices. *Trans. Amer. Math. Soc.* **365** (2013) 2569–2596.
- [5] C. de Seguins Pazzis, On Gerstenhaber’s theorem for spaces of nilpotent matrices over a skew field. *Linear Algebra Appl.* **438-11** (2013) 4426–4438.
- [6] C. de Seguins Pazzis, Spaces of matrices with few eigenvalues. *Linear Algebra Appl.* **449** (2014) 210–311.
- [7] C. de Seguins Pazzis, From primitive spaces of bounded rank matrices to a generalized Gerstenhaber theorem. *Quart. J. Math.* **65-2** (2014) 319–325.
- [8] C. de Seguins Pazzis, Primitive spaces of matrices with upper rank two over the field with two elements. *Linear Multilinear Algebra* **64** (2016) 1321–1353.
- [9] C. de Seguins Pazzis, The structured Gerstenhaber problem I. *Linear Algebra Appl.* **567** (2019) 263–298.
- [10] C. de Seguins Pazzis, The structured Gerstenhaber problem II. *Linear Algebra Appl.* **569** (2019) 113–145.
- [11] C. de Seguins Pazzis, The structured Gerstenhaber problem III. *Linear Algebra Appl.* **601** (2020) 134–169.

### 6.2 Bibliographie historique

- [Ger58] M. Gerstenhaber, On Nilalgebras and Linear Varieties of Nilpotent Matrices I. *Amer. J. Math.* **80** (1958) 614–622.

- [Ger60] M. Gerstenhaber, On Nilalgebras and Linear Varieties of Nilpotent Matrices II. *Duke Math. J.* **27** (1960) 21–32.
- [Ger59] M. Gerstenhaber, On Nilalgebras and Linear Varieties of Nilpotent Matrices II. *Ann. Math.* **70** (1959) 167–205.
- [Ger62] M. Gerstenhaber, On Nilalgebras and Linear Varieties of Nilpotent Matrices IV. *Amer. J. Math.* **75** (1962) 382–418.
- [Bot83] E.P. Botta, S. Pierce, W. Watkins, Linear transformations that preserve nilpotent matrices. *Pac. J. Math.* **104** (1983) 39–46.
- [Ser85] V.N. Serezhkin, Linear transformations preserving nilpotency (en russe). *Izv. Akad. Nauk BSSR, Ser. Fiz.-Mat. Nauk* **125** (1985) 46–50.
- [Mes89] R. Meshulam, On two extremal matrix problems. *Linear Algebra Appl.* **114-115** (1989) 261–271.
- [Mat91] B. Mathes, M. Omladič, H. Radjavi, Linear spaces of nilpotent matrices. *Linear Algebra Appl.* **149** (1991) 215–225.
- [Bru93] R.A. Brualdi, K.L. Chavey, Linear spaces of Toeplitz and nilpotent matrices. *J. Combin. Theory Ser. A* **63** (1993) 65–78.
- [Oml96] M. Omladič, P. Šemrl, Matrix spaces with a bounded number of eigenvalues. *Linear Algebra Appl.* **249** (1996) 29–46.
- [Mes98] R. Meshulam, N. Radwan, On linear subspaces of nilpotent elements in a Lie algebra. *Linear Algebra Appl.* **279** (1998) 195–199.
- [Loe98] R. Loewy, N. Radwan, On spaces of matrices with a bounded number of eigenvalues. *Electronic J. Linear Algebra* **3** (1998) 142–152.
- [Dra06] J. Draisma, H. Kraft, J. Kuttler, Nilpotent subspaces of maximal dimension in semisimple Lie algebras. *Compos. Math.* **142** (2006) 464–476.
- [Qui11] R. Quinlan, Spaces of matrices without non-zero eigenvalues in their field of definition, and a question of Szechtman. *Linear Algebra Appl.* **434** (2011) 1580–1587.
- [Mac12] G.W. MacDonald, J.A. MacDougall, L.G. Sweet, On the dimension of linear spaces of nilpotent matrices. *Linear Algebra Appl.* **142** (2012) 2010–2030.
- [Buk17] D. Kokol Bukovšek, M. Omladič, Linear spaces of symmetric nilpotent matrices. *Linear Algebra Appl.* **530** (2017) 384–404.

### 6.3 Panorama historique

Alors que les recherches sur les sous-espaces vectoriels de matrices de rang majoré ont été initiées par Dieudonné (1948), c’est Murray Gerstenhaber (1958) qui est l’auteur des premiers résultats substantiels sur les sous-espaces vectoriels de matrices avec condition portant sur le spectre. Dans une fameuse série d’articles [Ger58, Ger60, Ger59, Ger62], Gerstenhaber étudia les sous-espaces vectoriels et les sous-algèbres de matrices nilpotentes. Et dès sa première livraison il énonça et démontra un théorème majeur sur la question, qui est le pendant du théorème de Dieudonné sur les sous-espaces vectoriels de matrices singulières.



**Définition 8.** Un sous-espace vectoriel  $\mathcal{V}$  de  $M_n(\mathbb{K})$  est dit **nilpotent** lorsque tous ses éléments sont nilpotents.

**Théorème 6.1.** Soit  $\mathcal{V}$  un sous-espace vectoriel nilpotent de  $M_n(\mathbb{K})$ . On suppose que  $|\mathbb{K}| \geq n$ . Alors  $\dim \mathcal{V} \leq \frac{n(n-1)}{2}$ , et l'égalité est réalisée si et seulement si  $\mathcal{V}$  est conjugué à l'espace  $T_n^{++}(\mathbb{K})$  des matrices triangulaires supérieures strictes.

La démonstration de Gerstenhaber est considérée par beaucoup comme ardue. L'intérêt de l'étude du cas d'égalité réside dans l'observation qu'un sous-espace nilpotent n'est pas nécessairement trigonalisable, au sens qu'il n'est pas nécessairement conjugué à un sous-espace vectoriel de  $T_n^{++}(\mathbb{K})$ . L'exemple incontournable est celui des matrices de la forme  $\begin{bmatrix} 0 & 0 & x \\ 0 & 0 & y \\ -y & x & 0 \end{bmatrix}$  avec  $(x, y) \in \mathbb{K}^2$ .

Il est bien constitué de matrices nilpotentes, mais aucun vecteur non nul de  $\mathbb{K}^3$  n'est dans le noyau de tous ses éléments.

Gerstenhaber alla plus loin encore dans [Ger62], en majorant de manière optimale la dimension d'un sous-espace vectoriel nilpotent en fonction des réduites de Jordan possibles pour ses éléments (sur ce sujet, voir plus récemment [Mac12]).

La restriction sur le cardinal du corps de base est nécessaire pour faire fonctionner la démonstration de Gerstenhaber : la première chose que fait ce dernier est en effet de réduire le problème à celui des corps algébriquement clos, ce qui lui permet d'utiliser des méthodes fondées sur des calculs de polynômes. Lorsque  $\mathcal{V}$  est un sous-espace vectoriel nilpotent de  $M_n(\mathbb{K})$ , et  $\mathbb{L}$  une extension de  $\mathbb{K}$ , on peut garantir que  $\mathcal{V}_{\mathbb{L}} := \text{Vect}_{\mathbb{L}}(\mathcal{V})$  (sous-espace de  $M_n(\mathbb{L})$ ) est encore nilpotent tant que  $|\mathbb{K}| \geq n$  (considérer le polynôme caractéristique, dont les coefficients sont des fonctions homogènes de degré au plus  $n$  des coefficients de la matrice). En revanche, ce résultat est faux en général : par exemple, le sous-espace vectoriel de  $M_3(\mathbb{F}_2)$  constitué des matrices de la forme  $\begin{bmatrix} 0 & x & 0 \\ 0 & 0 & y \\ x+y & 0 & 0 \end{bmatrix}$  est nilpotent,

en revanche son extension à  $\mathbb{F}_4$  contient des matrices inversibles!

Il fallut attendre plus de vingt ans avant que les travaux de Gerstenhaber connussent un prolongement. En 1983, Botta, Pierce et Watkins [Bot83] appliquèrent le théorème 6.1 pour déterminer les automorphismes de l'espace vectoriel  $\mathfrak{sl}_n(\mathbb{K})$  qui stabilisent le cône nilpotent. Il s'agit, classiquement, des fonctions ayant l'une des formes  $M \mapsto \lambda P M P^{-1}$  ou  $M \mapsto \lambda P M^T P^{-1}$ , où  $\lambda \in \mathbb{K}^*$  et  $P \in \text{GL}_n(\mathbb{K})$ . Et ce théorème était toujours donné sous l'hypothèse  $|\mathbb{K}| \geq n$  de Gerstenhaber.

Ce n'est qu'en 1985 que l'hypothèse de cardinalité du théorème 6.1 fut enfin levée par Serezhkin [Ser85] dans un article publié en russe et dont les méthodes n'ont jamais été développées par la suite. Il est amusant de constater que c'est la même année que Meshulam avait fait tomber les restrictions de cardinalité pour le théorème de Flanders, qui suivait lui-même de peu les résultats de Gerstenhaber!

L'intérêt pour le sujet pris de l'ampleur dans les années 1990. D'abord, Mathes, Omladić et Radjavi [Mat91] donnèrent une démonstration nettement simplifiée de la majoration de la dimension dans le théorème 6.1 sans hypothèse de cardinalité. Leur méthode repose sur des considérations très élémentaires de trace. En prolongeant leurs méthodes, ils parvinrent à démontrer le cas d'égalité

en excluant seulement les corps à deux éléments. Leur démonstration utilise de manière très astucieuse le théorème de Jacobson de réduction simultanée, qui généralise le théorème d'Engel, dans sa version « algèbres de Jordan ». Plus clairement encore, Mathes, Omladič et Radjavi démontrent que si  $\mathcal{V}$  est un sous-espace vectoriel nilpotent de dimension maximale, alors il est stable par élévation au carré, puis par polarisation il est stable par le produit de Jordan  $(A, B) \mapsto AB + BA$ . Et classiquement, tout sous-espace vectoriel nilpotent vérifiant cette propriété est simultanément trigonalisable<sup>1</sup>.

Très peu de temps après, Brualdi et Chavey [Bru93] reprirent les problématiques de majoration de dimension en mettant en évidence l'importance des méthodes combinatoires, mais ils mirent de côté toute tentative de caractérisation des sous-espaces de dimension maximale.

Plus tard, des variantes du théorème de Gerstenhaber commencèrent à être explorées plus profondément. Dans [Oml96], Omladič et Šemrl déterminèrent la dimension maximale pour un sous-espace de matrices de  $M_n(\mathbb{C})$  ayant au plus  $k$  valeurs propres, pour  $k \in \{1, 2\}$ , et ils classifièrent les sous-espaces de dimension maximale, sauf lorsque  $n$  est pair et  $k = 2$ . Dans [Loe98], Loewy et Radwan prolongèrent ces travaux, en traitant le cas  $k = 3$  et en prolongeant même l'étude à n'importe quel corps de caractéristique nulle (à condition de s'intéresser aux valeurs propres dans une clôture algébrique du corps de base). Dans [Mes98], Meshulam et Radwan généralisèrent le théorème de Gerstenhaber en déterminant la dimension maximale d'un sous-espace vectoriel ad-nilpotent d'une algèbre de Lie semi-simple (sur un corps algébriquement clos de caractéristique nulle). En particulier, cela permet de déterminer la dimension maximale pour un sous-espace vectoriel de matrices antisymétriques nilpotentes complexes. Et ils complétèrent ce travail par une détermination de la dimension maximale pour un sous-espace vectoriel de matrices symétriques nilpotentes complexes, par une méthode différente qui s'inspirent fortement de celle de Mathes, Omladič et Radjavi. Leurs travaux et méthodes furent prolongés ultérieurement par Draisma, Kraft et Kuttler [Dra06], qui utilisèrent judicieusement le théorème du point fixe de Borel et parvinrent à classifier, sous certaines réserves sur la caractéristique du corps de base (toujours supposé algébriquement clos), les sous-espaces ad-nilpotents de dimension maximale.

Enfin, quelques années plus tôt, et au détour de ses pérégrinations sur les sous-espaces vectoriels de rang borné, Meshulam [Mes89] avait fait l'observation d'un lien entre les sous-espaces *affines* de matrices de rang au moins  $r$  et une variante du problème de Gerstenhaber. Si en effet on considère par exemple un sous-espace affine  $\mathcal{V}$  de  $M_n(\mathbb{K})$  qui est entièrement constitué de matrices inversibles, alors quitte à remplacer  $\mathcal{V}$  par un sous-espace équivalent on peut supposer  $I_n \in \mathcal{V}$ . Dans ce cas, pour tout élément  $M$  de la direction  $V$  de  $\mathcal{V}$ , la matrice  $tI_n - M = t(I_n - t^{-1}M)$  est inversible, si bien que la seule valeur propre possible pour  $M$  dans le corps  $\mathbb{K}$  est 0 : on dit que  $V$  est à spectre trivial. Lorsque le corps de base est algébriquement clos, les sous-espaces à spectre trivial se confondent avec les sous-espaces nilpotents, ce qui n'est pas vrai sur un corps arbitraire : par exemple,  $A_n(\mathbb{R})$  est à spectre trivial mais son seul élément nilpotent est la matrice nulle ! Meshulam conjectura que la majoration de la dimension dans le théorème 6.1 pourrait s'étendre à n'importe quel sous-espace de spectre trivial, mais il ne parvint pas à le démontrer.

---

1. H. Radjavi, The Engel-Jacobson theorem revisited. J. Algebra **111** (1987) 427–430.

## 6.4 Historique de mes travaux

Mes travaux sur le thème du présent chapitre remontent à l'automne 2006. Je sortais d'une première série d'études sur les sous-espaces vectoriels de rang majoré et les sous-espaces vectoriels de matrices diagonalisables. Pour tuer le temps alors que je surveillais mes élèves lors de leur composition précédant les vacances de la Toussaint, je décidai d'étudier le problème des sous-espaces vectoriels de matrices nilpotentes. Le lendemain, j'avais retrouvé le théorème de Gerstenhaber avec une démonstration du cas d'égalité pour les corps ayant au moins  $n$  éléments (mais sans restriction pour la majoration de la dimension). Ma démonstration pour la majoration recouvrait essentiellement la méthode de Mathes, Omladić et Radjavi (j'avais retrouvé en partie leur méthode de trace, me limitant toutefois à l'observation que  $\text{tr}(AB) = 0$  pour toutes les matrices  $A, B$  dont toutes les combinaisons linéaires sont nilpotentes). En revanche, ma démonstration du cas d'égalité se révéla originale. Elle utilisait en partie les techniques mises en œuvre par Flanders pour l'étude des sous-espaces de rang majoré, et reposait sur une récurrence sur la taille des matrices, où l'hypothèse de récurrence est utilisée deux fois. Cette démonstration ne fut jamais publiée, bien que je la mis au propre en discussion avec Saab Abou-Jaoudé.

Je n'étais toutefois pas satisfait de la limitation sur le cardinal du corps de base. Au début du printemps 2007, et alors que je venais de progresser dans mes méthodes sur les sous-espaces vectoriels de matrices de rang majoré, je tentai de revenir au problème. L'étude du rang majoré m'avait fait prendre conscience de l'intérêt d'examiner en priorité les matrices de tout petit rang dans les espaces que je manipulais, et j'essayai de transposer l'idée aux sous-espaces de matrices nilpotentes. En une journée de travail (prolongée tard dans la nuit), je parvins à lever toute restriction sur le cardinal du corps de base dans l'étude des sous-espaces nilpotents de dimension maximale, d'abord avec une méthode adaptée à tous les corps ayant plus de deux éléments, puis je parvins à traiter le cas général par une analyse encore plus fine de la situation. J'avais mis au point ce que j'appelle désormais la méthode de « compatibilité diagonale », que je décrirai dans le paragraphe 6.5.3. Je publiai enfin le tout à la RMS<sup>2</sup>, parfaitement ignorant du caractère novateur de plusieurs des méthodes que j'avais trouvées alors. Puis je soumis aussi la détermination des automorphismes de  $\mathfrak{sl}_n(\mathbb{K})$  conservant la nilpotence<sup>3</sup>, sans avoir consulté de près les travaux de Botta, Pierce et Watkins sur la question [Bot83].

Près de trois années s'écoulèrent avant que je ne me remisse au travail. En cette fin d'hiver 2010, mon intérêt était accaparé par les propriétés des grands espaces vectoriels de matrices (autrement dit, de ceux de petite codimension). Mon intérêt avait notamment été excité par un renforcement du théorème de Flanders : si l'on prend un sous-espace vectoriel  $\mathcal{V}$  de codimension au plus  $n-1$  de  $M_n(\mathbb{K})$ , non seulement il contient une matrice inversible, mais il possède une base de matrices inversibles (sauf dans le cas particulier  $n = |\mathbb{K}| = 2$ ). Initialement, j'avais tenté de démontrer cela de manière complètement élémentaire, rédigeant un article complet sur la question. Mais alors que la rédaction était bien avancée, je pris conscience que tout cela n'était qu'une simple conséquence de la version

---

2. C. de Seguins Pazzis, Sous-espaces vectoriels de matrices nilpotentes. RMS **118-1** (2007) 36–53.

3. C. de Seguins Pazzis, Automorphismes de l'espace vectoriel  $\mathcal{M}_n(\mathbb{K})$  préservant la nilpotence. RMS **120-1** (2009) 16–32.

affine du théorème de Flanders (voir la section 2.5.1). Et je m'intéressai ensuite à la question plus générale de savoir si  $\mathcal{V}$  est engendré par ses matrices de rang  $r$  pour un  $r \in \llbracket 1, n - 1 \rrbracket$  donné. Je pris alors conscience que ce problème était intimement lié à celui des sous-espaces vectoriels à spectre trivial. Et je me rappelai alors d'une remarque que j'avais faite dans mon article pour la RMS sur les sous-espaces vectoriels de matrices nilpotentes : un des lemmes de base me permettant d'attaquer le cas d'égalité sur un corps quelconque permettait aussi de majorer la dimension, et sa démonstration ne requérait que l'hypothèse voulant que le sous-espace envisagé est à spectre trivial<sup>4</sup>. J'avais donc, trois années auparavant, résolu sans le savoir la conjecture posée par Meshulam en 1989! Je m'empressai donc de soumettre l'article [1]. Et je fus presque devancé par Rachel Quinlan [Qui11], qui avait résolu un problème dual avec une méthode plus compliquée. La publication quasi-simultanée en 2011 de ces deux démonstrations piqua mon orgueil, et je me résolus alors à caractériser le cas d'égalité pour les sous-espaces de matrices à spectre trivial. Ma méthode, ainsi qu'une très bonne connaissance de ma démonstration du théorème de Gerstenhaber publiée dans la RMS, me permit de résoudre complètement la question en quelques jours [4], ne laissant en suspens que le cas des corps à deux éléments. Je généralisai ensuite les résultats pour caractériser complètement les sous-espaces affines de rang minimal  $r$  et de dimension maximale (problème posé initialement par Meshulam dans [Mes89]). Et je conclus cette série en étudiant la dimension maximale pour un sous-espace vectoriel de matrices possédant au plus une valeur propre dans le corps de base, et caractérisai les sous-espaces de dimension maximale parmi ceux constitués de matrices ayant une seule valeur propre dans une clôture algébrique du corps de base [3].

L'année suivante (2012), je poursuivis dans la même veine : d'une part je publiai une généralisation [5] du théorème de Gerstenhaber aux corps gauche de dimension finie sur leur centre (on considère donc des sous-espaces vectoriels sur un sous-corps d'indice fini dans ce centre), mettant en évidence, mieux encore que dans mes publications précédentes, la méthode de compatibilité diagonale ; et ensuite je réalisai le mastodonte, à savoir mon travail sur les sous-espaces vectoriels de matrices ayant au plus deux valeurs propres [6]. Ce travail de plus de 100 pages représente en quelque sorte l'aboutissement – voire l'épuisement – des méthodes imaginées des années auparavant.

Presque simultanément, et alors que je ne travaillais absolument pas dans cette direction, les idées avec lesquelles je jouais à l'époque pour étudier les espaces localement liés (voir le chapitre 3) m'ont rappelé de manière troublante le principe d'*intransitivité* que j'avais utilisé à répétition pour élucider la structure des sous-espaces de matrices à spectre trivial de dimension maximale dans [4]. Par une astuce de dualité, je parvins à montrer, moyennant la condition de cardinalité du théorème de Gerstenhaber, que le théorème de classification que j'avais établi pour ces derniers pouvait être déduit, avec une démonstration très courte, du théorème d'Atkinson sur la structure des sous-espaces primitifs de matrices de rang majoré vérifiant une condition extrême entre le nombre de ligne et le nombre de colonnes (voir la section 6.6). Et je publiai prestement cette découverte [7]. En réutilisant cette remarque, je parvins l'année suivante à élucider complètement la structure des sous-espaces de  $M_3(\mathbb{F}_2)$  de dimension 3 et à spectre trivial [8], en commençant par rectifier un théorème de Beasley

---

4. Voir l'exercice 4 page 46 de mon article RMS de 2007 op. cit.

sur la structure des sous-espaces de matrices de rang au plus 2 sur  $\mathbb{F}_2$ .

C'est à l'été 2016 que je lançai une dernière grande vague de travaux sur les sous-espaces de matrices nilpotentes. L'élément déclencheur fut un exposé de Rachel Quinlan à la conférence ILAS de Leuven : elle y présenta un étrange contre-exemple relatif à un espace de matrices alternées nilpotentes à coefficients dans  $\mathbb{F}_2$ . Cet espace violait une conjecture que l'on pouvait formuler par comparaison à des travaux de Meshulam et Radwan [Mes98] sur le corps des complexes. Cet exposé attira mon attention sur cet article, que je n'avais pas considéré avec suffisamment d'attention jusqu'alors. Je compris vite qu'à la fois l'exemple de Quinlan et les travaux de Meshulam et Radwan rentraient dans un cadre plus vaste, et j'eus très vite de bonnes idées pour résoudre les grandes questions qui se posaient dans ce cadre. Il s'agissait d'étudier le problème des sous-espaces d'endomorphismes nilpotents qui sont aussi autoadjoints ou anti-autoadjoints pour une forme bilinéaire symétrique ou alternée donnée, ce que j'ai appelé le problème de Gerstenhaber *structuré*, mais qu'on aurait aussi bien pu appeler problème de Gerstenhaber *quadratique*. Alors que j'avais proposé à Rachel Quinlan de travailler avec moi, nous avons rapidement dû convenir que je prendrai les choses en main seul, et j'ai totalement résolu le problème de la majoration de la dimension dans les trois articles [9, 10, 11] ainsi qu'une partie du problème sur la détermination des sous-espaces vectoriels de dimension maximale. Beaucoup reste encore à faire sur ce sujet . . .

## 6.5 Sous-espaces de matrices à spectre trivial

### 6.5.1 Des sous-espaces affines de rang minoré aux sous-espaces vectoriels à spectre trivial

**Définition 9.** Un sous-espace vectoriel  $\mathcal{V}$  de  $M_n(\mathbb{K})$  est dit à **spectre trivial** lorsqu'aucun élément de  $\mathcal{V}$  ne possède une valeur propre non nulle dans  $\mathbb{K}$ .

Grâce à la stabilité de  $\mathcal{V}$  par multiplication par les scalaires non nuls, il revient au même de dire qu'aucun élément de  $\mathcal{V}$  ne possède un vecteur fixe non nul.

Il s'agit là d'une généralisation de la notion de sous-espace vectoriel nilpotent, généralisation qui pourrait paraître un rien gratuite si ces espaces n'apparaissaient pas de façon naturelle dans la recherche des sous-espaces *affines* de rang minoré, comme cela a été observé en premier par Meshulam [Mes89]. Soit donc  $\mathcal{V} \subset M_{n,p}(\mathbb{K})$  un sous-espace affine dans lequel le rang minimal est un entier  $r > 0$ . Quitte à remplacer  $\mathcal{V}$  par un espace équivalent, on peut supposer que  $\mathcal{V}$  contient la matrice  $J_r = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ . Notons alors  $W$  le sous-espace vectoriel

de la direction  $V$  de  $\mathcal{V}$  formé des éléments de la forme  $M = \begin{bmatrix} K(M) & 0 \\ 0 & 0 \end{bmatrix}$  avec  $K(M) \in M_r(\mathbb{K})$ . On a donc  $I_r + \lambda K(M)$  inversible pour tout  $M \in W$  et tout  $\lambda \in \mathbb{K}^*$ , si bien que  $K(W)$  est à spectre trivial. En outre, le théorème du rang montre que  $\text{codim } \mathcal{V} \geq \text{codim } K(W)$ . Réciproquement, si l'on se donne un sous-espace vectoriel  $V'$  de  $M_r(\mathbb{K})$  à spectre trivial, alors l'ensemble des matrices de la forme

$$\begin{bmatrix} I_r + M' & [?]_{r \times (p-r)} \\ [?]_{(n-r) \times r} & [?]_{(n-r) \times (p-r)} \end{bmatrix} \quad \text{où } M' \in V',$$

est un sous-espace affine de  $M_{n,p}(\mathbb{K})$  où le rang minimal est  $r$ . Et cet espace affine est de codimension  $\text{codim } V'$  dans  $M_{n,p}(\mathbb{K})$ .

Ainsi, le problème de la codimension minimale pour un sous-espace affine de matrices de rang au moins  $r$  est entièrement lié au problème de la dimension maximale pour un sous-espace vectoriel à spectre trivial de  $M_r(\mathbb{K})$ .

Dans un deuxième temps, on peut bien sûr se demander quelle est la motivation pour étudier les sous-espaces affines de matrices de rang minoré, au-delà du caractère somme toute naturel de la question. Une motivation possible est dans la génération d'un sous-espace vectoriel  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  par ses matrices de rang  $r$  donné, et c'est dans cette optique que nous avons attaqué le problème dans notre première publication sur le sujet [1].

Donnons-donc trois entiers  $n \geq p \geq r > 0$ , et prenons un sous-espace vectoriel  $\mathcal{V}$  de  $M_{n,p}(\mathbb{K})$  de dimension suffisamment grande. Grande comment ? Tout simplement, assez grande pour assurer que  $\mathcal{V}$  possède des matrices de rang arbitraire. Par référence au théorème de Flanders-Dieudonné (voir le chapitre 2) nous prendrons tout simplement  $\text{codim } \mathcal{V} < n$  afin d'assurer que  $\mathcal{V}$  contienne une matrice de rang  $p$ . C'est non seulement suffisant pour que  $\mathcal{V}$  contienne des matrices de rang arbitraire, encore mieux ça l'est aussi pour que les matrices de rang  $r$  de  $\mathcal{V}$ , où  $r$  est fixé dans  $[[0, p]]$ , engendrent par combinaison linéaire toutes les matrices de  $\mathcal{V}$  de rang *inférieur* à  $r$ , sauf dans le cas critique  $n = p = r = 2$  et  $|\mathbb{K}| = 2$ , que l'on peut mettre de côté sans effroi. L'astuce est la suivante : on part de  $M$  dans  $\mathcal{V}$  de rang  $s \leq r$ . Quitte à remplacer  $\mathcal{V}$  par un espace équivalent, on peut supposer que  $M$  a toutes ses colonnes nulles à compter de la  $(s + 1)$ -ième. On considère alors le sous-espace vectoriel  $\mathcal{W}$  constitué des matrices de  $\mathcal{V}$  qui ont toutes leurs colonnes nulles à compter de la  $(r + 1)$ -ième, et on l'identifie à un sous-espace  $\mathcal{W}'$  de  $M_{n,r}(\mathbb{K})$ . Ce dernier est de codimension inférieure ou égale à celle de  $\mathcal{V}$ , donc un corollaire du théorème de Flanders-Dieudonné (généralisé par nos soins, voir le chapitre 2) montre que  $\mathcal{W}'$  est engendré par ses matrices de rang  $r$ , et en particulier  $M$  est combinaison linéaire de matrices de rang  $r$  de  $\mathcal{W}$  donc de  $\mathcal{V}$ .

Ensuite, on s'intéresse, et toujours sous les mêmes hypothèses, au sous-espace vectoriel de  $\mathcal{V}$  engendré par ses matrices de rang  $r$ . On vient de démontrer qu'il contient toutes les matrices de  $\mathcal{V}$  de rang au plus  $r$ . S'il n'est pas égal à  $\mathcal{V}$ , alors c'en est un sous-espace vectoriel strict ; tout sous-espace affine  $\mathcal{V}'$  de  $\mathcal{V}$  qui lui est strictement parallèle contient alors uniquement des matrices de rang strictement supérieur à  $r$ . On obtient donc une contradiction si la codimension de  $\mathcal{V}$  est strictement inférieure à  $d_{r+1,n,p} - 1$ , où  $d_{r+1,n,p}$  est la codimension minimale d'un sous-espace affine de  $M_{n,p}(\mathbb{K})$  dont les éléments sont de rang au moins  $r + 1$ .

## 6.5.2 Majoration de la dimension d'un sous-espace à spectre trivial : la méthode combinatoire

Presque tout est dit sur l'article [1], sinon la méthode pour majorer la dimension d'un sous-espace vectoriel de  $M_n(\mathbb{K})$  à spectre trivial. Comme indiqué dans le paragraphe 6.4, je savais établir la majoration optimale dès 2007 mais je ne l'avais pas publiée comme telle (je ne considérais pas alors les sous-espaces à spectre trivial comme des objets dignes d'intérêt). Il n'est pas inutile d'expliquer l'origine de ma démonstration. Il s'agissait à l'époque de donner une

démonstration du théorème de Gerstenhaber faisant fi de l'hypothèse de cardinalité sur le corps de base. Or, dans la première démonstration que j'étais parvenu à élaborer, cette hypothèse de cardinalité était indispensable pour montrer que le sous-espace  $\mathcal{V}$  envisagé (nilpotent, de dimension maximale) contenait nécessairement un élément de rang  $n - 1$ . Ensuite, on essayait tant bien que mal de démontrer que les vecteurs du noyau de cet élément étaient annulés par tous les autres éléments de  $\mathcal{V}$ . Mais les techniques de base utilisaient à fond la méthode de Flanders pour exploiter des majorations de rang (que l'on couplait avec des techniques de trace), et il n'en était pas question ici. Est alors venue l'idée qu'il ne fallait pas être trop pressé : puisque l'on cherchait en définitive un *drapeau* stable par tout élément de  $\mathcal{V}$ , on pouvait déjà chercher à changer la base de représentation matricielle de telle sorte que le dernier vecteur soit acceptable. Et il fallait pour cela que  $\mathcal{V}$  ne contienne aucun élément de rang 1 et ayant dans son image le dernier vecteur de base.

En outre, le théorème de la base incomplète indiquait que, pour peu que le résultat à démontrer soit vrai, le dernier vecteur d'une bonne base pourrait être choisi parmi les vecteurs de la base *canonique* de  $\mathbb{K}^n$ . Autrement dit, il s'agissait de montrer, pour un certain entier  $i \in \llbracket 1, n \rrbracket$ , la nullité du sous-espace vectoriel

$$\mathcal{L}_i(\mathcal{V}) := \{M \in \mathcal{V} : \forall k \in \llbracket 1, n \rrbracket \setminus \{i\}, L_k(M) = 0\}$$

des éléments de  $\mathcal{V}$  dont toutes les lignes sont nulles à l'exception éventuelle de la  $i$ -ème. Et, coup de chance, cet énoncé est non seulement vrai pour les sous-espaces à spectre trivial de dimension maximale, il vaut pour n'importe quel sous-espace à spectre trivial ! L'importance de ce résultat, combinée à la simplicité de la démonstration, mérite que nous détaillions cette dernière :

**Lemme 6.2** (Proposition 10 de [6]). *Soit  $\mathcal{V}$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  à spectre trivial. Il existe alors un indice  $i \in \llbracket 1, n \rrbracket$  tel que  $\mathcal{L}_i(\mathcal{V}) = \{0\}$ .*

*Démonstration.* On procède par récurrence sur  $n$ , le résultat étant évident si  $n = 1$ . Supposons  $n > 1$ . Raisonnons par l'absurde en supposant  $\mathcal{L}_i(\mathcal{V}) \neq \{0\}$  pour tout  $i \in \llbracket 1, n \rrbracket$ .

Notons  $\mathcal{V}'$  le sous-espace vectoriel de  $\mathcal{V}$  formé des matrices dont la dernière ligne est nulle, et écrivons tout élément de  $\mathcal{V}'$  sous la forme

$$M = \begin{bmatrix} K(M) & [?]_{(n-1) \times 1} \\ [0]_{1 \times (n-1)} & 0 \end{bmatrix}.$$

Le sous-espace  $K(\mathcal{V})$  de  $M_{n-1}(\mathbb{K})$  est visiblement à spectre trivial. L'hypothèse de récurrence donne donc un indice  $i \in \llbracket 1, n - 1 \rrbracket$  tel que  $\mathcal{L}_i(K(\mathcal{V})) = \{0\}$ . Puisque  $\mathcal{L}_i(\mathcal{V}) \neq \{0\}$ , il vient  $E_{i,n} \in \mathcal{V}$ .

En conjuguant  $\mathcal{V}$  par des matrices de permutation, on généralise le principe précédent et l'on obtient que pour tout  $j \in \llbracket 1, n \rrbracket$  on peut choisir un entier  $f(j) \in \llbracket 1, n \rrbracket \setminus \{j\}$  tel que  $E_{f(j),j} \in \mathcal{V}$ . Une contradiction est alors obtenue en remarquant qu'on peut prendre un cycle  $(j_1, \dots, j_r)$  pour la fonction  $f$  (de sorte que  $f(j_k) = j_{k+1}$  pour tout  $k \in \llbracket 1, r - 1 \rrbracket$ , et  $f(j_r) = j_1$ ), et qu'alors  $\sum_{k=1}^r E_{f(j_r),j_r}$  est dans  $\mathcal{V}$  et possède évidemment 1 pour valeur propre.  $\square$

Le lemme précédent permet ensuite de majorer la dimension des sous-espaces à spectre trivial, par récurrence.

**Théorème 6.3** (Théorème 9 de [6]). *Soit  $\mathcal{V}$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  à spectre trivial. Alors  $\dim \mathcal{V} \leq \frac{n(n-1)}{2}$ .*

*Démonstration.* On procède par récurrence sur  $n$ . Grâce au lemme précédent, on peut conjuguer  $\mathcal{V}$  par une matrice de permutation pour se ramener au cas où  $\mathcal{L}_n(\mathcal{V}) = \{0\}$ . Dans ce cas, on introduit le sous-espace  $\mathcal{W}$  de  $\mathcal{V}$  formé des matrices de  $\mathcal{V}$  de la forme  $M = \begin{bmatrix} K(M) & [0]_{(n-1) \times 1} \\ L(M) & ? \end{bmatrix}$ , si bien que  $\dim \mathcal{V} \leq \dim \mathcal{W} + (n-1)$ .

L'hypothèse  $\mathcal{L}_n(\mathcal{V}) = \{0\}$  montre que  $\dim \mathcal{W} = \dim K(\mathcal{W})$ .

Ensuite,  $K(\mathcal{W})$  est lui-même à spectre trivial, donc  $\dim K(\mathcal{W}) \leq \frac{(n-1)(n-2)}{2}$  par l'hypothèse de récurrence.

Il vient finalement  $\dim \mathcal{V} \leq \frac{(n-1)(n-2)}{2} + (n-1) = \frac{n(n-1)}{2}$ .  $\square$

En adaptant la démonstration, on peut même montrer par récurrence que pour tout sous-espace  $\mathcal{V}$  de  $M_n(\mathbb{K})$  à spectre trivial, il existe une matrice de permutation  $P$  telle que  $PVP^{-1}$  intersecte trivialement l'espace  $T_n^-(\mathbb{K})$  des matrices triangulaires inférieures. La majoration de la dimension en découle immédiatement.

### 6.5.3 La méthode de compatibilité diagonale

Les quatre travaux [3, 4, 5, 6] sont tous fondés sur la même stratégie, elle-même calquée sur la démonstration du théorème de Gerstenhaber que j'avais produite en 2017 pour la RMS. La première étape reprend ce qui est décrit dans le paragraphe précédent : elle consiste, pour un sous-espace vectoriel  $\mathcal{V}$  de  $M_n(\mathbb{K})$  vérifiant des conditions de limitation de spectre, à trouver un vecteur  $x \in \mathbb{K}^n \setminus \{0\}$  **adapté à  $\mathcal{V}$** , c'est-à-dire tel que  $\mathcal{V}$  ne contienne pas de matrice d'image  $\mathbb{K}x$  et de trace nulle. Je reviendrai brièvement sur la problématique d'une telle recherche dans le paragraphe 6.5.7. Dès que l'on a l'existence d'un tel vecteur, les majorations de dimension s'obtiennent par récurrence sur la taille des matrices, à nouveau comme dans le paragraphe précédent.

Ensuite, on cherche à analyser la structure des sous-espaces de dimension maximale. La méthode, que j'ai baptisée « compatibilité diagonale », est fondée sur une récurrence sur la taille des matrices, où l'on utilise l'hypothèse de récurrence à deux reprises. La meilleure façon d'expliquer et d'illustrer cette méthode est de se placer dans la situation la plus simple, autrement dit de revenir au théorème de Gerstenhaber sur les sous-espaces vectoriels nilpotents. Et pour mieux comprendre les choses, nous partons de la version géométrique du problème. Soit donc  $E$  un espace vectoriel de dimension  $n$ , et  $\mathcal{V}$  un sous-espace vectoriel nilpotent de  $\mathcal{L}(E)$ , que l'on suppose de dimension  $\frac{n(n-1)}{2}$ . Il s'agit de montrer que, dans une base appropriée, les éléments de  $\mathcal{V}$  sont représentés par les matrices triangulaires supérieures strictes. Autrement dit, on cherche un drapeau complet  $(E_0, \dots, E_n)$  de  $E$  stable par les éléments de  $\mathcal{V}$  (il n'y en aura d'ailleurs qu'un). L'idée est de construire très progressivement ce drapeau, en commençant par chercher un bon candidat pour un vecteur extérieur à l'hyperplan  $E_{n-1}$  du drapeau. Un tel vecteur devra évidemment être choisi parmi les vecteurs adaptés, et une fois qu'il sera choisi on n'aura logiquement plus besoin de le modifier. Prenons donc un vecteur  $e_n$  adapté à  $\mathcal{V}$  (il en existe un, d'après le lemme 6.2). Considérons alors le sous-espace  $\mathcal{V}' := \{u \in \mathcal{V} : u(e_n) = 0\}$ . Les éléments  $u$  de  $\mathcal{V}'$  induisent tous des endomorphismes nilpotents de  $E/\mathbb{K}e_n$  ; on



recupère ainsi un sous-espace nilpotent  $\mathcal{V}_h$  ( $h$  pour « haut ») de  $\mathcal{L}(E/\mathbb{K}e_n)$ . Le noyau de la projection canonique  $u \in \mathcal{V}' \mapsto \bar{u} \in \mathcal{V}_h$  est constitué des éléments de  $\mathcal{V}$  qui s'annulent en  $e_n$  et ont leur image incluse dans  $\mathbb{K}e_n$ . Comme  $e_n$  est adapté à  $\mathcal{V}$ , ce noyau est nul, donc  $\dim \mathcal{V}_h = \dim \mathcal{V}'$ . Arrêtons-nous brièvement pour illustrer matriciellement la situation en représentant  $\mathcal{V}$  par un espace de matrices  $\mathcal{M}$  dans la base  $(e_1, \dots, e_n)$ . L'hypothèse d'adaptation indique que  $\mathcal{M}$  ne contient aucune matrice non nulle de la forme

$$\begin{bmatrix} [0]_{(n-1) \times (n-1)} & [0]_{(n-1) \times 1} \\ [?]_{1 \times (n-1)} & 0 \end{bmatrix}.$$

Les éléments de l'espace haut  $\mathcal{V}_h$  sont ceux représentés dans  $(e_1, \dots, e_{n-1})$  par les matrices  $N$  telles que  $\mathcal{M}$  contienne une matrice de la forme

$$\begin{bmatrix} N & [0]_{(n-1) \times 1} \\ [?]_{1 \times (n-1)} & 0 \end{bmatrix}.$$

Reprenons maintenant la vision géométrique. Une double utilisation du théorème du rang donne

$$\dim \mathcal{V} = \dim(\mathcal{V}e_n) + \dim \mathcal{V}' = \dim(\mathcal{V}e_n) + \dim \mathcal{V}_h.$$

Par ailleurs  $\dim \mathcal{V}e_n < n$  car  $e_n \notin \mathcal{V}e_n$ , alors que  $\dim \mathcal{V}_h \leq \frac{(n-1)(n-2)}{2}$ . Comme on a supposé  $\dim \mathcal{V} = \frac{n(n-1)}{2}$ , on en déduit  $\dim \mathcal{V}e_n = n - 1$  et  $\dim \mathcal{V}_h = \frac{(n-1)(n-2)}{2}$ , puis une première utilisation de l'hypothèse de récurrence donne une base  $(\bar{e}_1, \dots, \bar{e}_{n-1})$  de  $V/\mathbb{K}e_n$  telle que  $\mathcal{V}_h$  soit l'ensemble des endomorphismes représentés dans cette base par un élément de  $\mathbb{T}_{n-1}^{++}(\mathbb{K})$ .

L'astuce est maintenant d'observer que l'hyperplan  $H := \text{Vect}(e_2, \dots, e_n)$  est adapté à  $\mathcal{V}$ , au sens où tout élément de  $\mathcal{V}$  nul sur  $H$  est identiquement nul (autrement dit, tout vecteur non nul du dual-orthogonal de  $H$  est adapté à l'espace transposé  $\mathcal{V}^t$ ). En effet, tout élément  $u \in \mathcal{V}$  nul sur  $H$  appartient en particulier à  $\mathcal{V}'$  et vérifie  $\bar{u} = 0$ . Prenons alors l'espace  $\mathcal{V}''$  formé des éléments de  $\mathcal{V}$  dont l'image est incluse dans  $H$ . Les éléments de  $\mathcal{V}''$  induisent des endomorphismes nilpotents de  $H$ . On peut noter  $\mathcal{V}_b$  l'espace formé par ces derniers ( $b$  pour « bas »). Et à nouveau, on trouve que  $\dim \mathcal{V}_b = \frac{(n-1)(n-2)}{2}$ . L'hypothèse de récurrence donne donc un drapeau complet de  $H$  stable par tous les éléments de  $\mathcal{V}_b$ .

Et c'est là qu'intervient la fameuse **compatibilité diagonale** : on va combiner les hypothèses sur les formes respectives de  $\mathcal{V}_h$  et  $\mathcal{V}_b$  pour obtenir que le drapeau correspondant à  $\mathcal{V}_b$  est très proche de celui associé à  $(e_2, \dots, e_n)$ . Dans le cas particulier qui nous préoccupe, on commence par une dernière simplification : en se rappelant que  $\mathcal{V}e_n$  est un hyperplan de  $E$  ne contenant pas  $e_n$ , on peut modifier  $e_1, \dots, e_{n-1}$  sans changer leur classe modulo  $\mathbb{K}e_n$ , de telle sorte que  $\mathcal{V}e_n = \text{Vect}(e_1, \dots, e_{n-1})$ . Représentons les choses matriciellement dans la base  $(e_1, \dots, e_n)$  à l'aide de découpages selon le format suivant

$$\begin{bmatrix} ? & [?]_{1 \times (n-2)} & ? \\ [?]_{(n-2) \times 1} & [?]_{(n-2) \times (n-2)} & [?]_{(n-2) \times 1} \\ ? & [?]_{1 \times (n-2)} & ? \end{bmatrix}.$$

Vu la forme des éléments de  $\mathcal{V}_h$  et l'hypothèse voulant que  $e_n$  est adapté à  $\mathcal{V}$ , on sait que pour toute ligne  $L \in M_{1, n-2}(\mathbb{K})$ , il existe dans  $\mathcal{V}$  un unique élément

représenté par une matrice de la forme

$$A_L = \begin{bmatrix} 0 & L & 0 \\ 0 & 0 & 0 \\ ? & ? & 0 \end{bmatrix}.$$

Vu l'hypothèse  $\mathcal{V}e_n = \text{Vect}(e_1, \dots, e_{n-1})$ , on sait aussi que pour toute colonne  $C \in \mathbb{K}^{n-2}$ , il existe dans  $\mathcal{V}$  un élément représenté par une matrice de la forme

$$\begin{bmatrix} ? & ? & 0 \\ ? & ? & C \\ ? & ? & 0 \end{bmatrix},$$

puis en ajoutant une matrice de type  $A_L$  et éliminant le coefficient en haut à gauche (par nilpotence), on peut raffiner davantage pour obtenir un élément de  $\mathcal{V}$  représenté par une matrice de la forme

$$\begin{bmatrix} 0 & 0 & 0 \\ ? & ? & C \\ ? & ? & 0 \end{bmatrix}.$$

En particulier, on trouve que  $\text{Vect}(e_2, \dots, e_{n-1}) \subset \mathcal{V}_b e_n$ , donc nécessairement  $\text{Vect}(e_2, \dots, e_{n-1})$  est l'hyperplan du drapeau associé à  $\mathcal{V}_b$ . Enfin, pour tout  $U \in \Gamma_{n-2}^{++}(\mathbb{K})$ , on sait que  $\mathcal{V}$  contient un élément représenté par une matrice de la forme

$$\begin{bmatrix} 0 & 0 & 0 \\ ? & U & 0 \\ ? & ? & ? \end{bmatrix},$$

et ce que l'on vient d'obtenir sur le drapeau associé à  $\mathcal{V}_b$  montre qu'en fait cette matrice est de la forme

$$H_U = \begin{bmatrix} 0 & 0 & 0 \\ ? & U & 0 \\ ? & 0 & 0 \end{bmatrix},$$

et même, vu la forme de  $\mathcal{V}_h$  :

$$H_U = \begin{bmatrix} 0 & 0 & 0 \\ 0 & U & 0 \\ ? & 0 & 0 \end{bmatrix}.$$

On reconnaît alors le restant du drapeau adapté à  $\mathcal{V}_b$  en faisant agir les endomorphismes associés à ces dernières matrices sur  $\text{Vect}(e_2, \dots, e_{n-1})$ . L'étape de compatibilité est donc franchie, et l'on sait désormais que, pour tout  $C \in \mathbb{K}^{n-2}$ , il existe dans  $\mathcal{V}$  un élément représenté par une matrice de la forme

$$B_C = \begin{bmatrix} 0 & 0 & 0 \\ ? & 0 & C \\ ? & 0 & 0 \end{bmatrix},$$

et une telle matrice est même unique car  $H$  est adapté à  $\mathcal{V}$  (au sens dual expliqué plus haut).

À ce stade, la base  $(e_1, \dots, e_n)$  obtenue n'est pas nécessairement convenable,  $e_1$  n'étant connu que modulo la droite  $\mathbb{K}e_n$ . Il restera donc une dernière transvection à réaliser, sur laquelle nous reviendrons sous peu.

Nous avons mis en évidence, dans l'espace de matrices  $\mathcal{M}$  représentant  $\mathcal{V}$  dans la base  $(e_1, \dots, e_n)$ , trois familles de matrices de forme assez simple (les familles  $(A_L)_L$ ,  $(B_C)_C$  et  $(H_U)_U$ ). Il ne manque, pour engendrer  $\mathcal{M}$ , qu'une matrice de la forme  $\begin{bmatrix} ? & ? & 1 \\ ? & ? & ? \\ ? & ? & 0 \end{bmatrix}$ . À partir de là, l'idée est de réduire progressivement la forme de ces matrices simples, en les combinant intelligemment et en utilisant l'hypothèse de nilpotence. En guise d'illustration, voyons par exemple comment simultanément simplifier  $A_L$  et  $B_C$  et reconnaître la dernière modification de base à réaliser. On sait que

$$A_L = \begin{bmatrix} 0 & L & 0 \\ 0 & 0 & 0 \\ ? & f(L) & 0 \end{bmatrix} \quad \text{et} \quad B_C = \begin{bmatrix} 0 & 0 & 0 \\ g(C) & 0 & C \\ ? & 0 & 0 \end{bmatrix}$$

pour des fonctions linéaires  $f$  et  $g$ . On observe que si  $LC = 0$  alors

$$(A_L + B_C)^2 = \begin{bmatrix} Lg(C) & 0 & 0 \\ ? & ? & 0 \\ ? & ? & f(L)C \end{bmatrix}$$

et l'on en tire que  $Lg(C) = 0 = f(L)C$ . Il est facile d'en déduire que  $f$  et  $g$  sont des homothéties. Puis, en appliquant par exemple le lemme de trace à  $A_L$  et  $B_C$  (voir la section 6.7), on trouve que ces rapports d'homothétie sont opposés. Et si  $f : L \mapsto \alpha L$ , alors il suffira de remplacer  $e_1$  par  $e_1 + \alpha e_n$  pour se ramener au cas où  $\alpha = 0$ . Dans ce cas réduit, on obtient  $f = g = 0$  et les matrices particulières trouvées prennent donc une forme encore plus simple. En poursuivant par une analyse fine des matrices obtenues (celles des trois familles, ainsi que la matrice surnuméraire) on parvient progressivement à montrer qu'elles sont toutes triangulaires supérieures strictes, ce qui achève la démonstration.

Bien sûr, dans ce qui précède on était dans le cadre le plus simple où cette stratégie s'applique (sous-espaces vectoriels de matrices nilpotentes). Lorsqu'on veut l'appliquer à des cas plus complexes, la technicité augmente sensiblement mais les grands principes restent. Dans les paragraphes suivants, je vais m'attacher à énoncer les principaux résultats qui ont été permis par cette technique, en évoquant exceptionnellement les petites adaptations dans la méthode (soit pour la recherche de vecteurs adaptés, soit pour la façon d'obtenir la compatibilité diagonale, soit enfin pour obtenir les simplifications de matrices).

#### 6.5.4 Sous-espaces vectoriels à spectre trivial de dimension maximale

On a vu dans la section 6.5.2 que tout sous-espace vectoriel de  $M_n(\mathbb{K})$  à spectre trivial est de dimension inférieure à  $\frac{n(n-1)}{2}$ . Bien sûr, en général un tel sous-espace, même s'il est de dimension maximale, n'est pas entièrement constitué de matrices nilpotentes. Pour simplifier le discours, un sous-espace vectoriel de  $M_n(\mathbb{K})$  à spectre trivial sera dit **optimal** lorsqu'il est de dimension  $\frac{n(n-1)}{2}$ . L'exemple standard d'espace optimal qui ne contient aucune matrice nilpotente non nulle est celui de l'espace  $A_n(\mathbb{R})$  des matrices antisymétriques réelles.

Voici une forme de généralisation de cette exemple, en version endomorphismes : on se donne une forme bilinéaire non dégénérée  $b : E \times E \rightarrow \mathbb{K}$ , où  $E$  est de dimension  $n$  (attention  $b$  n'est pas supposée symétrique ou antisymétrique). L'ensemble

$$\mathcal{A}_b := \{u \in \mathcal{L}(E) : \forall x \in E, b(x, u(x)) = 0\}$$

des endomorphismes  $b$ -alternés (à droite) est un sous-espace vectoriel de dimension  $\frac{n(n-1)}{2}$  de  $\mathcal{L}(E)$ . Il est bien clair que si  $b$  est anisotrope alors  $\mathcal{A}_b$  est à spectre trivial. Matriciellement, ces espaces correspondent à ceux de la forme  $PA_n(\mathbb{K})$  où  $P \in \text{GL}_n(\mathbb{K})$  est *anisotrope*, i.e.  $X^T P X \neq 0$  pour tout  $X \in \mathbb{K}^n \setminus \{0\}$ . En outre, on peut démontrer que ces espaces sont systématiquement irréductibles (au sens des espaces d'endomorphismes : il n'existe aucun sous-espace non trivial stable par tous ses éléments).

On peut ensuite construire d'autres exemples par l'opération de *jonction* : pour deux sous-ensembles  $\mathcal{A} \subset M_n(\mathbb{K})$  et  $\mathcal{B} \subset M_p(\mathbb{K})$ , on note

$$\mathcal{A} \vee \mathcal{B} := \left\{ \begin{bmatrix} A & C \\ 0 & B \end{bmatrix} \mid (A, B) \in \mathcal{A} \times \mathcal{B}, C \in M_{n,p}(\mathbb{K}) \right\} \subset M_{n+p}(\mathbb{K}).$$

On note que si  $\mathcal{A}$  et  $\mathcal{B}$  sont des sous-espaces à spectre trivial, alors  $\mathcal{A} \vee \mathcal{B}$  aussi, et que si en outre  $\mathcal{A}$  et  $\mathcal{B}$  sont optimaux alors  $\mathcal{A} \vee \mathcal{B}$  l'est aussi.

Ainsi, pour toute partition<sup>5</sup>  $(d_1, \dots, d_p)$  de l'entier  $n$ , et toute liste  $(P_1, \dots, P_p)$  de matrices anisotropes respectivement dans  $M_{d_1}(\mathbb{K}), \dots, M_{d_p}(\mathbb{K})$ , l'espace  $P_1 A_{d_1}(\mathbb{K}) \vee \dots \vee P_p A_{d_p}(\mathbb{K})$  est à spectre trivial et optimal. Le théorème majeur établi dans [4] énonce qu'il s'agit là des seules solutions à conjugaison près hormis pour les corps à deux éléments :

**Théorème 6.4** (Théorème 4 de [4]). *On suppose que  $|\mathbb{K}| > 2$ . Soit  $\mathcal{V}$  un sous-espace vectoriel à spectre trivial de  $M_n(\mathbb{K})$ , supposé optimal. Il existe alors une partition  $(d_1, \dots, d_p)$  de  $n$  et une liste  $(P_1, \dots, P_d_p) \in \text{GL}_{d_1}(\mathbb{K}) \times \dots \times \text{GL}_{d_p}(\mathbb{K})$  de matrices anisotropes telle que  $\mathcal{V}$  soit semblable (i.e. conjugué) à  $P_1 A_{d_1}(\mathbb{K}) \vee \dots \vee P_p A_{d_p}(\mathbb{K})$ .*

*La liste  $(d_1, \dots, d_p)$  est déterminée de manière unique, puis chaque  $P_k$  est déterminée de manière unique à congruence près et multiplication par un scalaire non nul près.*

Plus géométriquement, cet énoncé signifie que pour un espace vectoriel  $E$  de dimension finie et un sous-espace vectoriel  $\mathcal{V}$  de  $\mathcal{L}(E)$  à spectre trivial et optimal, il existe un drapeau (*a priori* incomplet)  $(E_0, \dots, E_p)$  de  $E$  (où  $E_0 = \{0\}$  et  $E_p = E$ ) et une liste  $(b_1, \dots, b_p)$  de formes bilinéaires anisotropes sur respectivement  $E_1/E_0, \dots, E_p/E_{p-1}$  telles que  $\mathcal{V}$  soit l'ensemble des endomorphismes  $u$  de  $E$  stabilisant le drapeau et induisant systématiquement sur  $E_k/E_{k-1}$  un endomorphisme  $b_k$ -alterné à droite. On peut même démontrer l'unicité d'un tel drapeau ainsi que l'unicité de chaque  $b_k$  à multiplication près par un scalaire non nul.

