UNIVERSITY OF LJUBLJANA
FACULTY OF MATHEMATICS AND PHYSICS
DEPARTMENT OF MATHEMATICS

Mathematics - 3rd cycle

Špela Špenko

Modifying the structure of associative algebras

Doctoral thesis

Supervisor: prof. dr. Matej Brešar

Ljubljana, 2015

UNIVERZA V LJUBLJANI
FAKULTETA ZA MATEMATIKO IN FIZIKO
ODDELEK ZA MATEMATIKO

Matematika - 3. stopnja

Špela Špenko

Preoblikovane strukture asociativnih algeber

Doktorska disertacija

Mentor: prof. dr. Matej Brešar

Ljubljana, 2015

Podpisana Špela Špenko izjavljam:

— da sem doktorsko disertacijo z naslovom *Preoblikovane strukture asociativnih algeber* izdelala samostojno pod mentorstvom prof. dr. Mateja Brešarja in

— da Fakulteti za matematiko in fiziko Univerze v Ljubljani dovoljujem objavo elektronske oblike svojega dela na spletnih straneh.

Ljubljana, 7. junij 2015 Podpis:

labirint

Hvala

Matej za prostor

Urban za kompas in daljnogled

Zala za drevesa

Boris za sonce

Igor za nove poti

Michel za zemljevid in metulje

Abstract

The structure of associative algebras can be modified by changing the operation of multiplication. In studying connections between the initial and the obtained modified category, the theory of functional identities has emerged. In this thesis we first study a subclass of functional identities – quasi-identities, which have played a fundamental role in the theory. They appear as linear relations among the noncommutative polynomial functions on algebras of matrices. We prove that quasi-identities follow from the Cayley-Hamilton identity if one allows central denominators, while the Cayley-Hamilton identity does not exhaust all quasi-identities globally. However, when considered in the class of all functional identities, every functional identity is a consequence of the Cayley-Hamilton identity.

The analysis depends heavily on the theory of generic matrix algebras and trace rings. These are universal objects in the category of algebras (resp. algebras with trace) satisfying all polynomial (resp. trace) identities of $n \times n$ matrices. Thus, they can be seen as analogues of polynomial rings from a noncommutative geometry standpoint. We explore some of their geometric properties. We prove a tracial Nullstellensatz and study the image of noncommutative polynomials and some special noncommutative polynomial maps on matrices. Moreover, we consider homological properties of trace rings and construct (twisted) noncommutative crepant resolutions of singularities for their centers.

We further apply the theory of identities on matrices and matrix invariants to free function theory. This enables a unified approach to an understanding of free maps and free maps with involution.

In Banach algebras we modify the multiplicative structure via the spectral function. We determine elements through their spectral functions and identify derivations through the spectra of their values. We investigate the stability of commuting maps, Lie maps and derivations, and obtain metric versions of Posner's theorems. We conclude by modifying the structure of C^*-algebras and especially algebras of matrices by introducing a multilinear multiplication induced by a noncommutative polynomial.

Math. Subj. Class. (2010): 13A50, 14L24, 15A86, 16E35, 16R10, 16R30, 16R60, 20G15, 32A05, 46H05, 46H15, 46J05, 46L05, 46L52, 46L57, 47A10, 47A56, 47A57, 47B47, 47B48.

Keywords: functional identities, quasi-identities, algebras with trace, Nullstellensatz, noncommutative polynomials, matrix invariants, noncommutative resolutions, length of a vector space, free analysis, Banach algebras, C^*-algebras, spectrum, commuting maps, Lie maps, linear preservers.

Izvleček

Strukturo asociativnih algeber lahko preoblikujemo s spremembo operacije množenja. Iz študija povezav med prvotno in preoblikovano strukturo je vzniknila teorija funkcijskih identitet. V disertaciji najprej proučujemo podrazred funkcijskih identitet - kvazi-identitete. Na algebrah matrik se pojavijo kot linearne relacije na nekomutativnih polinomskih funkcijah. Pokažemo, da kvazi-identitete izhajajo iz Cayley-Hamiltonove identitete, če dopustimo centralne imenovalce, globalno pa ta identiteta ne zaobjame vseh kvazi-identitet. Nasprotno je vsaka funkcijska identiteta znotraj celega razreda funkcijskih identitet posledica Cayley-Hamiltonove identitete.

Obravnava se močno nasloni na teorijo generičnih matričnih algeber in kolobarjev s sledjo. Generična matrična algebra in kolobar s sledjo sta univerzalna objekta v kategoriji algeber (oz. algeber s sledjo), ki zadoščajo vsem polinomskim identitetam (oz. identitetam s sledjo) $n \times n$ matrik. Torej sta s pogledom nekomutativne geometrije podobna polinomskim kolobarjem. Raziskujemo njune geometrijske lastnosti. Poiščemo Nullstellensatz s sledjo in postojimo pred slikami nekomutativnih polinomov in posebnimi nekomutativnimi polinomskimi preslikavami. Poglobimo se še v homološko naravo kolobarjev s sledjo in zgradimo nekomutativne krepantne odprave singularnosti njihovih centrov.

Teorijo identitet na matrikah in matričnih invariant prenesemo v okolje proste funkcijske teorije, kjer omogoči poenoten pristop k razumevanju prostih preslikav in prostih preslikav z involucijo.

V Banachovih algebrah preoblikujemo strukturo preko spektralne funkcije. Elemente prepoznamo po njihovih spektralnih funkcijah, odvajanja pa istovetimo preko spektrov njihovih vrednosti. Proučujemo stabilnost komutirajočih in liejevih preslikav ter odvajanj, in podamo metrične različice Posnerjevih izrekov. Zaključimo s preoblekami C^*-algeber in matričnih algeber z vpeljavo multilinearnega množenja, porojenega z nekomutativnim polinomom.

Math. Subj. Class. (2010): 13A50, 14L24, 15A86, 16E35, 16R10, 16R30, 16R60, 20G15, 32A05, 46H05, 46H15, 46J05, 46L05, 46L52, 46L57, 47A10, 47A56, 47A57, 47B47, 47B48.
Ključne besede: funkcijske identitete, kvazi-identitete, Cayley-Hamiltonov polinom, algebre s sledjo, izrek o ničlah, nekomutativni polinomi, matrične invariante, nekomutativne odprave singularnosti, dolžina vektorskega prostora, prosta analiza, Banachove algebre, C^*-algebre, spekter, komutirajoče preslikave, Liejeve preslikave, linearni ohranjevalci.

Contents

CHAPTER 1

Introduction

Algebras come equipped with linear and multiplicative structure. One may modify them by forgetting a part of the multiplicative structure. In this way one obtains two categories. The aim is to find a simpler modified category which still reflects the original to a large extent. To draw a parallel with "real life", one may be observing a running gazelle, and would like to gain understanding of its motion from its track.

More concretely, we can for example replace the original product by the Lie product induced by the commutator, or more generally by the product induced by an arbitrary multilinear noncommutative polynomial, and study linear maps that preserve this new multiplication. The multiplicative structure can also be replaced by the spectral function in the context of Banach algebras and one is left to explore how powerful the spectrum is as an invariant.

An effective method for handling such problems is the theory of functional identities. However, it has not been fully developed for finite dimensional algebras. In this thesis we try to fill this gap. We focus on identities in finite dimensional algebras, especially matrix algebras, in the first part. In the second we consider some modifications of the structure of Banach algebras.

Identities on matrices and the Cayley-Hamilton polynomial

Let us first give a brief background on the theory of functional identities. A functional identity is an identical relation in a ring that, besides arbitrary elements that appear in a similar fashion as in a polynomial identity, also involves arbitrary functions which are considered as unknowns. The goal of the general functional identity theory is to describe these functions. Starting with the solution of a long-standing Herstein's problem on Lie isomorphisms in 1993 [**Bre93b**], functional identities have since turned out to be applicable to various problems in noncommutative algebra, nonassociative algebra, operator theory, and functional analysis. We refer to the book [**BCM07**] for an account of functional identities and their applications.

Given a functional identity, one usually first finds its "obvious" solutions; i.e., those functions that satisfy this identity for formal reasons, independent of the structure of the ring in question. These are called the *standard solutions*. A typical result states that either the standard solutions are in fact the only possible solutions or the ring has some special properties, like satisfying a polynomial identity of a certain degree related to the number of variables. The existing theory of functional identities, as surveyed in [**BCM07**], thus gives definite results for a large class of noncommutative rings, but, paradoxically, tells nothing about the basic example of a noncommutative ring, i.e., the matrix algebra $M_n(F)$ (unless n is big enough). This is reflected in applications – one usually has to exclude $M_n(F)$ (for "small" n) in a variety of results whose proofs depend on the general theory of functional identities, although by the nature of these results one can conjecture that this exclusion is unnecessary (see [**BB09**,**BBS11**,**BBCM00**,**BC00b**,**BF99**] for typical examples). The problem with $M_n(F)$ is that it allows *nonstandard solutions*. Their description seemed to be a much harder problem than the description of standard solutions. However, using the coordinate-wise approach and the theory of algebras with trace we are able to make some progress in this direction.

We start by investigating a subclass of functional identities. Given a finite dimensional algebra A and an integer m (or ∞), let $\mathcal{C}$ be the commutative ring of polynomial functions on m copies of A and $\mathcal{C}\langle X\rangle$ the free algebra in m variables $X = \{x_1, \dots, x_m\}$. We call this algebra the algebra of *quasi-polynomials* of A. One can clearly evaluate elements in $\mathcal{C}\langle X\rangle$ on A and the *quasi-identities* of A are those quasi-polynomials that vanish at all evaluations.

We consider the fundamental case where $A = M_n(F)$, the algebra of matrices. Quasi-identities appear in a natural way as linear relations among the noncommutative polynomial functions on A. In this sense the theory of quasi-identities is a worthwhile generalization of the theory of polynomial identities of A.

Quasi-polynomials, also called Beidar polynomials in some papers, were introduced in 2000 by Beidar and Chebotar [**BC00a**], and have since played a fundamental role in the theory of functional identities and its applications. Standard solutions of quasi-identities can be very easily described: all coefficient functions must be 0 (cf. [**BCM07**, Lemma 4.4]). The Cayley-Hamilton theorem gives rise to a basic example of a quasi-identity on the matrix algebra $M_n(F)$ with nonstandard solutions. We call it the Cayley-Hamilton identity. The following question presents itself:

QUESTION. Is every quasi-identity of $M_n(F)$ a consequence of the Cayley-Hamilton identity?

An important motivation for this question is the well-known theorem, proved independently by Procesi [**Pro76**] and Razmyslov [**Raz74**], saying that the answer to such a question is positive for the related *trace identities*.

We show that also the quasi-identities are quite closely related to the Cayley-Hamilton identity.

THEOREM. *(See Theorem 2.2.7.) Let P be a quasi-identity of $M_n(F)$. For every central polynomial c of $M_n(F)$ with zero constant term there exists $m \in \mathbb{N}$ such that $c^m P$ is a consequence of the Cayley-Hamilton identity.*

The main object of study, the space $\mathfrak{I}_n/(Q_n)$ of quasi-identities modulo the subspace of those quasi-identities which follow from the Cayley-Hamilton identity, is an invariant of the quotient map of the action of the projective linear group acting, by simultaneous conjugation, on the space of m-tuples of matrices. This space appears as a module on the quotient variety and the previous theorem implies that is supported on the singular set. Still the complexity of the quotient map makes it difficult to describe this module and even to decide in a simple way if it is nonzero. We show that it is indeed nonzero by describing a particular subspace of this module. We find antisymmetric quasi-identities of $M_n(F)$ of degree n^2 that are not a consequence of the Cayley-Hamilton identity among those quasi-identities which transform under the general linear group as the adjoint representation.

THEOREM. *(See Theorem 2.3.7). There is a direct sum decomposition*

$$\mathbb{G}_n[n^2] = \mathbb{G}_n[n^2]_{CH} \oplus \bigwedge^{n^2-2} N_n^* X^2.$$

Here $G_n[n^2]$ denotes the isotypic component of antisymmetric quasi-identities of degree n^2 corresponding to the adjoint representation of GL_n on the space of trace zero matrices N_n, and $G_n[n^2]_{CH}$ is its submodule of those that follow from the Cayley-Hamilton identity. In order to achieve this result, we first show how the adjoint representation of the simple Lie algebra $\mathfrak{g}$ of traceless matrices sits in the exterior algebra $\bigwedge \mathfrak{g}$ of the same Lie algebra $\mathfrak{g}$. The discovery of this phenomenon has been the starting point for a general theorem for all simple Lie algebras, as shown in [**DCPP13**], and it also gives an insightful explanation of the basic theorem on identities of matrices, namely the Amitsur-Levitzki identity [**Pro13**].

While the structure of the module $\mathfrak{I}_n/(Q_n)$ remains quite unexplored, we prove that the module is finitely generated, which can be seen as a positive solution of the Specht problem for quasi-identities.

We next concentrate on two-sided functional identities; i.e., identities of the form

$$\sum_{k\in K} F_k(\overline{x}_m^k)x_k = \sum_{l\in L} x_l G_l(\overline{x}_m^l) \quad \text{for all } x_1\dots, x_m \in M_n(F),$$

where $\overline{x}_m^k = (x_1, \dots, \hat{x}_k, \dots, x_m)$, K and L are subsets of $\{1, \dots, m\}$, and F_k, G_l are arbitrary functions from $M_n(F)^{m-1}$ to $M_n(F)$.

The Cayley-Hamilton identity yields the fundamental example of a functional identity of $M_n(F)$ which does not have only standard solutions. In contrast to the case of quasi-identities all the nonstandard parts of solutions of two-sided functional identities follow from the Cayley-Hamilton identity.

Theorem. *(See Theorem 3.2.4, Theorem 3.3.7.) Every solution of a two-sided functional identity is standard modulo the Cayley-Hamilton identity.*

The idea is to interpret these functional identities coordinate-wise and then apply the theory of syzygies.

We separately consider functional identities of one variable, in particular the commuting maps. The latter we also consider in the framework of Banach algebras, therefore we give here a brief introduction of this notion in the setting of general associative rings.

Let R be a ring. A map $q : R \to R$ is said to be *commuting* if $[q(x), x] = 0$ for all $x \in R$. The study of commuting maps has a long and rich history, starting with Posner's 1957 theorem stating that there are no nonzero commuting derivations on noncommutative prime rings [**Pos57**]. We refer to the survey article [**Bre00**] for the general theory of commuting maps. We will be concerned with commuting traces of multilinear maps on prime algebras. A map q between additive groups (resp. vector spaces) A and B is called the trace of a d-additive (resp. d-linear) map if there exists a d-additive (resp. d-linear) map $M : A^d \to B$ such that $q(x) = M(x, \dots, x)$ for all $x \in A$ (for $d = 0$ this should be understood as that q is a constant). If R is a prime ring (resp. algebra) and $q : R \to R$ is a commuting trace of a d-additive map, is then q necessarily of a *standard form*, meaning that there exist traces of $(d-i)$-additive (resp. $(d-i)$-linear) maps μ_i from R into the extended centroid C of R such that $q(x) = \sum_{i=0}^{d} \mu_i(x)x^i$ for all $x \in R$? This question was initiated in 1993 by Brešar who obtained an affirmative answer for $d = 1$ [**Bre93a**], and also for $d = 2$ [**Bre93b**] provided that $\operatorname{char}(R) \neq 2$ and R does not satisfy S_4, the standard polynomial identity of degree 4. The $d = 3$ case was treated, in a slightly different context, by Beidar, Martindale, and Mikhalev [**BMM96**]. These three results have turned out to be applicable to various problems, particularly in the Lie algebra theory, and have played a crucial role in the development of the theory of functional identities [**BCM07**]. The next step was made in 1997 by Lee, Lin, Wang, and Wong [**LWLW97**], who answered the above question in affirmative for a general d, but under the assumption that $\operatorname{char}(R) = 0$ or $\operatorname{char}(R) > d$ and R does not satisfy the standard polynomial identity S_{2d}. The reason for the exclusion of rings satisfying polynomial identities (of low degrees) is the method of the proof; it is, therefore, natural to ask whether, by a necessarily different method, one can get rid of this assumption. For $d = 2$ this has turned out to be the case [**BŠ03**]. To the best of our knowledge, the question is still open for a general d (cf. [**Bre00**, p. 377] and [**BCM07**, p. 130]). We give an affirmative answer for centrally closed prime algebras and traces of d-linear maps. In this setting the problem can be reduced to the case where the algebra in question is a matrix algebra. This can be deduced from the description of non-standard parts of solutions of two-sided functional identities exposed in the previous theorem. However, we present an alternative approach interpreting the problem in the algebra of generic matrices and use some standard facts from commutative algebra.

THEOREM. *(See Theorem 4.2.3.) If A is a centrally closed prime algebra and $q : A \to A$ is a commuting trace of a multilinear map, then q is of a standard form.*

We also deal with considerably more general functional identities of one variable. That is, we consider functional identities for which $\sum_{i=0}^{m} x^i q_i(x) x^{m-i}$ is always 0 or always central, and where q_i are assumed to be the traces of multilinear maps. We handle the finite dimensional case, where the only non-standard solutions of functional identities again arise from the Cayley–Hamilton theorem.

Trace rings

Intimately connected with identities on matrices are the generic matrix algebra generated by m generic $n \times n$-matrices, and the trace ring obtained by adjoining to the generic matrix algebra the traces of products of generic matrices. They are universal objects in the category of algebras (resp. algebras with trace) satisfying polynomial (resp. trace) identities of $n \times n$ matrices. Therefore they play a role of a polynomial ring in noncommutative geometry. In Chapter 3 we investigate some of their geometric aspects.

Hilbert's Nullstellensatz is a classical result in algebraic geometry. Over an algebraically closed field it characterizes polynomials vanishing on the zero set of a set of polynomials. Due to its importance it has been generalized and extended in many different directions, including to free algebras. For instance, Amitsur's Nullstellensatz [**Ami57**] describes noncommutative polynomials vanishing on the zero set of a given finite set of noncommutative polynomials in a full matrix algebra. In another direction, the Nullstellensatz of Bergman [**HM04**] studies a weaker, directional notion of vanishing but in a dimension-independent context (see [**CHMN13**] for recent generalizations) allowing for a stronger conclusion. Namely, unlike in Hilbert's and Amitsur's Nullstellensatz, no powers are needed in the obtained algebraic certificate. We also refer to [**BK11**] for a survey of free Nullstellensätze.

We focus on *vanishing trace* of noncommutative polynomials. The relationship between sums of commutators and vanishing trace of a noncommutative polynomial is discussed e.g. in [**CGMS09**, **KS08**, **BK11**]. Our main result characterizes noncommutative polynomials polynomials f whose trace vanishes whenever the traces of polynomials $f_1, \dots, f_r$ vanish.

THEOREM. *(See Theorem 1.2.1.) Let $f_1, \dots, f_r, f \in F\langle X\rangle$. The implication*

$$\operatorname{tr}(f_1(A)) = \cdots = \operatorname{tr}(f_r(A)) = 0 \quad \Longrightarrow \quad \operatorname{tr}(f(A)) = 0$$

holds for every n and all $A \in M_n(F)^g$ if and only if f is cyclically equivalent to a linear combination of f_i's or a linear combination of f_i's is cyclically equivalent to a nonzero scalar.

The main ingredients in the proof are effective degree bounds on Hilbert's Nullstellensatz due to Kollár [**Kol88**] (see also Sombra [**Som99**] and Jelonek [**Jel05**]), as well as the theory of polynomial identities [**Row80**, **Pro76**]. Finally, we solve a tracial moment problem by dualizing the statement of the theorem.

While zeros of polynomials are the basic object in (classical) algebraic geometry, the image of a single polynomial is not very interesting assuming an algebraically closed base field. However, the image of elements in the trace rings or generic matrix algebras evaluated on $M_n(F)$ becomes an intriguing object which is not yet fully understood. One of the first to observe the importance of this object was Kaplansky [**KBMR12**]. An obvious necessary condition for a subset S of $M_n(F)$ to be equal to $\mathrm{im}(f)$ for some f is that S is closed under conjugation by invertible matrices, i.e., $tSt^{-1} \subseteq S$ for every invertible $t \in M_n(F)$. Chuang [**Chu90**] proved that if F is a finite field then this condition is also sufficient. This is not true for infinite fields. For example, the set of all square zero matrices cannot be the image of a polynomial [**Chu90**, Example, p. 294].

We consider the case where F is an algebraically closed field of characteristic 0. If f is a polynomial identity, then $\mathrm{im}(f) = \{0\}$. Another important situation where $\mathrm{im}(f)$ is "small" is when f is a central polynomial; then $\mathrm{im}(f)$ consists of scalar matrices. What are other possible small images? When considering this question, one has to take into account that if $a \in \mathrm{im}(f)$, then the similarity orbit of a is also contained in $\mathrm{im}(f)$. The images of many polynomials (for example the homogeneous ones) are also closed under scalar multiplication. Accordingly, let us denote $a^\sim = \{\lambda tat^{-1} \mid t \in GL_n, \lambda \in F\}$. Is it possible that $\mathrm{im}(f) \subseteq a^\sim$ for some nonscalar matrix a? In the $n = 2$ case the answer comes easily: $\mathrm{im}(x_1x_2 - x_2x_1)^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^\sim$. One can check this by an easy computation, but the concept behind this example is that the polynomial $(x_1x_2-x_2x_1)^2$ is central, so that $\mathrm{im}(x_1x_2 - x_2x_1)^3$ can consist only of those trace zero matrices whose determinant is nonzero. Let us also mention that $x_1x_2 - x_2x_1$ also has a relatively small image, namely $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^\sim \bigcup \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^\sim$. Return now to an arbitrary n, and let us make the following definition: A polynomial f is *finite on* $M_n(F)$ if there exist $a_1, \dots, a_k \in M_n(F)$ such that $\{0\} \neq \mathrm{im}(f) \subseteq a_1^\sim \bigcup \cdots \bigcup a_k^\sim$. Next, let j be a positive integer that divides n, choose a primitive j-th root of unity μ_j, denote by $\mathbf{1}_r$, $r = \frac{n}{j}$, the identity matrix in $M_r(F)$, and finally, denote by $\mathbf{w}_j$ the diagonal matrix in $M_n(F)$ having $\mathbf{1}_r, \mu_j\mathbf{1}_r, \dots, \mu_j^{j-1}\mathbf{1}_r$ on the diagonal.

THEOREM. *(See Theorem 2.2.3, Corollary 2.2.4.) A polynomial f is finite on $M_n(F)$ if and only if there exists a positive integer j dividing n such that f^j is central on $M_n(F)$ and f^i is not central for $1 \leq i < j$. In this case* $\mathrm{im}(f) \subseteq \mathbf{w}_j^\sim \bigcup n_2^\sim \bigcup \cdots \bigcup n_k^\sim$*, where n_i are nilpotent matrices. Moreover,* $\mathrm{im}(f^{j+1}) \subseteq \mathbf{w}_j^\sim$.

Recall that a subset U of F^{n^2} ($\cong M_n(F)$) is said to be a standard open set (with respect to the Zariski topology) if there exists $p \in F[z_1, \dots, z_{n^2}]$ such that $U = \{(u_1, \dots, u_{n^2}) \in F^{n^2} \mid p(u_1, \dots, u_{n^2}) \neq 0\}$. A simple concrete example of such a set U is GL_n. It can be easily seen that $U \cup \{0\}$ is the image of a noncommutative polynomial. The same is true for arbitrary standard open sets.

THEOREM. *(See Theorem 2.3.7.) If U is a standard open set in F^{n^2} that is closed under nonzero scalar multiplication and conjugation by invertible matrices, then $U \cup \{0\} = \mathrm{im}(f)$ for some polynomial f.*

We further consider the density of $\mathrm{im}(f)$ (with respect to the Zariski topology of $F^{n^2} \cong M_n(F)$). It is easy to see that the density of $\mathrm{im}(f)$ in $M_n(F)$ (resp. $M_n(F)^0$) is equivalent to $\mathrm{tr}(f), \mathrm{tr}(f^2), \dots, \mathrm{tr}(f^n)$ (resp. $\mathrm{tr}(f^2), \dots, \mathrm{tr}(f^n)$) being algebraically independent. Our original motivation for studying the density was the question by Lvov asking whether the image of a multilinear polynomial is a vector space. This was shown to be true for $n = 2$ and under some restrictions for $n = 3$ by Kanel-Belov, Malev and Rowen [**KBMR12, KBMR13**]. In general this problem is, to the best of our knowledge, open. If the answer was positive, then either $\mathrm{im}(f) = M_n(F)$ or $\mathrm{im}(f) = M_n(F)^0$ would hold for every multilinear polynomial f that is neither an identity nor central (see [**BK09**] or [**KBMR12**]). Establishing the density could be an important intermediate step for proving these equalities. On the other hand, as it will be apparent from the next paragraph, this would be sufficient for some applications.

Motivated by Lvov's problem we have posed ourselves the following two questions concerning a multilinear polynomial f. Is f central if there exists $k \geq 2$ such that f^k is central for $M_n(F)$, $n \neq 2$? Is f an identity if there exists $k \geq 2$ such that $\mathrm{tr}(f^k)$ vanishes on $M_n(F)$, $n \neq 2$? It has turned out that versions of the first one had already been discussed before (see [**Ler75, Row74**]). (Incidentally, the condition that $\mathrm{tr}(f^k)$ vanishes on $M_n(F)$ is equivalent to the condition that f^k is the sum of commutators and an identity [**Pro76, BK09**].) Note that an affirmative answer to Lvov's question implies that both questions have affirmative answers. Moreover, to establish the latter it would be enough to know only that $\mathrm{im}(f) \cap M_n(F)^0$ is dense in $M_n(F)^0$. Further, since $\mathbf{w}_j$ has trace zero, one can easily deduce from the last assertion of the above theorem that an affirmative answer to the second question implies an affirmative answer to the first. Unfortunately, we are unable to solve any of these two questions. We only show the dimension-free version; i.e., if f is a nonzero multilinear polynomial, then f^k, $k \geq 2$, is not a sum of commutators. However, we give a small evidence that the answer to Lvov's question may be affirmative: if f is a nonzero multilinear Lie polynomial of degree at most 4, then $\mathrm{im}(f) = M_n(F)^0$. More general result stating that any multilinear polynomial of degree at most 4 on $M_n(F)$ for $n \geq 3$ contains M_n^0 can be found in [**BW13**].

We further consider the image of special polynomial maps given by an n^2-tuple of words of degree d on $M_n(F)$. We prove that for the (almost) least possible d there exists an n^2-tuple in the image consisting of linearly independent matrices.

Theorem. *(See Theorem 3.2.1.) Let $g \geq 2$ and $d = \lceil \log_g n \rceil$. There exist $w_1, \dots, w_{n^2} \in \langle x_1, \dots, x_g \rangle_{2d}$ that are $M_n(F)$-locally linearly independent.*

The motivation for considering such maps stems also from the Paz conjecture. Let V be a vector subspace of $M_n(F)$. By V^k we denote the vector space spanned by the words of length at most k evaluated at V. The *length* $\ell(V)$ of V is an integer ℓ yielding a chain

$$V \subsetneq V^2 \subsetneq \cdots \subsetneq V^\ell = V^{\ell+1}.$$

Paz proved in [**Paz84**] that $\ell(V) \leq \lceil (n^2+2)/3 \rceil$, later Pappacena considerably improved this bound by giving an upper bound of order $O(n^{3/2})$ [**Pap97**]. Paz conjectured that the optimal bound would be $2n-2$. It is not difficult to find examples where this upper bound is actually attained. However, if V is a generic vector subspace of $M_n(F)$ one would expect that its length would be of order $O(\log n)$, which we call a generic Paz conjecture. An easy corollary of the above theorem establishes this natural guess.

Originally trace rings appeared in the classical invariant theory. Let V be a vector space of dimension n and let $G = \mathrm{GL}_n(\mathbb{C})$ act by simultaneous conjugation on $W = \mathrm{End}(V)^{\oplus m}$. Procesi proved in [**Pro76**] that the ring of invariants $(SW)^G$ is isomorphic to the center $\mathcal{T}_{m,n}$ of the trace ring $\mathcal{T}_{m,n}\langle \xi_k \rangle$ and that $\mathcal{T}_{m,n}\langle \xi_k \rangle = (SW \otimes \mathrm{End}(V))^G$ is the algebra of covariants. While $\mathcal{T}_{m,n}$ is singular for all (m,n) except for $(m,n) = (2,2)$ [**LBP87**, Proposition II.3.1], the homological properties of $\mathcal{T}_{m,n}\langle \xi_k \rangle$ are a bit better. It has finite global dimension if and only if $(m,n) = (2,2), (3,2), (2,3)$ (see [**LBVdB88**] for the if direction and [**LBP87**] for the only if direction).

A fairly recent development in the singularity theory is the appearance of "noncommutative resolutions". They were introduced by Van den Bergh in [**VdB04**].

Definition. Assume that S is a normal Noetherian Gorenstein domain. Then a *noncommutative crepant resolution* (NCCR) of S is an S-algebra of finite global dimension of the form $\Lambda = \mathrm{End}_S(M)$ where M is a non-zero finitely generated reflexive S-module and Λ is a Cohen-Macaulay S-module.

For the rationale behind this definition we refer to [**VdB04**]. In general the behaviour of NCCRs closely mimics that of (commutative) crepant resolutions as defined in algebraic geometry (see e.g. [**SVdB08**]). In fact the concepts are closely related, in dimension three the commutative and noncommutative notions are even equivalent [**VdB04**] and this is part of the motivation for the algebraic approach to the three dimensional minimal model program by Iyama and Wemyss [**IW14a**, **IW14b**, **Wem14**].

The notion of a NCCR is well understood in the case of quotient singularities for finite groups. Let G be a finite group, by $\hat{G}$ we denote the set of isomorphism classes of irreducible G-representations. For a finite dimensional G-representation U let $M(U) \stackrel{\text{def}}{=} (U \otimes F[X])^G$ be the corresponding $F[X]^G$*-module of covariants.* We put $U = \oplus_{V\in\hat{G}} V$. If W is a representation of G and $G \subset \mathrm{SL}(W)$, then $\Lambda = \mathrm{End}_{SW^G}(M(U))$ is an NCCR of SW^G.

For quotient singularities of infinite groups the noncommutative crepant resolutions have been constructed for determinantal varieties [**BLVdB10**,**BLVdB11**] and Pfaffian varieties of $n \times n$ skew-symmetric matrices of rank < 4 for odd n [**Kuz08**].

We show that quotient singularities for reductive groups always have noncommutative resolutions in an appropriate sense. As in the finite group case we will construct noncommutative resolutions using the properties of the category $\mathrm{mod}(G, SW)$. However if G is not finite the analysis is more complicated because of two non-trivial issues:

(1) the category $\mathrm{mod}(G, SW)$ does not have a projective generator as G has infinitely many irreducible representations;
(2) modules of covariants are usually not Cohen-Macaulay.

To handle the first issue we construct certain nice complexes which relate different projectives in $\mathrm{mod}(G, SW)$. The second issue is handled using the results in [**VdB91**, **VdB93**, **VdB99**].

We also exhibit a large class of such singularities which have (twisted) noncommutative crepant resolutions. With the developed methods it is possible to treat a number of examples, both new and old. In particular, NC(C)Rs exist in previously unknown cases for determinantal varieties of symmetric and skew symmetric matrices. Here we focus on centers of trace rings $\mathcal{T}_{m,n}\langle\xi_k\rangle$. A "twisted" NCCR is just like a NCCR except that is generically a central simple algebra rather than a matrix ring.

Theorem. *(See Theorem 4.7.1.) Assume $m \geq 2$, $n \geq 2$. Then $\mathcal{T}_{m,n}$ has a twisted NCCR.*

Free function theory

It seems reasonable to conclude the part on identities in matrices and corresponding universal objects with another instance of their applicability in the field of free function theory, which also serves as a bridge between considering our initial theme in the algebraic and analytic context.

Free maps are maps on g-tuples of matrices of arbitrary size that preserve simultaneous similarity and direct sums. The notion of a free map arises naturally in free probability, the study of noncommutative rational functions [**AD03**, **BGM06**, **HMV06**], and systems theory [**HBJP87**, **KVV12**]. The study of these maps is in the realm of free analysis [**AM15**, **AM14**, **AY14**, **AKV13**, **BV03**, **KVV14**, **HKM12**, **MS11**, **Pas14**, **PTD13**, **Pop06**, **Pop10**, **Tay73**, **Voi04**, **Voi10**].

In Chapter 4 we introduce powerful invariant-theoretic methods [**Pro76**] to free analysis. We present an alternative, algebraic approach to free function theory. While most of the current efforts in free analysis are focused on (involution-free) free maps – free analogs of analytic functions in several complex variables – where strong rigidity is observed, our main attention is to *free maps with involution*, e.g. noncommutative polynomials, rational function or power series in x, x^t. Our methods are uniform in that they work in both cases with only minimal adaptations needed. Thus we recover some of the existing results on (involution-free) free maps (cf. [**AM14**, **KVV14**, **Pas14**]).

We prove that a free map with involution f is a polynomial in x, x^t if and only if there is $d \in \mathbb{N}$ such that each of the level functions $f[n]$ is a polynomial of degree $\leq d$. This enables us to show

THEOREM. *(See Theorem 2.2.1.) Analytic free maps with involution admit convergent power series expansions about scalar points.*

We further show that analytic free maps with involution admit convergent power series expansions about non-scalar points, whose homogeneous parts are generalized polynomials. With this information we establish free inverse and implicit function theorems for differentiable free maps with involution.

We conclude the chapter by presenting several illustrating examples demonstrating non-rigidity properties of free maps with involution. For instance, we give an example of a bounded smooth free map with involution that is not analytic.

Spectrum as an invariant in Banach algebras

Having concentrated on the methods useful in studying our initial modification problem and their applications also in other directions, we focus in Chapter 5 on the problem itself in the context of Banach algebras. We are particularly interested in the spectral function and we try to extract from it some information about the elements in the algebra and derivations on it.

The spectrum of an element a of a Banach algebra A will be denoted by $\sigma(a)$. By $r(a)$ we denote the spectral radius of a. We write $Z(A)$ for the center of A. A Banach algebra A is *semisimple* if and only if the only element $a \in A$ with the property $\sigma(ax) = \{0\}$ for all $x \in A$ is the zero element. In other words, in semisimple Banach algebras only $a = 0$ has zero spectral function $x \mapsto \sigma(ax)$. We are interested whether also other elements can be determined by their spectral function. We thus address the following question.

QUESTION. Let A be a semisimple Banach algebra. Suppose that $a, b \in A$ satisfy

$$\sigma(ax) = \sigma(bx) \quad \text{for all } x \in A. \tag{0.1}$$

Does this imply $a = b$?

We do not know the answer in general. In various special cases, however, we are able to show that it is affirmative. Firstly, we establish this under the assumption that a can be written as the product of an idempotent and an invertible element. The proof is based on a spectral characterization of central idempotents. Secondly, we handle the case where A is a commutative Banach algebra, and thirdly, we prove that this is the case if A is a C^*-algebra.

THEOREM. *(See Theorem 1.1.6.) If A is a C^*-algebra and $a, b \in A$ satisfy $\sigma(ax) = \sigma(bx)$ for all $x \in A$, then $a = b$.*

We also treat a considerably more general condition that concerns the spectral radius. Let A be a semisimple Banach algebra. Suppose that $a, b \in A$ satisfy $r(ax) \leq r(bx)$ for all $x \in A$. What is the relation between a and b? The answer may depend on the algebra or on the elements in question. A special situation where $b = 1$ has been examined earlier by Pták [**Ptá78**] (and, independently, also in the recent paper [**BBR09**]). The conclusion in this case is that $a \in Z(A)$. We show that if A is a prime C^*-algebra, then a and b are necessarily linearly dependent.

Let B and A be Banach algebras, and let $\varphi : B \to A$ be a surjective linear map such that

$$\sigma\big(\varphi(x)\big) = \sigma(x) \quad \text{for all } x \in B.$$

Under what conditions is φ a Jordan homomorphism? This is a classical problem in the Banach algebra theory, initiated by Kaplansky in [**Kap70**], which was also a part of the motivation for considering (0.1). It is expected that a sufficient condition is that A is a C^*-algebra, or maybe even a general semisimple Banach algebra. In spite of considerable efforts of numerous authors, the problem seems to be out of reach at such level of generality; see, e.g., [**BŠ08**] for historic comments. One is therefore inclined to consider a stronger version of it that may perhaps shed some light on the classical situation.

In [**Mol02**] Molnar described not necessarily linear surjective maps φ satisfying $\sigma\big(\varphi(x)\varphi(y)\big) = \sigma(xy)$ for all $x, y \in B$ in the case $B = A = B(H)$ or $B = A = C(K)$. These results have been extended in different directions (see [**HLW08**, **LT07**, **TL09**] and references therein), but these generalizations also deal only with some special algebras. It seems that it is not easy to treat this condition in general classes of algebras. We consider similar, but more easily approachable conditions

$$\rho\big(\varphi(x)\varphi(y)\varphi(z)\big) = \rho(xyz) \quad \text{for all } x, y, z \in B,$$

where $\rho \in \{\sigma, r\}$, using previously mentioned results.

We investigate further what can be said about other pairs of linear maps on A having the same spectral function; i.e., such that the spectra of their values coincide at each point. We consider a pair of derivations d and g. In most of our results we will not need to assume the equality of the spectra, but only

$$\sigma(g(x)) \subseteq \sigma(d(x)) \quad \text{for all } x \in A. \tag{0.2}$$

By studying (0.2) we follow the line of investigation of spectral properties of values of derivations. Let us mention some topics in this area: derivations and their products that have quasinilpotent values [**CKL06**, **Lee05**, **TS87**], spectrally bounded derivations [**BM95a**], and derivations all of whose values have a finite spectrum [**BM04**, **BŠ10**, **BŠ96**]. A topic of a different kind, which, however, is closer to the problem considered here than it may seem at a first glance, is the study of derivations d and g such that the range of $g(x)$ is contained in the range of $d(x)$ for every $x \in A$. In their seminal work [**JW75**], Johnson and Williams considered such a range inclusion for the case where $A = B(H)$ and d is an inner derivation implemented by a normal operator. See also [**Fon84**, **KS01**] for further development.

A trivial possibility that 0.2 occurs is that $g = d$. If the range of d consists of non-invertible elements, e.g., if d is an inner derivation implemented by an element from a proper ideal, then we can take $g = 0$. There is another, less obvious possibility: $g = -d$, where d is inner and implemented by an algebraic element of degree 2. We show that the aforementioned three possibilities are also the only ones if A is a primitive Banach algebra with nonzero socle. Using this result we are able to handle (0.2) for a general semisimple Banach algebra A under the assumption that $d(x)$ has a finite spectrum for every $x \in A$. In the case of von Neumann algebras all derivations are inner and we have the following decomposition result.

THEOREM. *(See Theorem 2.3.4.) Let A be a von Neumann algebra and let $a, b \in A$. If $\sigma([b,x]) \subseteq \sigma([a,x]) \cup \{0\}$ for every $x \in A$, then $b = p_1 a - p_2 a + z$ for some orthogonal central projections p_1, p_2 and some $z \in Z(A)$.*

We specially consider the case where A is a C^*-algebra and derivations are inner, $d : x \mapsto [a,x]$ and $g : x \mapsto [b,x]$. We show that b lies in $\{a\}''$, the (relative) bicommutant of $\{a\}$, provided that a is normal. From (0.2) it clearly follows the condition $r([b,x]) \le Mr([a,x])$ with $M = 1$. If both a, b are selfadjoint, then the commutators $[a,x]$ and $[b,x]$ are anti-selfadjoint whenever x is selfadjoint. In this case it can be rephrased as

$$||[b,x]|| \le M||[a,x]|| \quad \text{for all selfadjoint } x \in A. \tag{0.3}$$

According to [**JW75**, Lemma 1.1], we can view (0.3) as a dual problem to the range inclusion problem. Following [**JW75**] and consecutive papers [**Fon84**,**KS01**] we consider the condition (0.3) for a normal element a in a C^*-algebra A. The most complete result, however, is obtained for selfadjoint elements a, b under the assumption that the equality (with $M = 1$) holds in (0.3), or equivalently $\rho([b,x]) = \rho([a,x])$ for all selfadjoint $x \in A$.

We further focus on derivations whose spectral function is zero on elements with trivial spectrum. By $Q = Q_A$ we denote the set of all quasinilpotent elements in A, i.e., $Q = \{q \in A \,|\, \sigma(q) = \{0\}\}$, and by $\mathrm{rad}(A)$ we denote the (Jacobson) radical of A. Recall that $\mathrm{rad}(A) = \{q \in A \,|\, qA \subseteq Q\}$.

It is well-known that $d(A) \subseteq \mathrm{rad}(A)$ if A is commutative; under the assumption that d is continuous this was proved by Singer and Wermer [**SW55**], and without this assumption considerably later by Thomas [**Tho88**]. This result has been extended to noncommutative algebras in various directions. For instance, Le Page [**LP67**] proved that $d(A) \subseteq Q$ implies $d(A) \subseteq \mathrm{rad}(A)$ in case d is an inner derivation. For a general derivation d this was established somewhat later by Turovskii and Shulman [**TS87**] (and independently in [**MM91**]). In [**BM95b**] it was proved that $d(A) \subseteq \mathrm{rad}(A)$ in case there exists $M > 0$ such that $r(d(x)) \le Mr(x)$ for all $x \in A$. Katavolos and Stamatopoulos [**KS08**] showed that if d is an inner derivation implemented by a quasinilpotent element, then $d(Q) \subseteq Q$ implies $d(A) \subseteq \mathrm{rad}(A)$.

The question whether $d(Q) \subseteq Q$ implies $d(A) \subseteq \mathrm{rad}(A)$ for an arbitrary derivation d seems natural since the condition $d(Q) \subseteq Q$ with d arbitrary covers all conditions from the preceding paragraph. However, in general the answer is negative since Q can be $\{0\}$ even when A is noncommutative [**DT75**], and in such a case every nonzero inner derivation of A gives rise to a counterexample. One is therefore forced to confine to special classes of Banach algebras.

THEOREM. *(See Theorem 3.3.1.) Let A be a Banach algebra with the property β, and let Q be the set of its quasinilpotent elements. If a derivation d of A satisfies $d(Q) \subseteq Q$, then $d(A) \subseteq \mathrm{rad}(A)$.*

The class of algebras satisfying the rather technical property β is quite large including C^*-algebras, group algebras on arbitrary locally compact groups, and Banach algebras generated by idempotents.

Analytic Lie maps

In Chapter 6 we continue by studying analytic analogues of derivations, commuting maps and Lie maps. A direct calculation shows that if T is a continuous linear (resp. quadratic) operator on a Banach algebra A, then

$$\sup\{\|T(a)a - aT(a)\|:\ a \in A, \|a\| = 1\} \leq 2\|T - S\|$$

for each commuting continuous linear (resp. quadratic) operator S on A. We show that the condition that $\sup_{a\in A,\|a\|=1}\|T(a)a - aT(a)\|$ is small implies that T is close to some commuting map. A natural framework for this question is the class of ultraprime Banach algebras, in which algebraic descriptions of commuting and related maps get particularly nice forms.

In that context we also study approximate Lie isomorphisms and approximate Lie derivations. Given Banach algebras A and B and continuous linear maps $\Phi\colon B \to A$ and $\Delta\colon A \to A$, we measure the Lie multiplicativity of Φ and the Lie derivativity of Δ through the constants

$$\operatorname{lmult}(\Phi) = \sup\{\|\Phi([a,b]) - [\Phi(a),\Phi(b)]\|:\ a,b \in B,\ \|a\| = \|b\| = 1\}$$

and

$$\operatorname{lder}(\Delta) = \sup\{\|\Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)]\|:\ a,b \in A,\ \|a\| = \|b\| = 1\},$$

respectively. We investigate whether the conditions of $\operatorname{lmult}(\Phi)$ and $\operatorname{lder}(\Delta)$ being small imply that Φ and Δ are near actual Lie homomorphisms and Lie derivations, respectively. We obtain the ultimate results for $\mathcal{L}(H)$, where H is a Hilbert space, relying on [**Joh88**] and [**AEV10**].

THEOREM. *(See Theorem 1.4.2, Proposition 1.4.6.) Let H be separable Hilbert space. For each $M,\varepsilon > 0$ there exists $\delta > 0$ such that if $\Phi\colon \mathcal{L}(H) \to \mathcal{L}(H)$ is a bijective continuous linear map with $\|\Phi\|, \|\Phi^{-1}\| \leq M$ and $\operatorname{lmult}(\Phi) < \delta$ then*

$$\min\left\{\operatorname{dist}\big(\Phi, \operatorname{Hom}(\mathcal{L}(H))\big), \operatorname{dist}\big(\Phi, -\operatorname{AHom}(\mathcal{L}(H))\big)\right\} < \varepsilon.$$

A related problem on approximate Jordan isomorphisms naturally appeared in the [**AEV11**] on approximately spectrum-preserving maps. In [**AEV11**] it is shown that the classical Herstein's theorem on Jordan epimorphisms is stable in the sense that approximate Jordan epimorphisms are either approximate epimorphisms or approximate anti-epimorphisms. Here we use similar techniques, but commuting and Lie maps are more demanding from the technical aspect because of the presence of central maps.

Turning back to derivations we give a metric version of two classical theorems on derivations proved by Posner in [**Pos57**]. A metric version of the first Posner's theorem is obtained in [**Bre91**] by estimating the distance from the composition D_1D_2 of two derivations D_1 and D_2 on an ultraprime Banach algebra A to the set of all generalized derivations on A. We measure the "derivativity" of a given continuous linear operator T on an ultraprime Banach algebra A through the constant

$$\operatorname{der}(T) = \sup\{\|T(ab) - T(a)b - aT(b)\|\colon \|a\| = \|b\| = 1\}.$$

We estimate $\|S\|\|T\|$ in terms of $\operatorname{der}(S)$, $\operatorname{der}(T)$, and $\operatorname{der}(ST)$ for arbitrary continuous linear operators S and T on A. Further, we present a metric version of the second Posner's theorem by estimating $\|T\|\sup\{\|ab - ba\|\colon \|a\| = \|b\| = 1\}$ in terms of $\operatorname{der}(T)$ and $\sup\left\{\operatorname{dist}\big([T(a),a], Z(A)\big)\colon \|a\| = 1\right\}$.

f-homomorphisms and f-derivations

In the last chapter we consider some generalizations of Lie maps and derivations, and try to understand how rigid the structure of algebras in question with respect to those is. Let f be a nonzero multinear noncommutative polynomial. We say that a map ϕ from an algebra A into itself *preserves zeros of f* if for all $a_1,\dots,a_d \in A$,

$$f(a_1,\dots,a_d) = 0 \implies f(\phi(a_1),\dots,\phi(a_d)) = 0.$$

The list of all maps on A that preserve zeros of f must certainly contain scalar multiples of automorphisms, for some polynomials it must also contain scalar multiples of antiautomorphisms (e.g. for $f = x_1x_2 + x_2x_1$), and for some even all maps of the form

$$\phi(x) = \alpha\theta(x) + \mu(x), \tag{0.4}$$

where $\alpha \in F$, $\theta : A \to A$ is either an automorphism or an antiautomorphism, and μ is a linear map from A into its center (e.g. for $f = x_1x_2 - x_2x_1$).

For certain simple polynomials, especially for $f = x_1x_2$ and $f = x_1x_2 - x_2x_1$, the problem has a long and rich history; see, for example, [**ABEV09**] and [**BCM07**] for historic comments and references. So far not much is known for general polynomials. For them the problem was explicitly posed by Chebotar et al. [**CFL05**] for the matrix algebra $A = M_n(F)$, and some partial solutions were obtained in two recent papers: [**GK09**] considers, in particular, the case where the sum of coefficients of f is a nonzero scalar (without assuming the linearity of ϕ), and [**DD10**] handles Lie polynomials of degree at most 4. Related but at first sight closer to homomorphisms are f-homomorphisms; i.e., linear maps that preserve all values of f, $\phi(f(a_1, \dots, a_d)) = f(\phi(a_1), \dots, \phi(a_d))$ for all $a_i \in A$. They can be described at a high level of generality by using functional identities, although for finite dimensional algebras (including $M_n(F)$) the obtained results are not optimal; see [**BF99**] and also [**BCM07**, Section 6.5].

We first consider maps preserving zeros of a polynomial for some special polynomials in the context of rather general classes of prime algebras and C^*-algebras, and prove that all are of the standard form (0.4). On matrix algebras $A = M_n(F)$ we show that every such map is of the desired form for an arbitrary polynomial under some mild technical restrictions.

THEOREM. *(See Theorem 2.2.1.) Let F be a field with* $\mathrm{char}(F) = 0$*, let $f \in F\langle X\rangle$ be a multilinear polynomial of degree $d \geq 2$, and let $\phi : M_n(F) \to M_n(F)$ be a bijective linear map that preserves zeros of f and satisfies $\phi(1) \in F \cdot 1$. Assume that $n \neq 2, 4$ and $d < 2n$. Then ϕ is of the standard form* (0.4)*.*

One needs to observe that the set of such maps is an algebraic subgroup of GL_{n^2} containing (inner) automorphisms of $M_n(F)$, and then use the description of those obtained by Platonov and Đoković in [**PĐ95**]. With a careful analysis of their proof one can also give a list of all algebraic Lie subalgebras of $\mathfrak{gl}_{n^2}$ that contain the Lie algebra of all (inner) derivations, and thereby describe f-derivations.

Notation and conventions

Notation and conventions wil be specified in each section separately. F always denotes a field. In some sections we additionally assume that it is of characteristic 0, and in some that it is algebraically closed.

Credits

The material presented in this thesis is an accumulation of notes that appeared in [**BPŠ15**, **BŠ14a**, **BŠ14b**, **KŠ14b**, **Špe13**, **KŠ15**, **ŠVdB15**, **KŠ14a**, **BŠ12a**, **BMŠ12**, **ABE$^+$14**, **AEŠV12b**, **AEŠV12a**, **ABŠV12**, **BŠ12b**].

CHAPTER 2

Identities on $M_n(F)$ and the Cayley-Hamilton polynomial

In this chapter we study various types of identities on the matrix algebra $M_n(F)$. We begin in Section 1 with a short overview of the most known types of identities, i.e., polynomial and trace identities. This introduction also serves as preliminaries for the rest of this chapter and for Chapters 3, 4.

In Section 2 we consider identities on the matrix algebra $M_n(F)$ that are defined similarly as the trace identities, except that the "coefficients" are arbitrary commutative polynomials, not necessarily those expressible by the traces. The main issue is whether such an identity is a consequence of the Cayley-Hamilton identity. We show that this holds for several special cases, and, moreover, for every such an identity P and every central polynomial c with zero constant term there exists $m \in \mathbb{N}$ such that the affirmative answer holds for $c^m P$. In general, however, the answer is negative. We prove that there exist antisymmetric identities that do not follow from the Cayley-Hamilton identity, and give a complete description of a certain family of such identities.

Complete solutions of functional identities $\sum_{k\in K} F_k(\overline{x}_m^k)x_k = \sum_{l\in L} x_l G_l(\overline{x}_m^l)$ on the matrix algebra $M_n(F)$ are given in Section 3. We use the coordinate-wise approach, which makes it possible to use the theory of syzygies on generic matrices due to [**Onn94**]. The nonstandard parts of these solutions turn out to follow from the Cayley-Hamilton identity.

In Section 4 we study functional identities in one variable on finite dimensional centrally closed prime algebra. Let $q : A \to A$ be the trace of a d-linear map; i.e., $q(x) = M(x, \dots, x)$ where $M : A^d \to A$ is a d-linear map. If $[q(x), x] = 0$ for every $x \in A$, then q is of the form $q(x) = \sum_{i=0}^{d} \mu_i(x)x^i$ where each μ_i is the trace of a $(d-i)$-linear map from A into F. For infinite dimensional algebras and algebras of dimension $> d^2$ this was proved by Lee, Lin, Wang, and Wong in 1997 [**LWLW97**]. We cover the remaining case where the dimension is $\leq d^2$. Using this result we are able to handle general functional identities in one variable on A; more specifically, we describe the traces of d-linear maps $q_i : A \to A$ that satisfy $\sum_{i=0}^{m} x^i q_i(x)x^{m-i} \in F$ for every $x \in A$.

This chapter is based on [**BPŠ15, BŠ14a, BŠ14b**].

1. Theory of identities

We recall some facts about the classical theory of identities that will be needed in the subsequent sections and chapters.

1.1. Polynomial identities. Polynomial identities appear in the formalism of *universal algebra*. Whenever we have some category of algebras which admits *free algebras* one has the concept of identities in m variables (where m can also be ∞), of an algebra A. That is the ideal of the free algebra $F\langle x_1, \dots, x_m\rangle$ in m variables x_k formed by those elements which vanish for all evaluations of the variables x_k into elements $a_k \in A$. An ideal of identities is a T-ideal; i.e., an ideal of the free algebra closed under substitution of the variables x_k with elements f_i of the same free algebra. It is easily seen that a T-ideal $I \subseteq F\langle x_1, \dots, x_m\rangle$ is automatically the ideal of identities of an algebra, namely $F\langle x_1, \dots, x_m\rangle/I$. Recall that we say that a T-ideal I is generated as T-ideal by a subset S if it is the minimal T-ideal containing S. That is, it is generated as an ideal by all subsets obtained from S applying substitution of variables with elements of the free algebra.

Of particular interest is the case of noncommutative associative algebras over a field F, for which we assume, for simplicity, that

$$\operatorname{char}(F) = 0$$

(this assumption will be used throughout this section and in the following section without further mention). In this case the free algebra in m variables x_k is the usual algebra of noncommutative polynomials with basis the *words* in the variables x_k. For $m = \infty$ we set

$$X := \{x_k \mid k = 1, 2, \dots\}$$

and let $F\langle X\rangle$ be the corresponding free algebra. In this case a particularly interesting example is the theory of polynomial identities of the algebra $M_n(F)$ of all $n \times n$ matrices over the field F. An implicit description of these identities is given through the *algebra of generic matrices.*

We fix an integer $n \geq 1$, and set

$$\mathcal{C} := F\big[x_{ij}^{(k)} \mid 1 \leq i, j \leq n, k = 1, 2, \dots\big]. \tag{1.1}$$

This commutative polynomial ring is the algebra of polynomial functions on sequences of matrices. Inside the matrix algebra $M_n(\mathcal{C})$ we can define the *generic matrices* ξ_k where ξ_k is the matrix with entries the variables $x_{ij}^{(k)}$. It is then easily seen, since F is assumed to be infinite, that the ideal of polynomial identities of $M_n(F)$ is the kernel of the evaluation map $x_k \mapsto \xi_k$.

The F-subalgebra of $M_n(\mathcal{C})$ generated by the ξ_k, i.e., the image of the free algebra under this evaluation, is the free algebra in the category of noncommutative algebras satisfying the identities of $M_n(F)$. This algebra has been extensively studied although a very precise description is available only for $n = 2$. We shall denote it by $F\langle \xi_k \mid n\rangle$ or just $F\langle \xi_k\rangle$ if the integer n is fixed, and call it the *algebra of generic matrices.* (It is also standard to denote it as GM_n, and sometimes we will use also this notation.)

1.2. Trace identities. When dealing with matrices in characteristic 0, it is useful to think that they form an algebra with a further unary operation the *trace*, $x \mapsto \operatorname{tr}(x)$. One can formalize this as follows.

An *algebra with trace* is an algebra R equipped with an additional structure, that is a linear map $\operatorname{tr} : R \to R$ satisfying the following properties

$$\operatorname{tr}(ab) = \operatorname{tr}(ba), \quad a \operatorname{tr}(b) = \operatorname{tr}(b)\, a, \quad \operatorname{tr}(\operatorname{tr}(a)b) = \operatorname{tr}(a) \operatorname{tr}(b)$$

for all $a, b \in R$. The notion of a morphism between algebras with trace is then obvious and such algebras form a category which contains free algebras.

In this case the free algebra is the algebra of noncommutative polynomials with basis the *words* in the variables x_k but over the polynomial ring $\mathfrak{T}$ in the infinitely many variables $\operatorname{tr}(M)$, where M runs over all possible words considered equivalent under cyclic moves (i.e., $ab \sim ba$). The elements of this free algebra are called *trace polynomials*, while the elements of $\mathfrak{T}$ are *pure trace polynomials.* The degree of a trace monomial $\operatorname{tr}(w_1) \cdots \operatorname{tr}(w_m)v$, $w_i, v \in \langle X\rangle$, equals $|v| + \sum_i |w_i|$, where $|u|$ denotes the length of a word u. The degree of a trace polynomial is the maximum of the degrees of its trace monomials.

In this setting again the *trace identities* of matrices are the kernel of the evaluation of the free algebra into the generic matrices, but now the image is the subalgebra $\mathcal{T}_n\langle\xi_k\rangle$ of $M_n(\mathcal{C})$ generated by the generic matrices and the algebra $\mathcal{T}_n$, the image of $\mathfrak{T}$, generated by all traces of the monomials in the ξ_k.

Let us further denote by $\mathcal{T}_{m,n}$ (resp. $\mathcal{T}_{m,n}\langle\xi_k\rangle$) the subalgebra of $\mathcal{C}$ (resp. $M_n(\mathcal{C})$) generated by all traces of monomials in ξ_k, $1 \le k \le m$, (and generic matrices ξ_k, $1 \le k \le m$). Let $\mathcal{T}_{\infty,n}$ (resp. $\mathcal{T}_{\infty,n}\langle\xi_k\rangle$) stand for $\mathcal{T}_n$ (resp. $\mathcal{T}_n\langle\xi_k\rangle$).

It is a remarkable fact that in this setting both the trace identities and the *free* algebra $\mathcal{T}_n\langle\xi_k\rangle$ can be interpreted in the language of the first and second fundamental theorem for matrices (FFT and SFT). The projective linear group $G := PGL_n(F)$ acts by conjugation on matrices and hence also on sequences of matrices, and for every $1 \le m \le \infty$ we have the following fundamental theorems (see [**Pro07**, Chapter 11]).

THEOREM 1.2.1.

FFT: *The algebra $\mathcal{T}_{m,n}$ is the algebra of G-invariant polynomial functions on the space of m-tuples of $n \times n$ matrices. The algebra $\mathcal{T}_{m,n}\langle\xi_k\rangle$ is the algebra of G-equivariant polynomial maps from m-tuples of $n \times n$ matrices to $n \times n$ matrices.*

SFT: *The ideal of trace identities on $n \times n$ matrices is generated, as a T-ideal, by the Cayley-Hamilton polynomial.*

Another way of stating the FFT is by noticing that G acts on $\mathcal{C}$ by $gf(x) := f(g^{-1}x)$ and on $M_n(F)$ by conjugation, hence it acts on $M_n(\mathcal{C}) = \mathcal{C} \otimes_F M_n(F)$ and we have

$$\mathcal{T}_n\langle\xi_k\rangle = M_n(\mathcal{C})^G, \quad \mathcal{T}_n = \mathcal{C}^G.$$

Notice that as soon as $n \ge 2$ the algebra $\mathcal{T}_{m,n}$ is the center of $\mathcal{T}_{m,n}\langle\xi_k\rangle$.

The FFT for matrices is essentially classical, as for the SFT, Procesi [**Pro76**] and Razmyslov [**Raz74**] proved that the T-ideal of (resp. pure) trace identities of $M_n(F)$ is generated by $Q_n(x_1, \dots, x_n)$ (resp. $\operatorname{tr}(Q_n(x_1, \dots, x_n)x_{n+1})$), where Q_n is the multilinear *Cayley-Hamilton polynomial.* Let us recall that the Cayley-Hamilton polynomial is

$$q_n = q_n(x_1) = x_1^n + \tau_1(x_1)x_1^{n-1} + \cdots + \tau(x_1).$$

As it is well-known, each $\tau_i(x_1)$ can be expressed (in characteristic 0) as a $\mathbb{Q}$-linear combination of the products of $\operatorname{tr}(x_1^j)$. Evaluating in $M_n(\mathcal{C})$ we have $\tau_1(\xi_1) = -\operatorname{tr}(\xi_1) = -(x_{11}^{(1)} + \cdots + x_{nn}^{(1)}), \dots, \tau_n(\xi_1) = (-1)^n \det(\xi_1)$. Now, $Q_n(x_1, \dots, x_n)$ denotes the multilinear version of $q_n(x_1)$ obtained by full polarization. Recall that it can be written as

$$Q_n := \sum_{\sigma \in S_{n+1}} \epsilon_\sigma \phi_\sigma(x_1, \dots, x_n) \tag{1.2}$$

where $\epsilon_\sigma = \pm 1$ denotes the sign of the permutation σ, while ϕ_σ is defined using the cycle decomposition of

$$\sigma = (i_1, \dots, i_{k_1})(j_1, \dots, j_{k_2}) \dots (u_1, \dots, u_h)(s_1, \dots s_k, n+1)$$

as

$$\phi_\sigma(x_1, \dots, x_n) = \operatorname{tr}(x_{i_1} \cdots x_{i_{k_1}}) \operatorname{tr}(x_{j_1} \cdots x_{j_{k_2}}) \cdots \operatorname{tr}(x_{u_1} \cdots x_{u_h}) x_{s_1} \cdots x_{s_k}.$$

Thus, for example,

$$Q_2(x_1, x_2) = x_1x_2 + x_2x_1 - \operatorname{tr}(x_1)x_2 - \operatorname{tr}(x_2)x_1 + \operatorname{tr}(x_1)\operatorname{tr}(x_2) - \operatorname{tr}(x_1x_2).$$

Note that $Q_n(x_1, \dots, x_n)$ is symmetric, i.e., $Q_n(x_1, \dots, x_n) = Q_n(x_{\sigma(1)}, \dots, x_{\sigma(n)})$ for every permutation σ, and that $q_n(x_1) = \frac{1}{n!}Q_n(x_1, \dots, x_1)$. By a slight abuse of terminology, we will call both Q_n and q_n the Cayley-Hamilton polynomial, or, when associated with $M_n(F)$, the *Cayley-Hamilton identity.* In view of the terminology introduced below, more accurate names in the setting of the next section may be the Cayley-Hamilton quasi-polynomial (resp. quasi-identity), but we omit "quasi". for simplicity.

1.3. Central polynomials. An element of the free algebra $F\langle X\rangle$ is a *central polynomial* for $n \times n$ matrices if it takes scalar values under any evaluation into matrices. It is then clear that the center, denoted $\mathcal{Z}_n$, of the algebra $F\langle \xi_k\rangle$ of generic matrices, is the image of the set of central polynomials. A basic discovery based on the existence of central polynomials found independently by Formanek [**For72**] and Razmyslov [**Raz73**] is that the center $\mathcal{Z}_n$ is rather large. Then fundamental theorems of PI theory tell us that the central quotient of $F\langle \xi_k\rangle$ and of $\mathcal{T}_n\langle \xi_k\rangle$ coincide and as soon as $n \geq 2$ give rise to a division algebra of rank n^2 over its center which is the field of quotients of both $\mathcal{Z}_n$ and $\mathcal{T}_n$.

2. Quasi-identities

2.1. Quasi-identities. The purpose of this subsection is to introduce the setting and record some easy results on *quasi-identities.* Let us point out that we will consider our problems exclusively on the algebra $M_n = M_n(F)$ over a field F of characteristic 0.

We will define a quasi-polynomial in a slightly different way than in [**BC00a**] and [**BCM07**]. Our definition is not restricted to the multilinear situation, and, on the other hand, is adjusted for applications to the matrix algebra M_n.

A *quasi-polynomial* is an element of the algebra $\mathcal{C}\big\langle X\big\rangle$, the free algebra in the variables X with coefficients in the polynomial algebra of functions on $M_n(F)^{|X|}$.

Thus, a quasi-polynomial is a polynomial in the noncommuting indeterminates x_k whose coefficients are ordinary polynomials in the commuting indeterminates $x_{ij}^{(k)}$, coordinates of the space $M_n(F)^{|X|}$. A quasi-polynomial P can be therefore uniquely written as

$$P = \sum \lambda_M M,$$

where M is a noncommutative monomial in the x_k's and λ_M is a commutative polynomial in the $x_{ij}^{(k)}$'s, that is a polynomial function on sequences of matrices. Of course, P depends on finitely many x_k's and finitely many $x_{ij}^{(k)}$'s. We can therefore write

$$P = P(x_{11}^{(1)}, \dots, x_{nn}^{(1)}, \dots, x_{11}^{(m)}, \dots, x_{nn}^{(m)}, x_1, \dots, x_m)$$

for some m. It is possible to put this setting in the framework of universal algebra, but we shall limit to the following easy facts.

2.1.1. *Substitutional rules.* Commutative indeterminates $x_{ij}^{(k)}$ have a *substitutional rule*, that is given as follows. We have a map $\Phi : x_k \mapsto \xi_k$ of $\mathcal{C}\big\langle X\big\rangle$ to $M_n(\mathcal{C})$ which maps x_k to the corresponding generic matrix and is the identity on $\mathcal{C}$, so for each choice of $f \in \mathcal{C}\big\langle X\big\rangle$ it makes sense to speak of $\Phi(f)_{ij}$, the (i,j) entry of $\Phi(f)$. The substitution in $\mathcal{C}\big\langle X\big\rangle$ should be understood as that one substitutes $x_k \mapsto f_k \in \mathcal{C}\big\langle X\big\rangle$ and simultaneously $x_{ij}^{(k)} \mapsto \Phi(f_k)_{ij}$, we thus define

Definition 2.1.1. A *T-ideal* of $\mathcal{C}\big\langle X\big\rangle$ is an ideal that is closed under all such substitutions.

It is convenient to use a more suggestive notation and write $\lambda_M(x_1, \dots, x_m)$ for $\lambda_M(x_{11}^{(1)}, \dots, x_{nn}^{(1)}, \dots, x_{11}^{(m)}, \dots, x_{nn}^{(m)})$, and hence $P(x_1, \dots, x_m)$ for P. We define the *evaluation* of P at an m-tuple $a_1, \dots, a_m \in M_n$, $P(a_1, \dots, a_m)$, by substituting a_k for x_k and $a_{ij}^{(k)}$ for $x_{ij}^{(k)}$, where $a_k = (a_{ij}^{(k)})$.

DEFINITION 2.1.2. A quasi-polynomial $P \in \mathcal{C}\langle X\rangle$ is a *quasi-identity* of M_n if $P(a_1, \ldots, a_m) = 0$ for all $a_1, \ldots, a_m \in M_n$.

The set $\mathfrak{I}_n$ of all quasi-identities of M_n clearly forms a T-ideal of $\mathcal{C}\big\langle X\big\rangle$. As for polynomial or trace identities, $\mathfrak{I}_n$ is the kernel of the $\mathcal{C}-$linear evaluation map from $\mathcal{C}\big\langle X\big\rangle$ to $M_n(\mathcal{C})$ mapping x_k to the generic matrix ξ_k. Let us give a proof of this simple fact. With I we denote the identity of $M_n(\mathcal{C})$.

LEMMA 2.1.3. *The algebra $\mathcal{C}\big\langle X\big\rangle/\mathfrak{I}_n$ is isomorphic to the subalgebra $\mathcal{C}\langle\xi_k\rangle$ of $M_n(\mathcal{C})$ generated by all generic matrices $\xi_k = (x_{ij}^{(k)})$, $k = 1, 2, \ldots$, and all λI, $\lambda \in \mathcal{C}$.*

PROOF. Let $\Phi : \mathcal{C}\big\langle X\big\rangle \to M_n(\mathcal{C})$ be the homomorphism determined by $\Phi(x_k) = (x_{ij}^{(k)})$ and $\Phi(\lambda) = \lambda I$ for $\lambda \in \mathcal{C}$. It is immediate that $\ker\Phi \subseteq \mathfrak{I}_n$. Given $P = P(x_1, \ldots, x_m) \in \mathfrak{I}_n$ we have $P(a_1, \ldots, a_m) = 0$ for all $a_i \in M_n$. Since $\mathrm{char}(F) = 0$, and hence F is infinite, a standard argument shows that $\Phi(P) = 0$. Thus, $\ker\Phi = \mathfrak{I}_n$, and the result follows. $\square$

In fact the evaluation ρ of the free algebra with trace to generic matrices with traces factors through $\mathcal{C}\big\langle X\big\rangle$

$$\rho : \mathfrak{T}\langle X\rangle \xrightarrow{\pi} \mathcal{C}\big\langle X\big\rangle \to M_n(\mathcal{C})$$

by evaluating the trace monomials $\mathrm{tr}(x_{i_1} \cdots x_{i_m})$ into $M_n(\mathcal{C})$ but keeping fixed the free variables. The image of π is the algebra $\mathcal{T}_n\big\langle X\big\rangle$ of invariants of the algebra $\mathcal{C}\big\langle X\big\rangle$ with respect to the action of the projective group on the coefficients $\mathcal{C}$ and fixing the variables X.

Thus the image through π of a trace polynomial can also be viewed as a quasi-polynomial $\sum \lambda_M M$, but such that every λ_M is an invariant and thus can be expressed as a linear combination of the products of $\mathrm{tr}(x_{i_1} \cdots x_{i_m})$.

Every trace identity gives rise to a quasi-identity of M_n, but a nontrivial element of $\mathfrak{T}\langle X\rangle$ may very well map to 0 under π, so a nontrivial trace identity may correspond to a trivial quasi-identity. In view of the SFT for matrices we may again consider the quasi-polynomial arising from the Cayley-Hamilton theorem, therefore it is natural to look in this context for a possible analogue of the SFT for quasi-identities (cf. Theorem 1.2.1).

DEFINITION 2.1.4. We shall say that a quasi-identity P of M_n is a *consequence of the Cayley-Hamilton identity* if P lies in the T-ideal of $\mathcal{C}\big\langle X\big\rangle$ generated by Q_n.

The question pointed out in the introduction thus asks the following:

QUESTION. Is the T-ideal $\mathfrak{I}_n$ generated by Q_n?

As we have already remarked, the ideal of quasi-identities $\mathfrak{I}_n$ is the kernel of the evaluation map Φ of $\mathcal{C}\big\langle X\big\rangle$ into $M_n(\mathcal{C})$ mapping the variables to the generic matrices. In view of Lemma 2.1.3 we have a sequence of inclusion maps

$$F\langle\xi_k\rangle \subset \mathcal{T}_n\langle\xi_k\rangle \subset \mathcal{C}\langle\xi_k\rangle.$$

Our first remark is that, unlike $F\langle\xi_k\rangle \cong F\langle X\rangle/\mathrm{id}(M_n)$ and $\mathcal{T}_n\langle\xi_k\rangle$, $\mathcal{C}\langle\xi_k\rangle \cong \mathcal{C}\big\langle X\big\rangle/\mathfrak{I}_n$ is not a domain. This can be deduced from Lemma 2.1.6 below, but let us give a simple concrete example.

EXAMPLE 2.1.5. Note that none of

$$P_1 = x_{12}^{(2)}x_1 - x_{12}^{(1)}x_2 + x_{12}^{(1)}x_{22}^{(2)} - x_{22}^{(1)}x_{12}^{(2)}$$

and

$$P_2 = x_{12}^{(2)}x_1 - x_{12}^{(1)}x_2 + x_{12}^{(1)}x_{11}^{(2)} - x_{11}^{(1)}x_{12}^{(2)}$$

lies in $\mathfrak{I}_2$, but P_1P_2 does.

The center of $\mathcal{C}\langle X\rangle/\mathfrak{I}_n$ is isomorphic to $\mathcal{C}$, which is a domain. We may therefore form the algebra of central quotients of $\mathcal{C}\langle X\rangle/\mathfrak{I}_n$, which consists of elements of the form αR where $R \in \mathcal{C}\langle X\rangle/\mathfrak{I}_n$ and α lies in

$$\mathcal{K} := F\big(x_{ij}^{(k)} \mid 1 \le i, j \le n, k = 1, 2, \ldots\big),$$

the field of rational functions in $x_{ij}^{(k)}$ (cf. [**Row80**, p. 54]). In order to describe this $\mathcal{K}$-algebra, we invoke the Capelli polynomials

$$C_{2k-1}(x_1, \ldots, x_k, y_1, \ldots, y_{k-1}) := \sum_{\sigma \in S_k} \epsilon_\sigma x_{\sigma(1)} y_1 x_{\sigma(2)} y_2 \cdots x_{\sigma(k-1)} y_{k-1} x_{\sigma(k)},$$

where ϵ_σ is the sign of the permutation σ. As it is well-known, C_{2n^2-1} is a polynomial identity of every proper subalgebra of $M_n(E)$ but not of $M_n(E)$ itself, for every field E [**Row80**, Theorem 1.4.8].

LEMMA 2.1.6. *The algebra of central quotients of $\mathcal{C}\langle X\rangle/\mathfrak{I}_n$ is isomorphic to $M_n(\mathcal{K})$.*

PROOF. Since C_{2n^2-1} is not a polynomial identity of $M_n(F)$, it is also not a polynomial identity of the $\mathcal{K}$-subalgebra of $M_n(\mathcal{K})$ generated by all generic matrices $(x_{ij}^{(k)})$, $k = 1, 2, \ldots$. But then this subalgebra is the whole algebra $M_n(\mathcal{K})$. Now we can apply Lemma 2.1.3. □

2.1.2. *Geometric viewpoint.* We give an indication of a geometric meaning of quasi-identities. The following theorem gathers together some known facts, but we recall them for completeness.

THEOREM 2.1.7. *If the constant term $c \in \mathcal{Z}_n$ is zero, then $F\langle\xi_k\rangle[c^{-1}] = \mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$ is a rank n^2 Azumaya algebra over its center $\mathcal{Z}_n[c^{-1}] = \mathcal{T}_n[c^{-1}]$. Moreover,*

$$M_n(\mathcal{C}[c^{-1}]) \cong \mathcal{C}[c^{-1}] \otimes_{\mathcal{Z}_n[c^{-1}]} \mathcal{T}_n\langle\xi_k\rangle[c^{-1}].$$

PROOF. As it is well known and easy to see, c is an identity of M_{n-1}. Since c is invertible in $F\langle\xi_k\rangle[c^{-1}]$ and $\mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$, these two algebras cannot have nonzero quotients satisfying the identities of M_{n-1}. It follows from the Artin-Procesi theorem that $F\langle\xi_k\rangle[c^{-1}]$ and $\mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$ are Azumaya algebras over their centers of rank n^2. These centers are clearly $\mathcal{Z}_n[c^{-1}]$ and $\mathcal{T}_n[c^{-1}]$. By general properties, the reduced trace of $x \in F\langle\xi_k\rangle[c^{-1}]$ is just the trace of x considered as a matrix in $F\langle\xi_k\rangle[c^{-1}]$, and $F\langle\xi_k\rangle[c^{-1}]$ is closed under the reduced trace. Hence every element in $\mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$ is contained in $F\langle\xi_k\rangle[c^{-1}]$. Accordingly, $F\langle\xi_k\rangle[c^{-1}] = \mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$ and $\mathcal{Z}_n[c^{-1}] = \mathcal{T}_n[c^{-1}]$.

Recall a standard fact (see [**AG60b**] and [**AG60a**] or [**Sal99**, Theorem 2.8]) that if $R \subseteq S$, R is an Azumaya algebra and the center $Z(R)$ of R is contained in the center $Z(S)$ of S, then $S \cong R \otimes_{Z(R)} R'$ where R' is the centralizer of R in S. Taking $\mathcal{T}_n\langle\xi_k\rangle[c^{-1}]$ for R and $M_n(\mathcal{C}[c^{-1}])$ for S we obtain the last assertion of the theorem. □

We should remark that this theorem has a geometric content. Let F be algebraically closed. If we fix the number of generic matrices to a finite number m, we have the action by simultaneous conjugation of $G := \mathrm{PGL}_n(F)$ on the affine space $M_n(F)^m$. By geometric invariant theory the algebra $\mathcal{T}_{m,n}$ is the coordinate ring of the categorical quotient $M_n(F)^m // \mathrm{PGL}_n(F)$, a variety parameterizing the closed orbits, which correspond to isomorphism classes of semisimple representations of dimension n of the free algebra in m generators (cf. [**Art69**]).

The action of the projective group G is free on the open set of irreducible representations and the complement of this open set is exactly the subvariety of

m-tuples of matrices where all central polynomials with no constant term vanish. The Azumaya algebra property reflects this geometry. Except for the special case $m = n = 2$ the variety $M_n(F)^m // \mathrm{PGL}_n(F)$ is smooth exactly on this open set and the quotient map $M_n(F)^m \to M_n(F)^m // \mathrm{PGL}_n(F)$ is not flat over the singular set (cf. [**LBP87**]). As we shall see these singularities are in some sense measured by the quasi-identities of matrices modulo those which are a consequence of the Cayley Hamilton identity, see the exact sequence (2.1). This will be described as a module $\mathfrak{I}_n/(Q_n)$ supported in the singular part of the quotient variety. On the other hand to prove that this module is indeed nontrivial is quite difficult and although we will show this, we only have a partial description of this phenomenon, the description of the *antisymmetric part of the module.*

2.2. Quasi-identities and the Cayley-Hamilton identity.

2.2.1. *Trace algebras and the Cayley-Hamilton identity.* We begin by reformulating our problem in the commutative algebra framework by using the result from [**Pro87**]. Let us, therefore, recall the content of that paper. We have already seen in §1.2 the notion of an algebra with trace, in particular for any commutative algebra $\mathcal{A}$ we consider $M_n(\mathcal{A})$ with the usual trace.