On notera que sur un corps quadratiquement clos, toute matrice anisotrope doit être de taille 1, et on obtient alors la même solution que dans le théorème de Gerstenhaber (avant de nous lancer dans la démonstration du théorème 6.4,

<sup>5</sup>. Une partition de  $n$  est une liste d'entiers naturels *non nuls* de somme  $n$ . Attention, on n'impose aucune contrainte sur l'ordre relatif des éléments d'une telle liste.

nous avons déjà fait un travail préliminaire où l'on démontrait précisément cela).

Voici un corollaire immédiat :

**Corollaire 6.5.** *Si  $|\mathbb{K}| > 2$  alors les sous-espaces à spectre trivial optimaux irréductibles de  $M_n(\mathbb{K})$  sont les espaces de la forme  $PA_n(\mathbb{K})$  avec  $P \in GL_n(\mathbb{K})$  anisotrope.*

La condition restrictive sur le cardinal de  $\mathbb{K}$  semble incontournable. Dans la dernière page de [4] figure un exemple de sous-espace à spectre trivial optimal de  $M_3(\mathbb{F}_2)$  qui est irréductible, alors que sur un corps fini une matrice anisotrope doit être de taille au plus 2. Et dans [8], nous avons même classifié tous les sous-espaces à spectre trivial optimaux de  $M_3(\mathbb{F}_2)$ . Il semble qu'aucune forme générale ne se dégage pour le cas de  $M_n(\mathbb{F}_2)$ . La démonstration du théorème 6.4 donnée dans [4] utilise à de très nombreuses reprises l'hypothèse  $|\mathbb{K}| > 2$  et il est peu probable qu'on puisse l'adapter pour traiter le cas de  $\mathbb{F}_2$ .

Quelques mots enfin sur la méthode de démonstration du théorème 6.4. Il s'agit d'une nouvelle adaptation de la méthode de compatibilité diagonale, décrite dans le paragraphe précédent pour les sous-espaces vectoriels de matrices nilpotentes. L'existence d'un vecteur adapté a été expliquée dans le paragraphe 6.5.2. En revanche, il y a ici des difficultés de deux ordres :

- D'abord l'hypothèse de récurrence est particulièrement lourde à manipuler, vu la forme générale qu'on peut avoir sur les espaces hauts et bas  $\mathcal{V}_h$  et  $\mathcal{V}_b$ . Un point critique est tout de même de se ramener systématiquement à des matrices facteurs  $P_i$  qui sont triangulaires *inférieures* (on l'obtient par orthogonalisation à droite de la forme bilinéaire anisotrope sous-jacente). Même comme cela, il faut faire très attention car il n'est pas toujours clair de savoir où aller (autrement dit à quel type on s'attend pour  $\mathcal{V}$  à partir des types respectifs de  $\mathcal{V}_h$  et  $\mathcal{V}_b$ ).
- Ensuite et surtout, on n'a plus aucun accès aux outils de trace, ni possibilité de jouer sur l'élévation des matrices à une certaine puissance, contrairement au cas des matrices nilpotentes. La réduction progressive des matrices *simples* trouvées par application de l'hypothèse de récurrence et des égalités de dimensions obtenues s'avère donc extrêmement laborieuse. Et, si l'on parvient parfois à repérer des vecteurs propres pour certaines matrices obtenues en combinant ces matrices simples, le plus souvent on opère largement à l'aveugle en utilisant un principe de **double-intransitivité**. En effet, la condition de spectre trivial a pour effet que  $\mathcal{V}X \subsetneq \mathbb{K}^n$  et  $\mathcal{V}^T X \subsetneq \mathbb{K}^n$  pour tout  $X \in \mathbb{K}^n \setminus \{0\}$ . Pour exploiter cette hypothèse à plein, on procède le plus souvent en considérant un grand nombre de matrices dans  $\mathcal{V}$ , en examinant l'effet de  $\mathcal{V}$  ou de  $\mathcal{V}^T$  sur un vecteur bien choisi et en calculant les contraintes linéaires issues de l'intransitivité.

Je n'en dirai pas plus sur cette longue démonstration, qui représente plus de quatorze pages très denses à partir de la mise en place de la compatibilité diagonale, sinon que j'ai davantage de satisfaction à savoir ce travail achevé par un aussi beau théorème que j'ai de plaisir à relire les détails de sa démonstration.

### 6.5.5 Applications aux sous-espaces affines de rang minoré

Venons-en maintenant à l'application des sous-espaces à spectre trivial aux sous-espaces affines de rang minoré. Le lien entre les deux concepts a été dûment expliqué dans le paragraphe 6.5.1 et nous n'y reviendrons pas. Citons immédiatement le principal résultat sur le cas critique des sous-espaces affines de matrices inversibles de  $M_n(\mathbb{K})$ .

**Théorème 6.6** (Théorème 7 de [4]). *On suppose  $|\mathbb{K}| > 2$ . Soit  $\mathcal{V}$  un sous-espace affine de  $M_n(\mathbb{K})$ , inclus dans  $GL_n(\mathbb{K})$  et de dimension  $\frac{n(n-1)}{2}$ . Il existe alors une partition  $(d_1, \dots, d_p)$  de l'entier  $n$  ainsi qu'une liste  $(P_1, \dots, P_p)$  de matrices anisotropes de tailles respectives  $d_1, \dots, d_p$  telles que  $\mathcal{V}$  soit équivalent à  $I_n + (P_1 A_{d_1}(\mathbb{K}) \vee \dots \vee P_p A_{d_p}(\mathbb{K}))$ .*

*En outre, la liste  $(d_1, \dots, d_p)$  est entièrement déterminée par  $\mathcal{V}$ , et pour tout  $k \in \llbracket 1, p \rrbracket$  la forme quadratique  $X \mapsto X^T P_k X$  est entièrement déterminée à équivalence près et à multiplication près par un scalaire non nul.*

On notera que si  $\chi(\mathbb{K}) \neq 2$ , les matrices  $P_k$  peuvent ainsi être prises symétriques anisotropes, et elles sont alors déterminées de manière unique à congruence près et à multiplication près par un scalaire non nul.

Ce théorème découle sans réelle difficulté du théorème 6.4. En revanche, la généralisation à un rang minimum quelconque est nettement plus délicate, et elle a donné lieu à un article séparé fort copieux [2]. Le résultat principal de ce dernier s'énonce assez facilement, en utilisant la notation suivante : pour trois entiers naturels  $n, p, r$  avec  $r \leq \min(n, p)$  et une partie  $\mathcal{W}$  de  $M_r(\mathbb{K})$ , on note  $i_{n,p}(\mathcal{W})$  l'ensemble des matrices de la forme

$$\begin{bmatrix} W & [?]_{r \times (p-r)} \\ [?]_{(n-r) \times r} & [?]_{(n-r) \times (p-r)} \end{bmatrix} \text{ avec } W \in \mathcal{W}.$$

Le théorème suivant ne devrait pas surprendre quiconque :

**Théorème 6.7** (Théorème 3 de [2]). *On suppose  $|\mathbb{K}| > 2$ . Soit trois entiers  $n, p, r$  avec  $2 \leq r \leq \min(n, p)$ . Soit  $\mathcal{V}$  un sous-espace affine de  $M_{n,p}(\mathbb{K})$  dans lequel toute matrice est de rang au moins  $r$ . Alors  $\text{codim } \mathcal{V} \geq \frac{r(r+1)}{2}$ . Si en outre  $\text{codim } \mathcal{V} = \frac{r(r+1)}{2}$ , alors  $\mathcal{V}$  est équivalent à  $i_{n,p}(\mathcal{W})$  pour un sous-espace affine  $\mathcal{W}$  de  $M_r(\mathbb{K})$  constitué de matrices inversibles et de dimension  $\frac{r(r-1)}{2}$ , déterminé à équivalence près de manière unique par  $\mathcal{V}$ .*

Dans le cas  $r = 1$ , le résultat ne tient plus car le théorème ne donnerait qu'une seule classe d'équivalence de solutions, alors qu'il y en a  $\min(n, p)$  : les solutions sont les hyperplans affines de  $M_{n,p}(\mathbb{K})$  ne passant pas par 0, ils sont déterminés à équivalence près par le rang des matrices non nulles de leur orthogonal pour la forme  $(A, B) \mapsto \text{tr}(AB)$ .

L'obtention du théorème 6.7 à partir du théorème 6.6 est loin d'être facile. Nous nous contenterons de donner les grandes lignes de la démonstration, et signalerons une erreur facile à rectifier dans la démonstration d'un lemme.

La mise en place est naturelle. On commence par supposer que  $\mathcal{V}$  contient la matrice  $\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$ , et dans ce cas on procède comme dans le paragraphe 6.5.1 pour

voir que l'ensemble des matrices  $N$  de  $M_r(\mathbb{K})$  telles que  $\mathcal{V}$  contienne  $\begin{bmatrix} N & 0 \\ 0 & 0 \end{bmatrix}$

est un sous-espace affine de  $M_r(\mathbb{K})$  entièrement constitué de matrices inversibles et de dimension  $\frac{r(r-1)}{2}$ . On le note  $\mathcal{W}$  et on cherche alors à montrer que  $\mathcal{V} \sim i_{n,p}(\mathcal{W})$ , sachant que la direction  $V$  de  $\mathcal{V}$  contient, pour tous  $B, C, D$  respectivement dans  $M_{n-r,r}(\mathbb{K})$ ,  $M_{r,p-r}(\mathbb{K})$  et  $M_{n-r,p-r}(\mathbb{K})$ , une matrice de la forme

$$\begin{bmatrix} ? & C \\ B & D \end{bmatrix}.$$

Presque toute la difficulté est concentrée dans le cas particulier où  $p = r$  et  $n = r + 1$ , cas où l'on doit montrer qu'après une série de transvections-lignes  $L_i \leftarrow L_i + \alpha_i L_n$  (où  $1 \leq i \leq r$ ) on peut réduire  $\mathcal{V}$  à  $i_{n,p}(\mathcal{W})$ . Cet énoncé de base se démontre par récurrence sur  $r$ , grâce à des techniques d'extraction/projection des blocs. Et il est absolument nécessaire ici d'avoir une compréhension fine de la structure de l'espace  $\mathcal{W}$ , requérant un appel au théorème 6.6. La démonstration de ce simple cas court sur douze pages et nous n'en dirons rien de plus.

Le passage au cas général est ensuite plutôt facile. Par exemple, on peut remarquer que le sous-espace affine des matrices de  $\mathcal{W}$  dont toutes les colonnes sont nulles à compter de la  $(r+1)$ -ième et toutes les lignes nulles à compter de la  $(r+2)$ -ième (on notera que cet ensemble est bien non vide) s'identifie à un espace vérifiant les hypothèses du cas particulier précédent. On peut faire de même en faisant varier les lignes nulles imposées (tant qu'on ne prend pas l'une des  $r$  premières), et faire aussi le même travail sur les colonnes. L'application du cas particulier montre qu'après des séries de transvections du type  $L_i \leftarrow L_i + \alpha L_j$  avec  $i \leq r < j \leq n$ , et  $C_i \leftarrow C_i + \alpha C_j$  avec  $i \leq r < j \leq p$ , on se réduit à la situation où, en plus des hypothèses de base, la direction  $V$  contient *toutes* les matrices de la forme

$$\begin{bmatrix} [0]_{r \times r} & [?]_{r \times (p-r)} \\ [?]_{(n-r) \times r} & [0]_{(n-r) \times (p-r)} \end{bmatrix}.$$

À partir de là, il n'est pas très difficile de conclure que  $V$  contient aussi toute matrice de la forme

$$\begin{bmatrix} [0]_{r \times r} & [0]_{r \times (p-r)} \\ [0]_{(n-r) \times r} & [?]_{(n-r) \times (p-r)} \end{bmatrix},$$

d'où la conclusion découle immédiatement ; comme la démonstration du lemme qui permet de conclure le dernier point non trivial (le lemme 7 de [2]) est entachée d'une petite erreur, nous allons la rectifier. Ce lemme consiste à étudier la situation où  $n = p = r + 1$  et à montrer dans ce cas que la direction  $V$  contient la matrice élémentaire  $E_{n,p}$ . Partons de l'observation que  $V$  contient une matrice de la forme  $A \oplus 1$  avec  $A \in M_r(\mathbb{K})$ . Si  $A$  est dans la direction  $W$  de  $\mathcal{W}$ , la conclusion est immédiate. Sinon, par maximalité de la dimension, le sous-espace affine  $\mathcal{W} + \mathbb{K}A$  contient une matrice singulière. L'erreur de la démonstration du lemme est de considérer que cette matrice doit être  $M + A$  pour un  $M \in \mathcal{W}$ . En réalité, elle est seulement de la forme  $B = M + \alpha A$  pour un  $\alpha \in \mathbb{K} \setminus \{0\}$  et un  $M \in \mathcal{W}$ , mais cette erreur est facilement rectifiable. Tout simplement, la matrice blocs  $\begin{bmatrix} B & C \\ L & \alpha \end{bmatrix}$  doit appartenir à  $\mathcal{V}$  pour tout  $L \in M_{1,r}(\mathbb{K})$  et tout  $C \in \mathbb{K}^r$ , donc elle doit être systématiquement de rang au moins  $r$ . Mais puisque  $\text{rg } B < r$  il est facile d'ajuster  $L$  et  $C$  pour que ce ne soit pas le cas : une fois ramené au cas où  $B = I_{r-1} \oplus 0$ , il suffit en effet de prendre  $L = [\alpha \ 0 \ \cdots \ 0]$

et  $C = [1 \ 0 \ \dots \ 0]^T$ . On notera enfin que c'est là qu'intervient de manière critique l'hypothèse voulant que  $r > 1$ .

### 6.5.6 Sous-espaces nilpotents : le cas d'un corps gauche

Un an après avoir conclu mes travaux sur les sous-espaces à spectre trivial et leur application aux sous-espaces affines de rang minoré, puis avoir étudié les sous-espaces de matrices ayant une seule valeur propre (j'y reviendrai dans la section 6.5.7), j'ai perçu une difficulté d'ordre pédagogique : la seule démonstration publiée alors dans des revues internationales à comité de lecture et mettant en scène la méthode de compatibilité diagonale était celle du théorème 6.4, et elle était tellement ardue qu'il était difficile de s'en servir comme point d'appui pour des développements ultérieurs.

J'ai donc entrepris, avant de poursuivre sur la question des sous-espaces vectoriels à spectre limité, de mettre en œuvre cette technique sur un cas simple généralisant de manière substantielle le théorème de Gerstenhaber. Il s'agissait d'examiner le cas des matrices à coefficients dans un corps gauche. Une difficulté immédiate dans ce cadre est de donner un sens à la notion de sous-espace vectoriel. J'ai donc choisi de me limiter au cas où le corps gauche considéré  $\mathbb{K}$  est de degré fini sur son centre, et où les scalaires considérés pour les combinaisons linéaires varient dans un sous-corps  $\mathbb{K}_0$  du centre de  $\mathbb{K}$  sur lequel  $\mathbb{K}$  est de degré fini. On dispose alors exactement du théorème attendu, sans restriction aucune :

**Théorème 6.8** (Théorème 2 de [5]). *Soit  $\mathbb{K}$  un corps gauche, et  $\mathbb{K}_0$  un sous-corps du centre de  $\mathbb{K}$  sur lequel  $\mathbb{K}$  est de degré fini  $d$ . Soit  $\mathcal{V}$  un sous- $\mathbb{K}_0$ -espace vectoriel nilpotent de  $M_n(\mathbb{K})$ . Alors  $\dim_{\mathbb{K}_0} \mathcal{V} \leq d \frac{n(n-1)}{2}$ , et en cas d'égalité  $\mathcal{V}$  est conjugué à  $T_n^{++}(\mathbb{K})$ .*

Le corollaire immédiat suivant, sur les corps finis, est particulièrement intéressant. On y prend pour  $\mathbb{K}_0$  le sous-corps premier de  $\mathbb{K}$ .

**Corollaire 6.9.** *Soit  $\mathbb{K}$  un corps fini de cardinal  $q$ . Soit  $\mathcal{V}$  un sous-groupe de  $(M_n(\mathbb{K}), +)$  dans lequel tout élément est nilpotent. Alors  $|\mathcal{V}| \leq q^{n(n-1)/2}$ , et en cas d'égalité  $\mathcal{V}$  est conjugué à  $T_n^{++}(\mathbb{K})$ .*

La stratégie de démonstration du théorème 6.8 est encore celle du vecteur adapté puis de la compatibilité diagonale. Ici, il y a tout de même bon nombre de difficultés techniques non triviales : impossibilité d'utiliser les méthodes de trace, absence de commutativité, structures de  $\mathbb{K}_0$ -espace vectoriel. La démonstration utilise notamment une bonne partie des idées d'intransitivité exploitées dans [4].

### 6.5.7 Sous-espaces à spectre de cardinal limité

Je vais terminer cette section en évoquant deux articles complétant les travaux sur les sous-espaces nilpotents et les sous-espaces à spectre trivial, avec des limitations légèrement moins drastiques sur le spectre.

Les auteurs qui m'avaient précédé sur la question étaient Omladič et Šemrl [Oml96] et Loewy et Radwan [Loe98]. Précisons un peu les résultats qu'ils avaient obtenus et qui préfigurent les miens. Dans [Oml96], les auteurs étudient les sous-espaces vectoriels de  $M_n(\mathbb{C})$  dans lesquels toute matrice possède au plus



$k$  valeurs propres, où  $k \in \{1, 2\}$ . Le cas  $k = 1$  se déduit trivialement de la situation des sous-espaces vectoriels nilpotents : en effet, lorsque  $M \in M_n(\mathbb{C})$  a une seule valeur propre, cette valeur propre est  $\frac{\text{tr } M}{n}$ , et on retrouve donc un sous-espace vectoriel nilpotent en prenant le noyau de la trace. Les espaces cherchés sont donc inclus dans les sommes des sous-espaces nilpotents avec la droite  $\mathbb{C}I_n$ . Le cas  $k = 2$  est beaucoup moins évident. Pour celui-ci, les auteurs se limitent très opportunément au cas où  $n$  est impair (ce qui force les multiplicités génériques des valeurs propres à être distinctes) et utilisent des arguments élémentaires pour établir que le sous-espace de matrices qu'ils considèrent possède la propriété L de Motzkin et Taussky, autrement dit que les valeurs propres peuvent s'écrire comme des fonctions linéaires de la matrice considérée. La réduction au théorème de Gerstenhaber est alors facile.

Dans [Loe98], Loewy et Radwan sont allés plus loin pour la majoration de la dimension. Ils ont démontré que sur un corps de caractéristique nulle, et pour tout  $k \in \{1, 2, 3\}$  avec  $k < n$ , tout sous-espace vectoriel de  $M_n(\mathbb{K})$  dont les éléments ont au plus  $k$  valeurs propres dans  $\overline{\mathbb{K}}$  est de dimension au plus  $\frac{n(n-1)}{2} + k$  si  $k \leq 2$ , et  $\frac{n(n-1)}{2} + 4$  si  $k = 3$ . Leur méthode est fondée sur l'analyse du polynôme caractéristique générique dans l'espace envisagé et de sa factorisation. Ils formulent enfin dans leur article une conjecture sur la généralisation à un  $k \in \llbracket 3, n-1 \rrbracket$  quelconque : pour tout corps  $\mathbb{K}$  de caractéristique 0, tout sous-espace vectoriel de  $M_n(\mathbb{K})$  dans lequel toute matrice admet au plus  $k$  valeurs propres dans  $\overline{\mathbb{K}}$  serait de dimension majorée par  $\binom{n}{2} + \binom{k}{2} + 1$ . Cette dimension limite est évidemment atteinte pour le sous-espace  $(\mathbb{K}I_{n-k+1} \oplus T_{n-k+1}^{++}(\mathbb{K})) \vee M_{k-1}(\mathbb{K})$ . Lorsque nous écrivions ces lignes, nous avons découvert que cette conjecture venait d'être résolue sur  $\mathbb{C}$  par Omladić et Šivic<sup>6</sup> grâce à des méthodes de groupes de Lie (le passage à un corps quelconque algébriquement clos s'obtient alors par la technique usuelle de transfert algébrique, puis le passage à un corps de caractéristique nulle par une simple extension du corps des scalaires).

Mon travail sur la question a été de deux ordres. D'abord, j'ai étudié le problème connexe où c'est non le spectre dans  $\overline{\mathbb{K}}$  mais celui dans  $\mathbb{K}$  qui est considéré. Et ensuite j'ai prolongé une partie des théorèmes de Loewy et Radwan à un corps (presque) quelconque. J'ai ainsi défini deux types d'espaces, un entier  $k \geq 0$  étant donné :

- les espaces  $k$ -spec sont ceux dans lesquels toute matrice a au plus  $k$  valeurs propres distinctes dans  $\mathbb{K}$  ;
- les espaces  $\overline{k}$ -spec sont ceux dans lesquels toute matrice a au plus  $k$  valeurs propres distinctes dans  $\overline{\mathbb{K}}$ .

Ainsi, les espaces  $\overline{0}$ -spec sont les espaces nilpotents, tandis que les espaces 0-spec sont les espaces à spectre trivial.

Dans [3] puis [6], j'ai obtenu la dimension maximale d'un sous-espace  $k$ -spec de  $M_n(\mathbb{K})$  pour  $k \leq 2$ , hormis dans le cas  $\chi(\mathbb{K}) = k = 2$  (et c'est la seule restriction sur le corps de base). Le résultat prolonge parfaitement le théorème de Loewy et Radwan, à l'exception du cas  $n = 2 = \chi(\mathbb{K})$ , où la dimension maximale d'un espace  $\overline{1}$ -spec est 3 et non pas 2 (la seule solution de dimension maximale étant  $\mathfrak{sl}_2(\mathbb{K})$ ).

J'ai aussi entièrement élucidé la structure des espaces  $\overline{1}$ -spec sur un corps

---

6. M. Omladić, K. Šivic, The solution of the Loewy-Radwan conjecture, prépublication ArXiv, septembre 2022, <https://arxiv.org/abs/2209.09416>

quelconque [3] puis celle des espaces  $\bar{2}$ -spec sur un corps quelconque de caractéristique différente de 2 (voir [6]). La stratégie générale d'étude est essentiellement celle qui a été décrite dans les paragraphes antérieurs : recherche d'un vecteur adapté, puis récurrence sur la taille des matrices avec double utilisation de l'hypothèse de récurrence sur les espaces haut et bas, et compatibilité diagonale. Une observation immédiate est que lorsqu'on considère l'espace haut (ou bas) dans cette méthode à partir d'un espace  $k$ -spec (respectivement  $\bar{k}$ -spec), on obtient mieux qu'un espace  $k$ -spec car les endomorphismes ayant servi à fabriquer l'espace haut (ou bas) ont déjà 0 dans leur spectre. On obtient donc une propriété intermédiaire qu'il a été nécessaire d'étudier pour elle-même, à savoir d'avoir au plus  $k$  valeurs propres *non nulles* (dans  $\mathbb{K}$  ou  $\bar{\mathbb{K}}$ , selon le contexte). On parle alors d'espace  $k^*$ -spec (respectivement,  $\bar{k}^*$ -spec). Dans [6], j'ai mené une étude complète des sous-espaces  $\bar{1}^*$ -spec de dimension maximale lorsque le corps de base n'est pas de caractéristique 2, comme étape préalable à la détermination des sous-espaces  $\bar{2}$ -spec de dimension maximale (sous la même hypothèse sur le corps de base).

Évoquons maintenant plus précisément les résultats que j'ai obtenus. D'abord sur les majorations de dimension :

**Théorème 6.10** (Théorème 1.3 de [3]). *Soit  $\mathcal{V}$  un sous-espace vectoriel 1-spec de  $M_n(\mathbb{K})$ , avec  $n \geq 2$ . Alors  $\dim \mathcal{V} \leq 1 + \frac{n(n-1)}{2}$  sauf si  $n = \chi(\mathbb{K}) = 2$ .*

Ce dernier résultat est même étendu aux espaces  $1^*$ -spec dans [6] (en mettant totalement de côté le cas  $n = 2$ , bien évidemment).

**Théorème 6.11** (Théorème 1.4 de [6]). *Soit  $\mathcal{V}$  un sous-espace vectoriel 2-spec de  $M_n(\mathbb{K})$ , avec  $n \geq 3$ . Alors  $\dim \mathcal{V} \leq 2 + \frac{n(n-1)}{2}$  si  $\chi(\mathbb{K}) \neq 2$ .*

En outre, toutes ces majorations sont optimales, y compris pour les espaces  $\bar{1}$ -spec et  $\bar{2}$ -spec.

Quant à la structure des sous-espaces réalisant la dimension critique, je me suis limité – comme déjà indiqué – aux sous-espaces  $\bar{k}$ -spec avec  $k \in \{1, 2\}$ . La difficulté pour obtenir la structure des espaces 0-spec ainsi que la complexité du résultat – voir le théorème 6.4 – sont telles qu'on imagine difficilement aller beaucoup plus loin dans la voie de l'étude des espaces  $k$ -spec.

Pour ce qui est des espaces  $\bar{1}$ -spec, on s'attendrait naturellement à ce que ceux de dimension maximale soient tous conjugués à  $\mathbb{K}I_n \oplus T_n^{++}(\mathbb{K})$ . Mais c'est faux, il y a des difficultés liées à la caractéristique. Bien sûr, et avec l'argument d'Omladič et Šemrl, c'est trivialement vrai (au sens d'être une conséquence immédiate du théorème de Gerstenhaber !) lorsque  $\chi(\mathbb{K})$  n'est pas un diviseur de  $n$ , mais sinon on ne voit pas bien comment établir que les matrices nilpotentes dans un sous-espace  $\bar{1}$ -spec forment un sous-espace vectoriel. Il a donc été nécessaire de revenir à ma méthode de démonstration du théorème de Gerstenhaber, ce qui m'a permis de mettre en évidence plusieurs contre-exemples remarquables. Si  $\mathbb{K}$  est de caractéristique 2, la matrice

$$\begin{bmatrix} 0 & l_1 & l_2 & x \\ c_2 & 0 & y & c_1 \\ c_1 & x & 0 & c_2 \\ y & l_2 & l_1 & 0 \end{bmatrix}$$

a pour polynôme caractéristique  $X^4 + a$  pour  $a = ((l_1 + l_2)(c_1 + c_2) + xy)^2 + l_1c_2(x+y)^2$ . En rajoutant la droite  $\mathbb{K}I_4$ , on construit ainsi un sous-espace  $\bar{1}$ -spec

de  $M_4(\mathbb{K})$  ayant la dimension critique 7. Et l'on peut démontrer que tout sous-espace  $\bar{1}$ -spec de  $M_4(\mathbb{K})$  et de dimension 7 est conjugué à l'espace ainsi créé ou à  $\mathbb{K}I_4 + T_4^{++}(\mathbb{K})$  (théorème 1.7 de [3]).

En dimension et caractéristique 3, on trouve deux grandes familles de contre-exemples, définies comme suit. Pour  $\delta \in \mathbb{K}$ , on note

$$\mathcal{F}_\delta := \text{Vect} \left( I_3, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \delta & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ -1 & -\delta & -1 \end{bmatrix} \right)$$

et

$$\mathcal{G}_\delta := \text{Vect} \left( I_3, \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ \delta & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -\delta & 0 \end{bmatrix} \right).$$

Il s'agit de sous-espaces de dimension 4, et on démontre que chacun d'entre eux est  $\bar{1}$ -spec. En outre, aucun d'entre eux n'est conjugué à  $\mathbb{K}I_3 + T_3^{++}(\mathbb{K})$ , et aucun espace de type  $\mathcal{F}_\delta$  n'est conjugué à un espace de type  $\mathcal{G}_\sigma$ . Enfin on dispose de conditions simples pour déterminer si deux espaces de type  $\mathcal{F}_\delta$  (respectivement  $\mathcal{G}_\delta$ ) sont conjugués :

- $\mathcal{F}_\delta$  et  $\mathcal{F}_\lambda$  sont conjugués si et seulement si  $\delta - \lambda$  appartient à l'image de l'homomorphisme de groupe (additif)  $x \in \mathbb{K} \mapsto x^3 - x$  ;
- $\mathcal{G}_\delta$  et  $\mathcal{G}_\lambda$  sont conjugués si et seulement s'il existe  $(a, b) \in \mathbb{K}^2$  tel que  $a \neq 0$  et  $\delta = a^3\lambda + b^3$ .

Et l'on démontre, toujours dans le cas où  $\chi(\mathbb{K}) = 3$ , qu'ont ainsi été mis en évidence à conjugaison près tous les sous-espaces  $\bar{1}$ -spec de  $M_3(\mathbb{K})$  de dimension 4 (théorème 4.6 de [3]). Les espaces précédents ont été obtenus par le calcul, et il serait fort intéressant de voir s'ils n'ont pas de lien avec des phénomènes connus en géométrie algébrique. Jusque-là, nous n'avons hélas rien trouvé dans cette direction.

Évoquons enfin les principaux résultats obtenus dans [6] sur la structure des espaces  $\bar{2}$ -spec de dimension maximale. Cette situation est très particulière au regard de la conjecture de Loewy-Radwan : en effet, le résultat obtenu très récemment par Omladič et Šivic est que, sur le corps des nombres complexes, et pour tout  $k \in \llbracket 3, n-1 \rrbracket$ , tout espace  $\bar{k}$ -spec ayant la dimension maximale

$\binom{n}{2} + \binom{k}{2} + 1$  est conjugué à  $\mathbb{C}I_n \oplus (T_p^{++}(\mathbb{C}) \vee M_{k-1}(\mathbb{C}) \vee T_{n-k+1-p}^{++}(\mathbb{C}))$  pour un certain  $p \in \llbracket 0, n-k+1 \rrbracket$ . Lorsque  $k = 2$ , on a une variété sensiblement plus grande de solutions : pour tout sous-ensemble non trivial  $I$  de  $\llbracket 1, n \rrbracket$ , on peut considérer l'ensemble  $\mathcal{D}_I$  des matrices triangulaires supérieures de  $M_n(\mathbb{K})$  ayant tous leurs coefficients diagonaux égaux pour les indices dans  $I$ , et tous leurs coefficients diagonaux égaux pour les indices dans  $\llbracket 1, n \rrbracket \setminus I$ . C'est un sous-espace  $\bar{2}$ -spec et il a la dimension maximale si  $n > 2$ . Mais il existe d'autres solutions étranges, dont une avait été citée dans [Oml96] : l'espace des matrices de  $M_4(\mathbb{K})$

de la forme  $\begin{bmatrix} A & [?] \\ 0 & A \end{bmatrix}$  avec  $A \in M_2(\mathbb{K})$  est aussi  $\bar{2}$ -spec et de dimension optimale. Pire encore, on peut aussi considérer l'espace des matrices de  $M_4(\mathbb{K})$  de la forme  $\begin{bmatrix} A & [?] \\ 0 & A^T \end{bmatrix}$  avec  $A \in M_2(\mathbb{K})$  ; et l'on peut démontrer que ce sous-espace n'est pas conjugué au précédent ! Pour les corps de caractéristique différente de 2 et 3, on peut établir que les espaces cités précédemment sont les seuls sous-espaces

$\overline{2}$ -spec de dimension maximale à conjugaison près. En caractéristique 3, d'autres espaces font leur apparition, liés aux sous-espaces  $\overline{1}$ -spec de  $M_3(\mathbb{K})$  dont nous avons parlé plus tôt. On pourra donc voir apparaître des espaces de la forme  $\mathcal{F}_\delta \vee (\mathbb{K}I_{n-3} + T_{n-3}^{++}(\mathbb{K}))$ , mais plus généralement encore  $\mathbb{K}I_n \oplus (T_i^{++}(\mathbb{K}) \vee \mathcal{F}_\delta \vee T_{n-i-3}^{++}(\mathbb{K}))$  avec  $i \in \llbracket 0, n-3 \rrbracket$ , mais aussi  $\mathcal{F}_\delta \vee \mathcal{F}_{\delta'}$  dans le cas  $n = 6$  etc.

Au vu de ce qui précède, on comprendra que, malgré un réel effort de synthèse, [6] s'étale sur plus de cent pages, dont une bonne moitié est consacrée au seul cas de la caractéristique 3 !

On a déjà évoqué le fait que la méthode générale pour obtenir ces résultats est une adaptation des techniques vues antérieurement avec de nombreuses difficultés techniques qu'il ne faut pas sous-estimer. J'évoquerai simplement la partie cruciale de mise en place de la démonstration, qui repose une nouvelle fois sur la recherche d'un vecteur adapté. Cette dernière est en effet nettement plus délicate pour les sous-espaces 1-spec ou 2-spec que pour les espaces 0-spec, et constitue pour moi la partie la plus intéressante de [6]. Alors que dans la situation vue dans le paragraphe 6.5.2 le raisonnement était essentiellement combinatoire, il est nécessaire ici d'aller plus loin et de rentrer plus pleinement dans la géométrie. À nouveau, on procède par récurrence et on démontre le lemme suivant :

**Lemme 6.12.** *Soit  $n \geq 2$ .*

- (a) *Tout sous-espace 1-spec de  $M_n(\mathbb{K})$  possède un vecteur adapté sauf si  $\chi(\mathbb{K}) = n = 2$ .*
- (b) *Si  $\chi(\mathbb{K}) \neq 2$  et  $n \geq 3$ , alors tout sous-espace 2-spec de  $M_n(\mathbb{K})$  possède un vecteur adapté.*

Ici, l'équivalent du lemme 6.2 tombe en défaut. Par exemple, si  $\chi(\mathbb{K})$  divise  $n$  alors  $\text{Vect}(E_{1,2}, E_{2,3}, \dots, E_{n-1,n}, E_{n,1})$  est un sous-espace  $\overline{1}$ -spec de  $M_n(\mathbb{K})$  mais aucun vecteur de la base canonique ne lui est adapté !

Je vais maintenant expliquer en détail l'adaptation aux sous-espaces 1-spec, puis plus succinctement celle aux sous-espaces 2-spec. Soit donc  $\mathcal{V}$  un sous-espace 1-spec de  $M_n(\mathbb{K})$ , supposé sans vecteur adapté. Avec les notations de la démonstration du lemme 6.2, le sous-espace  $K(\mathcal{V})$  est  $0^*$ -spec, autrement dit à spectre trivial. On peut donc appliquer le lemme 6.2 pour en trouver un vecteur adapté. Mais cela ne suffira pas de trouver seulement  $E_{i,n}$  dans  $\mathcal{V}$  pour un certain  $i \in \llbracket 1, n-1 \rrbracket$ . On peut être beaucoup plus malin que cela en jouant sur la géométrie : le lemme 6.2 ne montre en effet pas seulement que tout sous-espace  $\overline{0}$ -spec  $\mathcal{W}$  de  $M_{n-1}(\mathbb{K})$  possède un vecteur adapté, il montre que n'importe quelle base de  $\mathbb{K}^{n-1}$  contient au moins un vecteur  $\mathcal{W}$ -adapté, autrement dit que l'ensemble des vecteurs inadaptés à  $\mathcal{W}$  est contraint dans un hyperplan vectoriel de  $\mathbb{K}^{n-1}$ , et en particulier  $\mathbb{K}^{n-1}$  possède une base de vecteurs adaptés à  $\mathcal{W}$  ! En s'appuyant sur cette observation, on parvient après réduction à montrer que  $\mathcal{V}$  contient  $E_{i,n}$  pour *tout*  $i \in \llbracket 2, n \rrbracket$ . Et en généralisant on trouve ainsi que  $\mathcal{V}$  contient toutes les matrices nilpotentes de rang 1, puis  $\mathfrak{sl}_n(\mathbb{K}) \subset \mathcal{V}$ , ce qui aboutit à une contradiction.

Pour les espaces 2-spec, il faut raffiner encore l'approche : en s'inspirant du cas 1-spec, on fait donc des hypothèses sur la géométrie de l'ensemble des vecteurs adaptés. Voici le lemme pertinent :

**Lemme 6.13** (Lemme 2.2 de [6]). *On suppose  $\chi(\mathbb{K}) \neq 2$ . Soit  $E$  un  $\mathbb{K}$ -espace vectoriel de dimension  $n$ , et  $\mathcal{V}$  un sous-espace vectoriel 2-spec de  $\mathcal{L}(E)$ . Si  $n = 2$ ,*

on suppose en outre que  $\mathcal{V}$  est  $1^*$ -spec. L'ensemble des vecteurs  $\mathcal{V}$ -adaptés n'est alors inclus dans aucun ensemble de la forme  $E_1 \cup E_2 \cup \dots \cup E_n$ , où  $E_i$  est un sous-espace vectoriel de dimension  $\lfloor \frac{i+1}{2} \rfloor$  pour tout  $i \in \llbracket 1, n \rrbracket$ .

Ce lemme se démontre par récurrence en s'inspirant de la démonstration du lemme 6.2 : l'hypothèse de récurrence est exploitée pour fabriquer, lorsque  $\mathcal{V}$  ne vérifie pas la conclusion du lemme précédent, une grande quantité de matrices de rang 1 dans  $\mathcal{V}$ , et plus globalement tout un sous-espace totalement singulier de  $\mathcal{V}$  pour la forme bilinéaire  $(A, B) \mapsto \text{tr}(AB)$ , sous-espace de dimension excessive.

Notons enfin que, pour (b), l'hypothèse de caractéristique est indispensable, même si  $|\mathbb{K}| > 2$ . En caractéristique 2, l'espace  $(\mathbb{K}I_{n-2} + T_{n-2}^{++}(\mathbb{K})) \vee \mathfrak{sl}_2(\mathbb{K})$  est en effet  $\bar{2}$ -spec mais n'a aucun vecteur adapté ! Notre stratégie de base semble donc fortement compromise pour étudier ce cas, qui reste entièrement en suspens.

## 6.6 Lien avec les sous-espaces primitifs de matrices de rang majoré

L'observation dont il va s'agir dans cette courte section a été faite à l'automne 2012. Je travaillais à l'époque sur les espaces localement liés (voir le chapitre 3) et utilisait alors constamment des techniques fondées sur la dualité opérateur-vecteur. Des manipulations de cette dualité faisaient apparaître les questions d'intransitivité que j'avais rencontrées dans mes travaux sur les sous-espaces à spectre trivial. J'ai creusé la question et ai fait la découverte suivante : sur un corps de cardinal assez grand, le théorème 6.4 peut-être déduit à très peu de frais du théorème d'Atkinson (théorème 2.15) sur les sous-espaces semi-primitifs de matrices de rang majoré !

L'argument est tellement court que je vais le reproduire quasiment *in extenso*. Précisons que je vais ici donner l'idée sous une forme largement géométrique et que je donnerai la version transposée de l'argument donné dans [7]. Précisons aussi que l'on procède par récurrence sur la taille des matrices pour démontrer à la fois l'inégalité sur la dimension et régler le cas d'égalité. On prend un sous-espace  $\mathcal{V}$  à spectre trivial de  $\mathcal{L}(E)$ , pour un  $\mathbb{K}$ -espace vectoriel  $E$  de dimension  $n \geq 2$ , et l'on peut d'emblée supposer  $\mathcal{V}$  irréductible. On introduit alors l'espace dual-opérateur  $\widehat{\mathcal{V}} \subset \mathcal{L}(\mathcal{V}, E)$  constitué des opérateurs

$$\widehat{x} : u \in \mathcal{V} \mapsto u(x) \in E, \quad \text{où } x \in E.$$

La propriété d'intransitivité fait qu'aucun des éléments de  $\widehat{\mathcal{V}}$  n'est surjectif. En outre  $\widehat{\mathcal{V}}$  est réduit (voir le paragraphe 3.4) : en effet d'une part il est toujours vrai qu'un dual-opérateur est réduit au départ, et ensuite s'il existait un hyperplan  $H$  de  $E$  incluant l'image de tout élément de  $\widehat{\mathcal{V}}$  alors il inclurait l'image de tout élément de  $\mathcal{V}$  et serait en particulier stable par ceux-ci, contredisant l'hypothèse d'irréductibilité.

Maintenant, supposons très temporairement que  $\widehat{\mathcal{V}}$  soit de transposé semi-primitif. Le théorème d'Atkinson montre alors qu'il est équivalent à l'espace des opérateurs  $\{b \in \text{Alt}(E^2, \mathbb{R}) \mapsto b(x, -) \mid x \in E\}$ , où  $\text{Alt}(E^2, \mathbb{K})$  désigne l'espace des formes bilinéaires alternées sur  $E$ . En remontant l'équivalence, on obtient facilement une forme bilinéaire non dégénérée  $\omega$  sur  $E^2$  (a priori ni symétrique ni antisymétrique) telle que  $\mathcal{V}$  soit l'ensemble  $\mathcal{A}_\omega$  des endomorphismes  $\omega$ -alternés

à droite. Ensuite, comme  $\mathcal{V}$  est à spectre trivial il est facile de voir que  $\omega$  est anisotrope. Dans ce cas, on a retrouvé directement le corollaire 6.5, qui décrit les sous-espaces à spectre trivial optimaux irréductibles.

La démonstration ne s'arrête pas là car il n'y a pas de raison *a priori* pour que  $\widehat{\mathcal{V}}$  soit de transposé semi-primitif. On suppose donc que ce n'est pas le cas, et on va chercher une contradiction. On introduit alors un sous-espace non nul  $F$  de  $E$ , de dimension maximale et tel que, pour la projection canonique  $\pi : E \rightarrow E/F$ , l'espace  $\widehat{\mathcal{V}} \bmod F := \{\pi \circ f \mid f \in \widehat{\mathcal{V}}\}$  reste constitué d'applications linéaires non surjectives. Puis on prend l'intersection  $\mathcal{W}$  des noyaux des éléments de  $\widehat{\mathcal{V}} \bmod F$ , dont on note  $c$  la dimension, et on prend un supplémentaire  $\mathcal{V}'$  de  $\mathcal{W}$  dans  $\mathcal{V}$ . L'espace  $\mathcal{S}$  des opérateurs  $\pi \circ \widehat{x}|_{\mathcal{V}'}$ , où  $x \in E$ , est alors de transposé semi-primitif. On déduit alors du théorème d'Atkinson que si l'on note  $r := \dim(E/F)$ , alors  $\dim \mathcal{V}' \leq \frac{r(r-1)}{2}$  puis  $\dim \mathcal{V} \leq c + \frac{r(r-1)}{2}$ .

Interprétons maintenant la situation : l'espace  $\mathcal{W}$  est formé des éléments de  $\mathcal{V}$  qui ont leur image dans  $F$ . Leurs restrictions à  $F$  forment un sous-espace vectoriel  $\mathcal{W}_F$  de  $\mathcal{L}(F)$  à spectre trivial. En appliquant l'hypothèse de récurrence, il vient  $\dim \mathcal{W} \leq \frac{s(s-1)}{2} + rs$  pour  $s := \dim F = n - r$ . Ainsi

$$\dim \mathcal{V} = \dim \mathcal{W} + \dim \mathcal{V}' \leq \frac{s(s-1)}{2} + rs + \frac{r(r-1)}{2} = \frac{n(n-1)}{2}.$$

Supposons maintenant l'égalité réalisée. On apprend alors en particulier que  $\dim \mathcal{W} = \frac{s(s-1)}{2} + rs$ , et les observations ayant servi à obtenir l'inégalité garantissent que  $\mathcal{W}_F$  est optimal et que  $\mathcal{V}$  contient toute application linéaire nulle sur  $F$  et à valeurs dans  $F$ . C'est ce dernier point qui va emporter facilement le morceau : nous allons montrer que  $F$  est stable par tous les éléments de  $\mathcal{V}$ , ce qui contredira l'hypothèse d'irréductibilité.

Représentons matriciellement la situation dans une base adaptée à  $F$  (c'est-à-dire dont les premiers vecteurs engendrent  $F$ ). Les éléments de  $\mathcal{W}$  sont représentés par les matrices de la forme

$$\begin{bmatrix} A & [?]_{s \times r} \\ [0]_{r \times s} & [0]_{r \times r} \end{bmatrix}$$

où  $A$  vit dans un sous-espace vectoriel  $\mathcal{K}$  de  $M_s(\mathbb{K})$  à spectre trivial et optimal. En utilisant cette optimalité finement, on démontre alors que  $\mathcal{K}$  est exactement l'espace des blocs supérieurs gauches des matrices représentant les opérateurs de  $\mathcal{V}$  dans la base envisagée, *et c'est vrai quel que soit le choix des  $r$  derniers vecteurs de base*. Géométriquement, cela signifie qu'on a trouvé un sous-espace  $\mathcal{V}_0$  à spectre trivial (optimal) de  $\mathcal{L}(F)$  tel que, quelle que soit la projection  $\pi'$  de  $E$  sur  $F$  choisie, tous les opérateurs  $\pi' \circ u|_F$ , pour  $u \in \mathcal{V}$ , figurent dans  $\mathcal{V}_0$  et sont donc sans valeur propre non nulle dans  $\mathbb{K}$ .

Supposons pour finir qu'un certain opérateur  $u \in \mathcal{V}$  ne stabilise pas  $F$ . On trouverait donc un  $x \in F \setminus \{0\}$  tel que  $u(x) \notin F$ , et on pourrait alors trouver un projecteur  $\pi'$  de  $E$  sur  $F$  envoyant  $u(x)$  sur  $x$ . Alors  $(\pi' \circ u)(x) = x$ , contredisant le fait que  $\pi' \circ u|_F$  est censé être à spectre trivial. Cette contradiction montre que  $F$  est stable par tout élément de  $\mathcal{V}$ , ce qui contredit l'irréductibilité de  $\mathcal{V}$ .

Au vu de ce qui précède, on serait tenté de dire que les longues pages de la démonstration du théorème 6.4 peuvent être maintenant jetées aux orties, mais ce serait aller bien vite en besogne. Ce qui précède constitue certes une

démonstration beaucoup plus directe, mais elle ne vaut que sous l'hypothèse  $|\mathbb{K}| \geq n$ . Or on ne dispose d'aucune version du théorème d'Atkinson sans cette hypothèse de cardinalité. L'observation en petites dimensions permet certes de conjecturer que le nombre maximal  $\frac{r(r+1)}{2}$  de lignes pour un sous-espace primitif de rang au plus  $r$  reste vrai pour un corps quelconque, mais la caractérisation des cas d'égalité ne semble plus tenir.

Enfin, le potentiel de l'argument qui précède n'a probablement été qu'effleuré : je donnerai quelques indications à ce propos dans la section 6.8. Pour l'instant, en dehors de ce dont je viens de discuter, la seule autre utilisation a été, dans le même contexte, faite dans [8] pour classer à conjugaison près les sous-espaces optimaux à spectre trivial de  $M_3(\mathbb{F}_2)$  (rappelons que le théorème 6.4 est totalement aveugle au cas de  $\mathbb{F}_2$ ), avec application aux sous-espaces affines de  $M_3(\mathbb{F}_2)$  de dimension 3 et dont tous les éléments sont inversibles.

## 6.7 Le problème de Gerstenhaber structuré

### 6.7.1 Introduction et réduction au cas non dégénéré

La dernière série de travaux que j'ai réalisée sur les sous-espaces nilpotents a été initiée par un contre-exemple curieux présenté par Rachel Quinlan à la conférence ILAS 2016. Quinlan avait construit, sur n'importe quel corps  $\mathbb{K}$  de caractéristique 2 et pour n'importe quel entier  $m \geq 1$ , un sous-espace nilpotent de  $A_{2m+1}(\mathbb{K})$  de dimension  $m^2 + m$ . Or, quelques années plus tôt, Meshulam et Radwan [Mes98] avaient démontré que la dimension maximale pour un sous-espace nilpotent de  $A_{2m+1}(\mathbb{C})$  était  $m^2$ , avec une démonstration permettant de généraliser ce résultat à n'importe quel corps algébriquement clos de caractéristique différente de 2. Le contre-exemple de Quinlan m'a profondément intrigué, d'autant plus que j'étais passé totalement à côté de [Mes98].

J'ai très vite compris que le bon cadre conceptuel pour aborder la question était celui des algèbres d'endomorphismes plutôt que des matrices. On se donne un espace vectoriel  $E$  de dimension finie muni d'une forme bilinéaire symétrique ou alternée  $b$ . Un endomorphisme  $u$  de  $E$  est dit :

- **$b$ -symétrique** lorsque  $(x, y) \mapsto b(x, u(y))$  est symétrique ; lorsque  $b$  est non dégénérée, cela signifie que  $u$  est  $b$ -autoadjoint si  $b$  est symétrique, et  $b$ -antiautoadjoint si  $b$  est alternée ;
- **$b$ -alterné** lorsque  $(x, y) \mapsto b(x, u(y))$  est alternée ; lorsque  $b$  est non dégénérée, cela signifie que  $u$  est  $b$ -antiautoadjoint si  $b$  est symétrique, et  $b$ -autoadjoint si  $b$  est alternée.

En caractéristique 2, tout endomorphisme  $b$ -alterné est  $b$ -symétrique. On note  $\mathcal{S}_b$  (respectivement,  $\mathcal{A}_b$ ) le sous-espace de  $\mathcal{L}(E)$  formé des endomorphismes  $b$ -symétriques (respectivement,  $b$ -alternés).

Les deux questions naturelles sont les suivantes. Elles forment le *problème de Gerstenhaber structuré* :

- Quelle est la dimension maximale possible pour un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$  (respectivement, de  $\mathcal{A}_b$ ) ?
- Quels sous-espaces nilpotents de  $\mathcal{S}_b$  (respectivement, de  $\mathcal{A}_b$ ) réalisent la dimension maximale ?

Le cas où  $b$  est non dégénérée et  $\mathbb{K}$  est algébriquement clos de caractéristique 0 était presque entièrement connu en 2016 : la première question est résolue

dans [Mes98] pour  $b$  symétrique, la deuxième dans [Dra06] pour  $\mathcal{S}_b$  lorsque  $b$  est alternée, et pour  $\mathcal{A}_b$  lorsque  $b$  est symétrique (autrement dit, c'est le cas des endomorphismes  $b$ -antiautoadjoints, qui forment l'algèbre de Lie du groupe des isométries linéaires de  $b$ ). La situation considérée par Quinlan est celle des sous-espaces nilpotents d'endomorphismes  $b$ -alternés pour la forme bilinéaire canonique  $(X, Y) \mapsto X^T Y$  sur  $\mathbb{K}^n$ .

Dans le cas  $b = 0$ , on a directement  $\mathcal{S}_b = \mathcal{L}(E)$  et l'on retrouve ainsi le problème de Gerstenhaber. On pourrait donc s'inquiéter du fait de considérer une généralisation aussi large de ce dernier. En réalité, le problème peut entièrement se réduire à deux sous-problèmes : le cas  $b = 0$  (Gerstenhaber) et le cas où  $b$  est non dégénérée. Pour le voir, il suffit d'introduire le noyau  $\text{Rad}(b) = \{x \in E : b(x, -) = 0\}$  de  $b$  et d'observer que tout endomorphisme  $b$ -symétrique ou  $b$ -alterné le stabilise. Ainsi, tout  $u \in \mathcal{S}_b \cap \mathcal{A}_b$  induit des endomorphismes  $u_{\text{Rad}(b)}$  et  $\bar{u}$  respectivement de  $\text{Rad}(b)$  et de  $E/\text{Rad}(b)$ , et la nilpotence de  $u$  équivaut à la nilpotence de  $u_{\text{Rad}(b)}$  et  $\bar{u}$ . Enfin  $b$  induit sur  $E/\text{Rad}(b)$  une forme bilinéaire  $\bar{b}$  non dégénérée de même parité que  $b$  (elle est symétrique si  $b$  est symétrique, et alternée si  $b$  est alternée), et  $\bar{u}$  est  $\bar{b}$ -symétrique (respectivement,  $\bar{b}$ -alternée) dès que  $u$  est  $b$ -symétrique (respectivement,  $b$ -alternée). On est donc entièrement réduit à résoudre le problème de Gerstenhaber dans  $\mathcal{L}(\text{Rad}(b))$ , et le problème de Gerstenhaber structuré dans  $\mathcal{L}(E/\text{Rad}(b))$  pour la forme non dégénérée  $\bar{b}$ .

On peut donc entièrement se limiter au cas où  $b$  est non dégénérée, hypothèse que nous ferons désormais.

## 6.7.2 Les principaux résultats

Il est essentiel d'observer que la solution du problème structuré dépend intimement de la forme  $b$  envisagée. Si  $b$  est anisotrope alors un endomorphisme  $b$ -symétrique ou  $b$ -alterné n'est nilpotent que s'il est nul : en effet, noyau et image d'un endomorphisme  $b$ -symétrique ou  $b$ -alterné doivent être orthogonaux donc d'intersection nulle par anisotropie de  $b$ . À l'inverse, si  $b$  est isotrope et symétrique il y a toujours des endomorphismes  $b$ -symétriques nilpotents non nuls : voir le cas de la forme hyperbolique standard sur  $\mathbb{K}^2$  et de l'endomorphisme  $(x, y) \mapsto (y, 0)$ , qui est bien symétrique pour cette forme.

Comme on peut s'en douter, la dimension maximale d'un sous-espace nilpotent de  $\mathcal{S}_b$  ou  $\mathcal{A}_b$  est fonction de l'indice de Witt  $\nu(b)$  de la forme bilinéaire  $b$  (dont on rappelle qu'il s'agit de la dimension maximale pour un sous-espace vectoriel totalement  $b$ -singulier). Et voici le premier résultat obtenu dans [9], qui règle complètement la question en caractéristique différente de 2 :

**Théorème 6.14** (Théorème 1.7 de [9]). *Soit  $E$  un espace vectoriel de dimension finie sur un corps  $\mathbb{K}$  de caractéristique différente de 2, et  $b$  une forme bilinéaire symétrique ou alternée non dégénérée sur  $E$ . On note  $\nu$  l'indice de Witt de  $b$ , et  $n$  la dimension de  $E$ . Alors :*

- (a) *La plus grande dimension possible pour un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$  est  $\nu(n - \nu)$ .*
- (b) *La plus grande dimension possible pour un sous-espace vectoriel nilpotent de  $\mathcal{A}_b$  est  $\nu(n - \nu - 1)$ .*

Lorsque  $\mathbb{K}$  est algébriquement clos, on a  $\nu = \lfloor n/2 \rfloor$  et on retrouve facilement les bornes obtenues par Meshulam et Radwan dans [Mes98]. En particulier, si



$n = 2m + 1$  la deuxième dimension trouvée est le  $m^2$  cité en introduction de cette partie.

Pour trouver des sous-espaces réalisant ces dimensions maximales, on peut penser géométriquement : on part d'un sous-espace totalement  $b$ -singulier maximal  $F$  de  $E$  et on en prend un drapeau complet  $\mathcal{F} := (F_0, \dots, F_\nu)$ . À ce drapeau, on associe l'espace  $\mathcal{N}_{\mathcal{F},b}$  des endomorphismes envoyant  $F^\perp$  dans  $F$  et envoyant  $F_i$  dans  $F_{i-1}$  pour tout  $i \in \llbracket 1, \nu \rrbracket$ .