For an algebra with trace R and a number $n \in \mathbb{N}$, we define the universal map into $n \times n$ matrices as a pair of a commutative algebra $\mathcal{A}_R$ and a morphism (of algebras with trace) $j : R \to M_n(\mathcal{A}_R)$ with the following universal property: for any other map (of algebras with trace) $f : R \to M_n(\mathcal{B})$ with $\mathcal{B}$ commutative there is a unique map $\bar{f} : \mathcal{A}_R \to \mathcal{B}$ of commutative algebras making the diagram commutative

$$\begin{tikzcd} R \arrow[r, "j"] \arrow[dr, "f"'] & M_n(\mathcal{A}_R) \arrow[d, "M_n(\bar{f})"] \\ & M_n(\mathcal{B}) \end{tikzcd}$$

The existence of such a universal map is easily established, although in general it may be 0.

The main idea comes from category theory, that is, from representable functors. We take the functor from commutative algebras to sets which associates to a commutative algebra $\mathcal{B}$ the set of (trace preserving) morphisms $\hom(R, M_n(\mathcal{B}))$ and want to prove that it is representable, i.e., that there is a commutative algebra $\mathcal{A}_R$ and a natural isomorphism $\hom(R, M_n(\mathcal{B})) \cong \hom(\mathcal{A}_R, \mathcal{B})$. Then the identity map $1_{\mathcal{A}_R} \in \hom(\mathcal{A}_R, \mathcal{A}_R)$ corresponds to the universal map $j \in \hom(R, M_n(\mathcal{A}_R))$.

We have seen the category of algebras with trace has free algebras which are the usual free algebras in indeterminates x_k to which we add a commutative algebra $\mathfrak{T}$ of *formal traces.* Then we see that the commutative algebra associated to a free algebra is the polynomial algebra in indeterminates $x_{ij}^{(k)}$. The universal map maps x_k to the generic matrix with entries $x_{ij}^{(k)}$ and the formal traces $\mathrm{tr}(x_{i_1} x_{i_2} \dots x_{i_k})$ map to the traces of the corresponding monomials in the generic matrices. From a presentation of R as a quotient of a free algebra one obtains a presentation of $\mathcal{A}_R$ as a quotient of the ring of polynomials in the $x_{ij}^{(k)}$.

If we consider now algebras over a field F (which is of characteristic 0 by the above convention) we have that the group $G = \mathrm{GL}_n(F)$ of invertible $n \times n$ matrices (in fact the projective group $\mathrm{PGL}_n(F) = \mathrm{GL}_n(F)/F^*$) acts on the algebra $\mathcal{A}_R$ and it also acts by conjugation on $M_n(F)$, so it acts diagonally on $M_n(\mathcal{A}_R)$. The main theorem of [**Pro87**] says that

THEOREM 2.2.1. *The image of j is the invariant algebra $M_n(\mathcal{A}_R)^G$ and the kernel of j is the trace-ideal generated by the evaluations of the formal Cayley-Hamilton expression for the given n. In particular, if R satisfies the n-th Cayley-Hamilton identity, then j is injective.*

2.2.2. *The trace on $\mathcal{C}\langle X\rangle$.* Now we will apply this theory to $\mathcal{C}\langle X\rangle$. For this we need to make it into an algebra with trace. For reasons that will soon become clear, let us write $\mathcal{C}_x$ for $\mathcal{C}$ and hence $\mathcal{C}_x\langle X\rangle$ until the end of this section.

DEFINITION 2.2.2. We define the trace $\operatorname{tr} : \mathcal{C}_x\langle X\rangle \to \mathcal{C}_x$ as the $\mathcal{C}_x$-linear map satisfying $\operatorname{tr}(1) = n$ and mapping a monomial in the indeterminates x_k into the trace of the corresponding monomial in generic matrices ξ_k in the indeterminates $x_{ij}^{(k)}$.

In order to understand what is the universal map of this algebra with trace into $n \times n$ matrices we introduce a second polynomial algebra $\mathcal{C}_y = F\big[y_{ij}^{(k)} \mid 1 \le i, j \le n, k = 1, 2, \ldots\big]$.

The group G acts on $\mathcal{C}_x$ and $\mathcal{C}_y$, and, by the FFT, the invariants are in both cases the invariants of matrices, that is the algebra generated by the traces of monomials. We identify the two algebras of invariants and call this algebra $\mathcal{T}_n$.

Now we set

$$\mathcal{A}_n := \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y,$$

and let $\xi_k := (y_{ij}^{(k)})$ denote the generic matrix in $M_n(\mathcal{C}_y)$. Note that the algebra $\mathcal{T}_n\langle \xi_k, k = 1, 2, \ldots\rangle$ may be identified with the algebra $\mathcal{T}_n\langle \xi_k\rangle$ of equivariant maps studied in Theorem 1.2.1 and which has $\mathcal{T}_n$ as the center.

From now on let $j : \mathcal{C}\langle X\rangle \to M_n(\mathcal{A}_n)$ denote the $\mathcal{C}_x$-linear map which maps x_k to the generic matrix $\xi_k = (y_{ij}^{(k)})$, and let (Q_n) denote the T-ideal of $\mathcal{C}\langle X\rangle$ generated by Q_n. Since we are thinking of $\mathcal{C}_x$ as a coefficient ring, in the next proposition the action of G on $\mathcal{A}_n = \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y$, is by acting on the second factor $\mathcal{C}_y$. The action on $M_n(\mathcal{A}_n) = M_n(F) \otimes_F \mathcal{A}_n$ is the tensor product action.

PROPOSITION 2.2.3.
(1) The map $j : \mathcal{C}_x\langle X\rangle \to M_n(\mathcal{A}_n) = M_n(\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y)$ is the universal map into matrices.
(2) The algebra $\mathcal{C}\langle X\rangle/(Q_n)$ is isomorphic to the algebra $M_n(\mathcal{A}_n)^G = \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle \xi_k\rangle$.

PROOF. By Theorem 2.2.1, (2) follows from (1) so it is enough to prove that j is the universal map.

Let be $\mathcal{B}$ be an algebra with trace and let us compute the representable functor $\hom(\mathcal{C}\langle X\rangle, M_n(\mathcal{B}))$. In order to give a homomorphism $\phi : \mathcal{C}\langle X\rangle \to M_n(\mathcal{B})$ in the category of algebras with trace, we have to choose arbitrary elements $a_{ij}^{(k)} \in \mathcal{B}$ for the images of the elements $x_{ij}^{(k)}$, and matrices $b_k = (b_{ij}^{(k)})$ for the images of the elements x_k.

Moreover, if we consider the matrices $a_k = (a_{ij}^{(k)})$ we need to impose that the trace of each monomial formed by the a_k equals the trace of the corresponding monomial formed by the b_k.

Now to give the $a_{ij}^{(k)}$ is the same as to give a homomorphism of $\mathcal{C}_x$ to $\mathcal{B}$, and to give the $b_{ij}^{(k)}$ is the same as to give a homomorphism of $\mathcal{C}_y$ to $\mathcal{B}$. The compatibility means that the restrictions of these two homomorphisms to the algebra $\mathcal{T}_n$, which is contained naturally in both copies, coincide. This is exactly the description of a homomorphism of $\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y$ to $\mathcal{B}$. Thus, j is indeed the universal map.

Next observe that the action of G is only on the factor $\mathcal{C}_y$. By Theorem 2.2.1 it follows that the kernel of j is equal to (Q_n). Thus, it remains to find $M_n(\mathcal{A}_n)^G$, the image of j. Note that $M_n(\mathcal{A}_n) = \mathcal{C}_x \otimes_{\mathcal{T}_n} M_n(\mathcal{C}_y)$ and that G acts trivially on $\mathcal{C}_x$ while on $M_n(\mathcal{C}_y)$ it is the action used in the universal map of the free algebra with trace (see **[Pro87]** for details). By a standard argument on reductive groups we have $M_n(\mathcal{A}_n)^G = \mathcal{C}_x \otimes_{\mathcal{T}_n} M_n(\mathcal{C}_y)^G$, which is by the FFT equal to $\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle$. □

REMARK 2.2.4. The algebra $\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y$, a fiber product, contains a lot of the hidden combinatorics needed to understand quasi-identities. It appears to be a rather complicated object as pointed out by some experimental computations carried out by H. Kraft, which show that even for $n = 2$ as soon as the number of variables is ≥ 3 it is not an integral domain nor is it Cohen-Macaulay. This of course is due to the fact that the categorical quotient described by the inclusion $\mathcal{T}_n = \mathcal{C}^G$ has a rather singular behavior outside the open set parameterizing irreducible representations.

We have to introduce some more notation. We denoted by ξ_k the generic matrix $(y_{ij}^{(k)})$. Analogously, we write η_k for the generic matrix $(x_{ij}^{(k)})$. There is a canonical homomorphism

$$\pi : \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle \to \mathcal{C}_x\langle\eta_k, k = 1, 2, \ldots\rangle.$$
$$\pi : \lambda \otimes f(\xi_1, \ldots, \xi_d) \mapsto \lambda f(\eta_1, \ldots, \eta_d)$$

Note that by Lemma 2.1.3 the latter algebra is nothing but $\mathcal{C}\big\langle X\big\rangle/\mathfrak{I}_n$.

LEMMA 2.2.5. *A quasi-identity P of M_n is not a consequence of the Cayley-Hamilton identity if and only if $j(P)$ is a nonzero element of the kernel of π.*

PROOF. Let $\Phi : \mathcal{C}\big\langle X\big\rangle \to \mathcal{C}_x\langle\eta_k, k = 1, 2, \ldots\rangle$ be the homomorphism from Lemma 2.1.3, i.e., $\Phi(x_k) = \eta_k$ and $\Phi(\lambda) = \lambda I$ for $\lambda \in \mathcal{C}_x$, and let $j : \mathcal{C}\langle X\rangle \to \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle$ be the universal map. Note that $\pi j = \Phi$. Since, by Proposition 2.2.3, $\ker j$ is the T-ideal of $\mathcal{C}\big\langle X\big\rangle$ generated by Q_n, and $\ker\Phi = \mathfrak{I}_n$, this implies the assertion of the lemma. □

COROLLARY 2.2.6. *The space $\mathfrak{I}_n/(Q_n)$, measuring quasi-identities modulo the ones deduced from Q_n, is isomorphic under the map induced by $j : \mathcal{C}\langle X\rangle/(Q_n) \to C_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle$ to the kernel of the map π. That is, we have an exact sequence*

$$0 \longrightarrow \mathfrak{I}_n/(Q_n) \xrightarrow{\ j\ } \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle \xrightarrow{\ \pi\ } M_n(\mathcal{C}_x). \tag{2.1}$$

As an application of Theorem 2.1.7 we have the following theorem on quasi-identities.

THEOREM 2.2.7. *Let P be a quasi-identity of M_n. For every central polynomial c of M_n with zero constant term there exists $m \in \mathbb{N}$ such that $c^m P$ is a consequence of the Cayley-Hamilton identity.*

PROOF. Note that

$$(\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k\rangle)[c^{-1}] \cong \mathcal{C}_x[c^{-1}] \otimes_{\mathcal{T}_n[c^{-1}]} \mathcal{T}_n\langle\xi_k\rangle[c^{-1}] \cong M_n(\mathcal{C}_x[c^{-1}])$$

by Theorem 2.1.7 (the change of variables does not make any difference since $\mathcal{C}_x$ is canonically isomorphic to $\mathcal{C}_y$). This isomorphism is induced by π introduced before Lemma 2.2.5. Therefore $(\ker\pi)[c^{-1}] = 0$. Since every quasi-identity P lies in $\ker(\pi j)$ by Lemma 2.2.5, there exists $m \in \mathbb{N}$ such that $c^m P = 0$ in $\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle\xi_k, k = 1, 2, \ldots\rangle$, i.e., $c^m P$ is a consequence of the Cayley-Hamilton identity by Proposition 2.2.3. □

We have seen that $\ker\pi$ measures the space of quasi-identities modulo the ones deduced from Q_n. This is in fact a $\mathcal{T}_n-$module and, as we shall see, it is nonzero. What the previous theorem tells us is that this module is *supported* in the closed set of non-irreducible representations.

2.3. Antisymmetric quasi-identities.

2.3.1. *Antisymmetric identities derived from the Cayley-Hamilton identity.* By the *antisymmetrizer* we mean the operator that sends a multilinear expression $f(x_1, \dots, x_h)$ into the antisymmetric expression

$$\frac{1}{h!} \sum_{\sigma \in S_h} \epsilon_\sigma f(x_{\sigma(1)}, \dots, x_{\sigma(h)}),$$

where ϵ_σ is the sign of σ. For example, applying the antisymmetrizer to the noncommutative monomial $x_1 \cdots x_h$ we get the standard polynomial of degree h, $S_h(x_1, \dots, x_h) = \sum_{\sigma \in S_h} \epsilon_\sigma x_{\sigma(1)} \dots x_{\sigma(h)}$, and up to scalar this is the only multilinear antisymmetric noncommutative polynomial of degree h. Further, applying the antisymmetrizer to the quasi-monomial $x^{(1)}_{i_1,j_1} \cdots x^{(k)}_{i_k,j_k} x_{k+1} \cdots x_{n^2}$ we get an antisymmetric quasi-polynomial, which is nonzero as long as the pairs (i_l, j_l) are pairwise different, and is, because of the antisymmetry, an identity of every proper subspace of M_n, in particular of the space of trace zero $n \times n$ matrices. Replacing each variable x_k by $x_k - \frac{1}{n}\operatorname{tr}(x_k)$, we thus get a quasi-identity of M_n. Our ultimate goal is to show that *not every such quasi-identity is a consequence of the Cayley-Hamilton identity.* For this we need several auxiliary results. We start by introducing the appropriate setting.

Let A be a finite dimensional F-algebra with basis e_i, and let V be a vector space over F. The set of multilinear antisymmetric functions from V^k to A is given by functions $F(v_1, \dots, v_k) = \sum_i F_i(v_1, \dots, v_k) e_i$ with $F_i(v_1, \dots, v_k)$ multilinear antisymmetric functions from V^k to F, in other words $F_i(v_1, \dots, v_k) \in \bigwedge^k V^*$. This space can be therefore identified with $\bigwedge^k V^* \otimes A$. Using the algebra structure of A we have a wedge product of these functions: for $F \in \bigwedge^h V^* \otimes A$, $H \in \bigwedge^k V^* \otimes A$ we define

$$F \wedge H(v_1, \dots, v_{h+k}) := \frac{1}{h!k!} \sum_{\sigma \in S_{h+k}} \epsilon_\sigma F(v_{\sigma(1)}, \dots, v_{\sigma(h)}) H(v_{\sigma(h+1)}, \dots, v_{\sigma(h+k)})$$

It is easily verified that with this product the algebra of multilinear antisymmetric functions from V to A is isomorphic to the tensor product algebra $\bigwedge V^* \otimes A$. We will apply this to $V = A = M_n$. Again the group $G = \mathrm{PGL}_n(F)$ acts on these functions and it will be of interest to study the *invariant algebra*

$$A_n := (\bigwedge M_n^* \otimes M_n)^G.$$

If N_n denotes the Lie algebra of trace zero $n \times n$ matrices, the multilinear and antisymmetric trace expressions for such matrices can be identified with the invariants $(\bigwedge N_n^*)^G$ of $\bigwedge N_n^*$ under the action of G. By a result of Chevalley transgression [**Che52**] and Dynkin [**Dyn59**] this is the exterior algebra in the elements

$$T_h := \operatorname{tr}\big(S_{2h+1}(x_1, \dots, x_{2h+1})\big), \quad 1 \le h \le n-1.$$

In this subsection we will deal with $AT_n := (\bigwedge M_n^*)^G$ rather than with $(\bigwedge N_n^*)^G$. From this result it easily follows that, with a slight abuse of notation, the former is the exterior algebra in the elements $T_0 := \operatorname{tr}(S_1(x_1)), T_1, \dots, T_{n-1}$. We remark that we use only traces of the standard polynomials of odd degree since, as it is well-known, $\operatorname{tr}(S_{2h}(x_1, \dots, x_{2h})) = 0$ for every h, see [**Ros76**].

The group G obviously acts by automorphisms, thus $A_n = (\bigwedge M_n^* \otimes M_n)^G$ is indeed an associative algebra. The main known fact that we shall use is (see, e.g., [**Kos97**] or [**Ree97**, Corollary 4.2]):

PROPOSITION 2.3.1. *The dimension of A_n over F is $n2^n$.*

Inside A_n we have the identity map X which in the natural coordinates is the *generic matrix* $\sum_{h,k} x_{hk}e_{hk}$. By iterating the definition of wedge product we have the important fact (see also [**Pro13**]):

PROPOSITION 2.3.2. *As a multilinear function, each power* $X^a := X^{\wedge a}$ *equals the standard polynomial* S_a.

As a consequence we have $S_a \wedge S_b = S_{a+b}$ and by the Amitsur-Levitzki Theorem $X^{2n} = 0$. We summarize the rules:

$$S_a = X^a,\ T_h \wedge X = -X \wedge T_h,\ X^{2n} = 0,$$

where the powers of X should be understood with respect to the wedge product.

REMARK 2.3.3. Note that the elements

$$T_{h_1} \wedge T_{h_2} \ldots \wedge T_{h_i} \wedge X^k,$$

where $h_1 < h_2 < \ldots < h_i$ and k is arbitrary, form a linear basis of the algebra of multilinear and antisymmetric expressions in noncommutative variables and their traces.

We can consider this algebra as the exterior algebra in the variables T_h, and a variable X in degree 1 which anticommutes with the T_i. We now factor out the ideal of elements of degree $> n^2$ and T_h for $h \geq n$, and thus obtain a symbolic algebra which we call $\mathcal{TA}_n$. The algebra A_n of multilinear antisymmetric invariant functions on matrices to matrices is a quotient of this algebra. We have to discover the identities that generate the corresponding ideal, as for instance the Amitsur-Levitzki identity $X^{2n} = 0$, which is the basic *even* identity. The next lemma points out the basic *odd* identity.

LEMMA 2.3.4. *The element* $O_n := nX^{2n-1} - \sum_{i=0}^{n-1} X^{2i} \wedge T_{n-i-1} \in \mathcal{TA}_n$ *is an identity of* M_n. *Moreover,* O_n *is an antisymmetric trace identity of minimal degree.*

PROOF. We know by the SFT that every trace identity is obtained from Q_n by substitution of variables and multiplication, hence any antisymmetric identity is obtained by first applying such a procedure obtaining a multilinear identity and then antisymmetrizing. Since Q_n is symmetric this procedure gives zero if we substitute two variables by two monomials of the same odd degree. In particular this means that we can keep at most one variable unchanged and we have to substitute the others with monomials of degree ≥ 2, thus the minimal identity that we can develop in this way is by substituting $x_2, \ldots, x_n$ with distinct monomials M_i, $2 \leq i \leq n$, of degree 2 and then antisymmetrizing.

We use the formula (1.2). If a permutation σ contains a cycle $(i_1, \ldots, i_k)$ in which neither 1 nor $n+1$ appear, substituting and alternating into the corresponding element $\phi_\sigma(x_1, \ldots, x_n)$, we get that the antisymmetrization of the factor $\mathrm{tr}(M_{i_1} \ldots M_{i_k})$ is zero as $\mathrm{tr}(X^{2i}) = \mathrm{tr}(S_{2i}) = 0$. Thus the only terms of (1.2) which give a contribution are the ones where either σ is a unique cycle and they contribute to $(-1)^n n! X^{2n-1}$ or the ones with two cycles, one containing 1 and the other $n+1$; such permutations can be described in the form

$$\sigma = (1, i_1, \ldots, i_h)(i_{h+1}, \ldots, i_{n-1}, n+1).$$

For each h there are exactly $(n-1)!$ of these and they all have the sign $(-1)^{n-1}$. The antisymmetrization of ϕ_σ after substitution gives

$$\mathrm{tr}(X^{2h+1})X^{2(n-h-1)} = X^{2(n-h-1)} \wedge T_h,$$

and the claim follows. $\square$

By AT_{n-1} we denote the subalgebra of AT_n generated by the $n-1$ elements T_i, $0 \leq i \leq n-2$. This is an exterior algebra and has dimension 2^{n-1}.

THEOREM 2.3.5. *$A_n = (\bigwedge M_n^* \otimes M_n)^G$ is a free left module over the algebra AT_{n-1} with basis $\{1, X, \dots, X^{2n-1}\}$. The kernel of the canonical homomorphism from $\mathcal{TA}_n$ to $(\bigwedge M_n^* \otimes M_n)^G$ is generated by X^{2n} and O_n.*

PROOF. We have that $\dim(\bigwedge M_n^* \otimes M_n)^G = n2^n$ by Proposition 2.3.1. Moreover, by Remark 2.3.3 and the FFT we know that A_n as module over AT_n is generated by the elements $1, X, \dots, X^{2n-1}$.

Now consider the left submodule N of $(\bigwedge M_n^* \otimes M_n)^G$ generated over the algebra AT_{n-1} by the elements $1, X, \dots, X^{2n-1}$. Clearly $\dim(N) \leq (2n)2^{n-1} = n2^n$ and the equality holds if and only if N is a free module. By the dimension formula this is also equivalent to say that N coincides with $(\bigwedge M_n^* \otimes M_n)^G$.

So it is enough to show that N coincides with $(\bigwedge M_n^* \otimes M_n)^G$. For this it suffices to show that N is stable under multiplication by the *missing generator* T_{n-1}. Due to the commutation relations it is enough to use the right multiplication, which is an $AT_{n-1}-$linear map.

From the identity O_n we have

$$1 \wedge T_{n-1} = T_{n-1} = -\sum_{i=1}^{n-1} X^{2i} \wedge T_{n-i-1} + nX^{2n-1},$$

hence for all $i \geq 1$ we have

$$X^j \wedge T_{n-1} = -\sum_{i=1}^{n-[\frac{j}{2}]} X^{2i+j} \wedge T_{n-i-1},$$

which gives the matrix of such multiplication in this basis as desired. $\square$

2.3.2. *Antisymmetric quasi-identities that are not a consequence of the Cayley-Hamilton identity.* We have denoted by N_n the subspace of trace zero matrices and $G = \mathrm{PGL}_n(F)$ acts on M_n and N_n by conjugation.

We work with the associative algebra $(\bigwedge N_n^* \otimes M_n)^G$ of G-equivariant antisymmetric multilinear functions from N_n to M_n. We let Y be the element of $N_n^* \otimes M_n = \hom(N_n, M_n)$ corresponding to the inclusion. We note that $Y = X - \frac{\mathrm{tr}(X)}{n}$.

It easily follows from Theorem 2.3.5 and the previous formula that also $(\bigwedge N_n^* \otimes M_n)^G$ is a free module on the powers Y^i, $0 \leq i \leq 2n-1$, over the exterior algebra in the $n-2$ generators $\mathrm{tr}(Y^{2i+1})$, $1 \leq i \leq n-2$.

Finally, we know that the element $\mathrm{tr}(Y^{2n-1})$ acts on this basis by the basic formula:

$$Y^j \wedge \mathrm{tr}(Y^{2n-1}) = -\sum_{i=1}^{n-[\frac{j}{2}]} Y^{2i+j} \wedge \mathrm{tr}(Y^{2n-2i-1}).$$

We now construct the formal algebra of symbolic expressions by adding to $(\bigwedge N_n^* \otimes M_n)^G$ a variable X with the rules

$$XY = -YX, \quad X\,\mathrm{tr}(Y^{2i+1}) = -\,\mathrm{tr}(Y^{2i+1})X.$$

We place X in degree 1 and factor out all elements of degree $> n^2$. We call this formal algebra $\tilde{A}_n$. Its connection to quasi-identities will be revealed below.

Consider now the algebra $\mathbb{F}_n := \bigwedge N_n^*[X]$ with again X in degree 1, $X^{2n} = 0$ and X anticommutes with the elements of degree 1, that is with N_n^*. We also impose that the expressions of degrees $> n^2$ are zero in $\mathbb{F}_n$. Each element of this algebra induces an antisymmetric multilinear functions from N_n to M_n and the elements that give rise to the zero function are exactly the antisymmetric multilinear quasi-identities on N_n. As above, we set $T_i = \mathrm{tr}(Y^{2i+1}) \in \bigwedge^{2i+1} N_n^*$. Let us first identify the subspace of $\mathbb{F}_n$ of quasi-identities deduced from Q_n.

PROPOSITION 2.3.6. *The space of quasi-identities deduced from Q_n in $\mathbb{F}_n$ is the ideal generated by the element*

$$O_n := nX^{2n-1} - \sum_{i=0}^{n-2} X^{2i} \wedge T_{n-i-1}.$$

PROOF. By definition a quasi-identity is deduced from Q_n if it is obtained by first substituting the variables in Q_n with monomials, and then multiplying by monomials and polynomials in the coordinates. If it is multilinear this procedure passes only through steps in which all substitutions are multilinear, as for anti-symmetrizing we can first make it multilinear then antisymmetrize. Thus we see that the quasi-identities in $\mathbb{F}_n$ deduced from Q_n equal the ideal generated by the invariant antisymmetric quasi-identities deduced from Q_n. By Theorem 2.3.5 these are multiples of O_n, proving the result. (Note that we have slightly abused the notation – since we are dealing with trace zero matrices N_n we have $T_0 = 0$, unlike in Theorem 2.3.5.) □

We set $J := O_n\mathbb{F}_n$ to be the ideal generated by the element O_n. We will concentrate on degree n^2 where we know that all formal expressions are identically zero as functions on N_n. We want to describe in the space of the quasi-identities of degree n^2, $\mathbb{F}_n[n^2]$, the subspace $J \cap \mathbb{F}_n[n^2]$ of the elements which are a consequence of Q_n.

Restricting to an isotypic component. Let us notice that the group G acts on $\mathbb{F}_n$ through its action on $\bigwedge N_n^*$ and fixing X. Namely, we have a representation

$$\mathbb{F}_n = \oplus_{i=0}^{2n-1}(\oplus_{j=0}^{n^2-i} \bigwedge^{j} N_n^*)X^i, \quad \mathbb{F}_n[n^2] = \oplus_{i=1}^{2n-1} \bigwedge^{n^2-i} N_n^* X^i. \tag{2.2}$$

We now restrict to the subspace stable under G and corresponding to the isotypic component of type N_n. This is motivated by the fact that the component of $\mathbb{F}_n[n^2]$ relative to X^2 is $\bigwedge^{n^2-2} N_n^* X^2$, which is visibly isomorphic to N_n as a representation. It is explicitly described as follows: the space $\bigwedge^{n^2-2} N_n^*$ of multilinear antisymmetric functions of n^2-2 matrix variables can be thought of as the span of the determinants of the maximal minors (of size n^2-2) of the $(n^2-2) \times (n^2-1)$ matrix whose i^{th} row are the coordinates of the i^{th} matrix variable X_i which is assumed to be of trace 0.

Let us denote by $\mathbb{G}_n[n^2]$ the isotypic component of type N_n in $\mathbb{F}_n[n^2]$, and by $\mathbb{G}_n[n^2]_{CH}$ the part of this component deducible from Q_n. We are now in a position to state our main result.

THEOREM 2.3.7. *We have a direct sum decomposition*

$$\mathbb{G}_n[n^2] = \mathbb{G}_n[n^2]_{CH} \oplus \bigwedge^{n^2-2} N_n^* X^2.$$

In particular we have the following corollary.

COROLLARY 2.3.8. *The space $\bigwedge^{n^2-2} N_n^* X^2$ consists of quasi-identities which are not a consequence of the Cayley-Hamilton identity Q_n.*

REMARK 2.3.9. Corollary 2.3.8 shows that there exist quasi-identities on N_n which are not a consequence of the Cayley-Hamilton identity Q_n. However, by substituting the variable x_k with $x_k - \frac{1}{n}\mathrm{tr}(x_k)$ we readily obtain quasi-identities on

M_n that do not follow from Q_n. This corollary therefore answers our basic question posed in the introduction.

Before engaging in the proof of Theorem 2.3.7 we need to develop some formalism. First of all recall that for a reductive group G, a representation U, and an irreducible representation N, we have a canonical isomorphism

$$\hom_G(N,U) = (N^* \otimes U)^G, \quad j : (N^* \otimes U)^G \otimes N \xrightarrow{\cong} U_N;\ j[(\phi \otimes u) \otimes n] \mapsto \langle \phi \mid n \rangle u,$$

where U_N denotes the isotypic component of type N.

We want to apply this isomorphism to $U = \mathbb{F}_n$, or $\mathbb{F}_n[n^2]$ and $N = N_n \cong N_n^*$. In particular we have to start describing $(N_n \otimes \mathbb{F}_n)^G$. In fact it is necessary to work with

$$\tilde{A}_n = (M_n \otimes \mathbb{F}_n)^G = ((F \oplus N_n) \otimes \mathbb{F}_n)^G = \mathbb{F}_n^G \oplus (N_n \otimes \mathbb{F}_n)^G. \tag{2.3}$$

We have

$$\begin{aligned} \tilde{A}_n &= \mathbb{F}_n^G \oplus (N_n \otimes \mathbb{F}_n)^G \\ &= \oplus_{i=0}^{2n-1} \oplus_{j=0}^{n^2-i} (\bigwedge^j N_n^*)^G X^i \oplus_{i=0}^{2n-1} \oplus_{j=0}^{n^2-i} (N_n \otimes \bigwedge^j N_n^*)^G X^i, \end{aligned} \tag{2.4}$$

$$\begin{aligned} (M_n \otimes \mathbb{F}_n[n^2])^G &= \mathbb{F}_n[n^2]^G \oplus (N_n \otimes \mathbb{F}_n[n^2])^G \\ &= \oplus_{i=1}^{2n-1} (\bigwedge^{n^2-i} N_n^*)^G X^i \oplus_{i=1}^{2n-1} (N_n \otimes \bigwedge^{n^2-i} N_n^*)^G X^i. \end{aligned}$$

Now $\tilde{A}_n$ is still an algebra containing $\mathbb{F}_n^G$ as a subalgebra. This is the algebra described at the beginning of this subsection, where Y denoted the generic trace zero matrix.

LEMMA 2.3.10. *We have*

$$\begin{aligned} &(\mathbb{F}_n[n^2] \cap J)^G \oplus (N_n \otimes [\mathbb{F}_n[n^2] \cap J])^G \\ &= (M_n \otimes [\mathbb{F}_n[n^2] \cap J])^G = \tilde{A}_n[n^2] \cap (1 \otimes O_n)\tilde{A}_n \end{aligned} \tag{2.5}$$

and under the isomorphism $j : (N_n \otimes \mathbb{F}_n[n^2])^G \otimes N_n \to \mathbb{G}_n[n^2]$ *the space* $\mathbb{G}_n[n^2]_{CH}$ *corresponds to* $(N_n \otimes [\mathbb{F}_n[n^2] \cap J])^G \otimes N_n$.

PROOF. We have $(M_n \otimes [\mathbb{F}_n[n^2] \cap J])^G = \tilde{A}_n[n^2] \cap (M_n \otimes J)^G$. Since O_n is G- invariant, O_n acts (by multiplication with $1 \otimes O_n$) on the space of invariants $\tilde{A}_n = (M_n \otimes \mathbb{F}_n)^G$, that is

$$(M_n \otimes J)^G = (M_n \otimes O_n \mathbb{F}_n)^G = (1 \otimes O_n)(M_n \otimes \mathbb{F}_n)^G = (1 \otimes O_n)\tilde{A}_n,$$

proving (2.5).

By definition $\mathbb{G}_n[n^2]_{CH} = \mathbb{G}_n[n^2] \cap J = \mathbb{G}_n[n^2] \cap O_n \mathbb{F}_n$. Since by definition $\mathbb{G}_n[n^2]$ is the isotypic component of type $N_n \cong N_n^*$ in $\mathbb{F}_n[n^2]$, we have $(N_n \otimes \mathbb{F}_n[n^2])^G = (N_n \otimes \mathbb{G}_n[n^2])^G$. Thus clearly

$$(N_n \otimes \mathbb{G}_n[n^2]_{CH})^G = (N_n \otimes [\mathbb{G}_n[n^2] \cap J])^G = (N_n \otimes [\mathbb{F}_n[n^2] \cap J])^G.$$

$\square$

On the other hand, the elements $T_i \in (\bigwedge^{2i+1} N_n^*)^G$ equal $\operatorname{tr}(Y^{2i+1})$, so in particular $T_{n-1} \in \mathbb{F}_n^G$ acts on $\tilde{A}_n$ as

$$T_{n-1} = nY^{2n-1} - \sum_{i=1}^{n-2} Y^{2i} \wedge T_{n-i-1}.$$

Thus, we have the following lemma.

LEMMA 2.3.11. *On $\tilde{A}_n$ the element $1 \otimes O_n$ acts by multiplying by*

$$\bar{O}_n := n(X^{2n-1} - Y^{2n-1}) - \sum_{i=1}^{n-2} (X^{2i} - Y^{2i}) \wedge T_{n-i-1}.$$

Our goal is to understand $\mathbb{G}_n[n^2]_{CH}$. On the other hand, $\mathbb{F}_n[n^2]$ consists of all quasi-identities, hence $\mathbb{F}_n[n^2]^G$ is formed of trace identities and so it is contained in J. Thus, $(\mathbb{F}_n[n^2] \cap J)^G = \mathbb{F}_n[n^2]^G$ and from (2.5) we have

$$\tilde{A}_n[n^2] \cap \bar{O}_n \tilde{A}_n = \mathbb{F}_n[n^2]^G \oplus (N_n \otimes \mathbb{G}_n[n^2]_{CH})^G. \tag{2.6}$$

In order to study the isotypic component of type N_n in $\mathbb{F}_n[n^2] \cap J$ we therefore need to analyze

$$(M_n \otimes [\mathbb{F}_n[n^2] \cap J])^G = \tilde{A}_n[n^2] \cap \bar{O}_n \tilde{A}_n = \tilde{A}_n[n^2 - 2n + 1]\bar{O}_n. \tag{2.7}$$

LEMMA 2.3.12. *We have a* monomial basis *in $\tilde{A}_n$ made of elements of the form $\mathcal{T} X^i Y^j$ where $\mathcal{T}$ is a product of some of the elements T_k, $1 \leq k \leq n-2$, written in the increasing order. Its degree is $i + j$ plus the sum of the $2k + 1$ for the T_k appearing in $\mathcal{T}$.*

PROOF. This follows from (2.4) and Theorem 2.3.5. $\square$

From (2.7) we need to understand the monomials in degree n^2 and $n^2 - 2n + 1$ which are bases of $\tilde{A}_n$, $\tilde{A}_n[n^2 - 2n + 1]$, respectively, and consider the matrix in these bases of multiplication by $\bar{O}_n$ as a map

$$\pi_n : \tilde{A}_n[n^2 - 2n + 1] \to \tilde{A}_n[n^2].$$

In order to understand the image of π_n we construct a linear function ρ on $\tilde{A}_n[n^2]$ defined on the monomial $M := \mathcal{T} X^i Y^j$ of degree n^2 as follows.

(1) If $\mathcal{T}$ does not contain at least two of the factors T_h, T_k, we set $\rho(M) = 0$.
(2) If $\mathcal{T}$ does not contain only one factor T_h, we set $\rho(M) = (-1)^{h+n}$.
(3) If $\mathcal{T}$ contains all the factors T_k and $2 \leq i, j \leq 2n - 2$ are even, we set $\rho(M) = n$; otherwise we set $\rho(M) = 0$.

We denote by $\mathcal{S}$ the ordered product of all T_k, $1 \leq k \leq n-2$, an element of degree $n^2 - 2n$.

PROPOSITION 2.3.13. *The image of π_n equals the kernel of ρ. Moreover,*

$$\tilde{A}_n[n^2] = \operatorname{im} \pi_n \oplus F\mathcal{S}X^2Y^{2n-2}.$$

PROOF. First we prove that the image of π_n is contained in the kernel of ρ.

For this take any monomial $A = \mathcal{T} X^i Y^j \in \tilde{A}_n[n^2 - 2n + 1]$ and consider $A\bar{O}_n$.

i) Firstly, if $\mathcal{T}$ misses at least 3 of the elements T_i then all the terms in $A\bar{O}_n$ miss at least 2 of the elements T_i, thus ρ is 0 on all terms.

ii) Assume $\mathcal{T}$ misses two elements T_h, T_k. The terms $A\, n(X^{2n-1} - Y^{2n-1})$ in $A\bar{O}_n$ then miss at least 2 of the elements T_i, so ρ is 0 on these terms. The remaining nonzero terms come from $B := -[A(X^{2(n-1-h)} - Y^{2(n-1-h)}) \wedge T_h + A(X^{2(n-1-k)} - Y^{2(n-1-k)}) \wedge T_k]$. Observe first that $\sum_{i=1}^{n-2} 2i + 1 = n^2 - 2n$ and so the degree of $\mathcal{T}$ is $n^2 - 2n - 2(h+k) - 2$. The degree of A is $n^2 - 2n + 1$, so that $i + j = 2h + 2k + 3$. We may assume $h > k$, $i \geq j$.

If $i + 2(n-1-h) < 2n$ (and hence $j + 2(n-1-h) < 2n$), then $-A(X^{2(n-1-h)} - Y^{2(n-1-h)}) \wedge T_h$ is the difference of two monomials on which ρ attains the same value, so on this expression ρ vanishes, same for k.

If $i + 2(n-1-h) \geq 2n$, i.e., $i \geq 2h + 2$, and hence $j \leq 2k + 1 < 2h + 1$, we have

$$B = AY^{2(n-1-h)} \wedge T_h + AY^{2(n-1-k)} \wedge T_k = A \wedge T_h Y^{2(n-1-h)} + A \wedge T_k Y^{2(n-1-k)}$$

$$= (-1)^{i+j}(\mathcal{T} \wedge T_h X^i Y^{2(n-1-h)+j} + \mathcal{T} \wedge T_k X^i Y^{2(n-1-k)+j}).$$

When we place $\mathcal{T} \wedge T_h$ in the increasing order we multiply by $(-1)^u$ where u is the number of factors of index $> h$. Since we have $n-2$ factors, $(-1)^u = (-1)^{n-2-h}$ and the value of ρ on the first term is $(-1)^{i+j}(-1)^u(-1)^{k+n} = (-1)^{i+j+h+k}$. For the second term the number of terms we have to exchange is the number of terms of index bigger than k minus 1 so we get the sign $-(-1)^{i+j+h+k}$ and the two terms cancel.

iii) Assume $\mathcal{T}$ misses only one element T_h. In this case the degree of $\mathcal{T}$ is $n^2-2n-2h-1$, thus $i+j = 2h+2$. The two terms $A\,n(X^{2n-1} - Y^{2n-1})$ are 0 unless either $i=0$ or $j=0$, since we are assuming $i \geq j$ this implies $j=0, i=2h+2$ and

$$A\,n(X^{2n-1} - Y^{2n-1}) = -n\mathcal{T} X^{2h+2} Y^{2n-1}. \tag{2.8}$$

The other contribution to the product is $-A(X^{2(n-1-h)} - Y^{2(n-1-h)}) \wedge T_h$.

If we have $2(n-1-h)+i < 2n$, then on this contribution ρ vanishes. In this case $i \neq 2h+2$, so the contribution (2.8) does not appear.

If $2(n-1-h)+i \geq 2n$, i.e. $i \geq 2h+2$, we have $i = 2h+2, j=0$. The product $A\bar{O}_n$ is 0 unless $2h+2 < 2n$, in this case it equals

$$-n\mathcal{T} X^{2h+2} Y^{2n-1} + (-1)^{n-h} \mathcal{S} X^{2h+2} Y^{2(n-1-h)}, \tag{2.9}$$

where as before we have used $\mathcal{T} \wedge T_h = (-1)^{n-h}\mathcal{S}$. By definition the value of ρ on $\mathcal{T} X^{2h+2} Y^{2n-1}$ is $(-1)^{n+h}$, as for $\mathcal{S} X^{2h+2} Y^{2(n-1-h)}$ we have $2n+2 \geq 2$ and also $2(n-1-h) \geq 2$, thus the value of ρ on it equals n. The value on the sum is therefore $-n(-1)^{n+h} + (-1)^{n-h} n = 0$.

iv) Finally we consider the case in which $\mathcal{T}$ is the product of all the T_i's. In this case $i+j=1$, and since we assume $i \geq j$ we have $i=1, j=0$. The only possible terms in the product are in $A\,n(X^{2n-1} - Y^{2n-1}) = -n\mathcal{T} X Y^{2n-1}$. By definition ρ is 0 on this term.

We now want to prove that the image of π_n coincides with the kernel of ρ. For this we have to show that the image of π_n has codimension 1. It is enough to show that adding a single vector to the image of π_n we obtain the entire space. We define V to be the space spanned by $\mathcal{S} X^2 Y^{2n-2}$ and $\mathrm{im}(\pi_n)$. We want to show that $V = \tilde{A}_n[n^2]$. In the case iv) we have already seen that $\mathcal{S} X Y^{2n-1}, \mathcal{S} X^{2n-1} Y$ belong to the image of π_n.

Claim 1. For every h we have $\mathcal{S} X^{2h+1} Y^{2(n-h)-1} \in \mathrm{im}(\pi_n)$, $\mathcal{S} X^{2h} Y^{2(n-h)} \in V$.

To prove this claim, consider T_h of degree $2h+1$. We may remove T_h from $\mathcal{S}$ obtaining a product $\mathcal{S}^{(h)}$ and take the element

$$A := \mathcal{S}^{(h)} X^{2h+1} Y \in \tilde{A}_n[n^2-2n+1].$$

We have

$$A\bar{O}_n = \pm \mathcal{S} X^{2h+1} Y (X^{2(n-h-1)} - Y^{2(n-h-1)}) = \pm(\mathcal{S} X^{2n-1} Y - \mathcal{S} X^{2h+1} Y^{2n-2h-1}).$$

Since $\mathcal{S} X^{2n-1} Y \in \mathrm{im}(\pi_n)$ we deduce $\mathcal{S} X^{2h+1} Y^{2n-2h-1} \in \mathrm{im}(\pi_n)$.

For the other case consider $A := \mathcal{S}^{(n-h)} X^2 Y^{2(n-h)} \in \tilde{A}_n[n^2-2n+1]$. Then

$$A\bar{O}_n = \pm \mathcal{S} X^2 Y^{2(n-h)} (X^{2(h-1)} - Y^{2(h-1)}) = \pm(\mathcal{S} X^{2h} Y^{2(n-h)} - \mathcal{S} X^2 Y^{2n-2}).$$

Since $\mathcal{S} X^2 Y^{2n-2} \in V$ we have $\mathcal{S} X^{2h} Y^{2(n-h)} \in V$.

Claim 2. If $\mathcal{T} X^i Y^j \in V$, also $\mathcal{T} X^j Y^i \in V$.

By definition $\mathrm{im}(\pi_n)$ is invariant under the exchange of X, Y, while V is obtained from $\mathrm{im}(\pi_n)$ by adding $\mathcal{S} X^2 Y^{2n-2}$, but by Claim 1 we also have $\mathcal{S} X^{2n-2} Y^2 \in V$. This proves Claim 2.

CLAIM 3. All monomials $\mathcal{T}X^iY^j \in \tilde{A}_n[n^2]$, where $\mathcal{T}$ misses one element T_h, are in V.

We must have $i+j = 2(n+h)+1$. Apply (2.9) and Claim 1 to deduce that $\mathcal{T}X^{2h+2}Y^{2n-1} \in V$. By Claim 2 also $\mathcal{T}X^{2n-1}Y^{2h+2}$ belongs to V. We may assume $i \geq j$ by Claim 2. It thus suffices to consider only the case $i > 2h+2$. Note that in the case $h = n-2$, we have $i+j = 4n-3$, thus $i = 2n-1, j = 2n-2$, so in this case the previous argument establishes the claim.

Consider now the case $h < n-2$. We first consider the case $i = 2n-2$. (Note that the case $i = 2n-1$ has been considered above.) Take $A := \mathcal{T}^{(n-2)}X^{2n-2}Y^{2h+1}$, where $\mathcal{T}^{(n-2)}$ denotes the element obtained from $\mathcal{T}$ by removing T_{n-2}. Then

$$A\bar{O}_n = \pm\mathcal{T}X^{2n-2}Y^{2h+3} \pm \mathcal{S}^{(n-2)}X^{2n-2}Y^{2n-1}.$$

As $\mathcal{S}^{(n-2)}X^{2n-2}Y^{2n-1}$ has already been proven to belong to V, $\mathcal{T}X^{2n-2}Y^{2h+3} \in V$. We now prove by the decreasing induction that $\mathcal{T}X^iY^{2(n+h)+1-i}$ lies in V for $i > 2h+2$. Take $A := \mathcal{T}^{(n-2)}X^iY^{2h+2n-i-1} \in \tilde{A}_n[n^2-2n+1]$. We have

$$\begin{aligned} A\bar{O}_n =& \mathcal{T}^{(n-2)}X^iY^{2h+2n-i-1}\bar{O}_n \\ =& \mathcal{T}^{(n-2)}X^iY^{2h+2n-i-1}[-(X^2-Y^2)\wedge T_{n-2} - (X^{2(n-1-h)} - Y^{2(n-1-h)})\wedge T_h] \\ =& \pm\,\mathcal{T}(X^{i+2}Y^{2h+2n-i-1} - X^iY^{2h+2n-i+1}) \in V \end{aligned}$$

in case $i < 2n-2$. Since by the induction hypothesis $\mathcal{T}X^{i+2}Y^{2h+2n-i-1} \in V$, it follows that $\mathcal{T}X^iY^{2h+2n-i+1} \in V$, and we have the desired result.

CLAIM 4. All monomials $\mathcal{T}X^iY^j \in \tilde{A}_n[n^2]$, where $\mathcal{T}$ miss $m \geq 2$ elements T_h, are in V.

Assume that $\mathcal{T}$ misses elements $T_{h_1}, \ldots, T_{h_m}$. Let us denote $s = \sum_{i=1}^m (2h_i+1)$. We first show that $\mathcal{T}X^{2n-1}Y^{s+1} \in \mathrm{im}(\pi_n)$,$\mathcal{T}X^{2n-2}Y^{s+2} \in \mathrm{im}(\pi_n)$. Since $m \geq 2$ and we can assume that $s \leq 2(2n-1)-2n = 2n-2$, $\mathcal{T}$ cannot miss $\mathcal{T}_{n-2}$. Denote by $\mathcal{T}^{(n-2)}$ the element obtained from $\mathcal{T}$ by removing T_{n-2}. As all monomials in $\mathcal{T}^{(n-2)}\bar{O}_n$ miss at least two elements T_{k_1}, T_{k_2}, they cannot miss T_{n-2} by the previous argument, thus we have $\mathcal{T}^{(n-2)}X^{2n-1}Y^{s-1}\bar{O}_n = \pm\mathcal{T}X^{2n-1}Y^{s+1}$. The same argument shows that $\mathcal{T}X^{2n-2}Y^{s+2} \in \mathrm{im}(\pi_n)$. Arguing by the decreasing induction we may assume that $\mathcal{T}X^{2n-k+2}Y^{s+k-2} \in \mathrm{im}(\pi_n)$. We have

$$\mathcal{T}^{(n-2)}X^{2n-k}Y^{s+k-2}\bar{O}_n = \pm(\mathcal{T}X^{2n-k+2}Y^{s+k-2} - \mathcal{T}X^{2n-k}Y^{s+k}) \in \mathrm{im}(\pi_n).$$

By the induction hypothesis $\mathcal{T}X^{2n-k+2}Y^{s+k-2} \in \mathrm{im}(\pi_n)$ and thus $\mathcal{T}X^{2n-k}Y^{s+k} \in \mathrm{im}(\pi_n)$. □

PROOF OF THEOREM 2.3.7. Note that $\mathcal{S}X^2Y^{2n-2}$, which is not in the image of π_n by Proposition 2.3.13, is a generator of the 1-dimensional space $(N_n \otimes \bigwedge^{n^2-2} N_n^* X^2)^G \subset (N_n \otimes \mathbb{F}_n[n^2])^G$. The decomposition of $A_n[n^2]$ from Proposition 2.3.13 thus induces the decomposition $\mathbb{G}_n[n^2] = \mathbb{G}_n[n^2]_{CH} \oplus \bigwedge^{n^2-2} N_n^* X^2$. □

2.4. Quasi-identities that follow from the Cayley-Hamilton identity. Here we collect several results on quasi-identities of matrices and the Cayley-Hamilton identity, and give a positive solution for the Specht problem for quasi-identities of matrices.

2.4.1. *Quasi-identities and local linear dependence.* Let A be an F-algebra. Noncommutative polynomials $f_1, \ldots, f_t \in F\langle x_1, \ldots, x_m\rangle$ are said to be A-*locally linearly dependent* if the elements $f_1(a_1, \ldots, a_m), \ldots, f_t(a_1, \ldots, a_m)$ are linearly dependent in A for all $a_1, \ldots, a_m \in A$. This concept has actually appeared in Operator Theory **[CHSY03]**, and was recently studied from the algebraic point of view in **[BK13]**. We will see that it can be used in the study of quasi-identities.

Recall that C_m stands for the Capelli polynomial. The following well-known result (see, e.g., [**Row80**, Theorem 7.6.16]) was used in [**BK13**] as a basic tool.

THEOREM 2.4.1. *Let A be a prime algebra. Then $a_1, \dots, a_t \in A$ are linearly dependent over the extended centroid of A if and only if $C_{2t-1}(a_1, ..., a_t, r_1, ..., r_{t-1}) = 0$ for all $r_1, \dots, r_{t-1} \in A$.*

By using a similar approach as in the proof of [**BK13**, Theorem 3.1], just by applying Theorem 2.4.1 to the algebra of generic matrices instead of to the free algebra $F\langle X\rangle$, we get the following characterization of M_n-local linear dependence through the central polynomials.

THEOREM 2.4.2. *Noncommutative polynomials $f_1, \dots, f_t$ are M_n-locally linearly dependent if and only if there exist central polynomials $c_1, \dots, c_t$, not all polynomial identities, such that $\sum_{i=1}^t c_i f_i$ is a polynomial identity of M_n.*

PROOF. By Theorem 2.4.1, the condition that $f_1, \dots, f_t$ are M_n-locally linearly dependent is equivalent to the condition that

$$h := C_{2t-1}(f_1, \dots, f_t, y_1, \dots, y_{t-1})$$

is a polynomial identity of M_n. Since M_n and the algebra $F\langle \xi_k\rangle$ of $n \times n$ generic matrices satisfy the same polynomial identities, this is the same as saying that h is a polynomial identity of $F\langle \xi_k\rangle$. Using Theorem 2.4.1 once again we see that this is further equivalent to the condition that $f_1, \dots, f_t$, viewed as elements of $F\langle \xi_k\rangle$, are linearly dependent over the extended centroid of $F\langle \xi_k\rangle$. Since $F\langle \xi_k\rangle$ is a prime PI-algebra, its extended centroid is the field of fractions of the center of $F\langle \xi_k\rangle$; the latter can be identified with central polynomials, and hence the desired conclusion follows. □

COROLLARY 2.4.3. *If noncommutative polynomials $f_0, f_1, \dots, f_t$ are M_n-locally linearly dependent, while $f_1, \dots, f_t$ are M_n-locally linearly independent, then there exist central polynomials $c_0, c_1, \dots, c_t$, such that c_0 is nontrivial and $\sum_{i=0}^t c_i f_i$ is a polynomial identity of M_n.*

Later, in Remark 2.4.6, we will show that this result can be used to obtain an alternative proof of a somewhat weaker version of Theorem 2.2.7.

LEMMA 2.4.4. *If a quasi-polynomial $\sum_{i=1}^t \lambda_i M_i$ is a quasi-identity of M_n, then either each $\lambda_i = 0$ or $M_1, \dots, M_t$ are M_n-locally linearly dependent (and hence satisfy the conclusion of Theorem 2.4.2).*

PROOF. We may assume that $\lambda_i = \lambda_i(x_1, \dots, x_m)$ and $M_i = M_i(x_1, \dots, x_m)$. The set W of all m-tuples $(a_1, \dots, a_m) \in M_n^m$ such that $\lambda_i(a_1, \dots, a_m) = 0$ for all $i = 1, \dots, t$ is closed in the Zariski topology of $F^{n^2 m}$. Similarly, the set Z of all m-tuples $(a_1, \dots, a_m) \in M_n^m$ such that $M_1(a_1, \dots, a_m), \dots, M_t(a_1, \dots, a_m)$ are linearly dependent is also closed - namely, the linear dependence can be expressed through zeros of a polynomial by Theorem 2.4.1. If $Z = M_n^m$, then $M_1, \dots, M_t$ are M_n-locally linearly dependent. Assume therefore that $Z \neq M_n^m$. Suppose that $W \neq M_n^m$. Then, since $F^{n^2 m}$ is irreducible (as $\mathrm{char}(F) = 0$), the complements of W and Z in M_n^m have a nonempty intersection. This means that there exist $a_1, \dots, a_m \in M_n^m$ such that $\lambda_i(a_1, \dots, a_m) \neq 0$ for some i and $M_1(a_1, \dots, a_m), \dots, M_t(a_1, \dots, a_m)$ are linearly independent. However, this is impossible since $\sum_{i=1}^t \lambda_i M_i$ is a quasi-identity. Thus, $W = M_n^m$, i.e., each $\lambda_i = 0$. □

We conclude this subsection with a theorem giving a condition under which a quasi-identity is a consequence of the Cayley-Hamilton identity.

THEOREM 2.4.5. *Let* $P = \lambda_0 M_0 + \sum_{i=1}^t \lambda_i M_i \in \mathfrak{I}_n$. *If* $M_1, \ldots, M_t$ *are* M_n-*locally linearly independent, then* P *is a consequence of the Cayley-Hamilton identity.*

PROOF. We may assume that some $\lambda_i \neq 0$, and so $M_0, M_1, \ldots, M_t$ are M_n-locally linearly dependent by Lemma 2.4.4. Theorem 2.4.2 tells us that there exist central polynomials c_0, c_i, not all trivial, such that $c_0 M_0 + \sum_{i=1}^t c_i M_i$ is a polynomial identity. Multiplying this identity with λ_0 and then comparing it with the quasi-identity $c_0 P$ it follows that $\sum_{i=1}^t (c_0\lambda_i - c_i\lambda_0) M_i \in \mathfrak{I}_n$. Lemma 2.4.4 implies that $c_0\lambda_i = c_i\lambda_0$ for every i. Let us write $\lambda_i = \lambda_i^G \lambda_i'$ where λ_i^G is the product of all irreducible factors of λ_i that are invariant under the conjugation by $G = \mathrm{GL}_n(F)$, and λ_i' is the product of the remaining irreducible factors of λ_i. (Here we assume that the first nonzero coefficient of λ_i' in some order of $\mathcal{C}$ is 1.) Hence $c_i \lambda_0^G \lambda_0' = c_0 \lambda_i^G \lambda_i'$ and therefore $\lambda_0' = \lambda_i'$ for every i. We thus have

$$P = \lambda_0' \big(\lambda_0^G M + \sum_{i=1}^t \lambda_i^G M_i \big).$$

Since λ_i^G are invariant under G, they are trace polynomials. Therefore the desired conclusion follows from the SFT. $\square$

REMARK 2.4.6. Theorem 2.2.7 in particular tells us that for every quasi-identity P of M_n there exists a nontrivial central polynomial c with zero constant term such that cP is a consequence of the Cayley-Hamilton identity. Let us give an alternative proof of that, based on local linear dependence and the SFT.

We first remark that the condition that $c \in F\langle X \rangle$ is a central polynomial can be expressed as that there exists $\alpha_c \in \mathcal{C}$ such that $c - \alpha_c \in \mathfrak{I}_n$. Actually, $c - \alpha_c$ is a trace identity since $\alpha_c = \frac{1}{n}\mathrm{tr}(c)$. Therefore the SFT implies that for every central polynomial c of M_n there exists $\alpha_c \in \mathcal{C}$ such that $c - \alpha_c$ is a quasi-identity of M_n contained in the T-ideal of $\mathcal{C}\langle X \rangle$ generated by Q_n.

Now take an arbitrary $P \in \mathfrak{I}_n$, and let us prove that c with the aforementioned property exists. Write $P = \sum_{i=1}^{n^2} \lambda_i x_i + \sum \lambda_M M$ where each M in the second summation is different from $x_1, \ldots, x_{n^2}$. We proceed by induction on the number of summands d in the second summation. If $d = 0$, then $P = 0$ by Lemma 2.4.4 (since $x_1, \ldots, x_{n^2}$ are M_n-locally linearly independent). Let $d > 0$. Pick M_0 such that $M_0 \notin \{x_1, \ldots, x_{n^2}\}$ and $\lambda_{M_0} \neq 0$. Note that $M_0, x_1, \ldots, x_{n^2}$ are M_n-locally linearly dependent, while $x_1, \ldots, x_{n^2}$ are M_n-locally linearly independent. Thus, by Corollary 2.4.3 there exist central polynomials $c_0, c_1, \ldots, c_{n^2}$ such that c_0 is nontrivial and $f := c_0 M_0 + \sum_{i=1}^{n^2} c_i x_i$ is an identity of M_n. Let $\alpha_i \in \mathcal{C}$, $i = 0, 1, \ldots, n^2$, be such that $c_i - \alpha_i$ is a quasi-identity lying in the T-ideal generated by Q_n. Let us define $P' := \alpha_0 P - \lambda_{M_0}\alpha_0 M_0 - \lambda_{M_0} \sum_{i=1}^{n^2} \alpha_i x_i$. Writing each α_i as $c_i - (c_i - \alpha_i)$ we see that P' is a quasi-identity. Note that P' involves $d-1$ summands not lying in $\mathcal{C}x_i$, $i = 1, \ldots, n^2$. Therefore the induction assumption yields the existence of a nonzero central polynomial c' such that $c'P'$ lies in the T-ideal generated by Q_n. Setting $c = c_0 c'$ we thus have $c \neq 0$ and

$$\begin{aligned} cP =& (c_0 - \alpha_0)c'P + \alpha_0 c'P = (c_0 - \alpha_0)c'P + c'P' + \lambda_{M_0} c' \big(\alpha_0 M_0 + \sum_{i=1}^{n^2} \alpha_i x_i \big) \\ =& (c_0 - \alpha_0)c'P + c'P' - \lambda_{M_0} c' \big((c_0 - \alpha_0) M_0 + \sum_{i=1}^{n^2} (c_i - \alpha_i) x_i \big) + \lambda_{M_0} c' f. \end{aligned}$$

The T-ideal generated by Q_n contains $c_i - \alpha_i$, $0 \leq i \leq n^2$, $c'P'$, as well as f according to the SFT. Hence it also contains cP.

2.4.2. *Some special cases.* The purpose of this subsection is to examine several simple situations, which should in particular serve as an evidence of the delicacy of the problem of finding quasi-identities that are not a consequence of the Cayley-Hamilton identity.

We begin with quasi-polynomials P of one variable, i.e., $P = \sum_{i=0}^{m} \lambda_i(x)x^i$. Here we could refer to results on more general functional identities of one variable in Section 4, but an independent treatment is very simple.

PROPOSITION 2.4.7. *If a quasi-polynomial of one variable p is a quasi-identity of M_n, then there exists a quasi-polynomial $r(x)$ such that $p(x) = r(x)q_n(x)$.*

PROOF. Let $p(x) = \sum_{i=0}^{m} \lambda_i(x)x^i$. The proof is by induction on m. Since $1, x, \ldots, x^{n-1}$ are M_n-locally linearly independent, we may assume that $m \geq n$ by Lemma 2.4.4. Note that $p(x) - \lambda_m(x)x^{m-n}q_n(x)$ is a quasi-identity for which the induction assumption is applicable. Therefore $p(x) - \lambda_m(x)x^{m-n}q_n(x) = r_1(x)q_n(x)$ for some $r_1(x)$, and hence $p(x) = \big(\lambda_m(x)x^{m-n} + r_1(x)\big)q_n(x)$. $\square$

What about quasi-polynomials of two variables? At least for 2×2 matrices, the answer comes easily.

PROPOSITION 2.4.8. *If a quasi-polynomial of two variables $P = P(x, y)$ is a quasi-identity of M_2, then P is a consequence of the Cayley-Hamilton identity.*

PROOF. It is an easy to see that any quasi-polynomial of two variables $P = P(x, y)$ can be written as $P = \lambda_0 + \lambda_1 x + \lambda_2 y + \lambda_3 xy + R$ where R lies in the T-ideal generated by Q_2. Thus, if P is a quasi-identity, then each $\lambda_i = 0$ by Lemma 2.4.4, so that $P = R$ is a consequence of the Cayley-Hamilton identity. $\square$

Now one may wonder what should be the degree of a quasi-identity that may not follow from the Cayley-Hamilton identity. We first record a simple result which is a byproduct of the general theory of functional identities.

PROPOSITION 2.4.9. *If $\sum \lambda_M M$ is a quasi-identity of M_n and $\deg(\lambda_M) + \deg(M) < n$ for every M, then each $\lambda_M = 0$.*

PROOF. Apply, for example, [**BCM07**, Corollary 2.23, Lemma 4.4]. $\square$

The question arises what can be said about quasi-identities of degree n. A basic quasi-identity of degree n is the Cayley-Hamilton polynomial Q_n. In the next proposition we show that is essentially the only one.

PROPOSITION 2.4.10. *Every multilinear quasi-identity of M_n of degree n is a scalar multiple of Q_n.*

PROOF. Let $S_{n,k}$, $1 \leq k \leq n$, denote the set of all permutations $\sigma \in S_n$ such that $\sigma(1) < \sigma(2) < \cdots < \sigma(k)$ and $\sigma(k+1) < \sigma(k)$. For convenience we also set $S_{n,0} = S_n$. Note that a multilinear quasi-polynomial P of degree n can be written as

$$P(x_1, \ldots, x_n) = \sum_{k=0}^{n} \sum_{\sigma \in S_{n,k}} \lambda_{k\sigma}(x_{\sigma(1)}, \ldots, x_{\sigma(k)}) x_{\sigma(k+1)} \cdots x_{\sigma(n)}$$

(here, $\lambda_{0\sigma}$ are scalars). By e_{ij} we denote the matrix units in M_n.

We assume that P is a quasi-identity, and proceed by a series of claims.

CLAIM 1. For all $\sigma \in S_{n,k}$, $1 \leq k \leq n$, and all distinct $1 \leq i_1, \ldots, i_k \leq n$, we have

$$\lambda_{k\sigma}(e_{11}, \ldots, e_{kk}) = \lambda_{k\sigma}(e_{i_1 i_1}, \ldots, e_{i_k i_k}) = -\lambda_{k-1,\sigma}(e_{11}, \ldots, e_{k-1,k-1}).$$

The proof is by induction on k. First, take $0 \le j \le n-1$ and substitute

$$e_{1+j,1+j}, e_{1+j,2+j}, e_{2+j,3+j}, \dots, e_{n-1+j,n+j}$$

(with addition modulo n) for $x_{\sigma(1)}, \dots, x_{\sigma(n)}$ in P. Considering the coefficient at $e_{1+j,n+j}$ we get $\lambda_{0\sigma} + \lambda_{1\sigma}(e_{1+j,1+j}) = 0$. This implies the truth of Claim 1 for $k = 1$. Let $k > 1$ and take $\sigma \in S_{n,k}$. Choose a subset of $\{1, \dots, n\}$ with $k-1$ elements, $\{i_{n-k+2}, \dots, i_n\}$, and let $\{j_1, \dots, j_{n-k+1}\}$ be its complement. Let us substitute

$$e_{i_{n-k+2},i_{n-k+2}}, \dots, e_{i_n,i_n},\ e_{j_1,j_1}, e_{j_1,j_2}, e_{j_2,j_3}, \dots, e_{j_{n-k},j_{n-k+1}}$$

for $x_{\sigma(1)}, \dots, x_{\sigma(n)}$, respectively, in P. Similarly as above, this time by considering the coefficient at $e_{j_1,j_{n-k+1}}$, we obtain

$$\lambda_{k-1,\sigma}(e_{i_{n-k+2},i_{n-k+2}}, \dots, e_{i_n,i_n}) + \lambda_{k\sigma}(e_{i_{n-k+2},i_{n-k+2}}, \dots, e_{i_n,i_n}, e_{j_1,j_1}) = 0.$$

The desired conclusion follows from the induction hypothesis.

Claim 2. For all $\sigma, \tau \in S_{n,k}$, $0 \le k \le n-1$, and all distinct $1 \le i_1, \dots, i_k \le n$, we have

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = \lambda_{k\tau}(e_{i_1 i_1}, \dots, e_{i_k i_k}).$$

Evaluating P at $e_{11}, \dots, e_{nn}$ results in

$$\lambda_{n-1,\sigma_i}(e_{11}, \dots, e_{i-1,i-1}, e_{i+1,i+1}, \dots, e_{nn}) = \lambda_{n-1,\mathrm{id}}(e_{11}, \dots, e_{n-1,n-1})$$

for all $1 \le i \le n-1$, where σ_i stands for the cycle $(i\ i+1\ \dots\ n)$. Accordingly, since $S_{n,n-1}$ consists of id and all σ_i, $1 \le i \le n-1$, the case $k = n-1$ follows from Claim 1. We may now assume that $k < n-1$ and that Claim 2 holds for $k+1$. Take $\sigma \in S_{n,k}$. If $\sigma \in S_{n,k+1}$ then

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = -\lambda_{k+1,\sigma}(e_{11}, \dots, e_{k+1,k+1})$$

by Claim 1. If $\sigma \notin S_{n,k+1}$ there exists $1 \le i \le k$ such that $\sigma(k+1) < \sigma(i)$. Substituting

$$e_{11}, \dots, e_{kk}, e_{k+1,k+1}, e_{k+1,k+2}, e_{k+2,k+3}, \dots, e_{n-1,n}$$

for $x_{\sigma(1)}, \dots, x_{\sigma(n)}$ in P we infer that for a certain permutation τ (specifically, $\tau = \sigma \circ (k+1\ k\ \dots\ i+1\ i)$) we have

$$\lambda_{k\sigma}(e_{11}, \dots, e_{kk}) = -\lambda_{k+1,\tau}(e_{11}, \dots, e_{i-1,i-1}, e_{k+1,k+1}, e_{i+1,i+1}, \dots, e_{k-1,k-1}).$$

Since every $\lambda_{k\sigma}(e_{11}, \dots, e_{kk})$ is associated to an evaluation of $\lambda_{k+1,\tau}$, Claim 2 follows by the induction hypothesis and Claim 1.

Claim 3. $P = \lambda_{0,id} Q_n$.

By Claim 2 we have $\lambda_{0\sigma} = \lambda_{0\tau}$ for all $\sigma, \tau \in S_n$. Accordingly, $R := P - \lambda_{0,id} Q_n$ does not involve summands of the form $\mu x_{\sigma(1)} \dots x_{\sigma(n)}$, $\mu \in F$, and can be therefore written as

$$R(x_1, \dots, x_n) = \sum_{k=1}^{n} \sum_{\sigma \in S_{n,k}} \mu_{k\sigma}(x_{\sigma(1)}, \dots, x_{\sigma(k)}) x_{\sigma(k+1)} \cdots x_{\sigma(n)}.$$

We must prove that $R = 0$, i.e., each $\mu_{k\sigma} = 0$. We proceed by induction on k. For $k = 0$ this holds by the hypothesis, so let $k > 0$. It suffices to show that $\mu_{k\sigma}(e_{i_1 j_1}, \dots, e_{i_k j_k}) = 0$ for arbitrary matrix units $e_{i_1 j_1}, \dots, e_{i_k j_k}$. Choose distinct $l_1, \dots, l_{n-k}$ such that $l_s \ne i_t$ for all s, t. Substitute

$$e_{i_1 j_1}, \dots, e_{i_k j_k}, e_{l_1 l_2}, e_{l_2 l_3}, \dots, e_{l_{n-k} i_1}$$

for $x_{\sigma(1)}, \dots, x_{\sigma(n)}$ in P. There is only one way to factorize $e_{l_1 i_1}$ as a product of at most $n-k$ chosen matrix units, i.e., $e_{l_1 i_1} = e_{l_1 l_2} e_{l_2 l_3} \cdots e_{l_{n-k} i_1}$. By induction hypothesis it thus follows that $\mu_{k\sigma}(e_{i_1 j_1}, \dots, e_{i_k j_k}) = 0$. $\square$

2.5. Specht problem for quasi-identities. In this subsection we finally prove a version of Specht problem for quasi-identities, i.e., we show that $\mathfrak{I}_n$ is finitely generated as a T-ideal. As it is well-known, Kemer [**Kem87**] has shown that for polynomial identities such a question has a positive answer (in characteristic 0). In our case the answer is also positive since we can apply the classical method of *primary covariants* of Capelli and Deruyts, cf. [**Pro07**, Chapter 3], to which we also refer for the statements used in the proof.

THEOREM 2.5.1. *The ideal $\mathfrak{I}_n$ of all quasi-identities of M_n is finitely generated, as a T-ideal, by elements which depend on at most $2n^2$ variables.*

PROOF. First of all, if we impose the Cayley-Hamilton identity, we are reduced to study the problem for the space $\mathcal{C}\langle X\rangle/(Q_n)$ isomorphic to ker $\pi \subset \mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{T}_n\langle \xi_k\rangle \subset M_n(\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y)$. Instead of considering all possible we consider only linear substitutions of variables, that is, we consider all these spaces as representations of the infinite linear group.

Now we use the language of symmetric algebras; the space $\mathcal{C}_x$ equals $S[M_n^* \otimes V]$ where $V = \oplus_{i=1}^{\infty} Fe_i$ is an infinite dimensional vector space over which the infinite linear group G_∞ acts.

By Cauchy's formula we have

$$S[M_n^* \otimes V] = \oplus_\lambda S_\lambda(M_n^*) \otimes S_\lambda(V),$$

where λ runs over all partitions with at most n^2 columns and $S_\lambda(V)$ is the corresponding Schur functor. By representation theory the tensor product $S_\lambda(V) \otimes S_\mu(V)$ of two such representations is a sum of representations $S_\gamma(V)$ where γ runs over partitions with at most $2n^2$ columns.

Hence we have that $\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y$, which is a quotient of $\mathcal{C}_x \otimes_F \mathcal{C}_y$, is a sum of representations $S_\lambda(M_n^*) \otimes S_\mu(V)$ where μ has at most $2n^2$ columns. Now any representation $S_\gamma(V)$ is irreducible under G_∞ and generated by a highest weight vector. If γ has k columns such a highest weight vector on the other hand lies in $S_\gamma(\oplus_{i=1}^{k} Fe_i)$.

This means that under linear substitution of variables any element in $M_n(\mathcal{C}_x \otimes_{\mathcal{T}_n} \mathcal{C}_y)$, and hence also in ker π, is obtained from elements depending on at most $2n^2$ variables.

Finally, if we restrict the number of variables to a finite number m, the corresponding space ker π_m is a finitely generated module over a finitely generated algebra, and the claim follows. □

3. Functional identities

In this section we consider general functional identities on $M_n(F)$. We show that unlike the quasi-identities they follow from the Cayley-Hamilton identity. We start with a remark that generalized polynomial identities on $M_n(F)$ also share this property.

3.1. Generalized polynomial identities. Let us recall the necessary definitions on generalized polynomial identities (more details can be found in [**BMM96**]). The elements of the free product $M_n(F) * F\langle X\rangle$ over F are sometimes called *generalized polynomials.* Informally they can be viewed as sums of expressions of the form $a_{i_0}x_{j_1}a_{i_1}\dots a_{i_{k-1}}x_{j_k}a_{i_k}$ where $a_{i_\ell} \in M_n(F)$. Given $f = f(x_1,\dots,x_k) \in M_n(F) * F\langle X\rangle$ and $b_1,\dots,b_k \in M_n(F)$, we define the evaluation $f(b_1,\dots,b_k)$ in the obvious way. We say that f is a *generalized polynomial identity* (GPI) of $M_n(F)$ if $f(b_1,\dots,b_k) = 0$ for all $b_i \in M_n(F)$. For example, if e is a rank one idempotent in $M_n(F)$, then $[ex_1e, ex_2e]$ is readily a GPI.

An ideal I of $M_n(F) * F\langle X\rangle$ is a T-ideal if $f(x_1,\dots,x_k) \in I$ and $g_1,\dots,g_k \in M_n(F) * F\langle X\rangle$ implies $f(g_1,\dots,g_k) \in I$. The set of all GPI's of $M_n(F)$ is clearly a T-ideal. The next proposition was obtained, in some form, already in [**Lit31**], and later extended to considerably more general rings by Beidar (see e.g. [**BMM96**]). We will give a short alternative proof. Let us first recall the following elementary fact: If an algebra B contains a set of $n \times n$ matrix units; i.e., a set of elements e_{ij}, $1 \le i,j \le n$, satisfying

$$e_{ij}e_{kl} = \delta_{jk}e_{il}, \quad \sum_{i=1}^{n} e_{ii} = 1,$$

then B is isomorphic to the matrix algebra $M_n(A)$ where $A = e_{11}Be_{11}$.

PROPOSITION 3.1.1 ([**BMM96**, Proposition 6.5.5]). *Let e be a rank one idempotent in $M_n(F)$.*

(1) *If $|F| = \infty$, then the T-ideal of all GPI's of $M_n(F)$ is generated by $[ex_1e, ex_2e]$.*
(2) *If $|F| = q$, then the T-ideal of all GPI's of $M_n(F)$ is generated by $[ex_1e, ex_2e]$ and $(ex_1e)^q - ex_1e$.*

PROOF. Pick a set of matrix units e_{ij}, $1 \le i,j \le n$, of $M_n(F)$ so that $e_{11} = e$. As $M_n(F) * F\langle X\rangle$ contains $M_n(F)$ as a subalgebra, it also contains all e_{ij}. Accordingly, $M_n(F) * F\langle X\rangle \cong M_n(A)$ where $A = e_{11}\big(M_n(F) * F\langle X\rangle\big)e_{11}$. Since $M_n(F) * F\langle X\rangle$ is generated by the elements

$$e_{si}x_ke_{jt} = e_{s1}(e_{1i}x_ke_{j1})e_{1t},$$

it follows that A is generated by the elements

$$x_{ij}^{(k)} := e_{1i}x_ke_{j1}, \quad 1 \le i,j \le n, \ \ k = 1,2,\dots$$

Note that A is actually the free algebra on the set $\overline{x} := \{x_{ij}^{(k)} \,|\, 1 \le i,j \le n, \ k = 1,2,\dots\}$.

Let $\mathrm{Hom}_{M_n}(M_n(F) * F\langle X\rangle, M_n(F))$ denote the set of algebra homomorphisms from $M_n(F) * F\langle X\rangle$ to $M_n(F)$ that act as the identity on $M_n(F)$. Identifying $M_n(F) * F\langle X\rangle$ with $M_n(A)$ one easily sees that there is a canonical isomorphism

$$\mathrm{Hom}_{M_n}(M_n(F) * F\langle X\rangle, M_n(F)) \cong \mathrm{Hom}(A,F).$$

Now take a GPI f of $M_n(F)$. This means that

$$f \in \cap_{\phi\in \mathrm{Hom}_{M_n}(M_n(F)*F\langle X\rangle, M_n(F))} \ker\phi,$$

or equivalently,

$$e_{1i}fe_{j1} \in \cap_{\phi\in\mathrm{Hom}(A,F)} \ker\phi$$

for every $1 \le i,j \le n$. Since A is the free algebra on $\overline{x}$ it can be easily shown that $\cap_{\phi\in\mathrm{Hom}(A,F)} \ker\phi$ is generated by

$$[x_{ij}^{(k)}, x_{pq}^{(l)}] = [ex_{ij}^{(k)}e, ex_{pq}^{(l)}e]$$

and if $|F| = q$ additionally by

$$\left(x_{ij}^{(k)}\right)^q - x_{ij}^{(k)} = \left(ex_{ij}^{(k)}e\right)^q - ex_{ij}^{(k)}e.$$

Hence $f = \sum_{i,j} e_{i1}(e_{1i}fe_{j1})e_{1j}$ lies in the T-ideal generated by $[ex_1e, ex_2e]$, and additionally by $(ex_1e)^q - ex_1e$ if $|F| = q$. $\square$

Let us recall that

$$q_n(x_1) = x_1^n + \tau_1(x_1)x_1^{n-1} + \cdots + \tau_n(x_1)$$

denotes the Cayley-Hamilton polynomial and $Q_n = Q_n(x_1, \dots, x_n)$ denotes the multilinear version of $q_n(x_1)$ obtained by full polarization. Regarding $\mathrm{tr}(x_1 \dots x_k)$ as $\sum_{i,j} e_{ij}x_1 \dots x_k e_{ji}$, we may consider Q_n as a GPI of $M_n(F)$.

COROLLARY 3.1.2. *If* $\mathrm{char}(F) = 0$*, then the* T*-ideal of all GPI's of* $M_n(F)$ *is generated by the Cayley-Hamilton identity* Q_n*.*

PROOF. It suffices to show that the basic identity

$$[ex_1e, ex_2e] = ex_1ex_2e - ex_2ex_1e \in M_n(F) * F\langle X\rangle,$$

where e is a rank one idempotent, follows from the Cayley-Hamilton identity. To this end we insert $[ex_1e, ex_2e]$ for the first variable and e for the others in Q_n. Note that 1 needs to be in the last cycle of σ for $\phi_\sigma([ex_1e, ex_2e], e, \dots, e)$ to be nonzero. In this case $\phi_\sigma([ex_1e, ex_2e], e, \dots, e) = [ex_1e, ex_2e]$. Thus, we need to count the number of such permutations with the corresponding signs.

Take a cycle τ. Note that permutations with the corresponding signs having the last cycle τ do not sum to zero only if τ is of length n or $n+1$. For τ of length n we have $(n-1)(n-1)!$ permutations of sign $(-1)^{n-1}$, while for τ of length $n+1$ the number of permutation is $n!$ and all have sign $(-1)^n$. Hence,

$$Q_n\big([ex_1e, ex_2e], e, \dots, e\big) = (-1)^n(n-1)![ex_1e, ex_2e].$$

□

3.2. One-sided functional identities and syzygies. One-sided functional identities are intimately connected with syzygies on generic matrices. We first recall a result on syzygies that will be used in the sequel.

3.2.1. *Syzygies on a generic matrix.* First we introduce some auxiliary notation. Let $r, s \geq 1$, let $\mathcal{C}_y = F[y_{ij} \mid 1 \leq i \leq r, 1 \leq j \leq s]$ be a polynomial algebra, and let $u_1, \dots, u_r$ be generators of the free module $\mathcal{C}_y^n$. Let $\eta = (y_{ij})_{1\leq i\leq r, 1\leq j\leq s}$ be a generic $r \times s$-matrix. With a syzygy on η we mean a syzygy on the rows of η; i.e., an element $\sum_{i=1}^r f_iu_i \in \mathcal{C}_y^r$ with the property $\sum_{i=1}^r f_iy_{ij} = 0$ for $j = 1, \dots, s$. If $s \leq r$ we denote by $[j_1, \dots, j_s]$, where $1 \leq j_1 < \cdots < j_s \leq r$, the determinant of the $s \times s$-submatrix of η, obtained by restricting η to the rows indexed by $j_1, \dots, j_s$.

THEOREM 3.2.1 ([**Onn94**, Theorem 7.2]). *Let* $1 \leq s < r$ *and let* η *be a generic* $r \times s$*-matrix. The set of determinantal relations*

$$G = \left\{ \sum_{\ell=0}^{s} (-1)^\ell [j_0, \dots, \hat{j}_\ell, \dots, j_s] u_{j_\ell} \;\middle|\; 1 \leq j_0 < j_1 < \cdots < j_s \leq r \right\}$$

generates the module of syzygies on η *and is a Gröbner basis for it with respect to any lexicographic monomial order on* $\mathcal{C}_y^r$*.*

3.2.2. *Solving left-sided functional identities.* We restrict ourselves to the *left-sided functional identities*; i.e., functional identities of the form $\sum_{k\in K} F_k(\overline{x}_m^k)x_k = 0$. The right-sided functional identities $\sum_{l\in L} x_lG_l(\overline{x}_m^l) = 0$ can be of course treated in the same way. Besides, by applying the transpose operation to a right-sided functional identity we obtain a left-sided functional identity, so that the results that we will obtain are more or less directly applicable to right-sided functional identities.

The standard solution of $\sum_{k\in K} F_k(\overline{x}_m^k)x_k = 0$ is defined as $F_k = 0$ for each $k \in K$. If $|K| \leq n$, then there are no other solutions on $M_n(F)$ – this is an easy consequence of the general theory of functional identities (see e.g. [**BCM07**, Corollary 2.23]). To obtain a nonstandard solution we have to modify the Cayley-Hamilton identity. We take only its noncentral part and commute it with the product of a fixed matrix and a new variable. In this way we arrive at a basic functional identity on $M_n(F)$ in $n+1$ variables

$$\big[\tilde{Q}_n(a_1x_1,\dots,a_nx_n),a_{n+1}x_{n+1}\big]=0, \tag{3.1}$$

where $\tilde{Q}_n$ denotes the noncentral part of Q_n, i.e.,

$$\tilde{Q}_n(x_1,\dots,x_n)=\sum_{\sigma\in S_{n+1}\setminus S_n}\epsilon_\sigma\phi_\sigma(x_1,\dots,x_n),$$

and $a_i\in M_n(F)$. Note that (3.1) can be indeed interpreted as a left-sided functional identity. For example, in the case $n=2$ we have

$$\tilde{Q}_2(x_1,x_2)=x_1x_2+x_2x_1-\mathrm{tr}(x_1)x_2-\mathrm{tr}(x_2)x_1,$$

so that

$$\begin{aligned}\big[\tilde{Q}_2(a_1x_1,a_2x_2),a_3x_3\big]&=\Big(a_3x_3(-a_2x_2+\mathrm{tr}(a_2x_2))a_1\Big)x_1\\&\quad+\Big(a_3x_3(-a_1x_1+\mathrm{tr}(a_1x_1))a_2\Big)x_2\\&+\Big(\big(a_1x_1a_2x_2+a_2x_2a_1x_1-\mathrm{tr}(a_1x_1)a_2x_2-\mathrm{tr}(a_2x_2)a_1x_1\big)a_3\Big)x_3=0.\end{aligned}$$

If we multiply (3.1) by a scalar-valued function $\lambda(x_{n+2},\dots,x_m)$, we get a left-sided functional identity of $M_n(F)$ in m variables for an arbitrary $m>n$. Of course, by permuting the variables x_i we get further examples. We will see that in fact every left-sided functional identity of $M_n(F)$ is a sum of left-sided functional identities of such a type.

Before giving a full description of the unknown functions F_k satisfying

$$\sum_{k\in K}F_k(\overline{x}_m^k)x_k=0$$

we need to introduce some more notation. Let

$$\mathcal{C}=F\big[x_{ij}^{(k)}\mid 1\le i,j\le n,1\le k\le m\big].$$

We denote by

$$[(i_1,j_1),\dots,(i_n,j_n)]$$

the determinant of the matrix with the ℓ-th row equal to the j_ℓ-th row of the generic matrix $\xi_{i_\ell}=(x_{ij}^{(i_\ell)})$. For example, in the case $n=2$, $[(1,1),(2,1)]$ denotes the determinant of the matrix

$$\begin{pmatrix} x_{11}^{(1)} & x_{12}^{(1)}\\ x_{11}^{(2)} & x_{12}^{(2)}\end{pmatrix},$$

and thus equals $x_{11}^{(1)}x_{12}^{(2)}-x_{12}^{(1)}x_{11}^{(2)}$. Further, $\mathbb{N}_d$ denotes the set $\{1,\dots,d\}$. If $I=\{i_1,\dots,i_p\}\subseteq\mathbb{N}_m$ with $i_1<\dots<i_p$, then we denote

$$\overline{x}_m^I=(x_1,\dots,\hat{x}_{i_1},\dots,\hat{x}_{i_p},\dots,x_m).$$

Finally, as above by e_{ij} we denote the matrix units.

Without loss of generality we may assume that $K=\mathbb{N}_m$. When dealing with two-sided functional identities (3.2) it is often convenient to deal with the case where K and L are arbitrary subsets of $\mathbb{N}_m$, but for one-sided identities this would only cause notational complications.

PROPOSITION 3.2.2. *Let $\sum_{k=1}^m F_k(\overline{x}_m^k)x_k=0$ be a functional identity of $M_n(F)$, where $F_k:M_n(F)^{m-1}\to M_n(F)$ are multilinear functions. Then there exist multilinear scalar-valued functions $\lambda_{\ell IJ}$ such that*

$$F_k(\overline{x}_m^k)=\sum_{\ell,I,J}(-1)^s\lambda_{\ell IJ}(\overline{x}_m^I)\big[(i_1,j_1),\dots,\widehat{(i_s,j_s)},\dots,(i_{n+1},j_{n+1})\big]e_{\ell j_s},$$

where the sum runs over all $\ell \in \mathbb{N}_n$, all $I = \{i_1, \dots, i_{n+1}\} \subset \mathbb{N}_m$ such that $i_s = k$ for some s, $i_1 < \cdots < i_{n+1}$, and all $J = (j_1, \dots, j_{n+1}) \in \mathbb{N}_n^{n+1}$.

PROOF. We can view $\sum_{k=1}^m F_k(\overline{x}_m^k)x_k = 0$ as the identity in $M_n(\mathcal{C})$, since F_k are multilinear and hence polynomial maps. We then have an identity of the form

$$\sum_{k=1}^{m} H_k \xi_k = 0,$$

where $\xi_k = (x_{ij}^{(k)})$, $1 \le k \le m$, are generic matrices, and $H_k \in M_n(\mathcal{C})$ correspond to F_k. Note that this identity is equivalent to n identities

$$\sum_{k=1}^{m} e_{\ell\ell} H_k \xi_k = 0, \quad 1 \le \ell \le n.$$

We thus first find solutions of each of them. Without loss of generality we assume that $\ell = 1$ and $H_k = e_{11}H_k$, $1 \le k \le m$. We can further rewrite this identity as

$$\sum_{k=1}^{m} \sum_{j=1}^{n} H_{1j}^{(k)} x_{jr}^{(k)} = 0$$

for $1 \le r \le n$, which can be viewed as

$$\left(H_{11}^{(1)}, \dots, H_{1n}^{(1)}, \ \dots \ , H_{11}^{(m)}, \dots, H_{1n}^{(m)}\right) \begin{pmatrix} x_{11}^{(1)} & \dots & x_{1n}^{(1)} \\ \vdots & \ddots & \vdots \\ x_{n1}^{(1)} & \dots & x_{nn}^{(1)} \\ & & \\ & \vdots & \\ & & \\ x_{11}^{(m)} & \dots & x_{1n}^{(m)} \\ \vdots & \ddots & \vdots \\ x_{n1}^{(m)} & \dots & x_{nn}^{(m)} \end{pmatrix} = (0 \dots 0).$$

We write $\xi_{n,m}$ for the matrix on the right. By definition,

$$(H_{11}^{(1)}, \dots, H_{1n}^{(1)}, \ \dots \ , H_{11}^{(m)}, \dots, H_{1n}^{(m)})$$

is a syzygy on the rows of the generic matrix $X_{n,m}$. The Gröbner basis of the module of syzygies is described in Theorem 3.2.1. Since in our case $H_{ij}^{(k)}$ are multilinear as functions of $\xi_1, \dots, \xi_{k-1}, \xi_{k+1}, \dots, \xi_m$, we look for the elements in the Gröbner basis with the same property. Note that

$$\big[(i_1, j_1), \dots, (i_n, j_n)\big]$$

denotes the determinant of the submatrix of $\xi_{n,m}$ with the k-th row equal to j_k-th row of the generic matrix ξ_{i_k}. By Theorem 3.2.1 the desired generators are

$$\sum_{k=1}^{n+1} (-1)^k \big[(i_1, j_1), \dots, \widehat{(i_k, j_k)}, \dots, (i_{n+1}, j_{n+1})\big] u_{(i_k - 1)n + j_k}$$

for $1 \le i_1 < i_2 < \cdots < i_{n+1} \le m$, $1 \le j_1, \dots, j_{n+1} \le n$, and the proposition follows. $\square$

3.2.3. *Left-sided functional identity as a GPI.* Each function F_k is multilinear so it corresponds to a multilinear generalized polynomial $f_k \in M_n(F) * F\langle X\rangle$ such that the evaluation of f_k on $M_n(F)$ coincides with F_k. Note that this correspondence is uniquely determined up to the generalized polynomial identities. At any rate, the problem of describing functional identities $\sum_{k\in K} F_k(\overline{x}^k_m)x_k = 0$ is basically equivalent to the problem of describing GPI's of the form $\sum_{k\in K} f_k(\overline{x}^k_m)x_k$. In this section we will deal with the latter since the GPI setting seems to be more convenient for formulating the main result.

We start with a technical lemma. By $D_{j_1,\dots,j_n}$ we denote the determinant of the matrix whose k-th row is equal to the j_k-th row of the generic matrix ξ_k.

LEMMA 3.2.3. *Let $i_\ell, j_\ell \in \mathbb{N}_n$, $1 \le \ell \le n+1$. If $\{i_1,\dots,i_n\} = \{1,\dots,n\}$ then*

$$\begin{aligned}
&e_{i_{n+1},j_n}x_nQ_{n-1}(e_{i_1j_1}x_1,\dots,e_{i_{n-1},j_{n-1}}x_{n-1})e_{i_nj_{n+1}}\\
=&(-1)^\tau D_{j_1,\dots,j_n}e_{i_{n+1},j_{n+1}}\\
=&e_{i_{n+1},j_{n+1}}\tilde{Q}_n(e_{i_1j_1}x_1,\dots,e_{i_nj_n}x_n),
\end{aligned}$$

where $\tau \in S_n$ is given by $\tau(k) = i_k$, otherwise

$$Q_{n-1}(e_{i_1,j_1}x_1,\dots,e_{i_{n-1},j_{n-1}}x_{n-1})e_{i_n,j_n} = 0.$$

PROOF. The last assertion follows from the fact that Q_{n-1} is an identity of $M_{n-1}(F)$. Indeed, we may assume without loss of generality that $i_k = k$ for $1 \le k \le s < n$, $i_s = \dots = i_n = s$, and hence

$$Q_{n-1}(e_{1j_1}x_1,\dots,e_{sj_s}x_s,\dots,e_{sj_{n-1}}x_{n-1})e_{sj_n} \in \big(M_n(F)e_{nn}\big)e_{sj_n} = \{0\}.$$

The first equality clearly follows by using the identity

$$e_{i_{n+1},j_n}x_n\phi_\sigma(e_{i_1j_1}x_1,\dots,e_{i_{n-1},j_{n-1}}x_{n-1})e_{i_n,j_{n+1}} = x^{(1)}_{j_1i_{\sigma(1)}}\cdots x^{(n)}_{j_ni_{\sigma(n)}}e_{i_{n+1},j_{n+1}}$$

for $\sigma \in S_n$ in the expression (1.2) of Q_{n-1}, and the second one follows from

$$(-1)^\tau D_{j_1,\dots,j_n} = -\tilde{Q}_n(e_{i_1j_1}x_1,\dots,e_{i_nj_n}x_n),$$

which can be deduced in a similar way after applying the identity

$$\tilde{Q}_n(y_1,\dots,y_n) = -\sum_{\sigma\in S_n\subset S_{n+1}}(-1)^\sigma\phi_\sigma(y_1,\dots,y_n).$$

□

Let us call $g \in M_n(F) * F\langle X\rangle$ a *central generalized polynomial* if all its evaluations on $M_n(F)$ are scalar matrices. For instance, $\tilde{Q}_n$ is a central generalized polynomial.

THEOREM 3.2.4. *Let $f_1,\dots,f_m \in M_n(F) * F\langle X\rangle$ be multilinear generalized polynomials such that $P = \sum_{k=1}^m f_k(\overline{x}^k_m)x_k$ is a GPI of $M_n(F)$. Then P can be written as a sum of GPI's of the form*

$$g\big[\tilde{Q}_n(a_1x_{k_1},\dots,a_nx_{k_n}),a_{n+1}x_{k_{n+1}}\big],$$

where $k_i \ne k_j$ if $i \ne j$, $a_i \in M_n(F)$, and g is a multilinear central generalized polynomial (in all variables except x_{k_i}).

PROOF. Proposition 3.2.2 implies that P can be written as a sum of GPI's of the form

$$\sum_{k=1}^{n+1}(-1)^k\big[(i_1,j_1),\dots,\widehat{(i_k,j_k)},\dots,(i_{n+1},j_{n+1})\big]e_{\ell j_k}x_{i_k},$$

multiplied by central generalized polynomials; the determinants appearing in the identity can also be viewed as central generalized polynomials. It is thus enough to prove that this identity can be written in the desired way. We can assume without loss of generality that $\ell = 1$, $i_k = k$, $1 \le k \le n+1$. Hence we can write the identity as

$$\sum_{k=1}^{n+1}(-1)^k D_{j_1,\dots,\hat{j}_k,\dots,j_{n+1}} e_{1j_k}x_k,$$

where $D_{j_1,\dots,\hat{j}_k,\dots,j_{n+1}}$ stands for the determinant of the matrix obtained by removing the k-th row from the matrix with the ℓ-th row equal to the j_ℓ-th row of the generic matrix ξ_ℓ. Note that

$$\tilde{Q}_n(x_1,\dots,x_n) = \sum_{k=1}^{n} Q_{n-1}(x_1,\dots,\hat{x}_k,\dots,x_n)x_k.$$

Applying Lemma 3.2.3 we thus obtain

$$\begin{aligned}
&e_{1j_{n+1}}x_{n+1}\tilde{Q}_n(e_{1j_1}x_1,\dots,e_{nj_n}x_n)\\
&\qquad\qquad=e_{1j_{n+1}}x_{n+1}\sum_{k=1}^{n} Q_{n-1}(e_{1j_1}x_1,\dots,\widehat{e_{kj_k}x_k},\dots,e_{nj_n}x_n)e_{kj_k}x_k\\
&\qquad\qquad=\sum_{k=1}^{n}(-1)^{\tau_k} D_{j_1,\dots,\hat{j}_k,\dots,j_{n+1}} e_{1j_k}x_k,
\end{aligned}$$

where $\tau_k = (k, k+1,\dots,n)$, and thus $(-1)^{\tau_k} = (-1)^{n-k}$. Applying Lemma 3.2.3 once again we obtain

$$\begin{aligned}
\sum_{k=1}^{n+1}(-1)^k D_{j_1,\dots,\hat{j}_k,\dots,j_{n+1}} e_{1j_k}x_k &= (-1)^n e_{1j_{n+1}}x_{n+1}\tilde{Q}_n(e_{1j_1}x_1,\dots,e_{nj_n}x_n)\\
&\quad + (-1)^{n+1}Q_n(e_{1j_1}x_1,\dots,e_{nj_n}x_n)e_{1j_{n+1}}x_{n+1},
\end{aligned}$$

which yields the desired conclusion. $\square$

REMARK 3.2.5. Let us remark that the identity (3.1) can be written as

$$\sum_{k=1}^{n+1}\tilde{Q}_n(a_1x_1,\dots,\widehat{a_kx_k},\dots,a_nx_n,\tilde{a}_{n+1,k}x_{n+1})\tilde{a}_kx_k = 0$$

for some $\tilde{a}_k, \tilde{a}_{n+1,k} \in M_n(F)$. It is enough to establish the statement in the case when $a_k = e_{i_k,j_k}$, $1 \le k \le n+1$, are matrix units. Applying Lemma 3.2.3 we obtain, similarly as in the proof of Corollary 3.2.4, the identity

$$\begin{aligned}
&e_{i_{n+1}j_{n+1}}x_{n+1}\tilde{Q}_n(e_{i_1j_1}x_1,\dots,e_{i_nj_n}x_n)\\
&\quad= e_{i_{n+1}j_{n+1}}x_{n+1}\sum_k Q_{n-1}(e_{i_1j_1}x_1,\dots,\widehat{e_{i_kj_k}x_k},\dots,e_{i_nj_n}x_n)e_{i_kj_k}x_k\\
&\qquad\qquad= \sum_k(-1)^{\tau} D_{j_1,\dots,\hat{j}_k,\dots,j_{n+1}} e_{i_{n+1},j_k}x_k\\
&\qquad\quad= \sum_k(-1)^n\tilde{Q}_n(e_{i_1j_1}x_1,\dots,\widehat{e_{i_k,j_k}x_k},\dots,e_{i_kj_{n+1}}x_{n+1})e_{i_{n+1}j_k}x_k,
\end{aligned}$$

where $\tau \in S_n$, $\tau(k) = i_k$.

3.2.4. *An application: Characterization of the determinant.* The determinants have appeared throughout our discussion on one-sided identities. To point out their role more clearly, we record two simple corollaries at the end of the section.

Let A be an algebra over a field F. A function $T : A \to A$ is said to be the *trace of an r-linear function $F : A^r \to A$* if $T(x) = F(x,\dots,x)$ for all $x \in A$ (if $r = 0$ then T is a constant function). We remark that if F is symmetric and $\mathrm{char}(F)$ is 0 or greater than r, then F is uniquely determined by its trace T. This can be shown by applying the linearization process.

COROLLARY 3.2.6. *Let $m > n$ and let $T_k : M_n(F) \to M_n(F)$ be the trace of an $(m-1)$-linear function $F_k : M_n(F)^{m-1} \to M_n(F)$, $1 \le k \le m$. If $\sum_{k=1}^m F_k(\overline{x}_m^k)x_k = 0$ is a functional identity of $M_n(F)$, then each T_k can be written as $T_k(x) = \det(x)S_k(x)$ where S_k is the trace of an $(m-1-n)$-linear function.*

PROOF. Proposition 3.2.2 shows that $T_k(x_1)$ is a sum of the terms of the form

$$(-1)^s\lambda_{\ell IJ}(x_1,\ldots,x_1)\big[(1,j_1),\ldots,\widehat{(1,j_s)},\ldots,(1,j_{n+1})\big]e_{\ell j_s}.$$

By definition, $\big[(1,j_1),\ldots,\widehat{(1,j_s)},\ldots,(1,j_{n+1})\big] = \pm\det(x_1)$ if $j_1,\ldots,\hat{j}_s,\ldots,j_{n+1} \in \mathbb{N}_n$ are distinct, and 0 otherwise. This proves the corollary. $\square$

In the final corollary we add more assumptions in order to obtain an abstract characterization of the determinant.

COROLLARY 3.2.7. *Let $F : M_n(F)^n \to M_n(F)$ be an n-linear function satisfying $F(1,\ldots,1) = 1$, and let T be the trace of F. If $\mathrm{char}(F) = 0$ or $\mathrm{char}(F) > n$, then the following statements are equivalent:*

(i) *There exist multilinear functions $F_1,\ldots,F_n : M_n(F)^n \to M_n(F)$ such that*
$$\sum_{k=1}^n F_k(\overline{x}_{n+1}^k)x_k + F(\overline{x}_{n+1}^{n+1})x_{n+1} = 0$$
is a functional identity of $M_n(F)$.

(ii) $T(x) = \det(x)\cdot 1$.

PROOF. Corollary 3.2.6 shows that (i) implies (ii). To establish the converse, just note that $\det(x)\cdot 1$ is the trace of the function $\frac{1}{n!}\tilde{Q}_n(x_1,\ldots,x_n)$, and that $\frac{1}{n!}[\tilde{Q}_n(x_1,\ldots,x_n),x_{n+1}] = 0$ can be written in the desired form with $F(\overline{x}_{n+1}^{n+1}) = \frac{1}{n!}\tilde{Q}_n(\overline{x}_{n+1}^{n+1})$. $\square$

3.3. Two-sided functional identities. In this section we consider the general two-sided functional identities

$$\sum_{k\in K} F_k(\overline{x}_m^k)x_k = \sum_{l\in L} x_l G_l(\overline{x}_m^l). \tag{3.2}$$

Let us first examine what result can be expected.

3.3.1. *Solutions of two-sided functional identities.* Let us first consider (3.2) in an arbitrary algebra A with center Z. Suppose there exist multilinear functions

$$p_{kl} : A^{m-2} \to A,\ k \in K,\ l \in L,\ k \ne l,$$
$$\lambda_i : A^{m-1} \to Z,\ i \in K \cup L,$$

such that

$$\begin{aligned} F_k(\overline{x}_m^k) &= \sum_{l\in L,\, l\ne k} x_l p_{kl}(\overline{x}_m^{kl}) + \lambda_k(\overline{x}_m^k), \quad k \in K,\\ G_l(\overline{x}_m^l) &= \sum_{k\in K,\, k\ne l} p_{kl}(\overline{x}_m^{kl})x_k + \lambda_l(\overline{x}_m^l), \quad l \in L, \\ \lambda_i &= 0 \quad \text{if} \quad i \notin K\cap L. \end{aligned} \tag{3.3}$$

Note that then (3.2) is fulfilled. We call (3.3) a *standard solution* of (3.2). In a large class of algebras a standard solution is also the only possible solution of (3.2) [**BCM07**]. Especially in infinite dimensional algebras this often turns out to be the case. For $A = M_n(F)$, this holds provided only if $|K| \le n$ and $|L| \le n$

(see e.g. [**BCM07**, Corollary 2.23]). Indeed, if $|K| > n$ or $|L| > n$ then there exist nonstandard solutions of one-sided identities. Do all nonstandard solutions of (3.2) in $M_n(F)$ arise from the one-sided identities? More specifically, let us call a solution of (3.2) a *standard solution modulo one-sided identities* if there exist multilinear functions

$$\begin{aligned}
&\varphi_k, \psi_l : A^{m-1} \to A,\ k \in K,\ l \in L,\\
&p_{kl} : A^{m-2} \to A,\ k \in K,\ l \in L,\ k \neq l,\\
&\lambda_i : A^{m-1} \to Z,\ i \in K \cup L,
\end{aligned}$$

such that

$$\begin{aligned}
F_k(\overline{x}_m^k) &= \sum_{l \in L,\, l \neq k} x_l p_{kl}(\overline{x}_m^{kl}) + \lambda_k(\overline{x}_m^k) + \varphi_k(\overline{x}_m^k), \quad k \in K,\\
G_l(\overline{x}_m^l) &= \sum_{k \in K,\, k \neq l} p_{kl}(\overline{x}_m^{kl}) x_k + \lambda_l(\overline{x}_m^l) + \psi_l(\overline{x}_m^l), \quad l \in L,\\
&\lambda_i = 0 \quad \text{if} \quad i \notin K \cap L,\\
&\sum_{k \in K} \varphi_k(\overline{x}_m^k) x_k = \sum_{l \in L} x_l \psi_l(\overline{x}_m^l) = 0.
\end{aligned} \tag{3.4}$$

This notion is not vacuous. Namely, there do exist algebras admitting functional identities (3.2) having solutions that are not standard modulo one-sided identities. One can actually find algebras with this property in which all solutions of one-sided identities are standard; see e.g. [**BCM07**, Example 5.29].

In the rest of this section we show that every solution of (3.2) on $A = M_n(F)$ is standard modulo one-sided identities. As solutions of one-sided identities have been described in the preceding subsection, this will give a complete description of two-sided functional identities on $M_n(F)$.

3.3.2. *Gröbner basis of a module representing functional identities.* As in the case of one-sided functional identities we use the coordinate-wise approach and apply the theory of Gröbner bases.

First we introduce the necessary framework. We take $m, n \in \mathbb{N}$ and define $n^2 \times n^2$-matrices

$$\xi_k' = \begin{pmatrix}
x_{11}^{(k)} & \dots & x_{1n}^{(k)} & & & & \\
& \vdots & & & & & \\
x_{n1}^{(k)} & \dots & x_{nn}^{(k)} & & & & \\
& & & \ddots & & & \\
& & & & x_{11}^{(k)} & \dots & x_{1n}^{(k)} \\
& & & & & \vdots & \\
& & & & x_{n1}^{(k)} & \dots & x_{nn}^{(k)}
\end{pmatrix} = \xi_k \otimes 1 \in M_n(\mathcal{C}) \otimes M_n(\mathcal{C}),$$

$$\xi_k'' = \begin{pmatrix}
x_{11}^{(k)} & & & & x_{n1}^{(k)} & & \\
& \ddots & & \dots & & \ddots & \\
& & x_{11}^{(k)} & & & & x_{n1}^{(k)} \\
& & & \vdots & & & \\
x_{1n}^{(k)} & & & & x_{nn}^{(k)} & & \\
& \ddots & & \dots & & \ddots & \\
& & x_{1n}^{(k)} & & & & x_{nn}^{(k)}
\end{pmatrix} = 1 \otimes \xi_k^t \in M_n(\mathcal{C}) \otimes M_n(\mathcal{C}),$$

and write
$$\Xi = (\xi_1', \dots, \xi_m', \xi_1'', \dots, \xi_m'')^t.$$
Let M be the submodule of the module $\mathcal{C}^{n^2}$ generated by the rows of Ξ. We aim to find a Gröbner basis of M.

We need some more notation. Let $K = \{k_1, \dots, k_a\}$ and $L = \{l_1, \dots, l_b\}$ be subsets of $\mathbb{N}_m$, and let us write $Q = K \cap L = \{q_1, \dots, q_c\}$ (this set may be empty). Let $(d_1, f_1), \dots, (d_c, f_c)$ be such that $q_\ell = k_{d_\ell} = l_{f_\ell}$. We attach the tuples $V = (v_1, \dots, v_a) \in \mathbb{N}_n^a$ and $S = (s_1, \dots, s_b) \in \mathbb{N}_n^b$ to K and L, resp. For a subset $U \subset \mathbb{N}_a$, $|U| = a - c$, we write $U^{\mathsf{c}} = \{u_1', \dots, u_c'\}$, $u_1' < \dots < u_c'$, for the complement of U in $\mathbb{N}_a$. Let σ belong to $\operatorname{Sym} U^{\mathsf{c}}$, the permutation group of U^{c}. We choose $\lambda \in \mathbb{N}_n$ and write
$$D_\lambda^{\mathbf{c}}(Q_\sigma, L_S \setminus Q)$$
for the determinant of the $b \times b$-matrix $Y = (y_{ij})$, where
$$y_{i\ell} = \begin{cases} x_{is_\ell}^{(l_\ell)} & \text{for } 1 \le i \le b-1,\ \ell \in \mathbb{N}_b \setminus \{f_1, \dots, f_c\}, \\ x_{\lambda s_\ell}^{(l_\ell)} & \text{for } i = b,\ \ell \in \mathbb{N}_b \setminus \{f_1, \dots, f_c\}, \end{cases}$$
$$y_{if_\ell} = \begin{cases} x_{i\sigma(u_\ell')}^{(q_\ell)} & \text{for } 1 \le i \le b-1,\ 1 \le \ell \le c, \\ x_{\lambda\sigma(u_\ell')}^{(q_\ell)} & \text{for } i = b,\ 1 \le \ell \le c. \end{cases}$$
Analogously, let $\tau \in \operatorname{Sym} W^{\mathsf{c}}$, where $W^{\mathsf{c}} = \{w_1', \dots, w_c'\} \subseteq \mathbb{N}_b$, and write
$$D_\lambda^{\mathbf{r}}(Q_\tau, K_V \setminus Q)$$
for the determinant of the $a \times a$-matrix $Z = (z_{ij})$, where
$$z_{\ell j} = \begin{cases} x_{v_\ell j}^{(k_\ell)} & \text{for } 1 \le j \le a-1,\ \ell \in \mathbb{N}_a \setminus \{d_1, \dots, d_c\}, \\ x_{v_\ell \lambda}^{(k_\ell)} & \text{for } j = a,\ \ell \in \mathbb{N}_a \setminus \{d_1, \dots, d_c\}, \end{cases}$$
$$z_{d_\ell j} = \begin{cases} x_{\tau(w_\ell')j}^{(q_\ell)} & \text{for } 1 \le j \le a-1,\ 1 \le \ell \le c, \\ x_{\tau(w_\ell')\lambda}^{(q_\ell)} & \text{for } j = a,\ 1 \le \ell \le c. \end{cases}$$
In particular, we write $D_\lambda^{\mathbf{c}}(L_S)$, $D_\lambda^{\mathbf{r}}(K_V)$ for $D_\lambda^{\mathbf{c}}(\emptyset, L_S)$, $D_\lambda^{\mathbf{r}}(\emptyset, K_V)$, resp. (Note that Y and Z are formed from the columns (resp. rows) of certain matrices, which is the reason for using $^{\mathbf{c}}$ (resp. $^{\mathbf{r}}$) in the above notation.)

We further denote by
$$d_{\lambda, W^{\mathsf{c}}}^{\mathbf{c}}(Q_\sigma), \quad d_{\lambda, W}^{\mathbf{c}}(L_S \setminus Q)$$
the determinant of the submatrix of Y containing the columns labeled by $f_1, \dots, f_\ell$ (resp. by $\ell \in \mathbb{N}_b \setminus \{f_1, \dots, f_\ell\}$) and rows labeled by $i_\ell \in W^{\mathsf{c}}$ (resp. $i_\ell \in W$), and by
$$d_{\lambda, U^{\mathsf{c}}}^{\mathbf{r}}(Q_\tau), \quad d_{\lambda, U}^{\mathbf{r}}(K_V \setminus Q)$$
the determinant of the submatrix of Z containing the rows labeled by $d_1, \dots, d_\ell$ (resp. by $\ell \in \mathbb{N}_a \setminus \{d_1, \dots, d_\ell\}$) and columns labeled by $j_\ell \in U^{\mathsf{c}}$ (resp. $j_\ell \in U$). We let the determinant of the empty matrix be 1.

EXAMPLE 3.3.1. Let $n = 4$, $K = \{1, 2, 3\}$, $L = \{2, 3, 4, 5\}$, $V = (4, 1, 2)$, $S = (3, 4, 2, 1)$, $U = \{2\}$, $W = \{1, 3\}$, $\sigma = \mathrm{id}$, $\tau = (24)$, $\lambda = 4$. Then $Q = \{2, 3\}$,
$$D_\lambda^{\mathbf{c}}(Q_\sigma, L_S \backslash Q) = \begin{vmatrix} x_{11}^{(2)} & x_{13}^{(3)} & x_{12}^{(4)} & x_{11}^{(5)} \\ x_{21}^{(2)} & x_{23}^{(3)} & x_{22}^{(4)} & x_{21}^{(5)} \\ x_{31}^{(2)} & x_{33}^{(3)} & x_{32}^{(4)} & x_{31}^{(5)} \\ x_{41}^{(2)} & x_{43}^{(3)} & x_{42}^{(4)} & x_{41}^{(5)} \end{vmatrix}, \; D_\lambda^{\mathbf{r}}(Q_\tau, K_V \backslash Q) = \begin{vmatrix} x_{41}^{(1)} & x_{42}^{(1)} & x_{44}^{(1)} \\ x_{41}^{(2)} & x_{42}^{(2)} & x_{44}^{(2)} \\ x_{21}^{(3)} & x_{22}^{(3)} & x_{24}^{(3)} \end{vmatrix},$$
$$d_{\lambda, W^{\mathsf{c}}}^{\mathbf{c}}(Q_\sigma) = \begin{vmatrix} x_{21}^{(2)} & x_{23}^{(3)} \\ x_{41}^{(2)} & x_{43}^{(3)} \end{vmatrix}, \quad d_{\lambda, W}^{\mathbf{c}}(L_S \setminus Q) = \begin{vmatrix} x_{12}^{(4)} & x_{11}^{(5)} \\ x_{32}^{(4)} & x_{31}^{(5)} \end{vmatrix},$$

$$d^{\mathbf{r}}_{\lambda,U^c}(Q_\tau) = \begin{vmatrix} x^{(2)}_{41} & x^{(2)}_{44} \\ x^{(3)}_{21} & x^{(3)}_{24} \end{vmatrix}, \quad d^{\mathbf{r}}_{\lambda,U}(K_V \setminus Q) = \begin{vmatrix} x^{(1)}_{42} \end{vmatrix}.$$

Let $u_{\gamma,\delta}$ denote the $n(\gamma-1)+\delta$-th basis element in the $\mathcal{C}$-module $\mathcal{C}^{n^2}$. We write

$$G' = \Big\{ \sum_{\alpha=|K|}^{n} D^{\mathbf{r}}_{\alpha}(K_V) u_{\gamma,\alpha} \mid \gamma \in \mathbb{N}_n, K \subset \mathbb{N}_m, V \subset \mathbb{N}_n, |V| = |K| \Big\},$$

$$G'' = \Big\{ \sum_{\beta=|L|}^{n} D^{\mathbf{c}}_{\beta}(L_S) u_{\beta,\delta} \mid \delta \in \mathbb{N}_n, L \subset \mathbb{N}_m, S \subset \mathbb{N}_n, |S| = |L| \Big\}.$$

Note that every polynomial in $\mathcal{C}$ can be treated as a function on $M_n(F)^m$. Let $\mathcal{C}_{mult}$ denote the elements in $\mathcal{C}$ that are multilinear in some set of variables $x_{k_1}, \dots, x_{k_\ell}$, $1 \le k_i \ne k_j \le m$. We will say that G is a *multilinear Gröbner basis* of a $\mathcal{C}$-module $N \subset \mathcal{C}^r$ if there exists a Gröbner basis $\tilde{G}$ of N such that $G = \tilde{G} \cap \mathcal{C}_{mult}$.

Proposition 3.3.2 ([**Onn94**, Theorem 8.4]). *The set G' is a multilinear Gröbner basis of the submodule M' of M generated by the first mn^2 rows of Ξ, and the set G'' is a multilinear Gröbner basis of the submodule M'' of M generated by the last mn^2 rows of Ξ.*

We will use this proposition in order to show that one can obtain a multilinear Gröbner basis of M by simply joining G' and G''. To establish this result in Lemma 3.3.5 we need some preliminary lemmas.

For a subset $U \subset \mathbb{N}_a$ we set $U_\alpha = U$ if $a \notin U$, and $U_\alpha = (U \setminus \{a\}) \cup \{\alpha\}$ if $a \in U$.

Lemma 3.3.3. *Let $a, b, c \in \mathbb{N}$, $c \le a, b \le n$, $U \subset \mathbb{N}_a$, $W \subset \mathbb{N}_b$, $|U| = |W| = c$, $Q = \{q_1, \dots, q_c\}$. For $a \le \alpha$, $b \le \beta$ we have*

$$\sum_{\sigma \in \operatorname{Sym} U_\alpha} (-1)^\sigma d^{\mathbf{c}}_{\beta,W}(Q_\sigma) = \sum_{\tau \in \operatorname{Sym} W_\beta} (-1)^\tau d^{\mathbf{r}}_{\alpha,U}(Q_\tau). \tag{3.5}$$

Proof. Notice that we can assume without loss of generality that $U = W = Q = \{1, \dots, c\}$, $\alpha = \beta = c$. We compute

$$\begin{aligned} \sum_{\sigma \in \operatorname{Sym} U} (-1)^\sigma d^{\mathbf{c}}_{\beta,W}(Q_\sigma) &= \sum_{\sigma \in \operatorname{Sym} c} (-1)^\sigma \sum_{\rho \in \operatorname{Sym} c} (-1)^\rho \, x^{(\rho(1))}_{1,\sigma\rho(1)} \cdots x^{(\rho(c))}_{c,\sigma\rho(c)} \\ &= \sum_{\rho \in \operatorname{Sym} c} (-1)^{\rho^{-1}} \sum_{\sigma \in \operatorname{Sym} c} (-1)^{\sigma^{-1}} \, x^{(\sigma^{-1}(1))}_{\rho^{-1}\sigma^{-1}(1),1} \cdots x^{(\sigma^{-1}(c))}_{\rho^{-1}\sigma^{-1}(c),c} \\ &= \sum_{\tau \in \operatorname{Sym} W} (-1)^\tau d^{\mathbf{r}}_{\alpha,U}(Q_\tau). \end{aligned}$$

$\square$

Lemma 3.3.4. *Let $K = \{k_1, \dots, k_a\}$, $L = \{l_1, \dots, l_b\}$, $Q = K \cap L = \{q_1, \dots, q_c\}$. Let $(d_1, f_1), \dots, (d_c, f_c)$ be such that $q_\ell = k_{d_\ell} = l_{f_\ell}$, and let $V = (v_1, \dots, v_a) \in \mathbb{N}_n^a$, $S = (s_1, \dots, s_b) \in \mathbb{N}_n^b$. For $a \le \alpha \le n$, $b \le \beta \le n$ we have*

$$\begin{aligned} &(-1)^{\sum d_\ell} \sum_{U \subset \mathbb{N}_a, |U| = a-c} (-1)^{\sum_{U^c} u} \, d^{\mathbf{r}}_{\alpha,U}(K_V \setminus Q) \sum_{\sigma \in \operatorname{Sym}(U^c)_\alpha} (-1)^\sigma D^{\mathbf{c}}_{\beta}(Q_\sigma, L_S \setminus Q) \\ &= (-1)^{\sum f_\ell} \sum_{W \subset \mathbb{N}_b, |W| = b-c} (-1)^{\sum_{W^c} w} \, d^{\mathbf{c}}_{\beta,W}(L_S \setminus Q) \sum_{\tau \in \operatorname{Sym}(W^c)_\beta} (-1)^\tau D^{\mathbf{r}}_{\alpha}(Q_\tau, K_V \setminus Q). \end{aligned} \tag{3.6}$$

PROOF. Using the Laplace expansion by the columns $f_1, \ldots, f_c$ we obtain

$$D^{\mathbf{c}}_\beta(Q_\sigma, L_S \setminus Q) = \sum_{W \subset \mathbb{N}_b, |W|=b-c} (-1)^{\sum f_\ell}(-1)^{\sum_{W^c} w} d^{\mathbf{c}}_{\beta,W}(L_S \setminus Q) d^{\mathbf{c}}_{\beta,W^c}(Q_\sigma),$$

and analogously using the Laplace expansion by the rows $d_1, \ldots, d_c$ we have

$$D^{\mathbf{r}}_\alpha(Q_\tau, K_V \setminus Q) = \sum_{U \subset \mathbb{N}_a, |U|=a-c} (-1)^{\sum d_\ell}(-1)^{\sum_{U^c} u} d^{\mathbf{r}}_{\alpha,U}(K_V \setminus Q) d^{\mathbf{r}}_{\alpha,U^c}(Q_\tau).$$

By applying Lemma 3.3.3 we arrive at the desired conclusion. □

Before proceeding to the proof of the next lemma we make a little digression and recall some facts concerning Gröbner bases and syzygies (see e.g. [**Eis95**]). Let A be a polynomial algebra. By $u_1, \ldots, u_r$ we denote the generators of the free module A^r. Let N be a module over A and $\{g_1, \ldots, g_t\}$ its Gröbner basis with respect to any monomial order on A^r. Let $\mathrm{in}(g_i)$ stand for the initial term of g_i. If $\mathrm{in}(g_i)$ and $\mathrm{in}(g_j)$ involve the same basis element of A^r, set

$$m_{ij} = \frac{\mathrm{in}(g_i)}{\mathrm{GCD}(\mathrm{in}(g_i), \mathrm{in}(g_j))} \in A.$$

For each such pair i, j choose an expression

$$\sigma_{ij} := m_{ji} g_i - m_{ij} g_j = \sum_\ell h^{(ij)}_\ell g_\ell,$$

such that $\mathrm{in}(\sigma_{ij}) \geq \mathrm{in}(h^{(ij)}_\ell g_\ell)$ for every ℓ; it is called a standard expression of σ_{ij} in terms of the g_ℓ, and its existence is guaranteed by the fact that $\{g_1, \ldots, g_t\}$ is a Gröbner basis of N. For other pairs i, j set $m_{ij} = 0$, $h^{(ij)}_\ell = 0$. By Shreyer's theorem (see e.g. [**Eis95**, Theorem 15.10]) the module of syzygies on the Gröbner basis $\{g_1, \ldots, g_t\}$ of a module N is generated by

$$\tau_{ij} := m_{ji} u_i - m_{ij} u_j - \sum_\ell h^{(ij)}_\ell u_\ell,$$

$1 \leq i, j \leq t$.

Let us define a monomial order $>$ on the module C^{n^2}. On $\mathcal{C}$ we set $x^{(k_1)}_{i_1 j_1} > x^{(k_2)}_{i_2 j_2}$ if $(k_1, i_1, j_1) < (k_2, i_2, j_2)$ lexicographically (i.e. $k_1 < k_2$, or $k_1 = k_2$ and $i_1 < i_2$, or $k_1 = k_2$, $i_1 = i_2$ and $j_1 < j_2$). We define $p u_{\alpha,\beta} > q u_{\gamma,\delta}$ if $(\alpha, \beta, q) < (\gamma, \delta, p)$ lexicographically (i.e. $\alpha < \gamma$, or $\alpha = \gamma$ and $\beta < \delta$, or $\alpha = \gamma$, $\beta = \delta$ and $p > q$).

LEMMA 3.3.5. *The set $G' \cup G''$ is a multilinear Gröbner basis of M with respect to the order $>$ on $\mathcal{C}^{n^2}$.*

PROOF. Using the Buchberger's criterion (see e.g. [**Eis95**, Theorem 15.8]) together with Proposition 3.3.2 we see that it suffices to verify that σ_{ij} has a standard expression in terms of $g_\ell \in G' \cup G''$ for $g_i \in G'$, $g_j \in G''$, initial terms of which involve the same basis element in C^{n^2} and for which σ_{ij} is multilinear. Choose $g_i \in G'$, $g_j \in G''$ such that $\mathrm{in}(g_i)$ and $\mathrm{in}(g_j)$ involve the same basis element of $\mathcal{C}^{n^2}$. Define sets $K, L \subset \mathbb{N}_m$ such that g_i depends on the variables x_{k_ℓ} for $k_\ell \in K$, g_j on x_{l_ℓ} for $l_\ell \in L$. Then the initial term involves the variables in $Q = K \cap L$. For σ_{ij} to be multilinear the factors in the initial terms of g_i and g_j dependent on the variables in Q need to coincide. Hence,

$$g_i = \sum_{\alpha=a}^{n} D^{\mathbf{r}}_\alpha(Q_\tau, K_V \setminus Q) u_{b,\alpha},$$

$$g_j = \sum_{\beta=b}^{n} D^{\mathbf{c}}_\beta(Q_\sigma, L_S \setminus Q) u_{\beta,a},$$

where $|K| = a$, $|L| = b$, $|Q| = c$, $V \in \mathbb{N}_n^a$, $S \in \mathbb{N}_n^b$, and $W = \{d_1, \dots, d_c\}^{\mathsf{c}}$, $\tau = \mathrm{id}$, $U = \{f_1, \dots, f_c\}^c$, $\sigma = \mathrm{id}$, and we have

$$\mathrm{in}(g_i) = \prod_{\{\ell |\ k_\ell \in Q\}} x_{f_\ell, d_\ell}^{(q_\ell)} \prod_{\{\ell |\ k_\ell \in K \setminus Q\}} x_{v_\ell, \ell}^{(k_\ell)}, \qquad \mathrm{in}(g_j) = \prod_{\{\ell |\ l_\ell \in Q\}} x_{f_\ell, d_\ell}^{(q_\ell)} \prod_{\{\ell |\ l_\ell \in L \setminus Q\}} x_{\ell, s_\ell}^{(l_\ell)}.$$

By Lemma 3.3.4 we deduce

$$\begin{aligned}
&\sum_{\lambda=b}^{n} (-1)^{\sum f_\ell} \sum_{W \subset \mathbb{N}_b, |W| = b-c} (-1)^{\sum_{W^{\mathsf{c}}} w}\, d^{\mathbf{c}}_{\lambda, W}(L_S \setminus Q) \\
&\qquad\qquad \sum_{\tau \in \mathrm{Sym}\,(W^{\mathsf{c}})_\lambda} (-1)^\tau \sum_{\alpha=a}^{n} D^{\mathbf{r}}_\alpha(Q_\tau, K_V \setminus Q) u_{\lambda, \alpha} \\
&\qquad = \sum_{\lambda=a}^{n} (-1)^{\sum d_\ell} \sum_{U \subset \mathbb{N}_a, |U| = a-c} (-1)^{\sum_{U^{\mathsf{c}}} u}\, d^{\mathbf{r}}_{\lambda, U}(K_V \setminus Q) \\
&\qquad\qquad\qquad \sum_{\sigma \in \mathrm{Sym}\,(U^{\mathsf{c}})_\lambda} (-1)^\sigma \sum_{\beta=b}^{n} D^{\mathbf{c}}_\beta(Q_\sigma, L_S \setminus Q) u_{\beta, \lambda}.
\end{aligned}$$

Indeed, restricting to the basis element $u_{\gamma,\delta} \in \mathcal{C}^{n^2}$ on both sides of this identity we get the identity (3.6) for $\alpha = \delta, \beta = \gamma$. It remains to prove that this identity induces a standard expression for σ_{ij}. It is enough to check that the initial terms of the elements in the Gröbner basis (except for g_i and g_j) that appear in this identity and involve the same basis element of $\mathcal{C}^{n^2}$ are smaller or equal to $\mathrm{in}(\sigma_{ij})$. Those are

$$\mathrm{in}\Big(d^{\mathbf{c}}_{\lambda, W}(L_S \setminus Q) D^{\mathbf{r}}_a(Q_\tau, K_V \setminus Q)\Big) \leq \prod_{\{\ell |\ l_\ell \in L \setminus Q\}} x_{w_l, s_l}^{(l_\ell)} \prod_{\{\ell |\ k_\ell \in K \setminus Q\}} x_{v_\ell, \ell}^{(k_\ell)} \prod_{\{\ell |\ k_\ell \in Q\}} x_{\tau(w'_\ell), d_\ell}^{(q_\ell)}, \tag{3.7}$$

$$\mathrm{in}\Big(d^{\mathbf{r}}_{\lambda, U}(K_V \setminus Q) D^{\mathbf{c}}_b(Q_\sigma, L_S \setminus Q)\Big) \leq \prod_{\{\ell |\ k_\ell \in K \setminus Q\}} x_{v_l, u_l}^{(k_\ell)} \prod_{\{\ell |\ l_\ell \in L \setminus Q\}} x_{\ell, s_\ell}^{(l_\ell)} \prod_{\{\ell |\ l_\ell \in Q\}} x_{f_\ell, \sigma(u'_\ell)}^{(q_\ell)}, \tag{3.8}$$

where $W = \{w_1, \dots, w_{b-c}\} \subset \mathbb{N}_b$, $W^{\mathsf{c}} = \{w'_1, \dots, w'_c\}$, $U = \{u_1, \dots, u_{a-c}\} \subset \mathbb{N}_a$, $U^{\mathsf{c}} = \{u'_1, \dots, u'_c\}$, $\tau \in \mathrm{Sym}\, W^{\mathsf{c}}$, $\sigma \in \mathrm{Sym}\, U^{\mathsf{c}}$, and we need to exclude $W = \{d_1, \dots, d_c\}^{\mathsf{c}}$, $\tau = \mathrm{id}$, and $U = \{f_1, \dots, f_c\}^c$, $\sigma = \mathrm{id}$. Note that the equalities hold for $\lambda = b$ (resp. $\lambda = a$).

One easily infers that $\mathrm{in}(\sigma_{ij})$ equals the product

$$\prod_{\{\ell |\ k_\ell \in Q\}} x_{f_\ell, d_\ell}^{(q_\ell)} \prod_{\{\ell |\ k_\ell \in K \setminus Q\}} x_{v_\ell, \ell}^{(k_\ell)} \prod_{\{\ell |\ l_\ell \in L \setminus Q\}} x_{\ell, s_\ell}^{(l_\ell)}, \tag{3.9}$$

in which we replace the factor $x_{v_{a-1}, a-1}^{(k_{a-1})} x_{v_a, a}^{(k_a)}$ (resp. $x_{b-1, s_{b-1}}^{(l_{b-1})} x_{b, s_b}^{(l_b)}$) by $x_{v_a, a-1}^{(k_{a-1})} x_{v_{a-1}, a}^{(k_a)}$ (resp. $x_{b-1, s_b}^{(l_{b-1})} x_{b, s_{b-1}}^{(l_b)}$) if $k_{a-1} \geq l_{b-1}$ (resp. if $k_{a-1} < l_{b-1}$). The terms in (3.7) are obtained by permuting the indices corresponding to the rows of the elements $x_{\ell, s_\ell}^{(l_\ell)}$ appearing in (3.9) (notice that the elements $x_{f_\ell, d_\ell}^{(q_\ell)}$ are also of that form), while the terms in (3.8) are obtained by permuting the indices corresponding to the columns of the elements $x_{v_\ell, \ell}^{(k_\ell)}$ appearing in (3.9) (notice that the elements $x_{f_\ell, d_\ell}^{(q_\ell)}$ are also of that form). Since the permutation described in order to obtain the initial term of σ_{ij} leads to the biggest monomial among the monomials in (3.7), (3.8) with respect to the given order $>$ on $\mathcal{C}$, σ_{ij} has a standard expression, which concludes the proof. $\square$

We will need a slight generalization of Lemma 3.3.5. Let $K = \{k_1, \dots, k_a\} \subset \mathbb{N}_m$, $L = \{l_1, \dots, l_b\} \subset \mathbb{N}_m$. We denote

$$\Xi^{(KL)} = (\xi'_{k_1}, \dots, \xi'_{k_a}, \xi''_{l_1}, \dots, \xi''_{l_b}),$$

and write $M^{(KL)}$ for the module generated by the rows of $\Xi^{(KL)}$. Let $G'^{(K)} \subseteq G'$ be a multilinear Gröbner basis on the rows of $(\xi'_{k_1}, \dots, \xi'_{k_a})$, and $G''^{(L)} \subseteq G''$ be a multilinear Gröbner basis on the rows of $(\xi''_{l_1}, \dots, \xi''_{l_b})$. We state the next lemma without proof since one only needs to inspect the proof of Lemma 3.3.5, and notice that it carries over to a more general situation of the following lemma.

Lemma 3.3.6. *The set $G'^{(K)} \cup G''^{(L)}$ is a multilinear Gröbner basis of $M^{(KL)}$.*

3.3.3. *Reduction to standard and one-sided identities.* We are now in a position to establish our main result on two-sided functional identities, which together with Theorem 3.2.4 gives a full description of functional identities on $M_n(F)$.

Theorem 3.3.7. *Let $m \geq 2$ and $K, L \subseteq \mathbb{N}_m$. Every solution of the functional identity*

$$\sum_{k \in K} F_k(\overline{x}_m^k) x_k = \sum_{l \in L} x_l G_l(\overline{x}_m^l)$$

on $M_n(F)$ is standard modulo one-sided identities.

Proof. The functional identity in question can be written coordinate-wise as a system of equations. We will treat the situation where $K = L = \mathbb{N}_m$. Finding all solutions of our functional identity in this case, we will obtain the desired ones among those with $F_k = 0$ for $k \in \mathbb{N}_m \setminus K$, $G_l = 0$ for $l \in \mathbb{N}_m \setminus L$. For this let us first denote

$$F'_k = (F_{11}^{(k)}, \dots, F_{1n}^{(k)}, \dots, F_{n1}^{(k)}, \dots, F_{nn}^{(k)}), \quad G''_l = (G_{11}^{(l)}, \dots, G_{1n}^{(l)}, \dots, G_{n1}^{(l)}, \dots, G_{nn}^{(l)}),$$

$$H = \big(F'_1, \dots, F'_m, G''_1, \dots, G''_m\big).$$

Then the system of equations reads as

$$H\Xi = \big(F'_1, \dots, F'_m, G''_1, \dots, G''_m\big) \begin{pmatrix} \xi'_1 \\ \vdots \\ \xi'_m \\ \xi''_1 \\ \vdots \\ \xi''_m \end{pmatrix} = 0 \tag{3.10}$$

This system can thus be interpreted as a syzygy on the rows of the matrix Ξ.

In our case F_k, G_l are multilinear so we can restrict ourselves to the syzygies on the rows of Ξ which yield multilinear functions in $x_1, \dots, x_m$. These are generated by τ_{ij} for i, j such that g_i, g_j belong to a multilinear Gröbner basis G of M. By Lemma 3.3.5 we have $G = G' \cup G''$. If we take elements $g_i, g_j \in G'$ (resp. $g_i, g_j \in G''$), then $\sigma_{ij} = m_{ji} g_i - m_{ij} g_j$ can be expressed in terms of $g_\ell \in G'$ (resp. $g_\ell \in G''$), since G' (resp. G'') is a (multilinear) Gröbner basis of the module generated by the first (resp. last) mn^2 rows of Ξ. These τ_{ij} thus yield the one-sided functional identities. It remains to consider τ_{ij} for $g_i \in G'$, $g_j \in G''$. Both elements g_i, g_j are multilinear of degree at most n. Let g_i involve the variables appearing in ξ_k for $k \in K' \subset \mathbb{N}_m$, $|K'| \leq n$, and g_j those appearing in ξ_l for $l \in L' \subset \mathbb{N}_m$, $|L'| \leq n$. We can treat g_i, g_j as the elements of the Gröbner basis on the rows of the matrix

$\Xi^{(K'L')}$. By Lemma 3.3.6, σ_{ij} can be expressed in terms of those g_ℓ that form a Gröbner basis on the $\Xi^{(K'L')}$, which implies that the functional identity of the form

$$\sum_{k\in K'} F_k(\overline{x}_m^k)x_k = \sum_{l\in L'} x_l G_l(\overline{x}_m^l)$$

corresponds to the syzygy τ_{ij}. Since $|K'|, |L'| \leq n$, this identity has standard solutions by [**BCM07**, Corollary 2.23]. □

3.4. An application: Commuting traces. Let A be an algebra over a field F, and let $T : A \to A$ be the trace of an r-linear function $t : A^r \to A$ (see Subsection 3.2.4). We say that T is *commuting* if $[T(x), x] = 0$ for all $x \in A$. Such a map is said to be of a *standard form* if there exist traces of $(r-i)$-linear functions $\mu_i : A \to F$, $0 \leq i \leq r$, such that $T(x) = \sum_{i=0}^r \mu_i(x)x^i$ for all $x \in A$. The question whether every commuting trace of an r-linear function on A is of a standard form has been studied extensively for different classes of algebras A (see [**Bre04**, **BCM07**] for surveys), but, paradoxically, for the basic case where $A = M_n(F)$ this question remained open. Our goal now is to show that Theorem 3.3.7 can be used to fill this gap.

Corollary 3.4.1. *Let $T : M_n(F) \to M_n(F)$ be a commuting trace of an r-linear function t. If* $\mathrm{char}(F) = 0$ *or* $\mathrm{char}(F) > r+1$, *then T is of a standard form.*

Proof. The proof is by induction on r. The $r = 0$ case is trivial, so we may assume that $r > 0$ and that the result holds for $r - 1$.

Note that the complete linearization of $[T(x), x] = 0$ yields the functional identity

$$\sum_{k=1}^m F(\overline{x}_m^k)x_k = \sum_{l=1}^m x_l F(\overline{x}_m^l) \quad \text{for all } x_1 \ldots, x_m \in M_n(F),$$

where $m = r + 1$ and

$$F(x_1, \ldots, x_r) = \sum_{\sigma\in S_r} t(x_{\sigma(1)}, \ldots, x_{\sigma(r)}).$$

Applying Theorem 3.3.7 we see that for each $k = 1, \ldots, m = r+1$ there exist functions

$$\begin{aligned}
&\varphi_k : M_n(F)^{m-1} \to M_n(F),\ k \in \mathbb{N}_m,\\
&p_{kl} : M_n(F)^{m-2} \to M_n(F),\ k, l \in \mathbb{N}_m,\ k \neq l,\\
&\lambda_k : M_n(F)^{m-1} \to F,\ k \in \mathbb{N}_m,
\end{aligned}$$

such that

$$\sum_{k=1}^m \varphi_k(\overline{x}_m^k)x_k = 0 \tag{3.11}$$

and

$$F(\overline{x}_m^k) = \sum_{l\in\mathbb{N}_m,\, l\neq k} x_l p_{kl}(\overline{x}_m^{kl}) + \lambda_k(\overline{x}_m^k) + \varphi_k(\overline{x}_m^k).$$

Setting $x_1 = \cdots = x_m = x$ and using $F(x, \ldots, x) = r!T(x)$ we arrive at

$$T(x) = xP_k(x) + \Lambda_k(x) + \Phi_k(x), \quad 1 \leq k \leq m, \tag{3.12}$$

where P_k is the trace of $\frac{1}{r!}\sum_{l\in\mathbb{N}_m,\, l\neq k} p_{kl}$, Λ_k is the trace of $\frac{1}{r!}\lambda_k$, and Φ_k is the trace of $\frac{1}{r!}\varphi_k$. Note that (3.11) yields $\sum_{k=1}^m \Phi_k(x)x = 0$ for all $x \in M_n(F)$. Interpreting this identity as $A\xi = 0$ where ξ is a generic matrix and A is the matrix corresponding to $\sum_{k=1}^m \Phi_k$, it follows (since ξ is invertible) that $A = 0$, and hence that

$$\sum_{k=1}^{m} \Phi_k = 0. \tag{3.13}$$

Further, (3.12) shows that

$$\Phi_1(x) - \Phi_\ell(x) = x\big(P_\ell(x) - P_1(x)\big) + \Lambda_\ell(x) - \Lambda_1(x), \quad 2 \le \ell \le m.$$

Summing up these identities and using (3.13) we get

$$m\Phi_1(x) = x\Big(\sum_{\ell=2}^{m} P_\ell(x) - (m-1)P_1(x)\Big) + \sum_{\ell=2}^{m} \Lambda_\ell(x) - (m-1)\Lambda_1(x). \tag{3.14}$$

Accordingly,

$$\Phi_1(x) - x\tilde{P}_1(x) \in F \quad \text{for all } x \in M_n(F),$$

where $\tilde{P}_1 = \frac{1}{m}\Big(\sum_{\ell=2}^m P_\ell - (m-1)P_1\Big)$. Setting $P = P_1 + \tilde{P}_1$ we thus see from (3.12) that

$$T(x) - xP(x) \in F \quad \text{for all } x \in M_n(F). \tag{3.15}$$

Since T is commuting it follows that $[xP(x), x] = 0$ for every $x \in M_n(F)$, which can be written as $x[P(x), x] = 0$. Interpreting this identity through a generic matrix we see, similarly as above, that $[P(x), x] = 0$. Thus, P is a commuting trace of an $(r-1)$-linear function. By induction assumption, P is of a standard form. From (3.15) we thus see that T is of a standard form, too. □

A more careful analysis is needed if one wishes to further weaken the assumption on $\mathrm{char}(F)$. Let us examine only the case where $r = 2$, which is the one that plays the most prominent role in applications of functional identities. It naturally appears in the study of Lie isomorphisms, commutativity preserving maps, Lie-admissible maps, Poisson algebras, and several other topics (see [**Bre04**] and [**BCM07**, Section 1.4]).

We will thus consider the case where T is a commuting trace of a bilinear map F. We have to add the assumption that F satisfies the functional identity

$$F(x,y)z + F(z,x)y + F(y,z)x = zF(x,y) + yF(z,x) + xF(y,z). \tag{3.16}$$

One way of looking at (3.16) is that $(x, y) \mapsto F(x, y)$ is a nonassociative commutative product on $M_n(F)$ which is connected with the ordinary product through a version of the Jacobi identity: $[F(x, y), z] + [F(z, x), y] + [F(y, z), x] = 0$.

Note that (3.16) is actually equivalent to T being commuting provided that F is symmetric and $\mathrm{char}(F) \neq 2, 3$. However, we do not wish to impose these assumptions. The point of the next corollary is that it has no restrictions on $\mathrm{char}(F)$.

Corollary 3.4.2. *Let T be the trace of a bilinear function $F : M_n(F)^2 \to M_n(F)$. If T is commuting and F satisfies* (3.16)*, then T is of a standard form.*

Proof. If $\mathrm{char}(F) \neq 2$, then the result follows from [**BCM07**, Theorem 5.32 and Remark 5.33]. Assume, therefore, that $\mathrm{char}(F) = 2$. From now on we simply follow the proof of Corollary 3.4.1. Thus, we first derive (3.12) with $m = 3$, and after that (3.13) and (3.14). Note that $m\Phi_1(x) = \Phi_1(x)$ since $m = 3$. Therefore, (3.15) follows with P being a linear function. This implies that P is commuting, and hence is of a standard form [**BCM07**, Corollary 5.28]. Consequently, T is of a standard form as well. □

These corollaries can be extended, by using scalar extensions and other results on functional identities, to considerably more general algebras than $M_n(F)$. Also, some other special identities that have proved to be useful for applications could now be examined in greater detail. However, we only wanted to give an indication of the applicability of Theorem 3.3.7 and will not go further with addressing these questions.

4. Functional identities in one variable

In this section we describe functional identities in one variable. In the preceding section, see Corollary 3.4.1, we described commuting traces on $M_n(F)$. Here we present an alternative approach to commuting traces using some standard Commutative algebra facts, and show how the general case follows by reduction to this special case. We assume that $\mathrm{char}(F) = 0$.

4.1. Commutative algebra preliminaries. Let $\mathcal{C} = F[x_{ij} \mid 1 \leq i, j \leq n]$ be the polynomial algebra, and let $\mathcal{K}$ be its quotient field. We write ξ for the generic $n \times n$ matrix (x_{ij}), and $c(t)$ for the characteristic polynomial of ξ, i.e., $c(t) = \det(t - \xi) \in \mathcal{C}[t]$ (t We will use the same notation for scalars and scalar multiples of unity.

LEMMA 4.1.1. *The polynomial $c(t)$ is irreducible in $\mathcal{C}[t]$. Accordingly, $\mathcal{C}[t]/(c(t))$ is an integral domain.*

PROOF. We consider $c = c(t)$ as a polynomial in indeterminates $t, \xi_{ij}, 1 \leq i, j \leq n$. Suppose $c = fg$ for some $f, g \in F(\xi)[t]$ of positive degree in t. Then c is reducible already in $\mathcal{C}[t]$ (see, e.g., [**CLO07**, Corollary 3.5.4]), so we may assume that $f, g \in \mathcal{C}[t]$. Since c has degree n, f and g have to be of degree strictly less than n. Regarding c, f, g as polynomials in t, the constant term of c equals the product of the constant terms of f and g. Since the former equals $(-1)^n \det(\xi)$, which is irreducible [**Ber06**, pp. 7–8], and none of f and g can be constant due to the degree restriction, we have arrived at a contradiction. □

LEMMA 4.1.2. *The algebra $\mathcal{C}[t]/(c(t))$ is isomorphic to the subalgebra $\mathcal{C}[\xi]$ of the algebra $M_n(\mathcal{C})$.*

PROOF. Define a homomorphism from $\mathcal{C}[t]/(c(t))$ to $\mathcal{C}[\xi]$ according to $x_{ij} + (c(t)) \mapsto x_{ij},\ t + (c(t)) \mapsto \xi$. Since $\xi = (x_{ij})$ is a zero of its characteristic polynomial, this homomorphism is well-defined. It is obviously surjective. We can extend it to a homomorphism from $\mathcal{K}[t]/(c(t))$ to $M_n(\mathcal{K})$ which has trivial kernel as $\mathcal{K}[t]/(c(t))$ is a field by Lemma 4.1.1. □

Recall that an integral domain R is said to be integrally closed if it is integrally closed in its field of fractions Q. That is, if f is a monic polynomial with coefficients in R, then every root of f from Q actually lies in R.

PROPOSITION 4.1.3. *The algebra $\mathcal{C}[\xi]$ is integrally closed.*

PROOF. In view of Lemma 4.1.2, we have to show that $\mathcal{C}[t]/(c(t))$ is integrally closed. To this end, it is enough to prove that $B = \mathcal{C}/(\det(\xi))$ is integrally closed. Indeed, let us set $x'_{ii} = x_{ii} - t$ and $x'_{ij} = x_{ij}$ for $i \neq j$. Then $\xi - t$ is a generic matrix with entries x'_{ij}. If B is integrally closed, then so is $B' = k[x'_{ij} \mid 1 \leq i, j \leq n]/(\det(x'_{ij}))$, and therefore also $\mathcal{C}[t]/(c(t)) = B'[t] = k[x_{ij}, t \mid 1 \leq i, j \leq n]/(\det(x'_{ij}))$.

Let $D = k[\eta_{ij}, \mu_{ji} \mid 1 \leq i \leq n, 1 \leq j \leq n-1]$. We define the action of GL_{n-1} on D so that $\sigma \in \mathrm{GL}_{n-1}$ acts on an $n \times (n-1)$ matrix by $(a_{ij})^\sigma = (a_{ij})\sigma$, and on an $(n-1) \times n$ matrix by $(b_{ij})^\sigma = \sigma^{-1}(b_{ij})$. By the first fundamental theorem for the general linear group (see, e.g., [**KP96**, Section 2.1.]), the algebra of invariants $D^{\mathrm{GL}_{n-1}}$ is generated by the entries of the matrix $(\eta_{ij})(\mu_{ij})$. These entries parametrize the variety of matrices of rank at most $n-1$. Therefore the algebra $D^{\mathrm{GL}_{n-1}}$ is isomorphic to the coordinate ring $B = \mathcal{C}/(\det(\xi))$ of this variety. Thus, we can regard B as a subalgebra $D^{\mathrm{GL}_{n-1}}$ of D. If a lies in the field of fractions of B and is integral over B, then a lies in the field of fractions of D and is integral over D. But D is a polynomial algebra, so $a \in D$. Hence $a \in D^{\mathrm{GL}_{n-1}}$, i.e., a lies in the invariant subalgebra, namely B. $\square$

4.2. Commuting traces. Let $q : M_n(F) \to M_n(F)$ be the trace of a multilinear map. Identifying $M_n(F)$ with F^{n^2}, we can regard q as a homogeneous polynomial map from F^{n^2} to F^{n^2}. The identity $[q(x), x] = 0$ for every $x \in M_n(F)$ can be thus interpreted as $[q_\xi, \xi] = 0$ where ξ is, as above, a generic matrix, and q_ξ is the matrix in $M_n(\mathcal{C})$ corresponding to q. In the next lemma we will show that the centralizer $\mathrm{Cent}(\xi)$ of ξ in $M_n(\mathcal{C})$ is trivial. As we shall see, from this it can be easily derived that q is of a standard form. We remark that if $d \geq n$, then a standard form $\sum_{i=0}^{d} \mu_i(x)x^i$ can be in $M_n(F)$ written as $\sum_{i=0}^{n-1} \mu_i'(x)x^i$. This follows from the Cayley-Hamilton theorem.

Lemma 4.2.1. $\mathrm{Cent}(\xi) = \mathcal{C}[\xi]$.

Proof. Since ξ is a generic matrix, it has n distinct eigenvalues (see, e.g., [**Row80**, Lemma 1.3.12]). Its centralizer in $M_n(\mathcal{K})$ is therefore the $\mathcal{K}$-linear span of its powers. Thus, given $a \in \mathrm{Cent}(\xi)$, we have $a \in \mathcal{K}[\xi]$. By Lemma 4.1.2, $\mathcal{C}[\xi]$ is integrally closed. Since a is integral over $\mathcal{C}$ by the Cayley-Hamilton theorem and $\mathcal{K}[\xi]$ is the field of fractions of $\mathcal{C}[\xi]$, a thus belongs to $\mathcal{C}[\xi]$. This shows that $\mathrm{Cent}(\xi) \subseteq \mathcal{C}[\xi]$. The converse inclusion is trivial. $\square$

Lemma 4.2.2. *Every commuting trace of a multilinear map $q : M_n(F) \to M_n(F)$ is of a standard form.*

Proof. Since $q_\xi \in \mathrm{Cent}(\xi)$, Lemma 4.2.1 shows that

$$q_\xi = \sum_{i=0}^{n-1} p_i(x_{11}, x_{12}, \ldots, x_{nn})\xi^i$$

for some polynomials $p_i \in \mathcal{C}$. Since q is the trace of a d-linear map (for some $d \in \mathbb{N}$), q_ξ is homogeneous of degree d and therefore $\deg(p_i) = d-i$ (if $d < i$, then we simply take $p_i = 0$). Note that this is just another way of saying that p_i is the trace of a $(d-i)$-linear central map. Hence q is of a standard form. $\square$

In the proof of the next theorem we combine Lemma 4.2.2 with results from [**LWLW97**].

Theorem 4.2.3. *If A is a centrally closed prime F-algebra and $q : A \to A$ is a commuting trace of a multilinear map, then q is of a standard form.*

Proof. Let $M : A^d \to A$ be d-linear map such that $q(x) = M(x, \ldots, x)$, $x \in A$. In light of [**LWLW97**, Theorem 3.1] we may assume that A satisfies the standard polynomial identity S_{2d}. Since A is assumed to be centrally closed, this simply means that A is a central simple algebra over F of dimension at most d^2 (see, e.g., [**BCM07**, Theorem C.2]). Let $K \supseteq F$ be a splitting field of A. The scalar extension $A_K = A \otimes_F K$ of A to K is then isomorphic to $M_n(K)$ for some $n \leq d$. A standard argument shows that M can be extended to a d-linear (with respect to K) map $\tilde{M} : A_K^d \to A_K$ (so that $\tilde{M}(x_1 \otimes 1, \ldots, x_d \otimes 1) = M(x_1, \ldots, x_d) \otimes 1$). Let $\tilde{q}(x) := \tilde{M}(x, \ldots, x)$ be its trace. It is easy to verify that $\tilde{q}$ is commuting, and thus, by Lemma 4.2.2, of a standard form $\tilde{q}(x) = \sum_{i=0}^{n-1} \tilde{\mu}_i(x)x^i$, where $\tilde{\mu}_i$ is the trace of a $(d-i)$-linear map from A_K to K (i.e., a homogeneous polynomial of degree $d-i$). It remains to show that $\tilde{\mu}_i(A\otimes 1)\subseteq F$ to conclude that the restriction M of $\tilde{M}$ to $A\otimes 1$ is of a standard form. By [**LWLW97**, Theorem 1.4], $q(x)$ lies in the F-linear span of $1, x, \dots, x^{n-1}$. If the minimal polynomial of $a\in A$ has degree n, then it follows that $\tilde{\mu}_i(a)\in F$. We now follow the argument from the proof of [**Row80**, Theorem 3.1.49]. Let $h(x_1,\dots,x_n)=C_{2n-1}(1,x_1,\dots,x_1^{n-1},x_2,\dots,x_n)$, where C_{2n-1} is the Capelli polynomial. If $h(a,x_2,\dots,x_n)$ does not vanish on A, then the minimal polynomial of a has degree n (see, e.g., [**Row80**, Theorem 1.4.34]), and so in this case $\tilde{\mu}_i(a)\in F$. Since h is not an identity of A (see, e.g., [**Row80**, Proposition 3.1.6]), there exists $a\in A$ such that $h(a,x_2,\dots,x_n)$ is not a generalized polynomial identity. Take $b\in A$. A standard Vandermonde argument shows that $h(a+\alpha b,x_2,\dots,x_n)$ is a generalized polynomial identity for only finitely many $\alpha\in F$. Therefore $\tilde{\mu}_i(a+\alpha b)\in F$ for infinitely many $\alpha\in F$, and hence, again by a Vandermonde argument, $\tilde{\mu}_i(b)\in F$. □

4.3. Reduction to the commuting trace case. The goal of this section is to prove a lemma which will be used as an essential tool for reducing the study of general functional identities in one variable to commuting traces of multilinear maps.

Let A be a unital algebra over F. If $q: A\to A$ is the trace of a d-linear map $M: A^d\to A$, then we may assume that M is symmetric (otherwise we replace it by the map $\frac{1}{d!}\sum_{\sigma\in S_d} M(x_{\sigma(1)},\dots,x_{\sigma(d)})$). Then M is unique, which makes it possible for us to define $\partial q: A\to A$ by

$$\partial q(x)=M(x,\dots,x,1).$$

Obviously, ∂q is the trace of the symmetric $(d-1)$-linear map $M(x_1,\dots,x_{d-1},1)$. Alternatively, one can introduce ∂q by avoiding M as follows. First notice that $q(x+\lambda)$, where $x\in A$ and $\lambda\in F$, is a polynomial in λ with coefficients in A of degree at most d. The coefficient at λ is equal to $d\partial q(x)$. In fact, $q(x+\lambda)=q(x)+d\partial q(x)\lambda+\cdots+q(1)\lambda^d$ (and therefore, in a suitable setting, we have $\partial q(x)=\frac{1}{d}\lim_{\lambda\to 0}\frac{q(x+\lambda)-q(x)}{\lambda}$). Using this alternative definition, one easily checks that if q' is another trace of, say, d'-linear map, then we have

$$\partial\big(q(x)\cdot q'(x)\big)=\frac{d}{d+d'}\partial q(x)\cdot q'(x)+\frac{d'}{d+d'}q(x)\cdot\partial q'(x). \tag{4.1}$$

In the next lemma we consider the situation where the traces of d-linear maps q and r satisfy $[q(x)x-xr(x),x]=0$, in other words, the situation where $x\mapsto q(x)x-xr(x)$ is a commuting map. We are actually interested in the special case where $q(x)x-xr(x)\in F$. However, the proof runs more smoothly if we consider a slightly more general situation.

Lemma 4.3.1. *Let A be a unital F-algebra and let $q,r: A\to A$ be the traces of d-linear maps. If $[q(x)x-xr(x),x]=0$ for all $x\in A$, then there exists the trace of a $(d-1)$-linear map $p: A\to A$ such that $[[q(x)-xp(x),x],x]=0$ for all $x\in A$.*

Proof. We set $q_0=q$ and $q_t=\partial q_{t-1}$, $t=1,\dots,d$. If M is as above, then we have $q_t(x)=M(x,\dots,x,1,\dots,1)$ where x appears $d-t$ times and 1 appears t times. Analogously we introduce r_t, $t=0,\dots,d$. For the traces of multilinear maps f and g we will write $f(x)\equiv g(x)$ if $[f(x)-g(x),x]=0$ for every $x\in A$ (i.e., $f-g$ is commuting). In this case we also have $\partial f(x)\equiv\partial g(x)$. Indeed, this follows immediately by applying (4.1) to $f(x)x-g(x)x=xf(x)-xg(x)$. Accordingly, $q(x)x\equiv xr(x)$ implies $dq_1(x)x+q_0(x)\equiv dxr_1(x)+r_0(x)$, and, furthermore, by induction one easily checks that

$$(d-t)q_{t+1}(x)x+(t+1)q_t(x)\equiv(d-t)xr_{t+1}(x)+(t+1)r_t(x) \tag{4.2}$$

for $t = 0, 1, \ldots, d-1$. Next, we claim that

$$[q(x), x] \equiv (-1)^{t-1}\binom{d}{t}x(q_t - r_t)(x)x^t + \sum_{i=1}^{t}(-1)^{i-1}\binom{d}{i}x[r_i(x), x]x^{i-1} \tag{4.3}$$

for $t = 1, \ldots, d$. Indeed, from $q(x)x \equiv xr(x)$ we get $[q(x), x] \equiv x(r - q)(x)$, and hence, using (4.2) for $t = 0$, $[q(x), x] \equiv dx\big(q_1(x)x - xr_1(x)\big)$. Consequently,

$$[q(x), x] \equiv dx(q_1 - r_1)(x)x + dx[r_1(x), x],$$

which proves (4.3) for $t = 1$. Now assume that $1 < t < d$. Using (4.2) we obtain

$$\begin{aligned} & (-1)^{t-1}\binom{d}{t}x(q_t - r_t)(x)x^t \\ \equiv\;& (-1)^{t-1}\binom{d}{t}\frac{d-t}{t+1}x\big(xr_{t+1}(x) - q_{t+1}(x)x\big)x^t \\ =\;& (-1)^t\binom{d}{t+1}x[r_{t+1}(x), x]x^t + (-1)^t\binom{d}{t+1}x(q_{t+1} - r_{t+1})(x)x^{t+1}. \end{aligned}$$

This readily implies (4.3) for any t.