Pour simplifier le discours, nous prendrons  $b$  symétrique et nous intéresserons uniquement au cas des endomorphismes  $b$ -symétriques. Matriciellement, on peut se représenter les choses comme suit : on prend un supplémentaire  $G$  de  $F^\perp$  dans  $E$  et on note  $H := (F + G)^\perp$ . On prend une base  $(e_1, \dots, e_\nu)$  de  $F$  adaptée au drapeau  $\mathcal{F}$ , on la complète en une base hyperbolique  $(e_1, \dots, e_\nu, f_1, \dots, f_\nu)$  de  $F \oplus G$ , et on choisit une base  $(g_1, \dots, g_p)$  de  $H$ . La matrice de Gram de la forme bilinéaire  $b$  dans la base  $\mathbf{B} := (e_1, \dots, e_\nu, g_1, \dots, g_p, f_1, \dots, f_\nu)$  ainsi constituée est de la forme

$$B := \begin{bmatrix} 0 & 0 & I_\nu \\ 0 & P & 0 \\ I_\nu & 0 & 0 \end{bmatrix},$$

où  $P$  est symétrique anisotrope. On peut remarquer que pour un endomorphisme  $u$  quelconque stabilisant  $F$  et envoyant  $F^\perp$  dans  $F$ , sa matrice dans  $\mathbf{B}$  s'écrit

$$M := \begin{bmatrix} A & R & E \\ 0 & [0]_{p \times p} & C \\ 0 & 0 & A' \end{bmatrix},$$

où  $A$  et  $A'$  sont dans  $M_\nu(\mathbb{K})$ . On note que

$$BM = \begin{bmatrix} 0 & 0 & A' \\ 0 & [0]_{p \times p} & PC \\ A & R & E \end{bmatrix},$$

donc  $u$  est  $b$ -symétrique si et seulement si  $A' = A^T$ ,  $R = (PC)^T$  et  $E$  est symétrique. Ainsi, les éléments de l'espace vectoriel  $\mathcal{S}_b \cap \mathcal{N}_{\mathcal{F},b}$  sont précisément les endomorphismes représentés dans  $\mathbf{B}$  par les matrices de la forme

$$\begin{bmatrix} A & (PC)^T & E \\ 0 & [0]_{p \times p} & C \\ 0 & 0 & A^T \end{bmatrix},$$

avec  $A \in T_\nu^{++}(\mathbb{K})$ ,  $C \in M_{n-2\nu,\nu}(\mathbb{K})$  et  $E \in S_\nu(\mathbb{K})$ . Il est alors facile de voir que  $\mathcal{S}_b \cap \mathcal{N}_{\mathcal{F},b}$  est nilpotent et de dimension  $\nu^2 + \nu(n - 2\nu) = \nu(n - \nu)$ .

Mieux encore, nous avons établi sur presque tous les corps que les espaces ainsi trouvés sont les seuls de dimension maximale :

**Théorème 6.15** (Théorèmes 1.8 et 1.9 de [9]). *On suppose  $\chi(\mathbb{K}) \neq 2$ . Soit  $b$  une forme bilinéaire symétrique (respectivement, alternée) non dégénérée de rang  $n$  et d'indice de Witt  $\nu$ . Soit  $\mathcal{V}$  un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$  (respectivement, de  $\mathcal{A}_b$ ) de dimension  $\nu(n - \nu)$  (respectivement,  $\nu(n - \nu - 1)$ ).*

*Il existe alors un sous-espace totalement  $b$ -singulier maximal  $F$  et un drapeau complet  $\mathcal{F}$  de  $F$  tel que  $\mathcal{V} = \mathcal{S}_b \cap \mathcal{N}_{\mathcal{F},b}$  (respectivement,  $\mathcal{V} = \mathcal{A}_b \cap \mathcal{N}_{\mathcal{F},b}$ ).*

Ce résultat vaut sans restriction sur le corps de base hormis celle sur la caractéristique, mais laisse de côté les endomorphismes  $b$ -antiautoadjoints. Pour ce cas, qui généralise les résultats de Draisma, Kraft et Kuttler [Dra06], nous avons dû faire des hypothèses de cardinalité du même ordre que celles intervenant dans le théorème original de Gerstenhaber :

**Théorème 6.16** (Théorèmes 1.5 et 1.6 de [10]). *On suppose  $\chi(\mathbb{K}) \neq 2$ . Soit  $b$  une forme bilinéaire symétrique (respectivement, alternée) non dégénérée de rang  $n$  et d'indice de Witt  $\nu$ . On suppose que  $|\mathbb{K}| \geq \min(n, 2\nu + 1)$  (respectivement,  $|\mathbb{K}| \geq n$ ). Soit  $\mathcal{V}$  un sous-espace vectoriel nilpotent de  $\mathcal{A}_b$  (respectivement, de  $\mathcal{S}_b$ ) de dimension  $\nu(n - \nu - 1)$  (respectivement,  $\nu(n - \nu)$ ).*

*Il existe alors un sous-espace totalement  $b$ -singulier maximal  $F$  et un drapeau complet  $\mathcal{F}$  de  $F$  tel que  $\mathcal{V} = \mathcal{A}_b \cap \mathcal{N}_{\mathcal{F},b}$  (respectivement,  $\mathcal{V} = \mathcal{S}_b \cap \mathcal{N}_{\mathcal{F},b}$ ).*

Nous pensons que le résultat reste vrai sans hypothèse sur le cardinal de  $\mathbb{K}$ , mais les méthodes utilisées jusqu'à présent doivent être largement revues pour accéder aux corps de petit cardinal. En tout cas, les hypothèses de cardinalité sont naturelles : elles sont liées à l'indice de nilpotence maximal possible pour un endomorphisme nilpotent  $b$ -symétrique ou  $b$ -alterné ; et elles permettent de conserver la nilpotence et l'indice de nilpotence maximal par extension du corps de base.

Jusqu'à présent, nous avons écarté la caractéristique 2 qui était pourtant l'une des motivations de notre étude. Cette situation, qui occupe complètement [11], est nettement plus délicate. En effet, dans ce cas la dimension maximale ne dépend pas uniquement de l'indice de Witt et du rang de  $b$ , elle dépend aussi d'un autre invariant, qui est la dimension du *super-noyau* de  $b$ , défini comme l'intersection du cône isotrope  $\{x \in E : b(x, x) = 0\}$  (qui est un sous-espace vectoriel de  $E$ , caractéristique 2 oblige) avec son  $b$ -orthogonal.

**Théorème 6.17** (Théorème 1.3 de [9]). *On suppose  $\chi(\mathbb{K}) = 2$ . Soit  $b$  une forme bilinéaire symétrique non dégénérée. On note  $n$  son rang,  $\nu$  son indice de Witt, et  $d$  la dimension de son super-noyau. Alors :*

- (a) *La plus grande dimension possible pour un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$  est  $\nu(n - \nu)$ .*
- (b) *Si  $n \neq 2\nu + 1$  alors la plus grande dimension possible pour un sous-espace vectoriel nilpotent de  $\mathcal{A}_b$  est  $\nu(n - \nu - 1)$ .*
- (c) *Si  $n = 2\nu + 1$  alors la plus grande dimension possible pour un sous-espace vectoriel nilpotent de  $\mathcal{A}_b$  est  $\nu(n - \nu) - d$ .*

Au regard de la situation où  $\chi(\mathbb{K}) \neq 2$ , les deux premiers résultats n'ont rien de surprenant, et la construction de sous-espaces nilpotents réalisant ces dimensions s'obtient en appliquant la même méthode. Le cas  $n = 2\nu + 1$  est plus étonnant, et il généralise l'exemple découvert par Quinlan. Dans ce dernier, on a  $n = 2m + 1$ ,  $\nu = m$ , et le super-noyau de la forme  $(X, Y) \in (\mathbb{K}^{2m+1})^2 \mapsto X^T Y$  est *nul* (l'orthogonal du cône isotrope étant la droite vectorielle engendrée par  $(1, \dots, 1)$ , laquelle ne contient pas de vecteur isotrope non nul). Le théorème 6.17 donne la dimension maximale  $m^2 + m$  dans cette situation pour un sous-espace vectoriel nilpotent de  $\mathcal{A}_b$ , et il est remarquable que ce soit exactement la même dimension que pour un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$  !

Expliquons maintenant l'exemple de Quinlan avec un point de vue géométrique. On considère la forme  $b : (X, Y) \mapsto X^T Y$  sur  $\mathbb{K}^{2m+1}$ . Son cône isotrope est l'hyperplan  $H$  des vecteurs dont la somme des coordonnées est nulle, son orthogonal la droite vectorielle engendrée par  $(1, \dots, 1)$ . La forme  $b$  induit une forme *symplectique* sur  $H$ . On note  $\pi$  la projection  $b$ -orthogonale sur  $H$ . On démontre facilement que  $\bar{u} := \pi \circ u|_H$  est  $b_H$ -alternée pour tout  $u \in \mathcal{A}_b$ , et que la fonction  $u \in \mathcal{A}_b \mapsto \pi \circ u|_H \in \mathcal{A}_{b_H}$  est surjective. Le noyau de cette dernière est de dimension  $2m$ .

L'observation cruciale de Quinlan est la suivante :

**Lemme 6.18.** *Avec les notations précédentes, et pour tout  $u \in \mathcal{A}_b$ , l'endomorphisme  $u$  a pour polynôme caractéristique  $X \chi_\pi(X)$ .*

Déterminer les sous-espaces nilpotents de  $\mathcal{A}_b$  revient donc à déterminer ceux de  $\mathcal{A}_{b_H}$ , et la dimension maximale des premiers est  $2m$  plus la dimension maximale des derniers. On retrouve alors l'exemple fourni par Quinlan en piochant dans  $\mathcal{A}_{b_H}$  un sous-espace nilpotent de dimension  $m^2 - m$ , donné par le point (b) du théorème 6.14.

On peut enfin construire un exemple dans le cas plus général du point (c). On démontre, grâce à la théorie<sup>7</sup> des formes bilinéaires symétriques en caractéristique 2 qu'il existe une base  $(e_1, \dots, e_{2n+1})$  de  $E$  dans laquelle la matrice de Gram de  $b$  s'écrit

$$S = \begin{bmatrix} 0_d & [0]_{d \times (n-2d)} & I_d \\ [0]_{(n-2d) \times d} & \alpha I_{2m+1} & [0]_{(n-2d) \times d} \\ I_d & [0]_{d \times (n-2d)} & \Theta \end{bmatrix}$$

où  $\Theta \in M_d(\mathbb{K})$  est diagonale,  $\alpha \in \mathbb{K} \setminus \{0\}$  et  $n - 2d = 2m + 1$ . On choisit alors un sous-espace vectoriel nilpotent  $\mathcal{W}$  de  $A_{2m+1}(\mathbb{K})$  de dimension  $m^2 + m$ , et on considère l'ensemble des endomorphismes représentés dans la base précédente par une matrice de la forme

$$\begin{bmatrix} A & \alpha E^T & \Theta A^T + J \\ [0]_{(n-2d) \times d} & N & E \\ 0_d & [0]_{d \times (n-2d)} & A^T \end{bmatrix}$$

avec  $N \in \mathcal{W}$ ,  $A \in T_d^{++}(\mathbb{K})$ ,  $E \in M_{n-2d,d}(\mathbb{K})$  et  $J \in A_d(\mathbb{K})$ . On vérifie alors que c'est bien un sous-espace nilpotent ayant la dimension critique indiquée.

Dans le cadre de la caractéristique 2, on ne dispose pas encore de résultat sur la structure des sous-espaces nilpotents réalisant la dimension critique, mais il est probable que les méthodes utilisées pour obtenir les inégalités puissent être mises à profit pour élucider leur structure.

### 6.7.3 Les méthodes en caractéristique différente de 2

Il est temps d'évoquer les méthodes employées pour obtenir les résultats cités dans le paragraphe précédent. Elles diffèrent en effet assez fortement de celles que j'ai décrites dans la partie consacrée aux sous-espaces à spectre trivial. Je ne suis pas parvenu à transposer la technique des vecteurs adaptés au cadre

<sup>7</sup> Voir à cet effet le dernier chapitre de notre Invitation aux Formes Quadratiques, Calvage et Mounet, 2011.

structuré (i.e. quadratique), et j'ai donc dû opérer un retour aux sources de mes toutes premières idées sur le problème de Gerstenhaber, qui recouvrent essentiellement la méthode trouvée par Mathes, Omladič et Radjavi dans [Mat91]. Je vais donc décrire sommairement cette dernière et évoquer ensuite quelques adaptations et innovations qui ont été nécessaires pour traiter le cas structuré.

Revenons donc au problème de la majoration de la dimension pour un sous-espace vectoriel nilpotent  $\mathcal{V}$  de  $\mathcal{L}(E)$ , où  $E$  est de dimension finie  $n \geq 1$ . Une idée élémentaire est de procéder par récurrence. En fixant un vecteur non nul  $x$  de  $E$ , on peut considérer le noyau  $\mathcal{V}x$  de  $u \in \mathcal{V} \mapsto u(x)$ . Les éléments de ce noyau induisent des endomorphismes nilpotents de  $E/\mathbb{K}x$ , et l'on peut noter  $\overline{\mathcal{V}}$  l'espace des endomorphismes ainsi formés. Le noyau  $\mathcal{K}$  de la projection canonique  $u \in \mathcal{V} \mapsto \overline{u} \in \overline{\mathcal{V}}$  est constitué des éléments de  $\mathcal{V}$  dont l'image est incluse dans  $\mathbb{K}x$  : ils s'écrivent tous  $\varphi \otimes x : y \mapsto \varphi(y)x$  où  $\varphi \in E^*$  appartient à l'orthogonal  $\{x\}^\perp$  au sens de la dualité. Le théorème du rang donne ainsi

$$\dim \mathcal{V} = \dim \overline{\mathcal{V}} + \dim \mathcal{K} + \dim \mathcal{V}x,$$

et il suffira de montrer que  $\dim \mathcal{K} + \dim \mathcal{V}x \leq \dim E - 1$  pour obtenir par récurrence la majoration souhaitée. Cette dernière inégalité de dimension provient d'un principe d'orthogonalité qui repose sur le lemme suivant :

**Lemme 6.19** (Lemme de trace). *Soit  $u, v$  deux endomorphismes nilpotents de  $E$  dont toutes les combinaisons linéaires sont nilpotentes. Alors  $\text{tr}(uv) = 0$ .*

Ce résultat s'obtient facilement par polarisation à partir de la nullité du coefficient devant  $X^{n-2}$  dans le polynôme caractéristique.

En remarquant que  $u \circ (\varphi \otimes x) = \varphi \otimes u(x)$  est de trace  $\varphi(u(x))$  pour tout  $x \in E$ , on obtient alors que  $\mathbb{K}x \oplus \mathcal{V}x$  est dual-orthogonal à  $\{\varphi \in E^* : \varphi \otimes x \in \mathcal{V}\}$ , et l'inégalité voulue s'en déduit immédiatement.

En réalité, Mathes, Omladič et Radjavi esquivent toute récurrence : en représentant les choses matriciellement, ils introduisent le sous-espace  $\mathcal{V}_+$  des éléments triangulaires supérieures (et donc triangulaires supérieures strictes) de  $\mathcal{V}$ , le sous-espace  $\mathcal{V}_-$ , projeté de  $\mathcal{V}$  sur  $T_n^-(\mathbb{K})$  parallèlement à  $T_n^+(\mathbb{K})$ , et enfin ils utilisent l'argument de trace pour observer que  $\mathcal{V}_+$  est orthogonal à  $\mathcal{V}_-$  pour la forme bilinéaire  $B : (U, L) \in T_n^{++}(\mathbb{K}) \times T_n^{--}(\mathbb{K}) \mapsto \text{tr}(UL)$ , qui est non dégénérée par rapport à chaque variable. L'inégalité sur les dimensions est alors immédiate sans récurrence. J'avais déjà trouvé moi-même cette méthode dès mes premiers travaux sur la question en 2006, mais les trois auteurs précités étaient allés encore plus loin avec une remarque très astucieuse. D'abord, ils avaient généralisé le lemme de trace en démontrant le cas  $k = 2$  dans le lemme suivant :

**Lemme 6.20** (Lemme de trace généralisé). *Soit  $u, v$  deux endomorphismes nilpotents de  $E$  dont toutes les combinaisons linéaires sont nilpotentes, et soit  $k \in \mathbb{N}$ . Si  $\mathbb{K}$  a plus de  $k$  éléments alors  $\text{tr}(u^k v) = 0$ .*

L'hypothèse de cardinalité est tout à fait indispensable. Notons aussi qu'il n'y a aucune forme de généralisation à des traces de type  $\text{tr}(u^k v^l)$ , idée fausse que l'exemple fournit en page 95 suffit à dissiper.

La démonstration classique de ce lemme se fait à l'aide d'une réduction de Jordan de  $v$  : voir le théorème 2.4 de [Mac12]. Notons aussi qu'en caractéristique nulle il s'obtient facilement par dérivation en 0 de la fonction

polynomiale  $t \mapsto \text{tr}((u + tv)^{k+1})$ . Dans [9], j'ai donné une démonstration encore plus élémentaire que celle de [Mac12] : elle est fondée sur une expression des dérivées des coefficients du polynôme caractéristique de  $A + sB$  par rapport à  $s$  en 0 (voir le lemme 2.2 de [9]).

Mathes, Omladič et Radjavi supposent ensuite que  $|\mathbb{K}| > 2$ . Ils remarquent que, dans le cas d'égalité  $\dim \mathcal{V} = \frac{n(n-1)}{2}$ , les espaces  $\mathcal{V}_+$  et  $\mathcal{V}_-$  sont chacun l'orthogonal de l'autre pour la forme  $B$ . Le lemme de trace généralisé permet de voir que si l'on prend  $u$  dans  $\mathcal{V}_+$ , alors  $u^2$  (qui est encore triangulaire supérieure stricte) est orthogonale à  $\mathcal{V}_-$  et figure donc dans  $\mathcal{V}_+$  donc dans  $\mathcal{V}$ . Et comme tout élément de  $\mathcal{V}$  est trigonalisable, on trouve plus généralement que  $\mathcal{V}$  est stable par élévation au carré. En polarisant, on en déduit que  $\mathcal{V}$  est stable par le produit de Jordan  $(u, v) \mapsto uv + vu$ , puis la généralisation par Jacobson du théorème de Engel garantit que  $\mathcal{V}$  est trigonalisable.

Comment adapter ce qui précède au cadre structuré? Pour la majoration de la dimension par récurrence, le vecteur  $x$  doit être choisi  $b$ -isotrope. Tout endomorphisme nilpotent dans  $\mathcal{S}_b$  et  $\mathcal{A}_b$  et qui annule  $x$  stabilise  $\{x\}^\perp$  et induit ainsi un endomorphisme nilpotent du quotient  $\{x\}^\perp/\mathbb{K}x$ , que l'on munit de la forme non dégénérée  $\bar{b}$  induite par  $b$ . La récurrence se fera ainsi de deux en deux sur les dimensions. Les tenseurs  $\varphi \otimes x$  sont remplacés par :

- les tenseurs  $b$ -symétriques  $x \otimes_b y : z \mapsto b(y, z)x + b(x, z)y$  pour les endomorphismes  $b$ -symétriques ;
- les tenseurs  $b$ -alternés  $x \wedge_b y : z \mapsto b(y, z)x - b(x, z)y$  pour les endomorphismes  $b$ -alternés.

Le lemme de trace s'applique encore, et tout fonctionne bien à condition que  $\chi(\mathbb{K}) \neq 2$ . On observe cependant que  $\text{tr}(u \circ (x \otimes_b y)) = 2b(x, u(y))$  lorsque  $u$  est  $b$ -symétrique, ce qui rend inopérante la technique de trace pour traiter la caractéristique 2. Nous reviendrons sur ce point épineux ultérieurement.

Quant à la méthode de Mathes, Omladič et Radjavi pour analyser les sous-espaces de dimension maximale, elle peut encore être appliquée dans le cas  $b$ -autoadjoint (car le carré d'un endomorphisme  $b$ -autoadjoint l'est encore) mais échoue lamentablement dans le cas  $b$ -antiautoadjoint (car le carré d'un endomorphisme  $b$ -antiautoadjoint est un endomorphisme  $b$ -autoadjoint). Il n'y a donc aucun espoir de s'en tirer par cette voie. En revanche, dans les deux autres cas la méthode s'adapte sans trop de peine en introduisant, comme dans la description de sous-espaces nilpotents particuliers de dimension maximale, un sous-espace totalement  $b$ -singulier maximal et un drapeau complet de celui-ci (en lieu et place d'un drapeau complet de l'espace total). Cette observation avait également été faite en parallèle de nos recherches par Kokol Bukovšek et Omladič [Buk17], qui s'étaient limités au cas de la forme bilinéaire symétrique canonique sur  $\mathbb{C}^n$  et des endomorphismes symétriques pour celle-ci.

Dans le cas des endomorphismes  $b$ -antiautoadjoints, l'étude des sous-espaces nilpotents de dimension maximale s'est révélée beaucoup plus délicate. D'abord, on peut se limiter au cas où  $|\mathbb{K}| > 3$  et observer que la méthode de Mathes, Omladič et Radjavi ne permet pas de montrer qu'un tel sous-espace  $\mathcal{V}$  est stable par élévation au carré, mais cependant qu'il est stable par élévation au cube ! Las, cette propriété est insuffisante pour garantir le caractère simultanément trigonalisable d'un sous-espace nilpotent (voir à nouveau le contre-exemple en page 95).

Il m'a fallu réinventer l'approche du théorème de Gerstenhaber en revenant

à une idée élémentaire séduisante mais très difficile à mettre en œuvre en pratique. Revenons donc à la situation de Gerstenhaber pour fixer les idées, en reprenant un sous-espace nilpotent  $\mathcal{V}$  de  $\mathcal{L}(E)$  (avec  $E$  de dimension  $n$ ) ayant la dimension maximale  $\frac{n(n-1)}{2}$ . Pour majorer la dimension, on était parti d'un vecteur arbitraire  $x \in E \setminus \{0\}$ , et on peut être tenté de tirer tout ce qu'il faut des contraintes de dimension ainsi obtenues. Notamment, l'espace bas  $\overline{\mathcal{V}}$  associé est nilpotent et de dimension critique, et on peut lui appliquer l'hypothèse de récurrence. Cela donne en particulier que  $\mathcal{V}$  contient un élément de nilindice  $n$  ou  $n - 1$ . Dans ma toute première approche d'octobre 2006, ma stratégie était de montrer à partir de là, en utilisant notamment des techniques à la Flanders, qu'il existait en fait un élément de rang  $n - 1$  (donc de nilindice  $n$ ). Et ensuite l'idée naturelle est de choisir  $x$  dans le noyau d'un tel élément et de tenter de montrer que tous les éléments de  $\mathcal{V}$  s'annulent en  $x$  ; à partir de cela, l'application de l'hypothèse de récurrence à  $\overline{\mathcal{V}}$  permettrait de conclure immédiatement. Mais cette idée est fort difficile à mettre en œuvre. Une percée que j'ai faite dans le cadre de mes travaux sur le problème de Gerstenhaber structuré a consisté à retravailler cette idée en utilisant des méthodes plus fines. Notamment, plutôt que de m'intéresser au noyau d'un élément de nilindice maximal – nilindice que l'on notera  $p$  dans la suite et dont nous savons qu'il est au moins égal à  $n - 1$  – je me suis intéressé aux vecteurs appartenant à  $\text{Im } u^{p-1}$  pour un  $u \in \mathcal{V}$  réalisant le nilindice  $p$ . Il est en effet beaucoup plus aisé de manipuler les images des  $v^{p-1}$  que leurs noyaux, car ces images sont nécessairement (sauf cas très simple) de dimension 1.

Avant de poursuivre, il est important de noter que je supposerai  $|\mathbb{K}| \geq n$  à partir de maintenant, et il faut aussi évoquer quelques points élémentaires. D'abord, en revenant à un vecteur  $x \in E \setminus \{0\}$  arbitraire, les contraintes de dimension font que  $\mathbb{K}x \oplus \mathcal{V}x$  est le préduel-orthogonal de  $\{\varphi \in E^* : \varphi \otimes x \in \mathcal{V}\}$ . En utilisant le lemme de trace généralisé à la façon de Mathes, Omladič et Radjavi, on obtient la première conséquence utile suivante : pour tout  $v \in \mathcal{V}$ , tout  $x \in E$  et tout  $k \in \llbracket 0, n - 1 \rrbracket$ , le vecteur  $v^k(x)$  figure dans  $\mathbb{K}x \oplus \mathcal{V}x$ . Dans mes toutes premières recherches, je n'avais observé que le cas  $k = 1$ , mais la généralisation est d'un très grand secours. Une conséquence immédiate est que pour un  $v$  de  $\mathcal{V}$ , on aura  $v(x) = 0$  ou  $\text{Im } v^{p-1} \subset \mathbb{K}x \oplus \mathcal{V}x$  car  $\text{Im } v^{p-1}$  est de dimension 0 ou 1. Par un argument de densité algébrique (permis par l'hypothèse de cardinalité), l'alternative se globalise comme suit : ou bien  $\mathcal{V}x = \{0\}$ , auquel cas on conclut directement par récurrence, ou bien  $\text{Im } v^{p-1} \subset \mathbb{K}x \oplus \mathcal{V}x$  pour tout  $v \in \mathcal{V}$ .

La suite de la démonstration consiste à prendre un  $x$  appartenant à  $\text{Im } u^{p-1}$  pour un  $u$  dans  $\mathcal{V}$ , que l'on fixe. On examine ensuite en profondeur la situation défavorable où  $\mathcal{V}x \neq \{0\}$ . Il est naturel d'introduire la somme  $K(\mathcal{V})$  de tous les sous-espaces  $\text{Im } v^{p-1}$ , ce qui conduit à  $K(\mathcal{V}) \subset \mathbb{K}x \oplus \mathcal{V}x$ . La dernière clef est une idée issue de l'analyse : on part d'un  $y \in E$  tel que  $u^{p-1}(y) = x$ , et l'on étudie la courbe  $t \mapsto (u + tv)^{p-1}(y)$ , qui prend ses valeurs dans  $K(\mathcal{V})$ . Son vecteur tangent  $x_1$  en 0 est donc dans  $K(\mathcal{V})$ , ce qui donne

$$x_1 = \sum_{k=0}^{p-2} u^k v u^{p-2-k}(y) \in K(\mathcal{V}).$$

Et le même raisonnement donne

$$\sum_{k=0}^{p-1} u^k v u^{p-1-k}(y) = 0.$$

En appliquant  $u$  à la première égalité obtenue et en simplifiant grâce à la seconde, il vient  $v(x) = u(-x_1) \in u(K(\mathcal{V}))$ . On a donc établi l'inclusion  $\mathcal{V}x \subset u(K(\mathcal{V}))$ .

Le coup de grâce s'obtient en observant que cette inclusion, combinée à  $K(\mathcal{V}) \subset \mathbb{K}x \oplus \mathcal{V}x$ , débouche en fait sur  $\mathcal{V}x = \{0\}$ . En effet, partant de  $v \in \mathcal{V}$ , on écrit  $v(x) = u(\alpha x + x_1)$  avec  $x_1 \in K(\mathcal{V})$  et  $\alpha \in \mathbb{K}$ , puis  $v(x) = u(u(\beta x + x_2))$  avec  $\beta \in \mathbb{K}$  et  $x_2 \in K(\mathcal{V})$ . De proche en proche, on montre ainsi que  $v(x)$  est dans l'image de toute puissance itérée de  $u$ , et finalement  $v(x) = 0$  car  $u$  est nilpotente !

C'est la méthode que nous venons de décrire qui a été adaptée pour obtenir le théorème 6.16. Dans la situation de ce théorème, on observe que le vecteur  $x$  envisagé sera systématiquement  $b$ -isotrope (car  $\text{Im } u^{p-1} \subset \text{Ker } u \cap \text{Im } u$ , sauf dans le cas  $p = 2$  qui est facile à régler séparément). Le cas où  $b$  est alternée est alors assez facilement résolu, une des clefs étant d'utiliser les principes d'orthogonalité pour obtenir que les tenseurs symétriques  $x \otimes_b x$ , où  $x$  est dans un  $\text{Im } v^{p-1}$ , sont tous dans  $\mathcal{V}$ . Et on parvient à montrer que  $K(\mathcal{V}) \subset \mathbb{K}x \oplus \mathcal{V}x$  par un raisonnement par l'absurde subtil mais bref. À noter que ce cas pourrait se déduire du théorème obtenu par Draisma, Kraft et Kuttler dans [Dra06] au prix d'une extension à une clôture algébrique (ce qui fonctionne bien car  $b$  est alternée et l'on n'a donc à craindre aucune augmentation de l'indice de Witt).

Le cas où  $b$  est symétrique est beaucoup plus délicat mais repose sur les mêmes grands principes. En plus de ce qui précède, on utilise notamment la stabilité de  $\mathcal{V}$  par élévation au cube, et surtout la stabilité par l'opération binaire  $(u, v) \mapsto u^2v + uvu + uv^2$  qui en découle. Une autre clef importante a été de remarquer que l'inclusion  $\mathcal{V}x \subset u(K(\mathcal{V}))$  peut être précisée comme suit : en repartant d'un  $y \in E$  tel que  $u^{p-1}(y) = x$ , et en prenant arbitrairement  $v$  dans  $\mathcal{V}$ , on constate que le vecteur  $(u + tv)^{p-1}(y)$  est  $b$ -isotrope pour tout  $t \in \mathbb{K}$  (sauf cas  $p = 2$ , que l'on règle à nouveau séparément) ; en dérivant en 0 l'identité  $b((u + tv)^{p-1}(y), (u + tv)^{p-1}(y)) = 0$ , on obtient alors l'inclusion  $\mathcal{V}x \subset u(K(\mathcal{V}) \cap \{x\}^{\perp_b})$ .

#### 6.7.4 Les méthodes en caractéristique 2

Le point de départ de mes recherches était l'étrange exemple proposé par Rachel Quinlan sur les corps de caractéristique 2. Il serait donc inconvenant de ne pas terminer ce chapitre par quelques considérations sur cette situation exceptionnelle. En caractéristique différente de 2, nous avons vu que la méthode itérative pour la majoration de la dimension reposait sur le lemme de trace. Mais ce dernier est presque stérile en caractéristique 2 : on démontre en effet facilement que  $\text{tr}(uv) = 0$  lorsque  $u \in \mathcal{S}_b$  et  $v \in \mathcal{A}_b$  (quelle que soit la caractéristique, d'ailleurs), et ici c'est fatal car  $\mathcal{A}_b \subset \mathcal{S}_b$ .

J'ai néanmoins, et très tôt dans mes recherches, réussi à mettre en évidence quelque chose qui ressemble au lemme de trace lorsque l'un des endomorphismes est un tenseur  $b$ -antisymétrique. Cette version du lemme de trace fait intervenir un curieux objet que j'ai baptisé l'**a-transformée** de la forme  $b$  et dont je ne suis pas parvenu à trouver de trace dans la littérature. L' $a$ -transformée de  $b$  est

la fonction bilinéaire

$$b_a : (x, y) \in E^2 \mapsto b(x, y) + \sqrt{b(x, x)b(y, y)} \in \overline{\mathbb{K}},$$

où  $\sqrt{\phantom{x}}$  la fonction racine carrée sur  $\overline{\mathbb{K}}$ . L'a-transformée de  $b$  a pour particularité d'être  $\mathbb{K}$ -bilinéaire, alternée<sup>8</sup> ; elle n'est pas *a priori* à valeurs dans  $\mathbb{K}$ , mais elle l'est automatiquement si  $\mathbb{K}$  est parfait et en particulier si  $\mathbb{K}$  est fini. On peut d'ailleurs remarquer que le rang de l'extension de  $b_a$  en une forme  $\overline{\mathbb{K}}$ -bilinéaire se calcule en fonction de celui de  $b$  : c'est le même si  $b$  est de rang pair, sinon c'est  $\text{rg } b - 1$ . Le lemme de trace est alors remplacé par l'énoncé suivant :

**Lemme 6.21** (Lemme de trace en caractéristique 2). *On suppose  $\chi(\mathbb{K}) = 2$ . Soit  $b$  une forme bilinéaire symétrique non dégénérée sur un espace vectoriel  $E$ . Soit  $\mathcal{V}$  un sous-espace vectoriel nilpotent de  $\mathcal{S}_b$ . Pour tout  $u \in \mathcal{V}$  et tout  $(x, y) \in E^2$  tel que  $b(x, y) = 0$  et  $x \wedge_b y \in \mathcal{V}$ , on a  $b_a(x, u(y)) = 0$ .*

En examinant finement ce que l'on peut tirer de la  $b_a$ -orthogonalité de deux sous-espaces vectoriels de  $E$ , on peut alors utiliser ce lemme pour avancer dans une démonstration par récurrence du théorème 6.17. Une spécificité de la démonstration est de raisonner par l'absurde en supposant que  $\mathcal{V}$  dépasse la dimension critique et d'en tirer progressivement des conclusions fortes sur la structure de  $\mathcal{V}$ .

## 6.8 Bilan et perspectives

Au fil de mes recherches, je suis parvenu à trouver trois démonstrations profondément inédites du théorème de Gerstenhaber :

- la démonstration de [1] (majoration de la dimension) et [5] (espaces de dimension maximale), fondée sur la méthode des vecteurs adaptés et de la compatibilité diagonale, et qui s'adapte à des situations plus générales ;
- la démonstration fondée sur une réduction par dualité opérateur-vecteur au théorème d'Atkinson sur les sous-espaces semi-primitifs [7], sous hypothèse de cardinalité suffisante sur le corps de base ;
- et enfin une variante de la démonstration de Mathes, Omladič et Radjavi, fondée sur des considérations géométriques sur les images ultimes non nulles des éléments de nilindice maximal, variante ayant permis de percer la structure des sous-espaces d'antiautoadjoints nilpotents de dimension maximale, toujours sous hypothèse de cardinalité suffisamment grande du corps de base.

Les méthodes et travaux précédents méritent d'être prolongés, et de nombreuses directions sont possibles et souhaitables. Je vais en indiquer quelques-unes avant de clore définitivement ce chapitre.

- (a) La méthode de compatibilité diagonale peut très probablement être utilisée pour étudier les espaces 1-spec de dimension maximale, et élucider enfin la structure des espaces  $\overline{2}$ -spec de dimension maximale en caractéristique 2.
- (b) La méthode fondée sur le lien avec le théorème d'Atkinson a été très peu utilisée, et nous sommes convaincus qu'elle est loin d'avoir délivré tout son potentiel : jusqu'à maintenant, on n'en a utilisé que la version la plus simple,

---

8. D'où le « a » de la terminologie « a-transformée ».



et le théorème d'Atkinson décrivant explicitement des sous-espaces primitifs n'a été utilisée que dans la situation critique où le nombre de lignes est maximal. Or on a vu (théorème 2.17) que l'on dispose pour ces espaces, à la Atkinson-Lloyd, d'un grand intervalle de nombres de lignes proches du nombre maximal possible et où l'on comprend essentiellement la structure de ces espaces. Il semble donc possible d'aller beaucoup plus loin dans les applications aux sous-espaces nilpotents, notamment pour comprendre la structure de ceux dont la dimension est proche de la dimension critique.

- (c) Les travaux sur les sous-espaces nilpotents dans le cas structuré (endomorphismes symétriques ou alternés) sont inachevés. En caractéristique différente de 2, il reste à généraliser aux corps finis les théorèmes portant sur les sous-espaces de dimension maximale dans le cas antiautoadjoint. En caractéristique 2, nous n'avons actuellement aucun théorème décrivant les sous-espaces de dimension maximale, et il va bien falloir y arriver. Bien sûr, la notion d'a-transformée d'une forme bilinéaire symétrique, découverte à cette occasion, mérite d'être approfondie. Enfin, on peut remarquer que les techniques utilisées ont consisté en un retour à des idées classiques sur le théorème de Gerstenhaber. Il serait intéressant de trouver un équivalent de l'approche « vecteurs adaptés », afin par exemple de pouvoir prolonger les résultats aux sous-espaces à spectre trivial dans le cas structuré. Et, pourquoi pas, de trouver un équivalent de la méthode fondée sur le théorème d'Atkinson. Nous conjecturons par exemple que, pour une forme symplectique  $b$  sur un espace de dimension  $2n$ , tout sous-espace vectoriel  $\mathcal{V}$  de  $\mathcal{A}_b$  à spectre trivial est de dimension au plus  $n^2$ , avec égalité uniquement s'il existe un lagrangien  $L$  de  $b$  stable par tous les éléments de  $\mathcal{V}$  (auquel cas la détermination de ces sous-espaces peut être déduite de celle des sous-espaces à spectre trivial de  $\mathcal{L}(L)$ ).
- (d) Enfin, je me suis totalement tenu à l'écart des méthodes de Meshulam, Radwan, Draisma, Kraft et Kuttler [Mes98, Dra06] pour résoudre les problèmes de sous-espaces ad-nilpotents d'algèbres de Lie. Cependant, le cas  $b$ -autoadjoint dans le problème de Gerstenhaber structuré suggère qu'il serait utile de tenter de transposer ces méthodes aux algèbres de Jordan.

## Chapitre 7

# Espaces de matrices diagonalisables

### 7.1 Articles concernés

- [1] B. Randé, C. de Seguins Pazzis, The linear preservers of real diagonalisable matrices. *Linear Algebra Appl.* **435-6** (2011) 1257–1266.
- [2] C. de Seguins Pazzis, From trivial spectrum subspaces to spaces of diagonalisable real matrices. *Math. Proc. R. Ir. Acad.* **118-A** (2018) 5–8.

### 7.2 Bibliographie historique

- [Mot55] T.S. Motzkin, O. Taussky, Pairs of matrices with property L II. *Trans. Amer. Math. Soc.* **80** (1955) 387–401.
- [Wat76] W. Waterhouse, Self-adjoint operators and formally real fields. *Duke Math. J.* **43** (1976) 237–243.
- [Ran91] B. Randé, Espaces vectoriels de dimension  $\frac{n(n+1)}{2}$  constitués de matrices diagonalisables. *RMS* **102** (1991–1992) 345–348.
- [Oml98] M. Omladič, P. Šemrl, Preserving diagonalisability. *Linear Algebra Appl.* **285** (1998) 165–179.
- [Dob04] M. Dobovišek, Simultaneous symmetrization. *Linear Algebra Appl.* **383** (2004) 107–112.
- [Bog07] I.I. Bogdanov, A.È Guterman, Monotone matrix transformations defined by the group inverse and simultaneous diagonalizability. *Sbornik : Math.* **198 :1** (2007) 1–16.

## 7.3 Problématique et historique

Dans la classe des problèmes consacrés aux sous-espaces vectoriels de matrices particulières, celui des sous-espaces vectoriels de matrices diagonalisables a rencontré une attention assez limitée par le passé. Une des raisons de ce manque relatif de popularité pourrait résider dans l'observation que la structure arithmétique du corps de base joue un rôle crucial dans les résultats que l'on peut espérer obtenir.

Le premier résultat majeur obtenu sur la question est une conséquence du célèbre théorème de Motzkin et Taussky [Mot55] : lorsque toutes les combinaisons linéaires de deux matrices  $A$  et  $B$  de  $M_n(\mathbb{C})$  sont diagonalisables, les matrices  $A$  et  $B$  commutent et sont donc codiagonalisables. En conséquence, tout sous-espace vectoriel de  $M_n(\mathbb{C})$  formé de matrices diagonalisables est conjugué à un sous-espace vectoriel de l'espace  $D_n(\mathbb{C})$  des matrices *diagonales*. En particulier, un tel sous-espace est de dimension au plus  $n$ . *A contrario*, toute matrice carrée réelle est diagonalisable, si bien que  $S_n(\mathbb{R})$  est un sous-espace vectoriel de matrices diagonalisables de dimension  $\frac{n(n+1)}{2}$ . On est donc dans des ordres de grandeur de dimension très différents de la situation du corps des complexes. Mieux encore, il est tout à fait élémentaire de constater que  $\frac{n(n+1)}{2}$  majore la dimension de tout sous-espace vectoriel de matrices diagonalisables dans  $M_n(\mathbb{K})$  pour un corps *arbitraire*. Il suffit pour cela de réaliser qu'un tel sous-espace vectoriel doit intersecter trivialement l'espace  $T_n^{++}(\mathbb{K})$  des matrices triangulaires supérieures strictes, car ces dernières sont toutes nilpotentes.

Dans [Ran91], Bernard Randé a démontré que tout sous-espace vectoriel de  $M_n(\mathbb{R})$  formé de matrices diagonalisables et de dimension  $\frac{n(n+1)}{2}$  est conjugué à  $S_n(\mathbb{R})$ . J'ai retrouvé ce résultat de manière indépendante en 2006 au même moment que je travaillais sur le théorème de Gerstenhaber. Ce n'est qu'en 2010 que je me suis résolu à publier ces travaux, d'une part car ma démonstration me semblait plus percutante que celle de mon confrère, d'autre part car j'en avais trouvé une application à la détermination des automorphismes de  $M_n(\mathbb{R})$  qui conservent le caractère diagonalisable, et enfin car j'avais en réalité prolongé le résultat à une classe plus vaste de corps. J'ai invité Bernard à co-écrire l'article [1] puisqu'il était pour moi le premier découvreur du résultat. Et en 2016, alors que je préparais un exposé à l'occasion du départ à la retraite de Bernard, j'ai fait la découverte que le théorème avait un lien étroit avec mes travaux sur les sous-espaces vectoriels à spectre trivial (voir le chapitre 6), et j'ai publié la nouvelle démonstration trouvée sous la forme d'une très courte note [2]. Je découvris à l'époque que le théorème avait aussi été publié, avec une méthode semblable à la nôtre, par Dobovišek [Dob04] en 2004 mais qu'il était passé sous notre radar à cause d'un titre déroutant (parlant de symétrisation simultanée et non de diagonalisabilité).

## 7.4 Le principal théorème

Le principal résultat obtenu dans [1] s'énonce comme suit :

**Théorème 7.1** (Théorème 3, point (b), de [1]). *Soit  $\mathbb{K}$  un corps quelconque et  $\mathcal{V}$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  constitué de matrices diagonalisables et de dimension  $\frac{n(n+1)}{2}$ . Alors  $\mathcal{V} = P S_n(\mathbb{K}) P^{-1}$  pour une matrice  $P \in GL_n(\mathbb{K})$ .*

Commentons un peu ce résultat. Sous réserve de l'existence d'un sous-espace vectoriel de  $M_n(\mathbb{K})$  constitué de matrices diagonalisables et de dimension  $\frac{n(n+1)}{2}$ , il est automatique que toute matrice de  $S_n(\mathbb{K})$  est diagonalisable. On sait qu'en général ce dernier fait est faux. Mieux, on sait caractériser [Wat76] les corps pour lesquels toute matrice symétrique (de taille arbitraire) est diagonalisable : il s'agit des corps réels qui sont l'intersection de leurs clôtures réelles dans une clôture algébrique<sup>1</sup>. En revanche, on ne sait pas, pour un  $n \geq 2$  donné, caractériser de manière élégante les corps  $\mathbb{K}$  pour lesquels toute matrice de  $S_n(\mathbb{K})$  est diagonalisable. C'est cependant possible pour  $n = 2$  : pour que toute matrice de  $S_2(\mathbb{K})$  soit diagonalisable, il est nécessaire et suffisant que  $\mathbb{K}$  soit réel<sup>2</sup> et pythagoricien<sup>3</sup>.

Voyons maintenant les grandes étapes de la démonstration du théorème 7.1 donnée dans [1]. Cette dernière est présentée de manière très matricielle dans l'article, et nous adopterons ici, pour l'essentiel, une démonstration sensiblement plus géométrique. On se donnera donc un espace vectoriel  $E$  de dimension  $n$  sur  $\mathbb{K}$ , ainsi qu'un sous-espace vectoriel  $\mathcal{V}$  de  $\mathcal{L}(E)$ , de dimension  $\frac{n(n+1)}{2}$  et formé d'endomorphismes diagonalisables. Il s'agit de montrer qu'il existe une forme bilinéaire symétrique  $b$  sur  $E$  possédant une base orthonormale et pour laquelle  $\mathcal{V}$  est l'ensemble des endomorphismes  $b$ -autoadjoints. L'idée est tout simplement de construire  $b$  à partir d'une base bien choisie.

On laisse de côté le cas  $n \leq 2$ , qui est bien connu et se traite purement matriciellement, et on suppose  $n \geq 3$ . Nous partirons d'une base arbitraire  $(u_1, \dots, u_n)$  de  $V$ . Si l'on pense au procédé de Schmidt, on peut espérer qu'une base bien choisie  $(e_1, \dots, e_n)$  puisse être obtenue à partir de  $(u_1, \dots, u_n)$ , de telle sorte qu'elle engendre le même drapeau complet. On cherchera donc  $e_n$  dans  $E \setminus \text{Vect}(u_1, \dots, u_{n-1})$ , et on pourra aussi espérer trouver  $(e_1, \dots, e_{n-1})$  comme base de  $\text{Vect}(u_1, \dots, u_{n-1})$ , ce qui suggère bien sûr un raisonnement par récurrence.

Nous allons maintenant détailler la mise en place, et nous passerons ensuite rapidement sur les détails ultimes. On pose  $H := \text{Vect}(u_1, \dots, u_{n-1})$  et l'on considère le sous-espace  $\mathcal{V}_H := \{u \in \mathcal{V} : u(H) \subset H\}$ . Par le théorème du rang, on a  $\dim \mathcal{V}_H \geq \dim \mathcal{V} - (n-1)$ . Par restriction, les éléments de  $\mathcal{W}_H := \{u|_H \mid u \in \mathcal{V}_H\}$  sont tous diagonalisables, et ainsi  $\dim \mathcal{W}_H \leq \frac{n(n-1)}{2}$ . Par le théorème du rang, on a  $\dim \mathcal{W}_H + \dim \mathcal{G} = \dim \mathcal{V}_H$  pour  $\mathcal{G} := \{u \in \mathcal{V} : u|_H = 0\}$ . Or, tout élément non nul de  $\mathcal{G}$  est de rang 1, et s'il est de trace nulle il doit donc être nilpotent donc nul. Ainsi la trace est injective sur  $\mathcal{G}$ , donc  $\dim \mathcal{G} \leq 1$ . En mettant bout à bout toutes les inégalités de dimension obtenues, il vient  $\dim \mathcal{G} = 1$  et  $\dim \mathcal{W}_H = \frac{n(n-1)}{2}$ . Grâce au dernier point, l'hypothèse de récurrence s'applique donc et fournit une forme bilinéaire symétrique  $b'$  sur  $H$ , possédant une base orthonormée  $(e_1, \dots, e_{n-1})$ , et dont  $\mathcal{W}_H$  est précisément l'ensemble des endomorphismes autoadjoints. Ensuite, il existe dans  $\mathcal{G}$  un unique élément de trace 1. C'est donc un idempotent  $\pi$ , dont l'image s'écrit  $\mathbb{K}e_n$  pour un certain vecteur  $e_n$  de  $E \setminus H$ . Attention, à ce stade il reste *a priori* une modification à faire, car  $e_n$  n'est bien déterminé qu'à multiplication par un scalaire non nul près.

L'étape suivante, critique, consiste à observer que tous les éléments de  $\mathcal{V}_H$

- 
1. À noter que cette caractérisation est fort jolie mais pratiquement inutilisable.
  2. Autrement dit, si une somme de carrés dans  $\mathbb{K}$  est nulle alors ses termes sont nuls.
  3. Autrement dit, toute somme de carrés est un carré.

stabilisent  $\text{Im } \pi$ . Pour cela, on prend un  $u \in \mathcal{V}_H$  et on montre que le fait que  $u + \lambda\pi$  soit diagonalisable pour tout  $\lambda \in \mathbb{K}$  force l'inclusion  $u(\text{Im } \pi) \subset \text{Im } \pi$ . C'est fondé sur les deux lemmes élémentaires suivants, qui sont aussi utilisés plus tard dans la démonstration :

**Lemme 7.2.** *Soit  $N \in M_n(\mathbb{K})$ ,  $C \in \mathbb{K}^n$  et  $a \in \mathbb{K}$ . Pour que la matrice  $\begin{bmatrix} N & C \\ 0 & a \end{bmatrix}$  soit diagonalisable, il est nécessaire que  $C \in \text{Im}(N - aI_n)$ .*

**Lemme 7.3.** *Pour un endomorphisme diagonalisable  $u$  d'un espace vectoriel de dimension finie, on a  $\bigcap_{\lambda \in \mathbb{K}} \text{Im}(u - \lambda \text{id}) = \{0\}$ .*

À partir de là, on est ramené à la situation suivante : l'espace  $\mathcal{V}$  contient toutes les matrices représentées dans la base  $\mathbf{B} := (e_1, \dots, e_n)$  trouvée par une matrice de la forme

$$\begin{bmatrix} S & [0] \\ [0] & a \end{bmatrix} \quad \text{où } S \in S_{n-1}(\mathbb{K}) \text{ et } a \in \mathbb{K}.$$

Ensuite, et en revenant à la définition de  $\mathcal{V}_H$ , on démontre que  $\mathcal{V}$  contient, pour toute ligne  $L \in M_{1,n-1}(\mathbb{K})$ , un endomorphisme représenté dans  $\mathbf{B}$  par une matrice de la forme

$$\begin{bmatrix} [?] & ? \\ L & ? \end{bmatrix}.$$

Grâce à ce qui précède, on peut même s'arranger pour prendre le bloc supérieur gauche triangulaire supérieur strict, et le coefficient inférieur droit nul. Ensuite, par une utilisation judicieuse des lemmes précédents ainsi que des matrices déjà connues représentant des éléments de  $\mathcal{V}$ , on démontre que la matrice peut systématiquement être prise de la forme  $\begin{bmatrix} [0] & \lambda L \\ L & 0 \end{bmatrix}$  pour un scalaire non nul  $\lambda$  indépendant de  $L$ , dont on montre ensuite que c'est nécessairement un carré. Il suffira alors d'écrire  $\lambda = \mu^2$ , de remplacer  $e_n$  par  $\mu^{-1}e_n$ , puis de définir  $b$  comme la forme bilinéaire symétrique de base orthonormée  $(e_1, \dots, e_n)$ , pour que  $\mathcal{V}$  contienne tout endomorphisme  $b$ -autoadjoint. L'égalité des dimensions donne alors que  $\mathcal{V}$  est exactement l'ensemble des endomorphismes  $b$ -autoadjoints, et la démonstration est achevée.

Notons enfin un point important que nous avons passé sous silence : dans la deuxième partie de la démonstration, il est nécessaire à certaines étapes de savoir que  $\mathbb{K}$  est réel et pythagoricien, ce qui s'obtient grâce à l'hypothèse de récurrence, qui permet une descente jusqu'au cas  $n = 2$  (on a montré que  $\mathcal{W}_H$  est un sous-espace vectoriel d'endomorphismes diagonalisables de  $\mathcal{L}(H)$ , de dimension  $\frac{n(n-1)}{2}$ ). Bien sûr, cette subtilité n'a pas lieu de nous préoccuper si l'on ne s'intéresse d'emblée qu'au corps des réels.

## 7.5 Lien avec les sous-espaces à spectre trivial

La démonstration donnée plus tôt est très représentative du genre de technique mis en œuvre dans bon nombre de nos travaux sur les sous-espaces vectoriels de matrices particulières (voir les chapitres 2 et 6).

En 2016, à l'occasion de la préparation d'un exposé autour de cet article pour la journée en l'honneur de Bernard Randé, j'ai découvert un lien inattendu

entre le théorème 7.1 et les sous-espaces de matrices à spectre trivial, lien qui permet de démontrer ce résultat à partir de ma généralisation du théorème de Gerstenhaber (voir la section 6.5.4).

Tout est fondé sur une méthode de dualité, mais cette fois-ci c'est une dualité au sens de la trace. Je détaillerai l'argument plus largement que d'habitude car il est remarquablement court. Soit  $E$  un espace vectoriel de dimension  $n$ , et  $\mathcal{V}$  un sous-espace vectoriel de  $\mathcal{L}(E)$  de dimension  $\frac{n(n+1)}{2}$  et dont tout élément est diagonalisable. On considère l'orthogonal  $\mathcal{V}^\perp$  de  $\mathcal{V}$  pour la forme bilinéaire  $(u, v) \mapsto \text{tr}(uv)$  (qui est symétrique et non dégénérée). L'observation remarquable est que :

- (i)  $\mathcal{V}^\perp$  est irréductible;
- (ii)  $\mathcal{V}^\perp$  est à spectre trivial.

Le premier point est très facile et ne dépend aucunement de l'hypothèse sur la dimension de  $\mathcal{V}$ . Supposant  $\mathcal{V}^\perp$  non irréductible, on prend un sous-espace non trivial  $F$  de  $E$  stable par tout élément de  $\mathcal{V}^\perp$ . Il est alors facile de construire dans  $\mathcal{V}$  un élément nilpotent non nul. Matriciellement, si l'on dispose en effet d'un sous-espace  $\mathcal{W}$  de  $M_n(\mathbb{K})$  dans lequel, pour un certain  $p \in \llbracket 1, n-1 \rrbracket$ , tout élément est de la forme  $\begin{bmatrix} [?]_{p \times p} & [?]_{p \times (n-p)} \\ [0]_{(n-p) \times p} & [?]_{(n-p) \times (n-p)} \end{bmatrix}$ , alors la matrice nilpotente  $E_{1,n}$  est trivialement dans  $\mathcal{W}^\perp$ .

Le deuxième point est plus subtil, mais on peut l'obtenir facilement à partir des tous premiers éléments de notre démonstration esquissée dans le paragraphe précédent. On y a en effet montré que pour tout hyperplan  $H$  de  $E$ , il existe dans  $\mathcal{V}$  un projecteur de noyau  $H$ . Fixons alors une droite vectorielle  $D$  de  $E$ . Par dualité (appliquer le principe précédent au transposé de  $\mathcal{V}$  dans  $\mathcal{L}(E^*)$ ), on trouve que  $D$  est l'image d'un projecteur  $\pi$  (de rang 1) appartenant à  $\mathcal{V}$ . Soit alors  $u \in \mathcal{V}^\perp$  dont  $D$  est droite propre, pour une valeur propre  $\lambda$ . Alors  $\lambda = \text{tr}(\pi u) = 0$ . Ainsi,  $\mathcal{V}^\perp$  est à spectre trivial.

Enfin, on a évidemment  $\dim \mathcal{V}^\perp = \frac{n(n-1)}{2}$ .

Supposons maintenant  $|\mathbb{K}| > 2$ . Notre généralisation du théorème de Gerstenhaber (sous la forme du corollaire 6.5) fournit alors une forme bilinéaire non dégénérée (*a priori* ni symétrique ni alternée) telle que  $\mathcal{V}^\perp$  soit l'ensemble des endomorphismes  $u \in \mathcal{L}(E)$  pour lesquels  $(x, y) \mapsto b(x, u(y))$  est alternée. On en déduit ensuite une forme bilinéaire non dégénérée  $c$  (*a priori* ni symétrique ni alternée) telle que  $\mathcal{V}^\perp$  soit l'ensemble des endomorphismes  $u \in \mathcal{L}(E)$  pour lesquels  $(x, y) \mapsto c(x, u(y))$  est symétrique. En remarquant que  $\mathcal{V}$  doit évidemment contenir  $\text{id}_E$  (sinon  $\mathcal{V} + \text{Vect}(\text{id}_E)$  serait de dimension trop élevée), on trouve que  $c$  est symétrique. En travaillant sur des matrices simples, il n'est alors pas bien difficile de montrer qu'un multiple de  $c$  possède une base orthonormale (seul point que nous ne détaillerons pas).

Si enfin  $|\mathbb{K}| = 2$ , alors on démontre directement par récurrence que  $\mathcal{V}$  ne peut exister.

On notera enfin que l'argument précédent est magnifique mais hélas totalement inopérant pour explorer les sous-espaces de matrices diagonalisables en-deçà de la dimension  $\frac{n(n+1)}{2}$ .

## 7.6 Application à la préservation de la diagonalisabilité

Comme souvent, les résultats sur les espaces vectoriels de grande dimension formés de matrices vérifiant une propriété donnée peuvent être mis à profit pour étudier les préserveurs linéaires de cette propriété. Et c'est ce que nous avons fait dans la dernière partie de [1], qui vaut son nom à l'article bien que ce ne soit pas sa partie la plus intéressante à notre avis.

Il n'existe aucun énoncé général sur les automorphismes de l'espace vectoriel  $M_n(\mathbb{K})$  qui conservent la diagonalisabilité. Seuls les cas de certains corps sont connus. Omladič et Šemrl [Oml98] ont déterminé ces automorphismes pour le corps des nombres complexes, et plus généralement pour les corps algébriquement clos de caractéristique 0, en s'appuyant sur le théorème de Motzkin-Taussky [Mot55]. En effet, ce dernier théorème permet de voir que tout automorphisme de  $M_n(\mathbb{K})$  conservant le caractère diagonalisable conserve aussi le caractère *simultanément diagonalisable*. Plus récemment, Bogdanov et Guterman [Bog07] ont obtenu, en s'inspirant de [Oml98], la forme générale pour les automorphismes de l'espace vectoriel  $M_n(\mathbb{K})$  qui conservent le caractère simultanément diagonalisable pour les couples de matrices (plus précisément, tels que  $f(A)$  et  $f(B)$  soient simultanément diagonalisables quelles que soient  $A, B$  simultanément diagonalisables dans  $M_n(\mathbb{K})$ ). Accessoirement, ces préserveurs sont les applications de l'une des formes

$$M \mapsto \lambda PMP^{-1} + \theta(M)I_n \quad \text{ou} \quad M \mapsto \lambda PM^T P^{-1} + \theta(M)I_n,$$

avec  $\lambda \in \mathbb{K}^*$ ,  $P \in \text{GL}_n(\mathbb{K})$ , et  $\theta$  une forme linéaire sur  $M_n(\mathbb{K})$  telle que  $\theta(I_n) \neq -\lambda$  (condition garantissant l'injectivité).

Ainsi, sur un corps quelconque, la stratégie de base pour obtenir les préserveurs linéaires du caractère diagonalisable est de montrer la conservation du caractère simultanément diagonalisable.

Dans [1], nous nous sommes placés sur un corps  $\mathbb{K}$  pour lequel toute matrice de  $S_n(\mathbb{K})$  est diagonalisable (par exemple  $\mathbb{R}$ ), et avons précisément démontré cette propriété. L'idée est tout simplement de donner une caractérisation des conjugués du sous-espace  $D_n(\mathbb{K})$  des matrices diagonales en fonction des sous-espaces vectoriels de dimension maximale constitués de matrices diagonalisables. Cela s'obtient grâce à l'énoncé suivant :

**Proposition 7.4.** *Soit  $\mathbb{K}$  un corps, et  $n \geq 2$  un entier naturel tel que toute matrice de  $S_n(\mathbb{K})$  soit diagonalisable.*

- (a) *Il existe deux sous-espaces vectoriels  $\mathcal{V}_1$  et  $\mathcal{V}_2$  de matrices diagonalisables de  $M_n(\mathbb{K})$ , tous de dimension  $\frac{n(n+1)}{2}$  et tels que  $\mathcal{V}_1 \cap \mathcal{V}_2 = D_n(\mathbb{K})$ .*
- (b) *Étant donné deux sous-espaces vectoriels  $\mathcal{V}_1$  et  $\mathcal{V}_2$  de matrices diagonalisables de  $M_n(\mathbb{K})$ , tous de dimension  $\frac{n(n+1)}{2}$ , si  $\mathcal{V}_1 \cap \mathcal{V}_2$  est de dimension  $n$  alors il est conjugué à  $D_n(\mathbb{K})$ .*

*Idées de démonstration :* Le premier énoncé s'obtient en prenant  $\mathcal{V}_1 = S_n(\mathbb{K})$  et  $\mathcal{V}_2 = PS_n(\mathbb{K})P^{-1}$  pour  $P = \text{Diag}(1, 2, \dots, n)$ . Pour le deuxième, on se ramène au cas  $\mathcal{V}_1 = S_n(\mathbb{K})$  et  $\mathcal{V}_2 = PS_n(\mathbb{K})P^{-1}$  pour un  $P \in \text{GL}_n(\mathbb{K})$ , puis on montre que l'on peut réaliser une sorte de décomposition en valeurs singulières de  $P$  sous la forme  $P = ODU$  avec  $O, U$  dans  $O_n(\mathbb{K})$ , et  $D$  diagonale (cela fait

intervenir de manière critique l'hypothèse voulant que tout élément de  $S_n(\mathbb{K})$  est diagonalisable, ainsi que la conséquence voulant que  $\mathbb{K}$  est réel et pythagoricien). On se réduit ainsi au cas où  $P$  est diagonale. Dans cette situation, l'inclusion  $D_n(\mathbb{K}) \subset \mathcal{V}_1 \cap \mathcal{V}_2$  est immédiate, et on en déduit le point (b).  $\square$

En utilisant le principe précédent, il est facile de voir que pour un automorphisme  $f$  de l'espace vectoriel  $M_n(\mathbb{K})$  conservant la diagonalisabilité, et sous les hypothèses de la proposition précédente,  $f$  envoie tout conjugué de  $D_n(\mathbb{K})$  sur un conjugué de  $D_n(\mathbb{K})$ , autrement dit elle conserve (au sens faible) le caractère simultanément diagonalisable. Le théorème de Bogdanov et Guterman fait alors le reste.

## 7.7 Perspectives

Les espaces de matrices réelles diagonalisables ont fait l'objet de fort peu d'études par le passé, et l'on peut envisager de nombreux développements intéressants.

D'abord, à la façon d'Atkinson et Lloyd, on pourrait envisager d'étudier la structure de sous-espaces de dimension proche de la dimension maximale  $\frac{n(n+1)}{2}$ . On a cependant peu de pistes à l'heure actuelle pour y parvenir.

D'autres questions que la symétrisation simultanée pourraient être étudiées. Par exemple, on pourrait chercher la plus petite dimension  $d_n$  à partir de laquelle tout sous-espace de matrices diagonalisables contient une matrice non nulle possédant une valeur propre multiple. Il est probable que ce genre de considérations mette à contribution des outils de topologie algébrique, comme on l'observe dans certaines études sur les sous-espaces vectoriels de matrices réelles avec contraintes de rang.

Une autre direction possible est l'adaptation des théorèmes précédents au cadre hermitien. Par exemple, on conjecture que tout sous- $\mathbb{R}$ -espace vectoriel de  $M_n(\mathbb{C})$  formé de matrices diagonalisables à valeurs propres réelles est de dimension au plus  $n^2$ , et que les sous-espaces de dimension maximale sont les conjugués du sous-espace  $\mathcal{H}_n(\mathbb{C})$  des matrices hermitiennes. Et l'on pourrait aussi chercher les sous- $\mathbb{R}$ -espaces vectoriels de  $M_n(\mathbb{C})$  formés de matrices diagonalisables sur  $\mathbb{C}$ . Et enfin prolonger l'étude à des corps plus généraux munis d'une involution non triviale.



## Chapitre 8

# Décompositions à l'aide d'objets quadratiques

### 8.1 Articles concernés

- [1] C. de Seguins Pazzis, On linear combinations of two idempotent matrices over an arbitrary field. *Linear Algebra Appl.* **433-3** (2010) 625–636.
- [2] C. de Seguins Pazzis, On decomposing any matrix as a linear combination of three idempotents. *Linear Algebra Appl.* **433-4** (2010) 843–855.
- [3] C. de Seguins Pazzis, On sums of idempotent matrices over a field of positive characteristic. *Linear Algebra Appl.* **433-4** (2010) 856–866.
- [4] C. de Seguins Pazzis, Sums of two triangularizable quadratic matrices over an arbitrary field. *Linear Algebra Appl.* **436-9** (2012) 3293–3302.
- [5] C. de Seguins Pazzis, A note on sums of three square zero matrices. *Linear Multilinear Algebra* **65** (2017) 787–805.
- [6] C. de Seguins Pazzis, Sums of three quadratic endomorphisms of an infinite-dimensional vector space. *Acta Sci. Math. (Szeged)* **83** (2017) 83–111.
- [7] C. de Seguins Pazzis, Sums of quadratic endomorphisms of an infinite-dimensional vector space. *Proc. Edinburgh Math. Soc.* **61** (2018) 437–447.
- [8] C. de Seguins Pazzis, Products of involutions in the stable general linear group. *J. Algebra* **530** (2019) 235–289.
- [9] C. de Seguins Pazzis, Products of involutions of an infinite-dimensional vector space. *Canadian J. Math.* **73** (2021) 195–220.
- [10] C. de Seguins Pazzis, Sums and products of pairs of quadratic endomorphisms of countable-dimensional vector spaces. *Linear Multilinear Algebra*, accepté, en attente de publication
- [11] C. de Seguins Pazzis, The sum and the product of two quadratic matrices : regular cases. *Adv. Appl. Clifford Algebras* **32** (2022).
- [12] C. de Seguins Pazzis, The sum of two quadratic matrices : Exceptional cases. *Linear Algebra Appl.* **653** (2022) 357–394.
- [13] C. de Seguins Pazzis, The product of two invertible quadratic matrices : Exceptional cases. *Linear Algebra Appl.* **662** (2023) 67–109.

## 8.2 Bibliographie historique

- [Won66] M.J. Wonenburger, Transformations which are products of two involutions. *J. Math. Mech.* **16** (1966), 327–338.
- [Djo67] D.Ž. Djoković, Products of two involutions. *Arch. Math.* **18** (1967), 582–584.
- [Erd67] J. Erdos, On products of idempotent matrices. *Glasgow Math. J.* **8(2)** (1967), 118–122
- [Hof71] F. Hoffman, E.C. Paige, Products of two involutions in the general linear group. *Indiana Univ. Math. J.* **20** (1971), 1017–1020.
- [Bal76] C.S. Ballantine, Products of idempotent matrices. *Linear Algebra Appl.* **19** (1976), 81–86.
- [Gus76] W.H. Gustafson, P.R. Halmos, H. Radjavi. Products of involutions. *Linear Algebra Appl.* **13** (1976), 157–162.
- [Wu90] P.Y. Wu, Sums of idempotent matrices. *Linear Algebra Appl.* **142** (1990), 43–54.
- [Har90a] R.E. Hartwig, M.S. Putcha, When is a matrix a difference of two idempotents? *Linear Multilinear Algebra* **26** (1990), 267–277.
- [Har90b] R.E. Hartwig, M.S. Putcha, When is a matrix a sum of idempotents? **26** (1990), 279–286.
- [Wan91a] J.-H. Wang, P.Y. Wu, Sums of square zero operators. *Studia Math.* **99-2** (1991), 115–127.
- [Wan91b] J.-H. Wang, P.Y. Wu, Products of unipotent matrices with index 2. *Linear Algebra Appl.* **149** (1991), 111–123.
- [Wan95a] J.-H. Wang, Factorization of matrices into quadratic ones II. *Linear Algebra Appl.* **231** (1995), 111–152.
- [Wan95b] J.-H. Wang, Sums and products of two quadratic matrices. *Linear Algebra Appl.* **229-1** (1995), 127–149.
- [Wan96] J.-H. Wang, Factorization of matrices into quadratic ones III. *Linear Algebra Appl.* **240** (1996), 21–39.
- [Rab04] V. Rabanovich, Every matrix is a linear combination of three idempotents. *Linear Algebra Appl.* **13** (2004), 157–162.
- [Bot10] J.D. Botha, Product of two unipotent matrices of index 2. *Linear Algebra Appl.* **433** (2010), 1447–1451.
- [Bot12] J.D. Botha, Sums of two square-zero matrices over an arbitrary field. *Linear Algebra Appl.* **436** (2012), 516–524.
- [dLC15] R.J. de La Cruz, Each symplectic matrix is a product of four symplectic involutions. *Linear Algebra Appl.* **466** (2015), 382–400.
- [Shi16] Y. Shitov, Sums of square-zero endomorphisms of a free module. *Linear Algebra Appl.* **507** (2016), 191–197.
- [Bre18] S. Breaz, Endomorphisms of free modules as sums of four quadratic endomorphisms. *Linear Multilinear Algebra* **11** (2018), 2215–2217.
- [Ell20] E. Ellers, O. Villa, Generation of the symplectic group by involutions. *Linear Algebra Appl.* **591** (2020), 154–159.

### 8.3 Problématique

On considère ici un corps quelconque  $\mathbb{K}$ , dont on choisit une clôture algébrique  $\overline{\mathbb{K}}$ . On se donne également une  $\mathbb{K}$ -algèbre  $\mathcal{A}$  (associative, unitaire). Soit  $p \in \mathbb{K}[X]$  de degré 2. Un élément  $x$  de  $\mathcal{A}$  est dit  **$p$ -quadratique** lorsque  $p(x) = 0$ . Un élément  $x$  de  $\mathcal{A}$  est dit **quadratique** (tout court) lorsqu'il est  $p$ -quadratique pour au moins un polynôme  $p$  de degré 2 à coefficients dans  $\mathbb{K}$ . Parmi les éléments quadratiques remarquables, on peut citer les idempotents ( $x^2 = x$ ), les involutions ( $x^2 = 1_{\mathcal{A}}$ ), les éléments de carré nul ( $x^2 = 0$ ), les éléments unipotents d'indice 2 ( $(x - 1_{\mathcal{A}})^2 = 0$ ), les quarts-de-tour ( $x^2 = -1_{\mathcal{A}}$ ).

À partir des années 1960, de nombreux travaux ont porté sur la décomposition de matrices en somme ou produit de matrices quadratiques de type spécifié. Citons :

- La détermination par Erdos des matrices qui sont produits de matrices idempotentes [Erd67] (ce sont les matrices singulières et les matrices identité) ; et le raffinement par Ballantine, qui détermina les matrices qui sont produits d'exactly  $r$  matrices idempotentes [Bal76] (ce sont les matrices  $M$  vérifiant la condition  $\text{rg}(M - I) \leq r \dim \text{Ker } M$ ).
- La détermination par Wu, Hartwig et Putcha des matrices complexes qui sont sommes de matrices idempotentes [Wu90, Har90b] (ce sont les matrices  $M$  dont la trace est un entier naturel minoré par le rang).
- La caractérisation des matrices inversibles qui sont produits de deux involutions [Hof71], [Won66], [Djo67] (ce sont les matrices semblables à leur inverse). La démonstration par Gustafson, Halmos et Radjavi du fait que toute matrice carrée de déterminant  $\pm 1$  est produit de quatre involutions, mais pas de trois involutions en général [Gus76].
- La découverte par Wonenburger [Won66] du fait que tout élément d'un groupe orthogonal (d'une forme quadratique non dégénérée sur un corps de caractéristique différente de 2) est produit de deux symétries orthogonales.
- La caractérisation par Wang et Wu des matrices qui sont sommes de deux matrices de carré nul, et leur démonstration du fait que toute matrice de trace nulle est somme de quatre matrices de carré nul, mais pas de moins en général [Wan91a].
- Enfin, la détermination par Hartwig et Putcha des matrices qui peuvent s'écrire comme la différence (ou la somme, les deux problèmes étant équivalents) de deux matrices idempotentes [Har90a].

Nous simplifierons le discours grâce à la terminologie suivante : donnons-nous  $r$  polynômes  $p_1, \dots, p_r$  à coefficients dans  $\mathbb{K}$ , tous unitaires de degré 2. Dans la  $\mathbb{K}$ -algèbre  $\mathcal{A}$ , appelons  $(p_1, \dots, p_r)$ -**somme** tout élément  $x$  de  $\mathcal{A}$  se décomposant sous la forme  $x = \sum_{k=1}^r a_k$  où  $(a_1, \dots, a_r) \in \mathcal{A}^r$  et  $p_i(a_i) = 0$  pour tout  $i \in \llbracket 1, r \rrbracket$ . Appelons  $(p_1, \dots, p_r)$ -**produit** tout objet  $x$  se décomposant sous la forme  $x = \prod_{k=1}^r a_k$  où  $(a_1, \dots, a_r) \in \mathcal{A}^r$  et  $p_i(a_i) = 0$  pour tout  $i \in \llbracket 1, r \rrbracket$ .

L'ensemble des éléments de  $\mathcal{A}$  annulés par un polynôme donné est stable par conjugaison. Il vient que l'ensemble des  $(p_1, \dots, p_r)$ -sommés est stable par conjugaison dans l'algèbre  $\mathcal{A}$ , et il en est de même pour l'ensemble des  $(p_1, \dots, p_r)$ -produits. Il s'agit donc, dans des algèbres spécifiques, de caractériser les  $(p_1, \dots, p_r)$ -sommés ou les  $(p_1, \dots, p_r)$ -produits, si possible à l'aide des invariants attachés

aux classes de conjugaison.

Énoncé avec un tel degré de généralité, ce problème est tout à fait insurmontable (au sens où, dans l'immense majorité des situations, on ne sait pas donner une caractérisation à la fois intelligible et non triviale). Citons par exemple trois problèmes pour lesquels on n'a aucun espoir de trouver une solution raisonnable :

- la détermination des matrices qui sont produits de trois involutions (problème de Halmos) ;
- la détermination des matrices qui sont sommes de trois matrices de carré nul ;
- la détermination des matrices qui sont sommes de trois matrices idempotentes.

## 8.4 Historique de mes travaux

En 2005, mon collègue de l'époque, Hubert Quatreville, attira mon attention sur le problème dit des quatre projecteurs. Il s'agissait de trouver trois projecteurs  $p_1, p_2, p_3$  d'un espace vectoriel de dimension infinie sur un corps de caractéristique 0, dont la somme est un projecteur mais qui ne commutent pas deux à deux. Des exemples circulaient dans la littérature, et Hubert était fier de partager ce qu'il estimait (à raison) être la solution la plus simple. Il donnait régulièrement l'exemple à étudier à ses élèves de MP sous la forme d'un devoir à la maison.

Le problème est intéressant car le cas de deux projecteurs est un classique (sans hypothèse de dimension, mais il faut un corps de caractéristique différente de 2). Le problème à trois projecteurs peut être reformulé comme suit : trouver quatre projecteurs  $p_1, p_2, p_3, p_4$  tels que  $p_1 + p_2 = p_3 - p_4$  sans que  $p_1$  ne commute avec  $p_2$ . Autrement dit, on cherche un endomorphisme  $u$  qui n'est pas un projecteur mais se décompose à la fois comme la somme de deux projecteurs et comme la différence de deux projecteurs.

Or, la somme de deux projecteurs  $p$  et  $q$  s'écrit également  $\text{id} + (p - (\text{id} - q))$ . Ainsi, étudier les sommes de deux projecteurs revient à étudier les différences de deux projecteurs.

Cela me conduisit, lors de l'hiver 2007, à essayer de caractériser les différences de deux projecteurs. J'étais à l'époque totalement ignorant des travaux d'Hartwig et Putcha [Har90a], et je travaillai sans aide. Je parvins à améliorer les résultats connus, mes résultats valant sur tout corps (y compris en caractéristique 2, où il y a des difficultés tout à fait spécifiques). Je ne publiai pas ces travaux dans un premier temps. Expliquons quand même tout de suite comment un exemple simple au problème des trois projecteurs découle de cette étude (cet exemple reformule celui trouvé par Quatreville). Tant que  $u$  est diagonalisable, on peut montrer que c'est une différence de deux projecteurs dès que  $u$  et  $-u$  sont semblables (attention, cette condition n'est pas nécessaire). Grâce à l'astuce précédente, on trouve qu'un endomorphisme diagonalisable  $u$  est somme de deux projecteurs dès qu'il est semblable à  $\text{id} - u$ . L'exemple est alors obtenu comme suit. On considère un  $\mathbb{C}$ -espace vectoriel de dimension dénombrable  $V$  muni d'une base  $(e_k)_{k \in \mathbb{Z}}$  indexée par  $\mathbb{Z}$ , et l'endomorphisme  $u$  envoyant systématiquement  $e_k$  sur  $ke_k$ . On peut se représenter  $u$  par une matrice diagonale infinie dont les lignes et colonnes sont indexées par  $\mathbb{Z}$ , chaque entier apparaissant exactement une fois sur la diagonale. Il est alors tout à fait

clair que  $u$  est semblable à la fois à  $-u$  et à  $\text{id} - u$ . Via une remarque précédente, cela permet de décomposer  $u$  à la fois comme  $p_1 + p_2$  et  $p_3 - p_4$  où  $p_1, \dots, p_4$  sont des projecteurs de  $V$ . Ainsi,  $p_1 + p_2 + p_4 = p_3$ , mais  $p_1$  ne peut commuter avec  $p_2$  sinon toute valeur propre de  $u$  serait dans  $\{0, 1, 2\}$ . Et alors  $p_3$  donne le contre-exemple recherché.

Les choses avancèrent lorsque j'entrepris de démontrer que tout endomorphisme d'un espace vectoriel de dimension finie est combinaison linéaire de trois projecteurs. Je ne sais plus ce qui m'avait poussé à m'intéresser à ce résultat, qu'un article de Rabanovich [Rab04] avait démontré quelques années plus tôt pour les corps de caractéristique nulle. L'hypothèse de caractéristique posée par Rabanovich me semblait hors-sujet, et j'entrepris de réaliser une démonstration sur un corps quelconque. Pour les besoins de la démonstration, j'avais besoin de savoir caractériser les matrices de la forme  $\alpha P + \beta Q$ , où  $\alpha$  et  $\beta$  sont des scalaires non nuls imposés, et  $P$  et  $Q$  des matrices idempotentes variables. Cela utilisait essentiellement les mêmes ressorts que la détermination des différences de projecteurs. Je publiai alors une série de trois articles : le premier [1] résout le problème des décompositions en combinaison linéaire de deux idempotents à coefficients imposés, le deuxième [2] établit que toute matrice est combinaison linéaire de trois idempotentes (sur un corps arbitraire) ; le troisième [3] étudie le problème de la longueur de la décomposition d'une matrice en somme de matrices idempotentes sur un corps de caractéristique non nulle. Il y est en particulier établi que, sur un corps premier de caractéristique  $p > 0$ , toute matrice carrée assez grande (en fonction de  $p$ ) est somme de quatre matrices idempotentes.

Je me réintéressai à ces travaux quelques années plus tard. En 2012, Botha [Bot12] avait généralisé à un corps quelconque (y compris de caractéristique 2) les travaux de Wang et Wu sur les matrices sommes de deux matrices de carré nul. Je savais à l'époque établir son résultat, et le fait d'avoir été devancé dans la publication piqua mon orgueil. J'entrepris alors de généraliser le seul travail de Wang et Wu sur les sommes qui avait été oublié : la détermination des matrices qui sont sommes d'une matrice idempotente et d'une matrice de carré nul [Wan95b]. Ce fut fait très vite, l'article [4] combinant les idées de Wang et Wu à mes propres innovations. À la publication de l'article était alors achevée, dans le problème des  $(p_1, p_2)$ -sommes, le cas où les deux polynômes  $p_1$  et  $p_2$  sont scindés sur le corps de base. Alors que bien des mathématiciens de tradition anglo-saxonne pouvaient se satisfaire d'un tel résultat, je considérai qu'on ne pouvait en rester là et qu'il fallait s'intéresser au cas où l'un des polynômes est irréductible, mais pendant plusieurs années ce cas résista à mes assauts.

Ces travaux furent relancés quelques années plus tard, par hasard. On m'avait donné à évaluer un article sur les décompositions de matrices infinies en sommes de matrices infinies de carré nul. L'article à évaluer était de peu de valeur, mais le problème piqua immédiatement ma curiosité, notamment parce qu'en quelques dizaines de minutes j'étais parvenu, avec une approche complètement différente de celle de l'auteur, à améliorer très nettement les résultats de l'auteur. Quelques jours de travail supplémentaires me permirent de démontrer que tout endomorphisme d'un espace vectoriel de dimension infinie est somme de quatre endomorphismes de carré nul. Encore mieux, dès que l'on se donne quatre polynômes unitaires  $p_1, p_2, p_3, p_4$  de degré 2, dont au moins deux sont scindés sur le corps de base, tout endomorphisme d'un espace vectoriel de dimension infinie est une  $(p_1, p_2, p_3, p_4)$ -somme. L'une des principales clefs de ce résultat réside

dans la faculté à décomposer n'importe quel décalage à droite (fonction agissant comme  $e_k \mapsto e_{k+1}$  sur une base  $(e_k)_{k \in \mathbb{N}}$  d'un espace de dimension dénombrable) comme une  $(p_1, p_2)$ -somme dès que  $p_1$  ou  $p_2$  est scindé (et, bien sûr,  $p_1$  et  $p_2$  sont de degré 2).

Encore mieux, je parvins à donner une caractérisation totalement intelligible des endomorphismes d'un espace vectoriel de dimension infinie qui sont sommes de trois endomorphismes de carré nul : la caractérisation exploite un raffinement de ma méthode sur les décompositions en quatre termes, et elle nécessite également un détour par des considérations de dimension finie. On a besoin de démontrer une version « stable » du problème des sommes de trois termes déjà indiqué comme insoluble. En m'inspirant des grandes idées de mes travaux sur les combinaisons linéaires de trois projecteurs, je parvins à démontrer [5] que pour toute matrice  $A \in M_n(\mathbb{K})$  de trace nulle, la matrice augmentée  $A \oplus 0_n$  (de taille  $2n$ ) est somme de trois matrices de carré nul. Deux ans plus tard, la version « produit » fut établie, où je remplaçai les matrices de carré nul par les involutions ou les matrices unipotentes d'indice 2 (étudiant même des décompositions mixtes), et le bloc d'augmentation  $0_n$  par  $I_n$  (voire d'autres blocs scalaires) : voir [8].

Je revins enfin aux travaux sur les décompositions à deux termes à l'occasion de ceux sur les décompositions d'endomorphismes en dimension infinie. Je cherchais en effet à savoir si tout décalage à droite est une  $(p_1, p_2)$ -somme même lorsque  $p_1$  et  $p_2$  sont irréductibles, et je fis l'observation élémentaire suivante : lorsque  $u$  est un endomorphisme d'un espace vectoriel  $E$  et  $p$  est un polynôme irréductible de degré 2, si l'on a une décomposition  $u = a - b$  où  $p(a) = p(b) = 0$ , alors  $\text{Im } u$  doit être stable par  $a$  et  $b$ . Ainsi, si  $u$  est un décalage à droite alors une telle décomposition ne peut exister, car  $E/\text{Im } u$  est de dimension 1, et un endomorphisme d'une droite vectorielle ne peut être annulé par un polynôme irréductible de degré strictement supérieur à 1 !

À l'automne 2015, la découverte de ce contre-exemple me poussa à reprendre mes travaux sur les décompositions à deux termes, et je fis alors une avancée décisive en découvrant les structures quaternioniques sous-jacentes au problème et qui étaient cachées jusqu'à présent par les hypothèses de scindage ! En quelques semaines de travail, je parvins à une solution définitive du problème des  $(p_1, p_2)$ -sommes. Puis, à l'automne 2016, je constatai que les techniques pouvaient être adaptées pour la classification des  $(p_1, p_2)$ -produits dans la situation non dégénérée où  $p_1(0)p_2(0) \neq 0$ . Le problème des décompositions quadratiques à deux termes, entamé 25 ans plus tôt par d'autres auteurs, était essentiellement clos par trois articles [11, 12, 13]. Je parvins ensuite à adapter mes méthodes sur les sommes aux produits en dimension infinie, parvenant notamment à donner une caractérisation complète des automorphismes d'un espace vectoriel de dimension infinie qui sont produits de trois involutions [9].

Je pensais le problème loin derrière moi mais la découverte de travaux récents sur les décompositions en produit d'involutions dans les groupes symplectiques [dLC15, Ell20] me fit prendre conscience d'un autre champ où l'on pouvait étudier les décompositions à deux termes. Par exemple, Wonenburger [Won66] avait montré que tout élément d'un groupe orthogonal est produit de deux involutions (à condition que le corps de base soit de caractéristique différente de 2), et des travaux commençaient à apparaître sur la détermination des sommes de deux matrices hamiltoniennes ou antihamiltoniennes complexes de carré nul etc. Je compris alors qu'avait été largement négligé le problème des décompositions

en termes quadratiques dans les algèbres d'endomorphismes munies d'une involution, problème qui occupe l'essentiel de mes recherches actuelles.

## 8.5 Le problème à deux termes

### 8.5.1 Énoncé du problème

Je vais ici présenter les grandes idées de la résolution complète du problème des décompositions d'endomorphismes en somme de deux endomorphismes quadratiques à polynôme annulateur imposé. On fixe ici deux polynômes  $p$  et  $q$ , unitaires de degré 2 à coefficients dans  $\mathbb{K}$ . On rappelle que  $\text{tr } p$  est l'opposé du coefficient de  $p$  devant  $X$ . On note enfin  $\mathcal{Z}(p)$  l'ensemble des racines de  $p$  dans la clôture algébrique  $\overline{\mathbb{K}}$ .

On s'inquiète donc, un endomorphisme  $u$  d'un espace vectoriel  $E$  de dimension finie étant donné, de l'existence de deux endomorphismes  $a$  et  $b$  de  $E$  tels que

$$u = a + b, \quad p(a) = 0 \quad \text{et} \quad q(b) = 0.$$

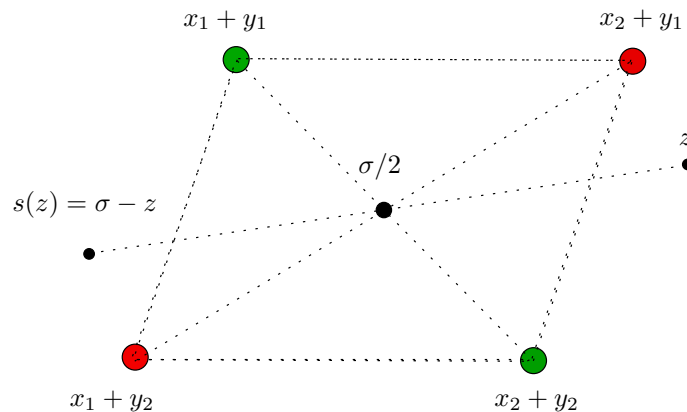
On cherche plus particulièrement à caractériser cette propriété d'existence en termes des invariants de similitude de  $u$ , ou encore de ses invariants primaires (les facteurs dans la décomposition des invariants de similitude en produits de puissances de polynômes irréductibles unitaires).

### 8.5.2 La dichotomie exceptionnel-régulier

Un rôle essentiel est joué par le scalaire

$$\sigma := \text{tr } p + \text{tr } q,$$

qui n'est autre que la somme des quatre racines de  $pq$  (dans  $\overline{\mathbb{K}}$ ). Un rôle tout aussi essentiel est, on s'en doute, joué par les éléments de  $\mathcal{Z}(p) + \mathcal{Z}(q)$ , qui forment (dans le cas générique) un parallélogramme.



Un principe important, qu'avaient déjà observé Wang, Wu, Hartwig et Putcha sur les cas particuliers qu'ils considéraient, est que lorsque  $u$  est une  $(p, q)$ -somme son spectre est presque symétrique par rapport à  $\frac{\sigma}{2}$ , ce qui s'exprime

très rigoureusement comme suit : lorsque  $u$  est une  $(p, q)$ -somme, l'ensemble  $\text{Sp}(u) \cap (\overline{\mathbb{K}} \setminus (\mathcal{Z}(p) + \mathcal{Z}(q)))$  est stable par la symétrie  $\lambda \mapsto \sigma - \lambda$ ; et mieux encore, pour tout  $\lambda$  dans  $\overline{\mathbb{K}} \setminus (\mathcal{Z}(p) + \mathcal{Z}(q))$ , la réduite de Jordan de  $u$  relative à la valeur propre  $\lambda$  a la même structure (i.e. les mêmes tailles de blocs) que celle relative à la valeur propre  $\sigma - \lambda$ . En revanche, et suivant les cas, la symétrie n'est plus qu'imparfaite lorsqu'on touche aux valeurs propres dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$ , et il est bien évident qu'on ne saurait avoir une symétrie parfaite en général. En effet lorsque  $\lambda \in \mathcal{Z}(p)$  et  $\mu \in \mathcal{Z}(q)$ , la matrice scalaire  $(\lambda + \mu)I_n$  est évidemment une  $(p, q)$ -somme dans  $M_n(\mathbb{K})$  mais  $\lambda + \mu$  est sa seule valeur propre ! Pour ces valeurs propres *exceptionnelles*, celles dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$  qui ne sont pas invariantes par  $\lambda \mapsto \sigma - \lambda$ , nous reviendrons sur cette symétrie imparfaite ultérieurement.

La dichotomie valeur propre dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$  contre valeur propre hors de  $\mathcal{Z}(p) + \mathcal{Z}(q)$  avait déjà été observée avant moi. J'ai néanmoins donné la version générale qui suit. D'abord, j'ai mis le doigt sur une généralisation de phénomènes de *commutation* que les autres auteurs avaient déjà observés et utilisés sur des cas particuliers. Ce phénomène vaut dans n'importe quelle  $\mathbb{K}$ -algèbre.

**Définition 10.** Soit  $p \in \mathbb{K}[X]$  unitaire de degré 2, et  $a$  un élément d'une  $\mathbb{K}$ -algèbre  $\mathcal{A}$  annulé par  $p$ . Le  $p$ -conjugué de  $a$  est défini comme  $a^* = (\text{tr } p)1_{\mathcal{A}} - a$ . Ainsi, il vérifie  $aa^* = a^*a = p(0)1_{\mathcal{A}}$ . Lorsque  $a$  n'est pas un élément scalaire,  $a^*$  est la seconde racine de  $p$  dans  $\mathbb{K}[a]$ .

**Lemme 8.1** (Lemme de commutation). *Soit  $a, b$  deux éléments quadratiques d'une  $\mathbb{K}$ -algèbre  $\mathcal{A}$ , annulés par des polynômes respectifs  $p$  et  $q$  unitaires de degré 2. On note  $a^*$  le  $p$ -conjugué de  $a$ , et  $b^*$  le  $q$ -conjugué de  $b$ . Alors, l'élément  $ab^* + ba^*$  est central dans l'algèbre engendrée par  $a$  et  $b$ .*

Considérons alors, sous les hypothèses du lemme, l'élément  $u := a + b$ . Il est d'une certaine manière logique d'introduire le *pseudo-adjoint*<sup>1</sup>

$$u^* := a^* + b^* = \sigma \text{id} - u$$

et la *pseudo-norme*

$$N(u) := uu^* = u^*u = \sigma u - u^2.$$

On note que  $N(u) = aa^* + bb^* + (ab^* + ba^*) = (p(0) + q(0))1_{\mathcal{A}} + (ab^* + ba^*)$ . Le lemme de commutation montre donc que la pseudo-norme commute avec  $a$  et  $b$ , ce qui s'utilise comme suit :

**Lemme 8.2** (Lemme de commutation pour les  $(p, q)$ -sommés). *Soit  $a, b$  deux éléments quadratiques d'une  $\mathbb{K}$ -algèbre  $\mathcal{A}$ , annulés par des polynômes respectifs  $p$  et  $q$  unitaires de degré 2. Posons  $u := a + b$ . Alors,  $a$  et  $b$  commutent tous deux avec  $u^2 - \sigma u$ .*

Ce phénomène de commutation est étroitement lié aux questions de symétrie évoquées plus tôt. Supposons en effet  $\mathbb{K}$  algébriquement clos, et considérons l'algèbre  $\mathcal{A}$  des endomorphismes d'un  $\mathbb{K}$ -espace vectoriel  $E$  de dimension finie. Sous les hypothèses du lemme précédent, pour tout  $\lambda$  dans  $\mathbb{K}$  le sous-espace caractéristique de  $u^2 - \sigma u$  associé à  $\lambda(\sigma - \lambda)$ , qui est la somme des sous-espaces

---

1. On prendra garde à cette notation séduisante mais qui peut être trompeuse : les objets  $u^*$  et  $N(u)$  ne sont pas fonctions de  $u$  seul, mais aussi de la décomposition choisie, autrement dit de  $a$  et  $b$ .



caractéristiques de  $u$  associés à  $\lambda$  et  $\sigma - \lambda$  (éventuellement égaux), est stable par  $a$  et  $b$ . Ce faisant, on décompose (toujours dans le cas particulier d'un corps algébriquement clos), le problème en se ramenant au cas où le spectre est inclus dans un ensemble de la forme  $\{\lambda, \sigma - \lambda\}$ .

Cette considération se généralise à un corps quelconque de la manière suivante. On introduit pour cela le polynôme

$$F_{p,q} := \prod_{1 \leq i, j \leq 2} (X - x_i - y_j)$$

où  $p = (X - x_1)(X - x_2)$  et  $q = (X - y_1)(X - y_2)$  ont été scindés sur  $\overline{\mathbb{K}}$ . On peut facilement voir que  $F_{p,q}$  est à coefficients dans  $\mathbb{K}$ . Mieux encore,  $F_{p,q}$  peut s'écrire  $\Lambda_{p,q}(X^2 - \sigma X)$  pour un polynôme  $\Lambda_{p,q}$  unitaire de degré 2 à coefficients dans  $\mathbb{K}$ .

Soit  $u$  un endomorphisme d'un espace vectoriel  $E$  de dimension finie. On peut alors s'appuyer sur la décomposition de Fitting

$$E = \text{Nil}(F_{p,q}(u)) \oplus \text{Co}(F_{p,q}(u))$$

où, pour un endomorphisme  $v$  de  $E$ , on a noté

$$\text{Nil}(v) := \bigcup_{n \in \mathbb{N}} \text{Ker } v^n \quad \text{et} \quad \text{Co}(v) := \bigcap_{n \in \mathbb{N}} \text{Im } v^n.$$

Si l'on dispose d'une décomposition  $u = a + b$  où  $a, b$  sont des endomorphismes de  $E$  annihilés respectivement par  $p$  et  $q$ , alors  $a$  et  $b$  commutent avec  $F_{p,q}(u)$ , car c'est un polynôme en  $u^2 - \sigma u$ , et  $a$  et  $b$  stabilisent donc les termes  $\text{Nil}(F_{p,q}(u))$  et  $\text{Co}(F_{p,q}(u))$  de sa décomposition de Fitting. Par considération des endomorphismes induits, on récupère alors que l'endomorphisme  $u_r$  induit par  $u$  sur  $\text{Co}(F_{p,q}(u))$  est une  $(p, q)$ -somme, et que l'endomorphisme induit  $u_e$  sur  $\text{Nil}(F_{p,q}(u))$  est aussi une  $(p, q)$ -somme. Et réciproquement, si  $u_e$  et  $u_r$  sont des  $(p, q)$ -sommés alors on reconstitue à partir de telles décompositions que  $u$  est une  $(p, q)$ -somme. Or,  $F_{p,q}(u_r)$  est inversible tandis que  $F_{p,q}(u_e)$  est nilpotent. Autrement dit, dans la clôture algébrique  $\overline{\mathbb{K}}$ ,  $u_r$  ne possède aucune valeur propre dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$ , alors que  $u_e$  a toutes ses valeurs propres dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$ .

Cela justifie la terminologie suivante :

**Définition 11.** Un endomorphisme  $u$  est dit  $(p, q)$ -régulier lorsqu'il ne possède aucune valeur propre dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$  (ici, par valeurs propres on entend les racines du polynôme caractéristique), autrement dit lorsque son polynôme minimal est premier avec  $F_{p,q}$ .

Un endomorphisme  $u$  est dit  $(p, q)$ -exceptionnel lorsque, dans  $\overline{\mathbb{K}}$ , toutes ses valeurs propres sont dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$ , autrement dit lorsque son polynôme minimal divise une puissance de  $F_{p,q}$ .

Avec les notations précédentes, on dit que  $u_r$  est la partie  $(p, q)$ -régulière de  $u$ , et que  $u_e$  est sa partie  $(p, q)$ -exceptionnelle.

On a ainsi établi qu'un endomorphisme  $u$  est une  $(p, q)$ -somme si et seulement si ses parties  $(p, q)$ -régulière et  $(p, q)$ -exceptionnelle sont toutes deux des  $(p, q)$ -sommés. Cette division de l'étude est tout à fait fondamentale, les outils et techniques pour étudier chacun de ces types de  $(p, q)$ -sommés étant profondément différents.

### 8.5.3 Intervention des algèbres de quaternions pour l'étude des $(p, q)$ -sommages régulières

Dans le cas où  $pq$  est scindé, l'étude des  $(p, q)$ -sommages régulières est obtenue par l'observation suivante : scindons  $p = (X - x)(X - y)$  et  $q = (X - x')(Y - x')$  ; pour toute matrice  $(p, q)$ -régulière  $M \in M_n(\mathbb{K})$  et toute décomposition  $M = A + B$  où  $p(A) = q(B) = 0$ , l'entier  $n$  est pair et les matrices  $A$  et  $B$  sont simultanément semblables, pour une certaine matrice  $N \in GL_{n/2}(\mathbb{K})$  inversible, à

$$\begin{bmatrix} xI_{n/2} & 0_{n/2} \\ I_{n/2} & yI_{n/2} \end{bmatrix} \quad \text{et} \quad \begin{bmatrix} x'I_{n/2} & N \\ 0_{n/2} & y'I_{n/2} \end{bmatrix}.$$

On est ainsi réduit à étudier, pour une matrice inversible  $N \in GL_{n/2}(\mathbb{K})$ , les invariants de similitude de la matrice blocs

$$\begin{bmatrix} (x + x')I_{n/2} & N \\ I_{n/2} & (y + y')I_{n/2} \end{bmatrix},$$

ce que le lemme suivant donne.

**Lemme 8.3** (Cas particulier du lemme 7 de [4], en fait démontré comme lemme 14 dans [1]). *Soit  $N \in M_n(\mathbb{K})$ , dont on note  $r_1, \dots, r_k$  les invariants de similitude, et soit  $\alpha$  et  $\beta$  dans  $\mathbb{K}$ . Alors les invariants de similitude de*

$$\begin{bmatrix} \alpha I_n & N \\ I_n & \beta I_n \end{bmatrix}$$

*sont  $r_1(X^2 - \sigma X), \dots, r_k(X^2 - \sigma X)$  pour  $\sigma := \alpha + \beta$ .*

La simplicité de cet argument a très longtemps bloqué mes efforts pour caractériser les  $(p, q)$ -sommages régulières lorsque l'un des polynômes  $p$  et  $q$  est irréductible. En effet, toute la méthode s'effondre dans ce cas, mais l'on peut bien sûr obtenir au moins une condition nécessaire par extension des scalaires et conservation des invariants de similitude dans une telle extension :

**Proposition 8.4.** *Les invariants de similitude d'une  $(p, q)$ -somme régulière sont nécessairement des polynômes en  $X^2 - \sigma X$ .*

La réciproque, qui est vraie dans le cas où  $pq$  est scindé, s'avère étonnamment fautive dans le cas général, et c'est sur ce point que mes recherches ont longtemps achoppées. Le contre-exemple suivant est tout à fait révélateur : considérons le corps  $\mathbb{K} = \mathbb{R}$  et les polynômes  $p = q = X^2 + 1$ . On s'intéresse donc aux sommages de deux quarts-de-tour sur le corps des réels. Considérons le polynôme  $X^2 + 2$ . Il est irréductible et c'est bien un polynôme en  $X^2 - \sigma X$  puisqu'ici  $\sigma = 0$ . Supposons qu'il existe un endomorphisme cyclique  $u$  (d'un  $\mathbb{R}$ -espace vectoriel  $V$ ) de polynôme minimal  $X^2 + 2$  qui admette une décomposition en somme de deux quarts-de-tour  $a$  et  $b$ . On note que  $0 = u^2 + 2 \text{id} = a^2 + b^2 + 2 \text{id} + ab + ba = ab + ba$ , donc  $a$  et  $b$  anticommulent. Or  $a^2 = b^2 = -\text{id}$ . On peut ainsi faire agir le corps gauche  $\mathbb{H}$  des quaternions sur  $V$  par  $i.x := a(x)$ ,  $j.x := b(x)$  et  $k.x := (ab)(x)$  pour tout  $x \in V$ , faisant donc de  $V$  un  $\mathbb{H}$ -espace vectoriel à gauche (induisant sa structure de  $\mathbb{R}$ -espace vectoriel). C'est absurde car  $V$  serait alors de dimension multiple de 4 sur  $\mathbb{R}$ , alors qu'il est censé être de dimension 2 !

Il est donc raisonnable d'imaginer que, plus généralement, les algèbres de quaternions vont apparaître dans l'étude du cas général. C'est effectivement le

cas, et je vais maintenant expliquer comment est naturellement venue la solution. Soit donc deux endomorphismes  $a$  et  $b$  d'un espace vectoriel de dimension finie, tels que  $p(a) = q(b) = 0$ . Posons  $u := a + b$  et supposons que, pour un certain polynôme unitaire  $R \in \mathbb{K}[X]$ , l'endomorphisme  $u$  soit cyclique de polynôme minimal  $R(X^2 - \sigma X)$ . D'après le lemme 8.3, il existe une base  $\mathbf{B}$  dans laquelle  $u$  est représenté par la matrice blocs  $\begin{bmatrix} 0 & C(R) \\ I_n & \sigma I_n \end{bmatrix}$  où  $C(R)$  est la matrice compagnon du polynôme  $R$ . Dans cette même base  $\mathbf{B}$ , on voit facilement que  $v := u^2 - \sigma u$  est représenté par  $\begin{bmatrix} C(R) & 0 \\ 0 & C(R) \end{bmatrix}$ . La commutation de  $a$  et  $b$  avec  $v$  indique alors que les matrices respectives  $A$  et  $B$  de  $a$  et  $b$  dans  $\mathbf{B}$  sont toutes dans  $M_2(D)$  où  $D$  est l'anneau commutatif  $\mathbb{K}[C(R)]$  des polynômes en  $C(R)$ , anneau naturellement isomorphe au quotient  $\mathbb{K}[X]/(R)$ .

Enfin,  $AB + BA - (\text{tr } q)A - (\text{tr } p)B$ , qui représente  $v + (p(0) + q(0)) \text{id}$  dans la base considérée, est donc la matrice scalaire  $(C(R) + (p(0) + q(0))I_d) \cdot I_2$  de  $M_2(D)$ , où  $d = \deg R$ . En n'oubliant pas que  $p(A) = 0$  et  $q(B) = 0$ , tout cela conduit naturellement à s'intéresser à la structure algébrique définie ci-après :

**Définition 12.** Soit  $D$  une  $\mathbb{K}$ -algèbre commutative unitaire, et  $x$  un élément de  $D$ . On note  $\mathcal{W}_D(p, q, x)$  l'anneau quotient de l'anneau  $D\{X, Y\}$  des polynômes non commutants en deux indéterminées  $X$  et  $Y$  à coefficients dans  $D$  par l'idéal bilatère engendré par les éléments  $p(X)$ ,  $q(Y)$  et  $XY - (\text{tr } q)X - (\text{tr } p)Y - x \cdot 1_{D\{X, Y\}}$ .

L'algèbre  $\mathcal{W}_D(p, q, x)$  est profondément liée aux algèbres de quaternions. Notons  $a$  et  $b$  les classes respectives des indéterminées  $X$  et  $Y$ , et notons que  $p(a) = q(b) = 0$  et  $ab + ba - (\text{tr } q)a - (\text{tr } p)b = 0$ . On peut alors naturellement définir un antiautomorphisme de conjugaison  $h \in \mathcal{W}_D(p, q, x) \mapsto h^* \in \mathcal{W}_D(p, q, x)$  qui envoie  $a$  sur son  $p$ -conjugué et  $b$  sur son  $q$ -conjugué. On démontre que cette conjugaison est une involution et que

$$\forall h \in \mathcal{W}_D(p, q, x), \quad hh^* = h^*h = N(h).1$$

pour une fonction  $N : \mathcal{W}_D(p, q, x) \rightarrow D$  dite norme. On peut démontrer que  $\mathcal{W}_D(p, q, x)$  est un  $D$ -module libre de rang 4 et de base  $(1, a, b, ab)$ . Tout cela n'est pas sans rappeler les algèbres de quaternions ! Et pour cause ...

**Théorème 8.5** (Théorème 2.4 de [11]). *On suppose que  $D$  est une extension de  $\mathbb{K}$  et que  $N$  est non dégénérée. Alors  $\mathcal{W}_D(p, q, x)$  est une algèbre de quaternions,  $h \mapsto h^*$  est sa conjugaison et  $N$  sa norme. Par suite, si  $N$  est anisotrope alors  $\mathcal{W}_D(p, q, x)$  est un corps gauche, et si  $N$  est isotrope alors  $\mathcal{W}_D(p, q, x)$  est isomorphe comme  $D$ -algèbre à  $M_2(D)$ .*

Remarquons en outre que dès que l'un des polynômes  $p$  ou  $q$  est scindé (mais peut-être pas l'autre), la forme  $N$  est isotrope. En effet, si l'on dispose par exemple d'une racine  $z$  de  $p$  dans  $\mathbb{K}$ , alors  $(a - z \cdot 1)(a - z \cdot 1)^* = (a - z \cdot 1)(z' \cdot 1 - a) = -p(a) = 0$  où  $z'$  est l'autre racine de  $p$ , et  $a - z \cdot 1$  n'est pas nul.

Donnons une application qui généralise l'exemple déjà vu pour le corps standard des quaternions : considérons un polynôme irréductible unitaire  $R \in \mathbb{K}[X]$  tel que  $R(X^2 - \sigma X)$  n'ait pas de racine dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$ , et supposons que  $\mathcal{W}_D(p, q, x)$  soit un corps gauche pour  $D := \mathbb{K}[X]/(R)$  et  $x$  la classe

de  $p(0) + q(0) + X$  modulo  $R$ . Il ne peut exister alors de morphisme de  $D$ -algèbres de  $\mathcal{W}_D(p, q, x)$  dans  $M_2(D)$  car tout  $\mathcal{W}_D(p, q, x)$ -module est de dimension au moins 4 comme  $D$ -espace vectoriel. Ainsi, la matrice compagnon du polynôme  $R(X^2 - \sigma X)$  n'est pas une  $(p, q)$ -somme! Inversement, si, avec les mêmes données,  $\mathcal{W}_D(p, q, x)$  est isomorphe à  $M_2(D)$ , alors un tel isomorphisme donne immédiatement que la matrice compagnon du polynôme  $R(X^2 - \sigma X)$  est une  $(p, q)$ -somme! Ces résultats se généralisent ensuite à des polynômes  $R$  qui sont des puissances d'irréductibles unitaires.

Nous nous contenterons de citer le résultat principal que nous avons obtenu dans [11].

**Théorème 8.6** (Théorème 3.10 de [11]). *Soit  $u$  un endomorphisme  $(p, q)$ -régulier d'un espace vectoriel de dimension finie.*

- (a) *Pour que  $u$  soit une  $(p, q)$ -somme, il est nécessaire que chacun de ses invariants de similitude soit un polynôme en  $X^2 - \sigma X$ .*
- (b) *On suppose que les invariants de similitude de  $u$  sont des polynômes en  $X^2 - \sigma X$ . Pour tout polynôme irréductible unitaire  $R \in \mathbb{K}[X]$  et tout entier  $k \geq 1$ , notons  $n_{R,k}(u)$  le nombre d'invariants de similitude de  $u$  de la forme  $S(X^2 - \sigma X)$  où  $S$  est de valuation  $k$  par rapport à l'irréductible  $R$ . Pour que  $u$  soit une  $(p, q)$ -somme, il est alors nécessaire et suffisant que le nombre  $n_{R,k}(u)$  soit pair pour tout entier  $k \geq 1$  et tout irréductible unitaire  $R \in \mathbb{K}[X]$  tel que  $R(X^2 - \sigma X)$  n'ait pas de racine dans  $\mathcal{Z}(p) + \mathcal{Z}(q)$  et tel que la norme de l'algèbre  $\mathcal{W}_D(p, q, x)$  soit anisotrope pour  $D := \mathbb{K}[X]/(R)$  et  $x$  la classe de  $p(0) + q(0) + X$  modulo  $R$ .*

### 8.5.4 Sur le cas des endomorphismes exceptionnels

Contrairement à la situation des endomorphismes  $(p, q)$ -réguliers, l'étude des endomorphismes  $(p, q)$ -exceptionnels nécessite un examen au cas par cas. Au total, on ne peut réduire la situation à moins de huit cas! Citons-les :

- (i)  $p$  et  $q$  sont scindés à racine double ;
- (ii)  $p$  et  $q$  sont scindés à racines simples ;
- (iii)  $p$  et  $q$  sont scindés et exactement l'un d'entre eux est à racine double ;
- (iv)  $p$  est irréductible,  $q$  est scindé et si l'on note  $y_1, y_2$  les racines de  $q$ , on a  $p(X - y_1) \neq p(X - y_2)$  ;
- (v)  $p$  est irréductible,  $q$  est scindé et si l'on note  $y_1, y_2$  les racines de  $q$ , on a  $p(X - y_1) = p(X - y_2)$  ;
- (vi)  $p$  et  $q$  sont irréductibles et ont le même corps de décomposition dans  $\overline{\mathbb{K}}$  ;
- (vii)  $p$  et  $q$  sont irréductibles, n'ont pas le même corps de décomposition dans  $\overline{\mathbb{K}}$  et ont des discriminants distincts ;
- (viii)  $p$  et  $q$  sont irréductibles, n'ont pas le même corps de décomposition dans  $\overline{\mathbb{K}}$  mais ont le même discriminant (ce qui ne peut se produire qu'en caractéristique 2).

Avant que je ne m'empare du sujet, rien n'était connu sur les cas d'irréductibilité de l'un des polynômes, et presque tout l'était lorsque les deux polynômes sont scindés, mais seulement sur les corps algébriquement clos de caractéristique différente de 2. Dans la suite, je vais principalement donner les grandes idées de

l'étude du cas où  $p$  et  $q$  sont scindés, ensuite j'expliquerai brièvement quelques stratégies de base pour traiter les cas d'irréductibilité.

On a vu que lorsqu'un endomorphisme  $u$  d'un espace vectoriel  $E$  de dimension finie est  $(p, q)$ -régulier, son spectre est invariant par la symétrie  $z \mapsto \sigma - z$ , autrement dit, pour toute valeur propre  $z \in \overline{\mathbb{K}}$  de  $u$  (nécessairement,  $z$  ne peut être la somme d'une racine de  $p$  et d'une racine de  $q$ ),  $\sigma - z$  est valeur propre de  $u$  et présente la même structure de blocs de Jordan (mieux, lorsque  $z = \sigma - z$ , le point (a) du théorème 8.6 permet de voir que  $u$  ne présente que des blocs de Jordan de taille paire pour la valeur propre  $z$ ). Ce phénomène de symétrie parfaite disparaît pour un endomorphisme  $(p, q)$ -exceptionnel.

Supposons  $p$  et  $q$  scindés sur  $\mathbb{K}$  et scindons-les :  $p = (X - x_1)(X - x_2)$  et  $q = (X - y_1)(X - y_2)$ . La matrice  $(x_1 + y_1).I_n$  est alors évidemment une  $(p, q)$ -somme exceptionnelle dans  $M_n(\mathbb{K})$ , mais hors du cas particulier où  $x_1 + y_1 = x_2 + y_2$ , son spectre n'est pas invariant par la symétrie  $z \mapsto \sigma - z$ . Il subsiste cependant, comme nous allons le voir, une forme de symétrie *imparfaite*. Pour cela, il faut examiner de plus près les blocs de Jordan. Le procédé suivant, qui synthétise des observations vues dans [Har90a, Wan95b], permet de construire des matrices à structure de Jordan assez simples et qui sont des  $(p, q)$ -sommages exceptionnelles.