Note that $q_d(x) = q(1)$ and $r_d(x) = r(1)$. Thus, the $t = d-1$ case of (4.2) reads as $q(1)x + dq_{d-1}(x) \equiv xr(1) + dr_{d-1}(x)$. Substituting 1 for x in this relation we obtain $(d+1)q(1) \equiv (d+1)r(1)$. That is, $q_d(x) \equiv r_d(x)$, and so the $t = d$ case of (4.3) reduces to

$$[q(x), x] \equiv \sum_{i=1}^{d}(-1)^{i-1}\binom{d}{i}x[r_i(x), x]x^{i-1}.$$

Since $x[r_i(x), x]x^{i-1} = [xr_i(x)x^{i-1}, x]$, we can rewrite this as

$$[q(x), x] \equiv \Big[\sum_{i=1}^{d}(-1)^{i-1}\binom{d}{i}xr_i(x)x^{i-1}, x\Big].$$

Thus,

$$p(x) := \sum_{i=1}^{d}(-1)^{i-1}\binom{d}{i}r_i(x)x^{i-1}$$

satisfies $[[q(x) - xp(x), x], x] = 0$. Note also that p is indeed the trace of a $(d-1)$-linear map. $\square$

4.4. An Engel condition with traces of multilinear maps. Lemma 4.3.1 gives rise to the question what can be said about the trace of a multilinear map s if it satisfies $[[s(x), x], x] = 0$. Fortunately, an answer comes easily if we combine our arguments with an old lemma by Jacobson **[Jac35]** saying that if $a, b \in M_n(F)$ are such that $[[b, a], a] = 0$, then $[b, a]$ is nilpotent.

PROPOSITION 4.4.1. *Let A be a centrally closed prime F-algebra. If $s : A \to A$ is the trace of a multinear map such that $[[s(x), x], x] = 0$ for all $x \in A$, then $[s(x), x] = 0$ for all $x \in A$ (and hence s is of a standard form $s(x) = \sum_i \mu_i(x)x^i$).*

PROOF. If A is infinite dimensional, then this follows easily from the general theory of functional identities **[BCM07]**; more explicitly, it is a special case of **[Bei98**, Theorem 4.6]. We may therefore assume that A is finite dimensional, in which case it is a central simple algebra. A standard scalar extension argument shows that with no loss of generality we may assume that $A = M_n(F)$ (cf. the proof of Theorem 4.2.3). Jacobson's lemma **[Jac35**, Lemma 2] implies that $\hat{s}(x) := [s(x), x]$ is nilpotent for every $x \in A$. As in Section 4.2, we may identify $\hat{s}$ with an element of $\mathcal{C}[\xi]$. Since, by Lemmas 4.1.1 and 4.1.2, $\mathcal{C}[\xi]$ is an integral domain, it does not contain nonzero nilpotents. Hence $\hat{s} = 0$. $\square$

In a seemingly more general, but in fact equivalent way we could formulate Proposition 4.4.1 so that the condition $[[s(x), x], x] = 0$ is replaced by an Engel condition $[\dots[[s(x), x], x], \dots, x] = 0$ (the number of brackets is arbitrary, but fixed). We remark that such a condition has been studied extensively in the case where s is a derivation or a related map (see, e.g., [**Lan93b**]), and also in the case where s is a general additive map [**BFLW97**, **Bre96**]. Proposition 4.4.1 thus takes a step further in this line of investigation.

4.5. A functional identity determined by the adjugation. Let A be a central simple F-algebra with $\dim A = n^2$. We define the trace of an element $a \in A$ as the trace of $a \otimes 1 \in A \otimes K \cong M_n(K)$, where K is a splitting field of A. It can be shown that the trace of a belongs to F and that this definition is independent of the choice of K [**Row80**, Theorem 3.1.49]. Hence, by using the fact that the coefficients of the characteristic polynomial of a matrix can be expressed through the traces of the powers of this matrix, one can also define $\mathrm{adj}(x)$, the *adjugate* of x. The map $x \mapsto \mathrm{adj}(x)$ is the trace of an $(n-1)$-linear map, in fact $\mathrm{adj}(x) = \sum_{i=0}^{n-1} \tau_i(x)x^i$ where τ_i is the trace of an $(n-1-i)$-linear map from A into F. Just as for matrices, we have

$$\mathrm{adj}(x^m) = \mathrm{adj}(x)^m$$

and

$$x\mathrm{adj}(x) = \mathrm{adj}(x)x = \det(x) \in F,$$

where $\det(x)$ can be also defined through the traces of powers of x. The goal of this section is to show that the functional identity $x^m q(x) \in F$ always arises from the adjugation.

PROPOSITION 4.5.1. *Let A be a centrally closed prime F-algebra, let $q \neq 0$ be the trace of a d-linear map, and let $m \in \mathbb{N}$. If $x^m q(x) \in F$ for all $x \in A$, then A is finite dimensional. Moreover, if $\dim A = n^2$, then $d \geq m(n-1)$ and there exists the trace of a $(d - m(n-1))$-linear map $\lambda : A \to F$, $\lambda \neq 0$, such that $q(x) = \lambda(x)\mathrm{adj}(x^m)$ for all $x \in A$. (In particular, $x^m q(x) \neq 0$ for some $x \in A$.)*

PROOF. We first treat the case where $A \cong M_n(F)$. Then, as in Section 4.2, q can be identified with $q_\xi \in M_n(\mathcal{C})$ and the condition $q(x)x^m \in F$ for every $x \in F$ reads as $q_\xi \xi^m = \alpha$ where $\alpha \in \mathcal{C}$. Multiplying by $\mathrm{adj}(\xi^m)$ from the right we obtain

$$\det(\xi)^m q_\xi = \alpha \mathrm{adj}(\xi^m). \tag{4.4}$$

Thus, $q_\xi \in \mathcal{K}[\xi]$. Since q_ξ is integral over $\mathcal{C}[\xi]$ by the Cayley-Hamilton theorem, and since $\mathcal{C}[\xi]$ is integrally closed by Proposition 4.1.3, it follows that $q_\xi \in \mathcal{C}[\xi]$. Using the Cayley-Hamilton theorem again, we can write $q_\xi = \sum_{i=0}^{n-1} \mu_i \xi^i$ and $\mathrm{adj}(\xi^m) = \sum_{i=0}^{n-1} \nu_i \xi^i$ with $\mu_i, \nu_i \in \mathcal{C}$. From (4.4) we get $\sum_{i=0}^{n-1} \big(\det(\xi)^m \mu_i - \alpha\nu_i\big)\xi^i = 0$. Since there exist matrices whose degree of algebraicity is n, and since the linear dependence of matrices can be expressed through zeros of a (Capelli) polynomial [**Row80**, Theorem 1.4.34], a Zariski topology argument shows that $\det(\xi)^m \mu_i = \alpha\nu_i$ for each i. Note that $\det(\xi)$ does not divide ν_i for some i. Namely, otherwise we would have $\mathrm{adj}(\xi^m) = \det(\xi)r(\xi)$ for some $r \in \mathcal{C}[\xi]$, implying that $\mathrm{adj}(a^m) = 0$ whenever $a \in M_n(F)$ is not invertible. But one can easily find matrices for which this is not true (say, diagonal matrices with exactly one 0 on the diagonal). Fix i such that $\det(\xi)$ does not divide ν_i. Since $F[\xi]$ is a unique factorization domain and $\det(\xi)$ is irreducible [**Ber06**], $\det(\xi)^m \mu_i = \alpha\nu_i$ implies that $\alpha = \det(\xi)^m \lambda_0$ for some $\lambda_0 \in \mathcal{C}$. From (4.4) it follows that $q_\xi = \lambda_0 \mathrm{adj}(\xi^m)$. We can interpret this as $q(x) = \lambda(x)\mathrm{adj}(x^m)$ for all $x \in A$ where $\lambda : A \to F$. Since $\mathrm{adj}(x^m)$ is the trace of a an $m(n-1)$-linear map, a simple homogeneity argument shows that $d \geq m(n-1)$ and that λ is the trace of a $(d - m(n-1))$-linear map.

Consider now the general case where A is an arbitrary centrally closed prime algebra. From the general theory of functional identities it follows that A is finite dimensional (specifically, one can use [**BCM07**, Corollary 5.13]). From now on we argue similarly as in the proof of Theorem 4.2.3, so we omit some details. Take a splitting field $K \supseteq F$ of A, and consider the scalar extension $A_K = A \otimes_F K \cong M_n(K)$. We extend q to $\tilde{q} : A_K \to A_K$ (so that $\tilde{q}(x \otimes 1) = q(x) \otimes 1$) and observe that $y^m \tilde{q}(y) \in K$ for all $y \in A_K$. Using the result of the previos paragraph we readily arrive at the situation where $q(x) \otimes 1 = \mathrm{adj}(x^m) \otimes \lambda(x)$ with $\lambda : A \to K$. If $\mathrm{adj}(x^m) \neq 0$ then $\lambda(x) \in F$; since such elements x exist, a Vandermonde argument easily yields $\lambda(x) \in F$ for every $x \in A$. □

4.6. General functional identities in one variable. Our final theorem will be derived from all the main results of the previous subsections.

THEOREM 4.6.1. *Let A be a centrally closed prime F-algebra, and let $q_0, q_1, \ldots, q_m : A \to A$ be the traces of d-linear maps. Suppose that*

$$q(x) := q_0(x)x^m + xq_1(x)x^{m-1} + \cdots + x^m q_m(x) \in F \text{ for all } x \in A.$$

Then there exist the traces of $(d-1)$-linear maps $p_0, p_1, \ldots, p_{m-1} : A \to A$ and the traces of d-linear maps $\mu_0, \mu_1, \ldots, \mu_{m-1} : A \to F$ such that

$$\begin{aligned} q_0(x) &= xp_0(x) + \mu_0(x), \\ q_i(x) &= -p_{i-1}(x)x + xp_i(x) + \mu_i(x), \quad i = 1, \ldots, m-1, \end{aligned}$$

for all $x \in A$. Moreover:

(a) *If $q(x) = 0$ for all $x \in A$, then*

$$q_m(x) = -p_{m-1}(x)x - \sum_{i=0}^{m-1} \mu_i(x)$$

for all $x \in A$.

(b) *If $q(x) \neq 0$ for some $x \in A$, then A is finite dimensional. If $\dim A = n^2$, then $d \geq m(n-1)$ and there exists the trace of a $(d - m(n-1))$-linear map $\lambda : A \to F$ such that $\lambda \neq 0$ and*

$$q_m(x) = \lambda(x)\mathrm{adj}(x^m) - p_{m-1}(x)x - \sum_{i=0}^{m-1} \mu_i(x)$$

for all $x \in A$.

PROOF. Let us first show that q_0 is of the desired form. Writing $q(x)$ as $(q_0(x)x^{m-1})x - xr(x)$, where $r(x) = -q_1(x)x^{m-1} - \cdots - x^{m-1}q_m(x)$, enables us to apply Lemma 4.3.1 and hence conclude that there exists the trace of a $(d+m-2)$-linear map $p : A \to A$ such that $[[q_0(x)x^{m-1} - xp(x), x], x] = 0$ for all $x \in A$. Proposition 4.4.1 (together with Theorem 4.2.3) therefore tells us that $q_0(x)x^{m-1} - xp(x) = \sum_i \alpha_i(x)x^i$ with $\alpha_i : A \to F$. Rearranging the terms we can write this as $q_0(x)x^{m-1} - xr'(x) = \alpha_0(x) \in F$ where r' is another trace of a $(d+m-2)$-linear map. If $m = 1$, then we are done. Otherwise, interpret the last relation as $(q_0(x)x^{m-2})x - xr'(x) \in F$, and repeat the above argument. In a finite number of steps we arrive at $q_0(x) = xp_0(x) + \mu_0(x)$.

We can now rewrite $q(x) \in F$ as

$$x\Big(\big(p_0(x)x + q_1(x)\big)x^{m-1} + xq_2(x)x^{m-2} + \cdots + x^{m-1}\big(q_m(x) + \mu_0(x)\big)\Big) \in F.$$

From Proposition 4.5.1 it follows that

$$\big(p_0(x)x + q_1(x)\big)x^{m-1} + xq_2(x)x^{m-2} + \cdots + x^{m-1}\big(q_m(x) + \mu_0(x)\big) = \sum \beta_i(x)x^i$$

for some traces of multilinear maps $\beta_i : A \to F$ – in fact, the right-hand side is either 0 or it can be expressed through $\operatorname{adj}(x)$, but the form we have stated is sufficient for our purposes. Namely, we can interpret the last identity as

$$\Big(\big(p_0(x)x + q_1(x)\big)x^{m-2}\Big)x - xs(x) = \big(p_0(x)x + q_1(x)\big)x^{m-1} - xs(x) = \beta_0(x) \in F$$

for some trace of a multilinear map s, which brings us to the situation considered in the first paragraph. The same argument therefore gives $p_0(x)x + q_1(x) = xp_1(x) + \mu_1(x)$, i.e., q_1 is of the desired form. We can now continue this line of argumentation. In the next step we write $q(x) \in F$ as

$$x^2\Big(\big(p_1(x)x + q_2(x)\big)x^{m-2} + xq_3(x)x^{m-3} + \cdots + x^{m-2}\big(q_m(x) + \mu_0(x) + \mu_1(x)\big)\Big) \in F.$$

First applying Proposition 4.5.1 and then the argument from the first paragraph it thus follows that q_2 has the desired form. In this manner we describe the form of all q_i with $i < m$. Consequently, $q(x) \in F$ can be written as

$$x^m\Big(p_{m-1}(x)x + q_m(x) + \mu_0(x) + \cdots + \mu_{m-1}(x)\Big) \in F.$$

The final conclusion now follows immediately from Proposition 4.5.1. □

In the special case where $d = 1$ and $q(x) = 0$ this theorem was proved by Beidar [**Bei98**, Theorem 4.4]. In fact, he considered general prime rings and additive maps q_i. The problem to extend Theorem 4.6.1 to prime rings and multiadditive maps remains open.

CHAPTER 3

Trace rings

We have seen in the previous chapter that trace rings play an important role in studying identities on matrices. In this chapter we study some of their geometric properties.

In the first three sections we take a classical geometric viewpoint. In Section 1 we give a tracial Nullstellensatz for noncommutative polynomials evaluated at tuples of matrices of all sizes: Suppose $f_1, \ldots, f_r, f$ are noncommutative polynomials, and $\operatorname{tr}(f)$ vanishes whenever all $\operatorname{tr}(f_j)$ vanish. Then either 1 or f is a linear combination of the f_j modulo sums of commutators. In Section 2 we study the image of noncommutative polynomials and explore which subsets of $M_n(F)$ can be realized as the images of noncommutative polynomials. We especially focus on multilinear polynomials in view of Lvov's conjecture stating that the image of a multilinear polynomial is a vector space. We further study noncommutative Veronese type polynomial maps in Section 3 and prove a generic version of Paz conjecture; i.e, the length of a generic g-dimensional vector subspace of $M_n(F)$ is $2d = 2\lceil \log_g n \rceil$.

As trace rings are free algebras in the category of algebras with trace satisfying the identities of $n \times n$-matrices they can be seen as analogues of polynomial rings from a noncommutative geometry standpoint. In Section 4 we recover some of their known homological properties; i.e, we show that $\mathcal{T}_{2,3}\langle \xi_k \rangle$, $\mathcal{T}_{2,3}\langle \xi_k \rangle$ have finite global dimension. We further construct (twisted) noncommutative crepant resolutions of singularities of their centers $\mathcal{T}_{m,n}$.

This chapter is based on [**KŠ14b**, **Špe13**, **KŠ15**, **ŠVdB15**].

1. A tracial Nullstellensatz

We begin this chapter by giving a dimension-free tracial Nullstellensatz.

1.1. Preliminaries.

We first introduce some notation and present an effective Nullstellensatz which plays a crucial role in the proof of Spurnullstellensatz.

Let F be a field of characteristic 0 and let $\mathcal{M}(F)^g$ stand for $\bigcup_n M_n(F)^g$. We denote the free associative algebra in the variables $x_1, \ldots, x_g$ by $F\langle X \rangle$. The noncommutative polynomials in $F\langle X \rangle$ of degree at most d are denoted by $F\langle X \rangle_d$, while $F\langle X \rangle_d'$ is the vector subspace of $F\langle X \rangle_d$ consisting of all elements with zero constant term. We denote by $\langle X \rangle$ the monoid generated by $x_1, \ldots, x_g$, and by $\langle X \rangle_d$ words in $\langle X \rangle$ of degree at most d.

We say that polynomials $f, h \in F\langle X \rangle$ are *cyclically equivalent* if $f - h$ is a sum of commutators in $F\langle X \rangle$ and write $f \overset{\text{cyc}}{\sim} h$. Observe that $f \overset{\text{cyc}}{\sim} h$ if and only if $\operatorname{tr}(f(A)) = \operatorname{tr}(h(A)) = 0$ for all $n \in \mathbb{N}$ and all $A \in M_n(F)^g$ (see [**KS08**, Proposition 2.3] or [**CD08**, Lemma 2.9]). More precisely, $f, h \in F\langle X \rangle_d$ are cyclically equivalent iff $\operatorname{tr}(f(A)) = \operatorname{tr}(h(A)) = 0$ for all $A \in M_{\lceil \frac{d+1}{2} \rceil}(F)^g$ [**BK09**, Corollary 4.7].

Let us recall an effective version of Hilbert's Nullstellensatz, giving bounds on the polynomials needed in the Bézout identity. We present a variant that combines

[**Kol88**, **Jel05**]. Define

$$N(n, d_1, \dots, d_r) = \begin{cases} d_1 \cdots d_r & \text{if } r \le n, \\ d_1 \cdots d_{n-1} d_r & \text{if } r > n > 1, \\ d_1 + d_r - 1 & \text{if } r > n = 1, \end{cases}$$

$$N'(n, d_1, \dots, d_r) = \begin{cases} N(d_1, \dots, d_r) & \text{if } r \le n, \\ N(d_1, \dots, d_r) & \text{if } r > n \ge 1 \text{ and } d_r > 2, \\ 2N(d_1, \dots, d_r) - 1 & \text{if } r > n > 1 \text{ and } d_r \le 2, \\ 2d_1 - 1 & \text{if } r > n = 1 \text{ and } d_r \le 2, \end{cases}$$

for $d_1 \ge \cdots \ge d_r$.

THEOREM 1.1.1 (Kollár–Jelonek). *Let F be an algebraically closed field and let $f_1, ..., f_r \in F[x_1, \dots, x_n]$ be commutative polynomials without a common zero. Let $d_i = \deg f_i$ and assume $d_1 \ge \cdots \ge d_r$. Then there exist $h_1, \dots, h_r \in F[x_1, ..., x_n]$ satisfying*

$$1 = h_1 f_1 + \cdots + h_r f_r,$$

with $\deg h_i f_i \le N'(n, d_1, \dots, d_r)$ *for* $1 \le i \le r$.

A core feature of this theorem we shall use is that the obtained degree bounds are independent of the number of variables n (for large enough n).

1.2. Spurnullstellensatz. Throughout this section let F be an algebraically closed field of characteristic 0. We remark that a very special case of the following tracial Nullstellensatz was obtained in [**BK11**] by different means.

THEOREM 1.2.1 (Spurnullstellensatz). *Let $f_1, \dots, f_r, f \in F\langle X\rangle$. The implication*

$$\operatorname{tr}(f_1(A)) = \cdots = \operatorname{tr}(f_r(A)) = 0 \quad \Longrightarrow \quad \operatorname{tr}(f(A)) = 0 \tag{1.1}$$

holds for every n and all $A \in M_n(F)^g$ if and only if f is cyclically equivalent to a linear combination of f_i's or a linear combination of f_i's is cyclically equivalent to a nonzero scalar.

REMARK 1.2.2. Note $\sum \lambda_i f_i \overset{\text{cyc}}{\sim} 1$ for some $\lambda_i \in F$ does not necessarily imply that $f \overset{\text{cyc}}{\sim} \sum \mu_i f_i$ for some $\mu_i \in F$. For example, take $f_1 = 1, f = x_1$.

1.2.1. *Proof of Theorem 1.2.1.* As a first step towards the proof of Theorem 1.2.1 we prove its weaker variant, characterizing sets of polynomials f_i whose traces do not have a common vanishing point.

PROPOSITION 1.2.3. *Let $f_1, \dots, f_r \in F\langle X\rangle$. If the equations*

$$\operatorname{tr}(f_i(x_1, \dots, x_g)) = 0,\ 1 \le i \le r, \tag{1.2}$$

do not have a common solution in $\mathcal{M}(F)^g$, then $\sum \lambda_i f_i \overset{\text{cyc}}{\sim} 1$ for some $\lambda_i \in F$.

PROOF. We can assume that $\operatorname{tr}(f_i)$'s are linearly independent as elements in the free algebra with trace and we also assume that $\operatorname{tr}(f_i)$ cannot be written as $\operatorname{tr}(f_i')$ for a polynomial f_i' with $\deg(f_i') < \deg(f_i)$.

For every fixed n we can evaluate a noncommutative polynomial f on the generic $n \times n$ matrices $\xi_k = (x_{ij}^{(k)})$, $1 \le k \le g$, and the trace of $f(\xi_1, \dots, \xi_g)$ is a commutative polynomial in the variables $x_{ij}^{(k)}$, $1 \le i, j \le n$, $1 \le k \le g$. Note that $\deg(\operatorname{tr}(f)) \le \deg(f)$ for every $f \in F\langle X\rangle$, where $\deg(\operatorname{tr}(f))$ denotes the degree of the commutative polynomial $\operatorname{tr}(f(\xi_1, \dots, \xi_g))$. By Theorem 1.1.1, the condition (1.2) implies that there exist $h_1^{(n)}, \dots, h_r^{(n)} \in F[x_{ij}^{(k)} : 1 \le i, j \le n, 1 \le k \le g]$ such that

$$1 = \sum h_i^{(n)} \operatorname{tr}(f_i) \tag{1.3}$$

with

$$\deg(\operatorname{tr}(f_i))\deg(h_i^{(n)}) \leq N'(d_1,\ldots,d_r,n^2g), \tag{1.4}$$

where $d_i = \deg(f_i)$ and $d_1 \geq \cdots \geq d_r$. Applying the Reynolds operator (for the usual action of $\mathrm{GL_n}$ on the polynomial ring $F[x_{ij}^{(k)} : 1 \leq i,j \leq n, 1 \leq k \leq g]$; i.e., $\sigma \in \mathrm{GL}_n$ sends a variable $x_{ij}^{(k)}$ into the (i,j)-entry of the matrix $\sigma^{-1}\ (x_{ij}^{(k)})\ \sigma$) to the equality (1.3) we can assume that the $h_i^{(n)}$ are invariant, and thus induced by pure trace polynomials by the FFT (Theorem 2.1.2.1). The above bound (1.4) is independent of n for $r \leq n^2 g$ (see Theorem 1.1.1). Thus, the degree of $h_i^{(n)}\operatorname{tr}(f_i)$ can be in this case bounded above by $d_1\cdots d_r$.

We thus get for every sufficiently large n; i.e, for $n \geq d_1\cdots d_r$ and $n \geq \sqrt{\frac{r}{g}}$, a trace identity for $M_n(F)$,

$$1 = \sum h_i^{(n)}\operatorname{tr}(f_i), \tag{1.5}$$

where $\deg(h_i^{(n)}\operatorname{tr}(f_i)) \leq n$. Let us fix n with these properties. Since any nontrivial pure trace identity on $n\times n$ matrices has degree at least $n+1$ by [**Pro76**, Theorem 4.5], the above identity (1.5) must be trivial, which means that it holds in the free algebra with trace. A little care is needed at this point. If $f_1,\ldots,f_r$ do not all have zero constant term, then (1.5) is an identity in the free algebra with trace if we replace $\operatorname{tr}(f_i)$ by

$$\tau_i = \operatorname{tr}(\overline{f_i}) + \alpha_i n, \quad 1 \leq i \leq r,$$

where $\overline{f_i}$ is the sum of all nonconstant terms of f_i, and α_i is its constant term, and thus obtain an identity

$$1 = \sum h_i^{(n)}\tau_i. \tag{1.6}$$

Since all polynomials that appear in (1.6) are pure trace polynomials, they belong to the commutative subalgebra $\mathfrak{T}$ of the free algebra with trace. Let us recall that $\mathfrak{T}$ is a polynomial algebra in infinitely many variables.

Let us denote by t_0 the empty word, i.e., the identity of $\mathfrak{T}$, and write

$$\tau_i = \sum_{j=1}^{m}\alpha_{ij}t_j + \alpha_{im+1}t_0.$$

Note that $\alpha_{im+1} = \alpha_i n$. As we assumed that $\operatorname{tr}(f_i)$ are linearly independent, also τ_i are linearly independent. Indeed, assume that $\sum\lambda_i\tau_i = 0$ for some $\lambda_i \in F$. Then $\sum\lambda_i\operatorname{tr}(f_i) = 0$ on $M_n(F)$ for the chosen $n \geq \deg(f_i)$. Thus, $\sum\lambda_i f_i$ has zero trace on $M_n(F)$, so it is cyclically equivalent to a polynomial identity on $M_n(F)$ [**BK09**, Theorem 4.5]. As $n \geq \deg(f_i)$ we can conclude that $\sum\lambda_i\operatorname{tr}(f_i) = 0$ is an identity in the free algebra with trace and due to the independence assumption, $\lambda_i = 0$ for all i.

Note that the independence on τ_i implies that the matrix (α_{ij}) has linearly independent rows and can be brought to the reduced row echelon form. In particular, we can express $t_{i_1},\ldots,t_{i_r}$ as linear combinations of $\tau_1,\ldots,\tau_r$ and $t_{i_{r+1}},\ldots,t_{i_m}$, for some $\{i_0,i_1,\ldots,i_m\} = \{0,1,\ldots,m\}$. If the last row in the reduced echelon form of this matrix has all zeros except for the last entry, then $1 = \sum\lambda_i\tau_i$ as the last column corresponds to t_0. In this case $\operatorname{tr}(\sum\lambda_i f_i - \frac{1}{n}) = 0$ on $M_n(F)$, which implies that $\sum\lambda_i f_i$ is cyclically equivalent to a nonzero scalar. Otherwise we can choose (free) generators of $\mathfrak{T}$ that include $\tau_1,\ldots,\tau_r$ and the above identity (1.6) cannot hold since it does not hold in the polynomial algebra. □

Proof of Theorem 1.2.1. To prove the nontrivial direction, assume that no linear combination of f_i's is cyclically equivalent to a nonzero scalar. The condition (1.1) implies that the equations

$$\operatorname{tr}(f_i(x_1,\dots,x_g))=0\ \ (1\le i\le r),\quad \operatorname{tr}(f(x_1,\dots,x_g)+1)=0$$

do not have a common solution in $\mathcal{M}(F)$. Hence

$$\sum\lambda_i f_i+\lambda(1+f)\overset{\text{cyc}}{\sim}1$$

for some $\lambda_i,\lambda\in F$ by Proposition 1.2.3. By our assumption, $\lambda\neq 0$. Thus,

$$f\overset{\text{cyc}}{\sim}\sum\mu_i f_i+\mu$$

for some $\mu_i,\mu\in F$. If $\mu\neq 0$, then the initial condition (1.1) is violated again by our assumption and Proposition 1.2.3. Hence $\mu=0$ and $f\overset{\text{cyc}}{\sim}\sum\mu_i f_i$. □

1.2.2. *Bounds on the size of matrices in Theorem 1.2.1.* The proof of Proposition 1.2.3 reveals a bound on the size of matrices for which it suffices to test the condition of this lemma and of Theorem 1.2.1 in order to draw the conclusion. If the implication (1.1) holds for all $A\in M_N(F)^g$ for $N=\max\left\{d_1\cdots d_r d,\sqrt{\frac{g}{r}}\right\}$, where $d_i=\deg f_i$, $d=\deg f$, then it holds for all $A\in M_n(F)^g$ for all $n\in\mathbb{N}$. In view of Theorem 1.1.1 we can sometimes sharpen this bound in the case that there exists m satisfying $d_1\cdots d_{m^2g}d_r\le m<\sqrt{\frac{g}{r}}$.

Example 1.2.4. Theorem 1.2.1 fails in the dimension-dependent context: For each $n\in\mathbb{N}$ we give an example of polynomials f_1,f_2 such that $\operatorname{tr}(f_1),\operatorname{tr}(f_2)$ do not have a common zero on $M_n(F)$, but for which there do not exist $\lambda_1,\lambda_2\in F$ and a polynomial identity p of $M_n(F)$ such that

$$\lambda_1 f_1+\lambda_2 f_2\overset{\text{cyc}}{\sim}1+p. \tag{1.7}$$

Let c be a homogeneous central polynomial of $M_n(F)$, which means that $c(A)\in F$ for all $A\in M_n(F)^g$ and c does not vanish identically on $M_n(F)$ (see e.g. [**Row80**]). Take

$$f_1=c,\ f_2=1+c^2.$$

Then $\operatorname{tr}(f_1),\operatorname{tr}(f_2)$ do not have a common zero on $M_n(F)$. If (1.7) holds for some $\lambda_1,\lambda_2\in F$ and a polynomial identity p of $M_n(F)$, then

$$\lambda_1\operatorname{tr}(c)+\lambda_2\Big(\frac{1}{n}\operatorname{tr}(c)^2+n\Big)=n.$$

As c is homogeneous, say of degree $k\ge 1$, then

$$\lambda_1\alpha^k\operatorname{tr}(c)+\lambda_2\Big(\frac{1}{n}\alpha^{2k}\operatorname{tr}(c)^2+n\Big)=n \tag{1.8}$$

for every $\alpha\in F$. But F is infinite, so (1.8) cannot hold for every $\alpha\in F$.

1.2.3. *Passing between a real closed field and its algebraic closure.*

Proposition 1.2.5. *Let R be a real closed field (e.g. $R=\mathbb{R}$) and let C be its algebraic closure. For polynomials $f_1,\dots,f_r,f\in R\langle X\rangle$ the following conditions are equivalent:*

(i) *For every $n\in\mathbb{N}$ and all $A\in M_n(C)^g$ we have* $\operatorname{tr}(f_1(A))=\cdots=\operatorname{tr}(f_r(A))=0$ *implies* $\operatorname{tr}(f(A))=0$*;*
(ii) *There exist $\lambda_i\in C$ such that $\sum\lambda_i f_i\overset{\text{cyc}}{\sim}1$ or $\sum\lambda_i f_i\overset{\text{cyc}}{\sim}f$;*
(iii) *There exist $\lambda_i\in R$ such that $\sum\lambda_i f_i\overset{\text{cyc}}{\sim}1$ or $\sum\lambda_i f_i\overset{\text{cyc}}{\sim}f$;*
(iv) *For every $n\in\mathbb{N}$ and all $A\in M_n(R)^g$ we have* $\operatorname{tr}(f_1(A))=\cdots=\operatorname{tr}(f_r(A))=0$ *implies* $\operatorname{tr}(f(A))=0$.

PROOF. By Theorem 1.2.1, (i) is equivalent to (ii). By taking real parts, it is easy to see that (ii) implies (iii). The implication (iii) to (iv) is trivial. We will prove that (ii) follows by assuming (iv). To obtain a contradiction suppose first that the equations

$$\operatorname{tr}(f_i(x_1,\dots,x_g)) = 0,\ 1 \le i \le r \tag{1.9}$$

do not have a common solution in $\mathcal{M}(R)$ but do have one in $\mathcal{M}(C)$. Let $a_1,\dots,a_g \in M_n(C)$ be such that $\operatorname{tr}(f_i(a_1,\dots,a_g)) = 0$ for $1 \le i \le r$. Write $a_j = b_j + \mathbb{i}\, c_j$, where $b_j, c_j \in M_n(R)$, and define

$$\widetilde{a_j} = \begin{pmatrix} b_j & c_j \\ -c_j & b_j \end{pmatrix}.$$

We have

$$\operatorname{tr}(f_i(\widetilde{a_1},\dots,\widetilde{a_g})) = 2\mathrm{Re}(\operatorname{tr}(f_i(a_1,\dots,a_g))) = 0$$

for every $1 \le i \le r$. Thus, $\widetilde{a_1},\dots,\widetilde{a_g}$ is a common solution of the equations (1.9) in $\mathcal{M}(R)$, a contradiction. By Proposition 1.2.3 we therefore have $\sum \lambda_i f_i \overset{\text{cyc}}{\sim} 1$ for some $\lambda_i \in C$.

If the system (1.9) does have a solution in $\mathcal{M}(R)$, then (iv) implies that the equations

$$\operatorname{tr}(f_i(x_1,\dots,x_g)) = 0,\ \operatorname{tr}(1 + f(x_1,\dots,x_g)) = 0$$

do not have a common solution in $\mathcal{M}(R)$, and by the previous step, applied to polynomials $f_1,\dots,f_r, 1+f$, they also do not have a solution in $\mathcal{M}(C)$. By Proposition 1.2.3, $f \overset{\text{cyc}}{\sim} \sum \lambda_i f_i + \lambda 1$ for some $\lambda_i, \lambda \in C$. Since we are assuming that the equations (1.9) have a common solution in $\mathcal{M}(R)$, (iv) implies $\lambda = 0$. □

1.3. A tracial moment problem. The main result of this section, Corollary 1.3.2, solves a constrained truncated tracial moment problem. For its proof we dualize the statement of Theorem 1.2.1. We refer to [**Bur11**] and the references therein for more details on tracial moment problems.

Let $F \in \{\mathbb{R}, \mathbb{C}\}$. We say that a linear functional $L : F\langle X\rangle_d \to F$ is *tracial* if it vanishes on sums of commutators, or equivalently, if $L(v) = L(w)$ for $v \overset{\text{cyc}}{\sim} w$. The simplest examples of such L are obtained as follows. For $A \in M_n(F)^g$ define

$$\phi_A : F\langle X\rangle_d \to F, \quad \phi_A(p) = \operatorname{tr}(p(A)).$$

Let us denote

$$\mathcal{C} = \operatorname{span}\{\phi_A \mid A \in M_n(F)^g, n \in \mathbb{N}\} \subseteq F\langle X\rangle_d^*.$$

Tracial linear functionals can be described in terms of moment sequences. A sequence $(\alpha_w)_{w\in\langle X\rangle_d}$ in F is a *truncated tracial moment sequence* if $\alpha_w = \alpha_v$ for $v \overset{\text{cyc}}{\sim} w$. Note that any element in $\mathcal{C}$ is tracial and defines a (truncated) tracial moment sequence.

PROPOSITION 1.3.1. *If L is a tracial linear functional on $F\langle X\rangle_d$, then $L \in \mathcal{C}$.*

PROOF. If $L \notin \mathcal{C}$, then there exists a linear functional $p \in (F\langle X\rangle_d^*)^* \cong F\langle X\rangle_d$ such that $p(\phi) = 0$ for every $\phi \in \mathcal{C}$ and $p(L) = 1$. By definition of $\mathcal{C}$ we have $\operatorname{tr}(p(A)) = 0$ for all $A \in M_n(F)^g$, $n \in \mathbb{N}$, which implies $p \overset{\text{cyc}}{\sim} 0$ (see e.g. [**Pro76**, Corollary 4.4]). As L is tracial, $p(L) = L(p) = 0$, a contradiction. □

To give an explicit representation for the tracial linear functional L as a linear combination of the ϕ_A with g-tuples A of $d \times d$ matrices, we present an alternative constructive proof of Proposition 1.3.1.

Alternative proof. Given a tracial moment sequence $(L(w))_{w\in\langle X\rangle_d}$ we show how to find g-tuples A_ℓ such that $L(p)=\sum\lambda_\ell\phi_{A_\ell}$. We proceed inductively. Assume that there exist A_ℓ, $1\le\ell\le\ell_m$, such that $L(p)=\sum\phi_{A_\ell}$ for all $p\in F\langle X\rangle'_m$. We want to find an element of $\mathcal{C}$ which coincides with L on $F\langle X\rangle'_{m+1}$. Define $L_{m+1}=L-\sum_{\ell=1}^{\ell_m}\phi_{A_\ell}$. It is enough to choose matrices $a_{\ell1},\dots,a_{\ell g}\in M_{m+1}(F)$, $\ell_m+1\le\ell\le\ell_{m+1}$, such that $\operatorname{tr}(p(a_{\ell1},\dots,a_{\ell g}))=0$ for $p\in F\langle X\rangle'_m$ and $L_{m+1}(p)=\sum_{\ell=\ell_m+1}^{\ell_{m+1}}\operatorname{tr}(p(a_{\ell1},\dots,a_{\ell g}))$ for p homogeneous of degree $m+1$.

Choose representatives $w_{\ell_m+1},\dots,w_{\ell_{m+1}}$ of cyclic equivalence classes of words in the variables $x_1,\dots,x_g$ of degree $m+1$. Let $w_\ell=x_{i_1}^{j_1}\cdots x_{i_s}^{j_s}$, where $\sum_k j_k=m+1$. We denote $s_k=\sum_{i=1}^k j_i$. Setting $a_{\ell i}=0$ at the beginning, we define matrices $a_{\ell i}\in M_{m+1}(F)$ $(1\le i\le g)$ as follows. We let k vary from 1 to s, and at step k we replace $a_{\ell i_k}$ by

$$\begin{cases} a_{\ell i_k}+\sum_{u=s_{k-1}+1}^{s_k}e_{u,u+1} & \text{if } k<s,\\ a_{\ell i_k}+\sum_{u=s_{k-1}+1}^{s_k-1}e_{u,u+1}+L_{m+1}(w_\ell)e_{m+1,1} & \text{if } k=s.\end{cases}$$

Here e_{ij} are the standard $(m+1)\times(m+1)$ matrix units.

We claim that the only word in $a_{\ell1},\dots,a_{\ell g}$ of degree $\le m{+}1$ with nonzero trace is cyclically equivalent to w_ℓ. A necessary condition for a word w in $a_{\ell1},\dots,a_{\ell g}$, $w=a_{\ell p_1}^{r_1}\cdots a_{\ell p_s}^{r_s}$, of degree m', $1\le m'\le m{+}1$, to have nonzero trace is the existence of a sequence $(\tilde e_j)_{j=1}^{m+1}$ of matrix units from the set $E=\{e_{u,u+1},e_{m+1,1}\mid 1\le u\le m\}$ such that $\operatorname{tr}(\tilde e_1\cdots\tilde e_{m+1})=1$, and if $\sum_{i=1}^{k-1}r_i<j\le\sum_{i=1}^k r_i$ then $\tilde e_j$ appears in $a_{\ell p_k}$. The product of the elements in E has nonzero trace only in a unique order (up to cyclic permutations). Since every element in E appears only in one $a_{\ell i}$, this order determines $p_1,\dots,p_s$. Thus, $\operatorname{tr}(w(a_{\ell1},\dots,a_{\ell g}))=L_{m+1}(w)$ for $w\overset{\text{cyc}}{\sim}w_\ell$ and 0 otherwise. Therefore, $a_{\ell1},\dots,a_{\ell g}$, $\ell_m+1\le\ell\le\ell_{m+1}$, have the desired properties.

We have thus found g-tuples $A_\ell\in M_{n_\ell}(F)^g$, $1\le\ell\le\ell_d$, such that $L(p)=\sum\phi_{A_\ell}(p)$ for every $p\in F\langle X\rangle'_d$. Take $A_0=(0,\dots,0)\in F^g$, and notice that $L(p)=\sum\phi_{A_\ell}(p)+(L(\emptyset)-n)\phi_{A_0}(p)$, where $n=\sum\phi_{A_\ell}(1)=\sum n_\ell$, for every $p\in F\langle X\rangle_d$. □

Let us fix polynomials $f_1,\dots,f_r\in F\langle X\rangle_d$ and write $f_i=\sum\lambda_{ij}w_j$. We say that a sequence $(L(w))_{w\in\langle X\rangle_d}$ is a *constrained truncated tracial moment sequence* if it is a truncated tracial moment sequence and if $L(f_i)=\sum_j\lambda_{ij}L(w_j)=0$ for $1\le i\le r$. We define a constrained analog of $\mathcal{C}$,

$$\mathcal{S}=\operatorname{span}\{\phi_A\in\mathcal{C}\mid\phi_A(f_i)=0,\ 1\le i\le r\}.$$

Note that every element of $\mathcal{S}$ defines a constrained truncated tracial moment sequence.

Corollary 1.3.2 (Constrained truncated tracial moment problem). *If a sequence $(L(w))_{w\in\langle X\rangle_d}$ is a constrained tracial moment sequence with $L(1)=1$, then $L\in\mathcal{S}$.*

Proof. If $L\notin\mathcal{S}$ then there exists an element $p\in(F\langle X\rangle_d^*)^*$ such that $p(\phi)=0$ for all $\phi\in\mathcal{S}$ and $p(L)=1$. We have $\operatorname{tr}(p(A))=0$ for all $A\in M_n(F)^g$ with the property $\operatorname{tr}(f_i(A))=0$ for all $1\le i\le r$. Thus, Theorem 1.2.1 and Proposition 1.2.5 imply that $p\overset{\text{cyc}}{\sim}\sum\lambda_if_i$ or $\sum\lambda_if_i\overset{\text{cyc}}{\sim}1$ for some $\lambda_i\in F$. In the former case we have $L(p)=\sum\lambda_iL(f_i)=0$, which contradicts $L(p)=1$, in the last case $L(1)=\sum\lambda_iL(f_i)=0$, which is contrary to the assumption $L(1)=1$. □

2. The image of a noncommutative polynomial

In this section we study the image of a noncommutative polynomial evaluated on $M_n = M_n(F)$ over an algebraically closed field F.

2.1. Preliminaries. We first introduce some notation and record some well known facts for easier reference. In this section we denote the algebra of generic matrices by GM_n. It is a domain by Amitsur's theorem [**Row80**, Theorem 3.26]. UD_n stands for the generic division ring. The trace of a matrix can be expressed as a quotient of two central polynomials and can be therefore viewed as an element of UD_n (see e.g. [**Row80**, Corollary 1.4.13, Exercise 1.4.9] or Theorem 2.1.7). Since we will need some properties of this expression we repeat here the form needed in the sequel.

PROPOSITION 2.1.1. *There exist central polynomials $c_1, \ldots, c_n$ and a multilinear central polynomial c_0, such that*

$$\mathrm{tr}(a^i)c_0(x_1, \ldots, x_t) = c_i(x_1, \ldots, x_t, a)$$

for every $a, x_1, \ldots, x_t \in M_n$, where $t = 2n^2$.

By replacing a by $f(y_1, \ldots, y_d)$ we can therefore determine the traces of evaluations of f.

Since the coefficients of the characteristic polynomial can be expressed through the traces (see (1.2)), a matrix is nilpotent if and only if the trace of each of its powers is zero.

Note that we have a bijective polynomial map from F^n to F^n, whose inverse is also a polynomial map, which maps coefficients of the characteristic polynomial of any matrix x into its "trace" tuple, $(\mathrm{tr}(x), \ldots, \mathrm{tr}(x^n))$. Let us record an easy lemma for future reference.

LEMMA 2.1.2. *Let p be a symmetric polynomial in n variables. If $f(x) = p(\lambda_1(x), \ldots, \lambda_n(x))$, where $\lambda_1(x), \ldots, \lambda_n(x)$ are the eigenvalues of a matrix $x \in M_n$, then $f(x) = q(\mathrm{tr}(x), \ldots, \mathrm{tr}(x^n))$ for some polynomial q.*

PROOF. Since p is a symmetric polynomial, it can be expressed as a polynomial in the elementary symmetric polynomials $e_1, \ldots, e_n$ by the fundamental theorem of symmetric polynomials. Thus,

$$f(x) = \tilde{p}\Big(e_1\big(\lambda_1(x), \ldots, \lambda_n(x)\big), \ldots, e_n\big(\lambda_1(x), \ldots, \lambda_n(x)\big)\Big).$$

Therefore it suffices to prove that $e_i(\lambda_1(x), \ldots, \lambda_n(x)) = q(\mathrm{tr}(x), \ldots, \mathrm{tr}(x^n))$ for some polynomial q. Since $\lambda_1(x), \ldots, \lambda_n(x)$ are the eigenvalues of a matrix x, they are the zeros of the characteristic polynomial of x, hence by Vieta's formulas $e_i(\lambda_1(x), \ldots, \lambda_n(x))$ equals the coefficient at x^{n-i} in the characteristic polynomial of x. The assertion of the lemma follows for every coefficient can be expressed as a polynomial in the traces of powers of x. □

2.2. Finite polynomials. In this subsection we want to find the "smallest possible" images of polynomials evaluated on M_n. If we want a set $S \subseteq M_n$ to be the image of a polynomial we have to require that it is closed under conjugation by invertible matrices. Hence, one possible criterion for the smallness of the image would be the number of similarity orbits contained in it. Therefore one may be inclined to study polynomials that have just a finite number of similarity orbits in their image. However, the images of many polynomials (for example homogeneous as F is algebraically closed) are closed under scalar multiplication, therefore we take only the orbits modulo scalar multiplication (by nonzero scalars) into account. (In this section the expression "modulo scalar multiplication" will always mean modulo scalar multiplication by nonzero scalars.) In this way we arrive at the following definition, where we denote $a^{\sim} = \{\lambda tat^{-1} \mid t \in \mathrm{GL}_n, \lambda \in F\}$.

DEFINITION 2.2.1. A polynomial f is *finite on* M_n if there exist $a_1, \ldots, a_k \in M_n$ such that $\{0\} \neq \mathrm{im}(f) \subseteq a_1^{\sim} \bigcup \cdots \bigcup a_k^{\sim}$.

Central polynomials of M_n have only one nonzero similarity orbit in their image modulo scalar multiplication and we will see that finite polynomials are in close relation with them.

LEMMA 2.2.2. *If f^j is a central polynomial of M_n for some $j \geq 1$, then f is finite on M_n.*

PROOF. If $b = f(a)$ for some $\underline{a} \in M_n^d$, then the Jordan form of b is either diagonal or nilpotent. If it is diagonal, then it is a scalar multiple of a matrix having j-th roots of unity on the diagonal. There are only finitely many such matrices modulo scalar multiplication. Also the number of similarity orbits of nilpotent matrices modulo scalar multiplication is finite. Thus, f is finite. $\square$

We aim to prove the converse of this simple observation. Let us first introduce a family of matrices that plays an important role in the next theorem. For every j dividing n choose a primitive j-th root of unity μ_j and define the matrix

$$\mathbf{w}_j = \begin{pmatrix} \mathbf{1}_r & & & \\ & \mu_j \mathbf{1}_r & & \\ & & \ddots & \\ & & & \mu_j^{j-1}\mathbf{1}_r \end{pmatrix},$$

where $r = \frac{n}{j}$ and $\mathbf{1}_r$ denotes the $r \times r$ identity matrix.

A polynomial f is said to be *j-central* on M_n if f^j is a central polynomial, while smaller powers of f are not central. We call a polynomial *power-central* if it is j-central for some $j > 1$.

THEOREM 2.2.3. *A polynomial f is finite on M_n if and only if there exists $j \in \mathbb{N}$ such that f is j-central on M_n. Moreover, in this case every nonnilpotent matrix in $\mathrm{im}(f)$ is similar to a scalar multiple of $\mathbf{w}_j$.*

PROOF. Let $a_1, \ldots, a_l$ be the representatives of distinct nonnilpotent similarity orbits of $\mathrm{im}(f)$ on M_n modulo scalar multiplication. For each a_i we set $j_i = \min\{j \mid \mathrm{tr}(a_i^j) \neq 0\}$ (such j exists since a_i is not nilpotent) and let $\alpha_{ik} = \frac{\mathrm{tr}(a_i^k)^{j_i}}{\mathrm{tr}(a_i^{j_i})^k}$ for $1 \leq k \leq n$; these scalars carry the information about the coefficients of the characteristic polynomial of a_i. Note that $\alpha_{ik} = 0$ for $1 \leq k \leq j_i - 1$. The trace polynomial

$$\sum_{k=1}^{n} (\mathrm{tr}(x^k)^{j_i} - \alpha_{ik}\,\mathrm{tr}(x^{j_i})^k) x_k$$

vanishes if we substitute a scalar multiple of a_i for x and arbitrary $b_1, \ldots, b_n \in M_n$ for $x_1, \ldots, x_n$. Since every nonnilpotent matrix in $\mathrm{im}(f)$ is similar to a scalar multiple of a_i for some i and the trace of powers of nilpotent matrices is zero, the following identity holds in UD_n (according to Proposition 2.1.1, all $\mathrm{tr}(f^k)$ lie in UD_n):

$$\prod_{i=1}^{l} \Big(\sum_{k=1}^{n} (\mathrm{tr}(f^k)^{j_i} - \alpha_{ik}\,\mathrm{tr}(f^{j_i})^k) x_k \Big) = 0.$$

Since UD_n is a division ring, one of the factors in the product equals zero in UD_n. Hence there exists i such that $\sum_{k=1}^{n} (\mathrm{tr}(f^k)^{j_i} - \alpha_{ik}\,\mathrm{tr}(f^{j_i})^k) x_k = 0$ and so $\mathrm{tr}(f^k)^{j_i} - \alpha_{ik}\,\mathrm{tr}(f^{j_i})^k = 0$ in UD_n for every $1 \leq k \leq n$. For simplicity of notation we write

j, α_k instead of j_i, α_{ik}, respectively, for $1 \le k \le n$. We will first consider the case when $\alpha_1 \neq 0$, i.e., $j = 1$ and $\mathrm{tr}(f) \neq 0$. Then the characteristic polynomial of f can be expressed as

$$f^n + \sum_{j=1}^{n} \beta_j \,\mathrm{tr}(f)^j f^{n-j} = 0 \tag{2.1}$$

for some $\beta_1, \dots, \beta_n \in F$. Let $\lambda_1, \dots, \lambda_n \in F$ be zeros of the polynomial $x^n + \sum_{j=1}^{n} \beta_j x^{n-j}$. Then we can factorize (2.1) in UD_n as

$$\prod_{k=1}^{n} (f - \lambda_k \,\mathrm{tr}(f)) = 0.$$

This is an identity in UD_n, hence $f - \lambda_k \,\mathrm{tr}(f) = 0$ for some $1 \le k \le n$, implying that f is a central polynomial. Now we consider the general case. We have $\alpha_j \neq 0$ for some $1 \le j \le n$ and $\alpha_k = 0$ for $1 \le k \le j-1$. Then $\mathrm{tr}(f^j) \neq 0$ and f^j is also finite, so we can just repeat the first part of the proof for f^j from which it follows that f^j is a central polynomial. In this case $\mathrm{tr}(f^k) = 0$ for all $1 \le k < j$, therefore f^k is not central.

So far we have proved that f is j-central for some $j \ge 1$ and $\mathrm{tr}(f^k) = 0$ for all $1 \le k < j$. It remains to prove that nonnilpotent matrices in $\mathrm{im}(f)$ are similar to a scalar multiple of $\mathbf{w}_j$. The values of f on M_n can be nilpotent matrices and matrices for which the Jordan form has (modulo scalar multiplication) just powers of the primitive j-th root μ_j of unity on the diagonal. For simplicity of notation we write μ instead of μ_j. Take a nonnilpotent matrix $a \in \mathrm{im}(f)$. We are reduced to proving that the eigenvalues of a are equal to $\lambda, \lambda\mu, \dots, \lambda\mu^{j-1}$ for some $0 \neq \lambda \in F$ (depending on a) and all have the same algebraic multiplicity $\frac{n}{j}$. Recall that $\mathrm{tr}(f^k) = 0$ for $k < j$. Hence, if k_i is the multiplicity of μ^i in the characteristic polynomial of $a \in \mathrm{im}(f)$, then we have

$$\begin{array}{l}
k_0 + \;\; k_1\mu \;\; + \dots + \;\; k_{j-1}\mu^{j-1} \;\; = 0 \\
k_0 + \;\; k_1\mu^2 \;\; + \dots + \; k_{j-1}(\mu^{j-1})^2 \; = 0 \\
\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
k_0 + k_1\mu^{j-1} + \dots + k_{j-1}(\mu^{j-1})^{j-1} = 0.
\end{array}$$

The above equations can be rewritten as $\sum_{i=1}^{j-1} k_i(\mu^t)^i = -k_0$, $1 \le t \le j-1$. Having fixed k_0, the system of equations in variables $k_1, \dots, k_{j-1}$ will have a unique solution if and only if the determinant of $((\mu^t)^i)$, $1 \le t, i \le j-1$, is different from zero. Since μ^t, $1 \le t \le j-1$, are distinct, the Vandermonde argument shows that it is nonzero indeed. Thus, $k_i = k_0$ for every $1 \le i \le j-1$ is the unique solution. Hence, every nonnilpotent matrix in $\mathrm{im}(f)$ is similar to a scalar multiple of the matrix $\mathbf{w}_j$.

The converse follows from Lemma 2.2.2. $\square$

Corollary 2.2.4. *If f is j-central on M_n for some $j \in \mathbb{N}$, then $\mathrm{im}(f^m)$ for $m \ge j$ consists of scalar multiples of exactly one similarity orbit generated by $\mathbf{w}_j^m$.*

Proof. Since every nonnilpotent matrix in $\mathrm{im}(f)$ is similar to a scalar multiple of $\mathbf{w}_j$, its m-th power is similar to a scalar multiple of $\mathbf{w}_j^m$. If $f(\underline{a})^m$, $m \ge j$, is nilpotent, so is $f(\underline{a})^j$. In this case $f(\underline{a})^j = 0$ due to the centrality of f^j. Hence, $\mathrm{im}(f^m)$, $m \ge j$, does not contain nonzero nilpotent matrices. $\square$

Power-central polynomials are important in the structure theory of division algebras. The question whether $M_p(\mathbb{Q})$ has a power-central polynomial for a prime p is equivalent to the long-standing open question whether division algebras of degree p are cyclic (see e.g. [**Row80**, Corollary 3.3.2]. This is known to be true for $p \le 3$. An example of 2-central polynomial on $M_2(K)$ for an arbitrary field K is

$[x, y]$, which is also multilinear. The truth of Lvov's conjecture would imply that there are no multilinear power-central polynomials on $M_n(K)$ for $n \geq 3$. While it is easy to see that multilinear j-central polynomials for $j > 2$ do not exist over $\mathbb{Q}$ (see, e.g., **[Ler75]**), the same question over an algebraically closed field F remains open.

REMARK 2.2.5. If f is j-central, then $\operatorname{tr}(f^2) = 0$ if $j > 2$, and $\operatorname{tr}(f^3) = 0$ if $j = 2$. Thus, if for multilinear polynomials f, g, the identity $\operatorname{tr}(f^2) = 0$ implies $f = 0$ (in UD_n) and the identity $\operatorname{tr}(g^3) = 0$ implies $g = 0$ (in UD_n), then it would follow that there do not exist multilinear noncentral power-central polynomials. (See also Subsection 2.5.)

2.3. Standard open sets as images of polynomials. We will show that if U is a Zariski open subset of F^{n^2}, defined as the nonvanishing set of a polynomial in $F[x_{11}, \ldots, x_{nn}]$ satisfying some natural conditions, then there exists a polynomial f such that $\operatorname{im}(f) = U \cup \{0\}$. We will first prove that this is true for the most prominent example of such a set, GL_n. We follow the standard notation and denote by $V(p)$ the set of zeros of a polynomial p, $V(p) = \{(u_1, \ldots, u_k) \in F^k | \, p(u_1, \ldots, u_k) = 0\}$, and by $D(p) = \{(u_1, \ldots, u_k) \in F^k | \, p(u_1, \ldots, u_k) \neq 0\}$ the complement of $V(p)$. For a subset V of F^k we define $I(V)$ to be the ideal of all polynomials vanishing on V, $I(V) = \{p \in F[z_1, \ldots, z_k] \mid p(u) = 0 \text{ for all } u \in V\}$.

PROPOSITION 2.3.1. *There exists a noncommutative polynomial f such that* $\operatorname{im}(f) = \mathrm{GL}_n \cup \{0\}$ *on M_n.*

PROOF. As $\det(x)$ is a polynomial in the traces of powers of x it can be expressed as the quotient of two central polynomials due to Proposition 2.1.1. We can write $\det(x) = \frac{c(x_1, \ldots, x_t, x)}{c_0(x_1, \ldots, x_t)^n}$, where c, c_0 are central polynomials, c_0 is multilinear and $t = 2n^2$. Note that if we choose $a_1, \ldots, a_t$ such that $c_0(a_1, \ldots, a_t) \neq 0$ then $\det(x) = 0$ if and only if $c(a_1, \ldots, a_t, x) \neq 0$. Define $f = c(x_1, \ldots, x_t, x)x$. As c_0 is multilinear, c is homogeneous in the first variable. Therefore $a \in \operatorname{im}(f)$ forces $Fa \subseteq \operatorname{im}(f)$ because F is algebraically closed. Hence, the image of f consists of all invertible matrices and the zero matrix. □

Two polynomials that differ by a polynomial identity of M_n have the same image when evaluated on M_n, thus we can identify noncommutative polynomials (resp. trace polynomials) with their images in the relatively free algebras GM_n (resp. $\mathcal{T}_n\langle \xi_k \rangle$). With $\xi = (x_{ij}), \xi_k = (x_{ij}^{(k)})$, $\eta_k = (y_{ij}^{(k)})$ we denote the generic matrices in $M_n(\mathcal{C})$ for $\mathcal{C} = F[x_{ij}, x_{ij}^{(k)}, y_{ij}^{(k)} \mid 1 \leq i, j \leq n, 1 \leq k < \infty]$.

LEMMA 2.3.2. *If V is the zero set in F^{n^2} of trace polynomials $p_1, \ldots, p_l \in \mathcal{T}_n$ and if V is closed under scalar multiplication, then there exists a noncommutative homogeneous polynomial f such that* $\operatorname{im}(f) = V^{\mathsf{c}} \cup \{0\}$.

PROOF. Let $p_i = P_i(\xi, \operatorname{tr}(\xi)\mathbf{1}, \ldots, \operatorname{tr}(\xi^n)\mathbf{1})$ for a polynomial $P_i(z_0, z_1, \ldots, z_n)$, $1 \leq i \leq l$. Let us write $\operatorname{tr}(\xi^i) = \frac{c_i(\xi_1, \ldots, \xi_t, \xi)}{c_0(\xi_1, \ldots, \xi_t)}$ where c_0, c_i, $1 \leq i \leq n$, are polynomials from Proposition 2.1.1. We replace $P_i(\xi, \operatorname{tr}(\xi)\mathbf{1}, \ldots, \operatorname{tr}(\xi^n)\mathbf{1})$, $1 \leq i \leq l$, with $Q_i(\xi, \eta_i) = \operatorname{tr}(P_i(\xi, \operatorname{tr}(\xi)\mathbf{1}, \ldots, \operatorname{tr}(\xi^n)\mathbf{1})\eta_i)$, $1 \leq i \leq l$, which map to F. Let $r_i - 1$ be the degree of the polynomial $P_i(z_0, z_1, \ldots, z_n)$ treated as a polynomial in the last n variables, $z_1, \ldots, z_n$. Then $c_0(\xi_1, \ldots, \xi_t)^{r_i} Q_i(\xi, \eta_i)$ is a central polynomial.

We denote $\underline{\xi}_i = (\xi_{i1}, \ldots, \xi_{it})$, $1 \leq i \leq l$, $\underline{\eta} = (\eta_1, \ldots, \eta_l)$. Then

$$c(\underline{\xi}_1, \ldots, \underline{\xi}_l, \xi, \underline{\eta}) = \sum_{i=1}^{l} c_0(\underline{\xi}_i)^{r_i} Q_i(\xi, \eta_i)$$

is a sum of central polynomials and therefore a central polynomial. If $a \in V$ we have $c(\underline{a}_1, \ldots, \underline{a}_l, a, \underline{b}) = 0$ for any choice of matrices a_{ij}, b_i. On the other hand, suppose that $b \notin V$, hence $p_i(a_{11}, \ldots, a_{nn}) \neq 0$ for some i. Consequently, there exists $b \in M_n$ such that $Q_i(a, b) \neq 0$. If we choose $a_{i1}, \ldots, a_{it}$ such that $c_0(\underline{a}_i) \neq 0$ and write $\underline{b}$ for the l-tuple that has b on the i-th place and zero elsewhere, then $c(0, \ldots, 0, \underline{a}_i, 0, \ldots, 0, a, \underline{b}) = \mu \mathbf{1}$ for some $0 \neq \mu \in F$. Let

$$f(\underline{\xi}_1, \ldots, \underline{\xi}_l, \xi, \underline{\eta}) = c(\underline{\xi}_1, \ldots, \underline{\xi}_l, \xi, \underline{\eta})\xi.$$

By construction, $\mu a \in \operatorname{im}(f)$ and since $Q_i(\xi, \eta_i)$ is linear in η_i, all scalar multiples of a belong to the image of f (indeed, $\lambda \mu a = c_0(a_1, \ldots, a_t)^{r_i} Q_i(a, \lambda b) a = f(0, \ldots, 0, \underline{a}_i, 0, \ldots, 0, a, \lambda \underline{b})$ for every $\lambda \in F$). Hence the image of f equals $V^c \cup \{0\}$.

Since V is closed under scalar multiplication we can assume that p_i, $1 \leq i \leq l$, are homogeneous, since otherwise we can replace them by their homogeneous components. These also belong to $I(V)$, which can be easily seen by the Vandermonde argument. The homogeneous components of p_i are also trace polynomials, which follows by comparing both sides of the equality $p_i(\lambda x_{11}, \ldots, \lambda x_{nn}) = P_i(\lambda\xi, \operatorname{tr}(\lambda\xi), \ldots, \operatorname{tr}((\lambda\xi)^n))$. Hence, we can assume that $P_i(\xi, \operatorname{tr}(\xi), \ldots, \operatorname{tr}(\xi^n))$ are homogeneous polynomials of degree d_i. We denote $d = \max\{d_i + r_i t,\ 1 \leq i \leq l\}$. If we replace $Q_i(\xi, \eta_i)$ in the above construction by $Q_i(\xi, \eta_i^{d-d_i-r_i t+1})$ then f becomes a homogeneous polynomial of degree $d+1$. Noting that F being algebraically closed guarantees that $\operatorname{im}(f)$ is closed under scalar multiplication it is easy to verify that the above proof remains valid with polynomials $Q_i(\xi, \eta_i^{d-d_i-r_i t+1})$ replacing polynomials $Q_i(\xi, \eta_i)$. $\square$

We illustrate this result with some examples of sets that can be realized as images of noncommutative polynomials.

Example 2.3.3. (a) The union of matrices that are not nilpotent of the nilindex less or equal to k and the zero matrix is the image of a noncommutative polynomial. The matrices whose k-th power equals zero are closed under conjugation by GL_n and under scalar multiplication, and they are the zero set of the (trace) polynomial X^k. Hence, we can apply Lemma 2.3.2.

(b) Matrices with at most k distinct eigenvalues, $0 \leq k \leq n-1$, are also the zero set of trace polynomials. Define polynomials $p_0 = \xi$, $q_l(z_1, \ldots, z_{l+1}) = \prod_{1 \leq i < j \leq l+1} (z_i - z_j)^2$ and

$$p_l = \sum_{1 \leq i_1 < \cdots < i_{l+1} \leq n} q_l(\lambda_{i_1}(\xi), \ldots, \lambda_{i_{l+1}}(\xi)), \qquad 1 \leq l \leq n-1,$$

where $\lambda_i(\xi)$, $1 \leq i \leq n$, are the eigenvalues of a matrix ξ. Note that the polynomials on the right-hand side of the above definition of p_l, $1 \leq l \leq n-1$, are symmetric polynomials in the eigenvalues of the matrix ξ, and thus pure trace polynomials by Lemma 2.1.2. The polynomials p_l, $k \leq l \leq n-1$, define the desired variety. Indeed, $p_{n-1}(\underline{x})$ is the discriminant of ξ and a matrix a is a zero of p_{n-1} if and only if a has at most $n-1$ distinct eigenvalues. Then we can proceed by reverse induction to show that the common zeros of $p_{n-1}, \ldots, p_k$ are the matrices that have at most k distinct eigenvalues supposing that the common zeros of $p_{n-1}, \ldots, p_{k+1}$ are the matrices that have at most $k+1$ distinct eigenvalues. If a is a zero of $p_{n-1}, \ldots, p_{k+1}$, i.e. a has at most $k+1$ distinct eigenvalues by the induction hypothesis, then $p_k(a)$ is a scalar multiple of $q_k(\lambda_1, \ldots, \lambda_{k+1}) \neq 0$ where $\lambda_1, \ldots, \lambda_{k+1}$ are possible distinct eigenvalues of a. Therefore, a is a zero of p_k if the evaluation of q_k in this $k+1$-tuple is equal to zero, i.e. if a has at most k distinct eigenvalues. By Lemma 2.3.2, the matrices with at least k distinct eigenvalues together with the zero matrix form the image of a noncommutative polynomial for every $1 \leq k \leq n$.

(c) Define trace polynomials $t_i = \mathrm{tr}(\xi^i)\xi - \mathrm{tr}(\xi)\xi^i \in \mathcal{T}_n\langle\xi_k\rangle$ for $2 \le i \le n$. Let a matrix a be a zero of $t_2, \dots, t_n$. Since a is a zero of t_2, a is a scalar multiple of an idempotent or $\mathrm{tr}(a) = 0$. In the second case, $\mathrm{tr}(a^i) = 0$, $1 \le i \le n$, since a is a zero of t_i, $2 \le i \le n$. Thus, the variety defined by t_i, $2 \le i \le n$, contains precisely the scalar multiples of idempotents and nilpotent matrices (only these have the trace of all powers equal to zero). Consequently, the complement of this variety, matrices that are not scalar multiples of an idempotent and not nilpotent, together with the zero matrix equals the image of a noncommutative polynomial.

The proof of the following theorem is based on an idea of Klemen Šivic. For another proof see the first proof of [**Špe13**, Theorem 4.5]. We first introduce some notation and prove a lemma that will play a role also in the subsequent subsection.

Let $\phi : M_n \to F^n$ be the map that assigns to every matrix the coefficients of its characteristic polynomial. More precisely, if $x^n + \alpha_1 x^{n-1} + \cdots + \alpha_n$ is the characteristic polynomial of a matrix a, then $\phi(a) = (\alpha_1, \dots, \alpha_n)$. Note that ϕ is a surjective polynomial map.

Lemma 2.3.4. *If Z is a proper closed subset of M_n that is closed under conjugation by GL_n, then $\phi(Z)$ is contained in a proper closed subset of F^n.*

Proof. Since the closure of similarity orbits of the set D of all diagonal matrices equals M_n, $Z \cap D$ is also a proper closed subset of $D \cong F^n$. Hence $\dim(Z \cap D) < n$. Therefore $\dim(\overline{\phi(Z \cap D)}) < n$, which implies that $\overline{\phi(Z \cap D)}$ is a proper closed set of F^n. Denote by $\tilde{D}$ the set of all diagonalizable matrices. As Z is closed under conjugation by GL_n, $\phi(Z \cap \tilde{D}) = \phi(Z \cap D)$. Decompose $\phi(Z) = \phi(Z \cap \tilde{D}) \cup \phi(Z \cap \tilde{D}^{\mathrm{c}})$ and notice that $\phi(Z \cap \tilde{D}^c)$ is a subset of the proper closed subset of the variety defined by the discriminant, $V(\mathrm{disc})$. Hence the closure of $\phi(Z)$ is a proper closed subset of F^n. □

Theorem 2.3.5. *Let p be a commutative polynomial in n^2 variables. If $V(p) \subset F^{n^2}$ is closed under conjugation by invertible matrices then p is a pure trace polynomial.*

Proof. Firstly, we can assume that p is irreducible. To see this we only need to observe that all irreducible components V_i of $V(p) = \bigcup V_i$ are closed under conjugation by invertible matrices. Take $a \in V_i$, then the variety $V_a = \overline{\{\sigma a \sigma^{-1},\ \sigma \in \mathrm{GL}_n\}}$ is rationally parametrized, and therefore irreducible (see, e.g., [**CLO07**, Proposition 4.5.6]). Hence, we have $V_a \subseteq V_i$ for every $a \in V_i$, so V_i is closed under conjugation by invertible matrices. In the rest of the proof we therefore assume p to be irreducible.

We fix an invertible matrix σ and define a polynomial $p_\sigma(x_{11}, \dots, x_{nn}) = p(\sigma\xi\sigma^{-1})$, which means that we first conjugate the matrix ξ with σ and then apply p on the n^2-tuple corresponding to the matrix $\sigma\xi\sigma^{-1}$. According to the assumption of the theorem, p and p_σ have equal zeros. Hence, $V(p) = V(p_\sigma)$. As p and hence also p_σ are irreducible, we have $p_\sigma = \alpha_\sigma p$ for some scalar $\alpha_\sigma \in F$ by Hilbert's Nullstellensatz. We shall have established the lemma if we prove that $\alpha_\sigma = 1$ for every $\sigma \in\ \mathrm{GL}_n$. Indeed, then we can use the FFT. We have $p(\sigma a \sigma{-}1) = \alpha_\sigma p(a)$ for every $\sigma \in \mathrm{GL}_n$, $a \in M_n$. In particular, $p(\sigma) = \alpha_\sigma p(\sigma)$, which implies $\alpha_\sigma = 1$ for every $\sigma \in U = \mathrm{GL}_n \cap D(p)$. Then for every $a \in M_n$ the polynomials $p(sa)$ and $p(as)$ in n^2 variables $s_{11}, \dots, s_{nn}$ equal on U. Since U is a dense subset of F^{n^2}, they are equal. Thus $\alpha_\sigma = 1$ for every $\sigma \in \mathrm{GL}_n$. □

The next corollary rephrases the last statement in the language of invariant theory.

COROLLARY 2.3.6. *If for a polynomial* $p : F^{n^2} \to F$ *and for every* $a \in M_n, \sigma \in \mathrm{GL}_n$ *we have* $p(\sigma a \sigma^{-1}) = 0$ *if and only if* $p(a) = 0$*, then* p *is a matrix invariant.*

Having established Lemma 2.3.2 and Theorem 2.3.5, we can now state the main result of this subsection.

THEOREM 2.3.7. *Let* $U = D(p)$ *be a standard open set in* F^{n^2} *closed under conjugation by* GL_n *and nonzero scalar multiplication. There exists a noncommutative homogeneous polynomial* f *such that* $\mathrm{im}(f) = U \cup \{0\}$.

To generalize this theorem to arbitrary open subsets of F^{n^2} that are closed under conjugation by GL_n and scalar multiplication with the similar approach (employing Lemma 2.3.2), one would need to prove that every variety that is closed under conjugation by GL_n and under scalar multiplication can be determined by trace polynomials. Those trace polynomials may include some extra variables. It is easy to adjust the proof of Lemma 2.3.2 to that slightly more general context. See Example 2.3.8 below.

EXAMPLE 2.3.8. Let V be the set of all matrices having minimal polynomial of degree at most 2. This is a closed set since each of its element is a zero of the Capelli polynomial $C_5(1, x, x^2, a_1, a_2)$ for arbitrary $a_1, a_2 \in M_n$, and due to [**Row80**, Theorem 1.4.34] for $a_0 \notin V$ there exist $a_1, a_2 \in M_n$ such that $C_5(1, a_0, a_0^2, a_1, a_2) \neq 0$. Hence $c_0(x_1, \dots, x_t)\,\mathrm{tr}(C_5(1, x, x^2, y, y)y_1)x$ has in its image exactly the zero matrix and matrices whose minimal polynomial has degree at least 3.

2.4. Density. Each noncommutative polynomial f in d variables gives rise to a function $f : M_n^d \to M_n$. In this subsection we will consider this function as a polynomial map in $n^2 d$ variables. We will be concerned with some topological aspects of its image on M_n. We discuss the sufficient conditions for establishing the "dense counterpart" of Lvov's conjecture. By this we mean the question whether the image of a multilinear polynomial f on M_n is dense in M_n or in M_n^0, assuming that f is neither a polynomial identity nor a central polynomial of M_n.

Recall that the map $\phi : M_n \to F^n$, introduced in the previous subsection, assigns to every matrix the coefficients of its characteristic polynomial. The restriction of ϕ to M_n^0 will be denoted by ϕ_0. Identifying $\{0\} \times F^{n-1}$ with F^{n-1}, we may and we will consider ϕ_0 as a map into F^{n-1}.

By saying that $\mathrm{im}(f)$ is dense in F^{n^2-1} we mean that the image of f is dense in M_n^0, an $(n^2 - 1)$-dimensional space over F, with the inherited topology from F^{n^2}. Let us give a proof of the following two rather obvious facts.

LEMMA 2.4.1. *Let* f *be a noncommutative polynomial. Then* $\mathrm{im}(f)$ *is dense in* F^{n^2} *(resp.* F^{n^2-1} *if* $\mathrm{tr}(f) = 0$*) if and only if* $\mathrm{im}(\phi(f))$ *(resp.* $\mathrm{im}(\phi_0(f))$*) is dense in* F^n *(resp.* F^{n-1}*).*

PROOF. Assume that $\mathrm{im}(\phi(f))$ is dense in F^n. Denote by Z the Zariski closure of $\mathrm{im}(f)$. As $\mathrm{im}(f)$ is closed under conjugation by GL_n so is Z, thus we can apply Lemma 2.3.4 to derive that $Z = F^{n^2}$. Conversely, if $\mathrm{im}(f)$ is dense in F^{n^2} then $\mathrm{im}(\phi(f))$ is dense in F^n since ϕ is a surjective continuous map. The respective part can be handled in much the same way, the only difference being the analysis of respective maps within the framework of M_n^0. $\square$

Let f be a noncommutative polynomial depending on d variables. In the following corollary we regard $\mathrm{tr}(f)$ as a commutative polynomial in $n^2 d$ commutative variables.

PROPOSITION 2.4.2. *The image of a polynomial* f *is dense in* F^{n^2} *(resp.* F^{n^2-1} *if* $\mathrm{tr}(f) = 0$*) if and only if* $\mathrm{tr}(f), \dots, \mathrm{tr}(f^n)$ *(resp.* $\mathrm{tr}(f^2), \dots, \mathrm{tr}(f^n)$*) are algebraically independent.*

Proof. We have a bijective polynomial map from F^n to F^n (whose inverse is also a polynomial map), which maps the coefficients of the characteristic polynomial of an arbitrary matrix $a \in M_n$ to its "trace" tuple, $(\operatorname{tr}(a), \ldots, \operatorname{tr}(a^n))$ (see Section 2.1). Hence $\operatorname{tr}(f), \ldots, \operatorname{tr}(f^n)$ are algebraically independent if and only if the coefficients of the characteristic polynomial of f are algebraically independent.

Assume that the coefficients of the characteristic polynomial of f are algebraically dependent. Then the image of $\phi(f)$ is contained in a proper algebraic subvariety in F^n, which is in particular not dense in F^n, therefore $\operatorname{im}(f)$ cannot be dense in F^{n^2}. To prove the converse assume that the coefficients of the characteristic polynomial of f are algebraically independent. Then the closure of $\operatorname{im}(\phi(f))$ cannot be a proper subvariety and is thus dense in F^n. We can now apply Lemma 2.4.1 to conclude the proof of the first part.

The respective part of Lemma 2.4.1 yields in the same manner as above the respective part of this corollary. $\square$

Let V be an irreducible algebraic variety. Recall that the closure of the image of a polynomial map $p : V \to F^k$ is an irreducible algebraic variety. Thus, if $p(V) \cap (F^{k-1} \times \{0\})$ is dense in $F^{k-1} \times \{0\}$ and $p(V) \not\subseteq F^{k-1} \times \{0\}$ then the (Zariski) closure Z of $p(V)$ equals F^k. (Suppose on contrary that $Z = V(p_1, \ldots, p_l) \neq F^k$. We can assume that $p_1(z_1, \ldots, z_k) \neq \alpha z_k^r$ for $r \in \mathbb{N}$, $\alpha \in F$. Write $p_1(z_1, \ldots, z_k) = \sum_{i=0}^m q_i(z_1, \ldots, z_{k-1}) z_k^i$ and note that q_0 equals zero since $Z \cap (F^{k-1} \times \{0\}) = F^{k-1} \times \{0\}$. Thus, there exists the maximal $r \geq 1$ such that we can write $p_1(z_1, \ldots, z_k) = q(z_1, \ldots, z_k) z_k^r$ for some nonconstant polynomial q. Hence, $V(p_1) = V(q) \cup V(z_k)$ and, by assumptions and choice of r, $Z \neq V(q) \cap Z \neq \emptyset$, $Z \neq V(z_k) \cap Z \neq \emptyset$. We derived a contradiction, $Z = Z \cap V(p_1) = (Z \cap V(q)) \cup (Z \cap V(z_k))$.)