Fixons un entier naturel  $n > 0$ . On prend les matrices blocs  $K = \begin{bmatrix} x_1 & 0 \\ 1 & x_2 \end{bmatrix}$  et  $L = \begin{bmatrix} y_2 & 0 \\ 1 & y_1 \end{bmatrix}$ , et on observe (Cayley-Hamilton) que  $p(K) = 0 = q(L)$ . On pose alors, si  $n$  est pair,

$$A := \underbrace{K \oplus \cdots \oplus K}_{\frac{n}{2} \text{ copies}} \quad \text{et} \quad B := [y_1] \oplus \underbrace{L \oplus \cdots \oplus L}_{\frac{n}{2} - 1 \text{ copies}} \oplus [y_2],$$

et  $n$  est impair,

$$A := \underbrace{K \oplus \cdots \oplus K}_{\frac{n-1}{2} \text{ copies}} \oplus [x_1] \quad \text{et} \quad B := [y_1] \oplus \underbrace{L \oplus \cdots \oplus L}_{\frac{n-1}{2} \text{ copies}}.$$

Bien sûr,  $p(A) = 0 = q(B)$  dans tous les cas, si bien que  $A + B$  est une  $(p, q)$ -somme. En outre, on observe que  $A + B$  est triangulaire inférieure, de coefficients diagonaux  $x_1 + y_1, x_2 + y_2, x_1 + y_1, \dots$ . Enfin, c'est une matrice de Hessenberg régulière, autrement tous les coefficients de sa première sous-diagonale sont non nuls, et ceux situés en dessous sont nuls. Classiquement, il vient de cette dernière observation que  $A + B$  est cyclique, puis on déduit de la première que le polynôme minimal de  $A + B$  est  $(X - x_1 - y_1)^{s+\varepsilon}(X - x_2 - y_2)^s$  où  $s$  est le quotient dans la division euclidienne de  $n$  par 2, et  $\varepsilon$  le reste. Si  $x_1 + y_1 = x_2 + y_2$ , autrement dit si  $x_1 + y_1$  est fixé par  $z \mapsto \sigma - z$ , on parvient ainsi à montrer que tout bloc de Jordan pour la valeur propre  $x_1 + y_1$  est une  $(p, q)$ -somme. Dans ce cas, aucune contrainte n'est à attendre sur les blocs de Jordan d'une  $(p, q)$ -somme exceptionnelle pour la valeur propre  $x_1 + y_1$ . Si en revanche  $x_1 + y_1 \neq x_2 + y_2$ , alors  $A + B$  est semblable à la somme directe de deux blocs de Jordan de tailles respectives  $s + \varepsilon$  et  $s$  pour les valeurs propres  $x_1 + y_1$  et  $x_2 + y_2$ . Lorsque  $n = 1$  on reconnaît un cas banal de la remarque initiale de ce paragraphe, à savoir que l'on peut créer une  $(p, q)$ -somme diagonalisable avec pour seule valeur propre  $x_1 + y_1$ . Plus généralement cette construction permet donc, pour tout  $z \in \mathcal{Z}(p) + \mathcal{Z}(q)$  différent de  $\sigma - z$ , de créer une  $(p, q)$ -somme qui a exactement deux blocs de

Jordan, pour les valeurs propres  $z$  et  $\sigma - z$ , et dont les tailles respectives diffèrent d'au plus une unité. Dans le cas très particulier où  $y_1 = y_2$  et  $x_1 \neq x_2$ , on peut même aller plus loin en modifiant, dans le cas  $n$  pair, le dernier coefficient de  $B$  pour le prendre égal à  $y_1$ , et on observe alors que  $A + B$  possède deux blocs de Jordan, de tailles respectives  $\frac{n}{2} + 1$  et  $\frac{n}{2} - 1$ , pour les valeurs propres  $x_1 + y_1$  et  $x_2 + y_2$ . Dans ce cas, ces tailles diffèrent d'au plus deux unités!

Il est remarquable que la construction précédente, somme toute assez rudimentaire, fournit essentiellement toutes les solutions lorsque  $p$  et  $q$  sont scindés : par « toutes », on entend toutes les solutions *indécomposables au sens des*  $(p, q)$ -*sommes*. Ainsi, toute  $(p, q)$ -somme exceptionnelle est somme directe d'endomorphismes représentés par une matrice d'un des types précédents (bien sûr, moyennant une permutation des racines de  $p$  et  $q$ ). Il est néanmoins utile de faire remarquer que la démonstration de ce principe passe systématiquement par un stade intermédiaire dans lequel, à partir d'une  $(p, q)$ -somme, on cherche des résultats quantitatifs sur les nombres de blocs de Jordan associés aux valeurs propres  $z$  et  $\sigma - z$ . Notons, pour clarifier le discours,  $j_{z,k}(u)$  le nombre de blocs de Jordan de taille *au moins*  $k$  pour la valeur propre  $z$  dans la réduite de  $u$ . On adoptera aussi la terminologie suivante : pour un entier  $p \geq 1$ , deux suites décroissantes  $(u_k)_{k \geq 1}$  et  $(v_k)_{k \geq 1}$  sont dites  *$p$ -entrelacées* lorsque  $u_k \geq v_{k+p}$  et  $v_k \geq u_{k+p}$  pour tout  $k \geq 1$ .

On démontre alors les deux résultats suivants :

**Proposition 8.7.** *Supposons  $p$  et  $q$  scindés sur  $\mathbb{K}$ . Soit  $u \in \mathcal{L}(E)$  une  $(p, q)$ -somme, et  $z \in \mathcal{Z}(p) + \mathcal{Z}(q)$  tel que  $z \neq \sigma - z$ .*

- (i) *Si  $p$  et  $q$  sont scindés à racines simples, alors  $(j_{z,k}(u))_{k \geq 1}$  et  $(j_{\sigma-z,k}(u))_{k \geq 1}$  sont 1-entrelacées.*
- (ii) *Sinon,  $(j_{z,k}(u))_{k \geq 1}$  et  $(j_{\sigma-z,k}(u))_{k \geq 1}$  sont 2-entrelacées.*

Dans [Har90a, Wan95b], ces résultats critiques sont obtenus par une analyse globale fondée sur des calculs par blocs assez techniques. Pour la conclusion (i), on se réduit à une situation connue où intervient la notion de 1-entrelacement, qui est l'analyse des couples de matrices de la forme  $(AB, BA)$  avec  $A$  et  $B$  rectangulaires<sup>2</sup>. Il est en effet connu que pour un tel couple formé de matrices nilpotentes, les suites de nombres de Jordan associées pour la valeur propre 0 sont 1-entrelacées. Et Wang [Wan95b] a démontré un résultat intermédiaire *ad hoc* similaire pour traiter le point (ii) de la proposition précédente.

Une des petites innovations de mes travaux sur la question a été de simplifier les démonstrations précédentes par une analyse locale fine du problème. Plutôt que de traiter globalement le problème, on le réduit à une situation plus élémentaire. Prenons par exemple le point (i), fixons un entier  $k \geq 1$  et cherchons à démontrer que  $j_{z,k}(u) \geq j_{\sigma-z,k+1}(u)$ . Pour cela, ma technique consiste à ramener la situation au cas où  $u$  n'a que des blocs de Jordan de taille 1 et 2, où l'on cherche à établir que  $j_{z,1}(u) \geq j_{\sigma-z,2}(u)$ . Pour se ramener à cette situation, on utilise le lemme de commutation pour obtenir que l'endomorphisme  $\bar{u}$  induit par  $u$  sur  $\text{Ker}(u^2 - \sigma u - (z^2 - \sigma z))^{k+1} / \text{Ker}(u^2 - \sigma u - (z^2 - \sigma z))^{k-1}$  reste une  $(p, q)$ -somme. Or il est facile d'observer que  $\bar{u}$  est annulé par  $(X - z)^2(X - (\sigma - z))^2$  et qu'il présente exactement  $j_{k,z}(u)$  blocs de Jordan de taille au moins 1 pour la valeur propre  $z$ , et  $j_{k+1,\sigma-z}(u)$  blocs de Jordan de taille au moins 2 pour

<sup>2</sup> H. Flanders, Elementary divisors of  $AB$  and  $BA$ . Proc. Amer. Math. Soc. **2-6** (1951) 871–874.

la valeur propre  $\sigma - z$ . Ainsi, il suffira de montrer le résultat réduit pour  $\bar{u}$ . Dans [1], cette technique est utilisée pour l'analyse des combinaisons linéaires de deux idempotents à coefficients imposés, par une analyse par blocs assez originale. Dans [4], j'ai réalisé une fusion entre ma technique d'analyse locale et la méthode par blocs de Wang [Wan95b] (et en appendice, j'ai repris la technique pour donner une démonstration encore plus simple du résultat figurant dans [1]), ce qui simplifie les calculs de Wang et permet d'éviter les problèmes liés à la caractéristique 2. Et j'ai repris cette dernière méthode dans [13] pour analyser les  $(p, q)$ -produits exceptionnels. On notera qu'il s'agit de progrès somme toute mineurs, mais la prise de conscience de la possibilité de simplifier radicalement l'analyse de ces cas s'est révélée indispensable pour le cas irréductible.

Finalement, les conditions d'entrelacement permettent assez facilement de regrouper, dans la réduite de Jordan de  $u$ , les blocs de Jordan par paires. La démonstration se fait par récurrence sur la dimension, en commençant par extraire les blocs de plus grande taille, et ne mérite pas de commentaire particulier. Concluons pour le cas scindé : la symétrie parfaite qui prévalait pour les  $(p, q)$ -sommets réguliers se meut en une symétrie imparfaite, toujours pour  $z \mapsto \sigma - z$ , symétrie se traduisant par une forte proximité des réduites de Jordan associées aux valeurs propres respectives  $z$  et  $\sigma - z$ .

Examinons maintenant les cas où l'un des polynômes est irréductible. Celui où exactement l'un est irréductible, mettons  $p$ , n'est pas très difficile. Là, on adapte assez facilement la construction des matrices blocs particulières vue plus haut (mais uniquement avec  $n$  pair). Quant aux conditions d'entrelacements, elles s'obtiennent par un simple argument d'extension des scalaires au corps de décomposition de  $p$ .

La première véritable grande difficulté intervient lorsque  $p$  et  $q$  sont irréductibles avec le même corps de décomposition dans  $\bar{\mathbb{K}}$ . Scindons  $p = (X - x_1)(X - x_2)$  et  $q = (X - y_1)(X - y_2)$  dans ce corps de décomposition. Il y a essentiellement deux situations possibles pour la valeur propre  $x_1 + y_1$  : elle est dans  $\mathbb{K}$  si et seulement si elle est égale à sa symétrique  $x_2 + y_2$ , ce qui peut parfaitement se produire (par exemple si  $p = q = X^2 + 1$  sur le corps des réels, et  $x_1 = -y_1 = i$ ), et se produit si et seulement si  $q = p(d - X)$  pour  $d := x_1 + y_1$ . Des contraintes fort différentes apparaissent selon que  $x_1 + y_1$  est dans  $\mathbb{K}$  ou non.

- Si  $x_1 + y_1$  est dans  $\mathbb{K}$ , on peut démontrer que, pour tout entier  $k \geq 1$ , une  $(p, q)$ -somme exceptionnelle doit avoir un nombre *pair* de blocs de Jordan de taille  $k$  pour la valeur propre  $x_1 + y_1$ . Cela découle essentiellement de deux observations : d'une part, lorsque  $q = p(-X)$ , pour toute  $(p, q)$ -somme  $u = a + b$  (où  $p(a) = q(b) = 0$ ) les noyaux itérés  $\text{Ker } u^k$  sont tous stables par  $a$  et  $b$ ; d'autre part, un endomorphisme ne peut être annulé par  $p$  que si son espace vectoriel de départ est de dimension paire.
- Si en revanche  $z := x_1 + y_1$  n'est pas dans  $\mathbb{K}$ , le polynôme  $(X - z)(X - (\sigma - z))$  est irréductible sur  $\mathbb{K}$ , si bien que les valeurs propres  $z$  et  $\sigma - z$  d'une  $(p, q)$ -somme ont exactement les mêmes structures de Jordan associées ! Ainsi, dans ce cas une extension des scalaires au corps de décomposition de  $p$  ne donne aucune information dépassant cette observation banale. Et pourtant il existe des contraintes sur les nombres de blocs de Jordan généralisés associés à l'irréductible  $(X - z)(X - (\sigma - z))$ , sous une forme qui ressemble très fortement à ce qui a été vu dans le cas scindé. En effet, on peut démontrer que, pour une  $(p, q)$ -somme  $u$ , la réduite associée à

l'irréductible  $R := (X - z)(X - (\sigma - z))$  doit pouvoir se décomposer en somme directe de matrices, elles-mêmes toutes de la forme  $C(R^k) \oplus C(R^l)$ , où  $|k - l| \leq 1$  ! C'était exactement ce qui se présentait (mais avec deux valeurs propres distinctes) dans le cas scindé. Et la démonstration ne peut pas s'appuyer sur une simple réduction au cas scindé. En outre, elle est très subtile bien qu'elle repose à nouveau sur un principe de localisation (réduction au cas où  $(X - z)^2(X - (\sigma - z))^2$  annule  $u$ ). C'est le lemme suivant qui emporte essentiellement le morceau pour établir cette forme de symétrie *interne*.

**Lemme 8.8** (Lemme 4.1 dans [12]). *Supposons  $p$  et  $q$  irréductibles, de même corps de décomposition dans  $\overline{\mathbb{K}}$ . Scindons  $p = (X - x_1)(X - x_2)$  et  $q = (X - y_1)(X - y_2)$  dans  $\overline{\mathbb{K}}[X]$ . Soit  $a$  et  $b$  des endomorphismes d'un même espace vectoriel de dimension finie, supposés respectivement  $p$ -quadratique et  $q$ -quadratique. Notons  $a^*$  le  $p$ -conjugué de  $a$ , et  $b^*$  le  $q$ -conjugué de  $b$ , et considérons le produit croisé  $w := ab^* + ba^*$  et le scalaire  $z := x_1y_2 + x_2y_1$  (qui est dans  $\mathbb{K}$ ). Pour tout  $k \in \mathbb{N}^*$ , si  $w$  ne possède pas de bloc de Jordan de taille  $k$  pour la valeur propre  $z$ , alors  $j_{z, k+1}(w)$  est multiple de 4.*

Terminons par le cas où  $p$  et  $q$  sont irréductibles avec des corps de décomposition distincts dans  $\overline{\mathbb{K}}$ . Dans ce cas, et sauf dans un cas très particulier (celui où  $p$  est séparable et  $q$  est inséparable), les méthodes consistent pour l'essentiel en une réduction aux cas précédents par diverses techniques d'extension des scalaires. Soit on étend brutalement le corps des scalaires, soit on applique la méthode plus fine que nous allons maintenant décrire. Soit  $u$  une  $(p, q)$ -somme exceptionnelle, que l'on décompose  $u = a + b$  avec  $a, b$  annihilés respectivement par  $p$  et  $q$ . On sait que  $a$  et  $b$  commutent avec  $v := u^2 - \sigma u$ , tandis que  $v$  est annihilé par une puissance du polynôme  $\Lambda_{p,q}$ . Dans les cas favorables, ce dernier est séparable, si bien que  $v$  possède une décomposition de Jordan-Chevalley  $v = s + n$ , où  $s$  est annihilé par  $\Lambda_{p,q}$ ,  $n$  est nilpotent, et  $s \in \mathbb{K}[v]$ . Par suite,  $a$  et  $b$  commutent avec  $s$ , tandis que  $s$  induit une structure de  $\mathbb{K}[s]$ -espace vectoriel sur l'espace de départ  $E$  de  $u$ , et l'on peut voir que  $u$  reste une  $(p, q)$ -somme dans  $\mathcal{L}_{\mathbb{K}[s]}(E)$ . Cela permet pratiquement une réduction à l'un des cas déjà traités.

On notera enfin que le phénomène de symétrie interne observé dans le cas où  $p$  et  $q$  sont irréductibles avec même corps de décomposition se retrouve sous une forme très voisine lorsque  $p$  et  $q$  sont irréductibles avec des corps de décomposition distincts et des discriminants distincts. Dans ce cas, on démontre que  $F_{p,q}$  est irréductible et que les  $(p, q)$ -sommes exceptionnelles sont les endomorphismes pouvant être représentés par une matrice diagonale par blocs dont chaque bloc diagonal est de la forme  $C((F_{p,q})^k) \oplus C((F_{p,q})^l)$  avec des entiers naturels  $k, l$  (éventuellement nuls) tels que  $|k - l| \leq 1$ . On notera qu'ici les quatre éléments de  $\mathcal{Z}(p) + \mathcal{Z}(q)$  sont indistinguables en tant que racines de  $F_{p,q}$ .

### 8.5.5 Le problème des produits

Nous avons à ce stade considéré uniquement le problème des  $(p, q)$ -sommés. De manière remarquable, les méthodes précédentes s'adaptent naturellement aux  $(p, q)$ -produits à condition que  $p$  et  $q$  soient *non dégénérés*, autrement dit  $p(0)q(0) \neq 0$ . Le cas dégénéré est très différent du cas non dégénéré, et nous n'en dirons rien ici.



Supposons donc  $(pq)(0) \neq 0$  à partir de maintenant. Tout  $(p, q)$ -produit d'une algèbre est alors inversible. Dans le cas  $\mathcal{A} = \mathcal{L}(E)$  avec  $E$  de dimension finie, on s'intéresse donc uniquement aux automorphismes. On peut alors adapter les outils précédents :

- (i) L'ensemble pertinent des valeurs propres « exceptionnelles » est maintenant  $\mathcal{Z}(p)\mathcal{Z}(q)$  (formé des éléments de  $\overline{\mathbb{K}}$  qui sont produit d'une racine de  $p$  par une racine de  $q$ ).
- (ii) Le scalaire  $\pi := p(0)q(0)$ , produit des racines de  $pq$ , joue le rôle de  $\sigma$ .
- (iii) Un automorphisme  $u$  d'un  $\mathbb{K}$ -espace vectoriel de dimension finie est dit  $(p, q)$ -régulier lorsqu'il n'a aucune valeur propre dans  $\mathcal{Z}(p)\mathcal{Z}(q)$ , et  $(p, q)$ -exceptionnel lorsqu'il a toutes ses valeurs propres dans  $\mathcal{Z}(p)\mathcal{Z}(q)$  (par valeur propre, on entend ici une racine du polynôme caractéristique dans  $\overline{\mathbb{K}}$ ).
- (iv) L'inversion  $z \in \overline{\mathbb{K}} \setminus \{0\} \mapsto \frac{\pi}{z}$  joue le rôle de la symétrie  $z \mapsto \sigma - z$ .
- (v) Le polynôme fondamental associé au couple  $(p, q)$  est maintenant  $F_{p,q} := \prod_{1 \leq i, j \leq n} (X - x_i y_j)$  si l'on scinde  $p = (X - x_1)(X - x_2)$  et  $q = (X - y_1)(X - y_2)$  sur  $\overline{\mathbb{K}}$ .
- (vi) Étant donné un  $(p, q)$ -produit  $x = ab$ , où  $p(a) = q(b) = 0$ , son pseudo-adjoint est naturellement défini comme  $x^* := b^* a^*$ , et il vérifie  $xx^* = x^*x = \pi 1_{\mathcal{A}}$ , tandis que sa pseudo-trace est  $x + x^* = a(b^*)^* + b^* a^*$ . La pseudo-trace commute avec  $a$  et  $b^*$  donc aussi avec  $b$ , et elle vaut aussi  $x + \pi x^{-1}$ .
- (vii) Le polynôme fondamental  $F_{p,q}$  s'écrit aussi  $R_\pi(\Lambda_{p,q})$  pour  $\Lambda_{p,q} := (X - (x_1 y_1 + x_2 y_2))(X - (x_1 y_2 + x_2 y_1))$ , qui se trouve être à coefficients dans  $\mathbb{K}$ , et la transformation  $R_\pi$  des polynômes définie par  $R_\pi(S) = X^d S(X + \pi X^{-1})$  pour tout polynôme  $S$  de degré  $d$ . On en déduit une notion correspondante de partie  $(p, q)$ -régulière et de partie  $(p, q)$ -exceptionnelle pour un endomorphisme d'un  $\mathbb{K}$ -espace vectoriel de dimension finie.

On notera que lorsque  $\pi = 1$  (ce qui arrive par exemple dans le cas remarquable des produits de deux symétries, où  $p = q = X^2 - 1$ ), les polynômes de la forme  $R_\pi(S)$  avec  $S \neq 0$  sont exactement les palindromes de degré pair ! On appellera plus généralement  $\pi$ -palindrome tout polynôme de la forme  $R_\pi(S)$  avec  $S$  non nul (on peut montrer qu'un polynôme  $p = \sum_{k=0}^n a_k X^k$  de degré  $n$  est un  $\pi$ -palindrome si et seulement si  $n$  est pair et  $a_{n-k} = \pi^{n/2-k} a_k$  pour tout  $k \in \llbracket 0, n \rrbracket$ ).

Voyons par exemple comment s'adapte la classification des  $(p, q)$ -produits réguliers :

**Théorème 8.9** (Théorème 4.10 de [11]). *Soit  $u$  un automorphisme  $(p, q)$ -régulier (au sens des produits) d'un espace vectoriel de dimension finie.*

- (a) *Pour que  $u$  soit un  $(p, q)$ -produit, il est nécessaire que chacun de ses invariants de similitude soit un  $\pi$ -palindrome.*
- (b) *On suppose que les invariants de similitude de  $u$  sont des  $\pi$ -palindromes. Pour tout polynôme irréductible unitaire  $S \in \mathbb{K}[X]$  et tout entier  $k \geq 1$ , notons  $n_{S,k}(u)$  le nombre d'invariants de similitude de  $u$  de la forme  $R_\pi(R)$  où  $R$  est de valuation  $k$  par rapport à l'irréductible  $S$ . Pour que  $u$  soit un*

$(p, q)$ -produit, il est alors nécessaire et suffisant que le nombre  $n_{S,k}(u)$  soit pair pour tout entier  $k \geq 1$  et tout irréductible unitaire  $S \in \mathbb{K}[X]$  tel que  $R_\pi(S)$  n'ait pas de racine dans  $\mathcal{Z}(p)\mathcal{Z}(q)$  et tel que la norme de l'algèbre  $\mathcal{W}_D(p, q, x)$  soit anisotrope pour  $D := \mathbb{K}[X]/(S)$  et  $x$  la classe de  $q(0)^{-1}X$  modulo  $R$ .

Enfin, l'analyse du cas exceptionnel se réalise avec des méthodes identiques à celles déployées pour l'étude des  $(p, q)$ -sommes exceptionnelles, et certains résultats fondamentaux sont mêmes communs aux deux situations. Par exemple, le lemme 8.8 intervient dans les deux études.

## 8.6 Le problème à trois termes en dimension finie

Si l'essentiel de mes travaux récents a porté sur les décompositions à deux termes, je ne me suis pas limité à celles-ci. Mes premiers travaux ont en effet concerné les décompositions en combinaisons linéaires de trois matrices idempotentes [2], ainsi qu'aux décompositions en sommes d'idempotentes sur un corps de caractéristique non nulle [3] (voir la section 8.7). Ultérieurement, j'ai repris mes méthodes afin de travailler sur les matrices « augmentées », travaux motivés par l'étude en dimension infinie (voir la section 8.8).

### 8.6.1 Décompositions en sommes ou combinaisons linéaires à trois termes

Pour fixer les idées, je vais ici commencer par évoquer la question de la décomposition d'une matrice carrée en somme de matrices de carré nul, ce qui est l'un des problèmes les plus simples. Évidemment, une telle matrice doit être de trace nulle. Réciproquement, Wang et Wu [Wan91a] ont montré que toute matrice de trace nulle est somme de quatre matrices de carré nul (sur le corps des complexes). La généralisation à un corps quelconque figure dans [5]. Ce résultat fait écho au spectaculaire théorème de Gustafson, Halmos et Radjavi [Gus76], qui stipule que toute matrice de déterminant  $\pm 1$  est produit de quatre matrices de symétrie (i.e. de quatre involutions).

Rappelons aussi, à toutes fins utiles, le résultat de Wang, Wu et Botha :

**Théorème 8.10.** *Une matrice  $M \in M_n(\mathbb{K})$  est somme de deux matrices de carré nul si et seulement si ses invariants de similitude sont tous pairs ou impairs.*

En particulier, toute matrice nilpotente est somme de deux matrices de carré nul, et en caractéristique différente de 2 les sommes de deux matrices de carré nul sont les matrices semblables à leur opposée.

Une matrice de trace nulle est toujours somme de quatre matrices de carré nul, mais pas nécessairement de trois matrices de carré nul. Pour le voir, supposons  $\chi(\mathbb{K}) \neq 2$  et considérons  $\alpha \in \mathbb{K} \setminus \{0\}$  ainsi qu'une matrice  $M \in M_n(\mathbb{K})$  telle que  $\dim \text{Ker}(M - \alpha I_n) > \frac{3n}{4}$  et  $\text{tr } M = 0$  (une telle matrice existe évidemment si  $n > 4$ , prendre par exemple  $M = \alpha I_{n-1} \oplus -(n-1)\alpha I_1$ ). Supposons que  $M = A + B + C$  pour des matrices  $A, B, C$  de carré nul. On aurait  $(M - A - \alpha I)^2 = (M - \alpha I)((M - \alpha I) - A) - A(M - \alpha I)$  donc  $\text{rg}(M - A - \alpha I)^2 < \frac{n}{2}$ ,

si bien que  $\alpha$  serait valeur propre de  $M - A$  de multiplicité (algébrique) strictement supérieure à  $\frac{n}{2}$  ; par ailleurs  $M - A$  serait semblable à son opposée puisque somme de deux matrices de carré nul, donc  $-\alpha$  serait aussi valeur propre de  $M - A$  de multiplicité algébrique strictement supérieure à  $\frac{n}{2}$ , ce qui est impossible puisque  $\alpha \neq -\alpha$ . L'argument qui précède simplifie assez substantiellement celui proposé par Wang et Wu dans [Wan91a].

Évoquons maintenant la démonstration du fait que toute matrice  $M$  de trace nulle est somme de quatre matrices de carré nul. La preuve la plus simple connue (voir l'appendice de [5]) est la suivante : on suppose d'abord  $M$  non scalaire. Dans ce cas, un théorème classique de Fillmore<sup>3</sup> indique que  $M$  est semblable à une matrice  $M'$  de diagonale nulle. On peut ensuite casser  $M'$  en  $M'_1 + M'_2$  où  $M'_1$  est triangulaire inférieure stricte et  $M'_2$  est triangulaire supérieure stricte. Ainsi,  $M'_1$  et  $M'_2$  sont toutes deux nilpotentes donc sommes de deux matrices de carré nul. Dans le cadre traité par Wang et Wu, qui est celui du corps  $\mathbb{C}$ , il ne resterait qu'à traiter le cas la matrice nulle . . . Cependant, en caractéristique  $p > 0$ , il reste quand même à démontrer que  $I_p$  est somme de quatre matrices de carré nul. Cela revient à produire une matrice  $N \in M_p(\mathbb{K})$  telle que  $N$  et  $I_p + N$  soient toutes deux sommes de deux matrices de carré nul. Une telle matrice  $N$  devra avoir son spectre invariant par  $z \mapsto -z$  et par  $z \mapsto 2 - z$ , donc par la composée  $z \mapsto z + 2$ . Il apparaît alors naturel de prendre pour  $N$  une matrice diagonale ayant pour coefficients diagonaux tous les éléments du sous-corps premier de  $\mathbb{K}$ . La caractérisation des sommes de deux matrices de carré nul montre alors que  $N$  et  $I_p + N$  sont de telles matrices, et ainsi  $I_p$  est somme de quatre matrices de carré nul.

Je vais maintenant évoquer une variante de démonstration pour le cas d'une matrice non scalaire. Les idées introduites pour cette variante vont se révéler extrêmement fructueuses. Mon point de départ est un lemme que l'on trouve déjà dans la démonstration originale de Wang et Wu. Pour simplifier le discours, nous dirons désormais que deux matrices carrées (de même format)  $A$  et  $B$  sont **n-adjacentes** lorsque  $A$  est somme d'une matrice semblable à  $B$  et d'une matrice de carré nul. On notera que c'est une relation symétrique et réflexive, et que toute matrice n-adjacente à la somme de  $k$  matrices de carré nul est la somme de  $k + 1$  matrices de carré nul.

**Lemme 8.11** (Lemme d'adaptation cyclique). *Soit  $M \in M_n(\mathbb{K})$  une matrice cyclique et  $P \in \mathbb{K}[X]$  unitaire de degré  $n$  tel que  $\text{tr } P = \text{tr } M$ . Alors  $M$  est n-adjacente à  $C(P)$ .*

Pour le démontrer, on se ramène immédiatement au cas où  $M$  est une matrice compagnon  $C(Q)$  (pour un certain polynôme  $Q$  unitaire de degré  $n$ ), on fixe un vecteur arbitraire  $Y \in \mathbb{K}^{n-1}$  et on considère la matrice  $A_Y := \begin{bmatrix} [0]_{n-1} & Y \\ [0]_{1 \times (n-1)} & 0 \end{bmatrix}$ . La matrice  $M - A_Y$  est alors évidemment une matrice compagnon, et en ajustant  $Y$  on peut imposer tous les coefficients de son polynôme caractéristique devant les exposants inférieurs à  $n-2$ . Puisque  $\text{tr } P = \text{tr } M$ , on peut ajuster ce polynôme caractéristique de telle sorte qu'il vaille  $P$ .

Pour généraliser cette idée, nous allons travailler plus largement sur des matrices de **Hessenberg**. Rappelons qu'il s'agit des matrices  $M \in M_n(\mathbb{K})$  telles

3. P.A. Fillmore, On similarity and the diagonal of a matrix. Amer. Math. Monthly **76** (1969) 167-169.

que  $m_{i,j} = 0$  dès que  $i > j + 1$  (coefficients nuls sous la première diagonale), et qu'une telle matrice est dite **régulière** lorsqu'en outre  $m_{i+1,i} \neq 0$  pour tout  $i \in \llbracket 1, n-1 \rrbracket$ . Toute matrice de Hessenberg régulière est cyclique avec pour vecteur cyclique le premier vecteur de la base canonique. Voici une sorte de généralisation du lemme précédent :

**Lemme 8.12** (Lemme d'adaptation cyclique généralisé). *Soit  $M \in M_n(\mathbb{K})$  une matrice de Hessenberg régulière, et  $P \in \mathbb{K}[X]$  unitaire de degré  $n$  tel que  $\text{tr } P = \text{tr } M$ . Il existe alors  $Y \in \mathbb{K}^{n-1}$  tel que  $M + \begin{bmatrix} [0]_{n-1} & Y \\ [0]_{1 \times (n-1)} & 0 \end{bmatrix}$  soit cyclique de polynôme minimal  $P$ .*

On notera que, quel que soit le choix de  $Y$ , la matrice  $M + \begin{bmatrix} [0]_{n-1} & Y \\ [0]_{1 \times (n-1)} & 0 \end{bmatrix}$  reste de Hessenberg régulière, et à ce titre est cyclique. Il s'agit donc de démontrer qu'un choix judicieux de  $Y$  permet d'ajuster le polynôme caractéristique de la somme, ce qui est facile à l'aide d'un développement du déterminant caractéristique selon la dernière colonne (le cofacteur de la matrice caractéristique  $XI_n - M$  en position  $(i, n)$  est un polynôme de degré  $i - 1$ ).

Prenons maintenant  $M \in M_n(\mathbb{K})$  non scalaire de trace nulle. Après réduction de Frobenius, on peut se ramener au cas où  $M$  est diagonale par blocs  $M = M_1 \oplus \dots \oplus M_r$  dont les blocs diagonaux sont des matrices compagnons et le dernier bloc diagonal est de taille au moins 2 (prendre la matrice compagnon du polynôme minimal de  $M$ ). La matrice  $M$  est alors de Hessenberg (mais non régulière en général). L'astuce va alors être de rajouter à  $M$  deux matrices de carré nul pour à la fois la transformer en une matrice de Hessenberg régulière et imposer son polynôme caractéristique (à l'exception de la trace de ce dernier, qui devra manifestement être nulle). Pour cela, on va utiliser un principe de *double-connexion*. On se représente  $M$  comme suit

$$M = \begin{bmatrix} M_1 & (?) & ? \\ 0 & M_2 & (?) \\ & \ddots & \ddots & ? \\ (0) & & 0 & M_r \end{bmatrix}$$

où chaque bloc  $M_i$  est une matrice de Hessenberg régulière, et la dernière matrice  $M_r$  est de taille au moins 2. Il nous faut parvenir à remplacer les termes nuls de la première sous-diagonale par des termes non nuls. Notons donc  $M'$  la matrice obtenue à partir de  $M$  en remplaçant tous les termes nuls de sa première sous-diagonale par des 1, et en conservant tous les autres termes. Plus généralement, fixons  $Y \in \mathbb{K}^{n-1}$  et notons  $M'_Y$  la matrice obtenue à partir de  $M'$  en rajoutant le vecteur  $\begin{bmatrix} Y \\ 0 \end{bmatrix}$  à la dernière colonne. D'après le lemme d'adaptation généralisé, comme  $\text{tr } M' = \text{tr } M = 0$ , le vecteur  $Y$  peut être ajusté de telle sorte que le polynôme caractéristique de  $M'_Y$  soit égal à  $X^n$ . On observe alors que la différence  $M'_Y - M$  peut systématiquement être décomposée comme la somme de deux matrices  $A$  et  $B$  de carré nul. Pour cela, l'idée est de placer alternativement les coefficients manquants de la première sous-diagonale dans l'une des matrices, puis dans l'autre, et de procéder à une séparation inversée pour les coefficients de la dernière colonne. Soyons plus précis : on note  $I := \{i \in \llbracket 1, n-1 \rrbracket : m_{i+1,i} = 0\}$  et  $p$  son cardinal, et on prend la bijection croissante  $f : \llbracket 1, p \rrbracket \xrightarrow{\cong} I$ . On définit :

- la matrice  $A \in M_n(\mathbb{K})$  par  $a_{f(2i)+1, f(2i)} = -1$  pour tout  $i \in \mathbb{N}^*$  tel que  $2i \leq p$ ,  $a_{k,n} = y_k$  pour tout  $k \in \llbracket 1, n \rrbracket \setminus \{f(2i) \mid 1 \leq i \leq \lfloor p/2 \rfloor\}$  et tous les autres coefficients sont nuls ;
- la matrice  $B \in M_n(\mathbb{K})$  par  $b_{f(2i+1)+1, f(2i+1)} = -1$  pour tout  $i \in \mathbb{N}$  tel que  $2i+1 \leq p$ ,  $b_{k,n} = y_k$  pour tout  $k \in \{f(2i) \mid 1 \leq i \leq \lfloor p/2 \rfloor\}$  et tous les autres coefficients sont nuls.

Il est bien clair alors que  $A+B = M'_Y - M$ , et il reste à vérifier que  $A^2 = 0 = B^2$ . Ce dernier point se contrôle facilement, le point critique étant le fait que les deux dernières colonnes de  $A^2$  et  $B^2$  soient nulles, ce qui découle de l'inégalité  $f(p) < n-1$  (c'est là qu'intervient le fait que le dernier bloc diagonal de  $M$  est de taille supérieure à 2).

On notera que la même méthode permet d'écrire  $(J_n)^T$  comme la somme de deux matrices de carré nul (on prend ici  $Y = 0$ , la condition  $f(p) < n-1$  n'est alors plus remplie mais on peut s'en passer fort bien dans ce cas particulier). Et donc, si  $Y$  est ajustée de telle sorte que  $M'_Y$  soit de polynôme caractéristique  $X^n$ , cette matrice est semblable à  $(J_n)^T$  et est donc somme de deux matrices de carré nul. En définitive, pour un tel choix de  $Y$ , on voit que  $M = (M - M'_Y) + M'_Y$  est somme de quatre matrices de carré nul.

Nous reverrons cette technique de double-connexion dans la synthèse de nos travaux sur les décompositions en dimension infinie (section 8.8) ainsi que dans celle de nos travaux sur les sommes de projecteurs en caractéristique non nulle (section 8.7). Je vais maintenant exposer une variante de cette technique, variante dans laquelle on ne s'autorise qu'une seule connexion. On comprend immédiatement que, si l'on se limite à un seul élément de carré nul, il faut déjà imposer à tous les blocs intermédiaires  $M_2, \dots, M_{r-1}$  d'être de taille strictement supérieure à 1. Les blocs extrêmes peuvent quant à eux être de taille 1. Mais à ce stade la difficulté reste d'imposer le polynôme caractéristique après connexion. Pour cela, j'ai introduit la notion de matrice **bien partitionnée**. Une telle matrice est une matrice diagonale par blocs de la forme

$$C(P_1) \oplus \dots \oplus C(P_r) \oplus C(Q_1) \oplus \dots \oplus C(Q_s),$$

où  $r > 0$ ,  $s > 0$ ,  $P_2, \dots, P_r, Q_1, \dots, Q_{s-1}$  sont tous de degré au moins 2, et chaque  $P_i$  est premier avec chaque  $Q_j$ . Une telle matrice sera dite **très bien partitionnée** lorsqu'en outre l'un des polynômes  $P_1$  et  $Q_s$  est de degré au moins 2. Supposons notre matrice  $M$  bien partitionnée, et notons  $A := C(P_1) \oplus \dots \oplus C(P_r)$  et  $B := C(Q_1) \oplus \dots \oplus C(Q_s)$ .

Bien sûr  $M = A \oplus B$ , mais il y a mieux :  $M$  est semblable à n'importe quelle matrice de la forme  $\begin{bmatrix} A & ? \\ 0 & B \end{bmatrix}$ , ce qui découle du fait que  $\chi_A$  et  $\chi_B$  sont premiers entre eux (et peut donc par exemple se déduire d'un célèbre théorème de Roth<sup>4</sup>). Nous allons alors jouer sur ce degré de liberté grâce à une généralisation subtile du lemme d'adaptation :

**Lemme 8.13** (Lemme 11 de [2]). *Soit  $M \in M_n(\mathbb{K})$  une matrice de Hessenberg régulière, et  $p \in \llbracket 1, n-1 \rrbracket$ . On décompose  $M = \begin{bmatrix} A & [?] \\ B & C \end{bmatrix}$ . Soit  $P \in \mathbb{K}[X]$  unitaire de degré  $n$  et de trace  $\text{tr } A + \text{tr } C$ . Il existe alors  $D \in M_{p, n-p}(\mathbb{K})$  telle*

4. W. Roth, The equations  $AX - YB = C$  and  $AX - XB = C$  in matrices. Proc. Amer. Math. Soc. **3** (1952) 392–396.

que  $N := \begin{bmatrix} A & D \\ B & C \end{bmatrix}$  soit cyclique de polynôme minimal  $P$ .

*Idée de démonstration.* En remarquant que  $N$  est elle-même de Hessenberg régulière quel que soit le choix de  $D$ , on est ramené à une question de calcul de polynôme caractéristique, que l'on résout par des techniques d'opérations élémentaires, puis de division euclidienne.  $\square$

Nous allons en déduire le résultat suivant :

**Lemme 8.14** (Lemme d'adaptation pour les matrices bien partitionnées). *Soit  $M \in M_n(\mathbb{K})$  une matrice bien partitionnée, et  $P \in \mathbb{K}[X]$  unitaire de degré  $n$  tel que  $\text{tr } P = \text{tr } M$ . La matrice  $M$  est alors  $n$ -adjacente à  $C(P)$ .*

*Démonstration.* Reprenons donc les notations précédentes pour  $M$ . Prenons d'abord la matrice  $M'$  obtenue à partir de  $M$  en remplaçant tous les termes nuls sur sa première sous-diagonale par des 1. On obtient une matrice de Hessenberg régulière de la forme  $\begin{bmatrix} A' & 0 \\ C' & B' \end{bmatrix}$ , où  $A' \in M_p(\mathbb{K})$  etc. Par le lemme 8.13, on peut

donc choisir une matrice  $D \in M_{p,n-p}(\mathbb{K})$  telle que  $M'' = \begin{bmatrix} A' & D \\ C' & B' \end{bmatrix}$  ait pour polynôme caractéristique  $P$ . Vu la construction de  $A'$  et  $B'$  à partir de  $A$  et  $B$ , on note que  $\begin{bmatrix} A & D \\ 0 & B \end{bmatrix} - M''$  est la matrice dont tous les termes sont nuls, sauf ceux situés aux positions de la première sous-diagonale de  $M$  où il y avait des coefficients nuls, où les coefficients valent  $-1$ . Vu l'hypothèse sur les tailles des blocs intermédiaires de  $M$ , il est alors clair que  $\begin{bmatrix} A & D \\ 0 & B \end{bmatrix} - M''$  est de carré nul.

Ainsi,  $\begin{bmatrix} A & D \\ 0 & B \end{bmatrix}$  est  $n$ -adjacente à  $C(P)$ , donc  $M$  aussi.  $\square$

Si par exemple  $M$  était de trace nulle, cela permettrait de conclure directement que  $M$  est somme de trois matrices de carré nul, en choisissant  $P = X^n$  (de sorte que  $C(P)$  soit nilpotente donc somme de deux matrices de carré nul). Cependant, on a déjà vu qu'il n'est pas vrai que n'importe quelle matrice de trace nulle soit somme de trois matrices de carré nul. Notre principale contribution à la question, permise par la technique des matrices bien partitionnées, est la suivante :

**Théorème 8.15.** *Soit  $M \in M_n(\mathbb{K})$  de trace nulle. Alors  $M \oplus 0_n$  (de taille  $2n$ ) est somme de trois matrices de carré nul dans  $M_{2n}(\mathbb{K})$ .*

On notera que, contrairement à la caractérisation des sommes de deux matrices de carré nul, qui est totalement invariante par extension à une taille supérieure, ce principe ne vaut plus pour les sommes de trois matrices de carré nul. Autrement dit,  $M$  peut fort bien ne pas être la somme de trois matrices de carré nul dans  $M_n(\mathbb{K})$ , mais la matrice augmentée  $M \oplus 0_n$  l'est systématiquement (dans  $M_{2n}(\mathbb{K})$ ) dès que  $\text{tr } M = 0$ .

Nous allons maintenant rentrer plus avant dans la démonstration du théorème 8.15, car elle illustre bien les méthodes de mes trois articles [2, 5, 8], et c'est la plus simple de toutes. Je vais commencer par le cas d'un corps de caractéristique nulle (hypothèse simplifiant la démonstration), puis j'expliquerai brièvement comment adapter les idées aux caractéristiques premières (la démonstration

donnée en [5] n'établit pas cette distinction, au prix d'une complexité artificielle pour le traitement de la caractéristique 0). Supposons donc  $\chi(\mathbb{K}) = 0$ . Dans un premier temps (lemme 3.1 de [5]), on démontre que  $M \oplus 0_n$  est semblable à une matrice de la forme  $N \oplus \alpha I_r \oplus 0_q$  avec  $\alpha$  un scalaire non nul (éventuellement  $r = 0$ ),  $q \geq r$  et  $N$  bien partitionnée ou nilpotente (éventuellement vide). L'idée est de factoriser chaque invariant de similitude de  $M \oplus 0_n$  sous la forme  $X^{n_k} Q_k$  avec  $Q_k(0) \neq 0$  et  $n_k > 0$  (cette dernière inégalité est permise par la présence du bloc  $0_n$  dans l'augmentation) puis d'examiner la suite des  $Q_k$  ainsi obtenue. Le bloc  $\alpha I_r$  s'obtient en prenant le plus grand nombre des  $Q_k$  qui sont de degré 1 (ils s'écrivent alors tous  $X - \alpha$  pour le même scalaire  $\alpha$ ), et tous les autres sont regroupés pour former la deuxième partie de la matrice bien partitionnée, la première partie étant formée par le plus grand nombre possible de termes de la forme  $C(X^{n_k})$  avec  $n_k > 1$ , et exactement un  $C(X)$  s'il n'existe aucun  $n_k$  strictement supérieur à 1 (si tous les  $Q_k$  sont de degré 1, la partie  $N$  sera donc nilpotente).

Il reste à travailler sur la forme précédente. D'abord, par une astuce d'homothétie, et en remarquant que toute matrice nulle est somme de trois matrices de carré nul, on est ramené uniquement au cas de matrices de la forme  $N \oplus I_r \oplus 0_r$  avec  $N$  bien partitionnée ou nilpotente. Il faut alors montrer qu'une telle matrice est nécessairement somme de trois matrices de carré nul, et on va travailler pour cela de manière séparée sur les blocs  $N$  et  $I_r \oplus 0_r$ . Attention, bien évidemment aucun d'entre eux n'est en principe de trace nulle. On va retrancher (ou rajouter, cela revient au même) une matrice de carré nul à chacun d'entre deux afin que la *somme directe* des résultats soit somme de deux matrices de carré nul. Bien sûr, le fait de savoir repérer d'un clin d'oeil une somme de deux matrices de carré nul nous sera bien utile. Là va intervenir la technique des *chaînes de blocs*. Si nous ajoutons à l'idempotent  $P := I_r \oplus 0_r$  une matrice de carré nul, nous obtiendrons une  $(X^2 - 1, X^2)$ -somme, dont le spectre doit donc être invariant par la symétrie  $x \mapsto 1 - x$  à l'exception éventuelle de 0 et 1. L'idée est toute simple : on rajoute à  $P$  une bonne matrice de carré nul (c'est possible!) pour obtenir une matrice diagonalisable  $D$  de valeurs propres  $-(r-1), \dots, 0, 1, \dots, r$ . Pour le voir, on peut tout simplement appliquer le lemme d'adaptation cyclique : pour tout  $x \in \mathbb{Z}$ , il permet de voir que  $\text{Diag}(x, 1-x) \simeq C((X-x)(X-x+1))$  est n-adjacente à un idempotent de rang 1 ; on applique alors ce principe en prenant successivement pour  $x$  les éléments  $1, 2, \dots, r$ . Si  $N$  est nilpotente, alors  $r = 0$  (calcul de trace) donc immédiatement  $D$  est somme de deux matrices de carré nul, tandis que  $N$  l'est elle-même. Supposons  $N$  non nilpotente, donc bien partitionnée. Le lemme d'adaptation associé permet alors de voir que  $N$  est n-adjacente à la matrice compagnon de  $X^{t-1}(X+r)$ , elle-même semblable à  $C(X^{t-1}) \oplus (-rI_1)$ . Par suite,  $M$  est n-adjacente à  $C(X^{t-1}) \oplus \text{Diag}(-r, -r+1, \dots, 0, \dots, r-1, r)$ , et chaque terme de cette somme directe est bien somme de deux matrices de carré nul, ce qui permet de conclure.

Où est le problème dans le cas de la caractéristique non nulle ? On considère bien sûr que les termes diagonaux recherchés sont les éléments  $k.1_{\mathbb{K}}$  où  $k \in \llbracket -(r-1), r \rrbracket$ . Le point critique réside alors dans le fait que la similitude  $\text{Diag}(x, 1-x) \simeq C((X-x)(X-x+1))$  n'est valable que lorsque  $2x \neq 1$  (de façon plus mineure, le fait que  $N$  soit nilpotente n'implique pas ici que  $r = 0$ , mais suffit à avoir  $r.1_{\mathbb{K}}=0$ , ce qui permet de reconnaître directement le fait que  $\text{Diag}(-(r-1).1_{\mathbb{K}}, \dots, 0_{\mathbb{K}}, 1_{\mathbb{K}}, \dots, r.1_{\mathbb{K}})$  est somme de deux matrices de carré nul). Dans la démonstration précédente, on évitait cette situation grâce à la

caractéristique 0, mais en caractéristique *impair* il est fort possible de tomber sur cet os ! Quant à la caractéristique 2, la situation est pire encore car la symétrie spectrale par rapport à 0 n'apporte rien : la matrice  $\text{Diag}(-1, 1)$  sur  $\mathbb{F}_2$  n'est pas somme de deux matrices de carré nul. La bonne adaptation de la démonstration précédente consiste alors à utiliser non des matrices diagonalisables mais des matrices dont tous les blocs de Jordan sont de taille 2 ! Pour s'en sortir, on revient à la décomposition de départ, en montrant que l'on peut en outre choisir  $r$  *pair*, que l'on écrit  $r = 2s$ . On montre alors que  $I_{2s} \oplus 0_{2s}$  est n-adjacente à la somme directe des blocs de Jordan  $C((X - k.1_{\mathbb{K}})^2)$  pour  $k$  de  $-(s-1)$  à  $s$ . Si  $\mathbb{K}$  n'est pas de caractéristique 2 et  $2s.1_{\mathbb{K}} \neq 0$ , on utilise le lemme d'adaptation pour voir que  $N$  est n-adjacente à  $C(X^{t-2}) \oplus C((X + s.1_{\mathbb{K}})^2)$ , et on utilise enfin la caractérisation des sommes de deux matrices de carré nul pour voir que  $\bigoplus_{k=-s}^s C((X - k.1_{\mathbb{K}})^2)$  est de ce type, ce qui permet de conclure. En caractéristique 2, on observe directement que  $I_2$  est somme de trois matrices de carré nul, et comme  $N$  est de trace nulle on voit qu'elle est n-adjacente à une matrice cyclique nilpotente. Dans tous les cas, la conclusion est facile.

Examinons maintenant brièvement comment adapter les idées précédentes pour obtenir le théorème suivant, qui est le résultat principal de [2] :

**Théorème 8.16.** *Toute matrice de  $M_n(\mathbb{K})$  est combinaison linéaire de trois matrices idempotentes.*

Partons de  $M \in M_n(\mathbb{K})$ . En s'inspirant des idées précédentes, on voit qu'ou bien  $M$  a pour polynôme caractéristique une puissance d'un polynôme irréductible, ou bien  $M$  est semblable à  $N \oplus \alpha I_r \oplus \beta I_s$ , où  $\alpha$  et  $\beta$  sont des scalaires distincts, et  $N$  une matrice très bien partitionnée. Le premier cas se traite en utilisant la réduction de Jordan généralisée (pour récrire plus intelligemment les matrices cycliques associées aux puissances d'un polynôme) et nous n'en dirons rien. Examinons le deuxième cas. Dans celui-ci, le travail sur le bloc  $N$  est légèrement différent de celui effectué pour les décompositions en sommes de matrices de carré nul. Au lieu de retrancher à  $N$  la matrice  $S$  dont tous les coefficients sont nuls hormis ceux situés sur la première sous-diagonale, aux positions en lesquels ceux de  $N$  sont nuls, et qui valent  $-1$ , on remplace  $S$  par la matrice  $P$  ayant les mêmes coefficients que  $S$ , hormis aux positions  $(i, i)$  pour lesquelles  $s_{i+1, i} \neq 0$ , où l'on impose  $p_{i, i} = 1$ , ce qui permet d'avoir  $P^2 = P$  ; rappelons que, dans la démonstration du lemme 8.14, on retranche non pas  $S$  à  $N$ , mais une certaine matrice semblable à  $S$  à  $N$ . De même ici, pour n'importe quel polynôme unitaire  $R$  dont le degré est la taille  $d$  de  $N$ , et pour n'importe quel  $\gamma \in \mathbb{K} \setminus \{0\}$  tel que  $\text{tr } R = \text{tr } N - \gamma \text{tr } P$ , on pourra trouver une matrice  $P'$  semblable à  $P$  et telle que  $N - \gamma P'$  soit cyclique de polynôme minimal  $R$ . On comprend bien qu'on a tout intérêt à éviter la situation où  $\text{tr } P = 0$ , qui empêcherait la moindre liberté sur la trace finalement obtenue. Et c'est là qu'intervient le fait de supposer  $N$  très bien partitionnée : dans ce cas, on peut en effet modifier à la marge la définition de  $P$  : si le premier bloc diagonal compagnon dans  $N$  est de taille au moins 2, alors on peut prendre  $p_{1,1} = 1$  au lieu de  $p_{1,1} = 0$ , et sinon le dernier est de taille au moins 2 et l'on peut prendre  $p_{n,n} = 1$  au lieu de  $p_{n,n} = 0$ . Ce faisant, on voit que l'on peut toujours imposer à  $P$  d'être de trace non nulle, et finalement on peut choisir pour  $R$  n'importe quel polynôme unitaire de degré  $n$  et de trace différente de celle de  $N$ . Lorsque  $\alpha\beta \neq 0$ , on peut alors démontrer que  $R$  peut



être choisi avec les contraintes précédentes pour que sa matrice compagnon soit une  $(X^2 - \alpha X, X^2 - \beta X)$ -somme, grâce aux effets de symétrie imparfaite permis pour ces sommes (voir le paragraphe 8.5.4). Dans le cas où  $\alpha \neq 0$  et  $\beta = 0$ , on montre plutôt que  $R$  peut être choisi avec les contraintes précédentes pour que sa matrice compagnon soit une  $(X^2 - \alpha X, X^2 + \alpha X)$ -somme.

## 8.6.2 Décompositions en produits à trois termes

Maintenant que nous avons exposé l'essentiel des idées figurant dans [2] et [5], parlons des adaptations effectuées pour la question des décompositions en produits. Dans [8], nous montrons que pour toute matrice inversible  $M \in \text{GL}_n(\mathbb{K})$  de déterminant  $\pm 1$ , la matrice augmentée  $M \oplus I_n$  est produit de trois éléments d'ordre 1 ou 2 (i.e. d'involutions) du groupe  $\text{GL}_n(\mathbb{K})$ . Nous obtenons aussi des décompositions similaires où certains facteurs précis doivent être des involutions, et les autres des unipotents d'indice 2 (autrement dit,  $(A - I)^2 = 0$ ). Nous obtenons aussi d'autres résultats précieux avec d'autres formes d'augmentation de la matrice  $M$ , résultats tous motivés par les applications à la dimension infinie. Les méthodes utilisées dans cet article sont fondées sur une adaptation de la stratégie mise en œuvre dans [5], avec cependant des difficultés spécifiques qui furent loin d'être faciles à résoudre. Il y avait deux grandes adaptations à réaliser. La première est celle de la méthode des matrices bien partitionnées, la deuxième le traitement des matrices de la forme  $N \oplus \alpha I_n \oplus I_n$  avec  $N$  bien partitionnée ou unipotente, tant qu'on dispose d'un équivalent du lemme d'adaptation pour les matrices bien partitionnées (voir le lemme 8.14).

Passons rapidement sur la seconde adaptation : on a vu que, pour les décompositions en sommes de trois matrices de carré nul, on faisait apparaître des chaînes de blocs de Jordan de taille 1 ou 2, avec des phénomènes de cyclicité qui apparaissent en caractéristique non nulle. Ici, les phénomènes de cyclicité sont incontournables, même en caractéristique nulle, car on travaille dans le groupe multiplicatif  $(\mathbb{K}^*, \times)$  et non dans le groupe additif  $(\mathbb{K}, +)$ . On est contraint de regarder de très près les ordres des éléments de ce groupe. Nous n'en dirons rien de plus, sinon que tout cela nécessite un travail beaucoup plus précis et méticuleux que dans [5].