In the next lemma we will see how the image of a polynomial f evaluated on M_{n-1} impacts $\operatorname{im}(f)$ on M_n. In order to distinguish between these images we write $\operatorname{im}_k(f)$ for $\operatorname{im}(f)$ evaluated on M_k. We identify M_{n-1} with $\begin{pmatrix} M_{n-1} & 0 \\ 0 & 0 \end{pmatrix}$ inside M_n.

Lemma 2.4.3. *If* $\operatorname{im}_{n-1}(f) \cap M_{n-1}^0$ *is dense in* M_{n-1}^0 *then* $\operatorname{im}_n(f)$ *is dense is* M_n^0. *If, additionally,* $\operatorname{im}_n(f) \not\subseteq M_n^0$ *then* $\operatorname{im}_n(f)$ *is dense in* M_n.

Proof. Assume that $\operatorname{im}_{n-1}(f) \cap M_{n-1}^0$ is dense in M_{n-1}^0. Therefore $\operatorname{im}_n(\phi(f)) \cap (\{0\} \times F^{n-2} \times \{0\})$ is dense in $\{0\} \times F^{n-2} \times \{0\}$. (The last component of the polynomial map $\phi(f)$ is $\det(f)$.) According to the discussion preceding the lemma, $\operatorname{im}(\phi(f)) \cap (\{0\} \times F^{n-1})$ is dense in $\{0\} \times F^{n-1}$ if it contains an invertible matrix. The later was observed in [**LZ09**, Theorem 2.4]. Thus, $\operatorname{im}_n(f) \cap M_n^0$ is dense in $M_n^0 \cong F^{n^2-1}$ by Lemma 2.4.1. If, additionally, there exists a matrix in the image of f with nonzero trace, $\operatorname{im}_n(f)$ is dense in M_n by the above discussion identifying M_n^0 with F^{n^2-1}. $\square$

Corollary 2.4.4. *If a multilinear polynomial f is neither a polynomial identity nor a central polynomial of M_2, then* $\operatorname{im}(f)$ *is dense in M_n for every $n \geq 2$.*

Proof. Apply [**KBMR12**, Theorem 2] and Lemma 2.4.3. $\square$

In view of Lemma 2.4.3 it would suffice to verify the density version of Lvov's conjecture for a polynomial f evaluated on M_n for such n that f is a polynomial identity or a central polynomial of M_{n-1} and is not a polynomial identity or a central polynomial of M_n. The first step in this direction may be to establish the density of the image of the standard polynomials.

The following questions arise when trying to establish a connection between Lvov's conjecture and its dense counterpart. Does the density of $\operatorname{im}(f)$ in M_n or in M_n^0 for a multilinear polynomial f imply that $\mathrm{im}(f) = M_n$ or M_n^0, respectively? Is the image of a multilinear polynomial closed in F^{n^2}?

REMARK 2.4.5. The image of a homogeneous polynomial is not necessarily closed in F^{n^2}. Lemma 2.3.2 provides examples of such homogeneous polynomials.

REMARK 2.4.6. We were dealing with the Zariski topology, however, if the underlying field F equals $\mathbb{C}$, the field of complex numbers, all statements remain valid when we replace the Zariski topology with the (more familiar) Euclidean topology. This rests on the result from algebraic geometry (see, e.g., [**Har77**, Exercise II.3.19]) asserting that the image of a polynomial map g contains a Zariski open set of its closure. Thus, if the image of $g : \mathbb{C}^m \to \mathbb{C}^k$ is dense in the Zariski topology, then it contains a dense open subset, which is clearly open and also dense in $\mathbb{C}^k$ in the Euclidean topology. (Indeed, its complement, which is a set of zeros of some polynomials, cannot contain an open set.) Consequently, the image of a polynomial map $g : \mathbb{C}^m \to \mathbb{C}^k$ is dense in the Zariski topology in $\mathbb{C}^k$ if and only if it is dense in the Euclidean topology. However, the question whether the image of a multilinear polynomial $f : M_n(\mathbb{C})^d \to M_n(\mathbb{C})$ is closed in the Euclidean topology in $M_n(\mathbb{C})$ might be approachable with tools of complex analysis.

2.5. Zero trace squares of polynomials. Let f be a polynomial that is not an identity of M_n. The simplest situation where the conditions of Corollary 3.2.2 are not fulfilled is when $\mathrm{tr}(f^2) = 0$. Let us first show that this can actually occur. The proof of the next proposition is due to Igor Klep.

PROPOSITION 2.5.1. *Let $n = 2^m\ell$, where $\ell > 1$ is odd. Then there exists a multihomogeneous polynomial f which is not a polynomial identity of M_n with $\mathrm{tr}(f^2) = 0$ on M_n.*

PROOF. Consider the universal division algebra $\mathcal{D} = \mathrm{UD}_n$. $\mathcal{D}$ comes equipped with the reduced trace $\mathrm{tr} : \mathcal{D} \to \mathcal{Z} = Z(\mathcal{D})$.

We claim that the quadratic trace form $q : x \mapsto \mathrm{tr}(x^2)$ on $\mathcal{D}$ is isotropic. Since $\ell > 1$, there is an odd degree extension K of $\mathcal{Z}$ such that $\mathcal{D} \otimes_{\mathcal{Z}} K = M_l(K \otimes_{\mathcal{Z}} \mathcal{D}')$ where $\mathcal{D}'$ is a division ring (see e.g. [**Row80**, Theorem 3.1.40]). The natural extension q_K of q to $M_l(K \otimes_{\mathcal{Z}} \mathcal{D}')$ is obviously isotropic, i.e., there is $A \in M_l(K \otimes_{\mathcal{Z}} 1)$ with $q_K(A) = \mathrm{tr}(A^2) = 0$. Hence by Springer's theorem [**EKM08**, Corollary 18.5], q is isotropic as well. There exists $0 \neq y \in \mathcal{D}$ with $\mathrm{tr}(y^2) = 0$. We have $y = fc^{-1}$ for some $f \in \mathrm{GM}_n$ and $c \in Z(\mathrm{GM}_n)$. Replacing y by cy, we may assume without loss of generality that $y \in \mathrm{GM}_n$. There is $f \in F\langle X\rangle$ whose image in GM_n coincides with y.

By the universal property of the reduced trace on $\mathcal{D}$, $\mathrm{tr}(y^2) = 0$ translates into $\mathrm{tr}(f(\underline{a})^2) = 0$ for all tuples $\underline{a}$ of $n \times n$ matrices over F. By (multi)homogenizing we can even achieve that $0 \neq f$ is multihomogeneous. $\square$

As explained in the introduction, one would expect that multilinear polynomials that are not identities cannot satisfy $\mathrm{tr}(f^k) = 0$ for $k \geq 2$. Unfortunately, we are able to prove this only in the dimension-free setting. That is, we consider the situation where f satisfies $\mathrm{tr}(f^k) = 0$ on M_n for every $n \geq 1$. This is equivalent to the condition that f^k is a sum of commutators [**BK09**, Corollary 4.8].

PROPOSITION 2.5.2. *If $f \in F\langle X\rangle$ is a nonzero multilinear polynomial, then f^k, $k \geq 2$, is not a sum of commutators.*

PROOF. To avoid notational difficulties, we will consider only the case where $k = 2$. The modifications needed to cover the general case are rather obvious.

If f^2 is a sum of commutators then $\mathrm{tr}(f^2) = 0$ in $M_n(F)$ for all $n \in \mathbb{N}$. Let us write $f = f(x_1, \ldots, x_d) = \sum f_i x_1 g_i$. Since $\mathrm{tr}(f(x + y, x_2, \ldots, x_d)^2 -$

$f(x, x_2, \dots, x_d)^2 - f(y, x_2, \dots, x_d)^2) = 0$, we have $\mathrm{tr}(x(\sum_{i,j} g_i f_j y g_j f_i)) = 0$. This implies that $\sum_{i,j} g_i f_j x_1 g_j f_i$ is a polynomial identity for every matrix algebra, so it has to be trivial. Denote by f^* the Razmyslov transform of f according to x_1, $f^* = \sum g_i x_1 f_i$. We have $f^*(f(x_1, x_2, \dots, x_d), x_2, \dots, x_d) = 0$ in the free algebra $F\langle \underline{X} \rangle$, which further yields $f^* = 0$. Indeed, suppose $f^* \neq 0$ and choose monomials m_1, m_2 with nonzero coefficients in f and f^*, respectively, which are minimal due to the first appearance of x_1. Then the coefficient of the monomial $m_2(m_1, x_2, \dots, x_d)$ in the polynomial $f^*(f(x_1, x_2, \dots, x_d), x_2, \dots, x_d)$ is nonzero, a contradiction. Hence, f^* has to be zero, which leads to the contradiction $f = 0$ ($f^* = 0$ if and only if $f = 0$, see, e.g., [**For91**, Proposition 12]). $\square$

Remark 2.5.3. From the proof we deduce that only the linearity in one variable is needed. However, for general polynomials we were not able to find out whether f^k can be a sum of commutators. In any case, this problem can be just a test for a more general next question.

Let M_∞ denote the algebra of all infinite matrices with finitely many nonzero entries. We write M_∞^0 for the set of elements in M_∞ with zero trace, where the trace is defined as the sum of diagonal entries.

Question. Is the image of an arbitrary noncommutative polynomial with zero constant term f on M_∞ a dense subset (in the Zariski topology) of M_∞ or of M_∞^0? Is $\mathrm{im}(f) = M_\infty$ or $\mathrm{im}(f) = M_\infty^0$?

If f^k is a sum of commutators for some polynomial f and some $k > 1$, then $\mathrm{im}(f)$ on M_∞ is neither dense in M_∞ nor in M_∞^0, hence such a polynomial f would provide a counterexample to the above question.

The question about the density in the sense of the Jacobson density theorem was settled in [**CL10**].

2.6. Lie polynomials of degree $2, 3, 4$. We prove that Lvov's conjecture holds for multilinear Lie polynomials of degree less or equal to 4. We use the right-normed notation, $[x_n, \dots, x_1]$ denotes $[x_n, [x_{n-1}, [\dots [x_2, x_1]] \dots]$.

Lemma 2.6.1. *If f is of the form $f(x_1, \dots, x_d) = [x_{i_1}, x_{i_2}, \dots, x_{i_{k-1}}, x_1]$, where $2 \leq i_j \leq d$, then $\mathrm{im}(f) = M_n^0$.*

Proof. Choose a diagonal matrix s with distinct diagonal entries λ_i, $1 \leq i \leq n$. Then $f(x, s, \dots, s)_{ij} = \pm(\lambda_i - \lambda_j)^{k-1} x_{ij}$, where $x = (x_{ij})$. Thus, $\mathrm{im}(f)$ contains all matrices with zero diagonal entries. Since $\mathrm{im}(f)$ is closed under conjugation and every matrix with zero trace is similar to a matrix with zero diagonal (see, e.g., [**Sho37**]), we have $\mathrm{im}(f) = M_n^0$. $\square$

If f is a Lie polynomial of degree 2, $f = \alpha[x_1, x_2]$, $\alpha \neq 0$, it has been known for a long time [**Sho37, AM57**] that $\mathrm{im}(f) = M_n^0$. We list this as a lemma for the sake of reference.

Lemma 2.6.2. *If f is a Lie polynomial of degree 2, then $\mathrm{im}(f) = M_n^0$.*

Lemma 2.6.3. *If f is a multilinear Lie polynomial of degree 3, then $\mathrm{im}(f) = M_n^0$.*

Proof. We can assume that $f(x, y, z) = [z, y, x] + \alpha[y, z, x]$. If we take $x = z$, we have $f(x, y, x) = [x, y, x] = -[x, x, y]$. We apply Lemma 2.6.1 to conclude $M_n^0 = \mathrm{im}(f(x, y, x)) \subseteq \mathrm{im}(f) \subseteq M_n^0$. $\square$

Lemma 2.6.4. *If f is a multilinear Lie polynomial of degree 4, then $\mathrm{im}(f) = M_n^0$.*

Proof. It is easy to see that the monomials $[x_i, x_j, x_k, x_1]$, $\{i, j, k\} = \{2, 3, 4\}$, form a basis of multilinear Lie polynomials of degree 4. We can assume that

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= [x_4, x_3, x_2, x_1] + \alpha_1[x_3, x_4, x_2, x_1] + \alpha_2[x_4, x_2, x_3, x_1] \\ &\quad +\alpha_3[x_2, x_4, x_3, x_1] + \alpha_4[x_3, x_2, x_4, x_1] + \alpha_5[x_2, x_3, x_4, x_1]. \end{aligned}$$

Consider $f(x, x, x, y) = (\alpha_4 + \alpha_5)[x, x, y, x] = -(\alpha_4 + \alpha_5)[x, x, x, y]$. Due to Lemma 2.6.1 we can assume $\alpha_4 + \alpha_5 = 0$. Similarly, setting $x_1 = x_2 = x_4$ and $x_1 = x_3 = x_4$ yields $\alpha_2 + \alpha_3 = 0$ and $1 + \alpha_1 = 0$, respectively. Hence we can write

$$\begin{aligned} f(x_1, x_2, x_3, x_4) &= [x_4, x_3, x_2, x_1] - [x_3, x_4, x_2, x_1] + \alpha_2([x_4, x_2, x_3, x_1] \\ &\quad -[x_2, x_4, x_3, x_1]) + \alpha_4([x_3, x_2, x_4, x_1] - [x_2, x_3, x_4, x_1]) \\ &= [[x_4, x_3], x_2, x_1] + \alpha_2[[x_4, x_2], x_3, x_1] + \alpha_4[[x_3, x_2], x_4, x_1]. \end{aligned}$$

Then $f(x, y, y, z) = (1 + \alpha_2)[[z, y], y, x] = (1 + \alpha_2)[[z, y], [y, x]]$. It follows from the proof of **[Smi76**, Theorem 1] that any matrix with zero trace except for the rank-one matrix is similar to a commutator of two matrices with zero diagonal. Choose a diagonal matrix s with distinct diagonal entries. Take $a \in M_n^0$ with rank at least 2. Since $\mathrm{im}(f)$ is closed under conjugation by GL_n, we may assume that $a = [b, c]$ where $b, c \in M_n$ have zero diagonal. Hence, we can write $b = [b', s]$, $c = [c', s]$ for some $b', c' \in M_n$. Thus $(1 + \alpha_2)a = f(b', s, s, c')$. If $a \in M_n^0$ has rank one, then a is similar to the matrix unit e_{12} and $(1 + \alpha_2)e_{12} = f(e_{21}, e_{12}, e_{12}, -\frac{1}{2}e_{11})$. Hence, $1 + \alpha_2 = 0$ or $\mathrm{im}(f)$ contains all matrices with zero trace. Therefore we can assume $1 + \alpha_2 = 0$. If we set $x_1 = x_4$ we get in a similar way that $1 - \alpha_2 = 0$, which leads to a contradiction. Hence, $M_n^0 \subseteq \mathrm{im}(f) \subseteq M_n^0$ yields the desired conclusion. □

Proposition 2.6.5. *If f is a nonzero multilinear Lie polynomial of degree at most* 4*, then* $\mathrm{im}(f) = M_n^0$.

Proof. Apply Lemmas 2.6.2, 2.6.3, 2.6.4. □

3. Length of a generic vector space

In this section we assume that F is an infinite field. By $\langle x_1, \dots, x_g\rangle$ we denote the free monoid generated by $x_1, \dots, x_g$, and by $\langle x_1, \dots, x_g\rangle_d$ words in $\langle x_1, \dots, x_g\rangle$ of length d. In case $g = 2$ we write x, y instead of x_1, x_2. The set $\{1, \dots, d\}$ is denoted by $\mathbb{N}_d$.

3.1. Generic matrices and the discriminant.

We define the **discriminant** $\Delta(A^{(1)}, \dots, A^{(n^2)})$ of $n \times n$ matrices $A^{(1)}, \dots, A^{(n^2)}$ to be the determinant of the $n^2 \times n^2$ matrix whose k-th column $v^{(k)}$ is the vectorized matrix $A^{(k)}$; i.e., $v^{(k)}_{(n-1)i+j} = A^{(k)}_{ij}$.

Since we assume that F is infinite we have the following easy observation.

Lemma 3.1.1. *Words $w_1, \dots, w_{n^2} \in \langle x_1, \dots, x_g\rangle$ are $M_n(F)$-locally linearly independent if and only if the discriminant of $w_1(X_1, \dots, X_g), \dots, w_{n^2}(X_1, \dots, X_g)$ is nonzero.*

3.2. Main result on words.

Theorem 3.2.1. *Let $g \geq 2$ and $d = \lceil \log_g n\rceil$. Then there exist $M_n(F)$-locally linearly independent words $w_1, \dots, w_{n^2} \in \langle x_1, \dots, x_g\rangle_{2d}$. That is, for some $A \in M_n(F)^g$ the matrices $w_1(A), \dots, w_{n^2}(A)$ are linearly independent and thus span $M_n(F)$.*

3.2.1. *Length of a vector space.* Let V be a vector subspace of $M_n(F)$. By V^k we denote the vector space spanned by the words of length at most k evaluated at V. The **length** of V is the integer ℓ yielding a stationary chain

$$V \subsetneq V^2 \subsetneq \cdots \subsetneq V^\ell = V^{\ell+1}.$$

We say that words $w_1, \dots, w_t \in \langle x_1, \dots, x_g \rangle$ **sweep** $M_n(F)$ if there exists $A \in M_n(F)^g$ such that $w_1(A), \dots, w_t(A)$ span $M_n(F)$. Given a subset S of $M_n(F)^g$, a vector space $V \subseteq M_n(F)$ of dimension g is S**-generic** if it can be spanned by elements $A_1, \dots, A_g$ satisfying $(A_1, \dots, A_g) \in S$.

COROLLARY 3.2.2 (Generic version of Paz's conjecture). *Let $g \geq 2$ and let F be an infinite field. There exists a nonempty Zariski open subset $S \subseteq M_n(F)^g$ such that the length of an S-generic vector subspace of $M_n(F)$ is of order $O(\log n)$.*

Note that if $F \in \{\mathbb{R}, \mathbb{C}\}$, then a nonemprty Zariski open subset of F^m is automatically dense in the Euclidean topology.

3.2.2. *Words in random matrices span the full matrix algebra.* By Corollary 3.2.2, given a g-tuple A of random $n \times n$ matrices, words of length $O(\log n)$ in A span $M_n(F)$. In particular, we have:

COROLLARY 3.2.3. *For each g satisfying $n^2 \leq g^{2d}$ there exists a set of g matrices such that words of length $2d$ in those matrices span $M_n(F)$.*

Corollary 3.2.3 partially answers a question posed in [**?**], where this result is established in the case $d = 1$. The answer is complete for $n \in \mathbb{N}$ satisfying $g^{2d-1} < n^2 \leq g^{2d}$. The question whether the words of length $2d - 1$ sweep $M_n(F)$ in the case $n^2 \leq g^{2d-1}$ remains.

3.3. Proofs. The proof of Theorem 3.2.1 reduces to the study of a special kind of graphs which we introduce in the first subsection. The main ideas can be revealed already in the case of two variables, and the general case is only notationally more difficult, so we focus on $g = 2$. In the next proposition we state for convenience this special case separately.

PROPOSITION 3.3.1. *Let $d = \lceil \log_2 n \rceil$. There exist $M_n(F)$-locally linearly independent words $w_1, \dots, w_{n^2} \in \langle x, y \rangle_{2d}$.*

3.3.1. *Graphs.* We recursively construct a family of graphs $(G_d)_{d \in \mathbb{N}}$. Let G_0 be a graph with one vertex labeled by 1 and no edges. We let G_d be a (directed) graph with 2^d vertices labeled by $\mathbb{N}_{2^d}$, and define the edges as follows. There is a (directed) edge from i to j in G_d of multiplicity $4e$, $1 \leq i, j \leq 2^{d-1}$, if there is an edge of multiplicity e from i to j in G_{d-1}. Moreover, each vertex i, $1 \leq i \leq 2^{d-1}$, has additionally 2^{d+1} loops, and there are 2^d edges from i to $i + 2^{d-1}$ and back for $1 \leq i \leq 2^{d-1}$. We label the loops by x, and other edges by y. Note that G_d contains $2^{2d+1}d$ edges, half of them labeled by x and the other half by y.

For example, the figures below show G_1 and G_2 with the numbers on edges corresponding to their respective multiplicities, and instead of the labels x and y we use dashed (resp. solid) edges for the edges corresponding to x (resp. y).

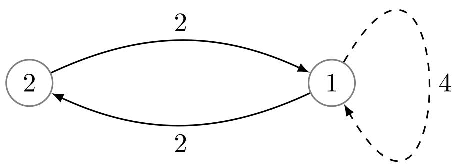

FIGURE 3.1. G_1

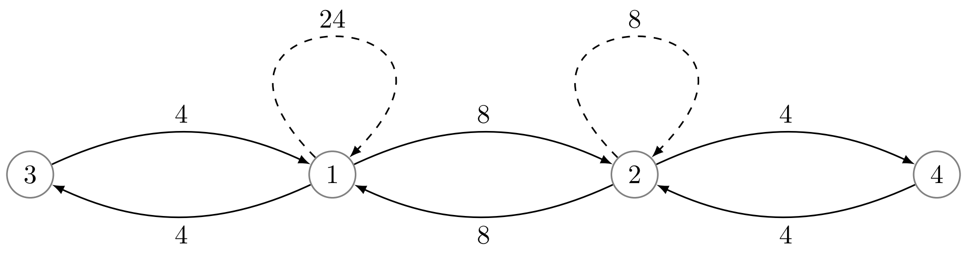

FIGURE 3.2. G_2

If p is a walk in the graph G_d then we associate to it a word corresponding to the labels on edges passed by p in the respective order.

We now state a technical lemma that will be extensively used in the proof of Proposition 3.3.1.

LEMMA 3.3.2. *There is a unique way of partitioning the graph G_d in 2^{2d} (edge-disjoint) walks p_{ij} of length $2d$, $1 \le i, j \le 2^d$, such that p_{ij} starts at i and ends at j, which yield all the words in $\langle x, y\rangle_{2d}$.*

PROOF. We prove the lemma by induction on d. Let us denote by $G_d^{(m)}$ the graph obtained from G_d by multiplying the multiplicity of each edge by m.

We claim that $G_d^{(m)}$ can be in only one way partitioned into $m2^{2d}$ walks of length $2d$ such that m walks start at i and end at j, $1 \le i, j \le 2^d$, and such that each word in $\langle x, y\rangle_{2d}$ corresponds to m walks. Consider first $G_1^{(m)}$. Then the only way of obtaining the desired partition is to take m walks $\{2 \to 1, 1 \to 2\}$, m walks $\{1 \to 1, 1 \to 2\}$, m walks $\{2 \to 1, 1 \to 1\}$ and m walks $\{1 \to 1, 1 \to 1\}$ as can be easily seen. Suppose that the claim holds for all graphs $G_\ell^{(m)}$, $\ell < d$. Consider now $G_d^{(m)}$. Since there are no loops on the vertices labeled by i, $2^{d-1}+1 \le i \le 2^d$, all words with the starting or ending point in these vertices need to begin, respectively end, with y. By the condition on the partition there are exactly half of walks with this property, thus words with other starting, resp. ending, point need to begin, resp. end, with x. Removing the edges starting or ending at i, $2^{d-1}+1 \le i \le 2^d$, and $m2^{d+1}$ loops on vertices i, $1 \le i \le 2^{d-1}$, we obtain a graph on 2^{d-1} vertices labeled by $\mathbb{N}_{2^{d-1}}$ (ignoring the isolated points) which coincides with $G_{d-1}^{(4m)}$ by construction, and which we need to partition into $4m2^{2(d-1)}$ walks of length $2(d-1)$ (as we have already removed the starting and the ending edge of walks in $G_d^{(m)}$) such that $4m$ walks start at i and end at j, $1 \le i, j \le 2^{d-1}$, and each word in $\langle x, y\rangle_{2d-2}$ corresponds to $4m$ walks. By the induction hypothesis, there is only one such a partition. The lemma thus follows by taking $m = 1$. □

For the proof of Theorem 3.2.1 we will need a slight generalization of the previous lemma. We thus introduce a graph G_d^g which has g^d vertices and is defined recursively by setting G_0^g to be the graph with 1 vertex labeled by 1 and no edges. Having constructed G_{d-1}^g we let G_d^g be a directed graph with g^d vertices labeled by $\mathbb{N}_{g^d}$, and having a (directed) edge from i to j of multiplicity g^2e, $i, j \in \mathbb{N}_{g^{d-1}}$, if there is an edge of multiplicity e from i to j in G_{d-1}^g, and is labeled as the corresponding edge in G_{d-1}^g, and there are g^d edges from i to $i+(k-1)g^{d-1}$ and back for $1 \le i \le g^{d-1}$, labeled by x_k, $1 \le k \le g$ (for $k = 1$ every loop has multiplicity g^{2d}). Note that G_d^g contains $2dg^{2d}$ edges.

LEMMA 3.3.3. *There is a unique partition of the graph G_d^g in g^{2d} (edge-disjoint) walks p_{ij} of length $2d$, $1 \le i,j \le 2^d$, such that p_{ij} starts at i and ends at j, which yield all the words in $\langle x_1,\dots,x_g\rangle_{2d}$.*

The proof of Lemma 3.3.3 is a straightforward modification of Lemma 3.3.2 and is omitted.

3.3.2. *Proof of Theorem 3.2.1.*

PROOF OF PROPOSITION 3.3.1. By Lemma 3.1.1 we need to show that

$$p(x_{11},\dots,x_{nn},y_{11},\dots,y_{nn}) := \Delta(w_1(X,Y),\dots,w_{n^2}(X,Y))$$

is nonzero for some $w_1,\dots,w_{n^2} \in \langle x,y\rangle_{2d}$, where $X=(x_{ij})$, $Y=(y_{ij})$ are generic $n\times n$ matrices. (We may and we will assume that X is diagonal [**Row80**, Proposition 1.3.15].) By the definition of the discriminant,

$$p(x_{11},\dots,x_{nn},y_{11},\dots,y_{nn}) = \sum_{\sigma\in S_{n^2}} (-1)^\sigma \prod_{1\le i,j\le n} w_{\sigma(k_{ij})}(X,Y)_{ij}, \tag{3.1}$$

where $k_{ij}=(i-1)n+j$, and $w_k(X,Y)_{ij}$ denotes the commutative polynomial at the entry (i,j) of the word w_k evaluated at the generic matrices X,Y.

Let us define the lexicographic order on $\langle x,y\rangle$ with $x>y$ and denote by v_s the vector of the words of length s listed decreasingly with respect to this order. Let

$$a^{(n)} = e_n v_d v_d^t e_n = \begin{pmatrix} x v_{d-1} v_{d-1}^t x & x v_{d-1} v_{d-1}^t y e_{n'} \\ e_{n'} y v_{d-1} v_{d-1}^t x & e_{n'} y v_{d-1} v_{d-1}^t y e_{n'} \end{pmatrix} \tag{3.2}$$

be the block matrix consisting of words, where $n' = n - 2^{d-1}$ and $e_{n'} = e_{11} + \cdots + e_{n'n'}$, with blocks of the size $2^{d-1}\times 2^{d-1}$, $2^{d-1}\times n'$, $n'\times 2^{d-1}$, and $n'\times n'$, respectively.

We proceed to find a monomial that appears in the product on the right-hand side of (3.1) for a unique $\sigma\in S_{n^2}$. Let $2^{d-1} < n \le 2^d$ and $x_i = x_{ii}$. We define

$$\overline{m}_n^{(ij)} = \begin{cases} x_i x_j & \text{if } 1\le i,j\le 2^{d-1}, \\ x_i y_{j-2^{d-1},j} & \text{if } 1\le i\le 2^{d-1}, 2^{d-1}<j\le n, \\ x_j y_{i,i-2^{d-1}} & \text{if } 2^{d-1}<i\le n, 1\le j\le 2^{d-1}, \\ y_{i,i-2^{d-1}} y_{j-2^{d-1},j} & \text{if } 2^{d-1}<i,j\le n. \end{cases}$$

We further inductively define

$$m_1^{(11)} = 1, \quad m_n^{(ij)} = m_{2^{d-1}}^{(i_d,j_d)} \overline{m}_n^{(ij)},$$

where $i_d \equiv i \bmod 2^{d-1}$, $j_d \equiv j \bmod 2^{d-1}$, $1 \le i_d, j_d \le 2^{d-1}$. Consider the monomial defined by

$$m_n = \prod_{1\le i,j\le n} m_n^{(ij)}.$$

In particular, in the case $n = 2^d$ we have

$$m_{2^d} = (m_{2^{d-1}})^4 \prod_{1\le i,j\le 2^d} \overline{m}_{2^d}^{(ij)}.$$

Let $w_{(i-1)n+j} = a_{ij}^{(n)}$ for $a^{(n)}$ defined in (3.2). We claim that m_n appears in

$$P_n^\sigma = \prod_{1\le i,j\le n} w_{\sigma(k_{ij})}(X,Y)_{ij}$$

only for $\sigma = \mathrm{id}$. By the construction of $m_n^{(ij)}, m_n$ and (3.2), $m_n^{(ij)}$ has a nonzero coefficient in the commutative polynomial $a_{ij}^{(n)}(X,Y)_{ij}$ and thus the same holds for the monomial m_n in

$$\prod_{1\le i,j\le n} w_{k_{ij}}(X,Y)_{ij} = \prod_{1\le i,j\le n} a_{ij}^{(n)}(X,Y)_{ij}.$$

It remains to show that m_n does not appear in P_n^σ for $\sigma \neq \mathrm{id}$. For this we use a graph-theoretic language.

We first consider the case $n = 2^d$. We can present the monomial m_n as a walk in a graph on n vertices, in which there is a directed edge of multiplicity s_{ij} between vertices i and j labeled by y if s_{ij} is the degree of y_{ij} in the monomial m_n, and there are s_i loops on the vertex i labeled by x if s_i is the degree of x_i in m_n. Since m_n needs to be written as a product of n^2 monomials u_{ij}, $1 \leq i, j \leq n$, arising from monomials in $w_{\sigma(k_{ij})}(X,Y)_{ij}$, $w_{k_{ij}} \in \langle x, y\rangle_{2d}$, our problem reduces to finding partitions of the graph associated to m_n into n^2 walks p_{ij}, $1 \leq i, j \leq n$, of length $2d$ that yield all the words in $\langle x, y\rangle_{2d}$. Lemma 3.3.2 asserts that there is only one such partition, and thus concludes the proof in the case $n = 2^d$.

For arbitrary n we observe that if

$$\sum_{\sigma\in S_{n^2}} (-1)^\sigma \prod_{1\leq i,j\leq n} a^{(n)}_{\sigma((i-1)n+j)}(X,Y)_{ij}$$

equaled 0 then m_n would appear in P_n^{id} and in some other P_n^ρ with $\rho \in S_{n^2}$. As $m_n^{(ij)}$ can be identified with $m_{2^d}^{(ij)}$ for $1 \leq i, j \leq n$, m_{2^d} would have a nonzero coefficient in the product $P_{2^d}^\sigma$ for $\sigma = \mathrm{id}$ and $\sigma = \tilde\rho \in S_{2^d}$ (here $\tilde\rho$ is the permutation in S_{2^d} induced by ρ and fixing all $i > n$), which is impossible by the claim proved in the previous paragraph. Thus the words $a_{11}^{(n)}, \dots, a_{nn}^{(n)}$ are $M_n(F)$-locally linearly independent. □

Proof of Theorem 3.2.1. One follows the steps of the proof of Proposition 3.3.1 where initially one needs to consider the case $n = g^d$, and defines the monomial m_n inductively corresponding to the matrix $a^{(n)} = e_n v_d v_d^t e_n$ equal to

$$\begin{pmatrix} x_1 v_{d-1} v_{d-1}^t x_1 & \dots & x_1 v_{d-1} v_{d-1}^t x_{g-1} & x_1 v_{d-1} v_{d-1}^t x_g e_{n'} \\ \vdots & \ddots & \vdots & \vdots \\ x_{g-1} v_{d-1} v_{d-1}^t x_1 & \dots & x_{g-1} v_{d-1} v_{d-1}^t x_{g-1} & x_{g-1} v_{d-1} v_{d-1}^t x_g e_{n'} \\ e_{n'} x_g v_{d-1} v_{d-1}^t x_1 & \dots & e_{n'} x_g v_{d-1} v_{d-1}^t x_{g-1} & e_{n'} x_g v_{d-1} v_{d-1}^t x_g e_{n'} \end{pmatrix}, \tag{3.3}$$

where $g^{d-1} < n \leq g^d$, $n' = n - g^{d-1}$, and v_s denotes the vector of g^s words of length s listed decreasingly in the monomial order induced by setting $x_1 > \dots > x_g$. Instead of applying Lemma 3.3.2 one concludes the proof by applying Lemma 3.3.3. □

Proof of Corollary 3.2.2. Let $w_1, \dots, w_{n^2}$ be the $M_n(F)$-locally linearly independent words of length $2d = 2\lceil \log_g n\rceil$ whose existence was established in Theorem 3.2.1. As then the discriminant $\Delta := \Delta(w_1(X_1,\dots,X_g),\dots,w_{n^2}(X_1,\dots,X_g))$ is nonzero, the subset S of $A \in M_n(F)^g$ where Δ does not vanish is a nonempty Zariski open subset, and therefore dense in $M_n(F)^g$. In the case $F \in \{\mathbb{R}, \mathbb{C}\}$, S is also dense in the Euclidean topology. By the definition of S it follows that every S-generic vector subspace of $M_n(F)$ is of length $O(\log n)$. □

Proof of Corollary 3.2.3. Choose $A_1, \dots, A_{\overline{g}}$ such that $(A_1, \dots, A_{\overline{g}}) \in S \subseteq M_n(F)^{\overline{g}}$, where $\overline{g}$ is the smallest among $m \in \mathbb{N}$ satisfying $m^d \geq n$, and S is the set described in the proof of Corollary 3.2.2. For the remaining $A_{\overline{g}+1}, \dots, A_g$ arbitrary $n \times n$ matrices will do. □

Remark 3.3.4. (a) It is not difficult to see that one can take the matrices in Corollary 3.2.3 to be symmetric. One only needs to note that the proof of Lemma 3.3.3 also works for the undirected version of the graphs G_d^g and then use these in the proof of Theorem 3.2.1.

(b) The proof of Proposition 3.3.1 also leads to an explicit construction of $n \times n$ matrices such that words of degree $2d = \lceil 2\log_2 n\rceil$ in these matrices span

$M_n(F)$. We give an example in characteristic 0. Keep the notation from the proof of Corollary 3.2.3. Let $M = n!(n^{2d-1})^n$. We set all variables that do not appear in m_n to zero, and denote by $\mathcal{C}'$ the polynomial algebra in the remaining variables. Let us order the variables as follows: $x^{(k)}_{i,i+(k-1)\bar{g}^{s-1}} < x^{(\ell)}_{j,j+(k-1)\bar{g}^{t-1}}$, (resp. $x^{(k)}_{i,i+(k-1)\bar{g}^{s-1}} < x^{(\ell)}_{j+(k-1)\bar{g}^{t-1},j}$), if $(s,k,i) < (t,\ell,j)$ (resp. $(s,k,i) \le (t,\ell,j)$) in the lexicographic order, and take the corresponding lexicographic ordering on $\mathcal{C}'$. Let

$$c_1 = 3, \quad c_s = 2\bar{g}^{s-1}(\bar{g}-1) + \bar{g}^{s-2} \ \ (s>1), \quad c = \sum_{s=1}^{d} c_s.$$

We further define for $g^{s-2} < i \le \bar{g}^{s-1}$ and $1 \le j \le \bar{g}^{s-1}$,

$$f_{1,s,i^+} = f_{1,s,i^-} = \sum_{t=0}^{s-1} c_t + j \ \ (\bar{g}^{s-2} < j \le \bar{g}^{s-1}),$$

$$f_{k,s,j^+} = \sum_{t=0}^{s-1} c_t + \bar{g}^{s-2} + 2\bar{g}^{s-1}(k-2) + j,$$

$$f_{k,s,j^-} = \sum_{t=0}^{s-1} c_t + \bar{g}^{s-2} + 2\bar{g}^{s-1}(k-2) + \bar{g}^{s-1} + j.$$

We set

$$A^{(k)}_{i,i+(k-1)\bar{g}^{s-1}} = M^{2d(c-f_{k,s,i^+})},$$
$$A^{(k)}_{i+(k-1)\bar{g}^{s-1},i} = M^{2d(c-f_{k,s,i^-})}.$$

Since the monomial m_n is the maximal monomial in $\mathcal{C}'$, the degree of monomials appearing in $\Delta(w_1(X_1,\ldots,X_g),\ldots,w_{n^2}(X_1,\ldots,X_{\bar{g}}))$ is $2d$, and there appear at most M monomials in Δ (counted with multiplicity). It is easy to see that the constructed $A^{(k)}$, $1 \le k \le \bar{g}$, and arbitrary $A^{(k)}$, $\bar{g} < k \le g$, have the desired property of Corollary 3.2.3.

(c) It would be interesting to know whether *arbitrary* n^2 words in x, y of fixed length $d \ge \lceil 2\log_2 n \rceil$ sweep $M_n(F)$. If the answer were positive then we could deduce that a quasi-identity of $M_n(F)$ (see [**?**] for the definition) $\sum_M \lambda_M M$ with $\deg M = d$ cannot be a sum of fewer than n^2 monomials, and this bound is sharp. This should be seen in contrast with [**Row80**, Exercise 7.2.3], stating that a multilinear polynomial identity of $M_n(F)$ cannot be a sum of fewer than 2^n monomials. However, the sharp bound is to the best of our knowledge not known.

4. Noncommutative crepant resolutions

In this section we construct twisted noncommutative crepant resolutions for centers of trace rings.

4.1. Noncommutative resolutions. We first need to adopt some standard definitions. Let S be a normal Noetherian domain with quotient field K. A finitely generated S-module M is said to be *reflexive* if the canonical map $M \mapsto M^{**}$ is an isomorphism. This implies in particular that M is torsion free. Reflexive modules are not affected by codimension two phenomena. For example a morphism $\phi : M \to N$ between reflexive modules is an isomorphism if this is the case for all $\phi_P : M_P \to N_P$ where P runs through the height one primes in S.

The category $\mathrm{ref}(S)$ of reflexive S-modules is a rigid symmetric monoidal category with tensor product $(M,N) \mapsto (M \otimes_S N)^{**}$. This implies that many concepts for S-modules and S-algebras have a natural "reflexive" analogue. For example a *reflexive Azumaya algebra* [**LB89**] A is a non-zero S-algebra A which is a reflexive S-module such that the natural map $A \otimes_S A^\circ \to \mathrm{End}_S(A,A) : a \otimes b \mapsto (x \mapsto axb)$

becomes an isomorphism after applying $(-)^{**}$. Such reflexive notions will be used without further comment below. A reflexive Azumaya algebra A is said to be trivial if it is of the form $\operatorname{End}_S(M)$ for M a reflexive S-module. In that case $\operatorname{ref}(S)$ and $\operatorname{ref}(A)$ are equivalent. This is a particular case of "reflexive Morita equivalence" which is defined in the obvious way.

DEFINITION 4.1.1. A *twisted noncommutative resolution* of S is a reflexive Azumaya algebra A over S such that $\operatorname{gl}\dim A < \infty$. If A is trivial then A is said to be a *noncommutative resolution (NCR)* of S.

DEFINITION 4.1.2. Assume that S is Gorenstein. A *twisted noncommutative crepant resolution* A of S is a twisted NCR of S which is in addition a Cohen-Macaulay S-module. If A is trivial then such A is said to be a *noncommutative crepant resolution (NCCR)* of S.

4.2. Notation. We first introduce some notations which will remain in force for this section except when overruled locally. Throughout this section we assume that F is an algebraically closed field. Let G be a connected reductive group. Let $T \subset B \subset G$ be respectively a maximal torus and a Borel subgroup of G with $\mathcal{W} = N(T)/T$ being the corresponding Weyl group. Put $X(T) = \operatorname{Hom}(T, G_m)$ and let $\Phi \subset X(T)$ be the roots of G. By convention the roots of B are the negative roots Φ^- and $\Phi^+ = \Phi - \Phi^-$ is the set of positive roots. We write $\bar{\rho} \in X(T)_\mathbb{R}$ for half the sum of the positive roots. Let $X(T)^+_\mathbb{R}$ be the dominant cone in $X(T)_\mathbb{R}$ and let $X(T)^+ = X(T)^+_\mathbb{R} \cap X(T)$ be the set of dominant weights. For $\chi \in X(T)^+$ we denote the simple G-representation with highest weight χ by $V(\chi)$.

Let W be a finite dimensional G-representation of dimension d and put $R = SW$, $X = \operatorname{Spec} SW = W^*$. Let $(\beta_i)_{i=1}^d \in X(T)$ be the T-weights of W.

Put
$$\Sigma = \left\{ \sum_i a_i \beta_i \mid a_i \in]-1, 0] \right\} \subset X(T)_\mathbb{R}.$$

4.3. Modules of covariants.

4.3.1. *Preliminaries.* Let G be a reductive group and let R be the coordinate ring of a smooth connected affine G-variety X with function field K. Let $\operatorname{mod}(G, R)$ be the category of G-equivariant finitely generated R-modules. The following is standard.

LEMMA 4.3.1. *(1) The objects $U \otimes R$ with $U \in \hat{G}$ form a family of projective generators for* $\operatorname{mod}(G, R)$.

(2) If $P \in \operatorname{mod}(G, R)$ is projective as R-module then it is a projective object in $\operatorname{mod}(G, R)$.

(3) $\operatorname{gl}\dim \operatorname{mod}(G, R) \leq \operatorname{gl}\dim R$.

(4) If X has a fixed point then $\operatorname{gl}\dim \operatorname{mod}(G, R) = \operatorname{gl}\dim R$.

PROOF. The first statement is clear and the third statement follows from the second and the first combined with the fact that R has finite global dimension. For the second statement choose a G-equivariant surjection $\phi : U \otimes R \to P$ with U a representation of G and an R-linear splitting $\theta : P \to U \otimes R$ for ϕ. Applying the Reynolds operator R to the identity $\phi\theta = \mathrm{id}$ yields $\phi R(\theta) = R(\phi\theta) = R(\mathrm{id}) = \mathrm{id}$. Hence $R(\theta)$ is a G-equivariant splitting for ϕ and thus P is projective in $\operatorname{mod}(G, R)$. Finally to prove (4) let x be the fixed point. Then we see that a G-equivariant projective resolution of $k(x)$ must have at least length $\dim X$ since this is true if we forget the G-action. This proves $\operatorname{gl}\dim \operatorname{mod}(G, R) \geq \dim X = \operatorname{gl}\dim R$. $\square$

Now let $\bar{G} \to G$ be a central extension of G with kernel A where $\bar{G}$ is also reductive. Put $X(A) = \operatorname{Hom}(A, G_m)$. For $\chi \in X(A)$ let $\operatorname{mod}(\bar{G}, R)_\chi$ be the abelian category of $\bar{G}$-equivariant finitely generated R-modules on which A acts through the character χ. Note $\mathrm{mod}(\bar{G}, R)_0 = \mathrm{mod}(G, R)$ and furthermore $\mathrm{mod}(\bar{G}, R) = \bigoplus_{\chi \in X(A)} \mathrm{mod}(\bar{G}, R)_\chi$. Clearly Lemma 4.3.1 extends to $\mathrm{mod}(\bar{G}, R)_\chi$.

Definition 4.3.2. We say that G acts *generically* on a smooth affine variety X if

(1) X contains a point with closed orbit and trivial stabilizer.
(2) If $X^{\mathbf{s}} \subset X$ is the locus of points that satisfy (1) then $\mathrm{codim}(X - X^{\mathbf{s}}) \geq 2$.

Lemma 4.3.3. *Assume that G acts generically on X. Let $\mathrm{ref}(G, R)$ be the category of G-equivariant R-modules which are reflexive as R-modules. Then the functors*

$$\mathrm{ref}(G, R) \mapsto \mathrm{ref}(R^G) : M \mapsto M^G,$$
$$\mathrm{ref}(R^G) \mapsto \mathrm{ref}(G, R) : N \mapsto (R \otimes_{R^G} N)^{**}$$

are inverse equivalences between the symmetric monoidal categories $\mathrm{ref}(G, R)$ and $\mathrm{ref}(R^G)$.

Proof. The hypothesis imply that $X \to X /\!\!/ G$ contracts no divisor. This implies that $(-)^G$ preserves reflexive modules by [**Bri93**, Prop. 1.3].

For the remaining part of the lemma we have to show that for $M \in \mathrm{ref}(G, R)$ the map $(R \otimes_{R^G} M^G)^{**} \to M$ is an isomorphism, or equivalently that $\phi : R \otimes_{R^G} M^G \to M$ is an isomorphism outside a closed subset of X of codimension two. We will show that ϕ is an isomorphism in a neighborhood of any point x of $X^{\mathbf{s}}$ (see Definition 4.3.2).

Up to restricting to a suitable G-invariant affine etale neighborhood of x we may assume by the Luna slice theorem [**Lun73**] that $X = G \times S$. But then the result is clear by descent. □

Let U be finite dimensional G-representation. Then

$$M(U) \overset{\mathrm{def}}{=} (U \otimes R)^G$$

is a finitely generated R^G-module which is called the *module of covariants* associated to U. Sometimes we use additional decorations such as $M_G(U)$, $M_{G,R}(U)$ to indicate context.

Corollary 4.3.4. *Assume that G acts generically on X. Then*

$$\mathrm{mod}(G, k) \mapsto \mathrm{ref}(R^G) : U \mapsto M(U)$$

is a symmetric monoidal functor.

Remark 4.3.5. In the examples we have to verify the genericity condition. This is routine but sometimes a bit messy. Here we outline how one may do the verification in general but in practice there are usually short cuts. Assume that $X = W^*$ is a G-representation. Let X^s be the locus of points in X which are *stable*, i.e. which have closed orbit and finite stabilizer. Then $X - X^s$ may be described by the usual numerical criterion in terms of one-parameter subgroups [**MF82**]. This may be used to bound the dimension of $X - X^s$.

The locus X_p of points in X which are stabilized by an element of prime order p may also be described numerically using one parameter subgroups. Indeed assume that $g \in G$ has order p and $x \in X$ is such that $gx = x$. The element g is in particular semi-simple and hence it is contained in a maximal torus and therefore in the image of an injective one-parameter subgroup $\lambda : G_m \to G$. Let $(\mu_i)_i \in \mathbb{Z}$ be the weights for a diagonalization of the λ-action on X. Then since x is a fixed point for $\lambda^{-1}(g)$ which has order p we see that if $p \nmid \mu_i$ we must have $x_i = 0$. In this way one may bound the dimension of X_p. Since $X^{\mathbf{s}} = X^s - \bigcup_p X_p$ we are done.

LEMMA 4.3.6. *Assume that G acts generically on X and that X contains a fixed point. If $M(U) \cong M(U')$ as R^G-modules then $U \cong U'$ as G-representations.*

PROOF. By Lemma 4.3.3 $U \otimes R \cong U' \otimes R$ in $\mathrm{ref}(G, R)$. Specializing at the fixed point yields what we want. □

Since $\mathrm{End}(U)$ is a G-equivariant k-algebra we obtain that $M(\mathrm{End}(U))$ is an R^G-algebra. We will call it an *algebra of covariants*. From Lemma 4.3.3 we obtain that if G acts generically then

$$M(\mathrm{End}(U)) = \mathrm{End}_{R^G}(M(U)).$$

Now assume that $U \in \mathrm{mod}(\bar{G}, k)_\chi$. If $\chi \neq 0$ then $(U \otimes R)^{\bar{G}} = 0$. However A acts trivially on $\mathrm{End}(U)$ so the algebra of covariants $M(\mathrm{End}(U))$ is still interesting.

For further reference recall the following result.

THEOREM 4.3.7. ***[Kno86]*** *Assume that W is a generic unimodular ($\det W \cong k$) G-representation and $R = SW$. Then R^G is Gorenstein.*

4.3.2. *Cohen-Macaulayness of modules of covariants.* We will be interested in sufficient criteria for $M(U)$ to be Cohen-Macaulay for U a finite dimensional representation of G.

A relevant conjecture in the connected case was stated in [**Sta79**] and this conjecture was almost completely proved in [**VdB91**]. Below we use those results to obtain easy to verify criteria for Cohen-Macaulayness in the cases that interest us.

DEFINITION 4.3.8. The elements of the intersection

$$X(T)^+ \cap (-2\bar{\rho} + \Sigma)$$

are called *strongly critical (dominant) weights* for G.

LEMMA 4.3.9. *Assume that χ is strongly critical for G and let the T-weights of $U = V(\chi)$ be given by $(\chi_i)_i$. Then for any $S \subset \Phi$ and for any i we have that $\chi_i + \sum_{\rho \in S} \rho$ is strongly critical for T.*

PROOF. Let Γ be the convex polygon

$$\Gamma = \{\sum_{\rho \in \Phi} u_\rho \rho \mid u_\rho \in [0, 1]\}.$$

For every $S \subset \Phi$ we have $\sum_{\rho \in S} \rho \in \Gamma$ and moreover by Corollary 4.B.3 below Γ is the convex hull of the $\mathcal{W}$-orbit of $2\bar{\rho}$.

Similarly all χ_i are contained in the convex hull of the $\mathcal{W}$-orbit of χ by [**FH91**, Thm 14.18]. Hence we have to prove that for all $v, w \in \mathcal{W}$ one has $v\chi + w(2\bar{\rho}) \in \Sigma$. This follows from Lemma 4.D.1 below since Σ is convex and $\mathcal{W}$-invariant, $2\bar{\rho}$ and χ are dominant and finally by hypothesis $\chi + 2\bar{\rho} \in \Sigma$. □

Recall that a stable point is a point with closed orbit and finite stabilizer.

THEOREM 4.3.10. *Assume X contains a stable point. Let $\chi \in X(T)^+$ be a strongly critical weight and $U = V(\chi)$. Then $M(U^*)$ is a Cohen-Macaulay R^G-module.*

PROOF. Let R_U be the isotypical component of R corresponding to U. I.e. R_U is the sum of all subrepresentations of R isomorphic to U. One has $R_U = U \otimes M(U^*)$. Hence $M(U^*)$ is Cohen-Macaulay if and only if R_U is Cohen-Macaulay.

Let $(\chi_i)_i$ be the T-weights of U. According to [**VdB91**, Thm 1.3] R_U will be Cohen-Macaulay if for every i and for every $S \subset \Phi$ one has that $\chi_i + \sum_{\rho \in S} \rho \in \Sigma$, or equivalently if $\chi_i + \sum_{\rho \in S} \rho$ is strongly critical for T. This condition holds by Lemma 4.3.9. □

PROPOSITION 4.3.11. *Assume that X contains a stable point. Let $\chi_1, \chi_2 \in X(T)^+$. If $\chi_1+\chi_2$ is strongly critical then $M((V(\chi_1)\otimes V(\chi_2))^*)$ is Cohen-Macaualay.*

PROOF. Assume that $V(\chi)$ with $\chi \in X(T)^+$ is a summand of $V(\chi_1) \otimes V(\chi_2)$. Then by [**FH91**, Ex. 25.33] $\chi = \chi_1+\chi'$ with χ' a weight of $V(\chi_2)$. Hence by [**FH91**, Thm 14.18] χ' is in the convex hull of the $\mathcal{W}$-orbit of χ_2. Hence by Theorem 4.3.10 we have to prove that $\chi_1 + w\chi_2 + 2\bar{\rho} \in \Sigma$ for $w \in \mathcal{W}$. This follows again from Lemma 4.D.1 below since $\chi_1 + 2\bar{\rho}$ and χ_2 are dominant and by hypothesis $\chi_1 + \chi_2 + 2\bar{\rho} \in \Sigma$. $\square$

4.4. Noncommutative resolutions and noncommutative crepant resolutions. We first state results, proofs will follow in the next subsection.

THEOREM 4.4.1. *(See §4.5.3 below.) Let Δ be a $\mathcal{W}$-invariant bounded closed convex subset of $X(T)_{\mathbb{R}}$. Let*

$$\begin{aligned}\mathcal{L} &= X(T)^+ \cap (-\bar{\rho} + \Sigma + \Delta),\\ U &= \bigoplus_{\chi\in\mathcal{L}} V(\chi).\end{aligned}$$

Then one has $\operatorname{gl}\dim M(\operatorname{End}(U)) < \infty$.

Note that we may always take Δ in such a way that $0 \in \mathcal{L}$ (e.g. let Δ be the convex hull of $\mathcal{W}\cdot\bar{\rho}$). In that case $U \neq 0$.

COROLLARY 4.4.2. *If W is generic then for Δ such that $U \neq 0$ in Theorem 4.4.1 one has that $M(U)$ yields a NCR of R^G.*

PROOF. This follows Theorem 4.4.1 together with the fact that if the action is generic then by Lemma 4.3.3

$$M(\operatorname{End}(U)) = \operatorname{End}_{R^G}(M(U))$$

and furthermore $M(U)$ is reflexive. $\square$

Let the notations be as in the previous section. We will say that W is *quasi-symmetric* if for every line $\ell \subset X(T)_{\mathbb{R}}$ through the origin we have

$$\sum_{\beta_i\in\ell} \beta_i = 0.$$

This implies in particular that W is unimodular (i.e. $\wedge^d W \cong k$) and hence R^G is Gorenstein if W is generic by Theorem 4.3.7.

The following result strengthens Theorem 4.4.1 in the quasi-symmetric case.

THEOREM 4.4.3. *(See §4.6.1 below.) Assume W is quasi-symmetric. Let Δ be a $\mathcal{W}$-invariant bounded closed convex subset of $X(T)_{\mathbb{R}}$.*

Put

$$\begin{aligned}\mathcal{L} &= X(T)^+ \cap (-\bar{\rho} + (1/2)\bar{\Sigma} + \Delta),\\ U &= \bigoplus_{\chi\in\mathcal{L}} V(\chi).\end{aligned}$$

Then one has $\operatorname{gl}\dim M(\operatorname{End}(U)) < \infty$.

If W is generic and $\mathcal{L} \neq \emptyset$ then this yields again a NCR as in Corollary 4.4.2. However our main concern in the quasi-symmetric case will be the construction of NCCRs rather than just NCRs. For this we need the concept of a *half open polygon* which generalizes the notion of a half open interval. Let $\Delta \subset \mathbb{R}^n$ be a bounded closed convex polygon. For $\varepsilon \in \mathbb{R}^n$ parallel to the linear space spanned by Δ put

$$\begin{aligned}\Delta_\varepsilon &= \bigcup_{r>0} \Delta \cap (r\varepsilon + \Delta),\\ \Delta_{\pm\varepsilon} &= \Delta_\varepsilon \cap \Delta_{-\varepsilon}.\end{aligned}$$

So Δ_ε is obtained from Δ by removing the boundary faces which are moved inwards by ε and $\Delta_{\pm\varepsilon}$ is obtained from Δ by removing the boundary faces not parallel to ε.

We will say that $\varepsilon \in \mathbb{R}^n$ is *generic* for Δ if it is a non-zero vector which is parallel to Δ, but not parallel to any of its boundary faces. In that case $\Delta_{\pm\varepsilon}$ is the relative interior of Δ.

Our main result concerning NCCRs is Theorem 4.4.5 below, It will be convenient to expand our setting slightly. So we assume that in addition to the connected G that there is a surjective morphism $\phi : \bar{G} \to G$ where $\bar{G}$ is a connected reductive group with $\dim \bar{G} = \dim G$. Then $A \overset{\text{def}}{=} \ker \phi$ is a finite subgroup of the center of $\bar{G}$. Let $\bar{T} \subset \bar{G}$ be the inverse image of T in G. Then $\bar{T}$ is still a maximal torus and $A \subset \bar{T}$. We have an exact sequence

$$0 \to X(T) \to X(\bar{T}) \to X(A) \to 0$$

and hence a corresponding coset decomposition

$$X(\bar{T}) = \bigcup_{\bar{\mu} \in X(A)} X(T)_{\bar{\mu}},$$

where $X(T)_{\bar{\mu}} = X(T) + \mu$. We set $X(T)^+_{\bar{\mu}} = X(T)^+_{\mathbb{R}} \cap X(T)_{\bar{\mu}}$. Our first result is a strengthening of Theorem 4.4.3.

THEOREM 4.4.4. *(See §4.6.2 below.) Let notation be as above. Assume W is quasi-symmetric. Let $\varepsilon \in X(T)_{\mathbb{R}}$ be $\mathcal{W}$-invariant and $\bar{\mu} \in A$. Put*

$$\begin{aligned}\mathcal{L} &= X(T)^+_{\bar{\mu}} \cap (-\bar{\rho} + (1/2)\bar{\Sigma}_\varepsilon),\\ U &= \bigoplus_{\chi \in \mathcal{L}} V(\chi),\\ \Lambda &= M(\operatorname{End}(U)).\end{aligned}$$

Then one has $\operatorname{gl\,dim} \Lambda < \infty$. *If W is is in addition generic and $\mathcal{L} \neq \emptyset$ then Λ is a twisted NCR for R^G.*

We now give the criterion for the existence of (twisted) NCCRs we will use.

THEOREM 4.4.5. *(See §4.6.3 below.) Assume that W is quasi-symmetric and generic. Assume that $\varepsilon \in X(T)_{\mathbb{R}}$ is $\mathcal{W}$-invariant and let $\bar{\mu}$ be such that*

$$X(T)_{\bar{\mu}} \cap (-\bar{\rho} + (1/2)(\bar{\Sigma}_{\pm\varepsilon} - \Sigma)) = \emptyset. \tag{4.1}$$

Put

$$\begin{aligned}\mathcal{L} &= X(T)^+_{\bar{\mu}} \cap (-\bar{\rho} + (1/2)\bar{\Sigma}_\varepsilon),\\ U &= \bigoplus_{\chi \in \mathcal{L}} V(\chi),\\ \Lambda &= M(\operatorname{End}(U)).\end{aligned}$$

If $\mathcal{L} \neq \emptyset$ then Λ is a twisted NCCR for R^G.

4.5. Proofs for noncommutative resolutions.

4.5.1. *Preliminaries.* For more unexplained notation and terminology regarding root systems we refer to Appendix 4.D.

Let $\mathcal{A} = \operatorname{mod}(G, SW)$ be the category of finitely generated G-equivariant SW-modules. By Lemma 4.3.1(4) $\operatorname{gl}\dim \mathcal{A} = d$. For $\chi \in X(T)^+$ we write $P_\chi = V(\chi) \otimes_F R$. By Lemma 4.3.1(1) $\mathcal{A}$ has a distinguished set of indecomposable projective generators given by P_χ for $\chi \in X(T)^+$, as well as a distinguished set of simple objects $S_\chi = V(\chi) \otimes_F SW/SW_{>0}$ also with $\chi \in X(T)^+$. The projectives and simples are dual in the following sense

$$\mathcal{A}(P_{\chi_1}, S_{\chi_2}) = \delta_{\chi_1,\chi_2} \cdot k \quad \text{for } \chi_1, \chi_2 \in X(T)^+.$$

Note that we have

$$\mathcal{A}(P_{\chi_1}, P_{\chi_2}) = M(\operatorname{Hom}_k(V(\chi_1), V(\chi_2))).$$

Fix a finite subset $\mathcal{L}$ of $X(T)^+$ and put

$$P_{\mathcal{L}} = \bigoplus_{\chi \in \mathcal{L}} P_\chi,$$

$$\Lambda_{\mathcal{L}} = \mathcal{A}(P_{\mathcal{L}}, P_{\mathcal{L}}).$$

We want to find conditions on $\mathcal{L}$ under which one has $\operatorname{gl}\dim \Lambda_{\mathcal{L}} < \infty$. For $\chi \in X(T)^+$ put

$$\tilde{P}_{\mathcal{L},\chi} = \mathcal{A}(P_{\mathcal{L}}, P_\chi).$$

This is a right projective $\Lambda_{\mathcal{L}}$-module if $\chi \in \mathcal{L}$. Similarly we put

$$\tilde{S}_{\mathcal{L},\chi} = \mathcal{A}(P_{\mathcal{L}}, S_\chi).$$

The graded simple right $\Lambda_{\mathcal{L}}$-modules are of the form $\tilde{S}_{\mathcal{L},\chi}$ for $\chi \in \mathcal{L}$. Note that if $\chi \notin \mathcal{L}$ then $\tilde{S}_{\mathcal{L},\chi} = 0$.

LEMMA 4.5.1. *The ring $\Lambda_{\mathcal{L}}$ has finite global dimension if and only for all $\chi \in X(T)^+$ one has*

$$\operatorname{pdim}_{\Lambda_{\mathcal{L}}} \tilde{P}_{\mathcal{L},\chi} < \infty. \tag{4.2}$$

PROOF. The $\Rightarrow$ direction is trivial so let us consider the $\Leftarrow$ direction. So we assume that (4.2) holds and we have to prove that $\operatorname{pdim}_{\Lambda_{\mathcal{L}}} \tilde{S}_{\mathcal{L},\chi} < \infty$ for $\chi \in \mathcal{L}$.

The Koszul complex gives us a resolution of S_χ:

$$0 \to P_d \to \cdots \to P_i \to \cdots \to P_0 \to S_\chi \to 0$$

with

$$P_i = \bigoplus_\mu P_\mu^{m_{\mu,i}}$$

where $m_{\mu,i}$ is the multiplicity of $V(\mu)$ in $V(\chi) \otimes_F \wedge^i W$.

Applying $\mathcal{A}(P_{\mathcal{L}}, -)$ we get an exact sequence in $\operatorname{mod}(\Lambda_{\mathcal{L}}^\circ)$

$$0 \to \tilde{P}_d \to \cdots \to \tilde{P}_i \to \cdots \to\to \tilde{P}_0 \to \tilde{S}_{\mathcal{L},\chi} \to 0$$

with

$$\begin{aligned} \tilde{P}_i &= \mathcal{A}(P_{\mathcal{L}}, P_i) \\ &= \bigoplus_\mu \tilde{P}_{\mathcal{L},\mu}^{m_{\mu,i}}. \end{aligned}$$

Since by (4.2) each of the $\tilde{P}_i$ has finite projective dimension over $\Lambda_{\mathcal{L}}$, the same holds for $\tilde{S}_{\mathcal{L},\chi}$. $\square$

4.5.2. *Creating complexes.* By $Y(T)$ we denote the group of one parameter subgroups of T. We let $Y(T)^-_{\mathbb{R}}$ be the subset of $Y(T)_{\mathbb{R}}$ consisting of all λ such that for all $\rho \in \Phi^+$ we have $\langle \lambda, \rho \rangle \leq 0$. Since $\Phi^- = -\Phi^+$ this implies $\langle \lambda, \rho \rangle \geq 0$ for all $\rho \in \Phi^-$. We also put $Y(T)^- = Y(T)^-_{\mathbb{R}} \cap Y(T)$.

For $0 \neq \lambda \in Y(T)$ we put

$$Z_\lambda = \{x \in X \mid \lim_{t \to 0} \lambda(t)x \text{ exists }\},$$

$$Q_\lambda = \{g \in G \mid \lim_{t \to 0} \lambda(t)g\lambda(t)^{-1} \text{ exists }\}.$$

Then Z_λ is a linear subspace of X cut out in $X \cong W^*$ by the subspace K_λ of W spanned by the weight vectors w_j such that $\langle \lambda, \beta_j \rangle > 0$. Moreover Q_λ is the parabolic subgroup of G containing T and having roots $\rho \in \Phi$ such that $\langle \lambda, \rho \rangle \geq 0$. If $\lambda \in Y(T)^-$ this implies $B \subset Q_\lambda$ and then Z_λ is preserved by B.

The descriptions of Q_λ, Z_λ using roots and weights still make sense if $\lambda \in Y(T)_{\mathbb{R}}$. Note that if $\lambda \in Y(T)^-_{\mathbb{R}}$ then there is always a $\lambda' \in Y(T)^-$ such that $Z_\lambda = Z_{\lambda'}$. So that if $\lambda \in Y(T)^-_{\mathbb{R}}$ it is still true that Q_λ contains B and Z_λ is preserved by Q_λ and hence by B.

For $\lambda \in Y(T)^-_{\mathbb{R}}$ we consider the usual "Springer type" diagram

$$\begin{tikzcd} G \times^B Z_\lambda \arrow[r, hook] \arrow[d] & G \times^B X \arrow[r, "\pi"] \arrow[d] & G/B \arrow[d, "\pi"] \\ GZ_\lambda \arrow[r] & X \arrow[r] & \bullet \end{tikzcd} \tag{4.3}$$

Denote the category equivalence from B-representations to G-equivariant bundles on G/B by $\widetilde{?}$. The inverse is given by taking the fiber in $e := [B] \in G/B$.

Since $G \times^B X \to G/B$ and $G \times^B Z_\lambda \to G/B$ are vector bundles we see that the left most top arrow in (4.3) is obtained by applying $\underline{\mathrm{Spec}}$ to the sheaves of $\mathcal{O}_{G/B}$-algebras

$$SW \otimes_F \mathcal{O}_{G/B} \to S_{G/B}((W/K_\lambda)\tilde{}).$$

We have the corresponding Koszul resolution of sheaves of $\mathcal{O}_{G/B}$-modules:

$$0 \to \wedge^{d_\lambda} K_\lambda\tilde{} \otimes_k SW \to \wedge^{d_\lambda - 1} K_\lambda\tilde{} \otimes_k SW \to \cdots$$
$$\cdots \to \mathcal{O}_{G/B} \otimes_F SW \to S_{G/B}((W/K_\lambda)\tilde{}) \to 0$$

for $d_\lambda = \dim_k K_\lambda$.

Let $\chi \in X(T)^+$. Then we still have an exact sequence

$$0 \to (\chi \otimes_F \wedge^{d_\lambda} K_\lambda)\tilde{} \otimes_k SW \to (\chi \otimes_F \wedge^{d_\lambda - 1} K_\lambda)\tilde{} \otimes_k SW \to \cdots$$
$$\cdots \to \tilde{\chi} \otimes_F SW \to \tilde{\chi} \otimes_F S_{G/B}((W/K_\lambda)\tilde{}) \to 0.$$

Then by Theorem 4.A.3 we get a G-equivariant quasi-isomorphism

$$\begin{aligned} C_{\lambda,\chi} &\overset{\text{def}}{=} \left(\bigoplus_{p \leq 0, q \geq 0} H^q(G/B, (\chi \otimes_F \wedge^{-p} K_\lambda)\tilde{}) \otimes_k SW[-p-q], d \right) \\ &\cong R\Gamma(G/B, \tilde{\chi} \otimes_F S_{G/B}((W/K_\lambda)\tilde{})) \end{aligned} \tag{4.4}$$

where d is obtained from a "horizontal twisted differential" (see Appendix 4.A). We have also used here that the G-equivariant and non-equivariant $R\Gamma$ coincide by [**VV10**, Lemma 1.5.9].

As above we fix a subset $\mathcal{L}$ of $X(T)^+$. We denote by $C_{\mathcal{L},\lambda,\chi}$ the complex of right $\Lambda_{\mathcal{L}}$-modules given by applying $\mathcal{A}(P_{\mathcal{L}}, -)$ to $C_{\lambda,\chi}$.

We say that $\chi \in X(T)^+$ is separated from $\mathcal{L}$ by $\lambda \in Y(T)^-_{\mathbb{R}}$ if

$$\langle \lambda, \chi \rangle < \langle \lambda, \mu \rangle \qquad \text{for every } \mu \in \mathcal{L}.$$

If there exists $w \in \mathcal{W}$ such that $\mu = w * \chi = w(\chi + \bar{\rho}) - \bar{\rho}$ is dominant (see Appendix 4.D) then we write $\chi^+ = \mu$.

LEMMA 4.5.2. *Assume that $\chi \in X(T)^+$ is separated from $\mathcal{L}$ by $\lambda \in Y(T)^-_{\mathbb{R}}$. Then $C_{\mathcal{L},\lambda,\chi}$ is acyclic. Furthermore, forgetting the differential $C_{\mathcal{L},\lambda,\chi}$ is a sum of right $\Lambda_{\mathcal{L}}$-modules of the form $\tilde{P}_{\mathcal{L},\mu}$-modules where the μ are among the weights*

$$(\chi + \beta_{i_1} + \beta_{i_2} + \cdots + \beta_{i_{-p}})^+$$

(with each such expression occurring at most once) where $\{i_1, \ldots, i_{-p}\} \subset \{1, \ldots, d\}$, $i_j \neq i_{j'}$ for $j \neq j'$ and $\langle \lambda, \beta_{i_j} \rangle > 0$.

PROOF. The claim about the $\tilde{P}_{\mathcal{L},\mu}$ that appear is a straightforward application of Bott's theorem after filtering the B-representations $\chi \otimes_F \wedge^{-p} K_\lambda$ by T-representations. So we only have to prove acyclicity. In other words we have to prove that the righthand side of (4.4) is acyclic after applying $\mathcal{A}(P_{\mathcal{L}}, -)$. This amounts to showing that the simple G-representations that appear as summand of

$$H^*(G/B, \tilde{\chi} \otimes_F S((W/K_\lambda)\tilde{\ }))$$

are not of the form $V(\mu)$ for $\mu \in \mathcal{L}$.

The weights of $\chi \otimes_F S(W/K_\lambda)$ are of the form

$$\chi + \beta_{i_1} + \cdots + \beta_{i_q}$$

where $i_j \in \{1, \ldots, d\}$ (repetitions are allowed) and $\langle \lambda, \beta_{i_j} \rangle \leq 0$. By Bott's theorem it follows that $H^*(G/B, \tilde{\chi} \otimes_F S((W/K_\lambda)\tilde{\ }))$ is a direct sum of representations of the form

$$V((\chi + \beta_{i_1} + \cdots + \beta_{i_q})^+).$$

We have

$$\langle \lambda, \chi + \beta_{i_1} + \cdots + \beta_{i_q} \rangle \leq \langle \lambda, \chi \rangle.$$

Thus it suffices to show that if $\mu \in X(T)$ and μ^+ exists then $\langle \lambda, \mu^+ \rangle \leq \langle \lambda, \mu \rangle$. This follows immediately from Corollary 4.D.3 (with $y = -\lambda$, $x = \chi$). $\square$

4.5.3. *Proof of Theorem 4.4.1.* Put $\Gamma = -\bar{\rho} + \{\sum_i a_i \beta_i \mid a_i \leq 0\} + \Delta$. For $\chi \in \Gamma$ let r_χ be minimal with respect to the property $\chi \in -\bar{\rho} + r_\chi \bar{\Sigma} + \Delta$. Note that $\chi \notin -\bar{\rho} + r_\chi \Sigma + \Delta$ for otherwise we could reduce r_χ.

For $\chi \in \Gamma$ let p_χ be the minimal number of a_i equal to $-r_\chi$ among all ways of writing $\chi = -\bar{\rho} + \sum_i a_i \beta_i + \delta$ with $a_i \in [-r_\chi, 0]$, $\delta \in \Delta$. The following properties follow directly from the definitions.

(1) Both r_χ and p_χ depend only on the $\mathcal{W}$-orbit of χ for the $*$-action.
(2) If $\chi \in]\chi', \chi''[$ and $r_\chi = r_{\chi'} = r_{\chi''}$ then $p_\chi \leq \min(p_{\chi'}, p_{\chi''})$.

Assume $\operatorname{gl\,dim} \Lambda_{\mathcal{L}} = \infty$. Then by Lemma 4.5.1 there is some $\chi \in X(T)^+$ such that $\operatorname{pdim}_{\Lambda_{\mathcal{L}}} \tilde{P}_{\mathcal{L},\chi} = \infty$. Then by Lemma 4.5.3 below χ must be in Γ for otherwise $\tilde{P}_{\mathcal{L},\chi} = 0$.

We pick χ such that first r_χ is minimal and then p_χ is minimal. We have $r_\chi \geq 1$ (for otherwise $\chi \in \mathcal{L}$ and hence $\operatorname{pdim} \tilde{P}_{\mathcal{L},\chi} = 0$). We find by Lemma 4.C.2 below (changing the sign of λ) that there exists λ such that for all $\mu \in -\bar{\rho} + r_\chi \Sigma + \Delta$ we have $\langle \lambda, \chi \rangle > \langle \lambda, \mu \rangle$. Hence also $\langle \lambda, \chi + \bar{\rho} \rangle > \langle \lambda, \mu + \bar{\rho} \rangle$. Let $w \in \mathcal{W}$ be such that $w\lambda$ is dominant. Then since $r_\chi \Sigma + \Delta$ is $\mathcal{W}$-invariant we still have for all $\mu \in -\bar{\rho} + r_\chi \Sigma + \Delta$: $\langle w\lambda, w(\chi + \bar{\rho}) \rangle > \langle w\lambda, \mu + \bar{\rho} \rangle$. Moreover by Corollary 4.D.3 below we also have $\langle w\lambda, w(\chi + \bar{\rho}) \rangle \leq \langle w\lambda, \chi + \bar{\rho} \rangle$. Replacing λ then by $-w\lambda$ we find $\lambda \in Y(T)^-_{\mathbb{R}}$ such that $\langle \lambda, \chi + \bar{\rho} \rangle < \langle \lambda, \mu + \bar{\rho} \rangle$, so that finally we have shown

$$\forall \mu \in -\bar{\rho} + r_\chi \Sigma + \Delta : \langle \lambda, \chi \rangle < \langle \lambda, \mu \rangle \tag{4.5}$$

Since $\mathcal{L} \subset -\bar{\rho}+\Sigma+\Delta \subset -\bar{\rho}+r_\chi\Sigma+\Delta$ we obtain in particular χ is separated from $\mathcal{L}$ by λ and hence by Lemma 4.5.2 we have an exact sequence $C_{\mathcal{L},\lambda,\chi}$. Let

$$\mu = (\chi + \beta_{i_1} + \beta_{i_2} + \cdots + \beta_{i_{-p}})^+$$

as in Lemma 4.5.2. If $p = 0$ then $\chi^+ = \chi$ and hence $\tilde{P}_{\mathcal{L},\mu} = \tilde{P}_{\mathcal{L},\chi}$. If $p < 0$ then we claim that either $r_\mu < r_\chi$, or else $p_\mu < p_\chi$. To start put

$$\mu' = \chi + \beta_{i_1} + \beta_{i_2} + \cdots + \beta_{i_{-p}}. \tag{4.6}$$

By Claim (1) above it is sufficient to prove that either $r_{\mu'} < r_\chi$ or else $p_{\mu'} < p_\chi$. This follows easily from the following observation

(3) Write $\chi = -\bar{\rho} + \sum_i a_i\beta_i + \delta$ with $a_i \in [-r_\chi, 0]$, $\delta \in \Delta$. If $\langle \lambda, \beta_i\rangle > 0$ then $a_i = -r_\chi$.

If this claim is false then there is an $\epsilon > 0$ such that $\chi - \epsilon\beta_i \in -\bar{\rho} + r_\chi\Sigma + \Delta$ but $\langle \lambda, \chi - \epsilon\beta_i\rangle = \langle \lambda, \chi\rangle - \epsilon\langle \lambda, \beta_i\rangle < \langle \lambda, \chi\rangle$ which contradicts (4.5).

So we conclude that the indecomposable projective right $\Lambda_{\mathcal{L}}$-modules $\tilde{P}_{\mathcal{L},\mu}$ occurring in $C_{\mathcal{L},\lambda,\chi}$, which are different from the single copy of $\tilde{P}_{\mathcal{L},\chi}$, satisfy either $r_\mu < r_\chi$ or else $p_\mu < p_\chi$. By the minimality assumptions on χ we have $\operatorname{pdim} \tilde{P}_{\mathcal{L},\mu} < \infty$. This implies that $\operatorname{pdim} \tilde{P}_{\mathcal{L},\chi} < \infty$ as well, which is a contradiction.

Lemma 4.5.3. *Assume that $\mathcal{A}(P_{\mathcal{L}}, P_\chi) \neq 0$. Then $\chi \in \Gamma$.*

Proof. If $\mathcal{A}(P_{\mathcal{L}}, P_\chi) \neq 0$ then some $V(\mu)$ for $\mu \in \mathcal{L}$ is a summand of $V(\chi) \otimes S^dW$ for some d. This is equivalent to $V(\chi)^*$ being a summand of some $V(\mu)^* \otimes S^dW$. We have $V(\mu)^* = V(\mu^*)$ with $\mu^* = -w_0\mu$ where w_0 is the longest element of $\mathcal{W}$. So in particular $\mu^* \in -\Gamma - 2\bar{\rho}$.

Since the weights of S^dW are in $\{\sum_i a_i\beta_i \mid a_i \geq 0\}$ we conclude by [**FH91**, Ex. 25.33] that $V(\mu^*) \otimes S^dW$ is a sum of representations of the form $V(\theta)$ with $\theta \in (-\Gamma - 2\bar{\rho}) \cap X(T)^+$. Hence if $V(\chi)^* = V(-w_0\chi)$ appears as a summand of $V(\mu^*) \otimes S^dW$ then $-w_0\chi \in -\Gamma - 2\bar{\rho}$ and hence $\chi \in \Gamma$. □

Remark 4.5.4. The proof Theorem 4.4.1 may be converted into a kind of algorithm to recognize algebras of covariants $\Lambda_{\mathcal{L}}$ that have finite global dimension. It is only a pseudo-algorithm in the sense that *if* it gives a positive answer then definitely $\operatorname{gl\,dim} \Lambda_{\mathcal{L}} < \infty$ but the algorithm is not applicable to every $\Lambda_{\mathcal{L}}$ of finite global dimension. Nonetheless for small examples the algorithm is quite efficient and can even be carried out manually. Furthermore we have implemented it for two-dimensional tori and in that case it has been very useful in our investigations.