Attardons-nous un peu plus sur l'adaptation du travail sur les matrices bien partitionnées. Repartons d'une telle matrice  $N$ . Dans le cas des décompositions en sommes de matrices de carré nul, et après modification du bloc supérieur droit de  $N$ , on se retrouvait à retrancher, à une matrice semblable à  $N$  obtenue en modifiant son bloc supérieur droit, une matrice ayant uniquement des termes non nuls sur sa sous-diagonale. Une adaptation naïve consisterait à prendre quelque chose comme suit : définir  $S$  comme prenant la valeur 1 en tout coefficient de la sous-diagonale de  $N$  qui est nul, 0 en tout autre élément de la sous-diagonale ou hors de la diagonale,  $-1$  en tout élément de la diagonale qui est situé au-dessus d'un terme valant déjà 1, et 1 partout ailleurs sur la diagonale. Mais bien que la matrice  $S$  ainsi construite soit une involution, le résultat du produit  $SN$  n'a pas du tout une forme sympathique manipulable ! Par exemple, si  $N$  a uniquement deux blocs carrés diagonaux dans sa forme bien partitionnée, de tailles respectives  $a$  et  $b$ , le bloc "inférieur gauche" de  $SN$  a pour première ligne  $[0 \ \cdots \ 0 \ 1 \ ?]$ , ce qui laisse perplexe sur la possibilité d'utiliser à bon escient la matrice  $S$ .

La bonne idée que j'ai eu, après tâtonnement, a consisté à construire  $S$  sous

la forme

$$S = \begin{bmatrix} D_1 & 0 & & [0] \\ F_1 & D_2 & \ddots & \\ & \ddots & \ddots & 0 \\ (0) & & F_{N-1} & D_N \end{bmatrix}$$

où chaque  $D_i$  est de la forme  $\text{Diag}(1, \dots, 1, -1)$  et les  $F_i$  sont toutes des ma-

trices élémentaires de la forme  $\begin{bmatrix} 0 & \dots & 0 \\ [0] & [0] & \vdots \\ 1 & [0] & 0 \end{bmatrix}$ . Le produit  $SN$  n'est alors pas de

Hessenberg, mais on peut néanmoins montrer qu'il est cyclique! L'innovation technique importante de ce travail réside dans la reconnaissance du caractère cyclique de certaines matrices inversibles de forme particulière : je les ai baptisées les matrices **quasi-compagnon par blocs**. J'ai d'abord introduit la notion de matrice quasi-compagnon : ce sont les matrices de la forme

$$\begin{bmatrix} 0 & & (0) & ? \\ ? & 0 & & ? \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & ? & 0 & ? \\ ? & \dots & ? & ? & ? \end{bmatrix}.$$

Une matrice est dite quasi-compagnon par blocs lorsqu'elle est de la forme

$$M = \begin{bmatrix} D_1 & ? & & & (?) \\ \beta_1 K_{d_2, d_1} & D_2 & & & \\ 0 & \beta_2 K_{d_3, d_2} & \ddots & & \\ \vdots & & \ddots & D_{N-1} & ? \\ (0) & \dots & 0 & \beta_{N-1} K_{d_N, d_{N-1}} & D_N \end{bmatrix},$$

où les  $D_i$  sont des matrices quasi-compagnon, les  $\beta_j$  des scalaires *non nuls*, et

$K_{a,b} = \begin{bmatrix} 0 & \dots & 0 \\ \vdots & (0) & \vdots \\ 0 & \dots & 1 \end{bmatrix} \in M_{a,b}(\mathbb{K})$  pour tout  $(a,b) \in (\mathbb{N}^*)^2$ . Le lemme clef (lemme

2.5 de [8]) stipule que lorsqu'une telle matrice de taille  $n$  est inversible, elle est automatiquement cyclique et, mieux, le  $d_1$ -ième vecteur de la base canonique est cyclique pour  $M$ , tandis que le  $(n - d_N + 1)$ -ième vecteur de la base canonique est cyclique pour  $M^T$ . Et lorsqu'on part d'une matrice bien partitionnée  $N$ , que l'on modifie le bloc supérieur droit arbitrairement puis que l'on multiplie à gauche par la matrice  $S$  construite précédemment, on peut observer que le résultat est une matrice quasi-compagnon par blocs (où les tailles des blocs diagonaux correspondent à ceux de la forme bien partitionnée de  $N$ ).

Il est alors possible d'avancer grâce aux versions pour les produits des lemmes d'adaptation :

**Lemme 8.17.** *Soit  $A \in M_n(\mathbb{K})$  et  $B \in M_m(\mathbb{K})$  des matrices de Hessenberg à coefficients tous égaux à 1 sur la première sous-diagonale, et  $R(X)$  un polynôme*

unitaire de degré  $n + m$  de norme  $(\det A)(\det B)$ . Il existe alors une matrice  $D \in M_{n,m}(\mathbb{K})$  telle que

$$\begin{vmatrix} XI_n - A & -D \\ XH_{m,n} & XI_m - B \end{vmatrix} = R(X) \quad \text{où} \quad H_{m,n} = \begin{bmatrix} 0 & \cdots & 1 \\ \vdots & (0) & \vdots \\ 0 & \cdots & 0 \end{bmatrix} \in M_{m,n}(\mathbb{K}).$$

Ce lemme peut se déduire du résultat correspondant pour les traces (lemme 8.13) par une astuce de développement selon la  $n$ -ième colonne. On obtient ensuite une généralisation plus directement manipulable :

**Lemme 8.18.** *Soit  $A \in M_n(\mathbb{K})$  et  $B \in M_m(\mathbb{K})$  des matrices cycliques, et  $Z$  et  $Y$  des vecteurs cycliques associés respectivement à  $B$  et  $A^T$ . Soit  $R(X)$  un polynôme unitaire de degré  $n + m$  et de norme  $(\det A)(\det B)$ . Il existe alors une matrice  $D \in M_{n,m}(\mathbb{F})$  telle que*

$$\begin{vmatrix} XI_n - A & -D \\ XZY^T & XI_m - B \end{vmatrix} = R(X).$$

À l'aide des idées précédentes, on obtient alors le lemme d'adaptation pour les matrices bien partitionnées :

**Lemme 8.19** (Théorème 3.6 de [8]). *Soit  $M \in GL_n(\mathbb{K})$  bien partitionnée. Il existe alors  $\eta \in \{\pm 1\}$  tel que, pour tout polynôme  $R \in \mathbb{K}[X]$  unitaire de norme  $\eta \det M$ , il existe une involution  $S \in GL_n(\mathbb{K})$  telle que  $SM$  soit cyclique de polynôme minimal  $R$ . Si en outre  $M$  est très bien partitionnée,  $\eta$  peut être choisi arbitrairement dans  $\{\pm 1\}$ .*

À partir de là, la stratégie est claire même si les détails demandent un travail minutieux. On dispose ainsi d'un résultat du même type (et encore plus simple) pour la multiplication par une matrice unipotente d'indice 2.

## 8.7 Sommes d'idempotents en caractéristique non nulle

Dans la foulée de mon travail sur les combinaisons linéaires de trois projecteurs, j'ai étudié les sommes de projecteurs en caractéristique non nulle [3]. Il est indispensable d'abord de rappeler les résultats en caractéristique nulle. Étant donné des projecteurs  $p_1, \dots, p_r$  d'un espace vectoriel de dimension finie, leur somme  $u$  vérifie  $\text{tr}(u) = \sum_{k=1}^r \text{tr}(p_k)$ , et cette trace est donc un entier naturel (identifié à un élément de  $\mathbb{K}$  puisque  $\chi(\mathbb{K}) = 0$ ) ; par ailleurs

$$\text{rg}(u) \leq \sum_{k=1}^r \text{rg}(p_k) = \sum_{k=1}^r \text{tr}(p_k) = \text{tr}(u).$$

Wu [Wu90] et Hartwig-Putcha [Har90b] ont démontré que les conditions  $\text{tr } u \in \mathbb{N}$  et  $\text{rg}(u) \leq \text{tr}(u)$  sont suffisantes pour qu'un endomorphisme  $u$  d'un espace vectoriel de dimension finie sur un corps de caractéristique nulle soit somme de projecteurs.

Il faut maintenant immédiatement dissiper une tentation éventuelle : l'examen des problèmes précédents pourrait par exemple laisser penser que l'on puisse drastiquement limiter le nombre de projecteurs nécessaire dans une décomposition possible, mais il n'en est rien. D'abord, il y a évidemment des contraintes liées à la trace et à la dimension  $n$  de l'espace vectoriel sous-jacent. En effet, tout projecteur de cet espace est de trace au plus  $n$ . Si  $u$  est somme de  $k$  projecteurs, alors on doit avoir  $\text{tr } u \leq nk$ . Il n'existe donc pas d'entier  $N$ , dépendant uniquement de  $n$ , tel que toute somme de projecteurs dans  $M_n(\mathbb{K})$  soit en réalité somme de  $N$  projecteurs. Ensuite, et même si l'on sait caractériser les sommes de deux projecteurs, on ne sait pas caractériser les sommes de trois projecteurs (on tombe sur des difficultés du même type que pour les sommes de trois matrices de carrés nuls, ou que pour les produits de trois matrices involutives).

Plaçons-nous en caractéristique  $p > 0$  et notons  $\mathbb{F}_p$  le sous-corps premier de  $\mathbb{K}$ . La première question à se poser est celle de la nature des endomorphismes qui sont des sommes de projecteurs. Il n'y a plus aucune contrainte de rang car on ne peut plus identifier les entiers avec des éléments du corps de base, mais il subsiste une contrainte évidente de trace : la trace d'un tel endomorphisme doit appartenir à  $\mathbb{F}_p$ . Cette contrainte est d'ailleurs vide si  $\mathbb{K}$  est lui-même premier. Le caractère suffisant de cette condition s'obtient facilement en remarquant par exemple que la matrice  $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$  est somme de  $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$  et de  $p - 1$  copies de  $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ . Plus généralement, cela permet de voir que toute matrice élémentaire nilpotente est somme de matrices de projecteurs, puis le théorème de Fillmore permet de voir que toute matrice non scalaire de trace nulle est somme de matrices de projecteurs. Et ajoutant à une matrice scalaire de trace nulle une matrice nilpotente non nulle, on obtient une matrice non scalaire de trace nulle, ce qui ramène au cas précédent. Enfin, à partir d'une matrice dont la trace appartient au sous-corps premier, il suffit de retrancher un nombre entier de fois une matrice de projecteur de rang 1 pour obtenir une matrice de trace nulle.

La question suivante est celle du contrôle sur le nombre minimal de matrices idempotentes nécessaire pour écrire n'importe quel élément de  $M_n(\mathbb{K})$  à trace dans  $\mathbb{F}_p$  comme somme de telles matrices. En adaptant la méthode ci-dessus, on voit que ce nombre  $s_n(\mathbb{K})$  vérifie  $s_n(\mathbb{K}) \geq \frac{p-1}{n}$  avec inégalité stricte si  $n > 1$ , ce qui n'a d'intérêt que pour les petites valeurs de  $n$ . Par ailleurs, et c'est là que les choses sérieuses commencent, une adaptation de la technique de double-connexion permet d'obtenir l'inégalité  $s_n(\mathbb{K}) \leq 5 + \lfloor \frac{p-1}{n} \rfloor$  (théorème 7 de [3]). L'idée est de partir de  $A \in M_n(\mathbb{K})$  de trace dans  $\mathbb{F}_p$ , mise sous forme réduite de Frobenius. Par la technique de double-connexion, on parvient alors à trouver deux idempotentes  $Q_1$  et  $Q_2$  telles que  $A - Q_1 - Q_2$  soit cyclique, toujours à trace dans  $\mathbb{F}_p$ . Ensuite, et c'est là une subtilité, on n'utilise pas directement le lemme d'adaptation pour les matrices cycliques ; plutôt, lorsque  $C(R)$  est une matrice compagnon de taille  $n$ , on lui retranche d'abord un certain nombre de fois (c'est là qu'intervient le nombre  $\lfloor \frac{p-1}{n} \rfloor$ ) la matrice identité pour obtenir une trace de la forme  $(k + \varepsilon) \cdot 1_{\mathbb{K}}$  avec  $k$  compris entre 1 et  $p - 1$ , et  $\varepsilon \in \{0, -1\}$ . On récrit par blocs le résultat sous la forme  $\begin{bmatrix} [?]_{k \times k} & [?]_{k \times (n-k)} \\ [?]_{(n-k) \times k} & [?]_{(n-k) \times (n-k)} \end{bmatrix}$ , et on cherche à lui retrancher un idempotent de la forme  $\begin{bmatrix} I_k & [?]_{k \times (n-k)} \\ 0 & 0 \end{bmatrix}$ , sachant

que le résultat restera cyclique de trace 0 ou 1 et que l'on peut jouer à volonté sur le choix du bloc supérieur droit de l'idempotent. Il reste alors à utiliser le lemme 8.13 pour montrer que le dernier idempotent retranché peut être choisi de manière à ajuster très librement le polynôme caractéristique du résultat (à l'exception de sa trace et de son degré, qui sont imposés). On reconnaît alors, par un choix convenable, la somme de deux idempotents.

Une conséquence du résultat précédent est que,  $\mathbb{K}$  restant fixé de caractéristique  $p$ , dès que  $n \geq p$  toute matrice de  $M_n(\mathbb{K})$  à trace dans  $\mathbb{F}_p$  est somme de cinq matrices idempotentes. On ne sait pas si l'on peut descendre à quatre matrices idempotentes, mais nous sommes parvenus au résultat suivant dans le cas particulier des corps premiers :

**Théorème 8.20.** *Soit  $p$  un nombre premier. Il existe un entier naturel  $n_0 > 0$  tel que, pour tout  $n \geq n_0$ , toute matrice de  $M_n(\mathbb{F}_p)$  soit somme de quatre matrices idempotentes.*

La démonstration de ce résultat utilise quelques techniques vues antérieurement mais ne s'y limite pas. Une première remarque élémentaire, mais qui simplifie assez bien les choses, consiste à se ramener à considérer des décompositions en *sommes de deux différences d'idempotents* (qui sont plus faciles à reconnaître).

Ensuite, on peut utiliser finement le principe de double-connexion : supposons  $M \in M_n(\mathbb{F}_p)$  non scalaire et réduisons-nous à la situation où elle est sous forme de Frobenius (inversée)  $M = C(P_1) \oplus \dots \oplus C(P_r)$  avec  $\deg P_r > 1$ . Comme toujours, on fabrique des idempotents  $Q, Q'$  de Hessenberg en jouant sur la position des coefficients sous-diagonaux nuls ainsi que sur un choix judicieux de dernière colonne, en essayant d'ajuster au mieux le polynôme caractéristique de  $M - (Q - Q')$ . Dans ce choix, on conserve de réels degrés de liberté sur le choix des coefficients diagonaux de  $Q$  et  $Q'$  (se résumant à la position des 1 et des 0) : si  $\sum_{k=1}^r (\deg(P_k) - 1) \geq p$  alors  $\text{tr}(Q - Q')$  peut être ajusté très librement dans  $\mathbb{F}_p$ . Cela donne directement le résultat lorsque les invariants de similitude de  $M$  sont de degré assez grand.

L'idée est ensuite de s'intéresser aux autres matrices, ce qui pousse plus généralement à décomposer  $M$  sous la forme  $N \oplus \alpha I_k$  où  $N$  a tous ses invariants de similitude de degré au moins 2, et  $\alpha \in \mathbb{F}$ . Et si  $N$  est de taille au moins  $2p$ , la méthode précédente donne directement que  $N \oplus \alpha I_k$  est somme de quatre idempotentes.

Supposons ensuite  $\alpha \neq 0$ . Si  $k$  est assez grand, on peut toujours casser  $N \oplus \alpha I_k = (N \oplus \alpha I_l) \oplus \alpha I_{k-l}$  en choisissant judicieusement  $l$  pour ajuster la trace de  $N \oplus \alpha I_l$  comme on veut, ce qui n'a pas d'impact sur les traces possibles pour les matrices  $(N \oplus \alpha I_l) - (Q - Q')$  que l'on construit par le procédé de double-connexion ; et dans ce cas  $k - l$  restera assez grand. Il reste dans ce cas à montrer que, pour tout  $n$  assez grand, toute matrice scalaire de  $M_n(\mathbb{F}_p)$  est somme de quatre idempotentes : cela s'obtient par des arguments ... d'arithmétique (!), l'idée de base étant, pour n'importe quel  $\alpha \in \mathbb{F}_p$ , de montrer que si  $n$  est assez grand on peut trouver une matrice diagonale  $D$  telle que  $D$  et  $D - \alpha I_n$  soient des différences de deux idempotentes. L'argument repose sur l'effet de symétrie spectrale pour les différences de deux idempotentes et se réduit à un choix convenable d'une partition de l'entier  $p$  vérifiant des conditions de symétrie imposées (voir le lemme 12 de [3]).

Ne reste alors qu'à traiter le cas où la taille de  $N$  est assez petite,  $k$  est assez grand et  $\alpha = 0$ . Dans ce cas, il y a peu d'invariants de similitude pour  $N$ , et

on groupe chacun d'entre eux avec un bloc  $0_l$  avec  $l$  assez grand à ajuster. On utilise alors une technique de chaîne de blocs de Jordan en jouant sur l'asymétrie entre les valeurs propres exceptionnelles 1 et  $-1$  pour la différence de deux idempotentes. Typiquement, la chaîne  $C(X+1) \oplus C(X-1) \oplus C((X+1)^2) \oplus \dots \oplus C((X-1)^{k-1}) \oplus C((X+1)^k)$  est différence de deux idempotentes et de trace  $-k$ , et si on lui rajoute le bloc  $C((X-1)^k)$ , qui est de trace  $k$ , on conserve une différence de deux idempotentes (avec parfaite symétrie spectrale par rapport à 0).

Par ailleurs, la technique de double-connexion permet de trouver, pour n'importe quel polynôme  $P \in \mathbb{F}_p[X]$  unitaire (de degré  $d$ ), un entier  $l \in \llbracket 0, p-1 \rrbracket$  tel que  $C(P) \oplus 0_l$  donne, après soustraction d'une différence de deux idempotentes bien choisie, une matrice cyclique de polynôme minimal  $(X-1)^{d+l}$ . En prenant la chaîne précédente associée à  $k := d+l$ , on obtient alors que  $C(P) \oplus 0_{l+a}$  est somme de deux différences d'idempotentes pour  $a := k^2$ .

Pour conclure, lorsque  $\alpha = 0$ ,  $N$  de format assez petit et  $k$  assez grand, on casse le bloc  $0_k$  en sous-blocs suffisamment gros pour accompagner chaque bloc compagnon de  $N$  et utiliser la technique précédente.

À noter enfin que les résultats précédents ont été améliorés pour les corps de toute petite caractéristique non nulle. En caractéristique  $p \in \{2, 3\}$ , toute matrice carrée qui est somme d'idempotentes est en fait somme de quatre idempotentes (proposition 10 de [3]). Enfin, sur  $\mathbb{F}_p$ , et lorsque  $p \in \{2, 3\}$ , toute matrice carrée est somme de trois matrices idempotentes (proposition 9 de [3]), résultat qui ne tient plus à partir de  $p = 5$ .

## 8.8 Décompositions en dimension infinie

### 8.8.1 Décomposition en sommes de quatre termes

La plupart des études sur les décompositions en objets quadratiques portent sur l'algèbre des endomorphismes d'un espace vectoriel de dimension finie, ou l'algèbre des opérateurs bornés d'un espace de Hilbert de dimension infinie. Mais au moment où j'ai pris le sujet, presque rien n'était connu sur la situation de l'algèbre des endomorphismes d'un espace vectoriel  $V$  de dimension infinie (sur un corps arbitraire). Le seul article publié donnait des résultats assez décevants sur les matrices infinies (autrement dit, sur le cas où  $V$  est de dimension dénombrable) : il y était par exemple démontré que dans ce cas toute matrice infinie sur un corps est la somme de douze (!) matrices de carré nul. J'ai rapidement démontré [7] que l'on pouvait descendre à quatre termes, et plus généralement que tout endomorphisme d'un espace vectoriel de dimension infinie est somme de quatre endomorphismes de carré nul. Et j'ai généralisé le résultat sous la forme suivante :

**Théorème 8.21** (Théorème 2.3 de [7]). *Soit  $p_1, p_2, p_3, p_4$  quatre polynômes de degré 2 scindés sur  $\mathbb{K}$ , et  $V$  un espace vectoriel de dimension infinie sur  $\mathbb{K}$ . Tout endomorphisme de  $V$  est alors une  $(p_1, p_2, p_3, p_4)$ -somme dans  $\mathcal{L}(V)$ .*

En particulier, tout endomorphisme de  $V$  est somme de quatre projecteurs, mais aussi somme de trois projecteurs et d'une symétrie, somme de deux projecteurs et de deux endomorphismes de carré nul etc. Alors qu'en dimension finie on aurait immédiatement des contraintes liées à la trace, rien de tel ici.

En outre, le théorème précédent est optimal comme on le verra plus tard (trois termes à polynômes annulateurs imposés ne suffisent pas en général).

Je vais immédiatement donner les grandes idées de la démonstration, qui repose sur une adaptation de la méthode de double-connexion vue dans la section 8.6. La clef est de s'intéresser à des endomorphismes particuliers, ceux qui sont des sommes directes de décalages à droite. Un **décalage à droite** est un endomorphisme  $u$  pour lequel il existe une base  $(e_n)_{n \in \mathbb{N}}$  de l'espace vectoriel sous-jacent telle que  $u(e_n) = e_{n+1}$  pour tout  $n \in \mathbb{N}$  (on dit alors que  $e_0$  est un **vecteur cyclique associé**). Et il est particulièrement utile de remarquer qu'est un décalage à droite tout endomorphisme  $u$  pour lequel il existe une base  $(e_n)_{n \in \mathbb{N}}$  telle que  $u(e_n) = e_{n+1}$  modulo  $\text{Vect}(e_0, \dots, e_n)$  pour tout  $n \in \mathbb{N}$ . En utilisant sous forme de matrice illimitée l'astuce décrite dans le paragraphe 8.5.4 pour fabriquer des  $(p, q)$ -sommes exceptionnelles à l'aide de sommes directes de blocs de taille 1 ou 2, on obtient le lemme suivant :

**Lemme 8.22.** *Soit  $p_1, p_2$  deux polynômes de degré 2 à coefficients dans  $\mathbb{K}$ , dont  $p_1$  scindé. Alors tout décalage à droite d'un espace vectoriel  $V$  est une  $(p_1, p_2)$ -somme dans  $\mathcal{L}(V)$ .*

Plus généralement, appelons **endomorphisme élémentaire** toute somme directe de décalages à droite. Il est alors immédiat qu'un tel endomorphisme est aussi une  $(p_1, p_2)$ -somme quels que soient  $p_1, p_2$  deux polynômes de degré 2 dans  $\mathbb{K}[X]$ , avec  $p_1$  scindé.

Dans [Shi16], qui a été soumis après la pré-publication de [7], Yaroslav Shitov a montré que tout endomorphisme d'un espace vectoriel de dimension infinie est somme de deux endomorphismes élémentaires. À partir de là, le théorème 8.21 découle immédiatement du lemme précédent, et on en obtient même une version légèrement améliorée, dans laquelle au plus deux des polynômes  $p_i$  peuvent être irréductibles (nous reviendrons plus tard sur ce point).

La méthode que nous avons utilisée dans [7] est substantiellement différente de celle de Shitov, car elle s'appuie sur la technique de double-connexion déjà vue dans la section 8.6. Et l'intérêt de cette méthode est qu'elle peut être adaptée pour étudier les décompositions à trois termes, alors que les auteurs des démonstrations alternatives<sup>5</sup> du théorème 8.21 n'ont à ce jour pas pu adapter leurs méthodes pour aller plus loin.

En dimension infinie, on dispose de très peu d'outils de réduction des endomorphismes. Nous nous sommes néanmoins appuyés sur une généralisation de la réduction à une forme de Hessenberg. La construction est ordinaire : on part d'un endomorphisme  $u$  de  $V$  et d'un vecteur non nul  $x_0$ , on examine la famille  $(u^k(x_0))_{k \geq 0}$  et l'espace  $V_0$  qu'elle engendre, et on extrait de  $(u^k(x_0))_{k \geq 0}$  une base de  $V_0$  formée de vecteurs consécutifs démarrant à  $x_0$ . Si  $V_0 \neq V$ , on prend un vecteur  $x_1$  de  $V \setminus V_0$ , on examine le sous-espace  $V_1 := V_0 + \text{Vect}(u^k(x_1))_{k \in \mathbb{N}}$  et on extrait de  $(u^k(x_1))_{k \in \mathbb{N}}$  une famille dont la projection modulo  $V_0$  soit une base de  $V_1/V_0$ , et partant de  $x_1$  en prenant des termes consécutifs. On poursuit indéfiniment si possible, étendant la constructions à des ordinaux. Pour clarifier le discours, nous adoptons la terminologie suivante : nous munissons  $V$  de la structure de  $\mathbb{K}[X]$ -module attachée à l'endomorphisme  $u$ , et nous appelons **stratification** de  $V$  toute famille  $(V_\alpha)_{\alpha \in D}$  strictement croissante de sous-modules

5. Outre Shitov, il faut aussi citer [Bre18], qui s'appuie sur un résultat connu stipulant que tout endomorphisme d'un espace vectoriel de dimension infinie est de la forme  $uv - vu$ .

de  $V$  indexée par un ensemble bien ordonné  $D$  (par exemple, un ordinal), de réunion  $V$  et telle que  $V_\alpha/(\sum_{\beta<\alpha} V_\beta)$  soit non nul et possède un générateur pour tout  $\alpha \in D$ . Une famille de vecteurs  $(e_\alpha)_{\alpha \in D}$  est dite adaptée à cette stratification lorsque  $e_\alpha$  se projette en un générateur de  $V_\alpha/(\sum_{\beta<\alpha} V_\beta)$  pour tout  $\alpha \in D$ .

Pour fixer les idées, examinons la situation où l'on a pu construire une telle stratification  $(V_n)_{n \in \mathbb{N}}$  indexée par  $\mathbb{N}$ , avec une famille  $(e_n)_{n \in \mathbb{N}}$  adaptée. Notons  $d_n$  la dimension de l'espace vectoriel  $V_n/V_{n-1}$  pour tout  $n \in \mathbb{N}$ , en convenant que  $V_{-1} = \{0\}$  et que  $d_n = +\infty$  si cette dimension est infinie dénombrable. Ainsi,  $(u^k(e_n))_{0 \leq k < d_n}$  est une base de  $V_n$  modulo  $V_{n-1}$  pour tout  $n \in \mathbb{N}$ , et la famille double  $(u^k(e_n))_{n \in \mathbb{N}, 0 \leq k < d_n}$  est une base de  $V$ . Donnons-nous maintenant deux polynômes scindés  $p_1 = (X - \alpha)(X - \alpha')$  et  $p_2 = (X - \beta)(X - \beta')$ . On peut construire des endomorphismes  $a$  et  $b$ , respectivement  $p_1$ -quadratique et  $p_2$ -quadratique, comme suit :

- $a(u^k(e_n)) = \alpha u^k(e_n)$  pour tout  $n \in \mathbb{N}$  pair et tout  $k < d_n - 1$  ;
- $a(u^{d_n-1}(e_n)) = \alpha u^{d_n-1}(e_n) + e_{n+1}$  pour tout  $n \in \mathbb{N}$  pair tel que  $d_n < +\infty$  ;
- $a(u^k(e_n)) = \alpha' u^k(e_n)$  pour tout  $n \in \mathbb{N}$  impair et tout  $k < d_n$  ;
- $b(u^k(e_n)) = \beta u^k(e_n)$  pour tout  $n \in \mathbb{N}$  impair et tout  $k < d_n - 1$  ;
- $b(u^{d_n-1}(e_n)) = \beta u^{d_n-1}(e_n) + e_{n+1}$  pour tout  $n \in \mathbb{N}$  impair tel que  $d_n < +\infty$  ;
- $b(u^k(e_n)) = \beta' u^k(e_n)$  pour tout  $n \in \mathbb{N}$  pair et tout  $k < d_n$ .

On peut alors démontrer que  $u - a - b$  est élémentaire (et c'est même un décalage à droite si tous les  $d_n$  sont finis), ce qui est intuitivement évident mais demande une démonstration précise.

La construction précédente s'adapte facilement au cas plus général d'une stratification  $(V_\alpha)_{\alpha \in D}$  où  $D$  est dénué de maximum. Et lorsque  $V$  est de dimension indécomposable, il est toujours possible de construire une telle stratification grâce à des raisonnements de cardinalité : tout simplement, on note  $\kappa$  la dimension de  $V$  (c'est un cardinal), on choisit une base  $(f_\alpha)_{\alpha \in \kappa}$  de l'espace vectoriel  $V$  ; on construit alors par récurrence transfinitive une suite strictement croissante de sous-module  $(V_\alpha)_{\alpha \in \kappa}$ , tous de dimension strictement inférieure à celle de  $V$  comme  $\mathbb{K}$ -espace vectoriel, comme suit : pour un ordinal  $\alpha$ , le sous-module  $\sum_{\beta < \alpha} V_\beta$  ne peut être égal à  $V$  (limitation de la dimension), on choisit donc un vecteur  $e_\alpha$  dans  $V \setminus \sum_{\beta < \alpha} V_\beta$ , vecteur que l'on prend même égal à  $f_\alpha$  si possible et on note  $V_\alpha := \mathbb{K}[X]e_\alpha + \sum_{\beta < \alpha} V_\beta$ . Le choix des  $e_\alpha$  force alors  $V$  à être la réunion de  $V_\alpha$ , et on dispose donc de la stratification recherchée.

Lorsqu'en revanche  $V$  est de dimension dénombrable il est possible que de telles stratifications n'existent pas. Dans ce cas j'ai montré qu'on pouvait toujours construire une stratification  $(V_\alpha)_{\alpha \in D}$  dans laquelle  $V_{\min D}$  est de dimension infinie, et que si  $D$  a un maximum alors il suffit d'adapter légèrement la construction de  $a$  et  $b$  (de manière à renverser la situation pour qu'au bout du compte, le premier "bloc" de la stratification fasse office de dernier "bloc" après soustraction de  $a$  et  $b$ ) pour obtenir le résultat voulu, à savoir que  $u - a - b$  est élémentaire.

On notera que la technique précédente n'est pas aussi performante que la démonstration de Shitov, en ce qu'elle nécessite déjà que les polynômes  $p_1$  et  $p_2$  soient scindés. Cependant, cette méthode s'est montrée beaucoup plus à même de traiter la question des décompositions à trois termes, que je vais maintenant évoquer. J'ai en effet étudié les obstructions à décomposer un endomorphisme



$u$  comme une  $(p_1, p_2, p_3)$ -somme lorsque  $p_1, p_2, p_3$  sont de degré 2 et scindés. Quelques obstructions de ce type apparaissaient déjà dans [7]. Par exemple, j'y montre qu'un endomorphisme de rang fini n'est somme de trois endomorphismes de carré nul que s'il est de trace nulle, et en particulier un projecteur de rang 1 n'est jamais la somme de trois endomorphismes de carré nul (proposition 2.5 de [7]). J'ai aussi montré que lorsque  $\lambda$  est un scalaire tel que  $2\lambda \neq 3$  et  $\lambda \notin \{0, 1, 2, 3\}$ , l'homothétie  $\lambda \text{id}$  n'est pas la somme de trois projecteurs (proposition 2.6 de [7]). J'évoquerai les généralisations de ces résultats dans le paragraphe 8.8.3.

## 8.8.2 Décompositions à trois termes

Je me suis intéressé à la possibilité de n'utiliser qu'un seul élément quadratique pour passer d'un endomorphisme  $u$  à un endomorphisme élémentaire par soustraction. Pour l'essentiel, la nécessité d'en utiliser deux, dans l'exemple précédent, est liée au risque que certaines dimensions  $d_n$  soient égales à 1. Repartons de l'idée naïve d'une simple connexion avec une stratification  $(V_\alpha)_{\alpha \in D}$  indexée par un ordinal  $D$  sans maximum, avec une famille de vecteurs adaptée  $(e_\alpha)_{\alpha \in D}$ , et notons  $d_\alpha$  la dimension de l'espace vectoriel quotient  $V_\alpha / (\sum_{\beta < \alpha} V_\beta)$  pour tout  $\alpha \in D$ . Partant d'un polynôme scindé  $p = (X - x)(X - y)$ , on définirait naturellement  $a(u^k(e_\alpha)) = x u^k(e_\alpha)$  si  $\alpha \in D$  est un ordinal *pair* (c'est-à-dire somme d'un ordinal sans prédécesseur et d'un entier pair) et  $k < d_\alpha - 1$ ,  $a(u^{d_\alpha-1}(e_\alpha)) = x u^{d_\alpha-1}(e_\alpha) - e_{\alpha+1}$  si  $\alpha \in D$  est un ordinal pair et si  $d_\alpha < +\infty$ , et procéder de même pour les ordinaux impairs en inversant les rôles de  $x$  et  $y$ . Et, une nouvelle fois, la différence  $u - a$  s'avère élémentaire!

Mais, pour que  $p(a) = 0$ , il est alors nécessaire et suffisant que  $d_\alpha > 1$  pour tout  $\alpha \in D$  qui a un prédécesseur et qui vérifie  $d_{\alpha-1} < +\infty$ . Il est donc naturel de s'intéresser à l'existence d'une stratification vérifiant ces propriétés. Cela exclut immédiatement les homothéties, et plus généralement on peut voir que cela exclut les endomorphismes possédant une valeur propre *dominante* selon la définition suivante :

**Définition 13.** Soit  $u$  un endomorphisme d'un espace vectoriel  $V$  de dimension infinie, et  $\lambda$  un scalaire. On dit que  $\lambda$  est une valeur propre **dominante** de  $u$  lorsque  $\text{rg}(u - \lambda \text{id}) < \dim V$ .

Attention, il s'agit ici de comparaison entre dimensions, donc entre cardinaux éventuellement infinis. Par exemple, si  $V$  est de dimension infinie dénombrable, alors  $\lambda$  est une valeur propre dominante de  $u$  si et seulement si  $u - \lambda \text{id}$  est de rang fini. Si  $V$  a pour dimension la puissance du continu,  $\lambda$  est une valeur propre dominante de  $u$  si et seulement si  $u - \lambda \text{id}$  est de rang strictement inférieur à la puissance du continu. Cette notion est intimement liée à la structure d'anneau de  $\mathcal{L}(V)$ , l'ensemble des endomorphismes de rang strictement inférieur à  $\dim V$  étant l'idéal bilatère maximal de  $V$ . On peut ainsi voir qu'un endomorphisme admet au plus une valeur propre dominante. Et on peut démontrer qu'une stratification du type précédent ne peut exister lorsque  $u$  possède une valeur propre dominante (dans ce cas, il existe forcément un terme  $V_\alpha$  de la stratification tel que le quotient  $V/V_\alpha$  soit non nul et  $u$  induise dessus l'homothétie de rapport  $\lambda$ , où  $\lambda$  est la valeur propre voulue). Et inversement, pour peu que  $V$  soit de dimension infinie indénombrable et que  $u$  n'ait pas de valeur propre dominante, il est possible de construire une bonne stratification. Pour cela, on

adapte la construction précédente : on prend une base  $(f_\gamma)_\gamma$  de  $V$  indexée par l'ensemble des ordinaux *sans prédécesseur* de la dimension  $\kappa$  de  $V$  (c'est possible car cet ensemble est en bijection avec  $\kappa$ , et c'est en particulier là qu'intervient l'indénombrabilité de  $\kappa$ ). Ensuite, dans le choix de la famille  $(e_\alpha)_\alpha$  adaptée à la stratification à construire, on choisit systématiquement  $e_\alpha = f_\alpha$  lorsque  $\alpha$  est sans prédécesseur et  $f_\alpha$  n'appartient à aucun des sous-modules déjà construits ; sinon et si  $\alpha$  admet un prédécesseur, l'endomorphisme induit par  $u$  sur le quotient  $V/V_{\alpha-1}$  n'est pas une homothétie (par absence de valeur propre dominante pour  $u$ ) et on peut donc choisir un vecteur  $e_\alpha$  tel que  $(e_\alpha, u(e_\alpha))$  soit libre modulo  $V_{\alpha-1}$  (ce qui fournit un bloc associé de taille au moins 2). On en déduit le résultat suivant :

**Théorème 8.23.** *Soit  $p \in \mathbb{K}[X]$  scindé de degré 2, et  $u \in \mathcal{L}(V)$  sans valeur propre dominante, où  $V$  est de dimension infinie indénombrable. Il existe alors  $a \in \mathcal{L}(V)$  tel que  $u - a$  soit élémentaire et  $p(a) = 0$ .*

Ce résultat reste vrai lorsque  $V$  est de dimension dénombrable, mais la démonstration est beaucoup plus technique. Dans cette situation, le cas où le  $\mathbb{K}[X]$ -module  $V$  est de torsion (pour la structure associée à  $u$ ) se traite par la même méthode que précédemment (en construisant une *bonne* stratification et en utilisant la technique de connexion). Quant au cas où  $V$  n'est pas de torsion, il est très subtil mais nous pouvons néanmoins mentionner qu'il repose sur une combinaison astucieuse de la technique de connexion avec l'énoncé suivant, qui représente la situation de blocage la plus emblématique possible :

**Lemme 8.24** (Lemme de couture, lemme 28 de [6]). *Soit  $u$  un endomorphisme d'un espace vectoriel  $V$ . On suppose que  $u$  est la somme directe d'une homothétie et d'un décalage à droite de vecteur cyclique  $x$ . Soit  $p \in \mathbb{K}[X]$  scindé de degré 2. Il existe alors  $a \in \mathcal{L}(V)$  tel que  $u - a$  soit un décalage à droite, de vecteur associé  $x$ , et  $p(a) = 0$ .*

On obtient alors le puissant théorème :

**Théorème 8.25.** *Soit  $p \in \mathbb{K}[X]$  scindé de degré 2, et  $u \in \mathcal{L}(V)$  sans valeur propre dominante, où  $V$  est de dimension infinie. Il existe alors  $a \in \mathcal{L}(V)$  tel que  $u - a$  soit élémentaire et  $p(a) = 0$ .*

En conséquence du résultat précédent, on obtient :

**Théorème 8.26.** *Soit  $p_1, p_2, p_3$  des polynômes de degré 2 à coefficients dans  $\mathbb{K}$ , dont  $p_2$  et  $p_3$  scindés. Alors tout endomorphisme  $u$  d'un espace vectoriel  $V$  de dimension infinie sans valeur propre dominante est une  $(p_1, p_2, p_3)$ -somme dans  $\mathcal{L}(V)$ .*

Le théorème précédent donne en effet un élément  $p_3$ -quadratique  $a$  tel que  $u - a$  soit élémentaire, et partant de là  $u - a$  est une  $(p_1, p_2)$ -somme.

On peut noter qu'à partir de là, on retrouverait facilement le théorème 8.21 et même une version renforcée de celui-ci dans laquelle on peut permettre à au plus deux des polynômes  $p_1$  et  $p_2$  d'être irréductibles sur le corps de base. Pour cela, et même si l'approche de Shitov est plus rapide, on part d'un quadruplet  $(p_1, p_2, p_3, p_4)$  de polynômes de degré 2 dans lequel  $p_3$  et  $p_4$  sont scindés, et on se donne  $u \in \mathcal{L}(V)$ . Dans tous les cas, on n'a pas de mal à trouver un endomorphisme  $p_1$ -quadratique  $a$  tel que  $u - a$  n'ait pas de valeur propre dominante et soit donc une  $(p_2, p_3, p_4)$ -somme.

### 8.8.3 Cas d'un endomorphisme à valeur propre dominante

On a vu que l'existence d'une valeur propre dominante bloquait les méthodes précédentes. Il est temps de constater que ces blocages ne sont pas artificiels et dus à une méthode inadaptée, mais qu'au contraire ils sont inévitables.

Nous fixons tout du long trois polynômes unitaires  $p_1, p_2, p_3$  scindés sur  $\mathbb{K}$ . Le résultat de base consiste en un examen des éléments centraux qui sont des  $(p_1, p_2, p_3)$ -sommés (ici, nous revenons à la situation de polynômes scindés). Un peu de terminologie sera utile :

**Définition 14.** Un scalaire  $\lambda$  est dit  $(p_1, p_2, p_3)$ -acceptable lorsqu'il vérifie l'une au moins des conditions suivantes :

- (i) il s'écrit  $x_1 + x_2 + x_3$  pour des racines respectives  $x_1, x_2, x_3$  de  $p_1, p_2, p_3$  ;
- (ii) il vérifie  $2\lambda = \text{tr } p_1 + \text{tr } p_2 + \text{tr } p_3$ .

La deuxième condition est liée à l'observation (non triviale) du fait que la matrice  $2\lambda I_2$  est une  $(p_1, p_2, p_3)$ -somme dans  $M_2(\mathbb{K})$  : pour le voir on démontre  $2\lambda I_2 - C(P_1)$  est bien une  $(p_2, p_3)$ -somme, ce qui peut s'obtenir grâce aux théorèmes évoqués dans la section précédente consacrée au problème à deux termes en dimension finie.

On a alors le résultat fondamental suivant :

**Lemme 8.27.** *Soit  $\lambda \in \mathbb{K}$  et  $\mathcal{A}$  une  $\mathbb{K}$ -algèbre non triviale.*

- (i) *Pour que  $\lambda.1_{\mathcal{A}}$  soit une  $(p_1, p_2, p_3)$ -somme dans  $\mathcal{A}$ , il est nécessaire que  $\lambda$  soit  $(p_1, p_2, p_3)$ -acceptable.*
- (ii) *Si  $\mathcal{A}$  est l'algèbre des endomorphismes d'un espace vectoriel de dimension infinie sur  $\mathbb{K}$ , il est suffisant que  $\lambda$  soit  $(p_1, p_2, p_3)$ -acceptable pour que  $\lambda.1_{\mathcal{A}}$  soit une  $(p_1, p_2, p_3)$ -somme dans  $\mathcal{A}$ .*

Le point (i) s'obtient par des identités algébriques élémentaires. Le point (ii) s'obtient presque trivialement pour le premier cas d'élément acceptable, et par des décompositions de matrices  $2 \times 2$  pour le deuxième.

On a donc déjà des contraintes sur ces éléments particuliers. Et ces contraintes rejaillissent aussi sur les valeurs possibles pour la valeur propre dominante éventuelle d'un endomorphisme  $u$  afin qu'il puisse être une  $(p_1, p_2, p_3)$ -somme. Cela peut se déduire en travaillant dans le quotient de  $\mathcal{L}(V)$  par son idéal bilatère maximal, mais on peut également utiliser le lemme suivant, qui est d'usage beaucoup plus général :

**Lemme 8.28** (Lemme du sous-espace stable commun). *Soit  $p_1, p_2, p_3$  trois polynômes de degré 2 dans  $\mathbb{K}[X]$ , et  $a, b, c$  des endomorphismes d'un espace vectoriel  $V$  de dimension infinie, tels que  $p_1(a) = p_2(b) = p_3(c) = 0$ . Soit  $W$  un sous-espace vectoriel de  $V$  stable par  $u := a + b + c$  et tel que  $\dim W < \dim V$ . Il existe alors un sous-espace vectoriel  $\overline{W}$  stable par  $a, b$  et  $c$ , tel que  $W \subset \overline{W}$  et  $\dim \overline{W} < \dim V$ , et même  $\dim \overline{W} = \dim W$  si  $W$  est de dimension infinie.*

L'idée, assez simple, est de prendre pour  $\overline{W}$  la somme de  $W$  et de toutes ses images par les mots à une ou deux lettres formés sur  $a, b, c$ . Par des astuces algébriques, on parvient péniblement à constater que  $\overline{W}$  est stable par  $a, b, c$ , les conditions de dimension étant alors évidemment remplies.

Grâce à ce lemme, on obtient en particulier que si  $u$  est une  $(p_1, p_2, p_3)$ -somme et a une valeur propre dominante  $\lambda$ , alors  $\lambda$  est  $(p_1, p_2, p_3)$ -acceptable.

En effet, on peut d'abord trouver un sous-espace  $W$  stable par  $u$  et tel que  $u$  induise l'homothétie de rapport  $\lambda$  sur le quotient  $V/W$ , puis le lemme précédent permet de remplacer  $W$  par un sous-espace qui, mieux que stable par  $u$ , est maintenant stable par  $a, b, c$ , si bien que l'homothétie de rapport  $\lambda$  sur le quotient (qui reste de dimension infinie) apparaît comme une  $(p_1, p_2, p_3)$ -somme. Mieux encore, lorsque  $u$  admet une valeur propre dominante  $\lambda$  telle que  $u - \lambda \text{id}$  ne soit pas de rang fini (ce qui nécessite que  $V$  soit de dimension indénumérable), une utilisation un peu plus subtile du lemme précédent permet aussi de trouver un découpage de  $u$  en la somme directe d'un endomorphisme sans valeur propre dominante (d'un espace de dimension infinie) et d'une homothétie de rapport  $\lambda$  d'un espace vectoriel de dimension infinie. À partir de là, si  $\lambda$  est  $(p_1, p_2, p_3)$ -acceptable, on déduit du lemme 8.27 que  $u$  est une  $(p_1, p_2, p_3)$ -somme.

Ainsi, pour caractériser les  $(p_1, p_2, p_3)$ -sommés dans  $\mathcal{L}(V)$ , il ne reste plus qu'à étudier le cas où  $u = \lambda \text{id} + v$  avec  $v$  de rang fini et  $\lambda$  un scalaire  $(p_1, p_2, p_3)$ -acceptable. Une nouvelle utilisation du lemme du sous-espace stable commun permet alors de se ramener, pour un  $\lambda$  qui soit  $(p_1, p_2, p_3)$ -acceptable, à étudier les matrices  $A \in M_n(\mathbb{K})$  telles qu'il existe un  $r \geq 0$  pour lequel  $A \oplus \lambda I_r$  est une  $(p_1, p_2, p_3)$ -somme dans  $M_{n+r}(\mathbb{K})$ . Pour un triplet général  $(p_1, p_2, p_3)$  (même avec des conditions de scindage), il n'y a malheureusement aucun espoir de savoir caractériser ces matrices, mais dans un tout petit nombre de cas particuliers cette caractérisation est possible : il s'agit du cas  $p_1 = p_2 = p_3 = X^2$  (i.e. les termes sont de carré nul) et du cas  $p_1 = p_2 = p_3 = X^2 - X$  sur les corps de caractéristique 2 (sommés de trois idempotents). C'est là que les théorèmes difficiles établis dans [5] s'avèrent indispensables. Citons pour finir la caractérisation des sommés de trois endomorphismes de carré nul en dimension infinie (nous nous limiterons à la caractéristique différente de 2, l'énoncé est un peu plus délicat dans le cas complémentaire).

**Théorème 8.29** (Théorème 5 de [6]). *Soit  $u$  un endomorphisme d'un espace vectoriel de dimension infinie  $V$ , où  $\chi(\mathbb{K}) \neq 2$ . Pour que  $u$  soit somme de trois endomorphismes de carré nul de  $V$ , il est nécessaire et suffisant que soient vraies les propriétés suivantes :*

- (i) *si  $u$  a une valeur propre dominante, celle-ci est nulle ;*
- (ii) *si  $u$  est de rang fini, alors  $\text{tr}(u) = 0$ .*

Avec les mêmes méthodes, j'ai également prolongé à un espace de dimension infinie mon théorème stipulant que tout endomorphisme est combinaison linéaire de trois projecteurs (théorème 9 de [6]).

### 8.8.4 Décompositions en produits

L'adaptation des idées précédentes pour les décompositions en produits n'a pas été tout à fait évidente. Cependant, comme on le verra, les résultats ont une similitude frappante avec le cas des sommés (et il est d'autant plus frustrant qu'il soit manifestement impossible de déduire les uns des autres!).

Tout d'abord, et selon le même principe que pour la dimension finie, il convient pour les décompositions en produits de ne considérer que des éléments quadratiques *inversibles*, si bien que l'on ne considère que des polynômes *non dégénérés*, autrement dit des polynômes  $p$  tels que  $p(0) \neq 0$ . Les décompositions

se feront donc dans le groupe linéaire  $GL(V)$  d'un espace vectoriel  $V$  de dimension infinie. Une difficulté est bien sûr l'absence de commutativité, mais elle n'a pas de conséquence excessivement fâcheuse bien qu'elle complique certaines démonstrations. Nous pouvons maintenant immédiatement citer les deux théorèmes majeurs pour les produits, qui font échos aux théorèmes 8.21 et 8.26 :

**Théorème 8.30** (Théorème 1.6 de [9]). *Soit  $u \in GL(V)$  sans valeur propre dominante, où  $V$  est un espace vectoriel de dimension infinie. Alors, quels que soient les polynômes non dégénérés et scindés  $p_1, p_2, p_3$  de degré 2, l'automorphisme  $u$  est un  $(p_1, p_2, p_3)$ -produit dans  $GL(V)$ .*

Par exemple, tout automorphisme d'un espace vectoriel de dimension infinie sans valeur propre dominante est produit de trois involutions.

**Théorème 8.31** (Théorème 1.2 de [9]). *Soit  $u \in GL(V)$ , où  $V$  est un espace vectoriel de dimension infinie. Alors, quels que soient les polynômes non dégénérés et scindés  $p_1, p_2, p_3, p_4$  de degré 2, l'automorphisme  $u$  est un  $(p_1, p_2, p_3, p_4)$ -produit dans  $GL(V)$ .*

Par exemple, tout automorphisme d'un espace vectoriel de dimension infinie est produit de quatre involutions.

Dans le théorème 8.30, on peut même affaiblir légèrement l'hypothèse de scindage en la faisant porter sur deux des polynômes seulement, de même dans le théorème 8.31 où il suffit de supposer que deux des quatres polynômes sont scindés.

La stratégie de démonstration du théorème 8.31 présente une différence fondamentale avec celle décrite au paragraphe 8.8.1 pour démontrer le théorème 8.21 : on n'y applique plus la technique de double-connexion (elle ne semble pas bien fonctionner dans le contexte présent). *A contrario*, on déduit directement le théorème 8.31 du théorème 8.30 (ce n'est pas très difficile).

Passons aux grandes lignes de la démonstration du théorème 8.30. Nous avons tout simplement adapté la stratégie de démonstration du théorème 8.26, avec quelques difficultés particulières. D'abord, il a fallu adapter la notion de stratification, car ce n'est plus la structure de  $\mathbb{K}[X]$ -module sur  $V$  héritée de  $u$  qui est pertinente (cela n'exploiterait pas bien l'hypothèse d'inversibilité de  $u$ ). Il faut plutôt utiliser la structure de  $\mathbb{K}[X, X^{-1}]$ -module, où  $\mathbb{K}[X, X^{-1}]$  est l'anneau des polynômes de Laurent à coefficients dans  $\mathbb{K}$ . Dans le cadre additif, l'objet de base était le décalage à droite, qui correspond aux  $\mathbb{K}[X]$ -modules libres de rang 1. Les  $\mathbb{K}[X, X^{-1}]$ -modules libres de rang 1 correspondent quant à eux aux décalages *bilatéraux*, autrement dit aux automorphismes  $u$  pour lesquels il existe une base  $(e_k)_{k \in \mathbb{Z}}$  telle que  $u(e_k) = e_{k+1}$  pour tout  $k \in \mathbb{Z}$ . Les automorphismes élémentaires sont alors les sommes directes de décalages bilatéraux, et on peut démontrer que pour tout couple  $(p_1, p_2)$  de polynômes non dégénérés de degré 2 dont au moins un est scindé, tout automorphisme élémentaire de  $V$  est un  $(p_1, p_2)$ -produit dans  $\mathcal{L}(V)$ .

Ensuite, dans la définition d'une stratification pour un automorphisme, on se place du point de vue de la structure de  $\mathbb{K}[X, X^{-1}]$ -module et non de celle de  $\mathbb{K}[X]$ -module. Et lorsqu'il existe une telle stratification vérifiant les « bonnes » conditions de dimension, on adapte la technique de connexion comme on l'avait adaptée pour les décompositions de matrices (voir le paragraphe 8.6.2) : ainsi, et avec les notations du paragraphe 8.8.2, on enverra  $e_\alpha$  sur  $xu^{d_\alpha-1}(e_\alpha) +$

$u^{d_{\alpha+1}-1}(e_{\alpha+1})$  si  $d_{\alpha}$  et  $d_{\alpha+1}$  sont finis, et sur  $xu^{d_{\alpha}-1}(e_{\alpha}) + e_{\alpha+1}$  si  $d_{\alpha}$  est fini mais  $d_{\alpha+1}$  est infini. Ce choix, relativement contre-intuitif en première approche, fonctionne à merveille. Et on comprend normalement assez bien ce qui coince pour tenter, à partir de cette idée, d'adapter le principe de double-connexion.

Ensuite, les idées se déroulent assez bien, le théorème 8.26 s'obtenant à partir de l'énoncé plus précis qui suit :

**Théorème 8.32.** *Soit  $u$  un automorphisme sans valeur propre dominante d'un espace vectoriel  $V$  de dimension infinie, et  $p \in \mathbb{K}[X]$  scindé de degré 2 tel que  $p(0) \neq 0$ . Il existe alors  $a \in \text{GL}(V)$  tel que  $au$  soit élémentaire et  $p(a) = 0$ .*

Ce résultat s'obtient directement par la technique de connexion lorsque  $V$  est de dimension indénombrable ou lorsque  $V$  est de torsion comme  $\mathbb{K}[X]$ -module (pour la structure associée à  $u$ ). Dans le cas complémentaire, une approche plus subtile est nécessaire, la démonstration se révélant même substantiellement plus simple que celle réalisée pour le cas des sommes (voilà l'une des rares situations où travailler sur les produits s'avère plus agréable). Nous ne donnerons volontairement aucun détail.

Terminons par l'adaptation des résultats sur le cas de valeurs propres dominantes. Tout se déroule naturellement. Pour les notations, on rappelle que la norme d'un polynôme  $p$  de degré  $d$ , notée  $N(p)$ , est définie comme  $(-1)^d \frac{p(0)}{\alpha}$ , où  $\alpha$  est le coefficient dominant de  $p$  (cette norme est le produit des racines de  $p$ , comptées avec multiplicité, dans une extension de décomposition de  $\mathbb{K}$ ).

**Définition 15.** Soit  $p_1, p_2, p_3$  trois polynômes non dégénérés de degré 2. Un scalaire  $\lambda$  est dit  $(p_1, p_2, p_3)$ -acceptable lorsque l'une au moins des conditions suivantes est vérifiée :

- (i)  $\lambda$  est un  $(p_1, p_2, p_3)$ -produit dans  $\mathbb{K}$  ;
- (ii)  $\lambda^2 = N(p_1)N(p_2)N(p_3)$ .

Par exemple, dans le problème des produits de trois involutions, les scalaires acceptables sont les racines quatrièmes de l'unité dans  $\mathbb{K}$ . Pour les produits de trois unipotents d'indice 2, les scalaires acceptables sont 1 et  $-1$ .