The basis of the algorithm is to verify (4.2) for all $\chi \in X(T)^+$. If we have verified (4.2) for a certain finite set of weights $\chi \in \mathcal{L}'$ (initially $\mathcal{L} = \mathcal{L}'$) then we attempt to enlarge this set using Lemma 4.5.2 (as in the proof of Theorem 4.4.1). If this turns out to be impossible then the algorithm returns no answer. Otherwise we attempt to continue enlarging $\mathcal{L}'$ until we arrive at a situation where

$$\mathcal{L}' \supset X(T)^+ \cap (-\bar{\rho} + r\Sigma + \Delta) \supset \mathcal{L} \tag{4.7}$$

for a suitable $r \geq 1$ and a suitable bounded closed convex $\mathcal{W}$-invariant set Δ (usually we may take $\Delta = \emptyset$). In that case we return a positive answer. We may really stop at this stage since now we may proceed as in the proof of Theorem 4.4.1 to enlarge $\mathcal{L}'$ to $X(T)^+$.

There are situations where we can keep enlarging $\mathcal{L}'$ without (4.7) ever becoming true. It is not so clear how to recognize this situation algorithmically. So we simply set a bound on the running time of the algorithm and return no answer if that bound is reached.

4.6. Proofs for noncommutative crepant resolutions.

4.6.1. *Proof of Theorem 4.4.3.* The proof runs parallel with the one of Theorem 4.4.1. We only highlight the differences. Instead of $r_\chi \geq 1$, we now have $r_\chi > 1/2$. Under this condition we have to show that if

$$\mu' = \chi + \sum_{i \in S} \beta_i,$$

$S = \{i_1, \ldots, i_{-p}\} \neq \emptyset$ as in (4.6) then $r_{\mu'} < r_\chi$ or $p_{\mu'} < p_\mu$ where the following additional conditions are satisfied

(1) $\chi = -\bar{\rho} + \sum_i a_i\beta_i + \delta$ with $a_i \in [-r_\chi, 0]$ and $\delta \in \Delta$ with the number of a_i satisfying $b_i = -r_\chi$ being minimal.
(2) All β_i for $i \in S$ are in an open half space $\langle \lambda, - \rangle > 0$.
(3) $a_i = -r_\chi$ for $i \in T_\lambda = \{i \mid \langle \lambda, \beta_i\rangle > 0\} \supset S$ (by "observation (3)" in the proof of Theorem of 4.4.1).

In addition we may and we will assume

(4) For every line $\ell \subset X(T)_\mathbb{R}$ through the origin the $\beta_i \in \ell$ with $a_i \neq 0$ are all *on the same side of the origin.*

This uses the fact $\{\beta_i \in \ell\}$ contains only the zero weight, or else contains weights β_i on both sides of the origin (since otherwise $\sum_{\beta_i \in \ell} \beta_i \neq 0$).

We have

$$\mu' = -\bar{\rho} + \sum_i a_i'\beta_i + \delta,$$

$$a_i' = \begin{cases} a_i & i \notin S \\ a_i + 1 & i \in S \end{cases}$$

We now write

$$\mu' = -\bar{\rho} + \sum_\ell \sum_{\beta_i \in \ell} a_i'\beta_i + \delta \tag{4.8}$$

where the sum is over the lines $\ell \subset X(T)_\mathbb{R}$ through the origin. Fix such a line. If $\ell \cap \{\beta_i \mid i \in S\} = \emptyset$ then $a_i' = a_i$ for all i such that $\beta_i \in \ell$ and hence $a_i' \in [-r_\chi, 0]$.

We assume now that $\ell \cap \{\beta_i \mid i \in S\} \neq \emptyset$ (note that it is clear that there are ℓ for which this holds). In particular $\langle \lambda, - \rangle$ is non-zero on ℓ. Let γ_u be a unit vector on ℓ such that $\langle \lambda, \gamma_u \rangle > 0$.

Put $T_\lambda^\ell = \{i \in T_\lambda \mid \beta_i \in \ell\}$, $S^\ell = S \cap T_\lambda^\ell$. For $\beta_i \in \ell$ put $\beta_i = c_i\gamma_u$. Then $c_i > 0$ for all $i \in T_\lambda^\ell$. If $i \in S^\ell$ then by (3) $a_i = -r_\chi \neq 0$. By (4) we deduce from this that if $\beta_i \in \ell - \{0\}$ is such that $a_i \neq 0$ then $i \in T_\lambda^\ell$. We then compute

$$\begin{aligned}\sum_{\beta_i \in \ell} a_i'\beta_i &= \sum_{i \in S^\ell} (1 - r_\chi)\beta_i + \sum_{i \in T_\lambda^\ell \setminus S^\ell} (-r_\chi)\beta_i + \sum_{\beta_i \in \ell, i \notin T_\lambda^\ell} b_i\beta_i \\ &= \sum_{i \in S^\ell} (1 - r_\chi)\beta_i + \sum_{i \in T_\lambda^\ell \setminus S^\ell} (-r_\chi)\beta_i \\ &= \pi\gamma_u\end{aligned}$$

with

$$\pi = \sum_{i \in S^\ell} (1 - r_\chi)c_i + \sum_{i \in T_\lambda^\ell \setminus S^\ell} (-r_\chi)c_i$$

where there is at least one term of the form $(1 - r_\chi)c_i$. Set $c = \sum_{i \in T_\lambda^\ell} c_i$. Since $c_i > 0$ for $i \in T_\lambda^\ell$ we have

$$\pi > \sum_{i \in S^\ell} (-r_\chi)c_i + \sum_{i \in T_\lambda^\ell \setminus S^\ell} (-r_\chi)c_i = -r_\chi c.$$

On the other hand as $r_\chi > 1/2$ we have $1 - r_\chi < r_\chi$, $-r_\chi < r_\chi$ and hence

$$\pi < \sum_{i\in S^\ell} r_\chi c_i + \sum_{i\in T^\ell_\lambda \setminus S^\ell} r_\chi c_i = r_\chi c.$$

Assume $\pi < 0$ and put $a' = \pi/c$. Then we have $a' \in]-r_\chi, 0]$ and

$$\sum_{\beta_i\in\ell} a'_i\beta_i = \pi\gamma_u = ca'\gamma_u = \sum_{i\in T^\ell_\lambda} a'\beta_i. \tag{4.9}$$

Similarly assume $\pi \geq 0$ and put $a' = -\pi/c$. Then again we have $a' \in]-r_\chi, 0]$ and

$$\sum_{\beta_i\in\ell} a'_i\beta_i = \pi\gamma_u = -ca'\gamma_u = -\sum_{i\in T^\ell_\lambda} a'\beta_i = \sum_{\beta_i\in\ell, i\notin T^\ell_\lambda} a'\beta_i. \tag{4.10}$$

Note that in the last equality we finally used the full force of the hypothesis $\sum_{\beta_i\in\ell}\beta_i = 0$. Plugging the righthand sides of (4.9)(4.10) into (4.8) we conclude that either $r_{\chi'} < r_\chi$ or else $p_{\chi'} < p_\chi$, contradicting the minimality of χ.

4.6.2. *Proof of Theorem 4.4.4.* We first prove that $\operatorname{gl\,dim}\Lambda < \infty$. For small strictly positive r we have

$$\begin{aligned}
X(T)^+_{\bar\mu} \cap (-\bar\rho + (1/2)\bar\Sigma_\varepsilon) &= X(T)^+_{\bar\mu} \cap (-\bar\rho + (1/2)(\bar\Sigma \cap (r\varepsilon + \bar\Sigma))) \\
&= X(T)^+_{\bar\mu} \cap (-\bar\rho + (1/2)(r\varepsilon + \bar\Sigma)) \\
&= X(T)^+_{\bar\mu} \cap (-\bar\rho + r\varepsilon/2 + (1/2)\bar\Sigma))
\end{aligned}$$

since for such small r one has that $X(T)^+_{\bar\mu} \cap (-\bar\rho + (1/2)((r\varepsilon + \bar\Sigma) \setminus \bar\Sigma))$ is empty. We now apply Theorem 4.4.3 with $\Delta = \{r\epsilon/2\}$.

4.6.3. *Proof of Theorem 4.4.5.* Using Theorem 4.4.4 it is sufficient to prove that Λ is Cohen-Macaulay. I.e. if $\chi_1, \chi_2 \in X(T)^+_{\bar\mu} \cap (-\bar\rho + (1/2)\bar\Sigma_\varepsilon)$ then $M(V(\chi_1)^* \otimes V(\chi_2)) = M(V(\chi_1)^* \otimes V(-w_0\chi_2)^*)$ is Cohen-Macaulay. By Proposition 4.3.11 it is sufficient to prove that $\chi_1 - w_0\chi_2$ is strongly critical.

We have $-w_0\chi_2 \in -\bar\rho + (1/2)(-\bar\Sigma)$. Since $\sum_i \beta_i = 0$ it easy to see that $-\bar\Sigma = \Sigma$. Thus $\chi_1 - w_0\chi_2 \in -2\bar\rho + (1/2)(\bar\Sigma + \bar\Sigma)) = -2\bar\rho + \bar\Sigma$.

If $\chi = \sum_i a_i\beta_i$ with $a_i \in]-1, 0]$ then by subtracting a small multiple of the identity $\sum_i \beta_i = 0$ we may assume that $a_i \in]-1, 1[$. Hence Σ is relatively open and thus it is equal to $\bar\Sigma - \partial\bar\Sigma$. Assume that $\chi_1 - w_0\chi_2 \notin -2\bar\rho + \Sigma$. This is only possible if $\bar\rho + \chi_1$ and $\bar\rho - w_0\chi_2$ are elements of the same boundary face F of $(1/2)\bar\Sigma$. Then F must be a boundary face of $(1/2)\Sigma_\varepsilon$ and a boundary face of $(1/2)(-w_0)(\Sigma_\varepsilon) = (1/2)\Sigma_{-\varepsilon}$. So F is in fact a boundary face of $(1/2)\Sigma_{\pm\varepsilon}$. Now $\chi_1 \in X(T)_{\bar\mu} \cap (-\bar\rho + F)$ which is empty by the hypothesis (4.1). This is a contradiction.

4.7. Trace rings. In this section we prove that centers of trace rings admit twisted NCCRs.

THEOREM 4.7.1. *Assume $m \geq 2$, $n \geq 2$. Then $\mathcal{T}_{m,n}$ has a twisted NCCR.*

To illustrate the proof goes we will first handle the case $(m, n) = (2, 3)$ graphically. From this discussion we will obtain a new proof for the fact that $\mathcal{T}_{2,3}\langle\xi_k\rangle$ has finite global dimension (see [**LBVdB88**]).

Put $G = \mathrm{PGL}_3$, $\bar G = \mathrm{SL}_3$, $V = k^3$. Put $W = \operatorname{End}(V)^{\oplus 2}$. It is easy to see that W is generic (cfr. Remark 4.3.5). Let $\bar T$ be the diagonal maximal torus in $\bar G$ and let T be its image in PGL_3. We have $A = \ker(\bar T \to T) \cong \mathbb{Z}/3\mathbb{Z}$. In Figure 4.1 we have drawn the elements of $X(\bar T)$, color coded according the character of A they correspond to. The red weights are those in $X(T)$. We have also indicated the simple roots $\alpha_{1,2}$ and the fundamental weights $\phi_{1,2}$ for $\bar G$.

The dominant cone has been colored brown. Finally we have also drawn part of the open regular hexagon $-\bar{\rho}+(1/2)\Sigma$. From this picture we see that $-\bar{\rho}+(1/2)(\bar{\Sigma}-\Sigma)$ contains no dark green or light green weights. Hence either of these colors yields a twisted NCCR by Theorem 4.4.5. The picture shows that $-\bar{\rho}+(1/2)\bar{\Sigma}$ contains a single dominant dark green and a single dominant light green weight. These are precisely the fundamental weights so they correspond to V and $\wedge^2 V = V^*$ respectively. The corresponding twisted NCCRs are $M(\operatorname{End}(V)) \cong \mathcal{T}_{2,3}\langle\xi_k\rangle$ and $M(\operatorname{End}(V^*)) \cong \mathcal{T}^\circ_{2,3}\langle\xi_k\rangle$.

Proof of Theorem 4.7.1. $\mathcal{T}_{2,2}$ is its own NCCR because $\operatorname{gl}\dim \mathcal{T}_{2,2} = 5$. Hence without loss of generality we may assume $(m,n) \neq (2,2)$ which is the same as saying that $W = \operatorname{End}(V)^{\oplus m}$ is generic.

Inspecting Figure 4.1 for the case $(m,n) = (2,3)$ we see that in fact the dark green and light green weights are missed by the entire boundary lines of $-\bar{\rho}+(1/2)\bar{\Sigma}$ and not just by the faces. We will show that this pattern persists for (m,n) arbitrary: the boundary hyperplanes of $-\bar{\rho}+(1/2)\bar{\Sigma}$ will always miss certain "colors" and these colors then yield a twisted NCCR by Theorem 4.4.5.

Put $G = \mathrm{PGL}_n$, $\bar{G} = \mathrm{SL}_n$. We let $\bar{T}$ be the standard maximal torus in $\bar{G}$ given by $\{\operatorname{diag}(z_1,\dots,z_n) \mid z_1\cdots z_n = 1\}$ and T its image in G. We let $L_i \in X(\bar{T})$ be the projection $\bar{T} \mapsto G_m : (z_1,\dots,z_n) \mapsto z_i$. Then $\sum_i L_i = 0$ and in fact $X(\bar{T}) = \bigoplus_i \mathbb{Z}L_i/(\sum_i L_i)$.

The group $A = \ker(\bar{T} \to T) = Z(\bar{G})$ consists of the scalar diagonal matrices $\operatorname{diag}(\xi,\dots,\xi)$ where $\xi^n = 1$. Hence canonically $X(A) \cong \mathbb{Z}/n\mathbb{Z}$ where $\bar{1}$ corresponds to the inclusion of the n-th roots of unity in G_m. Keeping the above terminology we will refer to the image of a weight χ in $X(A)$ as its "color" and denote it by $c(\chi)$. Of course color is preserved by the Weyl group action. One finds that the color of L_i is equal to $\bar{1}$.

We have

$$\begin{aligned}\bar{\rho} &= (n-1)/2L_1 + (n-3)/2L_2 + \cdots + (-n+1)/2L_n \\ &= n/2L_1 + (n-2)/2L_2 + \cdots + (-n+2)/2L_n\end{aligned}$$

We conclude

$$c(\bar{\rho}) = \begin{cases} \bar{0} & \text{if } n \text{ is odd} \\ \overline{n/2} & \text{if } n \text{ is even} \end{cases}$$

The non-zero weights of W are given by the root system $\Phi = A_{n-1}$. I.e. they are of the form $(L_i - L_j)_{i\neq j}$ where each such weight occurs with multiplicity m. Using the description of $\partial\bar{\Sigma}$ in Lemma 4.B.1 below one sees that the boundary hyperplanes of $\bar{\Sigma}$ correspond to maximal subroot systems of Φ and these are given by $A_{p-1}\times A_{q-1}$ for a partition $n = p+q$ (with $A_0 = 0$). More precisely, one verifies that up to the action of the Weyl group the boundary hyperplanes of $(1/2)\bar{\Sigma}$ are given by

$$H_{p,q} = m\rho_{p,q} + \sum_{1\le i<j\le p} \mathbb{R}(L_i - L_j) + \sum_{p+1\le i<j\le n} \mathbb{R}(L_i - L_j)$$

where

$$\begin{aligned}\rho_{p,q} &= \frac{1}{2}\sum_{i=1,\dots,p,\,j=p+1,\dots,n} (L_i - L_j) \\ &= \frac{q}{2}\sum_{i=1,\dots,p} L_i - \frac{p}{2}\sum_{i=p+1,\dots,n} L_i\end{aligned}$$

from which we compute in the same way as for $\bar{\rho}$:

$$c(m\rho_{p,q}) = \begin{cases} \bar{0} & \text{if } m \text{ even} \\ \bar{0} & \text{if } m \text{ odd, } p,q \text{ even} \\ \overline{n/2} & \text{if } m \text{ odd, } p,q \text{ odd} \\ \text{undefined} & \text{if } m \text{ odd, } n \text{ odd} \end{cases}$$

The last possibility is because if both m,n are odd then $m\rho_{p,q} \notin X(\bar{T})$.

We have

$$H_{p,q} = \{\sum_i a_i L_i \mid \sum_{i=1}^{p} a_i = mq/2, \sum_{i=p+1}^{n} a_i = -mp/2\} \tag{4.11}$$

Assume

$$\chi = \sum_i n_i L_i \in X(\bar{T}) \cap H_{p,q}$$

for $n_i \in \mathbb{Z}$. Then there exists $\epsilon \in \mathbb{R}$ such that for $a_i = n_i - \epsilon$ the conditions on $(a_i)_i$ on the righthand side of (4.11) are true. Hence in particular

$$\begin{aligned} mq/2 + p\epsilon &\in \mathbb{Z} \\ -mp/2 + q\epsilon &\in \mathbb{Z} \end{aligned}$$

which implies $m(p^2+q^2)/2 \in \mathbb{Z}$ which is impossible if m,n are both odd. Hence it this case we are done. So below we assume that m,n are not both odd. So in particular $m\rho_{p,q} \in X(\bar{T})$ and hence

$$X(\bar{T}) \cap H_{p,q} = m\rho_{p,q} + X(\bar{T}) \cap \{\sum_i a_i L_i \mid \sum_{i=1}^{p} a_i = 0, \sum_{i=p+1}^{n} a_i = 0\}.$$

Now assume

$$\chi' = \sum_i n'_i L_i \in X(\bar{T}) \cap \{\sum_i a_i L_i \mid \sum_{i=1}^{p} a_i = 0, \sum_{i=p+1}^{n} a_i = 0\}.$$

Then putting $a_i = n'_i - \epsilon$ as above we find $p\epsilon \in \mathbb{Z}$, $q\epsilon \in \mathbb{Z}$. It is easy to see that this implies ϵ may be written as $\epsilon = \frac{k}{d}$ for $d = \gcd(p,q)$ and $k \in \mathbb{Z}$. Furthermore we find

$$c(\chi') = c((p+q)\epsilon) = \overline{\frac{nk}{d}}$$

so that ultimately

$$c(\chi) = \overline{\frac{nk}{d}} + \begin{cases} \bar{0} & m \text{ even} \\ \bar{0} & m \text{ odd, } n \text{ even, } 2 \mid d \\ \overline{n/2} & m \text{ odd, } n \text{ even, } 2 \nmid d \end{cases}$$

with $d \mid n$, $d \neq n$. One now verifies the following statements.

1. If m is even then $c(\chi) \neq \bar{1}$.
2. If m is odd, n is even but $n/2$ is odd then $c(\chi) \neq \bar{2}$.
3. If m is odd, n is even but $n/2$ is even then $c(\chi) \neq \bar{1}$.

So as mentioned $c(-\bar{\rho}) + c(\chi)$ always misses certain colors, finishing the proof. $\square$

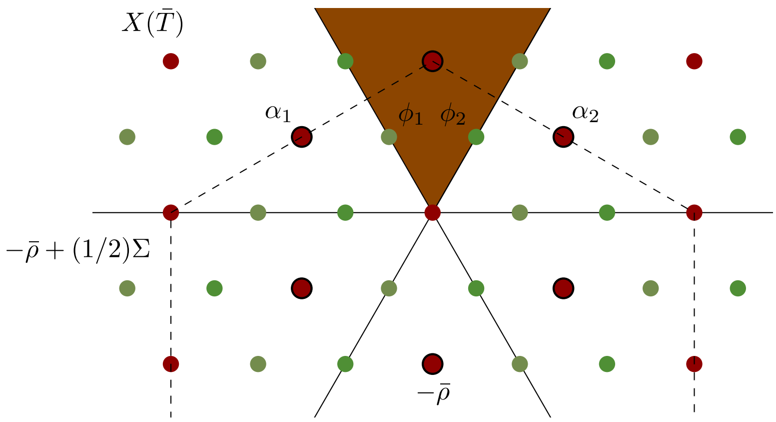

FIGURE 4.1. Relevant data for $\mathcal{T}_{2,3}$.

4.A. A refinement of the E_1-hypercohomology spectral sequence. Below $\mathcal{C}$ is an abelian category.

LEMMA 4.A.1. *Assume that $I^\bullet$ is a complex over $\mathcal{C}$ with projective homology. Then there is a quasi-isomorphism*

$$j : \left(\bigoplus_n H^n(I^\bullet)[-n], 0\right) \to I^\bullet.$$

PROOF. For each i choose a splitting β_i for the projection $Z^i(I^\bullet) \to H^i(I^\bullet)$ and let j_i be the composition

$$H^i(I^\bullet) \xrightarrow{\beta_i} Z^i(I^\bullet) \to I^i.$$

It now suffices to take $j = \oplus_i j_i$. $\square$

We now discuss a two-dimensional variant of this result. If $A^{\bullet\bullet}$ is a bigraded object in $\mathcal{C}$ then $\mathrm{Tot}_\oplus(A^{\bullet\bullet})$ is the graded object in $\mathcal{C}$ given by

$$\mathrm{Tot}_\oplus(A^{\bullet\bullet})^m \overset{\text{def}}{=} \bigoplus_{p+q=m} A^{pq}$$

provide this coproduct is finite. We will usually write $\mathrm{Tot}_\oplus(A) = \bigoplus_{pq} A^{pq}[-p-q]$.

A twisted differential on A^{pq} is a collection of maps $d_n^{pq} : A^{pq} \to A^{p+n,q-n+1}$ such that $d = \sum_{pqn} d_n^{pq}$ induces a differential on $\mathrm{Tot}_\oplus(A^{\bullet\bullet})$. In other word we require for all p, q, p', q'

$$\sum_{n+n'=p'-p} d_{n'}^{p+n,q-n+1} d_n^{pq} = 0. \tag{4.12}$$

Below we will make the adhoc definition that a twisted differential is horizontal (htd) if $d_n^{pq} = 0$ for $n \le 0$. Note that in that case $d_1^2 = 0$.

LEMMA 4.A.2. *Let $I^{\bullet\bullet}$ be a double complex over $\mathcal{C}$. Assume*

1. *There are some p_0, p_1 such that $I^{pq} = 0$ for all q and all $p \notin [p_0, p_1]$.*
2. *For each p, $I^{p,\bullet}$ has bounded cohomology.*
3. *For each p the cohomology of $I^{p,\bullet}$ is projective.*

Then there exists a quasi-isomorphism

$$\left(\bigoplus_{pq} H^q(I^{p,\bullet})[-p-q], d\right) \to \mathrm{Tot}_\oplus I^{\bullet\bullet}$$

where the differential on the left is obtained from a thd on $(H^q(I^{p,\bullet}))_{pq}$ with $d_1^{pq}: H^q(I^{p,\bullet}) \to H^q(I^{p+1,\bullet})$ being obtained from the differential $I^{pq} \to I^{p+1,q}$ in $I^{\bullet\bullet}$.

Proof. We have an exact sequence of bicomplexes

$$0 \to I^{\geq p_0+1,\bullet} \to I^{\geq p_0,\bullet} \to I^{p_0,\bullet} \to 0,$$

which is split if we ignore the horizontal differential. This yields a quasi-isomorphism

$$\operatorname{Tot}_\oplus I^{\geq p_0 \bullet} \cong \operatorname{cone}(\operatorname{Tot}_\oplus I^{p_0,\bullet}[-1] \xrightarrow{\theta} \operatorname{Tot}_\oplus I^{\geq p_0+1,\bullet}). \tag{4.13}$$

Using Lemma 4.A.1 we have in $D(\mathcal{C})$

$$\operatorname{Tot}_\oplus I^{p_0,\bullet} \cong \left(\bigoplus_q H^q(I^{p_0,\bullet})[-p_0-q], 0\right) \tag{4.14}$$

and by induction we also have

$$\operatorname{Tot}_\oplus I^{\geq p_0+1,\bullet} \cong \left(\bigoplus_{p\geq p_0+1,q} H^q(I^{p,\bullet})[-p-q], d'\right) \tag{4.15}$$

where d' is obtained from a thd. Substituting (4.14),(4.15) in (4.13) and noting that now θ becomes a map between bounded projective complexes, we find that θ is represented by an actual map of complexes in $\mathcal{C}$. The lemma now follows using the standard construction of the cone. □

The following is a refined version of [**BLVdB10**, Proposition 4.4].

Theorem 4.A.3. Let $F : \mathcal{A} \to \mathcal{B}$ be a left exact functor between abelian categories and assume in addition that $\mathcal{A}$ has enough injectives. Let $A^\bullet$ be a (literally) bounded complex in $\mathcal{A}$ and assume in addition that for each p, $R^qF(A^p) \in \operatorname{Ob}(\mathcal{B})$ is projective, and zero for $q \gg 0$. Then there is an isomorphism in $D(\mathcal{B})$

$$\left(\bigoplus_{p,q} R^qF(A^p)[-p-q], d\right) \cong RF(A^\bullet) \tag{4.16}$$

where d on the left is obtained from a thd such that $d_1^{pq}: RF^q(A^p) \to RF^q(A^{p+1})$ is equal to $RF^q(d^p_{A^\bullet})$.

Proof. Let $A^\bullet \to I^{\bullet\bullet}$ be an injective Cartan-Eilenberg resolution of $A^\bullet$ (with $I^{p,\bullet}$ resolving A^p). Then RFA is computed by $F(\operatorname{Tot}_\oplus(I^{\bullet\bullet}))$. It now suffices to use Lemma 4.A.2 (with $F(I^{\bullet\bullet})$ playing the role of $I^{\bullet\bullet}$). □

Remark 4.7.2. Note that (4.16) is a refined version of the standard hypercohomology spectral sequences

$$E_1^{pq} = RF^q(A^p) \Rightarrow R^{p+q}F(A^\bullet).$$

4.B. Faces of some polygons. Let E be a Euclidean space, $(\beta_i)_{i=1,\dots,d} \subset E$ a collection of points and let $a_i < b_i$, $i = 1, \dots, d$ be a collection of real numbers. Consider the closed polygon

$$\nabla = \left\{\sum_i g_i\beta_i \mid g_i \in [a_i, b_i]\right\}.$$

For $\lambda \in E^*$ write ∇_λ for the set of $\beta = \sum_i g_i\beta_i \in \nabla$ that satisfy

$$\langle \lambda, \beta_i\rangle > 0 \Rightarrow g_i = a_i, \tag{4.17}$$

$$\langle \lambda, \beta_i\rangle < 0 \Rightarrow g_i = b_i. \tag{4.18}$$

LEMMA 4.B.1. ∇_λ is a face of ∇ whose linear span is

$$H_\lambda = C_\lambda + \sum_{\langle\lambda,\beta_i\rangle=0} \mathbb{R}\beta_i$$

where

$$C_\lambda = \sum_{\langle\lambda,\beta_i\rangle>0} a_i\beta_i + \sum_{\langle\lambda,\beta_i\rangle<0} b_i\beta_i. \tag{4.19}$$

PROOF. If $\lambda = 0$ then $\nabla_\lambda = \nabla$. Assume $\lambda \neq 0$. Put $c_\lambda = \langle\lambda, C_\lambda\rangle$. It is an easy verification that ∇ is contained in the half space $H_\lambda = \{c_\lambda \leq \langle\lambda, -\rangle\}$ and moreover ∇_λ is the intersection ∇ with $\partial H_\lambda = \{c_\lambda = \langle\lambda, -\rangle\}$. This proves that ∇_λ is a face. The claim about the linear span is clear. $\square$

LEMMA 4.B.2. If F is a face in ∇ then $F = \nabla_\lambda$ for suitable λ.

PROOF. If $F = \nabla$ then $F = \nabla_0$. If $F \neq \nabla$ then there exists $\lambda \in E^*$, $c \in \mathbb{R}$ such that $\nabla \subset H_\lambda = \{\langle\lambda, -\rangle \geq c\}$ and $F = \nabla \cap \partial H_\lambda$. Let $c_\lambda = \langle\lambda, C_\lambda\rangle$ be as above. Then the minimum of $\langle\lambda, -\rangle$ attained on ∇ is both c and c_λ and these minima are achieved on F and ∇_λ respectively. Hence $c_\lambda = c$ and $F = \nabla_\lambda$. $\square$

COROLLARY 4.B.3. ∇ is the convex hull of C_λ where λ runs through those elements of E^* such that $\langle\lambda, \beta_i\rangle \neq 0$ for all i.

PROOF. By the above discussion the set of C_λ we have described is precisely the set of vertices of ∇. $\square$

4.C. Supporting hyperplanes of Minkowski sums. This section is related to Appendix 4.B. Presumably the following result is standard.

LEMMA 4.C.1. Let $(\Pi_i)_{i=1,\ldots,n}$ be closed convex sets in a finite dimensional vector space E and let $\Pi = \{\sum_i x_i \mid x_i \in \Pi_i\}$ be their Minkowski sum. Let $x \in \Pi$. Then there exists $\lambda \in E^*$ such that Π is contained in the half space $\langle\lambda, -\rangle \geq \langle\lambda, x\rangle$ and such that x can be written as $\sum_i x_i$ with $x_i \in \Pi_i$ in such a way that $\langle\lambda, -\rangle$ is constant on Π_i if and only if $x_i \notin \partial\Pi_i$.

PROOF. For $z \in \Pi$ let $p(z)$ be the minimal number of $z_i \in \partial\Pi_i$ among all ways of writing $z = \sum_i z_i$ with $z_i \in \Pi_i$. Note that for $z_1, z_2 \in \Pi$ and $z \in]z_1, z_2[$ we have

$$p(z) \leq \min(p(z_1), p(z_2)). \tag{4.20}$$

With x as in the statement of the lemma, write $x = \sum_i x_i$ such that the number of x_i in $\partial\Pi_i$ is mininal. Let Γ be the polyhedral cone spanned by all $y - x_i$ with $i \in \{1, \ldots, n\}$ and $y \in \Pi_i$. Note that if $\sigma \in \Gamma$ then for $\epsilon > 0$ small enough we have $x + \epsilon\sigma \in \Pi$ and moreover $p(x + \epsilon\sigma) \leq p(x)$.

Let L be the maximal linear subspace in Γ and let $\lambda \in E^*$ be such that $\langle\lambda, -\rangle > 0$ on $\Gamma - L$ and $\langle\lambda, -\rangle = 0$ on L.

If $x_i \notin \partial\Pi_i$ then $y - x_i$ with $y \in \Pi_i$ spans a linear subspace of Γ and hence $\langle\lambda, y - x_i\rangle = 0$ for $y \in \Pi_i$.

Assume $x_i \in \partial\Pi_i$. It is sufficient to prove that $y - x_i \in \Gamma - L$ for $y \in \operatorname{relint}\Pi_i$. If $y - x_i \in L$ then $-(y - x_i) \in \Gamma$ and hence as noted above, for $\epsilon > 0$ small enough we have $x' = x - \epsilon(y - x_i) \in \Pi$ as well as $p(x') \leq p(x)$. On the other hand we have for $x'' = x + (y - x_i) \in \Pi$: $p(x'') < p(x)$. Since $x \in]x_1, x_2[$ it now suffices to invoke (4.20) to obtain the contradiction $p(x) < p(x)$. $\square$

Now we use the notations of Appendix 4.B. We define in addition

$$\Sigma = \left\{ \sum_i g_i\beta_i \mid g_i \in]a_i, b_i] \right\}$$

such that $\nabla = \bar{\Sigma}$.

LEMMA 4.C.2. Let Δ be a closed convex subset of $X(T)_{\mathbb{R}}$. If $x \in (\bar{\Sigma}+\Delta)-(\Sigma+\Delta)$ then there exists $\lambda \in Y(T)_{\mathbb{R}}$ such that for all $z \in \Sigma+\Delta$ we have $\langle \lambda, z\rangle > \langle \lambda, x\rangle$.

PROOF. We apply the previous lemma (changing the indexing, and using opposite inequalities) with $\Pi_i = [a_i, b_i]\beta_i$ and $\Pi_0 = \Delta$. There exists $\lambda \in Y(T)_{\mathbb{R}}$ such that we may write x as

$$x = \sum_i g_i \beta_i + \delta$$

such that for all $\delta' \in \Delta$ we have $\langle \lambda, \delta\rangle \leq \langle \lambda, \delta'\rangle$ and moreover

$$\begin{aligned} \langle \lambda, \beta_i\rangle > 0 &\iff g_i = a_i, \\ \langle \lambda, \beta_i\rangle < 0 &\iff g_i = b_i, \\ \langle \lambda, \beta_i\rangle = 0 &\iff g_i \neq a_i, b_i. \end{aligned}$$

Since $x \notin \Sigma + \Delta$ there is at least one g_i equal to a_i. From this one easily deduces the claim in the statement of the lemma. $\square$

4.D. Some elementary facts about root systems. Let E be a finite dimensional real vector space equipped with a positive definite quadratic form $(-,-)$. Let Φ be a rootsystem in E spanning some subspace $E' \subset E$. Let $\Phi^+ \subset \Phi$ be a set of positive roots and let $S \subset \Phi^+$ be the corresponding simple roots. For ρ a root let $\check{\rho}$ be the corresponding coroot, given by $\check{\rho} = 2\rho/(\rho,\rho)$.

We say that $x \in E$ is dominant if $(\rho, x) \geq 0$ for all $\rho \in \Phi^+$. The reflection associated to a simple root α is

$$s_\alpha(x) = x - (\check{\alpha}, x)\alpha$$

By definition the reflections generate the Weyl group $\mathcal{W}$ of Φ. The Weyl group $\mathcal{W}$ preserves $E, (-,-), \Phi$.

LEMMA 4.D.1. Let Δ be a $\mathcal{W}$-invariant convex subset of E. Let $x, y \in E$ be dominant. If $x + y \in \Delta$ then for all $v, w \in \mathcal{W}$ we have $vx + wy \in \Delta$.

PROOF. Without loss of generality we may assume $v = 1$. Since y is dominant there exist simple roots $\alpha_1, \ldots, \alpha_n$ such that $w = s_{\alpha_n} \cdots s_{\alpha_1}$ and such that if we put $y_i = s_{\alpha_i} \cdots s_{\alpha_1} y$ then $(\alpha_{i+1}, y_i) > 0$. By induction on i we may assume $x + y_i \in \Delta$. Then we have $s_{\alpha_{i+1}}(x + y_i) \in \Delta$ and hence

$$\begin{aligned} x + y_i - 0 \cdot \alpha_{i+1} &\in \Delta, \\ x + y_i - (\check{\alpha}_{i+1}, x + y_i)\alpha_{i+1} &\in \Delta. \end{aligned}$$

Now note

$$x + y_{i+1} = x + y_i - (\check{\alpha}_{i+1}, y_i)\alpha_{i+1}$$

and hence since x is dominant and therefore $(\check{\alpha}_{i+1}, x) \geq 0$, we have

$$0 \leq (\check{\alpha}_{i+1}, y_i) \leq (\check{\alpha}_{i+1}, x + y_i).$$

Thus $x + y_{i+1}$ is in the convex hull of $x + y_i$ and $s_{\alpha_{i+1}}(x + y_i)$ and therefore $x + y_{i+1} \in \Delta$. $\square$

Let $\bar{\rho} = (1/2)\sum_{\rho \in \Phi^+} \rho$. Then it is well known that for each simple root we have $s_\alpha(\bar{\rho}) = \bar{\rho} - \alpha$. Put

$$E_{\mathbb{Z}} = \{x \in E \mid \forall \rho \in \Phi : (\hat{\rho}, x) \in \mathbb{Z}\}.$$

For $w \in \mathcal{W}$, $x \in E$ put $w{*}x = w(x + \bar{\rho}) - \bar{\rho}$. This defines an (affine) action of $\mathcal{W}$ on E, which preserves $E_{\mathbb{Z}}$.

LEMMA 4.D.2. Assume that $y \in E$ is dominant. If $x \in E$ and $\alpha \in S$ is such that $(\alpha, x) < 0$ then

$$(s_\alpha x, y) \geq (x, y). \tag{4.21}$$

Moreover if $x \in E_{\mathbb{Z}}$ then also

$$(s_\alpha * x, y) \geq (x, y). \tag{4.22}$$

PROOF. We compute

$$(s_\alpha x, y) = (x - (\check{\alpha}, x)\alpha, y) = (x, y) - (\check{\alpha}, x)(\alpha, y) \geq (x, y)$$

using the fact that y is dominant and hence $(\alpha, y) \geq 0$. Similarly

$$(s_\alpha * x, y) = (x - ((\check{\alpha}, x) + 1)\alpha, y) = (x, y) - ((\check{\alpha}, x) + 1)(\alpha, y) \geq (x, y)$$

using now in addition that $(\check{\alpha}, x) \in \mathbb{Z}$. $\square$

COROLLARY 4.D.3. Assume that $y \in E$ is dominant and $x \in E$ is arbitrary. If $w \in \mathcal{W}$ is such that wx is dominant then

$$(wx, y) \geq (x, y).$$

If $x \in E_{\mathbb{Z}}$ and $w * x$ is dominant then

$$(w * x, y) \geq (x, y).$$

PROOF. For the first inequality note that if wx is dominant then it can be written as $s_{\alpha_n} \cdots s_{\alpha_1} x$ such that for each $x_i = s_{\alpha_i} \cdots s_{\alpha_1} x$ the inequality $(\alpha_{i+1}, x_i) < 0$ is satisfied. It now suffices to invoke (4.21). The argument for $w * x$ is similar, now using (4.22). $\square$

CHAPTER 4

Free function theory through matrix invariants

In this chapter we present an application of the matrix invariant theory, in particular the theory of trace identities, in free analysis. In Section 1 we depict the language of free analysis. We also give involution analogues of certain notions introduced in Section 2.1, as the algebraic approach through matrix invariants enables us to study free maps with involution. These maps are free noncommutative analogues of real analytic functions of several variables.

In Section 2 and in Section 3 we present a characterization of polynomial free maps via properties of their finite-dimensional slices. This is used to establish power series expansions for analytic free maps about scalar and non-scalar points; the latter are given by series of generalized polynomials for which we obtain an invariant-theoretic characterization.

The existence of power series expansions makes it possible to develop some calculus for free maps. We give an inverse and implicit function theorem for free maps with involution in Section 4, where we also briefly discuss some possible topologies one may take on the base space.

In Section 5 we demonstrate with examples that free maps with involution do not exhibit strong rigidity properties of their involution-free free counterparts. While we focus throughout this chapter on involution given by transposition we indicate in Section 6 modifications needed for handling the conjugate transposition.

This chapter is based on [**KŠ14a**].

1. G-free sets and G-free maps

In this section we present preliminaries from free analysis needed in the sequel.

1.1. Notation. Let $F \in \{\mathbb{R}, \mathbb{C}\}$ and let $\mathcal{M}(F)^g$ stand for $\bigcup_n M_n(F)^g$. We write $\mathcal{M}(F)$ for $\mathcal{M}(F)^1$. We denote the monoid generated by $x_1, \ldots, x_g$ by $\langle x \rangle$, and the free algebra in the variables $x = \{x_1, \ldots, x_g\}$ by $F\langle x \rangle$. The free algebra with involution in the variables $x_1, x_1^t, \ldots, x_g, x_g^t$ is denoted by $F\langle x, x^t \rangle$. The elements of degree d in $F\langle x \rangle$ (resp. $F\langle x, x^t \rangle$) are denoted by $F\langle x \rangle_d$ (resp. $F\langle x, x^t \rangle_d$). We write

$$\mathcal{C} = F\big[x_{ij}^{(k)} \mid 1 \leq i, j \leq n, 1 \leq k \leq g\big]$$

for the commutative polynomial ring in gn^2 variables. We equip $M_n(\mathcal{C})$ with the transpose involution fixing $\mathcal{C}$ pointwise. Let $\xi_k = (x_{ij}^{(k)}) \in M_n(\mathcal{C})$, $1 \leq k \leq g$, be the generic matrices. By GM_n we denote the unital subalgebra of $M_n(\mathcal{C})$ generated by the generic matrices, and by $\mathrm{GM}_n^\dagger$ the subalgebra of $M_n(\mathcal{C})$ generated by generic matrices and their transposes.

Remark 1.1.1. We should remark that the notation standard for the topic addressed in this chapter differs slightly from the previous chapters. The main difference is the use of X, which in the previous chapters denoted the set of variables, and now it will stand for an element in $\mathcal{M}(F)^g$, while the set of variables will be denoted by x.

1.2. Free Sets and Free Maps. Let $G = (G_n)_n$ be a sequence of groups with $G_n \subseteq \mathrm{GL}_n(F)$, satisfying

$$G_n \oplus G_m = \begin{pmatrix} G_n & 0 \\ 0 & G_m \end{pmatrix} \subseteq G_{n+m}. \tag{1.1}$$

We will be primarily concerned with the case $G_n = \mathrm{GL}_n(F)$ for all n, or G_n is the orthogonal group $\mathrm{O}_n(\mathbb{R})$ for all n. The modifications needed for the case of the unitary groups $G_n = \mathrm{U}_n(\mathbb{C})$ will be discussed in Section 6. For simplicity of notation we write $\mathrm{GL}_n, \mathrm{O}_n, \mathrm{U}_n$ instead of $\mathrm{GL}_n(F), \mathrm{O}_n(\mathbb{R}), \mathrm{U}_n(\mathbb{C})$, respectively. Let us denote $\mathrm{GL} = (\mathrm{GL}_n)_{n\in\mathbb{N}}$, $\mathrm{O} = (\mathrm{O}_n)_{n\in\mathbb{N}}$, $\mathrm{U} = (\mathrm{U}_n)_{n\in\mathbb{N}}$. A subset $\mathcal{U} \subseteq \mathcal{M}(F)^g$ is a sequence $\mathcal{U} = (\mathcal{U}[n])_{n\in\mathbb{N}}$, where each $\mathcal{U}[n] \subseteq M_n(F)^g$. The set $\mathcal{U}$ is a *G-free set* if it is closed with respect to simultaneous G-similarity and with respect to direct sums; i.e., for every $m, n \in \mathbb{N}$:

$$\sigma X \sigma^{-1} = (\sigma X_1 \sigma^{-1}, \dots, \sigma X_g \sigma^{-1}) \in \mathcal{U}[n] \tag{1.2}$$

for all $X \in \mathcal{U}[n]$, $\sigma \in G_n$, and

$$X \oplus Y = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \in \mathcal{U}[m+n] \tag{1.3}$$

for all $X \in \mathcal{U}[m], Y \in \mathcal{U}[n]$.

Let $\mathcal{U}$ be a G-free set. We call a sequence of functions $f = (f[n])_{n\in\mathbb{N}} : (\mathcal{U}[n])_{n\in\mathbb{N}} \to \mathcal{M}(F)$ a *G-free map*, if it respects G-similarity and direct sums; i.e, for every $m, n \in \mathbb{N}$:

$$f[n](\sigma X \sigma^{-1}) = \sigma\, f[n](X)\, \sigma^{-1} \tag{1.4}$$

for all $X \in \mathcal{U}[n]$, $\sigma \in G_n$, and

$$f[m+n](X \oplus Y) = f[m](X) \oplus f[n](Y) \tag{1.5}$$

for all $X \in \mathcal{U}[m], Y \in \mathcal{U}[n]$.

In the language of invariant theory the condition (1.4) says that $f[n]$ is a G_n-concomitant for the action of G_n induced by conjugation. If f satisfies only (1.4) for all n (and not necessarily (1.5)) we call it a *free G-concomitant*. Sometimes a GL-free map is called simply a *free map* and an O-free map is a *free map with involution.*

Remark 1.2.1. Another standard name for a G-concomitant is a G-equivariant map, which is used in Section 1. However, here we decided to use this slightly more compact name.

With a slight abuse of notation we sometimes also refer to a map $f : \mathcal{U} \to \mathcal{M}$ as a G-free map if its domain $\mathcal{U}$ is only closed under direct sums, f respects direct sums and f respects G-similarity on $\mathcal{U}$; i.e, for every $n \in \mathbb{N}$:

$$f[n](\sigma X \sigma^{-1}) = \sigma\, f[n](X)\, \sigma^{-1}$$

for all $X \in \mathcal{U}[n]$, $\sigma \in G_n$ such that $\sigma X \sigma^{-1} \in \mathcal{U}[n]$. In this case we can canonically extend f to the similarity invariant envelope of $\mathcal{U}$ (cf. [**KVV14**, Appendix A]), and remain in the framework of the given definition:

Proposition 1.2.2. *Let $\mathcal{U} \subseteq \mathcal{M}(F)^g$ be closed under direct sums, and let $f : \mathcal{U} \to \mathcal{M}(F)$ respect direct sums and G-similarity on $\mathcal{U}$. Then*

$$\tilde{\mathcal{U}} = \{\sigma A \sigma^{-1} \mid A \in \mathcal{U}[n], \sigma \in G_n, n \in \mathbb{N}\}$$

is a G-free set, and there exists a unique G-free map $\tilde{f} : \tilde{\mathcal{U}} \to \mathcal{M}(F)$ such that $\tilde{f}|_{\mathcal{U}} = f$, defined by $\tilde{f}(\sigma X \sigma^{-1}) = \sigma f(X) \sigma^{-1}$ for $X \in \mathcal{U}[n]$, $\sigma \in G_n$.

REMARK 1.2.3. In [**KVV14**, Appendix A] the proof is given in the case $G = \mathrm{GL}$. The same proof with obvious modifications works also for any sequence of groups $G = (G_n)_n$ satisfying (1.1), in particular for $G \in \{\mathrm{O}, \mathrm{U}\}$.

A G-free map f is F-analytic around 0 if there exists a neighborhood

$$\mathcal{B}(0, \delta) = \bigcup_n \{X \in M_n(F)^g \mid \|X\| < \delta_n\} \tag{1.6}$$

of 0 in $\mathcal{M}(F)^g$ such that $f[n]_{ij}$ is F-analytic on $\mathcal{B}(0, \delta)[n]$, $\delta = (\delta_n)_n$, and $\delta_n > 0$ for every $n \in \mathbb{N}$. It is a polynomial map of degree m if $f[n]_{ij}$ are polynomials in $x_{ij}^{(k)}$ of degree $\leq m$ and at least one of the polynomials $f[n]_{ij}$ is of degree m; it is homogeneous of degree m if $f[n]_{ij}$ are homogeneous polynomials of degree m or zero polynomials, and $f[n]_{ij}$ is of degree m for at least one triple (n, i, j).

If the map f is induced by a noncommutative polynomial; i.e, $f(X)$ is equal to the evaluation of a noncommutative polynomial $\tilde{f}$ at X, then we will identify f with $\tilde{f}$ and say that f is a noncommutative polynomial. (Note that the choice of $\tilde{f}$ is unique. See Lemma 2.2.4.)

1.3. Trace polynomials with involution. The free $*$-algebra with trace $\mathfrak{T}^\dagger\langle x, x^t\rangle$ is the algebra of noncommutative polynomials in the variables x_k, x_k^t over the polynomial algebra $\mathfrak{T}^\dagger$ in the infinitely many variables $\mathrm{tr}(w)$, where w runs over all representatives of the $*$-cyclic equivalence classes of words in the variables x_k, x_k^t; i.e., words u and v are equivalent if $u \overset{\text{cyc}}{\sim} v$ or $u^t \overset{\text{cyc}}{\sim} v$. The elements of $\mathfrak{T}^\dagger\langle x, x^t\rangle$ are *trace polynomials with involution* and elements of $\mathfrak{T}^\dagger$ are *pure trace polynomials with involution. Trace identities with involution* of the matrix algebra $M_n(F)$ are the elements in the kernel of the evaluation map from the free algebra with involution with trace to $M_n(F)$.

The corresponding free algebra is isomorphic to the subalgebra $\mathcal{T}_{g,n}^\dagger\langle \xi_k, \xi_k^t\rangle$ of $M_n(\mathcal{C})$ generated by generic matrices and $\mathcal{T}_{g,n}^\dagger$, where $\mathcal{T}_{g,n}^\dagger$ denotes the subalgebra of $\mathcal{C}$ generated by the traces $\mathrm{tr}(\eta_{i_1} \cdots \eta_{i_k})$, $\eta_j \in \{\xi_j, \xi_j^t\}$.

As in the case of ordinary trace polynomials these algebras can be interpreted in terms of invariant theory. The action of GL_n restricts to the action of O_n, and Procesi proved the following first fundamental theorem of matrix invariants for the orthogonal group. As for the second fundamental theorem we only state its corollary which will be needed in the sequel.

THEOREM 1.3.1. *[**Pro76**, Theorem 7.2, Proposition 8.3]*

- *FFT: The algebra $\mathcal{T}_{g,n}^\dagger$ is the algebra of O_n-invariant polynomial functions on the space of g-tuples of $n \times n$ matrices. The algebra $\mathcal{T}_{g,n}^\dagger\langle \xi_k, \xi_k^t\rangle$ is the algebra of O_n-equivariant polynomial maps from g-tuples of matrices to matrices.*
- *SFT': Every nontrivial trace identity with involution on $n \times n$ matrices has degree at least n.*

2. Power series expansions about scalar points

In this section we investigate two distinguished classes of free maps, namely polynomials and analytic free maps. We characterize free maps which are polynomials in Subsection 2.1, and use this to show that analytic free maps admit power series expansions about scalar points in Subsection 2.2. These results are classical for $G = \mathrm{GL}$ (cf. [**KVV14, Tay73, Voi10**]) and are new for $G = \mathrm{O}$. Throughout this section $G \in \{\mathrm{GL}, \mathrm{O}\}$.

2.1. Polynomial free maps. We start by characterizing free polynomial maps f via their "slices" $f[n]$. For $G = \mathrm{GL}$ this result is due to Kaliuzhnyi-Verbovetskyi and Vinnikov [**KVV14**, Theorem 6.1] who deduce it from their power series expansion theorem for analytic free maps. In contrast to this we shall first characterize free polynomial maps and employ this in Subsection 2.2 to establish power series expansions for analytic G-free maps. Our proofs are uniform in that they work for both $G = \mathrm{GL}$ and $G = \mathrm{O}$, and are purely algebraic, depending only on the invariant theory of matrices [**Pro76**].

PROPOSITION 2.1.1. *Let $f : \mathcal{M}(F)^g \to \mathcal{M}(F)$ be a G-free map. If f is a polynomial map and $\max_n \deg f[n] = d$, then f is a free polynomial of degree d. That is, $f \in F\langle x\rangle_d$ if $G = \mathrm{GL}$ and $f \in F\langle x, x^t\rangle_d$ if $G = \mathrm{O}$.*

PROOF. Since $f[n] : M_n(F)^g \to M_n(F)$ is a concomitant, it follows by the first fundamental theorem (see Theorem 1.1.2.1 and Theorem 1.3.1) that $f[n]$ is a trace polynomial of degree $\leq d$ in the variables x_k (resp. x_k, x_k^t). Since there do not exist nontrivial trace identities for $M_n(F)$ of degree less than n by SFT and SFT', we can write $f[n]$ in the case $n \geq d+1$ uniquely as

$$f[n] = \sum_M \mathrm{tr}(h_M^n)M,$$

where M runs over all monomials of degree $\leq d$, h_M is a trace polynomial and $\deg \mathrm{tr}(h_M^n) + \deg M \leq d$. Choose $n \geq d+1$. As f is a free map, we have

$$\begin{aligned}\sum_M \mathrm{tr}\left(h_M^{2n}(X \oplus Y)\right)M(X) \oplus \sum_M \mathrm{tr}\left(h_M^{2n}(X \oplus Y)\right)M(Y) = f[2n](X \oplus Y)\\ = f[n](X) \oplus f[n](Y) = \sum_M \mathrm{tr}\left(h_M^n(X)\right)M(X) \oplus \sum_M \mathrm{tr}\left(h_M^n(Y)\right)M(Y).\end{aligned}$$

Comparing both sides of the above expression we obtain

$$\mathrm{tr}\left(h_M^n(X)\right) = \mathrm{tr}\left(h_M^{2n}(X \oplus Y)\right) = \mathrm{tr}\left(h_M^n(Y)\right)$$

since $M_n(F)$ does not satisfy a nontrivial trace identity of degree d. Thus,

$$\mathrm{tr}\left(h_M^n(X)\right) = \alpha = \mathrm{tr}\left(h_M^n(Y)\right)$$

for some $\alpha \in F$. Hence, for every $n > N$, $f[n] \in \mathrm{GM}_n$ (resp. $f[n] \in \mathrm{GM}_n^\dagger$) is represented by an element $\tilde{f} \in F\langle x\rangle$ (resp. $\tilde{f} \in F\langle x, x^t\rangle$) of degree d. Since f is a free map, we can identify it with a noncommutative polynomial in the variables x_k (resp. x_k, x_k^t). $\square$

REMARK 2.1.2. We note that Proposition 2.1.1 holds also if f is only defined on $\mathcal{B}(0,\delta)$ (cf. Proposition 1.2.2), since polynomial functions that agree on an open subset of $M_n(F)^g$ represent the same function on $M_n(F)^g$.

2.2. Analytic free maps. We next turn our attention to analytic G-free maps. We show they admit unique convergent power series expansions about scalar points $a \in F^g$, extending classical results for $G = \mathrm{GL}$, cf. [**Tay73**, **Voi04**, **Voi10**, **KVV14**, **HKM12**]. By a translation we may assume without loss of generality that $a = 0$.

THEOREM 2.2.1. *Let $\mathcal{U}$ be a G-free set and $f : \mathcal{U} \to \mathcal{M}(F)$ an F-analytic G-free map, and let $\mathcal{B}(0,\delta) \subseteq \mathcal{U}$, where $\delta = (\delta_n)_{n\in\mathbb{N}}$, $\delta_n > 0$ for every $n \in \mathbb{N}$. Then there exists a unique formal power series*

$$F = \sum_{m=0}^{\infty} \sum_{|w|=m} F_w w, \tag{2.1}$$

where $w \in \langle x \rangle$ (resp. $w \in \langle x, x^t \rangle$), which converges in norm on $\mathcal{B}(0,\delta)$, with $f(X) = F(X)$ for $X \in \mathcal{B}(0,\delta)$.

REMARK 2.2.2. If f is uniformly bounded, and $G = \mathrm{GL}$ then the convergence of the power series F in (2.1) is uniform, cf. [**HKM12**, Proposition 2.24], while this conclusion does not hold when $G = \mathrm{O}$. We present examples in Section 5.

We first prove the existence, the uniqueness will follow from Proposition 2.2.5 below.

PROOF OF THE EXISTENCE. Since f is analytic, there exists for every $X \in M_n(F)^g$ a neighbourhood of 0 such that the function $t \mapsto f[n](tX)$ is defined and analytic in that neighbourhood. Hence, $f[n](tX)$ can be expressed in that neighbourhood as a convergent power series of the form $\sum_{m=0}^{\infty} t^m f[n]_m(X)$, where $f[n]_m(X)$ is a function of X. Note that for $X \in \mathcal{B}(0,\delta)$, this power series converges for $t = 1$. The function $f[n]_m$ is a homogeneous polynomial function of degree m. Indeed, let $s \in F$, $X \in M_n(F)^g$ and choose δ' such that $tsX \in \mathcal{B}(0,\delta)$ for $|t| \leq \delta'$. Then

$$\sum_{m=0}^{\infty} t^m f[n]_m(sX) = f(tsX) = \sum_{m=0}^{\infty} (ts)^m f[n]_m(X),$$

and thus $f[n]_m(sX) = s^m f[n]_m(X)$.

Let us show that f_m defined by $f_m[n] := f[n]_m$ is an analytic free map. Choose δ' such that $tX, tY, \sigma tX\sigma^{-1} \in \mathcal{B}(0,\delta)$ for $|t| < \delta'$. As f is a free map we have

$$\begin{aligned}\sum_{m=0}^{\infty} t^m f[n+n']_m(X \oplus Y) &= f[n+n'](tX \oplus tY) \\ &= f[n](tX) \oplus f[n'](tY) = \sum_{m=0}^{\infty} t^m \big(f[n]_m(X) \oplus f[n']_m(Y)\big),\end{aligned}$$

and

$$\sum_{m=0}^{\infty} t^m \sigma\, f[n]_m(X)\, \sigma^{-1} = \sigma\, f[n](tX)\, \sigma^{-1} = f[n](t\sigma X\sigma^{-1}) = \sum_{m=0}^{\infty} t^m f[n]_m(\sigma X \sigma^{-1})$$

for all $|t| < \delta'$, which implies that f_m is a G-free map. By construction, f_m is a homogeneous polynomial function of degree m (or 0) for every m. By Proposition 2.1.1, f_m can be represented by a noncommutative polynomial in the variables x_k (resp. x_k, x_k^t) of degree m. Thus, f can be expressed as a power series in noncommuting variables, $F = \sum f_m$. By construction, this power series converges on $\mathcal{B}(0,\delta)$. $\square$

While the theories of GL- and O-free maps enjoy certain similarities, there are also major differences. For instance, for GL-free maps continuity implies analyticity and there is a very useful formula [**HKM11**, Proposition 2.5], [**KVV14**, Theorem 7.2] connecting function values with the derivative:

$$f\begin{pmatrix} X & H \\ 0 & X \end{pmatrix} = \begin{pmatrix} f(X) & \delta f(X)(H) \\ 0 & f(X) \end{pmatrix}, \tag{2.2}$$

where $\delta f(X)(H)$ denotes the Gâteaux (directional) derivative of f at X in the direction H; i.e.,

$$\delta f(X)(H) = \lim_{t \to 0} \frac{f(X+tH) - f(X)}{t}.$$

For O-free maps continuity does not imply differentiability; see Section 5 for examples. However, for differentiable O-free maps we do have an analog of formula (2.2), which can be deduced from [**PTD13**, Lemma 2.3, Proposition 2.5], but we prove it here for the sake of completeness. We write $\mathrm{D}f$ for a derivative of f, it can be either the Gâteaux or the Fréchet derivative. The Lie bracket $[a,B]$ stands for $([a,B_1],\dots,[a,B_g])$, where $a\in M_n(F)$, $B=(B_1,\dots,B_g)\in M_n(F)^g$.

LEMMA 2.2.3. *Let $f:\mathcal{U}\to\mathcal{M}(F)$ be a real differentiable G-free map. Then the identity*

$$\mathrm{D}f(X)([a,X])=[a,f(X)] \tag{2.3}$$

holds for all $X\in\mathcal{U}[n]$, $a^t=-a\in M_n(\mathbb{R})$. In particular,

$$\mathrm{D}f\begin{pmatrix}Y&0\\0&Z\end{pmatrix}\begin{pmatrix}0&Y-Z\\Y-Z&0\end{pmatrix}=\begin{pmatrix}0&f(Y)-f(Z)\\f(Y)-f(Z)&0\end{pmatrix}. \tag{2.4}$$

PROOF. Note that e^{sa} is orthogonal for $a^t=-a\in M_n(\mathbb{R})$ and $s\in\mathbb{R}$. Thus we have

$$f(e^{-sa}Xe^{sa})=e^{-sa}f(X)e^{sa}$$

for every $X\in\mathcal{U}[n]$. Differentating with respect to s at 0 yields

$$\mathrm{D}f(X)([a,X])=[a,f(X)].$$

Take

$$a=\begin{pmatrix}0&I_n\\-I_n&0\end{pmatrix}\in M_{2n}(\mathbb{R}),$$

where I_n denotes the identity in $M_n(\mathbb{R})$. Setting $X=\begin{pmatrix}Y&0\\0&Z\end{pmatrix}$ we get the identity (2.4). □

We now show that the power series expansion is unique for a G-free function and give a way to recover its coefficients.

LEMMA 2.2.4. *If $f(X)=\sum_{|w|\le m}F_w w$, where the sum is over words in the variables x_k (resp. x_k,x_k^t), then we can obtain the coefficients F_w by evaluations of f on $M_{m+1}(F)$.*

PROOF. We proceed inductively. Assume that we can obtain coefficients of $f(X)=\sum_{|w|\le k}F_w w$ for $k<m$ by evaluations of f on $M_{k+1}(F)$. The case $k=1$ is trivial. Suppose that $k=m$. Let us determine the coefficient at $w=u_{i_1}^{j_1}\cdots u_{i_s}^{j_s}$, where $\sum_{k=1}^s j_k=m$ and $u_{i_k}\in\{x_{i_k},x_{i_k}^t\}$. We denote $s_k=\sum_{i=1}^k j_i$. Setting $a_i=0$ at the beginning, we define a g-tuple $(a_i)\in M_{m+1}(F)^g$ as follows. We let k run from 1 to s, and at step k we replace a_{i_k} by

$$a_{i_k}=\begin{cases}a_{i_k}+\sum_{u=s_{k-1}+1}^{s_k}e_{u,u+1} & \text{if } u_{i_k}=x_{i_k},\\ a_{i_k}+\sum_{u=s_{k-1}+1}^{s_k}e_{u+1,u} & \text{if } u_{i_k}=x_{i_k}^t.\end{cases}$$

We shall show that $\operatorname{tr}(f(a_1,\dots,a_g)e_{m+1,1})=F_w$. We need to find the coefficient of $f(a_1,\dots,a_g)$ expressed in the standard basis e_{ij}, $1\le i,j\le m+1$, of $M_{m+1}(F)$ at $e_{1,m+1}$. According to the definition of the a_i's it suffices to show that $e_{1,m+1}$ can be obtained in only one way as a product of $\le m$ matrix units from the set $S=\{e_{i,i+1},e_{i+1,i}\mid 1\le i\le m\}$. Note that the multiplication on the right of any matrix unit e_{ij} by any element of S either increases or decreases j by 1. In order to obtain $e_{1,m+1}$ as a product of $\le m$ elements from S, we can thus only choose matrix units which increase the second subscript of the preceding matrix unit in the product. Hence, $e_{1,m+1}=e_{12}\cdots e_{m,m+1}$, and any other product of $\le m$ elements from S will be different from $e_{1,m+1}$. As each $e_{i,i+1}$ appears only in one of the a_i,a_i^t, $1\le i\le g$, the order $e_{12},\dots,e_{m,m+1}$ corresponds to exactly one order of the $a_i's$. By the definition of a_i this order corresponds to w. Now we can find the coefficients of $f - \sum_{|w|=m} F_w w = \sum_{|w|<m} F_w w$ by the induction hypothesis on $M_m(F) \subset M_{m+1}(F)$. $\square$

PROPOSITION 2.2.5. *Suppose that a G-free map f has a power series expansion in a neighbourhood $\mathcal{B}(0,\delta)$ of 0, $\delta = (\delta_n)_{n\in\mathbb{N}}$; i.e.,*

$$f(X) = \sum_{m=0}^{\infty} \sum_{|w|=m} F_w w(X),$$

for $X \in \mathcal{B}(0,\delta)$. Then F_w for $|w| = m$ is determined by the m-th derivative of the function $t \mapsto f[m+1](tX)$ at 0 and hence by its evaluation on $M_{m+1}(F)$.

PROOF. Let $|t| < 1$, then $tX \in \mathcal{B}(0,\delta)[n]$ for every $X \in \mathcal{B}(0,\delta)[n]$, and

$$f[n](tX) = \sum_{m=0}^{\infty} t^m f_m[n](X)$$

is a convergent power series in t, where f_m are homogeneous noncommutative polynomials of degree m. We can thus determine $f_m[n](X)$ as

$$\frac{1}{m!}\frac{\mathrm{d}}{\mathrm{d}t^m} f[n](tX)\bigg|_{t=0}.$$

Since $M_n(F)$ does not admit a nontrivial polynomial identity (with involution) of degree $< n$ (see e.g. [**Row80**, Lemma 1.4.3, Remark 2.5.14]), f_m is uniquely determined on $M_{m+1}(F)$. Hence we can recover f_m by the m-th derivative of the function $t \mapsto f[m+1](tX)$. The coefficients of the polynomial f_m can be constructively determined by evaluations on $M_{m+1}(F)$ by Lemma 2.2.4. $\square$

3. Power series expansions about non-scalar points

Theorem 2.2.1 gives a convergent power series expansion of a free analytic map about a *scalar point* $a \in F^g$. In this section we present power series expansions about non-scalar points $A \in M_n(F)^g$, whose homogeneous components are generalized polynomials. These are the topic of Subsection 3.1 and their obtained properties will be used in Subsection 3.2 to deduce the desired power series expansion. Our methods are algebraic, and work for $G = \mathrm{GL}$ and $G = \mathrm{O}$. For $G = \mathrm{GL}$ a similar result has been obtained earlier in [**KVV14**] with a different proof.

Throughout this section $G \in \{\mathrm{GL}, \mathrm{O}\}$.

3.1. Generalized Polynomials. Let us recall that we call the elements of the free product $M_n(F) * F\langle x\rangle$ *generalized polynomials.* They can be written in the form

$$\sum a_{i_0} x_{k_1} a_{i_1} x_{k_2} \cdots a_{i_{\ell-1}} x_{k_\ell} a_{i_\ell},$$

where $a_{i_j} \in M_n(F)$. Let e_{ij} denote the standard matrix units of $M_n(F)$. Then a basis of $M_n(F) * F\langle x\rangle$ consists of monomials

$$e_{i_0,j_0} x_{k_1} e_{i_1,j_1} x_{k_2} \cdots e_{i_{\ell-1},j_{\ell-1}} x_{k_\ell} e_{i_\ell,j_\ell}$$

for $\ell \in \mathbb{N}_0$, $I, J \in \{1,\ldots,n\}^{\ell+1}$, $K \in \{1,\ldots,g\}^{\ell}$, where $I = (i_0,\ldots,i_\ell)$, $J = (j_0,\ldots,j_\ell)$, $K = (k_1,\ldots,k_\ell)$.

The algebra $M_n(F) * F\langle x\rangle$ can be evaluated (as an algebra with unity) in $M_{ns}(F)$ for $s \in \mathbb{N}$ and we have an isomorphism

$$\mathrm{Hom}_{M_n}(M_n(F) * F\langle x\rangle, M_{ns}(F)) \cong \mathrm{Hom}(\mathfrak{W}_n(F\langle x\rangle), M_s(F)), \tag{3.1}$$

where $\mathfrak{W}_n$ denotes the matrix reduction functor (see [**Coh95**, Section 1.7]). The isomorphism is a consequence of the identity

$$M_n(F) * F\langle x\rangle \cong M_n(\mathfrak{W}_n(F\langle x\rangle)). \tag{3.2}$$

For the free algebra $F\langle x\rangle = F\langle x_1,\dots,x_g\rangle$ we have

$$\mathfrak{W}_n(F\langle x\rangle) = F\langle y_{ij}^{(k)} \mid 1\le i,j\le n, 1\le k\le g\rangle,$$

where $y_{ij}^{(k)}$, as the brackets suggest, denote free noncommutative variables. For example, the evaluation of the element

$$e_{11}x_1e_{12}x_2e_{22}\in M_2(F)*F\langle x\rangle$$

in $M_4(F)$, defined by mapping x_1,x_2 to $A,B\in M_4(F)$, is

$$\begin{pmatrix} I_2 & \\ & \end{pmatrix}\begin{pmatrix} A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\begin{pmatrix} & I_2\\ & \end{pmatrix}\begin{pmatrix} B_{11} & B_{12}\\ B_{21} & B_{22}\end{pmatrix}\begin{pmatrix} & \\ & I_2\end{pmatrix}=\begin{pmatrix} & A_{11}B_{22}\\ & \end{pmatrix}$$

where I_2 denotes the identity of $M_2(F)$, and A_{ij} (resp. B_{ij}) denotes the (i,j)-block entry of A (resp. B), or

$$(e_{11}\otimes I_2)A(e_{12}\otimes I_2)B(e_{22}\otimes I_2) = e_{12}\otimes A_{11}B_{22},$$

viewed as en element in $M_2(F)\otimes M_2(F)\cong M_4(F)$.

Note that (3.1) and (3.2) imply that no generalized polynomial vanishes on $M_{ns}(F)$ for all s. In fact, two generalized polynomials of degree $2d$ which agree on $M_{ns}(F)$ for some $s>d$ are equal. We denote by $\mathrm{g}J_{ns}$ the ideal of the elements in $M_n(F)*F\langle x\rangle$ that vanish when evaluated on $M_{ns}(F)$ and let

$$C_{ns} = F\big[x_{ij}^{(k)} \mid 1\le i,j\le ns, 1\le k\le g\big].$$

The quotient algebra $\mathrm{gGM}_{ns} = \big(M_n(F)*F\langle x\rangle\big)/\mathrm{g}J_{ns}$ is isomorphic to the image of

$$\phi : M_n(F)*F\langle x\rangle \to M_{ns}(C_{ns}),$$

defined by mapping x_k to the corresponding generic matrix $(x_{ij}^{(k)})$. We write gR_{ns} for the subalgebra of $M_{ns}(C_{ns})$ generated by gGM_{ns} and traces of the elements in gGM_{ns}. Note that every polynomial map $p: M_{ns}(F)^g\to M_{ns}(F)$ can be considered as an element $\tilde p\in M_{ns}(C_{ns})$.

Let GL_{ns} act on $M_{ns}(F)$ by conjugation. We will be interested in the action of its subgroup $I_n\otimes\mathrm{GL}_s$. In the next proposition we describe the invariants and concomitants of this action.

Proposition 3.1.1. *If $p: M_{ns}(F)^g\to M_{ns}(F)$ is an $I_n\otimes\mathrm{GL}_s$-concomitant, then $\tilde p\in\mathrm{gR}_{ns}$.*

Proof. We can assume that p is multilinear of degree d. Then p corresponds to an element in $(M_{ns}(F)^{\otimes d})^*\otimes M_{ns}(F)$, which is canonically isomorphic to $M_n(F)^{\otimes d+1}\otimes (M_s(F)^{\otimes d})^*\otimes M_s(F)$ as $I_n\otimes\mathrm{GL}_s$-module. The action of the group $I_n\otimes\mathrm{GL}_s$ reduces to the action of GL_s on $(M_s(F)^{\otimes d})^*\otimes M_s(F)$. The invariants of this action correspond to multilinear trace polynomials of degree d in $M_s(C_s)$ by FFT. Moreover, the elements of the form

$$\sum_{I,J} e_{i_1j_1}\otimes\cdots\otimes e_{i_dj_d}\otimes\tau_{IJ},$$

where $\tau_{IJ}\in(M_s(F)^{\otimes d})^*\otimes M_s(F)$ is a GL_s-concomitant map, can be identified with multilinear elements of degree d in gR_{ns}. $\square$

3.1.1. *Generalized polynomials with involution.* To consider the case of algebras with involution we need to introduce some additional notation. We call the elements of the algebra $M_n(F) * F\langle x, x^t\rangle$ *generalized polynomials with involution.* By $\mathrm{g}J_{ns}^{\dagger}$ we denote the ideal of elements in $M_n(F) * F\langle x, x^t\rangle$ that vanish on $M_{ns}(F)$. The quotient algebra is isomorphic to the subalgebra $\mathrm{gGM}_{ns}^{\dagger}$ of $M_{ns}(C_{ns})$ generated by gGM_{ns} and transposes of elements in gGM_{ns}. We write $\mathrm{gR}_{ns}^{\dagger}$ for the subalgebra of $M_{ns}(C_{ns})$ generated by $\mathrm{gGM}_{ns}^{\dagger}$ and traces of elements in $\mathrm{gGM}_{ns}^{\dagger}$.

We have the (usual) action of O_{ns} on $M_{ns}(C_{ns})$. The following proposition is the analog of Proposition 3.1.1 for the action of $I_n \otimes \mathrm{O}_s$ on $M_{ns}(C_{ns})$.

PROPOSITION 3.1.2. *If $p \in M_{ns}(F)^g \to M_{ns}(F)$ is an $I_n \otimes \mathrm{O}_s$-concomitant, then $\tilde{p} \in \mathrm{gR}_{ns}^{\dagger}$.*

PROOF. The proof goes along the same lines as that of Proposition 3.1.1, we only need to invoke FFT for the orthogonal group instead of FFT for the general linear group. $\square$

3.1.2. *Block and centralizing G-concomitants.* Let us denote

$$\mathcal{M}_n(F)^k = \bigcup_s M_{ns}(F)^k,$$

$k \in \mathbb{N}$. We say that a map $f : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)$ is $I_n \otimes G$-concomitant if

$$f[ns] : \big(M_n(F) \otimes M_s(F)\big)^g \to M_n(F) \otimes M_s(F)$$

is a $I_n \otimes G_s$-concomitant for every $s \in \mathbb{N}$.

PROPOSITION 3.1.3. *If $f : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)$ is a homogeneous polynomial map of degree d and $I_n \otimes \mathrm{GL}$-concomitant (resp. $I_n \otimes \mathrm{O}$-concomitant) that preserves direct sums, then $f \in M_n(F) * F\langle x\rangle$ (resp. $f \in M_n(F) * F\langle x, x^t\rangle$).*

PROOF. We prove the lemma only in the case $G = \mathrm{GL}$, the modifications needed to treat the case $G = \mathrm{O}$ are minor. We can assume that f is multilinear. Since $f[ns]$ is a $I_n \otimes \mathrm{GL}_s$-concomitant, $f[ns] \in \mathrm{gR}_{ns}$ by Proposition 3.1.1. We can view $f[ns]$ as an element in $M_n(F)^{\otimes d+1} \otimes (M_s(F)^{\otimes d})^* \otimes M_s(F)$ and write it in the form

$$f[ns] = \sum_{I,J} e_{i_1 j_1} \otimes \cdots \otimes e_{i_d j_d} \otimes e_{i_{d+1} j_{d+1}} \otimes \tau_{IJ}^{(s)},$$

where $\tau_{IJ}^{(s)}$ is a GL_s-concomitant. Let $s > d$. Since f preserves direct sums we have

$$f[ns](X) \oplus f[ns](Y) = f[2ns](X \oplus Y).$$

We obtain for all I, J an identity

$$\tau_{IJ}^{(s)}(X) \oplus \tau_{IJ}^{(s)}(Y) = \tau_{IJ}^{(2s)}(X \oplus Y). \tag{3.3}$$

Let us fix I, J. To simplify the notation we write $\tau^{(s)}$ instead of $\tau_{IJ}^{(s)}$. We have

$$\tau^{(s)} = \sum_M h_M^{(s)} M,$$

where h_M is a pure trace polynomial, M is a monomial in the variables x_k, and $\deg M + \deg h_M = d$. Then the identity (3.3) together with the fact that there are no trace identities of $M_s(F)$ of degree $< s$ yields

$$h_M^{(s)}(X) = h_M^{(2s)}(X \oplus Y) = h_M^{(s)}(Y)$$

for all monomials M, which implies that

$$\tau^{(s)} = \sum_M \alpha_M M$$

for some $\alpha_M \in F$. Thus, $f[ns] \in \mathrm{gGM}_{ns}$ for every $s > d$ is represented by the same generalized polynomial $\tilde{f}$. Since f respects direct sums, we can identify it with $\tilde{f}$. $\square$

For a subset B of $M_n(F)$ we denote by $C(B)$ its *centralizer* in $M_n(F)$; i.e.,

$$C(B) = \{c \in M_n(F) \mid cb = bc \text{ for all } b \in B\},$$

while $C_{G_n}(B)$ stands for $C(B) \cap G_n$. We say that a map $f : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)$ is a $(C_{G_n}(B), G)$*-concomitant* if $f[ns]$ is a $(C_{G_n}(B) \otimes M_s(F)) \cap G_{ns}$-concomitant for every $s \in \mathbb{N}$.