Et avec les mêmes techniques que pour les sommes, on obtient la caractérisation des homothéties qui sont des  $(p_1, p_2, p_3)$ -produits :

**Lemme 8.33.** *Soit  $\lambda \in \mathbb{K}$  et  $\mathcal{A}$  une  $\mathbb{K}$ -algèbre non triviale.*

- (i) *Pour que  $\lambda.1_{\mathcal{A}}$  soit un  $(p_1, p_2, p_3)$ -produit dans  $\mathcal{A}$ , il est nécessaire que  $\lambda$  soit  $(p_1, p_2, p_3)$ -acceptable.*
- (ii) *Si  $\mathcal{A}$  est l'algèbre des endomorphismes d'un espace vectoriel de dimension infinie sur  $\mathbb{K}$ , il est suffisant que  $\lambda$  soit  $(p_1, p_2, p_3)$ -acceptable pour que  $\lambda.1_{\mathcal{A}}$  soit un  $(p_1, p_2, p_3)$ -produit dans  $\mathcal{A}$ .*

On en déduit facilement que si  $u \in \text{GL}(V)$  admet une valeur propre dominante  $\lambda$  et est un  $(p_1, p_2, p_3)$ -produit, le scalaire  $\lambda$  est nécessairement  $(p_1, p_2, p_3)$ -acceptable. Il reste à examiner, lorsque  $u \in \text{GL}(V)$  admet une valeur propre dominante  $(p_1, p_2, p_3)$ -acceptable, si l'on peut démontrer que  $u$  est un  $(p_1, p_2, p_3)$ -produit. Cela repose à nouveau sur un lemme de sous-espace stable, qui est une simple adaptation du lemme 8.28 et que nous laissons au lecteur le soin de formuler. En particulier, on obtient que si  $\lambda$  est valeur propre dominante

$(p_1, p_2, p_3)$ -acceptable de  $u \in \text{GL}(V)$  et  $u - \lambda \text{id}_V$  n'est pas de rang fini, alors  $u$  est un  $(p_1, p_2, p_3)$ -produit dans  $\text{GL}(V)$ . Il reste enfin le cas des automorphismes de la forme  $\lambda \text{id}_V + v$  avec  $\lambda$  un scalaire  $(p_1, p_2, p_3)$ -acceptable et  $v$  un endomorphisme de rang fini. Comme dans le cas des sommes, cela nécessite de déterminer les matrices inversibles  $A \in \text{GL}_n(\mathbb{K})$  pour lesquelles il existe  $k \in \mathbb{N}$  tel que  $A \oplus \lambda I_k$  soit un  $(p_1, p_2, p_3)$ -produit dans  $M_{n+k}(\mathbb{K})$ . Pour un triplet général  $(p_1, p_2, p_3)$ , une telle quête est généralement vouée à l'échec (ce qui est illustré par plusieurs exemples donnés dans [8]), mais dans quelques cas emblématiques nous savons caractériser ces matrices  $A$  : il s'agit principalement du cas où chaque  $p_i$  est dans  $\{X^2 - 1, (X - 1)^2\}$ , autrement dit des problèmes de décomposition faisant intervenir des involutions et des unipotents d'indice 2. Ce cas, qui a motivé le travail [8], débouche sur le spectaculaire théorème suivant :

**Théorème 8.34** (Théorème 1.8 de [9]). *Soit  $u$  un automorphisme d'un espace vectoriel  $V$  de dimension infinie. Alors  $u$  est produit de trois involutions sauf dans les cas suivants :*

- (i)  *$u$  possède une valeur propre dominante qui n'est pas une racine quatrième de l'unité ;*
- (ii) *Il existe  $\lambda \in \mathbb{K}$  tel que  $\lambda^4 = 1$ ,  $u - \lambda \text{id}_V$  soit de rang fini, et  $\det(\lambda^{-1}u)$  n'appartienne pas au sous-groupe de  $\mathbb{K}^*$  engendré par  $\lambda$  et  $-1$ .*

Nous avons également établi des résultats tout aussi précis sur les décompositions mixtes faisant intervenir à la fois des involutions et des unipotents d'indice 2, et aussi sur les décompositions en produits de trois unipotents d'indice 2.

### 8.8.5 Perspectives

Nous terminons cette partie sur la dimension infinie en évoquant un certain nombre de problèmes ouverts auxquels notre récent travail [10] doit être vu comme une première contribution.

Dans le cadre des décompositions à trois ou quatre termes, restent en suspens deux choses :

- (i) les décompositions en produits dégénérées, i.e. le cas où l'un des polynômes annulateurs possède 0 pour racine ; ce problème exige très certainement des techniques assez différentes de celles que nous avons mises en œuvre jusque-là.
- (ii) le cas de polynômes annulateurs irréductibles sur le corps  $\mathbb{K}$ . Par exemple, on pourrait s'intéresser aux décompositions en produits de quarts-de-tour (automorphismes  $f$  tels que  $f^2 = -\text{id}$ ) sur le corps des réels.

On a mentionné précédemment le fait que nos théorèmes de décomposition cités précédemment autorisaient certains des polynômes annulateurs imposés à être irréductibles, mais jamais tous. En effet, la stratégie de base, qui consiste à se raccrocher à des endomorphismes élémentaires (la signification dépendant du contexte somme ou produit), tombe en défaut dans cette situation. Par exemple, pour les sommes, on peut démontrer que si  $p$  est irréductible unitaire de degré 2, alors un décalage à droite  $u$  ne peut pas se décomposer comme la différence de deux endomorphismes  $a$  et  $b$  annulés par  $p$ . Cela repose tout simplement sur l'observation que  $\text{Im } u$  serait, si une telle décomposition existait, stable par  $a$  (conséquence de l'identité  $au + ub = \text{tr}(p)u$ , facile à établir) si bien que  $a$  induirait un endomorphisme du quotient  $V/\text{Im } u$ , ce qui est impossible car ce

quotient est de dimension 1 et  $a$  admet un annulateur irréductible de degré 2. On démontre de manière similaire (mais c'est plus délicat car il faut en passer par des considérations sur  $u + u^{-1}$ ) que, sur le corps des réels, un décalage bilatéral  $u$  ne peut être le produit  $ab^{-1}$ , où  $a$  et  $b$  sont annulés par  $X^2 + 1$ .

Il est possible que l'on puisse établir des théorèmes équivalents aux théorèmes cités précédemment sans restriction de scindage des polynômes, mais la stratégie doit probablement être entièrement revue.

Jusque-là nous ne nous sommes intéressés en dimension infinie qu'à des décompositions à trois ou quatre termes. Il y a là une forme de paradoxe : en dimension finie, nous avons obtenu une caractérisation complète dans le problème à deux termes, mais ne disposons d'aucun résultat suffisamment général sur les décompositions à trois ou quatre termes. En dimension infinie, la situation est inversée, et c'est le problème à deux termes qui est redoutable ! Une raison est fort simple : si l'on observe la façon dont les caractérisations sont exprimées en dimension finie, on constate qu'interviennent quasi-systématiquement des considérations fines de réduction, portant sur des invariants qui caractérisent les matrices à similitude près (invariants de similitude, réduite de Jordan etc). Seuls échappent à cela un tout petit nombre de cas particuliers, où l'on peut formellement se passer de ces invariants : les produits de deux matrices involutives sont les matrices inversibles semblables à leur inverse, les sommes de deux matrices de carré nul sont les matrices semblables à leur opposée si  $\chi(\mathbb{K}) \neq 2$  ; les endomorphismes sommes de deux endomorphismes de carré nul sont les endomorphismes  $u$  pour lesquels il existe une décomposition  $V = V_1 \oplus V_2$  de l'espace vectoriel sous-jacent dont les termes sont échangés par l'action de  $u$ . Cependant ces cas sont très particuliers. Or, pour un endomorphisme d'un espace vectoriel de dimension infinie, on ne dispose en général d'aucun outil intelligible permettant de caractériser  $u$  à conjugaison près dans l'algèbre  $\mathcal{L}(V)$ . Alors, certes on pourra formuler les conjectures suivantes :

**Conjecture 1.** *Soit  $u$  un endomorphisme d'un espace vectoriel  $V$  de dimension infinie. Pour que  $u$  soit somme de deux endomorphismes de carré nul de  $V$ , il est nécessaire et suffisant qu'il existe une décomposition  $V = V_1 \oplus V_2$  telle que  $u(V_1) \subset V_2$  et  $u(V_2) \subset V_1$ .*

**Conjecture 2.** *Soit  $u$  un automorphisme d'un espace vectoriel  $V$  de dimension infinie. Pour que  $u$  soit produit de deux involutions dans  $\text{GL}(V)$ , il est nécessaire et suffisant que  $u$  soit conjugué à  $u^{-1}$  dans  $\text{GL}(V)$ .*

Cependant, l'observation des démonstrations des résultats correspondants en dimension finie indique que l'on ne peut pas se passer d'outils de réduction évolués. Il faut donc probablement se faire une raison, considérer les conjectures précédentes comme des horizons assez inaccessibles et se rabattre sur une classe particulière d'endomorphismes d'espaces vectoriels de dimension infinie, pour lesquels on dispose d'outils de réduction. J'ai ainsi choisi d'examiner la plus grande classe sur laquelle on puisse travailler, à savoir celle des endomorphismes *localement finis* d'un espace vectoriel de dimension *dénombrable*. Pour préciser immédiatement les choses, un endomorphisme  $u$  est localement fini lorsque le  $\mathbb{K}[X]$ -module associé est de torsion, autrement dit lorsque  $(u^k(x))_{k \in \mathbb{N}}$  est liée pour tout vecteur  $x$  de l'espace vectoriel sous-jacent. Dans ce cas, on dispose de l'équivalent des sous-espaces caractéristiques (pour chaque polynôme irréductible unitaire de  $\mathbb{K}[X]$ ), et on dispose d'une sorte d'équivalent



des nombres de Jordan. Lorsque  $u$  est un endomorphisme d'un espace de dimension finie,  $p$  un irréductible unitaire de  $\mathbb{K}[X]$ , et  $k \in \mathbb{N}^*$ , on note  $n_{p,k}(u)$  le nombre d'invariants primaires de  $u$  égaux à  $k$ , c'est-à-dire le nombre d'invariants de similitude de  $u$  dont la valuation  $p$ -adique est  $k$ . Ce nombre peut se calculer comme  $d^{-1}(2 \operatorname{rg} p(u)^k - \operatorname{rg} p(u)^{k+1} - \operatorname{rg} p(u)^{k-1})$ , où  $d := \deg p$ . C'est aussi la dimension, comme  $\mathbb{K}[X]/(p)$ -espace vectoriel, du noyau du morphisme (surjectif)  $\operatorname{Im} p(u)^k / \operatorname{Im} p(u)^{k-1} \rightarrow \operatorname{Im} p(u)^{k+1} / \operatorname{Im} p(u)^k$  induit par  $p(u)$ . En dimension quelconque, on dispose d'un équivalent de ces nombres  $n_{p,k}(u)$ , via la *théorie de Kaplansky*. Au lieu de se limiter à la suite  $(\operatorname{Im} p(u)^k)_{k \in \mathbb{N}}$  des images itérées, on prolonge cette suite au transfini, obtenant ainsi une famille  $(\operatorname{Im} p(u)^\alpha)_\alpha$  indexée par les ordinaux (où l'on pose  $\operatorname{Im} p(u)^0 = V$ ,  $\operatorname{Im} p(u)^\alpha = p(u)(\operatorname{Im} p(u)^{\alpha-1})$  lorsque  $\alpha$  admet un prédécesseur, et  $\operatorname{Im} p(u)^\alpha = \bigcap_{\beta < \alpha} \operatorname{Im} p(u)^\beta$  lorsque  $\alpha$  est un ordinal limite). On définit alors, pour tout ordinal  $\alpha$ , le nombre de Kaplansky  $\kappa_{p,\alpha}(u)$  comme la dimension sur  $\mathbb{K}[X]/(p)$  du noyau de l'application linéaire surjective

$$\operatorname{Im} p(u)^{\alpha+1} / \operatorname{Im} p(u)^\alpha \rightarrow \operatorname{Im} p(u)^{\alpha+2} / \operatorname{Im} p(u)^{\alpha+1}$$

induite par  $p(u)$  (noter le léger décalage avec les notations sur les nombres de Jordan, puisque les nombres de Kaplansky démarrent à 0 et non à 1). On définit enfin  $\operatorname{Im} p(u)^\infty$  comme l'intersection de tous les  $\operatorname{Im} p(u)^\alpha$ , ainsi que  $\kappa_{p,\infty}(u)$  comme la dimension sur  $\mathbb{K}[X]/(p)$  de  $\operatorname{Ker} p(u) \cap \operatorname{Im} p(u)^\infty$ . On a ainsi défini les **nombres de Kaplansky** de  $u$  (certains pouvant être des cardinaux infinis, mais ils seront au plus dénombrables si  $V$  est de dimension dénombrable).

Une conséquence d'un théorème général de Mackey et Kaplansky<sup>6</sup> est que, lorsque  $u$  est localement fini et  $V$  de dimension dénombrable, la collection des nombres  $\kappa_{p,\alpha}(u)$  caractérise  $u$  à conjugaison près dans  $\mathcal{L}(V)$ . Puisque la solution au problème à deux termes en dimension finie s'exprime en définitive uniquement à l'aide des nombres de Jordan, on peut espérer que sa résolution, dans le cas « localement fini - dimension dénombrable », puisse s'exprimer uniquement en termes des nombres de Kaplansky. Et nous avons réussi [10] à résoudre cette question dans quelques cas très particuliers :

- les sommes de deux endomorphismes de carré nul, où nous avons en particulier validé la conjecture 1 dans ce cas restreint ;
- les produits de deux involutions, et les produits de deux unipotents d'indice 2 (la conjecture 2 a, par le même coup, été validée).

Là, les méthodes sont semblables à celles mises en œuvre dans les travaux correspondants en dimension finie, et nous n'en dirons pas beaucoup plus, sinon qu'il est difficile de manipuler concrètement les endomorphismes localement finis (par exemple, on ne dispose pas de représentation aussi immédiatement claire et manipulable que la réduite de Jordan).

Nous en sommes manifestement, sur le problème évoqué, au tout début des travaux. Une étude préliminaire sur les situations moins simples, comme celle de la somme de deux idempotents, semble indiquer qu'une très grande complexité est à attendre, même en restant dans le cadre somme tout assez étroit « localement fini et dimension dénombrable ».

6. I. Kaplansky, *Infinite Abelian Groups*, University of Michigan Press, Ann Arbor, 1954.

## 8.9 Perspectives : le problème bi-quadratique

Pour conclure ce chapitre, nous revenons à notre point de départ, à savoir le problème à deux termes en dimension finie (que ce soit pour les sommes ou les produits). On a vu que dans l'algèbre des endomorphismes d'un espace vectoriel de dimension finie, ce problème est entièrement clos sauf pour la situation des produits dans le cas dégénéré. Dans ce dernier, il ne reste qu'à étudier les  $(p, q)$ -produits lorsque  $p = X^2$  ou  $p = X^2 - X$ , et  $q$  est irréductible (tous les autres cas dégénérés ont soit déjà été traités, soit se réduisent à ces situations-là). Cette situation en suspens n'est probablement pas d'une grande difficulté, et nous l'avons jusqu'ici laissée de côté.

Il reste en revanche énormément de travail sur un aspect jusque-là négligé, qui est celui des  $(p, q)$ -sommes et  $(p, q)$ -produits dans les algèbres à involutions. Considérons une  $\mathbb{K}$ -algèbre  $\mathcal{A}$  munie d'une involution  $x \mapsto x^*$ , c'est-à-dire d'un *anti*-automorphisme involutif. Notons immédiatement que lorsque  $\mathcal{A} = \mathcal{L}(E)$  pour un espace vectoriel  $E$  de dimension finie, une conséquence classique du théorème de Skolem-Noether est que l'involution envisagée est nécessairement l'adjonction associée à une forme bilinéaire symétrique ou antisymétrique non dégénérée sur  $E$  ! Dans la situation générale d'une algèbre à involution, nous pouvons considérer naturellement :

- le problème des  $(p, q)$ -sommes restreint aux éléments autoadjoints : on s'intéresse ainsi aux éléments  $x$  qui s'écrivent  $a + b$ , où  $a$  et  $b$  vérifient  $a^* = a$ ,  $b^* = b$  et  $p(a) = q(b) = 0$  ;
- le problème des  $(p, q)$ -sommes restreint aux éléments antiautoadjoints : on s'intéresse ainsi aux éléments  $x$  qui s'écrivent  $a + b$ , où  $a$  et  $b$  vérifient  $a^* = -a$ ,  $b^* = -b$  et  $p(a) = q(b) = 0$  ;
- le problème des  $(p, q)$ -produits restreint aux éléments unitaires : on s'intéresse ainsi, lorsque  $(pq)(0) \neq 0$ , aux éléments  $x$  qui s'écrivent  $ab$ , où  $a$  et  $b$  vérifient  $aa^* = a^*a = bb^* = b^*b = 1_{\mathcal{A}}$ , et  $p(a) = q(b) = 0$ .

Il existe enfin une variante du problème lorsque  $\mathcal{A}$  est munie d'une *semi-involution*, c'est-à-dire que l'on dispose déjà d'une involution non triviale du corps  $\mathbb{K}$ , une semi-involution de  $\mathcal{A}$  étant alors un antiautomorphisme semi-linéaire de la  $\mathbb{K}$ -algèbre  $\mathcal{A}$ . À nouveau, une conséquence du théorème de Skolem-Noether est que lorsque  $\mathcal{A} = \mathcal{L}(E)$ , une semi-involution est nécessairement l'adjonction relative à une forme hermitienne non dégénérée sur  $E$ .

En se limitant aux algèbres d'endomorphismes, on peut limiter le problème pour les antiautoadjoints car tous les couples  $(p, q)$  ne sont pas pertinents : en effet, comme tout endomorphisme antiautoadjoint (d'un espace de dimension finie) est semblable à son opposé, on peut se limiter au cas où  $p$  et  $q$  sont pairs (à condition que  $\mathbb{K}$  soit de caractéristique différente de 2, ce qu'on suppose systématiquement dans le problème bi-quadratique). De même, pour le problème des  $(p, q)$ -produits on se limite (sauf dans le cas hermitien) au cas où  $p$  et  $q$  sont des palindromes ou des antipalindromes.

Nous avons baptisé « bi-quadratique » le problème précédent, le premier « quadratique » étant pour les objets manipulés, le second pour les structures sous-jacentes (formes quadratiques, symplectiques ou hermitiennes).

Un théorème de Wonenburger [Won66] stipule que dans le groupe orthogonal d'une forme quadratique régulière (en caractéristique différente de 2), tout élément est produit de deux involutions, autrement dit tout élément est un  $(X^2 - 1, X^2 - 1)$ -produit. Il s'agit là d'un cas particulier du problème des

$(p, q)$ -produits unitaires ! De manière similaire, de La Cruz [dLC15] a étudié les éléments des groupes symplectiques qui sont produits de deux involutions sur le corps des complexes (Nielsen l'avait déjà fait sur un corps quelconque de caractéristique différente de 2, mais sans jamais publier ses résultats). Et de La Cruz a, dans le même article, étudié le problème des  $(p, q)$ -sommes d'autoadjoints et d'antiautoadjoints pour une forme symplectique, lorsque  $p = q = X^2$ , et toujours sur le corps des nombres complexes.

Cependant, alors que le problème à deux termes dans les algèbres de matrices avait mobilisé bon nombre de chercheurs dans les années 1990, le problème bi-quadratique n'a jusqu'à présent jamais été abordé de manière systématique. Une des difficultés réside dans ce que l'on cherche réellement. Dans le cas du problème à deux termes dans les algèbres d'endomorphismes, les résultats s'expriment en effet en termes des invariants pour la similitude, qui sont très bien connus. En revanche, si l'on prend par exemple le problème des  $(p, q)$ -sommes d'autoadjoints dans  $\mathcal{L}(E)$  lorsque  $E$  est muni d'une forme bilinéaire symétrique ou alternée non dégénérée  $b$ , il est nécessaire de disposer d'invariants des autoadjoints modulo l'action par conjugaison du groupe des isométries de  $(E, b)$  ! Fort heureusement, de tels invariants sont connus (ils ont été obtenus par divers auteurs dans les années 1970, mais le plus connu est Waterhouse). Mais, contrairement aux invariants de similitude et à la réduite de Frobenius associée (réduction par blocs qui sont des matrices compagnons), les formes canoniques pour les endomorphismes autoadjoints dans le cas général sont très peu usitées.

Depuis deux ans, je consacre l'essentiel de mes recherches au problème bi-quadratique. Ce problème est un défi technique stimulant : l'étude de quelques cas indique qu'il semble possible de le résoudre complètement sur un corps de caractéristique différente de 2, mais cette résolution sera nettement plus délicate que le problème « simple » dans les algèbres d'endomorphismes. Je me bornerai ici à citer les éléments qui sont maintenant résolus (on suppose ici que le corps de base est de caractéristique différente de 2) :

- les sommes de deux éléments de carré nul, dans toutes les situations (formes symétriques, formes alternées, formes hermitiennes ; endomorphismes autoadjoints ou antiautoadjoints) ;
- les produits de deux involutions dans les groupes orthogonaux, symplectiques et unitaires ;
- les produits de deux éléments unipotents d'indice 2 dans les groupes orthogonaux et dans les groupes symplectiques ;
- la totalité du problème des  $(p, q)$ -sommes pour les autoadjoints lorsque l'adjonction est associée à une forme symplectique ;
- les différences de deux idempotents autoadjoints.

On est donc très loin à ce stade d'avoir fait le tour du problème bi-quadratique. Le cas régulier devrait se traiter assez bien en utilisant les techniques liées aux algèbres de quaternions, et il est fort à parier que le cas exceptionnel sera redoutable. Si l'on prend par exemple le cas tout simple où  $p = q = X^2$ , on peut remarquer que la situation des algèbres d'endomorphismes se traite en trois ou quatre pages à peine, alors que les cinq variantes dans le cadre bi-quadratique nécessitent près de 80 pages d'exposition et de démonstration. Et l'essentiel de la difficulté, dans ce cas, réside dans l'analyse des endomorphismes nilpotents : alors que tout endomorphisme nilpotent d'un espace vectoriel de dimension finie est somme de deux endomorphismes de carré nul, il y a des contraintes subtiles sur un endomorphisme nilpotent qui est autoadjoint pour

une certaine forme bilinéaire symétrique non dégénérée  $b$  pour qu'il soit somme de deux endomorphismes  $b$ -autoadjoints de carré nul !

## Chapitre 9

# Travaux inclassables

- [1] C. de Seguins Pazzis, Invariance of simultaneous similarity and equivalence of matrices under extension of the ground field. *Linear Algebra Appl.* **433-3** (2010) 618–624.
- [2] C. de Seguins Pazzis, On small matrix subalgebras with a trivial centralizer. *Linear Multilinear Algebra* **59-7** (2011) 761–771.
- [3] C. de Seguins Pazzis, To what extent is a large space of matrices not closed under product? *Linear Algebra Appl.* **435-11** (2011) 2708–2721.
- [4] C. de Seguins Pazzis, When does a linear map belong to at least one orthogonal or symplectic group? *Linear Algebra Appl.* **436-5** (2012) 1385–1405.
- [5] C. de Seguins Pazzis, The semigroup generated by the similarity class of a singular matrix. *Comm. Algebra.* **40-9** (2012) 3414–3419.
- [6] C. de Seguins Pazzis, On commuting matrices and exponentials. *Proc. Amer. Math. Soc.* **141** (2013) 763–774.
- [7] C. de Seguins Pazzis, Commutators from a hyperplane of matrices. *Electron. J. Linear Algebra* **27** (2014) 39–54.
- [8] C. de Seguins Pazzis, The space of all  $p$ -th roots of a nilpotent complex matrix is path-connected. *Linear Algebra Appl.* **596** (2020) 106–116.

### 9.1 Une démonstration originale du lemme de Noether-Deuring [1]

Soit  $\mathbb{K} - \mathbb{L}$  une extension de corps, et  $(A_i)_{i \in I}$  et  $(B_i)_{i \in I}$  deux familles de matrices de  $M_n(\mathbb{K})$ , indexées par le même ensemble  $I$  (éventuellement infini). On dit que  $(A_i)_{i \in I}$  et  $(B_i)_{i \in I}$  sont  $\mathbb{K}$ -semblables lorsqu'il existe  $P \in \text{GL}_n(\mathbb{K})$  telle que  $B_i = PA_iP^{-1}$  pour tout  $i \in I$ . Dans ce cas,  $(A_i)_{i \in I}$  et  $(B_i)_{i \in I}$  sont évidemment  $\mathbb{L}$ -semblables. Il est extrêmement connu (utiliser typiquement la réduction de Frobenius) que la réciproque est vraie lorsque  $I$  est un singleton.

Il est beaucoup moins classique que la réciproque soit vraie, cependant elle est connue des spécialistes de théorie des représentations (d'algèbres) sous le nom de *lemme de Noether-Deuring*. Ce lemme s'énonce comme suit :

**Théorème 9.1** (Lemme de Noether-Deuring). *Soit  $\mathbb{L}$  un corps,  $\mathbb{K}$  un sous-corps de  $\mathbb{L}$  et  $n$  un entier naturel non nul. Soit  $\mathcal{A}$  une  $\mathbb{K}$ -algèbre, et  $\varphi : \mathcal{A} \rightarrow M_n(\mathbb{K})$*

et  $\psi : \mathcal{A} \rightarrow M_n(\mathbb{K})$  deux morphismes de  $\mathbb{K}$ -algèbres. On suppose que  $\varphi$  et  $\psi$  sont  $\mathbb{L}$ -conjugués, autrement dit qu'il existe  $P \in GL_n(\mathbb{L})$  telle que  $P\varphi(x) = \psi(x)P$  pour tout  $x \in \mathcal{A}$ . Alors  $\varphi$  et  $\psi$  sont  $\mathbb{K}$ -conjugués.

Pour montrer que ce résultat implique celui que l'on souhaite, il suffit de prendre la  $\mathbb{K}$ -algèbre libre  $\mathcal{A}$  sur une famille  $(x_i)_{i \in I}$  d'indéterminées non commutant indexée par  $I$ , et les morphismes  $\varphi$  et  $\psi$  définis par  $\varphi(x_i) = A_i$  et  $\psi(x_i) = B_i$  pour tout  $i \in I$ . On peut aussi démontrer que le résultat que l'on souhaite établir implique le lemme de Noether-Deuring, mais nous laisserons cet aspect de côté ici.

### 9.1.1 La démonstration classique

La démonstration classique du lemme de Noether-Deuring se déroule en deux étapes. D'abord, lorsque  $\mathbb{K}$  est infini, ou fini de cardinal supérieur ou égal à  $n$ , la méthode est extrêmement classique et reprend une méthode connue dans le cas d'un singleton : on part de  $Q \in GL_n(\mathbb{K})$  telle que  $Q\varphi(x) = \psi(x)Q$  pour tout  $x \in \mathcal{A}$ . On extrait une base  $(\lambda_1, \dots, \lambda_p)$  du sous- $\mathbb{K}$ -espace vectoriel de  $\mathbb{L}$  engendré par les coefficients de  $Q$ , et l'on décompose  $Q = \sum_{k=1}^p \lambda_k P_k$  où  $P_1, \dots, P_p$  sont dans  $M_n(\mathbb{K})$ . On obtient alors que, pour tout  $(u_1, \dots, u_p)$  dans  $\mathbb{K}^p$ , la matrice  $R(u_1, \dots, u_p) := \sum_{k=1}^p u_k P_k$  vérifie  $R(u_1, \dots, u_p)\varphi(x) = \psi(x)R(u_1, \dots, u_p)$  pour tout  $x \in \mathcal{A}$ . Pour conclure, il suffit de montrer que l'on peut choisir  $u_1, \dots, u_p$  de telle sorte que  $R(u_1, \dots, u_p)$  soit inversible. C'est classique en remarquant que  $(u_1, \dots, u_p) \in \mathbb{L}^p \mapsto \det R(u_1, \dots, u_p)$  est polynomiale homogène de degré  $n$  et non nulle.

Si  $\mathbb{K}$  est fini et  $\mathbb{L}$  infini, on passe par une clôture algébrique  $\overline{\mathbb{L}}$  de  $\mathbb{L}$ , puis on peut trouver une extension intermédiaire  $\mathbb{K} - \mathbb{K}' - \overline{\mathbb{L}}$  avec  $\mathbb{K}'$  fini de cardinal aussi grand que l'on souhaite. Il est donc clair qu'il ne reste qu'à traiter le cas d'une extension finie.

Supposons alors  $\mathbb{L}$  de degré fini  $d$  sur  $\mathbb{K}$ . Considérons le  $\mathcal{A}$ -module  $M$  (respectivement  $N$ ) dont l'ensemble sous-jacent est  $\mathbb{K}^n$  et la loi externe définie comme  $x.X := \varphi(x)X$  (respectivement  $x.X := \psi(x)X$ ) pour tout  $x \in \mathcal{A}$  et tout  $X \in \mathbb{K}^n$ . On définit de même des  $\mathcal{A}$ -modules  $M_{\mathbb{L}}$  et  $N_{\mathbb{L}}$  en partant de  $\mathbb{L}^n$  au lieu de  $\mathbb{K}^n$ . À l'aide d'une base de  $\mathbb{L}$  sur  $\mathbb{K}$ , on démontre sans peine que  $M_{\mathbb{L}}$  est isomorphe, en tant que  $\mathcal{A}$ -module, à la somme directe externe de  $d$  copies de  $M$ .

Un isomorphisme de  $\mathcal{A}$ -modules de  $M$  sur  $N$  est en particulier un isomorphisme pour les  $\mathbb{K}$ -espaces vectoriels sous-jacents. C'est donc une fonction de la forme  $\theta : X \mapsto QX$  où  $Q \in GL_n(\mathbb{K})$ , vérifiant

$$\forall x \in \mathcal{A}, \forall X \in \mathbb{K}^n, \theta(\varphi(x)X) = \psi(x)\theta(X)$$

ce qui se traduit par  $\forall x \in \mathcal{A}, Q\varphi(x) = \psi(x)Q$ . En faisant de même sur  $\mathbb{L}$ , on conclut provisoirement que l'hypothèse voulant que  $\varphi$  et  $\psi$  soient  $\mathbb{L}$ -conjugués suffit à garantir que les  $\mathcal{A}$ -modules  $M_{\mathbb{L}}$  et  $N_{\mathbb{L}}$  sont isomorphes; d'autre part il nous suffit de démontrer que les  $\mathcal{A}$ -modules  $M$  et  $N$  sont isomorphes pour obtenir  $P \in GL_n(\mathbb{K})$  telle que  $\forall x \in \mathcal{A}, P\varphi(x) = \psi(x)P$ .

On considère alors des décompositions  $M \simeq M_1 \oplus \dots \oplus M_p$  et  $N \simeq N_1 \oplus \dots \oplus N_q$  où  $M_1, \dots, M_p, N_1, \dots, N_q$  sont des  $\mathcal{A}$ -modules indécomposables. Ainsi,  $M_{\mathbb{L}}$  est isomorphe à la somme directe de  $d$  copies de  $M_1$ , de  $d$  copies de  $M_2$ , etc, de  $d$  copies de  $M_p$ . De même,  $N_{\mathbb{L}}$  est isomorphe à la somme directe de  $d$

copies de  $N_1$ , de  $d$  copies de  $N_2$ , etc, de  $d$  copies de  $N_q$ . Soit  $W$  un  $\mathcal{A}$ -module indécomposable arbitraire. Notons  $k$  (respectivement  $k'$ ) le nombre d'indices  $i \in \llbracket 1, p \rrbracket$  (respectivement  $j \in \llbracket 1, q \rrbracket$ ) tels que  $M_i$  (respectivement  $N_j$ ) soit isomorphe à  $W$ . Il y a donc, dans la décomposition précédente de  $M_{\mathbb{L}}$  (respectivement de  $N_{\mathbb{L}}$ ), exactement  $dk$  (respectivement  $dk'$ ) facteurs isomorphes à  $W$ . Le théorème de Krull-Schmidt montre alors que  $dk = dk'$ , d'où  $k = k'$ . En faisant varier  $W$  parmi  $M_1, \dots, M_p, N_1, \dots, N_q$ , on en déduit facilement que  $p = q$  et qu'il existe une permutation  $\sigma$  de  $\llbracket 1, p \rrbracket$  telle que  $M_i$  soit isomorphe à  $N_{\sigma(i)}$  pour tout  $i \in \llbracket 1, p \rrbracket$ . À partir de là, on reconstitue sans peine un isomorphisme de  $\mathcal{A}$ -modules de  $M$  sur  $N$ , ce qui montre que  $\varphi$  et  $\psi$  sont  $\mathbb{K}$ -conjugués.

### 9.1.2 Les grandes lignes de ma démonstration du lemme de Noether-Deuring

Il est remarquable que la démonstration classique du lemme de Noether-Deuring n'utilise aucun outil particulier de réduction des matrices, sinon les considérations basiques sur la décomposition de Fitting (qui est au cœur de la démonstration du théorème de Krull-Schmidt). Pourtant, pour le cas où  $I$  est un singleton, la démonstration classique passe par une compréhension fine des classes de similitude.

L'originalité de cette démonstration réside dans l'utilisation de considérations de réduction, non sur les matrices  $A_i, B_i$  elles-mêmes mais sur les matrices de passage!

On écarte immédiatement le cas où  $\mathbb{K}$  est de cardinal supérieur ou égal à  $n$ , cas que l'on traite exactement comme dans la démonstration classique. On se réduit alors au cas où  $\mathbb{K}$  est fini et on traite uniquement le cas d'une extension *quadratique* séparable  $\mathbb{L}$  de  $\mathbb{K}$ . En effet, si  $\mathbb{K}$  est fini, on peut prendre une tour d'extensions quadratiques séparables  $\mathbb{K} = \mathbb{K}_0 \subset \mathbb{K}_1 \subset \dots \subset \mathbb{K}_n \subset \overline{\mathbb{L}}$ , et  $\mathbb{K}_n$  est de cardinal au moins  $2^n$  donc au moins  $n$ . Si l'on dispose du résultat pour chaque extension quadratique séparable, alors on peut d'abord remonter de  $\mathbb{L}$  à  $\overline{\mathbb{L}}$ , puis redescendre à  $\mathbb{K}_n$ , et enfin descendre successivement à  $\mathbb{K}_{n-1}, \dots, \mathbb{K}_0 = \mathbb{K}$ , grâce à notre hypothèse de validité sur les extensions quadratiques séparables.

Il ne reste plus qu'à établir le lemme pour une extension quadratique séparable  $\mathbb{K} - \mathbb{L}$  (notons qu'il n'est pas nécessaire de supposer  $\mathbb{K}$  fini). Dans ce cas, on va jouer sur la réduction de la matrice de passage fournie. Introduisons l'unique  $\sigma$  dans  $\text{Gal}(\mathbb{L}/\mathbb{K}) \setminus \{\text{id}\}$ . Supposons disposer d'une matrice  $P \in \text{GL}_n(\mathbb{L})$  telle que  $P\varphi(x) = \psi(x)P$  pour tout  $x \in \mathcal{A}$ . En appliquant l'automorphisme  $\sigma$  coefficient par coefficient, on voit que  $\forall x \in \mathcal{A}$ ,  $P^\sigma \varphi(x) = \psi(x)P^\sigma$ , donc  $\varphi(x)$  commute avec  $(P^\sigma)^{-1}P$  pour tout  $x \in \mathcal{A}$ . Si le commutant de  $(P^\sigma)^{-1}P$  était inclus dans celui de  $P$ , on en déduirait que  $\varphi = \psi$ , ce qui donnerait immédiatement le résultat voulu : l'idée de [1] consiste à se ramener à cette situation après modification de  $P$ .

Remarquons que l'on peut, pour n'importe quel couple  $(R, S) \in \text{GL}_n(\mathbb{K})^2$ , remplacer  $\varphi$  et  $\psi$  respectivement par  $x \mapsto S\varphi(x)S^{-1}$  et  $x \mapsto R\psi(x)R^{-1}$ , ce qui a pour effet de remplacer  $P$  par  $RPS^{-1}$ .

Décomposons alors  $P = A + \varepsilon B$  où  $(A, B) \in \text{M}_n(\mathbb{K})^2$  et  $\varepsilon \in \mathbb{L} \setminus \mathbb{K}$ . Ce qui précède montre que l'on peut remplacer le couple  $(A, B)$  par  $(RAS^{-1}, RBS^{-1})$  pour un couple arbitraire  $(R, S) \in \text{GL}_n(\mathbb{K})^2$  sans rien changer à notre problème. Or les classes d'équivalence pour l'équivalence simultanée des couples de ma-

trices sont connues : elles relèvent de la *réduction de Kronecker-Weierstrass* (voir l'annexe de [1] pour une version simplifiée de cette réduction). Dans la classe d'équivalence de  $(A, B)$  existe un couple particulier  $(A_1, B_1)$  dit réduit, et pour un tel couple on démontre que la matrice  $P_1 := A_1 + \varepsilon B_1$  commute avec toute matrice commutant avec  $(P_1^\sigma)^{-1}P_1$ . Nous renvoyons à [1] pour les détails de démonstration.

### 9.1.3 Épilogue

Lorsque j'ai conçu la démonstration ci-dessus, j'étais totalement ignorant du lemme de Noether-Deuring. Et il s'avère que bien des chercheurs qui publient dans les revues d'algèbre linéaire semblent tout aussi ignorants que je l'étais alors du fait que ce résultat est connu depuis bien longtemps. En conséquence, [1] est régulièrement cité depuis dix ans comme la référence sur le sujet, ce qui me met un peu mal à l'aise ...

## 9.2 Sur les sous-algèbres de matrices carrées à commutant trivial [2]

Le commutant d'un sous-ensemble  $V$  de  $M_n(\mathbb{K})$  est défini comme l'ensemble

$$\{M \in M_n(\mathbb{K}) : \forall A \in V, AM = MA\}$$

des matrices qui commutent avec tous les éléments de  $V$ . C'est aussi le commutant de la sous-algèbre engendrée par  $V$ .

Il est bien connu que le commutant d'une matrice  $A \in M_n(\mathbb{K})$  est de dimension au moins  $n$ . Dès que l'on prend deux matrices  $A$  et  $B$  de  $M_n(\mathbb{K})$ , il est possible que le commutant de  $\{A, B\}$  soit trivial, c'est-à-dire réduit aux matrices scalaires. Tout simplement, il est suffisant pour cela que  $A$  et  $B$  engendrent l'algèbre  $M_n(\mathbb{K})$ , par ailleurs il est bien connu que de telles matrices existent, par exemple  $A = P_c$ , matrice de permutation associée à la permutation circulaire  $c = (1\ 2\ \dots\ n)$ , et  $B = E_{1,1}$  (matrice élémentaire), où encore  $A = J_n$  et  $B = (J_n)^T$ , où  $J$  désigne la matrice de Jordan nilpotente de taille  $n$  (triangulaire supérieure stricte).

Avec mon ancien collègue Saab Abou-Jaoudé, nous nous sommes interrogés en 2008 sur les *sous-algèbres* de  $M_n(\mathbb{K})$  dont le commutant est trivial. Plus précisément, quelle est d'abord la dimension minimale  $t_n$  d'une telle sous-algèbre (vue comme sous-espace vectoriel) ? Le premier exemple qui nous vint est celui de l'algèbre des matrices de la forme

$$\begin{bmatrix} \lambda & [?]_{1 \times (n-1)} \\ 0 & \mu I_n \end{bmatrix}, \quad \text{où } \lambda \text{ et } \mu \text{ sont dans } \mathbb{K},$$

dont il est facile de montrer la trivialité du commutant. Cette sous-algèbre est de dimension  $n+1$ . Cela suggérerait une dimension minimale fonction croissante de  $n$ , mais en réalité nous avons rapidement vu, grâce à une suggestion de Pierre Mazet, que l'on pouvait trouver des exemples de dimension bien plus petite. Le cas  $n = 2$  se traite de manière élémentaire et nous n'en dirons rien ici. Lorsque  $n \geq 3$ , la solution dépend uniquement de la parité de  $n$ . Les résultats suivants ont été établis dans [2].



**Théorème 9.2** (Théorème 1.3 de [2]). *Soit  $n \geq 3$  entier. Alors  $t_n = 4$  si  $n$  est impair, et  $t_n = 5$  si  $n$  est pair.*

**Proposition 9.3** (Proposition 1.4 de [2]). *Soit  $n = 2p$  un entier pair supérieur ou égal à 4. Le sous-espace vectoriel de  $M_n(\mathbb{K})$  engendré par les cinq matrices*

$$\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & I_p \end{bmatrix}, \quad \begin{bmatrix} 0 & I_p \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & J_p \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & (J_p)^T \\ 0 & 0 \end{bmatrix}$$

*est alors une sous-algèbre de commutant trivial.*

Le caractère de sous-algèbre est facile à obtenir. Quant à la trivialité du commutant, elle s'obtient rapidement comme suit : tout élément du commutant de la première matrice est de la forme  $X \oplus Y$  avec  $X, Y$  dans  $M_p(\mathbb{K})$ , puis une telle matrice commute avec la troisième si et seulement si  $X = Y$ , enfin dans ce cas  $X \oplus Y$  commute avec les quatrième et cinquième matrices si et seulement si  $X$  commute avec  $J_p$  et  $J_p^T$ , ce qui impose que  $X$  soit scalaire (voir nos toutes premières observations).

L'exemple fondamental dans le cas d'une dimension paire est un peu plus subtil, et nous n'en détaillerons pas la démonstration :

**Proposition 9.4** (Proposition 1.5 de [2]). *Soit  $n = 2p + 1$  un entier impair supérieur ou égal à 3. Notons  $C_p := \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}$  et  $D_p := \begin{bmatrix} 0 & I_p \\ 0 & 0 \end{bmatrix}$  dans  $M_{p,p+1}(\mathbb{K})$ . Le sous-espace vectoriel  $\mathcal{H}_{2p+1}$  de  $M_n(\mathbb{K})$  engendré par les quatre matrices*

$$\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & I_{p+1} \end{bmatrix}, \quad \begin{bmatrix} 0 & C_p \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & D_p \\ 0 & 0 \end{bmatrix}$$

*est alors une sous-algèbre de commutant trivial.*

Mieux, dans ce cas l'exemple donné est essentiellement le seul :

**Théorème 9.5** (Proposition 1.6 de [2]). *Soit  $n$  un entier impair supérieur ou égal à 3. Toute sous-algèbre de dimension 4 de  $M_n(\mathbb{K})$  à commutant trivial est alors conjuguée à  $\mathcal{H}_n$  ou à sa transposée.*

Il semble en revanche inabordable d'obtenir une caractérisation équivalente dans le cas où  $n$  est pair. Tout simplement, dans l'exemple proposé dans la proposition 9.3, on pourrait remplacer les blocs supérieurs droits  $J_p$  et  $J_p^T$  par les termes de n'importe quelle paire de matrices de  $M_p(\mathbb{K})$  qui engendre cette algèbre!

Donnons enfin un aperçu des méthodes de démonstration. Une première idée est de se ramener au cas d'un corps algébriquement clos : par extension du corps des scalaires, on peut en effet voir que le nombre  $t_n$  ne peut pas augmenter, et vu les exemples cités précédemment il est bien clair qu'il suffit d'établir le théorème 9.2 dans le cas algébriquement clos (en revanche, il y a une subtilité pour le théorème 9.5, car la conjugaison de deux sous-algèbres après extension des scalaires ne garantit nullement qu'elles étaient conjuguées au départ - la situation est tout à fait différente de celle évoquée dans le paragraphe 9.1 sur le lemme de Noether-Deuring).

Supposons donc  $\mathbb{K}$  algébriquement clos. On écarte alors le cas d'une sous-algèbre unispectrale (c'est-à-dire dans laquelle toute matrice possède une unique valeur propre). Dans ce cas, on démontre en effet assez facilement qu'une telle

sous-algèbre est trigonalisable, et lorsqu'elle est carrément constituée de matrices triangulaires supérieures à coefficients diagonaux tous égaux, il est facile de voir que la matrice élémentaire  $E_{1,n}$  figure dans son commutant. On peut donc se limiter au cas d'une sous-algèbre  $\mathcal{A}$  non unispectrale, puis (utilisation d'un projecteur spectral) en déduire que  $\mathcal{A}$  possède un idempotent non trivial, qu'après réduction on peut supposer égal à  $P = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}$  pour un certain entier  $k \in \llbracket 1, n-1 \rrbracket$ . Ensuite, on décompose tout élément de  $\mathcal{A}$  par blocs selon le même format que  $P$ , et on récupère quatre sous-espaces vectoriels  $PAP$ ,  $PAP'$ ,  $P'AP$  et  $P'AP'$  (pour  $P' := I - P$ ), qui reconstituent  $\mathcal{A}$  comme somme directe. En supposant  $\mathcal{A}$  de très petite dimension, on peut alors facilement analyser ces sous-espaces vectoriels, et en discutant selon les cas fabriquer une matrice non scalaire dans le commutant de  $\mathcal{A}$ .

La démonstration du théorème 9.5 est essentiellement fondée sur la même idée; intervient à la fin de celle-ci la classification de Kronecker-Weierstrass des pinceaux de matrices, le couple  $(C_p, D_p)$  vu dans le théorème 9.5 étant un pinceau indécomposable classique dans cette classification.

### 9.3 Sur l'instabilité par multiplication des grands sous-espaces de matrices carrées [3]

La problématique de ce travail [3] est un peu singulière dans ma bibliographie, mais ce travail est directement relié à mes recherches sur les sous-espaces de matrices singulières. Il s'inscrit aussi dans une étude générale des propriétés des sous-espaces de matrices de petite codimension.

Ici, le point de départ est une conséquence classique de théorèmes de classification des sous-algèbres irréductibles de  $M_n(\mathbb{K})$  (et en particulier du théorème de Burnside si  $\mathbb{K}$  est algébriquement clos). Cette conséquence classique est le fait que la plus petite codimension d'une sous-algèbre non triviale de  $M_n(\mathbb{K})$  est  $n - 1$ . Un exemple de sous-algèbre ayant cette codimension est celle formée des matrices de la forme

$$\begin{bmatrix} [?]_{(n-1) \times (n-1)} & [?]_{(n-1) \times 1} \\ [0]_{1 \times (n-1)} & ? \end{bmatrix}.$$

Une autre façon de présenter le résultat est la suivante : si l'on part d'un sous-espace vectoriel  $V$  de  $M_n(\mathbb{K})$  de codimension au plus  $n - 2$ , alors  $V$  est une partie génératrice de l'algèbre  $M_n(\mathbb{K})$ . Autrement dit, l'espace vectoriel  $M_n(\mathbb{K})$  est engendré par les *mots* formés sur  $V$  (c'est-à-dire les produits finis d'éléments de  $V$ ).

Dans [3] nous avons voulu raffiner le résultat précédent, dans deux directions principalement :

**Théorème 9.6** (Théorème 2 de [3]). *Soit  $V$  un sous-espace vectoriel de  $M_n(\mathbb{K})$  de codimension inférieure à  $n - 2$ . Alors  $M_n(\mathbb{K}) = \text{Vect}\{AB \mid (A, B) \in V^2\}$ .*

**Théorème 9.7** (Théorème 6 de [3]). *Soit  $\mathcal{V}$  un sous-espace affine de  $M_n(\mathbb{K})$  de codimension inférieure à  $n - 2$ . Alors*

$$M_n(\mathbb{K}) = \mathcal{V}^{(\infty)} := \{A_1 \cdots A_p \mid p \geq 1, (A_1, \dots, A_p) \in \mathcal{V}^p\},$$

*autrement dit  $\mathcal{V}$  engendre le monoïde multiplicatif  $(M_n(\mathbb{K}), \times)$ .*

La démonstration du théorème 9.6 est assez simple, le point crucial est tout de même le passage par la case intermédiaire des crochets de Lie. Il s'agit alors de démontrer que, sous les hypothèses du théorème, on a  $\mathfrak{sl}_n(\mathbb{K}) = \text{Vect}\{[A, B] \mid (A, B) \in V^2\}$ . Pour cela, on utilise un argument d'orthogonalité, en étudiant les  $C \in M_n(\mathbb{K})$  tels que  $\text{tr}([A, B]C) = 0$  pour tout  $(A, B) \in V^2$ . Le théorème du rang permet de voir, grâce au calcul de la dimension du commutant de  $C$  en fonction de ses invariants de similitude, que  $C$  est scalaire. Ensuite, il est facile de conclure en montrant que les produits de deux éléments de  $V$  ne sont pas tous de trace nulle (considérations d'orthogonalité pour une forme bilinéaire symétrique non dégénérée).

Le théorème 9.7 est beaucoup plus délicat. Il utilise de manière cruciale le théorème de Dieudonné sur les sous-espaces affines de matrices singulières (voir la section 2.5.1). On notera que l'on parle ici de sous-espaces affines, bien que la conclusion la plus intéressante porte sur les sous-espaces vectoriels. On a déjà souligné dans le chapitre 2 l'intérêt de cette technique d'extension aux sous-espaces affines, car elle permet des raisonnements par récurrence sur la taille des matrices. Il y a trois grandes étapes, dont la plus délicate est la première :

- (i) On montre que  $\mathcal{V}^{(\infty)}$  contient toutes les matrices inversibles.
- (ii) On montre que  $\mathcal{V}^{(\infty)}$  contient au moins une matrice de rang  $n - 1$ .
- (iii) On montre que  $\mathcal{V}^{(\infty)}$  contient toutes les matrices de rang  $n - 1$ , puis toutes les matrices singulières.

À partir de (i) et (ii), l'obtention de (iii) est très facile et nous n'en dirons rien. En revanche, les étapes (i) et (ii) sont parfaitement indépendantes. Nous commencerons par l'étape (ii), qui est la plus simple. On note  $V^\perp$  l'orthogonal de la direction  $V$  de  $\mathcal{V}$  pour  $(A, B) \mapsto \text{tr}(AB)$ . On observe d'abord que, quitte à conjuguer  $\mathcal{V}$  par une matrice de permutation, on peut supposer que  $V^\perp$  ne contient aucune matrice de rang 1 dont les  $n - 1$  premières lignes sont nulles (sinon on trouverait  $\dim V^\perp \geq n$ ), ce qui assure que  $M \mapsto L_n(M)$  surjecte  $\mathcal{V}$  sur  $M_{1,n}(\mathbb{K})$ . Il existe en particulier dans  $\mathcal{V}$  une matrice de dernière ligne nulle, et l'ensemble  $\mathcal{W}$  formé de ces matrices s'identifie naturellement, après suppression de la dernière ligne, à un sous-espace affine de  $M_{n-1,n}(\mathbb{K})$  dont la codimension dans  $M_{n-1,n}(\mathbb{K})$  est celle de  $\mathcal{V}$  dans  $M_n(\mathbb{K})$ . La généralisation du théorème de Dieudonné aux matrices rectangulaires (théorème 2.5) assure alors que  $\mathcal{W}$  contient une matrice de rang  $n - 1$ , ce qui assure l'étape (ii).

Pour l'étape (i), on procède par récurrence sur  $n$ . La situation semi-favorable est celle où  $M \mapsto C_1(M)$  surjecte  $\mathcal{V}$  sur  $\mathbb{K}^n$ . Dans ce cas, on introduit le sous-espace affine  $\mathcal{V}'$  formé des matrices de  $\mathcal{V}$  qui sont de la forme

$$M = \begin{bmatrix} 1 & L(M) \\ [0]_{(n-1) \times 1} & K(M) \end{bmatrix}.$$

On note que  $\text{codim } K(\mathcal{V}') < n - 1$ , donc le théorème de Dieudonné (théorème 2.2) donne que  $K(\mathcal{V}')$  contient une matrice inversible, puis  $\mathcal{V}'$  aussi. Par invariance du problème par multiplication à gauche par une matrice inversible, on trouve plus généralement que pour tout  $C \in \mathbb{K}^n \setminus \{0\}$  le sous-espace  $\mathcal{V}$  contient une matrice inversible de première colonne  $C$ . Et, à partir de là, il suffira d'établir l'énoncé suivant :

(E) : Pour tout  $L_1 \in M_{1,n-1}(\mathbb{K})$  et tout  $P \in \text{GL}_{n-1}(\mathbb{K})$ , la matrice inversible  $\begin{bmatrix} 1 & L_1 \\ [0]_{(n-1) \times 1} & P \end{bmatrix}$  est produit d'éléments de  $\mathcal{V}$ .

Poursuivons et remarquons en outre que si  $\text{codim } K(\mathcal{V}') < n - 2$ , alors l'hypothèse de récurrence s'applique et permet donc de voir que tout élément de  $M_{n-1}(\mathbb{K})$  est produit d'éléments de  $K(\mathcal{V}')$ . À partir de là, et en remarquant que la direction de  $\mathcal{V}$  contient nécessairement une matrice de la forme 
$$\begin{bmatrix} 0 & L_0 \\ [0]_{(n-1) \times 1} & [0]_{(n-1) \times (n-1)} \end{bmatrix}$$
 avec  $L_0 \neq 0$ , on s'offre le petit degré de liberté qu'il faut, et il n'est pas très difficile, en utilisant convenablement les différentes matrices accessibles, de valider l'énoncé (E).

Reste à s'assurer que  $\text{codim } K(\mathcal{V}') < n - 2$ , ce qui n'est pas vrai en général. Cependant, si c'est faux alors on est assuré (théorème du rang) que  $\text{codim } \mathcal{V} = n - 2$  et que la direction de  $\mathcal{V}$  contient absolument toutes les matrices de la forme 
$$\begin{bmatrix} 0 & L \\ [0]_{(n-1) \times 1} & [0]_{(n-1) \times (n-1)} \end{bmatrix}$$
 avec  $L \in M_{1, n-1}(\mathbb{K})$ . Et plus généralement, en conjuguant  $\mathcal{V}$ , on en déduit que la conclusion voulue s'obtient sauf si  $\text{codim } \mathcal{V} = n - 2$  et, pour tout  $x \in \mathbb{K}^n$  telle que  $M \in \mathcal{V} \mapsto Mx \in \mathbb{K}^n$  soit surjective, la direction  $V$  contient toutes les matrices de trace nulle et d'image  $\mathbb{K}x$ .

Pour conclure la démonstration, on se place dans ce cas défavorable et on en analyse les conséquences sur la structure de  $\mathcal{V}$ . D'abord, la condition de surjectivité est invalidée précisément pour les vecteurs  $x$  de  $\mathbb{K}^n \setminus \{0\}$  tels que  $V^\perp$  contienne au moins une matrice d'image  $\mathbb{K}x$ . La condition  $\dim V^\perp \leq n - 2$  coïncide donc ces vecteurs  $x$  dans un sous-espace de codimension  $n - 2$  de  $\mathbb{K}^n$ . Et à partir de là il est facile de montrer que  $V$  contient toutes les matrices de trace nulle, puis la condition  $\text{codim } \mathcal{V} = n - 2$  force  $n = 3$  (car le cas  $n = 2$  est évidemment trivial), puis  $\mathcal{V} = \{M \in M_3(\mathbb{K}) : \text{tr } M = a\}$  pour un certain  $a \in \mathbb{K}$ . Ce cas très particulier se traite alors de manière plus directe, en fabriquant ce qu'il faut de matrices de dilatations et de transvections dans  $\mathcal{V}^{(\infty)}$ . Et ainsi s'achève la démonstration du théorème 9.7.