LEMMA 3.1.4. *Let B be a subalgebra of $M_n(F)$. If $f : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)$ is a homogeneous polynomial map of degree d that is a $(C_{\mathrm{GL}_n}(B), \mathrm{GL})$-concomitant, then $f \in C(C(B)) * F\langle x\rangle$.*

PROOF. By Lemma 3.1.3, $f \in M_n(F) * F\langle x\rangle$. Since GL_n is dense in $M_n(F)$, the vector space spanned by $C_{\mathrm{GL}_n}(B)$ coincides with $C(B)$. Thus we can choose a basis $\{c_1, \dots, c_t\}$ of $C(B)$ with $c_\ell \in \mathrm{GL}_n$. Let $\{b_1, \dots, b_u\}$ be a basis of $C(C(B))$ and complete it to a basis $\{b_\ell \mid 1 \le \ell \le n^2\}$ of $M_n(F)$. We can write f uniquely as

$$f = \sum_{I,K} \alpha_{IK} b_{i_1} x_{k_1} b_{i_2} \cdots x_{k_d} b_{i_{d+1}},$$

where I runs over all $d+1$-tuples of elements in $\{1, \dots, n^2\}$, and K over all d-tuples of elements in $\{1, \dots, g\}$. Take $s > d$ and evaluate f on $M_{2nts}(F) \cong M_n(F) \otimes M_{2t}(F) \otimes M_s(F)$. Note that f on $M_{2nts}(F)$ can be identified with the evaluation of the generalized polynomial

$$\sum_{I,K} \alpha_{IK} \Big(\sum_{i=1}^{2t} b_{i_1} \otimes e_{ii}\Big) x_{k_1} \Big(\sum_{i=1}^{2t} b_{i_2} \otimes e_{ii}\Big) \cdots x_{k_d} \Big(\sum_{i=1}^{2t} b_{i_{d+1}} \otimes e_{ii}\Big).$$

in $M_{2nt}(F) * F\langle x\rangle = (M_n(F) \otimes M_{2t}(F)) * F\langle x\rangle$, and every element in $M_{2nt}(F) * F\langle x\rangle$ has a unique expression with the matrix coefficients $b_\ell \otimes e_{ij}$, $1 \le i, j \le 2t$, $1 \le \ell \le n^2$, on $M_{2nts}(F)$ as $s > d$. Let

$$\sigma = \Big(\alpha 1 \otimes 1 + \beta \sum_{\ell=1}^{t} (c_\ell \otimes e_{\ell,t+\ell} - c_\ell^{-1} \otimes e_{t+\ell,\ell})\Big) \otimes 1$$

for $\alpha^2 + \beta^2 = 1$, $\alpha, \beta \in \mathbb{R}$. Then $\sigma \in (C_{\mathrm{GL}_n}(B) \otimes M_{2t}(F) \otimes M_s(F)) \cap \mathrm{GL}_{2nts}$. Note that

$$\sigma^{-1} = \Big(\alpha 1 \otimes 1 - \beta \sum_{\ell=1}^{t} (c_\ell \otimes e_{\ell,t+\ell} - c_\ell^{-1} \otimes e_{t+\ell,\ell})\Big) \otimes 1. \tag{3.4}$$

Since f is a $(C_{\mathrm{GL}_n}(B), \mathrm{GL})$-concomitant we have

$$\sum_{I,K} \alpha_{IK} b_{i_1}^{\sigma} x_{k_1} b_{i_2}^{\sigma} \cdots x_{k_d} b_{i_{d+1}}^{\sigma} = \sum_{I,K} \alpha_{IK} b_{i_1} x_{k_1} b_{i_2} \cdots x_{k_d} b_{i_{d+1}},$$

where by a slight abuse of notation b_i denotes $b_i \otimes 1 \otimes 1$, and

(3.5)

$$\begin{aligned} b_i^{\sigma} = \sigma^{-1} b_i \sigma = \alpha^2 b_i \otimes 1 \otimes 1 + \sum_{\ell=1}^{t} \beta^2 c_\ell b_i c_\ell^{-1} \otimes e_{\ell\ell} \otimes 1 + \beta^2 c_\ell^{-1} b_i c_\ell \otimes e_{t+\ell,t+\ell} \otimes 1 \\ + \alpha\beta (b_i c_\ell - c_\ell b_i) \otimes e_{\ell,t+\ell} \otimes 1 - \alpha\beta (b_i c_\ell^{-1} - c_\ell^{-1} b_i) \otimes e_{t+\ell,\ell} \otimes 1. \end{aligned}$$

Since $s > d$ both sides of equation (3.5) have a unique expression as generalized polynomials in $M_{2tn} * F\langle x\rangle$ with the generalized coefficients $b_\ell \otimes e_{ij}$, $1 \le i, j \le 2t$, $1 \le \ell \le n^2$. We thus derive

$$\sum_k \alpha_{I_k^j K}(b_k c_\ell - c_\ell b_k) = 0 \tag{3.6}$$

for every $1 \le j \le d+1$, $1 \le \ell \le t$, where I_k^j denotes a tuple of $d+1$-elements in $\{1, \dots, n^2\}$ with k at the j-th position. Equation (3.6) implies that

$$\sum_k \alpha_{I_k^j K} b_k \in C(C(B)),$$

which is by the choice of b_ℓ, $1 \le \ell \le n^2$, only possible if $\alpha_{I_k^j K} = 0$ for $b_k \notin C(C(B))$. Therefore we have $f \in C(C(B)) * F\langle x\rangle$. □

LEMMA 3.1.5. *If B is a $*$-subalgebra of $M_n(\mathbb{R})$, then the subalgebra generated by $C_{\mathrm{O}_n}(B)$ is equal to $C(B)$, and $C(C_{\mathrm{O}_n}(B)) = C(C(B)) = B$.*

PROOF. Since B is a $*$-subalgebra of $M_n(\mathbb{R})$, $C(B)$ is also a $*$-subalgebra of $M_n(\mathbb{R})$, thus semisimple. Notice that in order to show that $\mathbb{R}\langle C_{\mathrm{O}_n}(B)\rangle$, the subalgebra of $C(B)$ generated by $C_{\mathrm{O}_n}(B)$, coincides with $C(B)$, we can assume that $C(B)$ is simple. We have $c^t - c \in \operatorname{span} C_{\mathrm{O}_n}(B)$, the vector subspace of $M_n(\mathbb{R})$ spanned by $C_{\mathrm{O}_n}(B)$, for every $c \in C(B)$. Indeed, $e^{\lambda(c^t - c)} \in C_{\mathrm{O}_n}(B)$ for every $\lambda \in \mathbb{R}$, $c \in C(B)$ yields $c^t - c \in \operatorname{span} C_{\mathrm{O}_n}(B)$. If $C(B)$ is isomorphic to $\mathbb{R}$, $M_2(\mathbb{R})$, $\mathbb{C}$, or $M_2(\mathbb{C})$, where the involution on $\mathbb{C}$ is the complex conjugation, then one can easily verify that $\operatorname{span} C_{O_n}(B) = C(B)$. Recall that a finite dimensional simple $\mathbb{R}$-algebra with involution which is not isomorphic to $\mathbb{R}$, $M_2(\mathbb{R})$, $\mathbb{C}$, or $M_2(\mathbb{C})$ coincides with its subalgebra generated by the skew-symmetric elements (see e.g. [**KMRT98**, Lemma 2.26]). Therefore $\mathbb{R}\langle C_{\mathrm{O}_n}(B)\rangle = C(B)$, which further implies $C(C_{\mathrm{O}_n}(B)) = C(C(B))$, and the identity $C(C(B)) = B$ follows from the double centralizer theorem (see e.g. [**KMRT98**, Theorem 1.5]). □

LEMMA 3.1.6. *Let B be a $*$-subalgebra of $M_n(\mathbb{R})$. If $f : \mathcal{M}_n(\mathbb{R})^g \to \mathcal{M}_n(\mathbb{R})$ is a homogeneous polynomial map of degree d that is a $(C_{\mathrm{O}_n}(B), \mathrm{O})$-concomitant, then $f \in B * \mathbb{R}\langle x, x^t\rangle$.*

PROOF. Since the proof is similar to that of Lemma 3.1.6 we omit some of the details. By Proposition 3.1.2 we have $f \in M_n(\mathbb{R}) * \mathbb{R}\langle x, x^t\rangle$. Let $c_1, \dots, c_t$ be a basis of $\operatorname{span} C_{\mathrm{O}_n}(B)$, the vector space spanned by $C_{\mathrm{O}_n}(B)$, with $c_\ell \in \mathrm{O}_n$. Let us write

$$f = \sum_{I,K} \alpha_{IK} b_{i_1} u_{k_1} b_{i_2} \cdots u_{k_d} b_{i_{d+1}},$$

where $u_k \in \{x_k, x_k^t\}$. Take $s > d$ and evaluate f on $M_{2nts}(F)$. Let

$$\sigma = \Big(\alpha 1 \otimes 1 + \beta \sum_{\ell=1}^{t} (c_\ell \otimes e_{\ell, t+\ell} - c_\ell^t \otimes e_{t+\ell,\ell})\Big) \otimes 1$$

for $\alpha^2 + \beta^2 = 1$, $\alpha, \beta \in \mathbb{R}$. Then $\sigma \in (C_{\mathrm{O}_n}(B) \otimes M_{2t}(F) \otimes M_s(F)) \cap \mathrm{O}_{2nts}$ and

$$\sigma^t = \Big(\alpha 1 \otimes 1 - \beta \sum_{\ell=1}^{t} (c_\ell \otimes e_{\ell, t+\ell} - c_\ell^t \otimes e_{t+\ell,\ell})\Big) \otimes 1. \tag{3.7}$$

Since f is a $(C_{\mathrm{O}_n}(B), \mathrm{O})$-concomitant we have

$$\sum_{I,K} \alpha_{IK} b_{i_1}^\sigma u_{k_1} b_{i_2}^\sigma \cdots u_{k_d} b_{i_{d+1}}^\sigma = \sum_{I,K} \alpha_{IK} b_{i_1} u_{k_1} b_{i_2} \cdots u_{k_d} b_{i_{d+1}},$$

where b_i denotes $b_i \otimes 1 \otimes 1$, and

$$b_i^\sigma = \sigma^t b_i \sigma = \alpha^2 b_i \otimes 1 \otimes 1 + \sum_{\ell=1}^{t} \beta^2 c_\ell b_i c_\ell^t \otimes e_{\ell\ell} \otimes 1 + \beta^2 c_\ell^t b_i c_\ell \otimes e_{t+\ell,t+\ell} \otimes 1+ \\ + \alpha\beta(b_i c_\ell - c_\ell b_i) \otimes e_{\ell,t+\ell} \otimes 1 - \alpha\beta(b_i c_\ell^t - c_\ell^t b_i) \otimes e_{t+\ell,\ell} \otimes 1.$$

As $s > d$ both sides of the last identity have a unique expression as generalized polynomials in $M_{2tn} * \mathbb{R}\langle x, x^t\rangle$ with the generalized coefficients $b_\ell \otimes e_{ij}$, $1 \leq i, j \leq 2t$, $1 \leq \ell \leq n^2$. Thus, $\alpha_{I_k^j K} = 0$ for $b_k \notin C(C_{\mathrm{O}_n}(B))$, where I_k^j denotes a tuple of $d+1$-elements in $\{1, \dots, n^2\}$ with k at the j-th position. Since $C(C_{\mathrm{O}_n}(B)) = B$ by Lemma 3.1.5, f belongs to $B * \mathbb{R}\langle x, x^t\rangle$. □

3.2. Power series expansions about non-scalar points. We next turn to analytic free maps and exhibit their power series expansions about a non-scalar point A. Homogeneous components of such an expansion will be generalized polynomials. For $G = \mathrm{GL}$ their matrix coefficients belong to the double centralizer $C(C(A))$, while for $G = \mathrm{O}$ they lie in the $*$-subalgebra $F\langle A, A^t\rangle$ generated by A.

Let us first introduce neighbourhoods of non-scalar points. Given $A \in M_n(F)^g$, set

$$\mathcal{B}(A, \delta) = \bigcup_{s=1}^{\infty} \big\{X \in M_{ns}(F)^g \mid \big\|X - \bigoplus_{i=1}^{s} A\big\| < \delta_s\big\},$$

where $\delta = (\delta_s)_{s\in\mathbb{N}}$, $\delta_s > 0$ for every $s \in \mathbb{N}$.

3.2.1. *GL-free maps.* The next theorem gives a power series expansion of a GL-free map f about $A = (A_1, \dots, A_g) \in M_n(F)^g$, whose matrix coefficients are elements of the double centralizer algebra $C(C(F\langle A\rangle)) \subseteq M_n(F)$ of the subalgebra $F\langle A\rangle$ generated by $A_1, \dots, A_g$.

THEOREM 3.2.1. *Let $\mathcal{U}$ be a* GL*-free set, $f : \mathcal{U} \to \mathcal{M}(F)$ be an F-analytic* GL*-free map, and let $\mathcal{B}(A, \delta) \subseteq \mathcal{U}$, where $A \in M_n(F)^g$, and $\delta = (\delta_s)_{s\in\mathbb{N}}$, $\delta_s > 0$ for every $s \in \mathbb{N}$. Then there exist unique generalized polynomials $f_m \in C(C(F\langle A\rangle))\rangle * F\langle x\rangle$ of degree m so that the formal power series*

$$F(X) = \sum_{m=0}^{\infty} f_m(X - A), \tag{3.8}$$

converges in norm on the neighbourhood $\mathcal{B}(A, \delta)$ of A to f.

PROOF. As $A \in \mathcal{U}[n]$ and $\mathcal{U}$ is a GL-free set we have

$$A^{\oplus s} = \bigoplus_{i=1}^{s} A \in \mathcal{U}[ns]$$

for every $s \in \mathbb{N}$. Since $f[ns]$ is analytic in a neighbourhood of $A^{\oplus s}$, the function

$$t \mapsto f[ns]\Big(A^{\oplus s} + t\big(X - A^{\oplus s}\big)\Big)$$

is defined and analytic for all $|t| < \delta_X$, where δ_X depends on $X \in M_{ns}(F)$. Thus, we can expand it in a power series

$$f[ns]\Big(A^{\oplus s} + t\big(X - A^{\oplus s}\big)\Big) = \sum_{m=0}^{\infty} t^m f[ns]_m\big(X - A^{\oplus s}\big) \tag{3.9}$$

that converges for $|t| < \delta_X$. If $X \in \mathcal{B}(A, \delta)$, then we have $\delta_X \geq 1$. We claim that $f[ns]_m$ is a homogeneous polynomial function of degree m. Indeed, as

$$\begin{aligned}\sum_{m=0}^{\infty} t_1^m f[ns]_m\Big(t_2\big(X - A^{\oplus s}\big)\Big) =& f[ns]\Big(A^{\oplus s} + t_1 t_2\big(X - A^{\oplus s}\big)\Big)\\ =& \sum_{m=0}^{\infty} t_1^m t_2^m f[ns]_m\big(X - A^{\oplus s}\big)\end{aligned}$$

for all t_1 that satisfy $|t_1|, |t_1 t_2| < \delta_X$, we obtain

$$f[ns]_m(tY) = t^m f[ns]_m(Y)$$

for all $t \in F$, $Y \in M_{ns}(F)^g$. Let us show that

$$f_m : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)$$

defined by $f_m[ns] := f[ns]_m$ is a $(C_{\mathrm{GL}_n}(B), \mathrm{GL})$-concomitant that preserves direct sums. Take $s \in \mathbb{N}$, $\sigma \in (C_{\mathrm{GL}_n}(F\langle A\rangle) \otimes M_s(F)) \cap \mathrm{GL}_{ns}$ and note that

$$\sigma A^{\oplus s} \sigma^{-1} = A^{\oplus s}.$$

Then the identity

$$\begin{aligned}\sum t^m \sigma f[ns]_m\big(X - A^{\oplus s}\big)\sigma^{-1} &= \sigma f[ns]\Big(A^{\oplus s} + t\big(X - A^{\oplus s}\big)\Big)\sigma^{-1}\\ &= f[ns]\Big(A^{\oplus s} + t\big(\sigma X \sigma^{-1} - A^{\oplus s}\big)\Big)\\ &= \sum t^m f[ns]_m\Big(\sigma\big(X - A^{\oplus s}\big)\sigma^{-1}\Big),\end{aligned}$$

for all small enough t yields the desired conclusion.

To conclude the proof of the existence we proceed as at the end of the proof of existence in Theorem 2.2.1. Thus, $f_m \in C(C(F\langle A\rangle)) * F\langle x\rangle$ by Lemma 3.1.4. Note that setting $t = 1$ in (3.9) establishes the existence of the desired power series.

For the uniqueness, we can also follow the proof of uniqueness in Theorem 2.2.1 carried out in Lemma 2.2.4 and Proposition 2.2.5, after recalling the identity (3.1). Hence we can recover f_m by the m-th derivative of the function $t \mapsto f[n(m+1)](t(X - A))$ at 0, and the matrix coefficients of the generalized polynomial f_m can be determined by evaluations on $M_{n(m+1)}(F)$. $\square$

REMARK 3.2.2. If f is a uniformly bounded GL-free map then the convergence of F in (3.8) is uniform, which can be proved in the same way as the analogous statement for $F = \mathbb{C}$ and power series expansion about scalar points in the last part of the proof of [**HKM12**, Proposition 2.24]. The only modification needed is to replace $\exp(\mathrm{i}t)I_{ns}, \exp(-\mathrm{i}mt)I_{ns} \in M_{ns}(\mathbb{C})$ in the equation

$$C \geq \left\| \frac{1}{2\pi} \int f(\exp(\mathrm{i}t)X)\exp(-\mathrm{i}mt)dt \right\| = \|f^{(m)}(X)\|$$

with the corresponding matrices in $M_{2ns}(\mathbb{R})$.

In general one cannot expect the matrix coefficients of the power series expansion of a GL-free map f about a non-scalar point A to lie in $F\langle A\rangle * F\langle x\rangle$. In this case one would have $f(A) \in F\langle A\rangle$, which is not always the case by [**AM14**, Theorem 7.7]. However, this does hold true in the case that A is a generic point. That is, if $g = 1$, then A is similar to a diagonal matrix with n distinct eigenvalues, and if $g > 1$ then $F\langle A\rangle = M_n(F)$.

COROLLARY 3.2.3. *Let $\mathcal{U}$ be a* GL*-free set, $f : \mathcal{U} \to \mathcal{M}(F)$ be an F-analytic* GL*-free map, and let $\mathcal{B}(A, \delta) \subseteq \mathcal{U}$, where $A \in M_n(F)^g$ is a generic point, and*

$\delta = (\delta_s)_{s\in\mathbb{N}}$, $\delta_s > 0$ *for every* $s \in \mathbb{N}$. *Then there exist generalized polynomials* $f_m \in M_n(F) * F\langle x\rangle$ *of degree* m *so that the formal power series*

$$F(X) = \sum_{m=0}^{\infty} f_m(X - A),$$

converges in norm on the neighbourhood $\mathcal{B}(A, \delta)$ *of* A *to* f.

3.2.2. O-*free maps.* In the case of free maps with involution the matrix coefficients in the power series expansion of an O-free map about $A = (A_1, \dots, A_g) \in M_n(F)^g$ lie in the $*$-subalgebra $F\langle A, A^t\rangle$ of $M_n(F)$ generated by $A_1, \dots, A_g$. This contrasts the analogous result for GL-free maps (Theorem 3.2.1) where the double centralizer of $F\langle A\rangle$ is required.

THEOREM 3.2.4. *Let* $\mathcal{U}$ *be an* O*-free set,* $f : \mathcal{U} \to \mathcal{M}(F)$ *be an* F*-analytic* O*-free map, and let* $\mathcal{B}(A, \delta) \subseteq \mathcal{U}$, *where* $A \in M_n(F)^g$, *and* $\delta = (\delta_s)_{s\in\mathbb{N}}$, $\delta_s > 0$ *for every* $s \in \mathbb{N}$. *Then there exist unique generalized polynomials* $f_m \in F\langle A, A^t\rangle * F\langle x, x^t\rangle$ *of degree* m *so that the formal power series*

$$F(X) = \sum_{m=0}^{\infty} f_m(X - A),$$

converges in norm on the neighbourhood $\mathcal{B}(A, \delta)$ *of* A *to* f.

PROOF. The proof resembles that of Theorem 3.2.1 with obvious modifications. One only needs to apply Lemma 3.1.6 instead of Lemma 3.1.4. □

4. Inverse function theorem

As an application of the tools and techniques developed we present an inverse and implicit function theorem for free maps. For $G = \mathrm{GL}$ these results have been obtained by Pascoe [**Pas14**], Agler and McCarthy [**AM14**], Kaliuzhnyi-Verbovetskyi and Vinnikov (private communication).

Following [**KVV14**] we recall two topologies on $\mathcal{M}(F)^g$. The first is the *finitely open topology*. Its basis are open sets U such that the intersection of U with $M_n(F)^g$ is open for every $n \in \mathbb{N}$. The second topology is the *uniformly open topology* and its basis consists of sets of the form

$$\mathcal{B}(A, r) = \bigcup_{s=1}^{\infty} \{X \in M_{ns}(F)^g \mid \|X - \bigoplus_{i=1}^{s} A\| < r\},$$

for $A \in M_n(F)^g$, $n \in \mathbb{N}$, $r \geq 0$. Further topologies in this free context are considered in [**AM15, AM14**].

Let us recall a version of the classical inverse function theorem, giving information on the injectivity domain (see e.g. [**Lan93a**, Theorem XIV.1.2], [**KP02**, Theorem 2.5.1], [**KK83**, Theorem 0.8.3]). We state it only in the case when $f : \mathcal{U} \to V$ for $\mathcal{U} \subset V$, 0 is in the domain of f, $f(0) = 0$, $\mathrm{D}f(0) = \mathrm{id}_V$, to which the general case can be reduced by replacing the function $f : \mathcal{U} \to V$ with the function $\overline{f}(x) = \mathrm{D}f(x_0)^{-1}(f(x + x_0) - f(x_0))$, if x_0 is the point in the domain of f. Here D denotes the Fréchet derivative. We say that $f \in \mathcal{C}^r$ if all $\mathrm{D}^k f$, $1 \leq k \leq r$, exist and are continuous.

THEOREM 4.0.1. *Let* V *be a Banach space,* $\mathcal{U} \subset V$ *an open set containing* 0, $f : \mathcal{U} \to V$, *and let* $f \in \mathcal{C}^r$ *for some* $r \in \mathbb{N}$ (*resp.* f *is analytic*). *Let* $\mathrm{D}f(0) : V \to V$ *be a continuous bijective linear map. If* $\mathrm{Ball}(0, 2\delta) \subseteq \mathcal{U}$ *and* $\|\mathrm{D}(x - f(x))\| < \frac{1}{2}$ *for* $\|x\| < 2\delta$, *then* f *is injective on* $\mathrm{Ball}(0, \delta)$, *and there exists* $h : \mathrm{Ball}(0, \frac{\delta}{2}) \to \mathcal{V}$, *where* $\mathcal{V}$ *is an open subset of* $\mathrm{Ball}(0, \delta)$, *such that* $hf = \mathrm{id}_{\mathcal{V}}$, $fh = \mathrm{id}_{\mathrm{Ball}(0, \frac{\delta}{2})}$, *and* $h \in \mathcal{C}^r$ (*resp.* h *is analytic*).

With a slight abuse of notation, we call a g'-tuple $f = (f_1, \dots, f_{g'})$ of G-free maps $f_i : \mathcal{U} \to \mathcal{M}(F)$, also a G-free map. Throughout this section we let $G \in \{\mathrm{GL}, \mathrm{O}\}$.

4.1. Uniformly open topology. In this subsection we work with the uniformly open topology. The Fréchet derivative $\mathrm{D}f$ is continuous in the uniformly open topology at $A \in M_n(F)^g$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\mathrm{D}f(X) - \mathrm{D}f(A^{\oplus s})\| < \varepsilon$ if $s \in \mathbb{N}$ and $X \in \mathcal{B}(A, \delta)[ns]$.

THEOREM 4.1.1 (Inverse free function theorem). *Let $\mathcal{U} \subset \mathcal{M}(F)^g$ be an open G-free set containing 0, $f : \mathcal{U} \to \mathcal{M}(F)^{g'}$ a G-free map, and let $f \in \mathcal{C}^r$ for $r \in \mathbb{N}$ (resp. f analytic), with $\mathrm{D}f(0)$ invertible as a continuous linear map. Then there exist open G-free sets $\mathcal{W} \subset \mathcal{M}(F)^g$, $\mathcal{W}' \subset \mathcal{M}(F)^{g'}$ containing 0, $f(0)$ respectively, and a G-free map $h : \mathcal{W}' \to \mathcal{W}$ so that $fh = \mathrm{id}_{\mathcal{W}'}$, $hf = \mathrm{id}_{\mathcal{W}}$, and $h \in \mathcal{C}^r$ (resp. h analytic). Moreover, h is analytic for every $r \in \mathbb{N}$ in the case $G = \mathrm{GL}$.*

PROOF. Since $\mathcal{M}(F)^g$ is not a Banach space we cannot directly apply Theorem 4.0.1. However, we can use it levelwise. Without loss of generality we can assume that $g = g'$, $f(0) = 0$ and $\mathrm{D}f(0) = \mathrm{id}_{\mathcal{M}(F)^g}$ by replacing f with the function

$$\overline{f} : \mathcal{M}(F)^g \to \mathcal{M}(F)^g, \quad \overline{f} = \mathrm{D}f(0)^{-1}(f - f(0)).$$

As $\mathrm{D}f$ is continuous on $\mathcal{U}$ and invertible at 0 with a continuous inverse in the uniformly open topology, there exists (by the definition of the topology) $\delta > 0$ such that $\mathcal{B}(0, 2\delta) \subseteq \mathcal{U}$ and $\|\mathrm{D}(x - f(x))\| < \frac{1}{2}$ for $\|x\| < 2\delta$. Theorem 4.0.1 therefore implies that f is injective on $\mathcal{B}(0, \delta)$, and provides a $\mathcal{C}^r$-map $h : \mathcal{B}(0, \frac{\delta}{2}) \to \mathcal{V}$, where $\mathcal{V}$ is an open subset of $\mathcal{B}(0, \delta)$, that satisfies the desired identities.

Let us first show that $\mathcal{V}$ is an O-free set and h is an O-free map. Let $u \in \mathrm{O}_n$, $Y \in \mathcal{B}(0, \frac{\delta}{2})[n]$. As $uYu^t \in \mathcal{B}(0, \frac{\delta}{2})[n]$ and f is a G-free map we have

$$f\big(h(uYu^t)\big) = uYu^t = uf\big(h(Y)\big)u^t = f\big(uh(Y)u^t\big). \tag{4.1}$$

Since $uh(Y)u^t \subset u\mathcal{V}u^t \subset \mathcal{B}(0, \delta)$ and f is injective on $\mathcal{B}(0, \delta)$, h respects O-similarity. In the same way one can show that h respects direct sums, so it is indeed an O-free map. In consequence, $\mathcal{V} = h(\mathcal{B}(0, \frac{\delta}{2}))$ is an O-free set. Thus, in the case $G = \mathrm{O}$, the proposition follows.

It remains to consider the case $G = \mathrm{GL}$. We claim that h is analytic in this case. In the case $F = \mathbb{C}$, f is analytic (see [**HKM11**, Proposition 2.5] or [**KVV14**, Theorem 7.2]). Our assumptions imply that f is (uniformly) bounded in $\mathcal{B}(0, \delta)$, therefore we can apply [**KVV14**, Theorem 7.23, Remark 7.35] to deduce that f is analytic also in the case $F = \mathbb{R}$. Thus, h is analytic by Theorem 4.0.1. Since h is an O-free map according to the previous paragraph, it can be expanded in a power series (2.1) in x, x^t about 0 by Theorem 2.2.1, which converges in $\mathcal{B}(0, \frac{\delta}{2})$. Note that (4.1) holds also if we replace u, u^t by σ, σ^{-1} respectively, for $\sigma \in \mathrm{GL}_n$ such that $\sigma Y \sigma^{-1} \in \mathcal{B}(0, \frac{\delta}{2})$, $\sigma h(Y) \sigma^{-1} \in \mathcal{B}(0, \delta)$. Note that for every $Y \in \mathcal{B}(0, \frac{\delta}{2})$ there exists $\delta_\sigma > 0$, such that $t\sigma Y \sigma^{-1} \in \mathcal{B}(0, \frac{\delta}{2}), \sigma h(tY) \sigma^{-1} \in \mathcal{B}(0, \delta)$ for every $|t| < \delta_\sigma$. Thus,

$$h(\sigma t Y \sigma^{-1}) = \sigma h(tY) \sigma^{-1}$$

for every $|t| < \delta_\sigma$. Writing this identity as a power series in t, we can deduce that each homogeneous part h_m of the power series H of h is a GL-concomitant. Thus, H is a power series in x, and h is a GL-free map on $\mathcal{B}(0, \frac{\delta}{2})$. Now notice that the GL-similarity invariant envelopes

$$\mathcal{W} = \widetilde{\mathcal{V}}, \quad \mathcal{W}' = \widetilde{\mathcal{B}(0, \frac{\delta}{2})}$$

are open sets since the function $X \mapsto \sigma X \sigma^{-1}$ is an (analytic) isomorphism. As $\mathcal{U}$ is a G-free set, $\mathcal{W}$ is contained in $\mathcal{U}$. Furthermore, $\tilde{h}$ (cf. Proposition 1.2.2) maps $\mathcal{W}'$ to $\mathcal{W}$. Thus, we only need to check that f and $\tilde{h}$ satisfy the desired identities. Let $\tilde{X} = \sigma X \sigma^{-1} \in \mathcal{W}$, where $X \in \mathcal{V}[n], \sigma \in \mathrm{GL}_n$. Then

$$\tilde{h}\Big(f\big(\sigma X \sigma^{-1}\big)\Big) = \tilde{h}\big(\sigma f(X)\sigma^{-1}\big) = \sigma h(f(X))\sigma^{-1} = \sigma X \sigma^{-1}$$

implies that $\tilde{h}f = \mathrm{id}_{\mathcal{W}}$. The identity $f\tilde{h} = \mathrm{id}_{\mathcal{W}'}$ can be checked similarly. $\square$

The proof used in the classical setting to derive the implicit function theorem from the inverse function theorem can be also utilized in the free setting. Thus, we obtain an implicit free function theorem. We denote by $\mathrm{D}_2 f(a,b)$, where $f : \mathcal{U} \times \mathcal{V} \to \mathcal{W}$, and $(a,b) \in \mathcal{U} \times \mathcal{V}$, the Fréchet derivative of the function $y \mapsto f(a,y)$ evaluated at b.

Corollary 4.1.2 (Implicit free function theorem). *Let $\mathcal{U}_1 \times \mathcal{U}_2 \subseteq \mathcal{M}(F)^g \times \mathcal{M}(F)^{g'}$ be an open G-free set, $f : \mathcal{U}_1 \times \mathcal{U}_2 \to \mathcal{M}(F)^{g'}$ a G-free map, and let $f \in \mathcal{C}^r$ for some $r \in \mathbb{N}$, with $\mathrm{D}_2 f(0,0)$ invertible. There exist an open G-free set $\mathcal{V}_1 \times \mathcal{V}_2$ containing $(0,0)$, and a G-free map $h : \mathcal{V}_1 \to \mathcal{V}_2$, $h \in \mathcal{C}^r$, such that $f(x,y) = 0$ for $(x,y) \in \mathcal{V}_1 \times \mathcal{V}_2$ if and only if $y = h(x)$.*

We now turn our attention to the inverse function theorem about neighbourhoods of non-scalar points. Let us denote

$$C_G(A) = \{\sigma \in G_n \mid \sigma A_i = A_i \sigma,\ 1 \le i \le g\}$$

for $A = (A_1, \dots, A_g) \in M_n(F)^g$. We say that $\mathcal{U} \subset \mathcal{M}_n(F)$ is a $C_G(A) \otimes G$-free set if it is closed under direct sums and simultaneous $C_G(A) \otimes G$-similarity. By

$$\tilde{\mathrm{D}} f(A) : \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)^{g'}$$

for $f : \mathcal{U} \to \mathcal{M}_n(F)^{g'}$, $A \in \mathcal{U} \subseteq \mathcal{M}_n(F)^g$, we denote the linear map defined levelwise for every $s \in \mathbb{N}$ as

$$\tilde{\mathrm{D}} f(A)[ns](H) := \mathrm{D} f(A^{\oplus s})(H).$$

The next theorem generalizes Theorem 4.1.1 to the case of non-scalar center points.

Theorem 4.1.3. *Let $\mathcal{U} \subset \mathcal{M}(F)^g$ be an open G-free set, $A \in \mathcal{U}[n]$, $f : \mathcal{U} \to \mathcal{M}(F)^{g'}$ a G-free map, and let $f \in \mathcal{C}^r$ for $r \in \mathbb{N}$, with $\tilde{\mathrm{D}} f(A)$ invertible as a continuous linear map. There exist open $C_G(A) \otimes G$-free sets $\mathcal{W} \subset \mathcal{M}_n(F)^g$, $\mathcal{W}' \subset \mathcal{M}_n(F)^{g'}$ containing A, $f(A)$ respectively, and a $C_G(A) \otimes G$-free map $h : \mathcal{W}' \to \mathcal{W}$ so that $fh = \mathrm{id}_{\mathcal{W}'}$, $hf = \mathrm{id}_{\mathcal{W}}$, and $h \in \mathcal{C}^r$.*

Proof. Note that

$$\mathrm{D} f(\sigma X \sigma^{-1})(\sigma H \sigma^{-1}) = \sigma \mathrm{D} f(X)(H)\sigma^{-1}$$

for every $X, H \in M_n(F)^g, \sigma \in G_n, n \in \mathbb{N}$. Since $A \in \mathcal{U}$, which is an open G-free set, there exists $\delta > 0$ such that $\mathcal{B}(A,\delta) \subset \mathcal{U}$. Then the function $\overline{f} : \mathcal{B}(0,\delta) \cap \mathcal{M}_n(F)^g \to \mathcal{M}_n(F)^g$ defined by

$$\overline{f}[ns](X) := \mathrm{D} f\big(A^{\oplus s}\big)^{-1}\Big(f\big(X + A^{\oplus s}\big) - f\big(A^{\oplus s}\big)\Big)$$

is $C_G(A) \otimes G$-free with $\overline{f}(0) = 0$, $\mathrm{D}\overline{f}(0) = \mathrm{id}_{\mathcal{M}_n(F)}$. A similar reasoning to that in the proof of Theorem 4.1.1 with obvious modifications and using Theorem 3.2.1 in the place of Theorem 2.2.1 now yields the desired conclusions. $\square$

4.2. Finitely open topology. Now we state a weak form of the inverse function theorem for the finitely open topology. The Fréchet derivative $\mathrm{D}f$ is continuous in the finitely open topology if $\mathrm{D}f[n]$ is continuous for every $n \in \mathbb{N}$.

PROPOSITION 4.2.1. *Let $\mathcal{U} \subseteq \mathcal{M}(F)^g$ be an open G-free set, $f : \mathcal{U} \to \mathcal{M}(F)^{g'}$ a G-free map, and let $f \in \mathcal{C}^r$ for some $r > 0$ with $\mathrm{D}f(0)$ be invertible. There exist finitely open sets $\mathcal{W}, \mathcal{V}$, containing 0, $f(0)$ respectively, and a free* O*-concomitant map $h : \mathcal{V} \to \mathcal{W}$ such that $fh = \mathrm{id}_{\mathcal{V}}$, $hf = \mathrm{id}_{\mathcal{W}}$, and $h \in \mathcal{C}^r$. In the case $F = \mathbb{C}$, h is a a free G-concomitant map.*

PROOF. By the classical inverse function theorem we can find for every $n \in \mathbb{N}$ neighbourhoods $\mathcal{V}_n$, $\mathcal{B}(0, \delta_n)$ of 0, $f[n](0)$ respectively, such that $f[n] : \mathcal{V}_n \to \mathcal{B}(0, \delta_n)$ is a diffeomorphism with the inverse $h[n] \in \mathcal{C}^r$. Since $\mathcal{B}(0, \delta_n)$ is O_n-invariant so is $\mathcal{V}_n$ for every $n \in \mathbb{N}$. As in the proof of Theorem 4.1.1 it is easy to show that $h(uYu^t) = uh(Y)u^t$ for every $u \in \mathrm{O}_n$, $Y \in \mathcal{V}_n$. By the definition of the finitely open topology, the sets $\mathcal{V} = \bigcup_n \mathcal{V}_n$, $\mathcal{W} = \bigcup_n \mathcal{B}(0, \delta_n)$ are finitely open. This establishes the proposition in the case $G = \mathrm{O}$. In the case $G = \mathrm{GL}_n$ and $F = \mathbb{C}$ we proceed as in the proof of Theorem 4.1.1, and replace $\mathcal{V}$, $\mathcal{W}$ by $\widetilde{\mathcal{V}}$, $\widetilde{\mathcal{W}}$ respectively. To show that $f, \tilde{h}$ satisfy the required identities one also only needs to follow the steps in the proof of Theorem 4.1.1. $\square$

We do not know whether $\mathcal{W}$ and $\mathcal{V}$ in Proposition 4.2.1 can be taken to be G-free sets, and consequently h would be a G-free map; cf. [**AM14**, Section 8].

4.3. Global free inverse function theorem. In [**Pas14**, Theorem 1.1] it is proved that if f is a GL-free map and $\mathrm{D}f(X)$ is nonsingular for every $X \in \mathcal{M}(\mathbb{C})$ then f is injective, cf. [**AM14**]. This also holds for O-free maps.

PROPOSITION 4.3.1. *If $f : \mathcal{M}(F)^g \to \mathcal{M}(F)^{g'}$ is a differentiable G-free map such that $\mathrm{D}f(X)$ is nonsingular for every $X \in \mathcal{M}(F)$ then f is injective. If $f \in \mathcal{C}^r$ for some $r \in \mathbb{N}$ then there exists a G-free map $h : f(\mathcal{M}(F)^g) \to \mathcal{M}(F)^{g'}$, $h \in \mathcal{C}^r$, such that $hf = \mathrm{id}|_{\mathcal{M}(F)^g}$, $fh = \mathrm{id}|_{f(\mathcal{M}(F)^g)}$.*

PROOF. Suppose that $f(X_1) = f(X_2)$ for some $X_1, X_2 \in M_n(F)^g$. Then (2.4) yields

$$\mathrm{D}f \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix} \begin{pmatrix} 0 & X_1 - X_2 \\ X_1 - X_2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Since $\mathrm{D}f \begin{pmatrix} X_1 & 0 \\ 0 & X_2 \end{pmatrix}$ is nonsingular we have $X_1 = X_2$, which implies the injectivity of f. The proof of the existence of the free map h satisfying the required properties is the same as that of Theorem 4.1.1. $\square$

REMARK 4.3.2. We remark that a free real Jacobian conjecture can be deduced from Proposition 4.3.1. (See e.g. [**Pas14**, Theorem 1.3].)

5. Examples of O-free maps

The theory of GL-free maps is very rigid to the point that many properties are stronger than for complex analytic functions [**KVV14**,**HKM11**,**HKM12**,**Voi10**]. In contrast to this is the theory of O-free maps as we shall now demonstrate. We start by presenting the following examples:

- a continuous O-free map which is not differentiable (Example 5.0.1); more generally,
- C^k-maps which are not C^{k+1} (Example 5.0.2);
- a smooth O-free map which is not analytic (Example 5.0.3).

EXAMPLE 5.0.1. Consider the O-free map $f_m : \mathcal{M}(\mathbb{R}) \to \mathcal{M}(\mathbb{R})$ defined by

$$f_m(x) = (xx^t)^{\frac{1}{m}} \qquad \text{for some } m \geq 2.$$

It is continuous by [**CH97**, Theorem 1.1]. Note that f_m is not differentiable at 0.

EXAMPLE 5.0.2. Let $k \in \mathbb{N}$ and

$$f : \mathcal{M}(\mathbb{R}) \to \mathcal{M}(\mathbb{R}) \quad f(x) = (xx^t)^{k+\frac{1}{2}}.$$

Then f is an O-free C^k-map [**CH97**, Theorem 1.1], but is not C^{k+1}.

EXAMPLE 5.0.3. For an example of a smooth nonanalytic O-free map consider the map

$$f : \mathcal{M}(\mathbb{R}) \to \mathcal{M}(\mathbb{R}), \quad f(x) = \sum_{j=0}^{\infty} e^{-\sqrt{2^j}} \cos\big(2^j(x+x^t)\big).$$

Since $\|\cos(2^j(A+A^t))\| \leq 1$ for every $A \in \mathcal{M}(\mathbb{R})$, the power series is convergent. We show that there exist derivatives of all orders in all directions at all points of $\mathcal{M}(\mathbb{R})$, but f is not analytic. Let us show first that f is not analytic at 0. This holds already for the function $f[1] : \mathbb{R} \to \mathbb{R}$. Indeed, since

$$\limsup_{n\to\infty} \frac{|f[1]^{(n)}(0)|}{n!} \leq \limsup_{n\to\infty} \frac{e^{-\sqrt{n}} n^n}{n!} = \infty,$$

the radius of convergence of the Taylor series of $f[1]$ at 0 is 0. Consider now the ℓ-th order derivative of the function $x \mapsto \cos(kx)$ at a point $A \in M_n(\mathbb{R})$ in the direction $H \in M_n(\mathbb{R})$. We define matrices

$$A_H^\ell = \begin{pmatrix} A & H & & \\ & \ddots & \ddots & \\ & & A & H \\ & & & A \end{pmatrix} \in M_{(\ell+1)n}(\mathbb{R}).$$

Let F be an analytic function around 0 with the radius of convergence ∞. The $\ell!$-multiple of the $(1,\ell+1)$-entry of the matrix $F(A_H^\ell)$ equals the ℓ-th order derivative of F at the point A in the direction H. By [**Hig08**, Theorem 4.25] we have

$$||\cos(kA_H^\ell)|| \leq (\ell+1)n\alpha k^{\ell n},$$

where α depends only on A, for $A = A^t, H = H^t \in M_n(\mathbb{R})$. This implies that

$$\sum_{j=0}^{\infty} e^{-\sqrt{2^j}} \Big\|\delta^\ell \cos\big(2^j(A+A^t)\big)(H+H^t)\Big\| \leq (\ell+1)!n\alpha \sum_{j=0}^{\infty} e^{-\sqrt{2^j}} 2^{j\ell n} < \infty.$$

Hence the ℓ-th order derivative of f at A in the direction H exists and equals

$$\sum_{j=0}^{\infty} e^{-\sqrt{2^j}} \delta^\ell \cos\big(2^j(A+A^t)\big)(H+H^t).$$

Let $f : \mathcal{U} \to \mathcal{M}(\mathbb{C})$ be an analytic GL-free map. If f is uniformly bounded on $\mathcal{U}$ then the m-th homogeneous part of the corresponding power series is also uniformly bounded (see e.g. the last part of the proof of [**HKM12**, Proposition 2.24]). In the case of O-free maps this is no longer the case.

EXAMPLE 5.0.4. The analytic O-free map

$$x \mapsto \sin(xx^t)$$

is uniformly bounded on $\mathcal{M}(F)$, however its $(4m+2)$-th homogeneous part

$$(-1)^m \frac{1}{(2m+1)!}(xx^t)^{2m+1}$$

is not uniformly bounded.

If an analytic GL-free map $f : \mathcal{U} \to \mathcal{M}(\mathbb{C})$ is uniformly bounded then it converges uniformly on $\mathcal{U}$ by [**HKM12**, Proposition 2.24]. The proof of the uniform convergence is easily established after noticing that the homogeneous parts of f are also uniformly bounded by the same constant. As the previous example shows this does not necessarily hold for O-free maps. Here is an explicit example of a uniformly bounded analytic O-free map, which does not converge uniformly in a neighborhood of 0.

EXAMPLE 5.0.5. We provide an example of a bounded analytic O-free map, such that the corresponding power series converges uniformly on $M_n(\mathbb{R})$ for all n but does not converge uniformly on $\mathcal{M}(\mathbb{R})$. Define the homogeneous polynomials $z_{ij} = x_3^2 x_2^{i-1} x_1^{j-1} - x_2^i x_1^j$ and let

$$h_k(x_1, x_2, x_3) = S_{2k}(z_{11}, z_{22}, z_{12}, z_{33}, \dots, z_{kk}, z_{k-1,k}, z_{k+1,k+1}),$$

where S_{2k} denotes the standard polynomial of degree $2k$; i.e.,

$$S_{2k}(x_1, \dots, x_{2n}) = \sum_{\sigma \in \mathrm{Sym}(2n)} (-1)^\sigma x_{\sigma(1)} \cdots x_{\sigma(2n)}.$$

We take

$$f(x_1, x_2, x_3) = \sin\Big(\sum_{k=1}^{\infty} k!\big(h_k(x_1, x_2, x_3) + h_k(x_1, x_2, x_3)^t\big)\Big). \tag{5.1}$$

Since S_{2k} is a polynomial identity of $M_n(\mathbb{R})$ for $k \geq n$ by the Amitsur-Levitzki theorem (see e.g. [**Row80**, Theorem 1.4.1]), $f[n]$ can be defined by taking only a finite sum in the argument of sin in (5.1). Since $x \mapsto \sin(x)$ is analytic on $M_n(\mathbb{R})$, $f[n]$ is real analytic on $M_n(\mathbb{R})$. Moreover, f is uniformly bounded by 1, since the argument of sin in f is symmetric. Note that the corresponding power series $F = \sum f_m$, where f_m is homogeneous of degree m, converges uniformly on $M_n(\mathbb{R})^3$ for every n, since the sum in the argument of sin in the definition of f is finite on $M_n(\mathbb{R})^3$ and the power series corresponding to sin restricted to symmetric matrices converges uniformly.

We will now show that F does not converge uniformly on $\mathcal{M}(\mathbb{R})$. Assume for the sake of contradiction that for every $\varepsilon > 0$ there exist N and $r > 0$ such that

$$\Big\| f(X) - \sum_{m=0}^{n} f_m(X) \Big\| < \varepsilon \quad \text{for every } \|X\| < r,\ n \geq N.$$

Fix $\varepsilon < 1$ and the corresponding N and r. Take $n > N$ such that

$$n! \left(\frac{r}{2}\right)^{2n^2+3n+1} > \frac{\pi}{2}. \tag{5.2}$$

Let

$$x_1 = \sum_{i=1}^{n} e_{i,i+1}, \quad x_2 = \sum_{i=1}^{n} e_{i+1,i}, \quad x_3 = \sum_{i=1}^{n+1} e_{ii} + e_{n,n+1}$$

be elements in $M_{n+1}(\mathbb{R})$. Note that $z_{ij} = e_{ij}$ for $1 \leq i, j \leq n+1$, $i < j$. and $z_{ii} = e_{ii} + e_{n,n+1}$. For $n > 2$ we thus have

$$h_k(x_1, x_2, x_3) = 0 \quad \text{for } k \neq n, \quad h_n(x_1, x_2, x_3) = (-1)^{n-1}(n+1)e_{1,n+1},$$

where the last identity follows by the identities

$$\begin{aligned} &S_{2n}(e_{11}, e_{22}, e_{12}, e_{33}, \dots, e_{k-2,k-1}, e_{n,n+1}, e_{k-1,k}, \dots, e_{n-1,n}, e_{n+1,n+1}) \\ &\qquad = S_{2n}(e_{n,n+1}, e_{22}, e_{12}, \dots, e_{n-1,n}, e_{n-1,n}, e_{n+1,n+1}) = (-1)^{n-1} e_{1,n+1} \end{aligned}$$

for $2 \leq k \leq n+1$, and setting $e_{01} = e_{11}$. By (5.2) there is $r' < r$ such that

$$(n+1)! \left(\frac{r'}{2}\right)^{2n^2+3n+1} = \frac{\pi}{2}.$$

Letting

$$y_i = \frac{r'}{2} x_i, \quad 1 \leq i \leq 3,$$

we have $||y|| < r$ and

$$h_n(y_1, y_2, y_3) = (-1)^{n-1} \left(\frac{r'}{2}\right)^{2n^2+3n+1} (n+1) e_{1,n+1},$$

whence

$$f(y_1, y_2, y_3) = (-1)^{n-1}(e_{1,n+1} + e_{n+1,1}).$$

Note that $f_m(A_1, A_2, A_3) = 0$ for $m < \ell$ if $h_k(A_1, A_2, A_3) = 0$ for $k < \ell$. Thus,

$$\sum_{m=0}^{N} f_m(y_1, y_2, y_3) = 0$$

and

$$\left\| f(y_1, y_2, y_3) - \sum_{m=0}^{n} f_m(y_1, y_2, y_3) \right\| = 1 > \varepsilon,$$

a contradiction.

6. U-free maps

In this section we give a sample of the minor modifications needed to handle the case $G = \mathrm{U} = (\mathrm{U}_n)_{n \in \mathbb{N}}$, $F = \mathbb{C}$. The free algebra with trace with involution over $\mathbb{C}$ consists of noncommutative polynomials in the variables x_k, x_k^* over the polynomial algebra T^* in the variables $\mathrm{tr}(w)$, where $w \in \langle x, x^* \rangle /_{\underset{\sim}{\mathrm{cyc}}}$, with the involution $\mathrm{tr}(w)^* := \mathrm{tr}(w^*)$, $\alpha^* = \overline{\alpha}$ for $\alpha \in \mathbb{C}$. The evaluation map from the free algebra with involution with trace to $M_n(\mathbb{C})$ respects involution, in particular, $\mathrm{tr}(A^{w^*}) = \mathrm{tr}(A^w)^* = \overline{\mathrm{tr}(A^w)}$.

It follows from [**Pro76**, Theorem 11.2] that a polynomial map in the variables $x_{ij}^{(k)}, (x_{ij}^{(k)})^*$ is a U_n-concomitant if and only if it is a trace polynomial in the variables x_k, x_k^*, and nontrivial trace identities in the variables x_k, x_k^* of $M_n(\mathbb{C})$ first appear in the degree n. Note that functions in commutative complex variables $x_{ij}^{(k)}$ that are real analytic can be expressed as power series in the variables $x_{ij}^{(k)}, (x_{ij}^{(k)})^*$. With this observation and the previous statements the proofs of the following proposition and theorem go along the same lines as the proofs of analogous results (Proposition 2.1.1, Theorem 3.2.1) in the cases $G = \mathrm{GL}$, $G = \mathrm{O}$.

PROPOSITION 6.0.1. *Let $f : \mathcal{M}(\mathbb{C})^g \to \mathcal{M}(\mathbb{C})$ be a* U*-free map such that $f[n]$ is a polynomial map in the variables $x_{ij}^{(k)}, (x_{ij}{}^{(k)})^*$ for every $n \in \mathbb{N}$, and* $\max_n \deg f[n] = d$, *then f is a noncommutative polynomial of degree d in the variables x_k, x_k^*.*

THEOREM 6.0.2. *Let $f : \mathcal{U} \to \mathcal{M}(\mathbb{C})$ be an $\mathbb{R}$-analytic* U*-free map, and let $\mathcal{B}(A, \delta) \in \mathcal{U}$, $A \in M_n(\mathbb{C})^g$, $\delta = (\delta_s)_{s \in \mathbb{N}}$, $\delta_s > 0$ for every $s \in \mathbb{N}$. There exist $f_m \in \mathbb{C}\langle A, A^* \rangle * \mathbb{C}\langle x \rangle$ and a formal power series*

$$F(X) = \sum_{m=0}^{\infty} f_m(X - A),$$

which converges in norm in a neighbourhood $\mathcal{B}(A, \delta)$ of A such that $F(X) = f(X)$ for $X \in \mathcal{B}(A, \delta)$.

REMARK 6.0.3. If f is a U-free polynomial map (i.e., for every $n \in \mathbb{N}$, $f[n]$ is a polynomial map in $x_{ij}^{(k)}$, $1 \leq i, j, \leq n$, $1 \leq k \leq g$) of bounded degree, then f is a polynomial in the variables x_k, x_k^* by Proposition 6.0.1. However, as f is a polynomial map, it does not involve conjugate variables, so f is a polynomial in the variables x_k. This also follows from the fact that U_n is Zariski dense in GL_n. Therefore U-free $\mathbb{C}$-analytic maps are fairly close to GL-free $\mathbb{C}$-analytic maps.

CHAPTER 5

Spectrum as an invariant in Banach algebras

In this chapter we start exploring our initial problem of extracting information from some particular invariants in the context of Banach algebras. We distinguish elements in Banach algebras and maps between them through their spectral properties.

Let A be a Banach algebra. By $\sigma(x)$ and $r(x)$ we denote the spectrum and the spectral radius of $x \in A$, respectively. In Section 1 we consider the relationship between elements $a, b \in A$ that satisfy one of the following two conditions: (1) $\sigma(ax) = \sigma(bx)$ for all $x \in A$, (2) $r(ax) \leq r(bx)$ for all $x \in A$. In particular we show that (1) implies $a = b$ if A is a C^*-algebra, and (2) implies $a \in \mathbb{C}b$ if A is a prime C^*-algebra. As an application of the results concerning the conditions (1) and (2) we obtain some spectral characterizations of multiplicative maps.

In Section 2 we are concerned with derivations d and g that satisfy $\sigma(g(x)) \subseteq \sigma(d(x))$ for every $x \in A$. It turns out that in some basic situations, say if $A = \mathcal{L}(X)$, the only possibilities are that $g = d$, $g = 0$, and, if d is an inner derivation implemented by an algebraic element of degree 2, also $g = -d$. The conclusions in more complex classes of algebras are not so simple, but are of a similar spirit. A rather definite result is obtained for von Neumann algebras. In general C^*-algebras we have to make some adjustments, in particular we restrict our attention to inner derivations implemented by selfadjoint elements. We also consider a related condition $\|[b, x]\| \leq M\|[a, x]\|$ for all selfadjoint elements x from a C^*-algebra A, where $a, b \in B$ and a is normal.

In the last section we show that the derivation d with zero spectral function restricted to quasi-nilpotent elements, i.e., $\sigma(d(x)) = 0$ for every $x \in A$ with $\sigma(x) = 0$, has its range in the radical of A (and in particular globally zero spectral function) if A belongs to a special class of Banach algebras, which includes C^*-algebras, group algebras on arbitrary locally compact groups, commutative algebras, $L(X)$ for any Banach space X.

This chapter is based on [**BŠ12a**, **BMŠ12**, **ABE$^+$14**].

1. Determining elements through their spectral properties

In this section we will be mainly concerned with the following two problems. Let A be a semisimple Banach algebra with the identity. Suppose that $a, b \in A$ satisfy

$$\rho(ax) = \rho(bx) \quad \text{for all } x \in A, \tag{1.1}$$

$\rho \in \{\sigma, r\}$. What is the relationship between a and b?

1.1. The condition $\sigma(ax) = \sigma(bx)$. We start by mentioning that if $A = \mathcal{L}(X)$, the algebra of all bounded linear operators on a Banach space X, (1.1) indeed implies $a = b$. One just has to take an arbitrary rank one operator for x in (1.1), and the desired conclusion easily follows (cf. [**TL09**, Lemma 1]). In more general Banach algebras, where we do not have appropriate analogues of finite rank operators, the spectrum is not so easily tractable and more sophisticated methods are necessary.

1.1.1. *Spectral characterization of central idempotents.* We begin by recording an elementary lemma which will be needed in the proofs of Theorems 1.1.2 and 1.2.7.

LEMMA 1.1.1. *Let X be a complex vector space and let $S, T : X \to X$ be linear operators such that $S\xi \in \mathbb{C}T\xi$ for every $\xi \in X$. Then $S \in \mathbb{C}T$.*

PROOF. This lemma can be proved directly by elementary methods. On the other hand, one can apply a more general result [**BŠ99**, Theorem 2.3] which reduces the problem to an easily handled situation where both S and T have rank one. $\square$

In our first theorem we consider a variation of the condition (1.1).

THEOREM 1.1.2. *Let A be a semisimple Banach algebra. The following conditions are equivalent for $e \in A$:*

(i) $\sigma(ex) \subseteq \sigma(x) \cup \{0\}$ for all $x \in A$.
(ii) e is a central idempotent.

PROOF. (i)$\Longrightarrow$(ii). Let π be an irreducible representation of A on a Banach space X. Suppose there exists $\xi \in X$ such that ξ and $\eta = \pi(e)\xi$ are linearly independent. By Sinclair's extension of the Jacobson density theorem [**Aup91**, Corollary 4.2.6] there exists an invertible $t \in A$ such that $\pi(t)\xi = -\eta$ and $\pi(t)\eta = \xi$. Accordingly,

$$\pi\big(et^{-1}et\big)\eta = \pi(e)\pi(t)^{-1}\pi(e)\pi(t)\eta = -\eta.$$

Hence

$$-1 \in \sigma\Big(\pi\big(et^{-1}et\big)\Big) \subseteq \sigma\big(et^{-1}et\big) \subseteq \sigma\big(t^{-1}et\big) \cup \{0\} = \sigma(e) \cup \{0\} \subseteq \sigma(1) \cup \{0\} = \{0,1\},$$

a contradiction. Therefore $\pi(e)\xi \in \mathbb{C}\xi$ for every $\xi \in X$. Lemma 1.1.1 implies that there exists $\lambda \in \mathbb{C}$ such that $\pi(e) = \lambda\pi(1)$. Thus $\lambda \in \sigma\big(\pi(e)\big) \subseteq \sigma(e) \subseteq \{0,1\}$, so that $\lambda = 0$ or $\lambda = 1$. Therefore $\pi(e^2) = \pi(e)$ and also $\pi(ex - xe) = 0$ for every $x \in A$. The semisimplicity of A implies that e is an idempotent lying in the center of A.

(ii)$\Longrightarrow$(i). Take $\lambda \notin \sigma(x)$ such that $\lambda \neq 0$. Then $ex - \lambda$ has an inverse, namely

$$(ex - \lambda)^{-1} = e(x - \lambda)^{-1} - \lambda^{-1}(1 - e).$$

Therefore $\lambda \notin \sigma(ex)$. $\square$

COROLLARY 1.1.3. *Let A be a semisimple Banach algebra. If $e \in A$ is such that*

$$\sigma(ex) \cup \{0\} = \sigma(x) \cup \{0\} \quad \text{for all } x \in A,$$

then $e = 1$.

PROOF. Theorem 1.1.2 says that e is an idempotent. By taking $1 - e$ for x we obtain $e = 1$. $\square$

1.1.2. *The unit-regular element case.* An element of a ring R that can be written as the product of an idempotent and an invertible element is called a *unit-regular element.*

THEOREM 1.1.4. *Let A be a semisimple Banach algebra and let $a, b \in A$ be such that $\sigma(ax) = \sigma(bx)$ for all $x \in A$. If a is a unit-regular element, then $a = b$.*

Proof. We have $a = et$ with e an idempotent and t invertible. Replacing x by $t^{-1}x$ in $\sigma(ax) = \sigma(bx)$ we get $\sigma(ex) = \sigma(b'x)$ for all $x \in A$, where $b' = bt^{-1}$. Hence we see that with no loss of generality we may assume that $a = e$ is an idempotent. Further, in view of Corollary 1.1.3 we may also assume that $e \neq 1$.

Replacing x by $(1-e)x$ in $\sigma(ex) = \sigma(bx)$ we get $\sigma\big(b(1-e)x\big) = \{0\}$, and therefore $b(1-e) = 0$ by the semisimplicity of A. Similarly, replacing x by $x(1-e)$ we get $\sigma\big(ex(1-e)\big) = \sigma\big(bx(1-e)\big)$, hence $\sigma\big((1-e)ex\big) \cup \{0\} = \sigma\big((1-e)bx\big) \cup \{0\}$, which gives $\sigma\big((1-e)bx\big) = \{0\}$. Consequently, $(1-e)b = 0$. Together with $b(1-e) = 0$ this yields $b \in eAe$.

It is easy to see that eAe is a Banach subalgebra of A with e as an identity element, and that $\sigma_A(y) = \sigma_{eAe}(y) \cup \{0\}$ for every $y \in eAe$ (see, e.g., [**Ric60**, Theorem 1.6.15]). The condition $\sigma(exe) = \sigma(b \cdot exe)$ for every $x \in A$ can be therefore rewritten as $\sigma_{eAe}(y) \cup \{0\} = \sigma_{eAe}(by) \cup \{0\}$ for every $y \in eAe$. Since the algebra eAe is also semisimple, we infer from Corollary 1.1.3 that $b = e$. □

1.1.3. *The commutative case.*

Theorem 1.1.5. *If A is a commutative semisimple Banach algebra and $a, b \in A$ satisfy $\sigma(ax) = \sigma(bx)$ for all $x \in A$, then $a = b$.*

Proof. By the Gelfand representation theorem we may consider A as a subalgebra of $C(K)$, the algebra of all continuous functions on a compact Hausdorff space K, which separates points and contains constants. Thus its closure $\overline{A}$ with respect to the uniform norm is a uniform algebra. Since the spectrum in commutative Banach algebras is continuous [**Aup91**, Theorem 3.4.1], $\sigma(ax) = \sigma(bx)$ holds for all $x \in \overline{A}$. Therefore $a = b$ follows from [**LT07**, Lemma 3]. □

1.1.4. *The C^*-algebra case.*

Theorem 1.1.6. *If A is a C^*-algebra and $a, b \in A$ satisfy $\sigma(ax) = \sigma(bx)$ for all $x \in A$, then $a = b$.*

Proof. The proof is divided into four steps.

Claim 1. If $a = a^*$, then $b = b^*$.

On the contrary, suppose that $b - b^* \neq 0$. Take an irreducible representation π of A on a Hilbert space H such that $\pi(b - b^*) \neq 0$, i.e., $\pi(b)$ is not self-adjoint. Then there exists $\xi \in H$, $||\xi|| = 1$, such that $\alpha = \langle \pi(b)\xi, \xi\rangle \in \mathbb{C} \setminus \mathbb{R}$. Then $\eta = \pi(b)\xi - \alpha\xi$ satisfies $\langle \eta, \xi \rangle = 0$. By Kadison's transitivity theorem (see, e.g., [**Mur90**, Theorem 5.2.2]) there exists $t \in A$ such that $\pi(t)\xi = \xi$, $\pi(t)\eta = 0$, and $t = t^*$. Therefore $\pi(t)\pi(b)\pi(t)\xi = \alpha\xi$, which gives

$$\alpha \in \sigma\big(\pi(t)\pi(b)\pi(t)\big) \subseteq \sigma(tbt) \subseteq \sigma(bt^2) \cup \{0\} = \sigma(at^2) \cup \{0\} = \sigma(tat) \cup \{0\}.$$

This is a contradiction since tat is self-adjoint and so its spectrum contains only real numbers.

Claim 2. If $a = a^*$, then $ab = ba$.

Replacing x by ax in $\sigma(ax) = \sigma(bx)$ we get $\sigma(a^2x) = \sigma(bax)$ for every $x \in A$. Since a^2 is self-adjoint, Claim 1 implies that ba is self-adjoint, too. Since b is also self-adjoint by Claim 1, it follows that $ab = ba$.

Claim 3. If $a = a^*$, then $a = b$.

Claims 1 and 2 imply that the C^*-subalgebra of A generated by a and b is commutative. Since $\sigma(ax) = \sigma(bx)$ of course holds for every x from this subalgebra, there is no loss of generality in assuming that A is commutative. Thus, the desired conclusion follows by Theorem 1.1.5.

CLAIM 4. If a is arbitrary, then $a = b$.

As special cases of $\sigma(ax) = \sigma(bx)$, $x \in A$, we have $\sigma(aa^*x) = \sigma(ba^*x)$, $x \in A$, and $\sigma(ab^*x) = \sigma(bb^*x)$, $x \in A$. From Claim 3 we infer that $aa^* = ba^*$ and $ab^* = bb^*$. Accordingly, $(a - b)(a^* - b^*) = 0$, which results in $a = b$. $\square$

1.2. The condition $r(ax) \leq r(bx)$. What to expect if elements a and b from a semisimple Banach algebra A satisfy this condition? An obvious possibility is that there exists $u \in Z(A)$ such that $r(u) \leq 1$ and $a = ub$. In fact, u does not need to be central, it is enough to assume that it commutes with all elements from the right ideal bA. We shall see that, unfortunately, the possibility $a = ub$ is not the only one in general; however, in two interesting special cases it is.

1.2.1. *The invertible element case.* If b is invertible, then the solution to our problem follows immediately from Ptak's result [**Ptá78**].

THEOREM 1.2.1. *Let A be a semisimple Banach algebra and let $a, b \in A$ be such that $r(ax) \leq r(bx)$ for all $x \in A$. If b is invertible, then there exists $u \in Z(A)$ such that $r(u) \leq 1$ and $a = ub$.*

PROOF. Set $u = ab^{-1}$. Our assumption can be written as $r(ux) \leq r(x)$ for all $x \in A$. Hence $u \in Z(A)$ by [**Ptá78**, Proposition 2.1] (see also [**BBR09**, Theorem 2.2]). Letting $x = 1$ we get $r(u) \leq 1$. $\square$

1.2.2. *Remarks on the C^*-algebra case.*

LEMMA 1.2.2. *Let A be a C^*-algebra and let $a, b \in A$. The following conditions are equivalent:*

(i) $r(ax) \leq r(bx)$ *for all* $x \in A$.
(ii) $||yaz|| \leq ||ybz||$ *for all* $y, z \in A$.

PROOF. (i)$\Longrightarrow$(ii). Note that yaz and ybz satisfy the same condition as a and b, that is,

$$r(yaz \cdot x) = r(azxy) \leq r(bzxy) = r(ybz \cdot x).$$

Therefore it suffices to show that (i) implies $||a|| \leq ||b||$. And this is easy:

$$||a||^2 = r(aa^*) \leq r(ba^*) = r\big((ba^*)^*\big) = r(ab^*) \leq r(bb^*) = ||b||^2.$$

(ii)$\Longrightarrow$(i). From (ii) we infer that

$$\begin{aligned} ||(ax)^n|| &= ||axaxax \dots ax|| \leq ||bxaxax \dots ax|| \\ &\leq ||bxbxax \dots ax|| \leq \dots \leq ||bxbxbx \dots bx|| \\ &= ||(bx)^n||. \end{aligned}$$

Therefore (i) follows from the spectral radius formula. $\square$

The following simple example indicates the delicacy of our problem.

EXAMPLE 1.2.3. Let A be the commutative C^*-algebra $C[-1, 1]$, and let $a, b \in A$ be given by $a(t) = t$, $b(t) = |t|$. Then

$$r(ax) = ||ax|| = ||bx|| = r(bx) \quad \text{for all } x \in A.$$

However, there does not exist $u \in A$ such that $a = ub$.

This suggests that in order to derive $a = ub$ with $u \in Z(A)$ from (1.1) for $\rho = r$ it might be reasonable to consider C^*-algebras whose center is small. In what follows we will deal with *prime C^*-algebras*, i.e., C^*-algebras with the property that that the product of any two of their nonzero ideals is nonzero. This is a fairly large class of C^*-algebras, which includes all primitive ones. It is known that such algebras have trivial centers, i.e., scalar multiples of 1 are their only central elements. Also, it is easy to see that only these elements commute with every element from a nonzero right ideal.

1.2.3. *Tools.* In the course of the proof we will use several tools which are not standard in spectral theory. For the clarity of the exposition we will therefore state them as lemmas. The first one is of crucial importance for our goal.

LEMMA 1.2.4. *Let A be a C^*-algebra and let X be a Banach space. If $\Phi : A \times A \to X$ is a continuous bilinear map such that $\Phi(y, z) = 0$ whenever $y, z \in A$ satisfy $yz = 0$, then $\Phi(yx, z) = \Phi(y, xz)$ for all $x, y, z \in B$.*

PROOF. This result actually holds for a large class of Banach algebras which includes C^*-algebras; see [**ABEV09**, Theorem 2.11 and Example 2, p. 137]. $\square$

LEMMA 1.2.5. *Let A be a prime C^*-algebra. Suppose $a, b, c, d \in A$ satisfy $axb = cxd$ for all $x \in A$. If $a \neq 0$, then $b \in \mathbb{C}d$. Similarly, if $b \neq 0$, then $a \in \mathbb{C}c$.*

PROOF. This result is basically due to Martindale [**Mar69a**] and it actually holds for general prime rings, just that $\mathbb{C}$ must be replaced by the so-called extended centroid (a certain extension of the center). It is a fact that the extended centroid of a prime C^*-algebra is equal to $\mathbb{C}$ [**AM03**, Proposition 2.2.10]. $\square$

In the next lemma we consider a special *functional identity* which can be handled by elementary means, avoiding the general theory [**BCM07**]. At the beginning of the proof we will use an idea from [**BCM07**, Example 1.4].

LEMMA 1.2.6. *Let A be a prime C^*-algebra. Suppose there exist a map $f : A \to A$ and $c \in A$ such that*

$$f(x)yc + f(y)xc = 0 \quad \text{for all } x, y \in A.$$

If $f \neq 0$ and $c \neq 0$, then there exists a faithful irreducible representation π of A on a Hilbert space H such that $\pi(A)$ contains $K(H)$, the algebra of all compact operators on H.

PROOF. Our assumption implies that for all $x, y, z \in A$ we have

$$f(y)xczc = -f(x)yczc = f(ycz)xc.$$

Fixing $y \in A$ such that $f(y) \neq 0$ we infer from Lemma 1.2.5 that for every $z \in A$ there exists $\lambda_z \in \mathbb{C}$ such that $czc = \lambda_z c$. Consequently, $(c^*c)^2 = \alpha c^*c$ for some $\alpha \in \mathbb{R} \setminus \{0\}$. Note that $e = \alpha^{-1}c^*c$ satisfies $e^2 = e = e^*$ and $eAe = \mathbb{C}e$, so we may identify eAe with $\mathbb{C}$. We endow Ae with an inner product $\langle ae, be \rangle = eb^*ae$. Note that the inner product norm coincides with the original norm on Ae, and is therefore complete. We denote the corresponding Hilbert space by H. Define $\pi : A \to B(H)$ according to $\pi(a)\xi = a\xi$, $a \in A$, $\xi \in H$, and note that π is an irreducible representation of A on H. Moreover, a faithful one since A is prime. Since $\pi(e)$ is a rank one operator, it follows that $K(H) \subseteq \pi(A)$ (see, e.g., [**Mur90**, Theorem 2.4.9]). $\square$

1.2.4. *The prime C^*-algebra case.* We now have enough information to prove the main result of this subsection.

THEOREM 1.2.7. *Let A be a prime C^*-algebra and let $a, b \in A$ be such that $r(ax) \leq r(bx)$ for all $x \in A$. Then there exists $\lambda \in \mathbb{C}$ such that $|\lambda| \leq 1$ and $a = \lambda b$.*

PROOF. Obviously it suffices to prove that $a \in \mathbb{C}b$. We divide the proof into four steps.

CLAIM 1. If $b = b^*$, then $ab = ba$.

Let B be the C^*-algebra generated by b. Define $\Phi : B \times B \to A$ by $\Phi(y, z) = yaz$. Since B is commutative, $yz = 0$ implies $ybz = 0$. According to Lemma 1.2.2 this further gives $\Phi(y, z) = 0$. Lemma 1.2.4 therefore tells us that $\Phi(yx, z) = \Phi(y, xz)$ for all $x, y, z \in B$. Setting $y = z = 1$ and $x = b$ we get $ab = ba$.

Claim 2. If $f : A \to A$ is defined by $f(x) = axb^*b - bxb^*a$, then $f(x)yb^* + f(y)xb^* = 0$ for all $x, y \in A$.

Take a self-adjoint $s \in A$. Substituting sb^*x for x in $r(ax) \le r(bx)$ we get $r(asb^*x) \le r(bsb^*x)$ for every $x \in A$. Since bsb^* is self-adjoint, Claim 1 implies that $(asb^*)(bsb^*) = (bsb^*)(asb^*)$, i.e., $f(s)sb^* = 0$ holds for an arbitrary self-adjoint $s \in A$. Replacing s by $s+t$ with both s, t self-adjoint it follows that $f(s)tb^* + f(t)sb^* = 0$. Since every element in A is a linear combination of two self-adjoint elements, the desired conclusion follows.

Claim 3. If $f \ne 0$, then $a \in \mathbb{C}b$.

Lemma 1.2.6 says that there exists a faithful representation π of A on a Hilbert space H such that $K(H) \subseteq \pi(A)$. By $\xi \otimes \eta$ we denote the rank one operator given by $(\xi \otimes \eta)\omega = \langle \omega, \eta \rangle \xi$. Note that $\sigma_{B(H)}(\xi \otimes \eta) = \{0, \langle \xi, \eta \rangle\}$ and that $A(\xi \otimes \eta) = A\xi \otimes \eta$ for every $A \in B(H)$. Of course, $\xi \otimes \eta \in \pi(A)$, and hence

$$r\big(\pi(a)(\xi \otimes \eta)\big) \le r\big(\pi(b)(\xi \otimes \eta)\big).$$

That is,

$$|\langle \pi(a)\xi, \eta \rangle| \le |\langle \pi(b)\xi, \eta \rangle|,$$

where ξ and η are arbitrary vectors in H. If $\pi(a)\xi$ was not a scalar multiple of $\pi(b)\xi$, then we could find η such that $\langle \pi(a)\xi, \eta \rangle \ne 0$ and $\langle \pi(b)\xi, \eta \rangle = 0$ – a contradiction. Therefore $\pi(a)\xi \in \mathbb{C}\pi(b)\xi$ for every $\xi \in H$, hence $\pi(a) \in \mathbb{C}\pi(b)$ by Lemma 1.1.1, and so $a \in \mathbb{C}b$.

Claim 4. If $f = 0$, then $a \in \mathbb{C}b$.

The result is trivial if $b = 0$, so let $b \ne 0$. We are assuming that $axb^*b = bxb^*a$ holds for every $x \in A$. Since $b^*b \ne 0$, we have $a \in \mathbb{C}b$ by Lemma 1.2.5. ☐

Corollary 1.2.8. *Let A be a prime C^*-algebra and let $a, b \in A$ be such that $r(ax) = r(bx)$ for all $x \in A$. Then there exists $\lambda \in \mathbb{C}$ such that $|\lambda| = 1$ and $a = \lambda b$.*

Remark 1.2.9. In (1.1) one can replace "$=$" by "$\subseteq$" for $\rho = \sigma$ and study the relationship between a and b in that setting. This problem was addressed in [**ABE$^+$13**], where also the result regarding the condition $r(ax) \le r(bx)$ was extended to not necessarily prime C^*-algebras. Since the generalizations make problems technically quite more involved we only state the results.

Theorem 1.2.10. *Let A be a unital C^*-algebra and let $a, b \in A$. Then*

(1) *$\sigma(ax) \subseteq \sigma(bx) \cup \{0\}$ for every $x \in A$ if and only if there exists a central projection $z \in A''$ such that $a = zb$,*
(2) *$r(ax) \le r(bx)$ for every $x \in A$ if and only if there exists $z \in Z(A'')$ such that $a = zb$ and $\|z\| \le 1$.*

1.3. Spectral characterizations of multiplicative maps.

1.3.1. *The condition $\sigma\big(\varphi(x)\varphi(y)\varphi(z)\big) = \sigma(xyz)$.* We begin with an application of Theorem 1.1.4.

Corollary 1.3.1. *Let B and A be Banach algebras with A semisimple. Let $\varphi : B \to A$ be a surjective map satisfying $\sigma\big(\varphi(x)\varphi(y)\varphi(z)\big) = \sigma(xyz)$ for all $x, y, z \in B$. Then $\varphi(1) \in Z(A)$, $\varphi(1)^3 = 1$, and $\varphi(xy) = \varphi(1)^2\varphi(x)\varphi(y)$ for all invertible $x, y \in B$.*

Proof. Set $u = \varphi(1)$. Taking $x = y = z = 1$ we get $\sigma(u^3) = \{1\}$. In particular, u is invertible. Next we have

$$\sigma\big(u\varphi(y)\varphi(z)\big) = \sigma(1yz) = \sigma(y1z) = \sigma\big(\varphi(y)u\varphi(z)\big)$$

for all $x, y \in B$. From Theorem 1.1.4 it follows that $\varphi(y)u = u\varphi(y)$ whenever $\varphi(y)$ is invertible. That is, u commutes with all invertible elements in A, and is therefore contained in $Z(A)$. Hence u^3 also belongs to $Z(A)$, and so $\sigma(u^3) = \{1\}$ implies that $u^3 = 1$.

From $\sigma(u^2\varphi(y)) = \sigma(y)$ we see that $\varphi(y)$ is invertible whenever y is invertible. Take invertible $x, y \in B$. Applying Theorem 1.1.4 to

$$\sigma\big(\varphi(x)\varphi(y)\varphi(z)\big) = \sigma(xyz) = \sigma\big(1(xy)z\big) = \sigma\big(u\varphi(xy)\varphi(z)\big)$$

we thus get $\varphi(x)\varphi(y) = u\varphi(xy)$. □

Adding the assumption that φ is linear we get a definite conclusion.

COROLLARY 1.3.2. *Let B and A be Banach algebras with A semisimple. Let $\varphi : B \to A$ be a surjective linear map satisfying $\sigma\big(\varphi(x)\varphi(y)\varphi(z)\big) = \sigma(xyz)$ for all $x, y, z \in B$. Then $\varphi(1) \in Z(A)$, $\varphi(1)^3 = 1$, and $\varphi(xy) = \varphi(1)^2\varphi(x)\varphi(y)$ for all $x, y \in B$.*

PROOF. If $x \in B$ is arbitrary, then $x - \lambda 1$ is invertible for some $\lambda \in \mathbb{C}$, and so $\varphi\big((x - \lambda 1)y\big) = \varphi(1)^2\varphi(x - \lambda 1)\varphi(y)$ for every invertible y. As φ is linear this clearly yields $\varphi(xy) = \varphi(1)^2\varphi(x)\varphi(y)$. A similar argument shows that the same is true if y is not invertible. □

In the C^*-algebra case we do not need to assume the linearity, which brings us closer to Molnar's results [**Mol02**].

COROLLARY 1.3.3. *Let B be a Banach algebra and A be a C^*-algebra. Let $\varphi : B \to A$ be a surjective map satisfying $\sigma\big(\varphi(x)\varphi(y)\varphi(z)\big) = \sigma(xyz)$ for all $x, y, z \in B$. Then $\varphi(1) \in Z(A)$, $\varphi(1)^3 = 1$, and $\varphi(xy) = \varphi(1)^2\varphi(x)\varphi(y)$ for all $x, y \in B$.*

PROOF. The same argument as in the proof of Corollary 1.3.1 works, except that at the end we may take arbitrary x and y and then apply Theorem 1.1.6 instead of Theorem 1.1.4. □

The conclusion can be read as that the map $x \mapsto \varphi(1)^2\varphi(x)$ is multiplicative. We remark that multiplicative maps on rings often turn out to be automatically additive [**Mar69b**, **Ric48**]; say, this is true in prime rings having nontrivial idempotents. Accordingly, by adding some assumptions to Corollary 1.3.3 one can get a more complete result. See also [**Mol02**].

1.3.2. *The condition $r\big(\varphi(x)\varphi(y)\varphi(z)\big) = r(xyz)$.* We conclude this subsection with a corollary to Theorem 1.2.7.