Terminons en citant le dernier résultat majeur de [3] :

**Théorème 9.8** (Théorème 14 de [3]). *Soit  $n \geq 3$ , et  $H_1$  et  $H_2$  deux hyperplans vectoriels de  $M_n(\mathbb{K})$ . Toute matrice de  $M_n(\mathbb{K})$  se décompose alors sous la forme  $AB$  où  $A \in H_1$  et  $B \in H_2$ .*

Et en particulier, pour n'importe quel hyperplan vectoriel  $H$  de  $M_n(\mathbb{K})$ , lorsque  $n \geq 3$  toute matrice de  $M_n(\mathbb{K})$  est un mot formé sur deux lettres dans  $H$ .

Le style de démonstration du théorème 9.8 est assez voisin de celui employé pour le théorème 9.7, et je n'en dirai rien de plus.

## 9.4 Sur la classe de similitude d'une matrice d'un groupe orthogonal ou symplectique en caractéristique 2 [4]

Ce travail porte sur les propriétés spectrales des endomorphismes remarquables pour une forme bilinéaire symétrique ou alternée non dégénérée. La thèse remonte à Jordan, qui avait commencé l'étude sur les corps finis.

Situons rapidement le problème : on se donne un automorphisme  $u$  d'un espace vectoriel  $E$  de dimension finie. On se demande s'il existe une forme bilinéaire symétrique (ou alternée) non dégénérée pour laquelle  $u$  est une isométrie.

La réponse ne dépend que de la classe de similitude de  $u$ , donc des invariants de similitude de  $u$ . Sur un corps de caractéristique différente de 2, il est probable que le résultat soit connu depuis le milieu du vingtième siècle mais nous ne sommes pas parvenus à trouver une référence précise.

Ensuite, il faut citer immédiatement que pour une isométrie (linéaire)  $u$  d'une forme bilinéaire symétrique non dégénérée, l'adjoint de  $u$  (pour cette forme) est conjugué à  $u^t$ , donc à  $u$  (énoncé classique en algèbre linéaire sur un corps), et ainsi  $u^{-1}$  est semblable à  $u$ . Mais cette dernière condition s'avère insuffisante. Rappelons donc les résultats connus avant mon travail :

**Théorème 9.9.** *Soit  $u$  un automorphisme d'un espace vectoriel  $E$  de dimension finie sur un corps  $\mathbb{K}$  de caractéristique différente de 2. Pour qu'il existe une forme bilinéaire symétrique (respectivement, alternée) non dégénérée pour laquelle  $u$  est une isométrie, il faut et suffit que soient validées les deux conditions suivantes :*

- (i)  $u$  est semblable à  $u^{-1}$  ;
- (ii) Pour tout entier  $k$  pair (respectivement, impair) et tout  $\eta = \pm 1$ , le nombre de blocs de Jordan de taille  $k$  de  $u$  pour la valeur propre  $\eta$  est pair.

Ce résultat fait en quelque sorte écho au théorème de Frobenius, qui stipule que tout endomorphisme d'un espace vectoriel de dimension finie est autoadjoint pour au moins une forme bilinéaire symétrique non dégénérée<sup>1</sup>, ainsi qu'aux résultats de Stenzel<sup>2</sup>, qui étudia les endomorphismes qui sont autoadjoints/antiautoadjoints pour au moins une forme bilinéaire symétrique/alternée non dégénérée.

En réalité, et toujours en caractéristique différente de 2, on dispose depuis Wall<sup>3</sup> d'une classification complète des classes de conjugaison dans un groupe orthogonal ou symplectique. En revanche, en caractéristique 2 on n'a aucun espoir d'obtenir une telle classification, la difficulté résidant dans la structure des isométries unipotentes (celles qui sont trigonalisables avec 1 pour seule valeur propre), dont on ne comprend pas bien les classes de conjugaison dans le groupe des isométries.

Mon modeste travail sur la question a consisté à établir l'équivalent du théorème 9.9 en caractéristique 2, ainsi que quelques raffinements. Dans ce cadre, il faut observer qu'il y a non pas deux mais trois problèmes, puisqu'on peut considérer des formes symplectiques, des formes bilinéaires symétriques non-dégénérées plus générales, mais aussi le problème de l'orthogonalité pour une forme quadratique régulière. Rappelons quelques faits élémentaires pour les non-spécialistes de formes quadratiques en caractéristique 2.

- Une forme quadratique est une fonction de la forme  $x \mapsto b(x, x)$  où  $b$  est une forme bilinéaire sur un espace vectoriel  $E$  de dimension finie.
- Pour une telle forme quadratique  $q$ , la forme polaire  $b_q : (x, y) \mapsto q(x + y) - q(x) - q(y)$  est bilinéaire et alternée (caractéristique 2 oblige). On dit que  $q$  est régulière lorsque  $b_q$  est non dégénérée (donc symplectique).

---

1. G. Frobenius, Über die mit einer Matrix vertauschbaren Matrizen. Sitzungsber. Preuss. Akad. Wiss. (1910) 3–15.

2. H. Stenzel, Über die Darstellbarkeit einer Matrix als Produkt von zwei symmetrischer Matrizen, als Produkt von zwei alternierenden Matrizen und als Produkt von einer symmetrischen und einer alternierenden Matrix. Math. Z. **15** (1922) 1–25.

3. G.E. Wall, On the conjugacy classes in orthogonal, symplectic and unitary groups. J. Austral. Math. Soc. **3-1** (1963) 1–62.

Plusieurs formes quadratiques partagent en général la même forme polaire (les formes quadratiques de forme polaire donnée forment un espace affine de même dimension que  $E$ ). Le groupe orthogonal de  $q$  est inclus dans celui de sa forme polaire, et ils ne se confondent pas en général.

Voici maintenant notre premier résultat :

**Théorème 9.10** (Théorèmes 2 et 5 de [4]). *Soit  $u$  un automorphisme d'un espace vectoriel  $E$  de dimension finie sur un corps  $\mathbb{K}$  de caractéristique 2. Pour que  $u$  soit une isométrie pour une forme symplectique sur  $E$ , il est nécessaire et suffisant que  $u$  vérifie les deux conditions suivantes :*

- (i)  $u$  est semblable à  $u^{-1}$  ;
- (ii) Pour tout entier impair  $k \geq 1$ , le nombre de blocs de Jordan de  $u$  de taille  $k$  pour la valeur propre 1 est pair.

*En outre, lorsque ces conditions sont vérifiées, il existe même une forme quadratique régulière sur  $E$  pour laquelle  $u$  est une isométrie.*

Il est remarquable que l'existence d'une forme quadratique régulière pour laquelle  $u$  est une isométrie n'apporte rien de plus que celle d'une forme symplectique. Ce n'est pas complètement étonnant lorsqu'on connaît l'astuce classique voulant que si  $u$  ne possède pas la valeur propre 1 et  $q$  est une forme quadratique de forme polaire  $b$  telle que  $u$  soit dans le groupe orthogonal de  $q$ , alors  $q(x - u(x)) = b(x, u(x))$  pour tout  $x \in E$ . Ainsi, dès que  $u$  ne possède pas 1 pour valeur propre et  $b$  est une forme symplectique pour laquelle  $u$  est une isométrie, il est naturel d'imaginer que  $u$  pourrait être une isométrie pour la forme quadratique  $q : x \mapsto b((u - \text{id})^{-1}(x), u((u - \text{id})^{-1}(x)))$  (ce qu'elle est, et cette forme quadratique est régulière).

Pour les groupes orthogonaux de formes bilinéaires symétriques, la condition est – on pouvait s'y attendre au vu du cas de la dimension 1 – légèrement moins restrictive :

**Théorème 9.11** (Théorème 3 de [4]). *Soit  $u$  un automorphisme d'un espace vectoriel  $E$  de dimension finie sur un corps  $\mathbb{K}$  de caractéristique 2. Pour que  $u$  soit une isométrie pour une forme bilinéaire symétrique non dégénérée sur  $E$ , il est nécessaire et suffisant que  $u$  vérifie les deux conditions suivantes :*

- (i)  $u$  est semblable à  $u^{-1}$  ;
- (ii) Pour tout entier impair  $k \geq 3$ , le nombre de blocs de Jordan de  $u$  de taille  $k$  pour la valeur propre 1 est pair.

Voyons maintenant quelques idées de démonstration. Le caractère nécessaire s'obtient par des arguments géométriques et des calculs d'invariants. Par exemple, dans le théorème 9.10, la condition (ii) s'obtient en démontrant que  $\text{rg}(u - \text{id})^{2k}$  est pair pour tout  $k \in \mathbb{N}$ , ce qui s'obtient en construisant astucieusement une forme bilinéaire alternée de même rang que  $\text{rg}(u - \text{id})^{2k}$  (plus précisément,  $(x, y) \mapsto b(u^k(x), (u - \text{id})^{2k}(y))$  est une telle forme, lorsque  $b$  est symplectique et  $u$  est une  $b$ -isométrie).

Pour la réciproque, on utilise la réduction primaire de  $u$ , ce qui permet de se limiter à l'une des trois situations *indécomposables* suivantes :

- (a)  $u$  est cyclique de polynôme minimal  $(X - 1)^{2k}$  pour un  $k \geq 1$  ;

- (b)  $u$  est cyclique de polynôme minimal  $P^k$  pour un palindrome irréductible unitaire  $P$  qui n'a pas 1 pour racine, et un entier  $k \geq 1$  ;
- (c)  $u$  est la somme directe de deux automorphismes cycliques dont les polynômes minimaux sont réciproques l'un de l'autre.

Le dernier cas est le plus classique, il se traite grâce à une simple astuce matricielle : pour toute matrice inversible  $A \in \text{GL}_n(\mathbb{K})$ , la matrice blocs  $A \oplus (A^T)^{-1}$  appartient au groupe symplectique usuel  $\text{Sp}_{2n}(\mathbb{K})$ , et est même orthogonale au sens de la forme quadratique  $q : \begin{bmatrix} X \\ Y \end{bmatrix} \in \mathbb{K}^{2n} \mapsto X^T Y$ . Pour la première situation, nous avons utilisé une démonstration matricielle, en résolvant l'équation linéaire  $J_{2k}(1)^T A J_{2k}(1) = A$  et en cherchant parmi les solutions une matrice alternée inversible, que l'on construit de manière itérative en descendant sur les antidiagonales (Proposition 11 de [4]).

La situation (b) est la plus complexe de toutes. Notre démonstration est astucieuse . . . peut-être même trop, car il y a plus élémentaire (mais nous ne l'avons découvert que plus tard)! D'abord, par une méthode d'extension des scalaires qui rappelle celle utilisée pour le lemme de Noether-Deuring (paragraphe 9.1), on peut réduire la situation au cas où le corps de base est fini (car sinon on passe dans une clôture algébrique, où la situation (b) ne se produit jamais). L'intérêt de cette réduction est que tous les polynômes sont séparables sur un corps fini, ce qui permet d'une part d'utiliser une décomposition de Jordan-Chevalley pour voir que  $C(P^k)$  est semblable au produit de Kronecker  $J_k(1) \otimes C(P)$  (cet argument ne fonctionne plus dans le cas inséparable). Ensuite, lorsque  $k = 1$  on peut observer que la résolution passe nécessairement par des considérations de formes hermitiennes. On peut en effet munir  $\mathbb{L} := \mathbb{K}[X]/(P)$  de l'involution non triviale  $\sigma$  envoyant la classe de  $X$  sur son inverse, et constater que l'endomorphisme de multiplication par la classe de  $X$  du  $\mathbb{K}$ -espace vectoriel  $\mathbb{L} := \mathbb{K}[X]/(P)$  est à la fois cyclique de polynôme minimal  $P$  et unitaire pour la forme sesquilinéaire  $(a, b) \mapsto \varepsilon \sigma(a)b$  où  $\varepsilon$  est un élément non nul à bien choisir ; et réciproquement on peut démontrer que toute forme symplectique sur  $\mathbb{L}$  pour laquelle la multiplication par la classe de  $X$  est une isométrie est de la forme  $(a, b) \mapsto \text{Tr}_{\mathbb{L}/\mathbb{K}}(\varepsilon \sigma(a)b)$  où  $\varepsilon$  est un élément non nul antihermitien (i.e.  $\sigma(\varepsilon) = -\varepsilon$ ). Par ailleurs, par calcul sur les coefficients de matrices, nous avons démontré que  $J_k(1)$  était unitaire pour au moins une forme hermitienne non dégénérée sur  $\mathbb{L}^n$ , ce qui permet ensuite de conclure.

Mon travail se conclut par deux raffinements. D'abord, sur un corps parfait de caractéristique 2 et en dimension paire, il existe deux types d'isométrie de formes bilinéaires symétriques non dégénérées : celles qui sont symplectiques, et celles qui ne le sont pas (alors qu'il en existe un unique type en dimension impaire). On peut donc affiner la question posée sur un tel corps : étant donné un automorphisme  $u$  d'un espace vectoriel de dimension finie, à quelle condition existe-t-il une forme bilinéaire symétrique non dégénérée mais non symplectique pour laquelle  $u$  est une isométrie? La réponse, très simple, indique qu'il est nécessaire et suffisant que  $u$  appartienne à un groupe symplectique (propriété caractérisée dans le théorème 9.10) et que 1 soit valeur propre de  $u$  (théorème 17 de [4]).

Le dernier raffinement étudié porte sur les groupes orthogonaux de formes quadratiques régulières. Lorsque  $\mathbb{K}$  est un corps fini de caractéristique 2 et  $n$  un entier naturel non nul, il existe à équivalence près deux types de formes qua-

dratiques régulières : celles qui sont hyperboliques, et les autres. Un invariant algébrique permet facilement de les distinguer, c'est l'invariant d'Arf, qui joue en caractéristique 2 le rôle que joue le discriminant en caractéristique différente de 2. On cherche alors naturellement, pour un automorphisme  $u \in \text{GL}(E)$ , à quelle condition il existe une forme quadratique régulière hyperbolique (respectivement, non hyperbolique) pour laquelle  $u$  est une isométrie. Terminons donc en citant complètement le théorème 24 de [4], qui répond complètement à cette question :

**Théorème 9.12** (Théorème 24 de [4]). *Soit  $\mathbb{K}$  un corps fini de caractéristique 2. Soit  $u$  un automorphisme d'un  $\mathbb{K}$ -espace vectoriel de dimension finie qui appartient à un groupe symplectique. On note  $N(u)$  le nombre d'invariants primaires de  $u$  de la forme  $P^k$  avec  $P$  palindrome irréductible de degré au moins 2, et  $k \in \mathbb{N}^*$  impair.*

- (a) *Si 1 est valeur propre de  $u$  alors il existe deux formes quadratiques régulières  $q_1$  et  $q_2$  sur  $V$ , dont exactement une est hyperbolique, telles que  $u$  soit une  $q_1$ -isométrie et une  $q_2$ -isométrie.*
- (b) *Si 1 n'est pas valeur propre de  $u$  et  $N(u)$  est pair (respectivement, impair) alors est hyperbolique (respectivement, non hyperbolique) toute forme quadratique régulière  $q$  pour laquelle  $u$  est une isométrie.*

Je terminerai par un commentaire global sur cet article. En relisant la plupart de mes travaux passés, je n'ai en général pas à rougir de ce que j'ai produit. Cet article ne m'a pas laissé la même impression. Le recul m'a en effet permis, notamment via mes travaux actuels sur le problème bi-quadratique, de beaucoup mieux comprendre la situation. Avec cette expérience accumulée, il est clair que je n'écrirais absolument pas cet article aujourd'hui comme à l'époque. Un point parmi d'autres : la démonstration du fait que la matrice compagnon  $C(P^k)$  soit dans un groupe symplectique dès que  $P$  est un palindrome irréductible de degré au moins 2, et  $k$  un entier naturel. D'abord, pour le cas  $k = 1$  on n'a nullement besoin de séparabilité sur  $P$  (rappelons que je m'étais réduit à un corps fini pour profiter de cette séparabilité); pour le voir, il suffit de remplacer la forme  $\text{Tr}_{\mathbb{L}/\mathbb{K}}$  (qui peut être nulle!) par  $f \circ \text{Tr}_{\mathbb{L}/\mathbb{K}_0}$ , où  $\mathbb{K}_0$  est le sous-corps de  $\mathbb{L}$  formé des éléments hermitiens (i.e. fixé par  $\sigma$ ), et  $f$  n'importe quelle forme  $\mathbb{K}$ -linéaire non nulle sur  $\mathbb{K}_0$ . L'utilisation de la décomposition de Jordan-Chevalley (nécessitant la séparabilité de  $P$ ) aurait aussi pu être contournée en introduisant le polynôme (irréductible)  $R$  de degré  $d$  tel que  $P = X^d R(X + X^{-1})$  et en raisonnant directement sur l'anneau quotient  $\mathbb{K}[X]/(P^k)$ , que l'on met en correspondance avec le quotient de  $\mathbb{K}[X, X^{-1}]$  par l'idéal engendré par  $R(X + X^{-1})^k$ .

## 9.5 Sur le sous-monoïde engendré par la classe de similitude d'une matrice singulière [5]

Le travail que je vais évoquer ici est vaguement relié à mes travaux sur les décompositions de matrices (chapitre 8). Il est plus directement lié au théorème de John Erdos<sup>4</sup>, qui stipule que toute matrice singulière (sur un corps) est

---

4. J. Erdos, On products of idempotent matrices. Glasgow Math. J. **8-2** (1967) 118–122.



produit de matrices idempotentes. Ce théorème a été raffiné par Ballantine<sup>5</sup> dans le sens d'une précision sur le nombre précis d'idempotents nécessaires dans une telle décomposition.

Ici, l'objet est légèrement différent : il s'agit d'explicitier le sous-monoïde de  $(M_n(\mathbb{K}), \times)$  engendré par la classe de similitude d'une matrice donnée. Précisons tout de suite que par sous-monoïde, on entend une partie stable par  $\times$  mais ne contenant pas nécessairement l'élément neutre  $I_n$  pour la multiplication. Lorsque la matrice  $M$  envisagée est inversible, on ne dispose pas de résultat intelligible sur la question, notamment parce qu'il y a des considérations de déterminant à prendre en compte. En revanche, lorsque  $M$  est singulière, on peut expliciter le sous-monoïde engendré par sa classe de similitude, autrement dit l'ensemble des matrices qui sont produits d'une liste finie *non vide* de matrices semblables à  $M$ .

**Théorème 9.13.** *Soit  $M \in M_n(\mathbb{K})$  non inversible, de rang  $p$ . Le sous-monoïde de  $(M_n(\mathbb{K}), \times)$  engendré par la classe de similitude de  $M$  est alors l'ensemble des matrices de  $M_n(\mathbb{K})$  de rang inférieur à  $p$ .*

Ce résultat était connu bien avant que je n'intervienne<sup>6</sup>, mais j'en ai donné une démonstration simplifiée utilisant de manière judicieuse les réduites de Frobenius et de Jordan.

Je me contenterai d'évoquer assez brièvement les grandes étapes de la démonstration, sachant que le résultat de Erdos cité plus haut intervient de manière cruciale. On part donc d'une matrice singulière  $M \in M_n(\mathbb{K})$  de rang  $p < n$ , et on note  $\langle M \rangle$  le sous-monoïde qu'elle engendre. On note que  $\langle M \rangle$  est stable par multiplication et conjugaison, et on établit successivement :

- (i) que  $\langle M \rangle$  contient une matrice de rang  $p$  pour laquelle 0 est valeur propre semi-simple ;
- (ii) que  $\langle M \rangle$  contient un idempotent de rang  $p$  ;
- (iii) que  $\langle M \rangle$  contient toutes les matrices de rang  $p$  pour lesquelles 0 est valeur propre semi-simple ;
- (iv) que  $\langle M \rangle$  contient toutes les matrices de rang  $p$  ;
- (v) et enfin que  $\langle M \rangle$  contient toutes les matrices de rang inférieur ou égal à  $p$ .

C'est l'étape (ii)  $\Rightarrow$  (iii) qui utilise de manière critique le théorème d'Erdos. L'implication (iv)  $\Rightarrow$  (v) est quant à elle tout à fait élémentaire. La démonstration du point (i) est fondée sur une réduction de Jordan de la partie nilpotente de  $M$ , sur l'observation que tout bloc de Jordan est semblable à son transposé, et que si  $J_k$  est le bloc de Jordan nilpotent (triangulaire supérieur) de taille  $k$ , alors  $J_k^T J_k = I_{k-1} \oplus 0$ . L'implication (i)  $\Rightarrow$  (ii) est établie en observant que  $C(Q)(KC(Q)K^{-1}) = I_{k-1} \oplus 0$  lorsque  $Q$  est unitaire de degré  $k$  et vérifie  $Q(0) = 0$ , et  $K$  désigne la matrice de permutation associée à  $i \mapsto k + 1 - i$ . Enfin, l'implication (iii)  $\Rightarrow$  (iv) repose principalement sur l'observation que  $C(X^k - X)(I_{k-1} \oplus 0) = J_k^T$  et que 0 est valeur propre semi-simple de  $C(X^k - X)$  et de  $I_{k-1} \oplus 0$ .

5. C.S. Ballantine, Products of idempotent matrices. *Linear Algebra Appl.* **19** (1976) 81–86.

6. J. Araújo, F.C. Silva, Semigroups of linear endomorphisms closed under conjugation. *Comm. Algebra.* **28-8** (2000) 3679–3689.

L. Grunenfelder, M. Omladič, H. Radjavi, A. Sourour, Semigroups generated by similarity orbits. *Semigroup Forum.* **62-3** (2001) 460–472.

## 9.6 Sur l'exponentielle de matrice [6]

Il est archi-classique que l'identité  $e^{x+y} = e^x e^y = e^y e^x$  est vraie pour deux éléments  $x$  et  $y$  d'une algèbre de Banach *pourvu qu'ils commutent*. On sait en revanche que la commutation n'est pas une condition nécessaire pour avoir cette propriété.

Mon travail sur la question est issu d'une discussion sur le forum les-mathematiques.net à l'automne 2010. Dans cette discussion fut cité un travail de Gérard Bourgeois<sup>7</sup> dans lequel celui-ci prétendait démontrer, pour des matrices  $A$  et  $B$  de  $M_3(\mathbb{C})$ , que l'identité  $e^{kA+B} = e^{kA} e^B = e^B e^{kA}$  suffit à obtenir la commutation de  $A$  et  $B$ . Pourtant, dans ce même fil de discussion, Jean-Louis Tu fournit un contre-exemple à cette affirmation (sans que j'arrive à comprendre par quelle magie il était parvenu à celui-ci). Fortement intrigué par ce phénomène, j'entrepris d'étudier la question plus avant, et après quelques jours je parvins au résultat suivant, qui est le cœur de [6] :

**Théorème 9.14.** *Soit  $A, B$  dans  $M_n(\mathbb{C})$  telles que  $e^{kA+lB} = e^{kA} e^{lB}$  pour tous entiers (relatifs)  $k$  et  $l$ . Alors  $A$  et  $B$  commutent.*

Bien sûr, un développement limité en 0 permet de voir que si  $e^{tA+tB} = e^{tA} e^{tB}$  pour tout réel  $t$ , alors  $A$  et  $B$  commutent. Ici, toute la difficulté est qu'on ne fait d'hypothèse que sur les combinaisons linéaires à coefficients entiers, ce qui semble interdire toute méthode de calcul différentiel. On notera aussi le renforcement subtil de la propriété de Bourgeois : on ne se limite pas à  $l = 1$  (noter aussi qu'on ne demande pas formellement l'identité  $e^{kA+lB} = e^{lB} e^{kA}$  mais qu'elle découle facilement par inversion de celle que l'on a supposée!).

Ce théorème est profondément difficile. La partie la plus délicate de la démonstration repose sur une adaptation d'arguments de la démonstration par Tosio Kato<sup>8</sup> du célèbre théorème de Motzkin et Taussky qui suit :

**Théorème 9.15** (Motzkin-Taussky<sup>9</sup>). *Soit  $A$  et  $B$  deux matrices de  $M_n(\mathbb{C})$  dont toute combinaison linéaire est diagonalisable. Alors  $A$  et  $B$  commutent.*

Avant de voir comment intervient ce théorème, je vais expliquer les grandes étapes de ma démonstration. Je restituerai le cheminement intellectuel réel que j'ai réalisé, si bien que les étapes que je vais ici décrire ne respectent pas l'ordre d'exposition de [6]. Ma première idée a été de réaliser une réduction à une situation plus simple, situation où  $A$  et  $B$  sont à spectre dans  $2i\pi\mathbb{Z}$ , en exploitant toutes sortes de commutations : pour peu que l'on puisse trouver une décomposition non triviale de  $\mathbb{C}^n$  en sous-espaces supplémentaires non triviaux stables à la fois par  $A$  et  $B$ , on peut en effet descendre sur les dimensions pour raisonner par récurrence. Il suffit donc de traiter la situation où  $(A, B)$  est *indécomposable*. Dans cette situation, et sous réserve que la fonction  $\gamma : (\lambda, \mu) \in \text{Sp}(e^A) \times \text{Sp}(e^B) \mapsto \lambda\mu$  soit injective, on démontre facilement, par la seule formule  $e^{A+B} = e^A e^B = e^B e^A$ , que tout sous-espace caractéristique de  $e^B$  est une somme de sous-espaces caractéristiques de  $e^{A+B}$ .

7. G. Bourgeois, On commuting exponentials in low dimensions. Linear Algebra Appl. **423** (2007) 277–286.

8. T. Kato. Perturbation Theory for Linear Operators. Grundlehren der Mathematischen Wissenschaften, Deuxième édition, Springer-Verlag, 1976.

9. T.S. Motzkin, O. Taussky, Pairs of matrices with property L (II). Trans. Amer. Math. Soc. **80** (1955) 387–401.

Toujours sous cette hypothèse d'injectivité, on en déduit que tous les sous-espaces caractéristiques de  $e^B$  sont stables par  $A+B$ , et donc aussi par  $A$  (puisqu'ils le sont par  $B$ ). Partant de là, on trouve que  $e^B$  n'a qu'une valeur propre (hypothèse d'indécomposabilité), et par symétrie le même argument donne que  $e^A$  aussi. On se ramène alors facilement au cas où cette valeur propre est 1, ce qui donne une première réduction. En général, la fonction  $\gamma$  n'est pas injective, mais il n'est pas difficile de se réduire à cette situation en remplaçant  $A$  par  $kA$  pour un entier  $k \geq 1$  bien choisi (ce qui conserve les hypothèses et fournit une conclusion suffisante pour obtenir la commutation voulue).

Ensuite, il m'a semblé naturel d'étudier un cas encore plus particulier : celui où non seulement  $e^A$  et  $e^B$  sont unipotentes, mais où mieux elles valent toutes deux  $I_n$ . Autrement dit, on se penche sur le cas où  $kA+lB$  est diagonalisable à spectre dans  $2i\pi\mathbb{Z}$  pour tout  $(k,l) \in \mathbb{Z}^2$ . Il s'agit alors de montrer que  $A$  et  $B$  commutent, et on est donc ramené après division par  $2i\pi$  à l'énoncé suivant, qui ne parle plus formellement d'exponentielle de matrice :

**Théorème 9.16** (Lemme 7 de [6]). *Soit  $A, B$  deux matrices de  $M_n(\mathbb{C})$  telles que  $kA+lB$  soit diagonalisable à spectre dans  $\mathbb{Z}$  pour tout  $(k,l) \in \mathbb{Z}^2$ . Alors  $AB=BA$ .*

C'est dans la démonstration de ce résultat qu'intervient la méthode de Kato pour le théorème de Motzkin-Taussky. Cette méthode permet en effet, comme je l'ai expliqué succinctement dans [6], d'obtenir le raffinement crucial suivant :

**Théorème 9.17** (Lemme 7 de [6]). *Soit  $A, B$  deux matrices de  $M_n(\mathbb{C})$  telles que  $B$  soit diagonalisable et  $A+z_0B$  soit diagonalisable pour tout point exceptionnel  $z_0$  du faisceau de matrices  $z \mapsto A+zB$ . Alors  $AB=BA$ .*

Il faut expliquer ce qu'est un tel point exceptionnel. Tout simplement, la matrice générique  $A+XB$  (à coefficients dans le corps de fractions rationnelles  $\mathbb{C}(X)$ ) possède un nombre  $p$  de valeurs propres dans une clôture algébrique de  $\mathbb{C}(X)$  que l'on appelle le nombre générique de valeurs propres du faisceau  $z \mapsto A+zB$ . Un nombre complexe  $z_0$  est dit exceptionnel pour ce faisceau lorsque  $A+z_0B$  a strictement moins de  $p$  valeurs propres (distinctes). La clef de la démonstration du théorème 9.16 réside dans l'observation (absolument non triviale!) que, sous les hypothèses de ce théorème, les points exceptionnels du faisceau  $z \mapsto A+zB$  sont tous des rationnels, la condition de diagonalisabilité en ces points étant alors immédiatement vérifiée. Cette rationalité des points exceptionnels fut obtenue en revenant à la source des arguments de Kato, via la *propriété L* de Motzkin et Taussky<sup>10</sup>. La propriété L stipule simplement, pour un couple  $(A, B)$  de matrices de  $M_n(\mathbb{C})$ , qu'il existe des fonctions  $\mathbb{C}$ -affines  $f_1, \dots, f_n$  de  $\mathbb{C}$  dans lui-même telles que les valeurs propres de  $A+zB$  soient systématiquement  $f_1(z), \dots, f_n(z)$  (en comptant les valeurs propres à mesure de leur multiplicité); c'est en quelque sorte une forme affaiblie de la co-trigonalisabilité. On peut noter que cette propriété est vérifiée dès que toute combinaison linéaire de  $A$  et  $B$  est nilpotente, ce qui ne nécessite pas que  $A$  et  $B$  soient co-trigonalisables (à cet effet, voir le chapitre 6). En outre, en général les valeurs propres de  $A+zB$  ne s'expriment même pas globalement comme des fonctions entières de la variable  $z$ , car il y a des phénomènes de ramification.

10. T.S. Motzkin, O. Taussky, Pairs of matrices with property L. Trans. Amer. Math. Soc. **73** (1952) 108–114.

**Lemme 9.18** (Proposition 5 de [6]). *Soit  $A, B$  dans  $M_n(\mathbb{C})$  telles que  $\text{Sp}(kA + B) \subset \mathbb{Z}$  pour tout  $k \in \mathbb{N}$ . Alors  $(A, B)$  possède la propriété L.*

La démonstration de ce lemme repose sur l'écriture des valeurs propres de  $A + zB$ , pour  $z$  non nul voisin de 0, grâce à des séries de Puiseux. En utilisant le fait que pour tout  $k \in \mathbb{N}^*$ , les matrices  $kA + B$  et  $(k + 1)A + B$  sont à valeurs propres entières, on parvient à simplifier ces séries de Puiseux et à montrer une version locale de la propriété L pour  $z$  voisin de 0, la globalisation étant alors facile par un argument de prolongement analytique.

Enfin, à partir de la propriété L, il est tout à fait facile d'obtenir la rationalité des points exceptionnels de  $z \mapsto A + zB$  sous l'hypothèse du lemme précédent, ce qui achève la démonstration du théorème 9.16.

À ce stade, on n'a pas totalement terminé car il reste un chaînon manquant entre le cas où  $e^A = e^B = I_n$  (théorème 9.16) et le cas plus général où  $e^A$  et  $e^B$  sont seulement supposées unipotentes, cas auquel on s'était ramené précédemment. Pour traiter le deuxième cas à l'aide du premier, on utilise assez naturellement les décompositions de Jordan-Dunford  $A = D + N$  et  $B = D' + N'$ . Il n'est alors pas très difficile de démontrer que le couple  $(D, D')$  vérifie les hypothèses du théorème 9.16 puis d'obtenir la commutation de  $A$  et  $B$ .

Notons enfin que notre travail ne se limite pas à établir le théorème 9.14 (dont le théorème 9.16 n'est qu'un cas particulier), il montre également que la condition de Bourgeois  $\forall k \in \mathbb{N}, e^{kA+B} = e^{kA}e^B = e^B e^{kA}$  est suffisante pour que  $(A, B)$  possède la propriété L, et j'ai généralisé cet énoncé à des listes plus longues de matrices.

## 9.7 Sur l'écriture d'une matrice de trace nulle comme commutateur [7]

Il est bien connu que le sous-espace vectoriel de  $M_n(\mathbb{K})$  engendré par les crochets de Lie  $[A, B] = AB - BA$ , où  $A$  et  $B$  parcourent  $M_n(\mathbb{K})$ , est l'hyperplan  $\mathfrak{sl}_n(\mathbb{K})$  des matrices de trace nulle. C'est aussi un exercice classique de taupe que d'observer que, pour peu que  $\mathbb{K}$  soit de caractéristique nulle, toute matrice de  $\mathfrak{sl}_n(\mathbb{K})$  est un crochet de Lie : classiquement, on montre dans ce cas que toute matrice de trace nulle est semblable à une matrice de diagonale nulle<sup>11</sup> puis on observe que toute matrice de diagonale nulle est dans l'image de  $\text{ad}_D := [D, -]$  pour  $D$  diagonale à valeurs propres simples. Il semble que la première publication de ce résultat remonte à Shoda<sup>12</sup>.

Plus de vingt ans après la publication de Shoda, Albert et Muckenhoupt<sup>13</sup> ont généralisé le résultat à n'importe quel corps. Leur stratégie est fondée sur l'étude de l'image de  $\text{ad}_C$  lorsque la matrice  $C$  est cyclique. En effet, on sait en général que l'image de  $\text{ad}_C$  est l'orthogonal pour  $(A, B) \mapsto \text{tr}(AB)$  du commutant de  $C$  (en effet,  $\text{ad}_C$  est antiautoadjointe pour cette forme bilinéaire symétrique non dégénérée). C'est une bonne idée de choisir  $C$  cyclique, car c'est ce qui produit des commutants de dimension minimale, et donc des images de

11. P.A. Fillmore, On similarity and the diagonal of a matrix. Amer. Math. Monthly **76** (1969) 167–169.

12. K. Shoda, Einige Sätze über Matrizen. Jap. J. Math. **13** (1936) 361–365.

13. A.A. Albert, B. Muckenhoupt, On matrices of trace zero. Michigan Math. J. **4** (1957) 1–3.

$\text{ad}_C$  de dimension maximale. En outre, dans ce cas le commutant de  $C$  est engendré par  $(C^k)_{0 \leq k \leq n-1}$ , si bien que l'image de  $\text{ad}_C$  est définie par le système d'équations linéaires  $\forall k \in \llbracket 0, n-1 \rrbracket, \text{tr}(C^k M) = 0$ . Pour une matrice de Hessenberg de trace nulle  $M$ , et  $J$  la matrice nilpotente de Jordan de  $M_n(\mathbb{K})$ , on peut ainsi voir que la condition  $M \in \text{Im}(\text{ad}_J)$  se résume à l'équation linéaire  $\sum_{k=1}^{n-1} m_{k+1,k} = 0$ . Et, sauf pour un tout petit nombre de cas particuliers – où il convient d'adapter légèrement la méthode – la matrice  $M$  à décomposer est systématiquement semblable à une matrice de Hessenberg vérifiant cette condition.

Ultérieurement, Thompson<sup>14</sup> a amélioré le théorème d'Albert et Muckenhoupt comme suit :

**Théorème 9.19.** *Soit un entier  $n \geq 3$ . Pour toute matrice  $M \in \mathfrak{sl}_n(\mathbb{K})$ , il existe  $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$  tel que  $M = [A, B]$ .*

Ainsi, non seulement toute matrice de trace nulle est un crochet de Lie, mais c'est un crochet de Lie de matrices de trace nulle. Inspiré par l'équivalent pour les produits (mon théorème 9.8), j'ai entrepris de généraliser le théorème de Thompson à n'importe quel hyperplan. Voici mon résultat, où pour simplifier les notations on écrit  $[A, B] := \{[A, B] \mid A \in \mathcal{A}, B \in \mathcal{B}\}$  pour des sous-ensembles arbitraires  $\mathcal{A}$  et  $\mathcal{B}$  de  $M_n(\mathbb{K})$ .

**Théorème 9.20.** *Soit  $n \geq 3$  un entier, et  $\mathbb{K}$  un corps ayant au moins 4 éléments. Soit  $\mathcal{H}$  un hyperplan vectoriel de  $M_n(\mathbb{K})$  et  $A \in \mathfrak{sl}_n(\mathbb{K})$ . Alors  $A \in [\mathcal{H}, \mathcal{H}]$ .*

Il semble très difficile de se passer de l'hypothèse de cardinalité sur  $\mathbb{K}$ , de nombreux points de démonstration l'utilisent en effet. En tout état de cause, il apparaît qu'une autre méthode de démonstration que la nôtre semble nécessaire pour traiter les corps à 2 ou 3 éléments, si tant est que le résultat soit vrai pour ceux-ci.

En abordant le théorème 9.20, j'étais en fait ignorant que Thompson avait déjà abordé le cas particulier de l'hyperplan  $\mathfrak{sl}_n(\mathbb{K})$ , et j'ai trouvé une démonstration originale de son résultat (mais je ne l'ai pas publiée).

Je vais maintenant aborder les grandes lignes de la démonstration du théorème 9.20. Commençons par quelques observations simples. D'abord on va systématiquement se représenter  $\mathcal{H}$  comme l'orthogonal d'une matrice  $B$  non nulle pour la forme bilinéaire  $(M, N) \mapsto \text{tr}(MN)$ . Le problème est alors invariant par similitude simultanée sur le couple  $(A, B)$ , autrement dit on peut remplacer à volonté  $(A, B)$  par  $(PAP^{-1}, PBP^{-1})$  pour  $P$  dans  $\text{GL}_n(\mathbb{K})$ . Ensuite, le théorème se déduit trivialement de celui d'Albert et Muckenhoupt si  $\mathcal{H}$  ne contient pas  $I_n$ . Dans ce cas, il suffit en effet de partir d'un couple  $(M, N) \in M_n(\mathbb{K})^2$  tel que  $A = [M, N]$ , et de décomposer  $M = \lambda I_n + M'$  et  $N = \mu I_n + N'$ , avec  $M', N'$  dans  $\mathcal{H}$ , pour obtenir  $A = [M', N']$ . On pourra donc se limiter au cas où  $\text{tr}(B) = 0$ .

Venons-en maintenant à la méthode de base pour décomposer la matrice  $A$ . Comme pour le théorème d'Albert et Muckenhoupt, on essaie de montrer que  $A$  est dans l'image de  $\text{ad}_C$  pour une matrice cyclique  $C$  à choisir judicieusement, et pour cela on suppose  $A$  triangulaire supérieure ou de Hessenberg. Le lemme de base est le suivant :

14. R.C. Thompson, Matrices with zero trace. Israel J. Math. 4 (1966) 33–42.

**Lemme 9.21** (Lemme 2.2 de [7]). *Soit  $(A, B) \in \mathfrak{sl}_n(\mathbb{K})^2$  avec  $B \neq 0$ . Posons  $\mathcal{H} = \{B\}^\perp$ . La matrice  $A$  appartient à  $[\mathcal{H}, \mathcal{H}]$  dans chacun des cas suivants :*

- (i)  $|\mathbb{K}| > 2$ ,  $A$  est triangulaire supérieure et  $B$  n'est pas de Hessenberg.
- (ii)  $|\mathbb{K}| > 3$ ,  $A$  est de Hessenberg et il existe un  $i \in \llbracket 2, n-1 \rrbracket$  et un  $j \in \llbracket 3, n \rrbracket \setminus \{i\}$  tels que  $a_{2,1} \neq 0$ ,  $a_{i+1,i} \neq 0$  et  $b_{j,1} \neq 0$ .

Dans chacun des cas, on raisonne par l'absurde en supposant  $A \notin [\mathcal{H}, \mathcal{H}]$ , et on cherche d'abord une matrice  $M \in \mathcal{H}$  telle que  $A \in \text{Im}(\text{ad}_M)$ . On cherche plus précisément une telle matrice sous la forme  $M = yE_{l',l} + \sum_{k=1}^{n-1} x_k E_{k,k+1}$ , où  $(l, l')$  est un couple tel que  $a_{l,l'} \neq 0$  et  $l - l'$  soit maximal pour ces couples (les hypothèses garantissent que  $l - l' > 1$ ). On vérifie alors facilement que  $\text{tr}(M) = 0$  et  $\text{tr}(M^k A) = 0$  pour tout  $k \in \llbracket 2, n-1 \rrbracket$  car  $M$  est triangulaire supérieure stricte et  $A$  de Hessenberg. Si l'on fixe une matrice de cette forme avec  $x_1, \dots, x_{n-1}$  tous non nuls, cette matrice est cyclique et l'appartenance  $A \in \text{Im}(\text{ad}_M)$  se traduit alors par la simple équation linéaire  $\text{tr}(AM) = 0$ , ce qui donne une relation entre les  $n$  paramètres. Lorsque cette relation est satisfaite, on en déduit  $\mathcal{C}(M) \subset \mathcal{H}$  et en particulier  $\text{tr}(M^k B) = 0$  pour tout  $k \in \llbracket 0, n-1 \rrbracket$  : en effet, sinon on pourrait trouver  $N \in M_n(\mathbb{K})$  tel que  $A = [M, N]$ , puis décomposer  $N = N_0 + N_1$  avec  $N_0 \in \mathcal{C}(M)$  et  $N_1 \in \mathcal{H}$ , et ainsi  $A = [M, N_1] \in [\mathcal{H}, \mathcal{H}]$ . En choisissant convenablement  $k$  dans chacune des situations (i) et (ii) pour exploiter l'identité  $\text{tr}(M^k B) = 0$ , on obtient à chaque fois une contradiction en faisant varier les paramètres définissant  $M$  (ce qui nécessite des considérations d'annulation de polynômes à plusieurs variables).

Expliquons maintenant comment exploiter le lemme précédent. On raisonne par l'absurde en supposant  $A \notin [\mathcal{H}, \mathcal{H}]$  (ce qui est une façon commode de rédiger, pour éviter une forme de raisonnement algorithmique aboutissant systématiquement à  $A \in [\mathcal{H}, \mathcal{H}]$  par étude de cas). Si par exemple il existe un vecteur  $x \in \mathbb{K}^n$  tel que  $(x, Ax, A^2x)$  soit libre, le point (ii) du lemme assure, après réduction simultanée de  $A$  et  $B$ , que  $Bx \in \text{Vect}(x, Ax)$ . L'idée est ensuite de montrer que, sauf dans un cas très particulier, la famille  $(x, Ax, Bx)$  est liée pour tout  $x \in \mathbb{K}^n$ . L'étude de cas est assez longue, et nous n'en dirons rien de plus, sinon qu'absolument tout est fondé sur des réductions judicieuses de  $A$  et des applications répétées du lemme 9.21. Pour être tout à fait exact, la conclusion obtenue à l'issue de cette partie de la démonstration est que l'une des deux situations suivantes se produit :

- ou bien le triplet  $(I_n, A, B)$  est localement lié (voir le chapitre 3) ;
- ou bien  $n = 3$  et  $A$  est semblable à  $\lambda I_3 + E_{2,3}$  pour un  $\lambda \in \mathbb{K}$ .

Le deuxième cas se traite de façon assez brutale, en essayant des matrices cycliques variées  $M$  de trace nulle, ajustées de telle sorte que  $A \in \text{Im}(\text{ad}_M)$  ; nous n'en dirons rien de plus.

Plaçons-nous pour finir dans le premier cas, qui est beaucoup plus intéressant puisque sa résolution s'appuie sur la classification des triplets d'opérateurs localement liés, qui était connue pour les corps ayant au moins 5 éléments et que nous avons généralisée pour n'importe quel corps à au moins 3 éléments : nous renvoyons le lecteur au théorème 3.9 de la section 3.5. Sachant que le triplet considéré ici contient une matrice inversible, l'application de la classification laisse seulement deux possibilités :

- Dans le premier cas,  $(I_n, A, B)$  est liée. Si  $A$  est scalaire, une application intelligente du point (i) du Lemme 9.21 permet alors de montrer que  $B$  est elle-aussi scalaire, autrement dit  $\mathcal{H} = \mathfrak{sl}_n(\mathbb{K})$ , ce qui ramène au théorème

de Thompson. Si  $A$  n'est pas scalaire, on décompose  $B = \lambda A + \mu I_n$ , on applique le théorème de Thompson pour décomposer  $A = [C_1, C_2]$  où  $\text{tr } C_1 = \text{tr } C_2 = 0$ , et on remarque alors que  $C_1$  et  $C_2$  sont en fait dans  $\mathcal{H}$  en observant que  $\text{tr}((B - \lambda A)C_i) = 0$  et  $\text{tr}(AC_i) = 0$  pour tout  $i \in \{1, 2\}$ .

Dans chacune de ces situations, on a essentiellement réduit l'étude à celle de l'hyperplan canonique  $\mathfrak{sl}_n(\mathbb{K})$  (autrement dit, au théorème de Thompson).

- Dans le deuxième et dernier cas,  $(I_n, A, B)$  est libre et il existe un plan vectoriel  $\mathcal{P}$  de  $\text{Vect}(I_n, A, B)$  et une droite vectorielle  $\mathcal{D}$  de  $\mathbb{K}^n$  tels que tout élément de  $\mathcal{P}$  ait son image dans  $\mathcal{D}$ . Nécessairement  $\mathcal{P} = \text{Vect}(A - \lambda I_n, B - \mu I_n)$  pour un certain couple  $(\lambda, \mu)$  de scalaires. En particulier  $A - \lambda I_n$  est de rang 1, si bien que la réduction de  $A$  est très simple. Une utilisation intelligente du point (i) du lemme 9.21 permet alors de trouver une contradiction.

## 9.8 Sur la topologie de l'ensemble des racines $p$ -ièmes d'une matrice nilpotente complexe [8]

Soit  $f : U \rightarrow \mathbb{C}$  une fonction analytique non constante sur un ouvert connexe  $U$  de  $\mathbb{C}$ . Étant donné une matrice  $B \in M_n(\mathbb{C})$ , on dispose en théorie d'une méthode, à partir du comportement local de  $f$  en les antécédents des valeurs propres de  $B$ , pour déterminer si l'équation  $f(X) = B$  d'inconnue  $X \in M_n(\mathbb{C})$  possède au moins une solution. Le travail dont je vais rendre compte ici est relatif à la situation canonique de l'équation  $X^p = N$ , où  $N$  est une matrice nilpotente complexe et  $p \in \mathbb{N}^*$ . On dispose dans ce cas de méthodes efficaces pour décider si l'équation a au moins une solution, à partir du profil de Jordan de  $N$ , c'est-à-dire de la suite  $j(N) = (j_k(N))_{k \geq 1}$  dans laquelle  $j_k(N)$  désigne le nombre de blocs de Jordan de taille  $k$  dans la réduite de  $N$ . Nous appellerons plus généralement profil toute suite presque nulle d'éléments de  $\mathbb{N}$  indexée par  $\mathbb{N}^*$ , l'ensemble  $\mathbb{N}^{(\mathbb{N}^*)}$  des profils étant donc naturellement muni d'une structure de monoïde pour l'addition.

Le travail que j'ai effectué, en réponse à une question posée sur le forum Questions-Réponses de la revue RMS, concerne la connexité par arcs de l'ensemble des solutions (lorsqu'il est non vide). Mon théorème est tout simplement le suivant :

**Théorème 9.22** (Théorème 1.2 de [8]). *Soit  $N \in M_n(\mathbb{C})$  une matrice nilpotente complexe, et  $p$  un entier naturel non nul. L'ensemble des racines  $p$ -ièmes de  $N$  est alors connexe par arcs.*

Il est en outre remarquable que ce résultat tombe en défaut sur  $\mathbb{R}$ . Si on considère en effet la matrice nilpotente élémentaire  $E_{1,3}$ , on calcule facilement (passer par le commutant) que ses racines carrées sont les matrices de la forme

$$\begin{bmatrix} 0 & x & y \\ 0 & 0 & x^{-1} \\ 0 & 0 & 0 \end{bmatrix}$$
 où  $(x, y) \in \mathbb{R}^* \times \mathbb{R}$ , et l'ensemble qu'elles forment est évidemment homéomorphe à  $\mathbb{R}^* \times \mathbb{R}$ , qui possède deux composantes connexes par arcs.

Notre démonstration du théorème 9.22 se déroule en trois grandes étapes. L'idée fondamentale est que les racines  $p$ -ièmes de  $N$  se répartissent *a priori* selon diverses classes de similitude, selon leur profil de Jordan. En général, il peut exister des racines  $p$ -ièmes de  $N$  n'ayant pas le même profil de Jordan. On peut noter  $R_{m,p}(N)$  l'ensemble de celles qui ont pour profil  $m$ . Ensuite, une autre remarque importante est la faculté de construire un autre profil de Jordan de racine  $p$ -ième de  $N$  à partir d'un profil connu. C'est fondé sur le lemme élémentaire suivant :

**Lemme 9.23.** *Soit  $a, k, l$  des entiers naturels vérifiant  $0 \leq pa \leq k < l \leq p(a+1)$ . Les matrices  $(J_k \oplus J_l)^p$  et  $(J_{k+1} \oplus J_{l-1})^p$  sont alors semblables.*

Cela conduit à introduire une notion de  $p$ -adjacence entre deux profils de Jordan. En notant  $(e_k)_{k \geq 1}$  la base canonique du  $\mathbb{Z}$ -module  $\mathbb{Z}^{(\mathbb{N}^*)}$ , nous dirons que deux profils  $m$  et  $m'$  sont  $p$ -adjacents lorsqu'il existe trois entiers  $a, k, l$  tels que  $pa \leq k < l \leq p(a+1)$  et  $m - m' = \pm(e_k + e_l - e_{k+1} - e_{l-1})$ . Si  $m$  et  $m'$  sont les profils de Jordan de deux matrices nilpotentes  $A$  et  $A'$ , ce qui précède signifie que l'on peut passer d'une réduite de Jordan de l'une à celle de l'autre par substitution d'un double-bloc  $J_k \oplus J_l$  à un double-bloc  $J_{k+1} \oplus J_{l-1}$  où  $a, k, l$  sont des entiers vérifiant  $0 \leq pa \leq k < l \leq p(a+1)$ .

La démonstration se déroule alors en trois temps :

- (i) On démontre que chaque ensemble  $R_{m,p}(N)$  est connexe par arcs. Cela passe par un argument très élémentaire de géométrie algébrique : pour deux éléments  $X$  et  $X'$  dans  $R_{m,p}(N)$ , il existe  $P \in \mathrm{GL}_n(\mathbb{C})$  telle que  $X' = PXP^{-1}$ . En élevant à la puissance  $p$ -ième on constate que  $P$  est dans l'algèbre commutante  $\mathcal{C}(N)$  de  $N$ . Et plus généralement  $QXQ^{-1}$  est dans  $R_{m,p}(N)$  pour tout  $Q$  inversible dans  $\mathcal{C}(N)$ . Ainsi  $R_{m,p}(N)$  est l'image de  $\mathcal{C}(N)^\times$  par une fonction continue. Et enfin  $\mathcal{C}(N)^\times$  est connexe par arcs car c'est le complémentaire d'un ensemble algébrique d'un  $\mathbb{C}$ -espace vectoriel de dimension finie.
- (ii) On démontre que lorsque  $m$  et  $m'$  sont deux profils  $p$ -adjacents, il existe un chemin de  $R_{m,p}(N)$  à  $R_{m',p}(N)$  restant dans  $R_{m,p}(N) \cup R_{m',p}(N)$  (et qui reste donc aussi dans l'ensemble des racines  $p$ -ièmes de  $N$ ). On utilise pour cela un argument de relèvement : avec les notations adoptées pour la  $p$ -adjacence, on construit un chemin très élémentaire de  $J_k \oplus J_l$  à une matrice semblable à  $J_{k+1} \oplus J_{l-1}$  et dont tous les points ont des puissances  $p$ -ièmes semblables. La clef est alors de relever ce chemin en matrices de passage, grâce au fait que  $P \in \mathrm{GL}_n(\mathbb{C}) \mapsto PMP^{-1} \in \mathcal{S}(M)$  est une fibration pour toute matrice  $M \in \mathrm{M}_n(\mathbb{C})$ , dont a noté  $\mathcal{S}(M)$  la classe de similitude (ce point est redémontré de manière très élémentaire dans l'appendice de [8]).
- (iii) Enfin, on démontre par un raisonnement combinatoire élémentaire que lorsque  $X$  et  $X'$  sont deux racines  $p$ -ièmes de  $N$ , il est possible de passer du profil de  $X$  à celui de  $X'$  par une succession de profils  $p$ -adjacents.

Reste évidemment en suspens l'obtention d'informations plus fines sur la structure homotopique de l'ensemble des racines  $p$ -ièmes de  $N$ , mais l'observation de quelques cas particuliers laisse craindre qu'il soit difficile de dégager des énoncés généraux élégants. Dans le cas réel, on pourrait tenter de majorer finement, en fonction ou non du profil de Jordan de  $N$ , le nombre de composantes connexes par arcs de l'ensemble de ses racines  $p$ -ièmes. Enfin, le cas des matrices quaternioniques est entièrement vierge d'exploration.





**Titre :** Géométrie affine des espaces de matrices, et questions de décompositions quadratiques

**Mots clés :** espaces de matrices, rang, nilpotence, dimension, objets quadratiques, décomposition

**Résumé :** Les travaux exposés ici s'articulent en deux grandes parties dont le point commun est le cadre : l'algèbre linéaire traditionnelle sur un corps arbitraire.

La première partie des travaux porte sur des problèmes de géométrie affine dans les espaces de matrices (ou d'opérateurs linéaires entre espaces vectoriels). Il s'agit essentiellement de prolongement de résultats dus à Dieudonné, Flanders et Gerstenhaber sur la structure linéaire des cônes de matrices définis par des propriétés particulières. Par exemple, le théorème de Dieudonné indique la dimension maximale pour un sous-espace affine de matrices carrées singulières, et donne la structure des espaces de dimension maximale.

Deux grandes classes de généralisations sont ici étudiées, la première portant sur des propriétés de rang, l'autre sur des propriétés de spectre (pour des matrices carrées). De nombreux ponts fructueux sont établis entre les différents problèmes par des techniques de dualité, principalement la dualité opérateur-vecteur.

La seconde partie des travaux étudie la décomposabilité d'un opérateur en somme ou produit d'opérateurs quadratiques dont on impose un polynôme annulateur de degré 2. Nous résolvons de tels problèmes pour des décompositions de petite longueur, en dimension finie ou infinie, et en particulier les décompositions à deux termes.

**Title :** On the affine geometry of matrix spaces, and quadratic decomposition issues

**Keywords :** matrix spaces, rank, nilpotence, dimension, quadratic objects, decomposition

**Abstract :** The works that are featured here are split into two main parts, and all deal with problems in standard linear algebra over an arbitrary field.

The first part deals with problems in the affine geometry of matrix spaces (or spaces of linear operators between vector spaces). It essentially consists of extensions of classical results of Dieudonné, Flanders and Gerstenhaber on the linear structures in cones of matrices defined by special properties. For example, Dieudonné's theorem gives the maximal dimension for an affine subspace of singular square matrices, as well as the structure of the spaces that attain the maximal possible dimension.

Two main sorts of generalizations are studied, that deal either with bounds on the rank or on the cardinality or nature of the spectrum (for square matrices). In particular, connections are found between many of these problems, thanks to the operator-vector duality.

The second part of the present work deals with the problem of decomposing a linear operator into a sum or product of matrices with prescribed annihilating polynomials of degree 2 (i.e. quadratic operators). We solve such problems for a small number of summands/factors, in finite or infinite dimension, with special care given to decompositions with two factors/summands.