COROLLARY 1.3.4. *Let B be a Banach algebra and A be a prime C^*-algebra. Let $\varphi : B \to A$ be a surjective map satisfying $r\big(\varphi(x)\varphi(y)\varphi(z)\big) = r(xyz)$ for all $x, y, z \in B$. Then for each pair $x, y \in B$ there exists $\lambda(x, y) \in \mathbb{C}$ such that $|\lambda(x, y)| = 1$ and $\varphi(xy) = \lambda(x, y)\varphi(x)\varphi(y)$.*

PROOF. We argue similarly as in the proof of Corollary 1.3.1. Set $u = \varphi(1)$. For all $y, z \in B$ we have

$$r\big(u\varphi(y)\varphi(z)\big) = r(1yz) = r(y1z) = r\big(\varphi(y)u\varphi(z)\big)$$

Corollary 1.2.8 tells us that $u\varphi(y)$ and $\varphi(y)u$ are equal up to a scalar factor of modulus 1. Thus, for every $x \in A$ there exists $\mu_x \in \mathbb{C}$ such that $|\mu_x| = 1$ and $ux = \mu_x xu$. Hence

$$\mu_{x+1}(x + 1)u = u(x + 1) = ux + u = \mu_x xu + u$$

for every $x \in A$. That is,

$$(\mu_{x+1} - \mu_x)xu = (1 - \mu_{x+1})u.$$

Therefore either $xu \in \mathbb{C}u$ or $\mu_{x+1} = 1$, i.e., $ux = xu$. In each of the two cases we have $uxu = xu^2$. Lemma 1.2.5 implies that $u \in \mathbb{C}$ (and so we can actually take $\mu_x = 1$ for every $x \in A$). From $r(u^3) = 1$ we see that $|u| = 1$. Finally, we have

$$r\big(\varphi(x)\varphi(y)\varphi(z)\big) = r(xyz) = r\big(1(xy)z\big) = r\big(u\varphi(xy)\varphi(z)\big)$$

and so the desired conclusion follows from Corollary 1.2.8. □

2. Identifying derivations through the spectra of their values

Following the previous subsection we continue by investigating the determination of maps by their spectral properties. We restrict to derivations and consider an analogue of (1.1). We are interested in the relationship between derivations d, g satifying

$$\sigma(g(x)) \subseteq \sigma(d(x)) \quad \text{for all } x \in A. \tag{2.1}$$

Also in this section we assume for simplicity that A has the identity.

2.1. Commutators with symmetric spectra. If an element a from a Banach algebra A is such that for every $x \in A$ the spectrum of $[a, x]$ is symmetric in the sense that $\sigma\big([a, x]\big) = -\sigma([a, x])$, then the spectral inclusion condition (2.1) is fulfilled for $g = -d$, $d(\,\cdot\,) = [a, \,\cdot\,]$. This situation is of special interest. Let us show that it can indeed occur.

Lemma 2.1.1. *Let A be a Banach algebra. If $e \in A$ is an idempotent and $x \in A$ is arbitrary, then $[e, x]$ is similar to $-[e, x]$. In particular, $\sigma([e, x]) = -\sigma([e, x])$.*

Proof. Take $s = 1 - 2e$. Then $s = s^{-1}$ and $s[e, x]s^{-1} = -[e, x]$. □

Lemma 2.1.2. *Let A be a Banach algebra in which for all $x, y \in A$, $xy = 1$ implies $yx = 1$. If $a \in A$ is algebraic of degree 2, then $\sigma([a, x]) = -\sigma([a, x])$ for all $x \in A$.*

Proof. Replacing a by $\alpha a + \beta$, where α and β are suitable scalars, the proof reduces to two cases: a is an idempotent and a is nilpotent of nilpotency degree 2. As the first one has already been treated in Lemma 2.1.1, we may suppose that $a^2 = 0$. It suffices to prove that $-1 \in \sigma([a, x])$ implies $1 \in \sigma([a, x])$ for every $x \in A$. We have

$$(1 + xa)(1 - ax) = 1 - [a, x],$$

and

$$-(1 - xa)(1 + ax) = -1 - [a, x].$$

If $1 - [a, x]$ is invertible, then, by our assumption, $1 + xa$ and $1 - ax$ are invertible. Consequently, (as it is well-known) $1 + ax$ and $1 - xa$ are invertible. By the above equality this implies the invertibility of $-1 - [a, x]$. □

Let us show that the assumption that $xy = 1$ implies $yx = 1$ is not redundant. We give an example of an element $a \in B(H)$ such that $a^2 = 0$ and there exists $x \in B(H)$ such that $\sigma([a, x])$ is not symmetric.

Example 2.1.3. Let $a, x \in B(H) \cong B(H \bigoplus H) \cong M_2(B(H))$ be of the form

$$a = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad x = \begin{pmatrix} y & 0 \\ u & 0 \end{pmatrix}.$$

Then

$$[a, x] = \begin{pmatrix} u & -y \\ 0 & -u \end{pmatrix},$$

and hence

$$1 + [a, x] = \begin{pmatrix} 1 + u & -y \\ 0 & 1 - u \end{pmatrix}, \quad 1 - [a, x] = \begin{pmatrix} 1 - u & y \\ 0 & 1 + u \end{pmatrix}.$$

It suffices to find u such that $1-u$ is right invertible, $1+u$ is left invertible and $\{0\} \neq \ker(1-u)$ is isomorphic to $\mathrm{im}(1+u)^{\perp}$. For then we take for y an operator which is an isomorphism from $\ker(1-u)$ onto $\mathrm{im}(1+u)^{\perp}$ and is zero on $\ker(1-u)^{\perp}$. Then $1+[a,x]$ is bijective (hence invertible), but $1-[a,x]$ is not injective since $\ker(1-u) \neq \{0\}$. One possible choice for u is $\begin{pmatrix} 1+s^* & 0 \\ 0 & -1+s \end{pmatrix}$, where s is the unilateral shift and s^* its adjoint. Indeed, $1-u = \begin{pmatrix} -s^* & 0 \\ 0 & 2-s \end{pmatrix}$ is right invertible (since s^* is right invertible and $2-s$ is invertible), $1+u = \begin{pmatrix} 2+s^* & 0 \\ 0 & s \end{pmatrix}$ is left invertible and $\ker(1-u) \cong \ker s^* = \mathrm{im} s^{\perp} \cong \mathrm{im}(1+u)^{\perp}$.

Roughly speaking, we have shown that algebraic elements of degree 2 often generate derivations whose values have symmetric spectra. In Lemma 2.2.2, we shall see that these are also the only natural examples of such elements.

2.2. Derivations on Banach algebras. We begin our discussion on the spectral inclusion condition (2.1) in algebras in which the spectrum can be easily computed at least for some elements. The prototype example we have in mind is $\mathcal{L}(X)$, the algebra of all bounded operators on a Banach space X. In this case the operators of finite rank have an easily approachable spectrum. Actually, we will work in a slightly more general setting of primitive Banach algebras with nonzero socle and we will replace (2.1) with a technically weaker condition $\sigma(g(x)) \subseteq \sigma(d(x)) \cup \{0\}$. This will be needed for further applications.

Recall that a semiprime Banach algebra A is said to have a nonzero socle if A has minimal one-sided ideals. In this case the socle of A, $\mathrm{soc}(A)$, is defined as the sum of all minimal left ideals of A (equivalently, the sum of all minimal right ideals of A). The socle has a particularly important role if A is a primitive algebra.

THEOREM 2.2.1. *Let A be a primitive Banach algebra with nonzero socle. If derivations $d, g : A \to B$ satisfy $\sigma(g(x)) \subseteq \sigma(d(x)) \cup \{0\}$ for all $x \in \mathrm{soc}(A)$, then $g = \lambda d$ with $\lambda \in \{-1, 0, 1\}$.*

PROOF. Let us recall some standard facts about the structure of A (see, e.g., [**BD73**, Section 31]). There exists an idempotent e such that $eAe = \mathbb{C}e$. We may identify eAe with $\mathbb{C}$. Denote the regular representation of A on Ae by π, which is faithful since A is primitive. If we define $\langle x, y\rangle = yx$ for $x \in Ae, y \in eA$, then $x \mapsto \langle x, v\rangle$ is a linear functional on Ae for every $v \in eA$. Write $X = Ae$ and $Y = \{f \in X^* : f(\,\cdot\,) = \langle\,\cdot\,, v\rangle \text{ for some } v \in eA\}$. If $\xi_1, \ldots, \xi_n \in X$ are linearly independent, there exists $f \in Y$ such that $f(\xi_1) = 1$ and $f(\xi_i) = 0$ for all $i > 1$. All operators of the form $\sum_{i=1}^n \xi_i \otimes f_i$ with $\xi_i \in X, f_i \in Y$ are contained in $\pi(\mathrm{soc}(A))$. We write $\{\xi\}^{\perp} = \{f \in Y : f(\xi) = 0\}$ for $\xi \in X$.

We may and we shall identify A with $\pi(A)$. Every derivation $\widetilde{d}$ on A is of the form $[\widetilde{a}, \cdot\,]$ for some $\widetilde{a} \in \mathcal{L}(X)$. The proof is the same as for $\mathcal{L}(X)$. (One just defines $\widetilde{a} : \xi \mapsto \widetilde{d}(\xi \otimes f)\eta$ for some $\eta \in X$, $f \in Y$ with $f(\eta) = 1$ and details can be easily verified.) Hence there exist $a, b \in \mathcal{L}(X)$ such that $d(\,\cdot\,) = [a, \cdot\,]$ and $g(\,\cdot\,) = [b, \cdot\,]$. We take $\xi \in X$, $f \in \{\xi\}^{\perp}$ and calculate the spectra of $[a, \xi \otimes f]$ and $[b, \xi \otimes f]$. These are operators of rank at most 2, and the only nonzero elements of their spectra appear as eigenvalues of their restrictions $\bar{a}, \bar{b}$ to the vector spaces $\mathrm{span}\{\xi, a\xi\}$ and $\mathrm{span}\{\xi, b\xi\}$, respectively. Suppose that $f(b\xi) \neq 0$. Then $-f(b\xi)$ is a nonzero eigenvalue of $\bar{b}$ (corresponding to the eigenvector ξ). Since $\bar{b}$ has trace zero, its eigenvalues are $-f(b\xi)$ and $f(b\xi)$. By the hypothesis of the theorem, $\bar{a}$ is nonsingular with nonzero eigenvalues $-f(a\xi)$ and $f(a\xi)$, and $\{-f(a\xi), f(a\xi)\} = \{-f(b\xi), f(b\xi)\}$. Hence $f(a\xi) = \pm f(b\xi)$.

Therefore, for all $\xi \in X, f \in \{\xi\}^\perp$ one of the following possibilities holds: $f((a-b)\xi) = 0$ or $f((a+b)\xi) = 0$ or $f(b\xi) = 0$. We shall have established the theorem if we prove that the same possibility holds for all $\xi \in X, f \in \{\xi\}^\perp$. Indeed, if $c \in L(X)$ has the property that $f(c\xi) = 0$ for all $\xi \in X, f \in \{\xi\}^\perp$, then for all $\xi \in X$, ξ and $c\xi$ are linearly dependent, which easily implies $c \in \mathbb{C}1$.

We first fix ξ and define

$$\begin{aligned} X_\xi^0 &= \{f \in \{\xi\}^\perp : f(b\xi) = 0\}, \\ X_\xi^- &= \{f \in \{\xi\}^\perp : f((a-b)\xi) = 0\}, \\ X_\xi^+ &= \{f \in \{\xi\}^\perp : f((a+b)\xi) = 0\}. \end{aligned}$$

From what has already been proved it follows that $\{\xi\}^\perp = X_\xi^0 \cup X_\xi^- \cup X_\xi^+$. Since X_ξ^0, X_ξ^- and X_ξ^+ are vector spaces, we may conclude that one of them equals $\{\xi\}^\perp$.

Let

$$\begin{aligned} X^0 &= \{\xi \in X : \{\xi\}^\perp = X_\xi^0\} = \{\xi \in X : b\xi \in \mathbb{C}\xi\}, \\ X^- &= \{\xi \in X : \{\xi\}^\perp = X_\xi^-\} = \{\xi \in X : (a-b)\xi \in \mathbb{C}\xi\}, \\ X^+ &= \{\xi \in X : \{\xi\}^\perp = X_\xi^+\} = \{\xi \in X : (a+b)\xi \in \mathbb{C}\xi\}. \end{aligned}$$

Then X^0, X^-, X^+ are closed sets with union X, therefore at least one of them, say X^δ with $\delta \in \{0, -, +\}$, has a nonempty interior by Baire's theorem. Let us prove that X^δ is a vector space. Since X^δ is closed under scalar multiplication, it suffices to show that $\xi + \xi' \in X^\delta$ for linearly independent $\xi, \xi' \in X^\delta$. As X^δ has a nonempty interior, there exists $\alpha \in \mathbb{C} \setminus \{0\}$ such that $\xi + \alpha\xi' \in X^\delta$. Then $(\delta a + b)(\xi + \alpha\xi') = \lambda_\alpha(\xi + \alpha\xi')$ for some $\lambda_\alpha \in \mathbb{C}$. As $\xi, \xi' \in X^\delta$, we have $(\delta a + b)(\xi + \alpha\xi') = (\delta a + b)\xi + (\delta a + b)\alpha\xi' = \lambda_\xi\xi + \alpha\lambda_{\xi'}\xi'$ for some $\lambda_\xi, \lambda_{\xi'} \in \mathbb{C}$. The linear independence of ξ, ξ' now gives $\lambda_\xi = \lambda_\alpha = \lambda_{\xi'}$, which implies that $\xi + \xi' \in X^\delta$. Hence X^δ is a vector subspace of X with nonempty interior and thus $X^\delta = X$, from which we easily deduce that $\delta a + b \in \mathbb{C}1$. $\square$

The following lemma shows that in the present context we have a kind of a converse to the results from the preceding subsection.

LEMMA 2.2.2. *Let A be a primitive Banach algebra with nonzero socle. If a derivation $d : A \to B$ satisfies $\sigma(d(x)) = -\sigma(d(x))$ for all $x \in A$, then d is implemented by an algebraic element of degree 2.*

PROOF. We adopt the notation of the preceding theorem. Let $d(\cdot) = [a, \cdot]$. Assume that a is not algebraic of degree 2. By Kaplansky's theorem (see e.g. [**Aup91**, Theorem 4.2.7]), we can find $\xi \in X$ such that $\xi, a\xi, a^2\xi$ are linearly independent. Take the element $\xi \otimes f + a\xi \otimes g$ where $f \in Y$ satisfies $f(\xi) \neq 0, f(a\xi) = 0, f(a^2\xi) \neq 0$ and $g \in Y$ satisfies $g(\xi) = 0, g(a\xi) = 0, g(a^2\xi) \neq 0$. Then the range of $[a, \xi \otimes f + a\xi \otimes g]$ is contained in the vector subspace of X spanned by the linearly independent vectors $\xi, a\xi, a^2\xi$. An easy computation shows that the restriction of $[a, \xi \otimes f + a\xi \otimes g]$ to this subspace is invertible. Hence it has three nonzero eigenvalues (counted by their multiplicity) whose sum is zero. Therefore $\sigma([a, \xi \otimes f + a\xi \otimes g])$ is not equal to $-\sigma([a, \xi \otimes f + a\xi \otimes g])$. $\square$

COROLLARY 2.2.3. *Let $a \in M_n(\mathbb{C})$. Then there exists a nonscalar $b \in M_n(\mathbb{C})$ such that $b \notin a + \mathbb{C}1$ and $\sigma([b, x]) \subseteq \sigma([a, x])$ for every $x \in M_n(\mathbb{C})$ if and only if a is algebraic of degree 2. In this case $b \in -a + \mathbb{C}1$.*

PROOF. Apply Lemma 2.1.2, Theorem 2.2.1, and Lemma 2.2.2. $\square$

We will need Theorem 2.2.1 in Section 4. Its first application, however, concerns derivations with the property that their values have a finite spectrum. Such derivations have been studied in a series of papers, started in [**BŠ96**] and ended in [**BŠ10**].

THEOREM 2.2.4. *Let d, g be derivations of a semisimple Banach algebra A. Suppose $\sigma(d(x))$ is finite for every $x \in A$. If $\sigma(g(x)) \subseteq \sigma(d(x))$ for every $x \in A$, then there exist derivations d_0, d_1, d_2 of A such that $d = d_0 + d_1 + d_2$, $g = d_1 - d_2$, and $d_i(B)Bd_j(B) = 0$ for $i \neq j$.*

PROOF. First we give an extraction from the aforementioned papers. It consists of the main results together with some details that are apparently not explicitly formulated in any of the papers from the series, but are evident from the proofs.

By [**BM04**, Theorem 2.4] there exist $a, b \in \operatorname{soc}(A)$ such that $d(\,\cdot\,) = [a, \cdot\,]$ and $g(\,\cdot\,) = [b, \cdot\,]$. Accordingly, each of $d(A)$ and $g(A)$ is contained in all but finitely many primitive ideals of A [**BŠ96**, Proposition 2.2]. If $d = 0$, then there are no such primitive ideals for $d(A)$. However, in this case g has only quasinilpotent values and hence it is 0 (see, e.g., [**TS87**]). We may therefore assume that $d \neq 0$. On the other hand, $g \neq 0$ can be assumed without loss of generality. Let $P_1, \ldots, P_m$ be the only primitive ideals such that $d(A) \not\subseteq P_i$, $i = 1, \ldots, m$, and similarly, let $Q_1, \ldots, Q_n$ be the only primitive ideals such that $g(A) \not\subseteq Q_j$, $j = 1, \ldots, n$. As noticed in the proof of [**BŠ10**, Theorem 2.1], we have $P_i \not\subseteq P_{i'}$ for all $i \neq i'$ and $Q_j \not\subseteq Q_{j'}$ for all $j \neq j'$. Therefore the proof of [**BM04**, Theorem 2.4] shows that there exist $a_1, \ldots, a_m \in A$ and $b_1, \ldots, b_n \in A$ such that

- $a = a_1 + \ldots + a_m$ and $b = b_1 + \ldots + b_n$.
- $d(x) + P_i = [a_i, x] + P_i$ and $g(x) + Q_j = [b_j, x] + Q_j$ for all $x \in A$.
- $a_i + P_i \in \operatorname{soc}(A/P_i)$ and $b_j + Q_j \in \operatorname{soc}(A/Q_j)$.
- $a_i \in \bigcap_{P \neq P_i} P$ and $b_j \in \bigcap_{P \neq Q_j} P$.

(Here, the intersection runs over primitive ideals P of A.) Note that each $a_i \neq 0$ and each $b_j \neq 0$. Therefore $\bigcap_{P \neq P_i} P$ and $\bigcap_{P \neq Q_j} P$ are nonzero ideals. Since $\left(\bigcap_{P \neq P_i} P\right) \bigcap P_i = 0$ and $\left(\bigcap_{P \neq Q_j} P\right) \bigcap Q_j = 0$ by the semisimplicity of A, it follows that $\bigcap_{P \neq P_i} P \not\subseteq P_i$ and $\bigcap_{P \neq Q_j} P \not\subseteq Q_j$.

We claim that $\{Q_1, \ldots, Q_n\} \subseteq \{P_1, \ldots, P_m\}$. Suppose that, say, Q_1 is none of the ideals P_i. Thus, $d(A) \subseteq Q_1$ and $g(A) \not\subseteq Q_1$. Set $I = \bigcap_{P \neq Q_1} P$ and take $x \in I$. In particular, x is contained in every P_i, hence $d(x)$ is contained in every P_i, and so $d(x)$ is actually contained in every primitive ideal of A. Therefore $d(x) = 0$, and, consequently, $\sigma(g(x)) = \{0\}$. Thus, the restriction of $g(\,\cdot\,) = [b, \cdot\,]$ to I is a continuous derivation of I with quasinilpotent values. As I is a closed ideal of a semisimple Banach algebra, it follows that $g(I) = 0$ [**TS87**]. Accordingly, for all $x \in A$ and $u \in I$ we have $g(x)u = g(xu) - xg(u) = 0 \in Q_1$. Since Q_1 is, in particular, a prime ideal, and since $g(x) \notin Q_1$ for some $x \in A$, we must have $I \subseteq Q_1$. However, at the end of the preceding paragraph we have shown that this is not true. Our claim is thus proved. Therefore $n \leq m$ and we may assume that

$$Q_1 = P_1,\ Q_2 = P_2, \ldots,\ Q_n = P_n.$$

Recall that $\sigma(y) = \bigcup_P \sigma(y + P)$ for every $y \in A$ [**Ric60**, Theorem 2.2.9]. Let $i \leq n$, pick $x \in \bigcap_{P \neq P_i} P$, and take $g(x)$ for y. Then we obtain $\sigma(g(x)) \cup \{0\} = \sigma(g(x)+P_i) \cup \{0\} = \sigma([b_i, x]+P_i) \cup \{0\}$. Similarly, $\sigma(d(x)) \cup \{0\} = \sigma([a_i, x]+P_i) \cup \{0\}$. Using the assumption of the theorem we thus have

$$\sigma([b_i, x] + P_i) \subseteq \sigma([b_i, x] + P_i) \cup \{0\} \subseteq \sigma([a_i, x] + P_i) \cup \{0\}$$

for all $x \in \bigcap_{P \neq P_i} P$. Since $\bigcap_{P \neq P_i} P$ is an ideal which is not contained in P_i, $\left(\bigcap_{P \neq P_i} P + P_i\right)/P_i$ is a nonzero ideal of A/P_i. It is well-known that the socle of a primitive algebra is contained in every other nonzero ideal. Therefore $\sigma([b_i +$

$P_i, y]) \subseteq \sigma([a_i + P_i, y]) \cup \{0\}$ holds for all $y \in \operatorname{soc}(A/P_i)$. This enables us to apply Theorem 2.2.1. Hence we conclude that $b_i + P_i = \lambda_i a_i + \mu_i 1 + P_i$ for some $\lambda_i \in \{-1, 1\}$ and $\mu_i \in \mathbb{C}$. Note that the case $\lambda_i = 0$ can not occur for $g(A) \not\subseteq P_i$. Now define $\tilde{a}_1$ as the sum of all a_i such that $i \leq n$ and $\lambda_i = 1$, $\tilde{a}_2$ as the sum of all a_i such that $i \leq n$ and $\lambda_i = -1$, and $\tilde{a}_0$ as the sum of all a_i such that $i > n$ (the sum over the empty set of indices should be read as 0). Setting $d_0(\,\cdot\,) = [\tilde{a}_0, \cdot\,]$, $d_1(\,\cdot\,) = [\tilde{a}_1, \cdot\,]$, and $d_2(\,\cdot\,) = [\tilde{a}_2, \cdot\,]$ we have $d = d_0 + d_1 + d_2$ and $g = d_1 - d_2$. Note that for any pair of different indices k and l we have $[a_k, A]A[a_l, A] \subseteq \big(\bigcap_{P \neq P_k} P\big) \bigcap \big(\bigcap_{P \neq P_l} P\big) = 0$. This clearly implies that $d_i(A)Ad_j(A) = 0$ if $i \neq j$. □

So far we have relied heavily on finite rank operators. In general Banach algebras the spectrum of a value of a derivation may not be so easily tractable. The next lemma reduces the treatment of the spectral inclusion condition (2.1) to another problem which may be of independent interest. It will play a fundamental role in the next subsection and in the first result of the last subsection.

Lemma 2.2.5. *Let A be a semisimple Banach algebra and $a, b \in A$. If $\sigma([b, x]) \subseteq \sigma([a, x]) \cup \{0\}$ for all $x \in A$, then for all $y, z \in A$, $yz = yaz = 0$ implies $ybz = 0$.*

Proof. From the assumptions $yz = 0$ and $yaz = 0$ we find by a short calculation that $[a, zxy]^3 = 0$ for all $x \in A$. Consequently, by the hypothesis of the lemma, $\sigma([b, zxy]) = \{0\}$ for all $x \in A$. Assume that $ybz \neq 0$ and, to obtain a contradiction, take an irreducible representation π of A on a Banach space X such that $\pi(ybz) \neq 0$. Choose $\xi \in X$ with $\pi(ybz)\xi \neq 0$. By irreducibility there exists $u \in A$ such that $\pi(u)\pi(ybz)\xi = \xi$. Then $\pi([b, zuy])\eta = -\eta$, where $\eta = \pi(z)\xi$. Hence we have $-1 \in \sigma([b, zuy])$, a contradiction. □

2.3. Derivations on von Neumann algebras. We begin with the study of the spectral inclusion condition (2.1) in von Neumann algebras. As derivations are automatically inner on these algebras (see, e.g., [**KR86**, Exercise 8.7.55]), we assume, throughout, that $d(\,\cdot\,) = [a, \cdot\,]$ and $g(\,\cdot\,) = [b, \cdot\,]$.

For factors the desired conclusion, which is the same as for primitive Banach algebras with nonzero socle, follows easily from Lemma 2.2.5 and the result on reflexivity from [**Mag93**]. But first we need a technical lemma.

Lemma 2.3.1. *Let A be a von Neumann algebra and I be a norm closed ideal in A. If an element $a \in A$ with $a + I \notin Z(A/I)$ satisfies $\lambda\sigma([a, x] + I) \subseteq \sigma([a, x] + I) \cup \{0\}$ for all $x \in A$ and for some $\lambda \in \mathbb{C}$, then $\lambda \in \{-1, 0, 1\}$.*

Proof. We first note that A/I is not necessarily a von Neumann algebra (since I is not necessarily weak* closed), but this does not make the proof more difficult since A/I has many projections. Since $a + I$ is not central in A/I, we can find a projection $p \in A$ such that $(1 - p)ap \notin I$. Let A act on a Hilbert space $H = pH + (1 - p)H$. According to this decomposition we can represent every $a \in A$ as a 2×2 matrix

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix},$$

where $a_1 = pap$, $a_2 = pa(1 - p)$ and so on. If we choose x in $pB(1 - p)$ so that x is represented by the matrix which has an element x_2 on the position $(1, 2)$ and zeros elsewhere, then a short computation shows that

$$[a, x] = \begin{pmatrix} -x_2 a_3 & a_1 x_2 - x_2 a_4 \\ 0 & a_3 x_2 \end{pmatrix}.$$

Since $a_3 = (1 - p)ap \notin I$, it follows that $a_3 a_3^* \notin I$ and there exists a closed subset in $\sigma(a_3 a_3^*)$ that does not contain 0 such that its characteristic function χ satisfies

$\chi(a_3a_3^*) \notin I$. Take a function g that satisfies $\chi(t) = g(t)t$ for every $t \in \sigma(a_3a_3^*)$ and $g(0) = 0$. If we choose $x_2 = a_3^*g(a_3a_3^*)$, then $[a, x]$ is of the form

$$[a, x] = \begin{pmatrix} -q_1 & y \\ 0 & q_2 \end{pmatrix},$$

where $q_1 = \chi(a_3^*a_3)$ and $q_2 = \chi(a_3a_3^*)$. (We have used the well-known fact that $g(a_3a_3^*)a_3 = a_3g(a_3^*a_3)$ which follows by approximating g by polynomials.) Since q_1 and q_2 are projections and not contained in I, we now see that $\{-1, 1\} \subseteq \sigma([a, x] + I) \subseteq \{-1, 0, 1\}$, from which the lemma is evident. □

Proposition 2.3.2. *Let A be a factor and let $a, b \in A$ satisfy $\sigma([b, x]) \subseteq \sigma([a, x]) \cup \{0\}$ for every $x \in A$. Then $b = \lambda a + \mu 1$ for some $\lambda \in \{-1, 0, 1\}$ and some $\mu \in \mathbb{C}$.*

Proof. The case $A = B(H)$ has already been handled by Theorem 2.2.1. Assume that $A \neq B(H)$ and $b \notin \operatorname{span}\{1, a\}$. According to [**Mag93**, Theorem 1.1], every finite dimensional subspace S in a factor A different from $B(H)$ is reflexive in the sense that for each $b \in A \setminus S$ there exist $y, z \in A$ such that $ySz = 0$ and $ybz \neq 0$. Thus, taking $S = \operatorname{span}\{1, a\}$, there exist $y, z \in A$ such that $yz = 0$ and $yaz = 0$ but $ybz \neq 0$, contradicting Lemma 2.2.5. Therefore $b \in \operatorname{span}\{1, a\}$. Now Lemma 2.3.1 with $I = 0$ yields the desired conclusion. □

We now proceed to general von Neumann algebras. Let us introduce some notation and list some standard results that will be needed in the sequel.

Denote by X the character space of $Z(A)$. Let At be the closed ideal in A generated by $t \in X$. We write $A(t)$ for the quotient algebra A/At and $x(t)$ for the coset $x + At \in A/At$. The function $t \mapsto \|x(t)\|$ is continuous for every $x \in A$ and the map $x \mapsto (x(t))_{t \in X}$ from A to $\Pi_{t \in X} A(t)$ is injective and hence an isometric embedding, in particular $\|x\| = \sup_{t \in X} \|x(t)\|$ (see [**Gli60**] for proofs). By [**Hal69**, Theorem 4.7], $A(t)$ is primitive for every $t \in X$.

The spectrum of elements relative to some subalgebra B of A will be denoted by $\sigma_B(\,.\,)$.

Lemma 2.3.3. *Let A be a von Neumann algebra, c an element in A, and let t be an element in the character space X of $Z(A)$. If P_t is the set of all projections in $Z(A)$ that correspond to the characteristic functions of those clopen sets that contain t, then $\sigma(c(t)) = \bigcap_{p \in P_t} \sigma_{pA}(pc)$.*

Proof. Observe that $A(t)$ is a quotient of pA for every $p \in P_t$. (Indeed, if p corresponds to the characteristic function χ of a clopen set containing t, then $1 - p \in t$, hence $A(1 - p) \subseteq At$, so $A(t)$ is a quotient of pA.) This immediately implies that $\sigma(c(t)) \subseteq \bigcap_{p \in P_t} \sigma_{pA}(pc)$.

For the reverse inclusion it suffices to show that

$$\sigma(c(t))^c \subseteq \cup_{p \in P_t} \sigma_{pA}(pc)^c.$$

Assume that $c(t)$ is invertible. It is enough to show that there exists a clopen set $U \subseteq X$ which contains t and that $c(s)$ is invertible for every $s \in U$. Then pc is invertible in pA where p is the projection that corresponds to the characteristic function of U.

Consider the polar decomposition $c = u|c|$ of c, hence $c(s) = u(s)|c(s)|$ for all $s \in X$. Since $c(t)$ is invertible, $u(t)$ is unitary. The continuous functions $s \mapsto \|u(s)u^*(s) - 1\|$ and $s \mapsto \|u^*(s)u(s) - 1\|$ equal zero at $s = t$. Hence, these functions are less than 1 on some neighborhood V of t in X. Since $u(s)$ is a partial isometry for every $s \in X$, it follows that $u(s)$ must be invertible (thus unitary) for all s in V. Hence we only need to show that $|c|$ is invertible, so we may assume that $c \geq 0$ on V. As $c(t)$ is invertible, $c(t) \geq m1$ for some $m \in \mathbb{R}^+$. We only need to show that

$c(s) > \frac{m}{2}1$ for all s is some neighborhood of t. Suppose the contrary that there exists a net $\{s_j\}_{j\in J}$ converging to t such that $\sigma(c(s_j))$ contains some $\lambda_j < \frac{m}{2}$ for every $j \in J$. Let $f : \mathbb{R} \to \mathbb{R}$ be the continuous map defined by

$$f(x) = \begin{cases} 1 & \text{if } x < \frac{m}{2}, \\ 2 - \frac{2}{m}x & \text{if } \frac{m}{2} \le x \le m, \\ 0 & \text{if } x > m. \end{cases}$$

Then $f(c(t)) = 0$ and $\|f(c(s_j))\| = 1$ for every $j \in J$. As $f(c)(s) = f(c(s))$, we have $f(c)(t) = 0$ and $\|f(c)(s_j)\| = 1$ for every $j \in J$. However, this contradicts the continuity of the map $s \mapsto \|f(c)(s)\|$. $\square$

THEOREM 2.3.4. *Let A be a von Neumann algebra and let $a, b \in A$. If $\sigma([b, x]) \subseteq \sigma([a, x]) \cup \{0\}$ for every $x \in A$, then $b = p_1a - p_2a + z$ for some orthogonal central projections p_1, p_2 and some $z \in Z(A)$.*

PROOF. There exist central orthogonal projections $z_1, z_2 \in A$ with $z_1 + z_2 = 1$ such that z_1A is of Type I while z_2A does not contain central portions of Type I. As $[z_ib, x] = [b, z_ix]$, we have the inclusion $\sigma([z_ib, x]) \subseteq \sigma([z_ia, x]) \cup \{0\}$ for every $x \in z_iA$, $i = 1, 2$, therefore $\sigma_{z_iA}([z_ib, x]) \subseteq \sigma_{z_iA}([z_ia, x]) \cup \{0\}$. Hence the proof will be divided into two cases, the one where A is of Type I and the one where A does not contain central portions of Type I.

Case 1. Let A be a von Neumann algebra of Type I. It suffices to show that the assertion of the theorem holds for $A = C(X)\overline{\otimes}B(H)$, the von Neumann algebra of continuous functions from a Stonean space X with values in $B(H)$, equipped with a weak operator topology, for a Hilbert space H. Indeed, A is a direct sum of such algebras.

We take $\xi, \eta \in H$ with $\langle \xi, \eta\rangle = 0$. Choose $\epsilon > 0$, $t_0 \in X$ and let U be a clopen neighborhood of t_0 with the property $|\langle(a(t) - a(t_0))\xi, \eta\rangle| < \epsilon$ and $|\langle(b(t) - b(t_0))\xi, \eta\rangle| < \epsilon$ for every $t \in U$. We calculate the spectra of $[a,(\xi \otimes \eta^*)_U]$ and $[b,(\xi\otimes\eta^*)_U]$, where $(\xi\otimes\eta^*)_U$ denotes the function $t \mapsto \chi_U(t)\xi\otimes\eta^*$ for every $t \in X$. We have

$$\sigma([a,(\xi \otimes \eta^*)_U]) \cup \{0\} = \cup_{t\in U}\sigma([a(t), \xi \otimes \eta^*]) \cup \{0\}$$

and

$$\sigma([b,(\xi \otimes \eta^*)_U]) \cup \{0\} = \cup_{t\in U}\sigma([b(t), \xi \otimes \eta^*]) \cup \{0\}.$$

From the second paragraph of the proof of Theorem 2.2.1 we see that

$$\sigma([a(t), \xi \otimes \eta^*]) \subseteq \{0, -\langle a(t)\xi, \eta\rangle, \langle a(t)\xi, \eta\rangle\}$$

and $\sigma([a(t), \xi \otimes \eta^*]) = \{0\}$ if and only if $\langle a(t)\xi, \eta\rangle = 0$. Thus,

$$\sigma([a,(\xi \otimes \eta^*)_U]) \subseteq \cup_{t\in U}\{0, -\langle a(t)\xi, \eta\rangle, \langle a(t)\xi, \eta\rangle\}$$

and by choice of U we have $|\langle a(t)\xi, \eta\rangle - \langle a(t_0)\xi, \eta\rangle| < \epsilon$. The same conclusions hold if we replace a by b. Since $\sigma([b,(\xi\otimes\eta^*)_U]) \subseteq \sigma([a,(\xi\otimes\eta^*)_U])\cup\{0\}$, it follows that $\langle b(t_0)\xi, \eta\rangle = 0$ or $|\langle b(t_0)\xi, \eta\rangle - \langle a(t_0)\xi, \eta\rangle| < \epsilon$ or $|\langle b(t_0)\xi, \eta\rangle + \langle a(t_0)\xi, \eta\rangle| < \epsilon$ for every $\epsilon > 0$. Consequently, we have $\langle b(t_0)\xi, \eta\rangle = 0$ or $\langle b(t_0)\xi, \eta\rangle = \pm\langle a(t_0)\xi, \eta\rangle$. Following the proof of Theorem 2.2.1 we may conclude that $b(t_0) \in \mathbb{C}1$ or $b(t_0) \pm a(t_0) \in \mathbb{C}1$ for every $t_0 \in X$.

Therefore, the union of the closed sets $F_0 = \{t \in X : b(t) \in \mathbb{C}1\}$, $F_1 = \{t \in X : b(t) - a(t) \in \mathbb{C}1\}$ and $F_2 = \{t \in X : b(t) + a(t) \in \mathbb{C}1\}$ equals X. Complements of these sets are open and the interiors G_0, G_1, G_2 of F_0, F_1, F_2, respectively, are clopen. Note that G_0^c is the closure $\overline{F_0^c}$ of F_0^c. Since the union of the sets F_j is X, F_0^c is contained in $F_1 \cup F_2$. But then also $G_0^c = \overline{F_0^c} \subseteq F_1 \cup F_2$ (since $F_1 \cup F_2$ is closed). The sets G_0, G_0^c then divide X in the disjoint union of two clopen sets, contained in F_0, $F_1 \cup F_2$, respectively. Similarly, we divide G_0^c in the union of two disjoint clopen sets $H_1 \subseteq F_1$, $H_1^c \subseteq F_2$. (Namely, the clopen set G_0^c is the union of the two closed sets $E_1 = F_1 \cap G_0^c$ and $E_2 = F_2 \cap G_0^c$, and we may take H_1 to be the interior of E_1 and $H_2 = G_0^c \setminus H_1$.) The characteristic functions of G_0, H_1, H_1^c yield central projections p_0, p_1, p_2 with sum 1. Moreover, the elements $z_0 = p_0 b$, $z_1 = p_1(b-a)$, $z_2 = p_2(b+a)$ are central. The result is $b = (p_0+p_1+p_2)b = p_1 a - p_2 a + z_0 + z_1 + z_2$.

Case 2. Assume now that A does not contain central portions of Type I. Let us examine the linear independence of $1, a, b$ in the quotient spaces $A(t)$ for $t \in X$. If the elements $1, a(t_0), b(t_0)$ are linearly independent for some $t_0 \in X$, then there exists a neighborhood U of t_0 such that $1, a(t), b(t)$ are linearly independent for every $t \in U$. (In order to prove this, we consider the continuous map $f : X \times S \to \mathbb{R}$ defined by $f(t, \alpha, \beta, \gamma) \mapsto \|\alpha 1 + \beta a(t) + \gamma b(t)\|$, where S is the unit sphere in $\mathbb{C}^3$. As $1, a(t_0), b(t_0)$ are linearly independent, $f(t_0, \alpha, \beta, \gamma) > m$ for some $m > 0$ and for all $(\alpha, \beta, \gamma) \in S$. Since S is compact we can find a neighborhood U of $t_0 \in X$ such that $f(t, \alpha, \beta, \gamma) > \frac{m}{2}$ for all $t \in U$, $(\alpha, \beta, \gamma) \in S$.)

Replacing U by its appropriate subset, if necessary, we may assume that U is clopen. Its characteristic function is continuous on X and corresponds to some projection $p \in Z(A)$. Thus $1, a(t), b(t)$ are linearly independent for every t in the character space X_p of $Z_p = Z(pA)$. From the hypothesis and since p is central, we have $\sigma([pb, x]) \subseteq \sigma([pa, x]) \cup \{0\}$ for every $x \in pB$.

Consider the map $f' : Z_p^2 \to pA$, defined by $(z_1, z_2) \mapsto z_1 + z_2 pa$. Since 1 and $a(t)$ are linearly independent for all $t \in X_p$, there exists $m' > 0$ such that $\|\alpha + \beta pa(t)\|^2 \geq m'(|\alpha|^2 + |\beta|^2)$ for every $\alpha, \beta \in \mathbb{C}$. Therefore, $\|z_1(t) + z_2(t)pa(t)\|^2 \geq m'(|z_1(t)|^2 + |z_2(t)|^2)$ for all $z_1, z_2 \in Z_p$. Taking the supremum over $t \in X_p$ gives $\|z_1 + z_2 pa\|^2 \geq m' \max\{\|z_1\|^2, \|z_2\|^2\}$. Thus the map f' is bounded from below, hence its range is norm-closed. Since the range of any weak* continuous linear map is norm closed if and only if it is weak* closed (see, e.g., [**Con90**, Chapter VI, Theorem 1.10]), $Z_p + Z_p a$ is weak* closed. Since pB does not contain central portions of Type I, by [**Mag93**, Theorem 1.1] the finitely generated weak*-closed Z_p-submodule $S := Z_p + Z_p a$ of pA is reflexive in the sense that for each $v \in pA \setminus S$ there exist $y, z \in pA$ such that $ySz = 0$ and $yvz \neq 0$. Thus, since $pb \notin S$, there exist $y, z \in pA$ such that $yz = 0$, $ypaz = 0$ but $ypbz \neq 0$, which contradicts Lemma 2.2.5. Thus, $b(t) \in \operatorname{span}\{1, a(t)\}$ for every $t \in X$ for which $1, a(t)$ are linearly independent. Hence in this case there exist $\lambda(t), \mu(t) \in \mathbb{C}$ such that $b(t) = \lambda(t)a(t) + \mu(t)1$. Further, from the hypothesis of the theorem applied to elements of the form qx where q is a projection in $Z(A)$ we have that $\sigma([qb, qx]) \subseteq \sigma([qa, qx]) \cup \{0\}$; then applying the formula from Lemma 2.3.3 we see that $\sigma([b(t), x(t)]) \subseteq \sigma([a(t), x(t)]) \cup \{0\}$ for all $x \in A$. Thus, now we have that $\lambda(t)\sigma([a(t), x(t)]) \subseteq \sigma([a(t), x(t)])$ for all $x \in A$, from which we conclude by Lemma 2.3.1 that $\lambda(t) \in \{-1, 0, 1\}$. In the remaining case, when $a(t)$ is a scalar for some t, $b(t)$ must also be a scalar by Lemma 2.3.3 and [**Aup91**, Theorem 5.2.1].

Therefore, the union of the closed sets $F_0 = \{t \in X : b(t) \in \mathbb{C}1\}$, $F_1 = \{t \in X : b(t) - a(t) \in \mathbb{C}1\}$ and $F_2 = \{t \in X : b(t) + a(t) \in \mathbb{C}1\}$ equals X. The proof can now be completed by the same argument as at the end of Case 1. □

2.4. Derivations on C^*-algebras. On C^*-algebras we consider inner derivations. Our first result is an easy consequence of Lemma 2.2.5 and a deeper result from [**ABEV09**].

THEOREM 2.4.1. *Let A be a unital C^*-algebra and $a, b \in A$ with a normal. If $\sigma([b, x]) \subseteq \sigma([a, x])$ for all $x \in A$, then $b \in \{a\}''$, the bicommutant of $\{a\}$ in A.*

PROOF. Let $A = \{a\}'$. Define $\phi : A \times A \to A$ by $\phi(y, z) = ybz$. For all $u, v \in A$, $uv = 0$ implies $uav = 0$. According to Lemma 2.2.5 this further gives $\phi(u, v) = 0$. Then $\phi(xy, z) = \phi(x, yz)$ for all $x, y, z \in A$ [**ABEV09**, Theorem 2.11 and Example 2, p. 137]. Setting $x = z = 1$ we get $yb = by$. □

We conclude this section by considering the norm inequality condition

$$(2.2)\qquad ||[b,x]|| \le M||[a,x]|| \quad \text{for all selfadjoint } x \in B,$$

which follows immediately from the spectral inclusion condition (2.1) for inner derivations induced by b, a, resp., if a and b are selfadjoint. When can (2.2) occur? For instance, if $b = a^n$, then we see from $[b,x] = a^{n-1}[a,x] + a^{n-2}[a,x]a + \cdots + [a,x]a^{n-1}$ that (2.2) holds with $M = n\|a\|^{n-1}$. Consequently, (2.2) is fulfilled whenever b is a polynomial in a. In the next theorem we will show directly that (2.2) implies that b is a Lipschitz function f of a, provided that A is a prime C*-algebra and a is a normal element. This is perhaps not the best possible conclusion, however, a complete description of the properties of the appropriate functions f could be too difficult. In [**JW75**] Johnson and Williams considered the special case where $A = B(H)$. By [**JW75**, Corollary 3.7], in this case the condition (2.2) is equivalent to the requirement that the range of $[b,x]$ is contained in the range of $[a,x]$ for every $x \in B(H)$. The description of appropriate functions f in the case $A = B(H)$ in [**JW75**, Theorems 3.6 and 4.1] is quite entangled. In a general C*-algebra A the range inclusion $\operatorname{im}[b,x] \subseteq \operatorname{im}[a,x]$ for all $x \in A$ implies the condition (2.2) by [**KS01**, Theorem 6.5].

THEOREM 2.4.2. *Let A be a prime C^*-algebra and let $a, b \in A$ satisfy $\|[b,x]\| \le M\|[a,x]\|$ for all selfadjoint $x \in A$ and some $M > 0$. If a is normal, then $b = f(a)$ where f is a Lipschitz function (with a Lipschitz constant M) on the spectrum of a.*

PROOF. We can assume $a \notin \mathbb{C}1$ without loss of generality.

Our assumption implies that a and b commute. Since a is normal, a commutes also with b^* by the Putnam-Fuglede theorem, and then the condition (2.2) implies that b is normal. Denote by A the C^*-algebra generated by a and b. The Gelfand transformation is an isomorphism between A and $C(\Omega)$, the algebra of continuous functions on the character space Ω of A, which can be identified with a compact subset K of $\mathbb{C}^2$ via the homeomorphism $\psi : \chi \mapsto (\chi(a), \chi(b))$. Let $\hat{a}$ and $\hat{b}$ denote the Gelfand transforms of a and b, regarded as functions on K. (These are just the restrictions to K of the two coordinate projections $\mathbb{C}^2 \to \mathbb{C}$.)

We can divide $\mathbb{C}^2 = \mathbb{R}^4$ into a grid of small closed cubes with sides parallel to the coordinate axes such that the intersection of any two cubes is either empty or a common face. Then it is not hard to see that there exists $p \in \mathbb{N}$ such that each cube intersects at most $p-1$ other cubes ($p = 3^4$; incidentally, $p - 1 = p_0 + p_1 + p_2 + p_3$, where p_k is the number of k-dimensional faces in the 4-dimensional cube). By taking slightly larger open cubes we can cover the compact set K by a finite family $\{P_i\}_{i=1}^n$ of such cubes, so that each intersects at most $p-1$ other cubes; moreover, for a given $\epsilon > 0$, by the uniform continuity we may assume that the cubes P_i are so small that $|\hat{a}(t) - \hat{a}(t')| < \epsilon$ and $|\hat{b}(t) - \hat{b}(t')| < \epsilon$ whenever $t, t' \in K$ are such that $t \in P_i$ and $t' \in P_j$ for some i, j with $P_i \cap P_j \ne \emptyset$.

Let $V_i = P_i \cap K$ and choose a partition of unity $\{\phi_i\}_{i=1}^n$ subordinate to the covering $\{V_i\}_{i=1}^n$ of K. (The functions ϕ_i will be also regarded as elements of A.) Then for each $t \in K$ and arbitrary $t_i \in V_i$, $|\hat{a}(t) - \sum_{i=1}^n \hat{a}(t_i)\phi_i(t)| \le \sum_{i:\, t \in V_i} |\hat{a}(t) - \hat{a}(t_i)|\phi_i(t) < \epsilon$, hence

$$\|\hat{a} - \sum_{i=1}^n \hat{a}(t_i)\phi_i\| < \epsilon,$$

and, similarly,

$$\|\hat{b} - \sum_{i=1}^n \hat{b}(t_i)\phi_i\| < \epsilon \text{ for arbitrary } t_i \in V_i.$$

Note also that $\phi_j \sum_{i:V_i\cap V_j\neq\emptyset} \phi_i = \phi_j$ for all $1 \le j \le n$. Moreover, if $V_i \cap V_j \neq \emptyset$ then $|\hat{a}(t_i) - \hat{a}(t_j)| < \epsilon$ for arbitrary $t_i \in V_i, t_j \in V_j$.

Choose $1 \le j, k \le n$ such that $V_j \cap V_k = \emptyset$. Since A is prime, there exists $y \in A$ such that $\|\phi_j y \phi_k\| = 1$. As $\phi_j\phi_k = 0$, we have $\|\phi_j y\phi_k \pm \phi_k y^*\phi_j\| = 1$. (Here we have used the equality $\|u + v\| = \max\{\|u\|, \|v\|\}$ which holds for two elements u and v in a C*-algebra satisfying $u^*v = 0 = uv^*$.) Set $x = \phi_j y\phi_k + \phi_k y^*\phi_j$. Take arbitrary $t_i \in V_i$. Let $L(j)$ denote the set of all $i \in \{1,\dots,n\}$ such that $V_i \cap V_j \neq \emptyset$; then $\phi_j \sum_{i\in L(j)} \phi_i = \phi_j$. For $p, q \in \mathbb{R}$ we write $p \approx_\epsilon q$ if $|p - q| < \epsilon$. According to the above observations we can estimate

$$\begin{aligned}
\|[a,x]\| \quad &\approx_{2\epsilon} \quad \left\| \left[\sum_{i=1}^n \hat{a}(t_i)\phi_i, x\right] \right\| \\
&= \quad \Big\| \sum_{L(j)} \hat{a}(t_i)(\phi_i\phi_j y\phi_k - \phi_k y^*\phi_j\phi_i) \\
&\qquad - \sum_{L(k)} \hat{a}(t_i)(\phi_j y\phi_k\phi_i - \phi_i\phi_k y^*\phi_j) \Big\| \\
&\approx_{4p\epsilon} \quad \Big\| \hat{a}(t_j)\phi_j \left(\sum_{L(j)} \phi_i\right) y\phi_k - \hat{a}(t_j)\phi_k y^*\phi_j \left(\sum_{L(j)} \phi_i\right) \\
&\qquad - \hat{a}(t_k)\phi_j y\phi_k \left(\sum_{L(k)} \phi_i\right) + \hat{a}(t_k)\phi_k \left(\sum_{L(k)} \phi_i\right) y^*\phi_j \Big\| \\
&= \quad \|\hat{a}(t_j)(\phi_j y\phi_k - \phi_k y^*\phi_j) - \hat{a}(t_k)(\phi_j y\phi_k - \phi_k y^*\phi_j)\| \\
&= \quad |\hat{a}(t_j) - \hat{a}(t_k)|.
\end{aligned}$$

In the same manner we show that $\|[b,x]\| \approx_{(4p+2)\epsilon} |\hat{b}(t_j) - \hat{b}(t_k)|$. Since in these estimates $\epsilon > 0$ and $t_i \in V_i$ are arbitrary, we may now conclude from our basic assumption (2.2) that

$$|\hat{b}(t) - \hat{b}(t')| \le M|\hat{a}(t) - \hat{a}(t')| \text{ for all } t, t' \in K.$$

From this we see in particular that $\hat{a}(t) = \hat{a}(t')$ implies $\hat{b}(t) = \hat{b}(t')$, hence b is a function of a, say $b = f(a)$. The above inequality means that this function f is Lipschitz. □

COROLLARY 2.4.3. *Let A be a prime C^*-algebra. The following conditions are equivalent for normal elements $a, b \in A$:*

(1) $\|[a,x]\| = \|[b,x]\|$ for every selfadjoint $x \in A$.
(2) $b = \lambda a + \mu 1$ or $b = \lambda a^* + \mu 1$ for some $\lambda, \mu \in \mathbb{C}$ with $|\lambda| = 1$.

PROOF. By replacing a with $a - t_0 1$ for some $t_0 \in \sigma(a)$ we may assume that $0 \in \sigma(a)$. From Theorem 2.4.2 we obtain $b = f(a)$ for a Lipschitz function f on $\sigma(a)$ and $a = g(b)$ for a Lipschitz function g on $\sigma(b)$, with both f and g having Lipschitz constants 1. Then $g \circ f = id_{\sigma(a)}$ yields $|f(t) - f(t')| = |t - t'|$ for all $t, t' \in \sigma(a)$, since $|t-t'| = |g(f(t)) - g(f(t'))| \le |f(t) - f(t')| \le |t-t'|$. Replacing f by $f - f(0)$ we may assume that $f(0) = 0$. It is easy to see that then f must take one of the following forms: $f(t) = e^{i\theta}t$ or $f(t) = e^{i\theta}\bar{t}$ for some $\theta \in [0, 2\pi)$. This establishes (2).

The converse is trivial. □

THEOREM 2.4.4. *Let A be a C^*-algebra on some Hilbert space and $\overline{A}$ be its weak* closure. The following conditions are equivalent for selfadjoint elements $a, b \in A$:*

(1) $r([a,x]) = r([b,x])$ for every $x \in A$.
(2) $\|[a,x]\| = \|[b,x]\|$ for every selfadjoint $x \in A$.
(3) For every primitive ideal P of A we have that $a + b + P \in \mathbb{C}1 + P$ or $a - b + P \in \mathbb{C}1 + P$.
(4) $b = ca + z$ for some $c, z \in Z(\overline{A})$ with $z = z^*, c = c^*, c^2 = 1$.

PROOF. Take a selfadjoint $x \in A$. Then $[a,x]$, $[b,x]$ are anti-selfadjoint and $\|[a,x]\| = r([a,x])$, $r([b,x]) = \|[b,x]\|$. Thus, (1) implies (2).

Assume that (2) holds. We will first observe that a passage to the quotient A/I for an arbitrary ideal I in A preserves the condition (2). This observation follows from the existence of a quasicentral approximate unit in I (see, e.g., [**KR86**, Exercise 10.5.6]). This is an increasing net $\{e_\lambda\}_{\lambda\in\Lambda}$ of positive elements in I which is an approximate unit for I and satisfies $\lim_\lambda \|e_\lambda z - ze_\lambda\| = 0$ for every $z \in A$. Since $\|y + I\| = \lim_\lambda \|(1-e_\lambda)y\|$ for all $y \in A$ (see, e.g., [**KR86**, Exercise 4.6.60]), we have

$$\begin{aligned}\|[a,x]+I\|\| &= \lim_\lambda \|(1-e_\lambda)[a,x]\| = \lim_\lambda \|[a,(1-e_\lambda)x]\| \\ &= \lim_\lambda \|[b,(1-e_\lambda)x]\| = \lim_\lambda \|(1-e_\lambda)[b,x]\| = \|[b,x]+I\|.\end{aligned}$$

Taking for I any primitive ideal P and taking into account that a and b are selfadjoint, we now deduce from Corollary 2.4.3 that $a+b+P \in \mathbb{C}1+P$ or $a-b+P \in \mathbb{C}1+P$. This proves (3).

If we assume (3), then $[a+b,x] \in P$ for every $x \in A$ or $[a-b,y] \in P$ for every $y \in A$. Hence $[a-b,A]A[a+b,A] \subseteq P$ for every primitive ideal P. Since A is semisimple, we obtain $[a-b,A]A[a+b,A] = \{0\}$. Write $S = [a-b,A]$, $T = [a+b,A]$. It follows that $S\overline{A}T = \{0\}$. Let $J = \{u \in \overline{A} : \ u\overline{A}T = \{0\}\}$. Note that J is a strongly closed ideal of $\overline{A}$. Therefore there exists a projection $p \in Z(\overline{A})$ such that $J = p\overline{A}$ (see, e.g., [**KR86**, Theorem 6.8.8]). Thus $pT = \{0\}$, and so $(1-p)t = t$ for every $t \in T$. On the other hand, $ps = s$ for every $s \in S$ since $S \subseteq J$. Therefore, $(1-p)[a-b,x] = 0$ and $p[a+b,x] = 0$ for all $x \in A$. Since p is central, $z_1 = (1-p)(a-b) \in Z(\overline{A})$ and $z_2 = p(a+b) \in Z(\overline{A})$. Hence, $b = (1-p)b+pb = (1-p)a-z_1+z_2-pa = (1-2p)a+z$ for an element $z = z^* \in Z(\overline{A})$. Taking $c = 1-2p$ establishes (4).

The implication (4)⇒(1) is clear. □

The elements c and z may be sometimes chosen from $Z(A)$, but in general we cannot expect this.

EXAMPLE 2.4.5. Take $A = C([0,1], M_2(\mathbb{C}))$, the C^*-algebra of continuous functions with matrix values. Define a and b by $a(t) = |1-2t|J$, $b(t) = (1-2t)J$, where J is an arbitrary selfadjoint non-scalar matrix. The two elements a and b satisfy (1), therefore $b = ca+z$ for some $c, z \in Z(\overline{A})$ with $z = z^*, c = c^*, c^2 = 1$. If c, z belonged to A, we would have $b(t) = c(t)a(t) + z(t)$ for every $t \in [0,1]$. Hence, $c(t) = 1$ on $[0, \frac{1}{2})$ and $c(t) = -1$ on $(\frac{1}{2}, 1]$, contradicting the continuity of c.

3. Derivations with zero spectral function

3.1. The property β. We will deal with the class of Banach algebras having the following property.

DEFINITION 3.1.1. A Banach algebra A is said to have the *property β* if there exists a family of continuous irreducible representations $(\pi_i)_{i\in I}$ of A on Banach spaces X_i such that

(a) $\bigcap_i \ker \pi_i = \operatorname{rad}(A)$.
(b) If $\dim X_i \geq 2$, then there exists $q \in A$ such that $q^2 = 0$ and $\pi(q) \neq 0$.

EXAMPLE 3.1.2. Every commutative Banach algebra obviously has the property β.

EXAMPLE 3.1.3. For every Banach space X, the algebra of all bounded linear operators on X, $\mathcal{L}(X)$, has the property β. Indeed, just take $\pi = 1$ and a nonzero finite rank nilpotent for q. More generally, a primitive Banach algebra with nonzero socle has the property β.

EXAMPLE 3.1.4. A Banach algebra A is said to have the *property* $\mathbb{B}$ if every continuous bilinear map $\varphi\colon A\times A\to X$, where X is an arbitrary Banach space, with the property that for all $a,b\in A$,

$$ab=0 \implies \varphi(a,b)=0,$$

necessarily satisfies

$$\varphi(ab,c)=\varphi(a,bc) \quad (a,b,c\in A).$$

This definition was introduced in [**ABEV09**]. The class of Banach algebras with the property $\mathbb{B}$ is quite large. It includes C^*-algebras, group algebras on arbitrary locally compact groups, Banach algebras generated by idempotents, and topologically simple Banach algebras containing a nontrivial idempotent. Furthermore, this class is stable under the usual methods of constructing Banach algebras.

We claim that

$$A \text{ has the property } \mathbb{B} \implies A \text{ has the property } \beta.$$

Indeed, take a continuous irreducible representation π of a Banach algebra A with the property $\mathbb{B}$ on a Banach space X with $\dim(X)\geq 2$. It is enough to show that there exist $a,b\in A$ such that

$$ab=0,\ \pi(a)\neq 0,\ \pi(b)\neq 0.$$

Namely, since $\pi(A)$ is a prime algebra, we can then find $c\in A$ such that $\pi(b)\pi(c)\pi(a)\neq 0$. Hence $q=bca$ satisfies $q^2=0$ and $\pi(q)\neq 0$, as required in Definition 3.1.1. Assume, therefore, that such a and b do not exist. That is, for all $a,b\in A$, $ab=0$ implies $\pi(a)=0$ or $\pi(b)=0$. Then the continuous bilinear mapping

$$\varphi\colon A\times A\to L(X)\widehat{\otimes}L(X),\quad \varphi(a,b)=\pi(a)\otimes\pi(b)\quad (a,b\in A)$$

satisfies the condition $ab=0\implies\varphi(a,b)=0$. Consequently, we have

$$\pi(a)\pi(b)\otimes\pi(c)=\pi(a)\otimes\pi(b)\pi(c) \quad (a,b,c\in A).$$

Let $\xi,\zeta\in X\setminus\{0\}$. There exist $a,b\in A$ such that $\pi(a)\xi=\zeta$ and $\pi(b)\zeta=\xi$. Then $\pi(a)\pi(b)\otimes\pi(a)=\pi(a)\otimes\pi(b)\pi(a)$ and both $\pi(a)$ and $\pi(b)\pi(a)$ are different from zero. This implies that there exists $\lambda\in\mathbb{C}$ such that $\pi(a)=\lambda\pi(b)\pi(a)$. Hence

$$\zeta=\pi(a)\xi=\lambda\pi(b)\pi(a)\xi=\lambda\pi(b)\zeta=\lambda\xi.$$

From this we conclude that $\dim(X)=1$, a contradiction.

EXAMPLE 3.1.5. Let A have the property β and let $(\pi_i)_{i\in I}$ be the corresponding representations. The following constructions will be used later.

(1) The quotient Banach algebra $A/\mathrm{rad}(A)$ also has the property β. Indeed, for every $i\in I$ the representation π_i drops to an irreducible representation ϖ_i of the quotient Banach algebra $A/\mathrm{rad}(A)$ on X_i by defining

$$\varpi_i(a+\mathrm{rad}(A))=\pi_i(a) \quad (a\in A).$$

It is clear that $(\varpi_i)_{i\in I}$ satisfies the required properties.

(2) Assume that A does not have an identity element. Let $A_{\mathbf{1}}$ be the Banach algebra formed by adjoining an identity to A, so that $A_{\mathbf{1}}=\mathbb{C}\mathbf{1}\oplus A$. For every $i\in I$, the representation π_i lifts to an irreducible representation ϖ_i of $A_{\mathbf{1}}$ on X_i by defining

$$\varpi_i(\alpha\mathbf{1}+a)\xi \;=\; \alpha\xi+\pi_i(a)\xi \quad (\alpha\in\mathbb{C},\, a\in A,\, \xi\in X_i).$$

Further, we adjoin the 1-dimensional representation $\varpi(\alpha\mathbf{1}+a)=\alpha$ ($\alpha\in\mathbb{C}, a\in A$) to the family $(\varpi_i)_{i\in I}$. Then the resulting family satisfies the requirements of Definition 3.1.1. That is, $A_{\mathbf{1}}$ has the property β.

3.2. Tools. The purpose of this section is to gather together the results needed for the proof of Theorem 3.3.1 below. We start with a simple lemma which indicates that it is enough to consider the condition $d(Q) \subseteq Q$ on semisimple Banach algebras.

LEMMA 3.2.1. *Let A be a Banach algebra and let d be a derivation of A such that $d(Q) \subseteq Q$. Then $d(\operatorname{rad}(A)) \subseteq \operatorname{rad}(A)$ and the derivation D of the semisimple Banach algebra $A/\operatorname{rad}(A)$, defined by $D(x + \operatorname{rad}(A)) = d(x) + \operatorname{rad}(A)$, satisfies $D(Q_{A/\operatorname{rad}(A)}) \subseteq Q_{A/\operatorname{rad}(A)}$.*

PROOF. Write $\mathcal{R}$ for $\operatorname{rad}(A)$. Then $\big(d(\mathcal{R}) + \mathcal{R}\big)/\mathcal{R}$ is a two-sided ideal of the semisimple Banach algebra $A/\mathcal{R}$. Since $d(Q) \subseteq Q$, it follows that $d(\mathcal{R}) \subseteq Q$ and so $\big(d(\mathcal{R}) + \mathcal{R}\big)/\mathcal{R}$ consists of quasinilpotent elements of $A/\mathcal{R}$. Therefore $\big(d(\mathcal{R}) + \mathcal{R}\big)/\mathcal{R} = \{0\}$, that is, $d(\mathcal{R}) \subseteq \mathcal{R}$.

On account of [**Dal00**, Proposition 1.5.29(i)], we have $Q_{A/\mathcal{R}} = Q_A/\mathcal{R}$ and this clearly implies that $D(Q_{A/\mathcal{R}}) \subseteq Q_{A/\mathcal{R}}$. □

We need two standard results on Banach algebra derivations (see, e.g., [**Dal00**, Proposition 2.7.22(ii) and Theorem 5.2.28(iii)]).

THEOREM 3.2.2. *Let d be a derivation on a Banach algebra A.*

(1) (Sinclair) If d is continuous, then $d(P) \subseteq P$ for each primitive ideal P of A.

(2) (Johnson and Sinclair) If A is semisimple, then d is automatically continuous.

Our main tool is the Jacobson density theorem together with its extensions. First we state a version of this theorem which includes Sinclair's generalization involving invertible elements (see, e.g., [**Aup91**, Theorem 4.2.5, Corollary 4.2.6]).

THEOREM 3.2.3. *Let π be a continuous irreducible representation of a unital Banach algebra A on a Banach space X. If $\xi_1, \dots, \xi_n$ are linearly independent elements in X, and $\eta_1, \dots, \eta_n$ are arbitrary elements in X, then there exists $a \in A$ such that $\pi(a)\xi_i = \eta_i$, $i = 1, \dots, n$. Moreover, if $\eta_1, \dots, \eta_n$ are linearly independent, then a can be chosen to be invertible.*

The next theorem is basically [**BB01**, Theorem 4.6], but stated in the analytic setting (alternatively, one can use [**BŠ99**, Theorem 3.6] together with Theorem 3.2.2).

THEOREM 3.2.4. *Let d be a continuous derivation on a Banach algebra A, and let π be a continuous irreducible representation of A on a Banach space X. The following statements are equivalent:*

(i) There does not exist a continuous linear operator $T : X \to X$ such that $\pi(d(x)) = T\pi(x) - \pi(x)T$ for all $x \in A$.

(ii) If $\xi_1, \dots, \xi_n$ are linearly independent elements and $\eta_1, \dots, \eta_n, \zeta_1, \dots, \zeta_n$, are arbitrary elements in X, then there exists $a \in A$ such that

$$\pi(a)\xi_i = \eta_i \quad \textit{and} \quad \pi(d(a))\xi_i = \zeta_i, \;\; i = 1, \dots, n.$$

3.3. The condition $d(Q) \subseteq Q$.

THEOREM 3.3.1. *Let A be a Banach algebra with the property β, and let Q be the set of its quasinilpotent elements. If a derivation d of A satisfies $d(Q) \subseteq Q$, then $d(A) \subseteq \operatorname{rad}(A)$.*

PROOF. We first assume that A is semisimple and has an identity element. Obviously $d(\mathbf{1}) = 0$. On account of Theorem 3.2.2, d is continuous and leaves the primitive ideals of A invariant.

Take an irreducible representation π of A on a Banach space X such as in Definition 3.1.1. We have to show that $\pi(d(A)) = \{0\}$.

Suppose first that $\dim X = 1$. Then $P = \ker \pi$ has codimension 1 in A, so that $A = \mathbb{C}\mathbf{1} \oplus P$. Hence $d(A) \subseteq P$, which gives $\pi(d(A)) = \{0\}$.

We now assume that $\dim X \geq 2$. According to Definition 3.1.1, there exists $q \in A$ such that $q^2 = 0$ and $\pi(q) \neq 0$. Let $\rho \in X$ be such that

$$\omega := \pi(q)\rho \neq 0.$$

Note that ω and ρ are linearly independent for $\pi(q)^2 = 0$. Also,

$$\pi(q)\omega = 0.$$

We now consider two cases.

Case 1. Let us first consider the possibility where conditions of Theorem 3.2.4 are fulfilled. Then there exists $a \in A$ such that

$$\pi(a)\rho = 0,\ \pi(a)\omega = 0,\ \pi(d(a))\rho = \omega,\ \pi(d(a))\omega = -\rho + \pi(d(q))\rho,$$

and

$$\pi(a)\pi(d(q))\rho = 0$$

(if $\pi(d(q))\rho$ lies in the linear span of ρ and ω, then this follows from the first two identities). Note that for any $n \geq 2$,

$$\pi(d(a^n))\rho = \pi(d(a))\pi(a)^{n-1}\rho + \cdots + \pi(a)^{n-1}\pi(d(a))\rho = 0,$$

and, similarly,

$$\pi(d(a^n))\omega = 0.$$

Both formulas trivially also hold for $n = 0$. Consequently,

$$\pi(d(e^a))\rho = \pi\Big(d\Big(\sum_{n=0}^{\infty} \frac{1}{n!}a^n\Big)\Big)\rho = \sum_{n=0}^{\infty} \frac{1}{n!}\pi(d(a^n))\rho = \pi(d(a))\rho = \omega.$$

Similarly,

$$\pi(d(e^a))\omega = \pi(d(a))\omega = -\rho + \pi(d(q))\rho.$$

By assumption, $d(e^{-a}qe^a) \in Q$, and hence also $e^a d(e^{-a}qe^a)e^{-a} \in Q$. Expanding $d(e^{-a}qe^a)$ according to the derivation law, and also using $e^a d(e^{-a}) + d(e^a)e^{-a} = d(\mathbf{1}) = 0$, it follows that

$$b := -d(e^a)e^{-a}q + d(q) + qd(e^a)e^{-a} \in Q.$$

However,

$$\begin{aligned}
\pi(b)\rho =& -\pi(d(e^a))\pi(e^{-a})\pi(q)\rho + \pi(d(q))\rho + \pi(q)\pi(d(e^a))\pi(e^{-a})\rho \\
=& -\pi(d(e^a))\pi(e^{-a})\omega + \pi(d(q))\rho + \pi(q)\pi(d(e^a))\rho \\
=& -\pi(d(e^a))\omega + \pi(d(q))\rho + \pi(q)\omega \\
=& \rho,
\end{aligned}$$

implying that $1 \in \sigma(\pi(b)) \subseteq \sigma(b)$ – a contradiction. This first possibility therefore cannot occur.

Case 2. We may now assume that there exists a continuous linear operator $T : X \to X$ such that

$$\pi(d(x)) = T\pi(x) - \pi(x)T$$

for each $x \in A$. Suppose there exists $\xi \in X$ such that ξ and $\eta := T\xi$ are linearly independent. By Theorem 3.2.3 then there is an invertible $a \in A$ such that $\pi(a)\rho =$

$-\eta$ and $\pi(a)\omega = \xi$. Put $c := d(aqa^{-1})$. Note that $c \in Q$ since $aqa^{-1} \in Q$. However,

$$\begin{aligned}\pi(c)\xi =& \big(T\pi(a)\pi(q)\pi(a)^{-1} - \pi(a)\pi(q)\pi(a)^{-1}T\big)\xi \\ =& T\pi(a)\pi(q)\omega - \pi(a)\pi(q)\pi(a)^{-1}\eta \\ =& \pi(a)\pi(q)\rho = \pi(a)\omega = \xi,\end{aligned}$$

and hence $1 \in \sigma(\pi(c)) \subseteq \sigma(c)$. This is a contradiction, so $T\xi$ and ξ are linearly dependent for every $\xi \in X$. It is easy to see that this implies that T is a scalar multiple of the identity, whence $\pi(d(A)) = 0$.

Finally, we consider the case when A is an arbitrary Banach algebra. On account of Lemma 3.2.1, $d(\mathrm{rad}(A)) \subseteq \mathrm{rad}(A)$ and therefore d drops to a derivation D on the semisimple Banach algebra $A/\mathrm{rad}(A)$ with the property that $D(Q_{A/\mathrm{rad}(A)}) \subseteq Q_{A/\mathrm{rad}(A)}$. According to Example 3.1.5, $A/\mathrm{rad}(A)$ has the property β. If this Banach algebra already has an identity element, then we apply what has previously been proved to show that $D(A/\mathrm{rad}(A)) = \{0\}$ and hence that $d(A) \subseteq \mathrm{rad}(A)$. If $A/\mathrm{rad}(A)$ does not have an identity element, then we consider its unitization B (considered in Example 3.1.5) and we extend D to a derivation Δ of B by defining $\Delta(\mathbf{1}) = 0$. It is clear that $Q_B = Q_{A/\mathrm{rad}(A)}$. Therefore $\Delta(Q_B) \subseteq Q_B$. We thus get $\Delta(B) = \{0\}$, which implies that $D(A/\mathrm{rad}(A)) = \{0\}$ and therefore that $d(A) \subseteq \mathrm{rad}(A)$. □

REMARK 3.3.2. From the proof of Theorem 4.6.1 it is evident that in the case where A is semisimple, the assumption that $d(Q) \subseteq Q$ can be replaced by a milder assumption that $d(q) \in Q$ for every square zero element $q \in A$.

COROLLARY 3.3.3. *Let A be a C^*-algebra and let Q be the set of its quasinilpotent elements. If a derivation d of A satisfies $d(Q) \subseteq Q$, then $d = 0$.*

COROLLARY 3.3.4. *Let G be a locally compact group and let Q be the set of the quasinilpotent elements of $L^1(G)$. If a derivation d of $L^1(G)$ satisfies $d(Q) \subseteq Q$, then $d = 0$.*

CHAPTER 6

Analytic Lie maps

In this chapter we investigate how rigid is the structure of Banach algebras with respect to its analytic Lie structure.

Let A be an ultraprime Banach algebra. We first prove that each approximately commuting continuous linear (resp. quadratic) map on A is near a commuting continuous linear (resp. quadratic) map. We use this analysis to study how close the approximate Lie isomorphisms and the approximate Lie derivations are to isomorphisms and derivations, respectively.

We further study metric versions of Posner's theorems. We show that if continuous linear maps S, T, and ST are approximate derivations, then either S or T approaches to zero. Moreover, if $[T(a), a]$ is near the center of A for each $a \in A$, then either T approaches to zero or A is nearly commutative. Further, we give quantitative estimates of these phenomena.

This chapter is based on [**AEŠV12a**, **AEŠV12b**].

1. Stability of commuting maps and Lie maps

1.1. Preliminaries. Let X be a Banach space. By X^* we denote the topological dual space of X. For a Banach space Y, let $\mathcal{L}(X,Y)$ denote the Banach space of all continuous linear operators from X into Y. We write $\mathcal{L}(X)$ for $\mathcal{L}(X,X)$. We write $\mathcal{L}^2(X)$ for the Banach space of all continuous bilinear maps from $X \times X$ into X. By a continuous quadratic map on X we mean a map $Q\colon X \to X$ of the form $Q(x) = F(x,x)$ $(x \in X)$ with $F \in \mathcal{L}^2(X)$. We write $\mathcal{Q}(X)$ for the set of all continuous quadratic maps on X.

We write $A + \mathbb{C}1$ for the "conditional unitization" of a given Banach algebra A, i.e., $A + \mathbb{C}1 = A$ if A has an identity and $A + \mathbb{C}1$ is formed by adjoining an identity to A otherwise.

The ultraprimeness is a metric version of the primeness which was introduced by Mathieu in [**Mat89**]. Let A be a Banach algebra. For each $a, b \in A$, we write $M_{a,b}$ for the two-sided multiplication operator on A defined by

$$M_{a,b}(x) = axb \ \ (x \in A).$$

Recall that A is prime if $M_{a,b} = 0$ implies $a = 0$ or $b = 0$. We define

$$\kappa(A) = \inf\{\|M_{a,b}\| : \ a, b \in A, \ \|a\| = \|b\| = 1\}.$$

The Banach algebra A is said to be *ultraprime* if $\kappa(A) > 0$. It is clear that each finite-dimensional prime Banach algebra is ultraprime. For each Banach space X, the Banach algebra $\mathcal{L}(X)$ is ultraprime and, more generally, every closed subalgebra A of $\mathcal{L}(X)$ containing the finite rank operators is ultraprime with $\kappa(A) = 1$ [**Mat89**]. Every prime C^*-algebra is ultraprime [**Mat88**].

Throughout we will use the ultraproduct of a sequence of Banach algebras as an important tool. From now on, $\mathcal{U}$ is a fixed free ultrafilter on $\mathbb{N}$. For a sequence of Banach spaces (X_n), we write $(X_n)^{\mathcal{U}}$ for the ultraproduct of (X_n) with respect to $\mathcal{U}$. This is the quotient Banach space $\ell^\infty(\mathbb{N}, X_n)/\mathcal{N}_{\mathcal{U}}$, where $\ell^\infty(\mathbb{N}, X_n)$ stands for the space of all bounded sequences (x_n) with $x_n \in X_n$ $(n \in \mathbb{N})$ and

$\mathcal{N}_{\mathcal{U}} = \{x \in \ell^\infty(\mathbb{N}, X_n) : \lim_{\mathcal{U}} \|x_n\| = 0\}$. With a slight abuse of notation we continue to write (x_n) for the equivalence class it represents; one can check that any definition we make is independent of the choice of representative of the equivalence class. The norm on $(X_n)^{\mathcal{U}}$ is given by $\|\mathsf{x}\| = \lim_{\mathcal{U}} \|x_n\|$ for each $\mathsf{x} = (x_n) \in (X_n)^{\mathcal{U}}$. If (A_n) is a sequence of Banach algebras, then the ultraproduct $(A_n)^{\mathcal{U}}$ turns into a Banach algebra with respect to the obvious product $\mathsf{ab} = (a_n b_n)$ for all $\mathsf{a} = (a_n), \mathsf{b} = (b_n) \in (A_n)^{\mathcal{U}}$. For each $n \in \mathbb{N}$, let $T_n \in \mathcal{L}(X_n, Y_n)$ be given for some Banach space Y_n and assume that $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$. Then we can define $(T_n)^{\mathcal{U}} \in \mathcal{L}\big((X_n)^{\mathcal{U}}, (Y_n)^{\mathcal{U}}\big)$ according to the rule $(x_n) \mapsto (T_n(x_n))$. Moreover,

$$\|(T_n)^{\mathcal{U}}\| = \lim_{\mathcal{U}} \|T_n\|.$$

In particular, if $f_n \in X_n^*$ $(n \in \mathbb{N})$ are given such that $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$, then $(f_n)^{\mathcal{U}} \in \big((X_n)^{\mathcal{U}}\big)^*$ is defined through $(x_n) \mapsto \lim_{\mathcal{U}} f_n(x_n)$. In the same manner, if $F_n \in \mathcal{L}^2(X_n)$ and $Q_n \in \mathcal{Q}(X_n)$ $(n \in \mathbb{N})$ are given with the property that $\sup_{n\in\mathbb{N}} \|F_n\| < \infty$ and $\sup_{n\in\mathbb{N}} \|Q_n\| < \infty$, then $(F_n)^{\mathcal{U}} \in \mathcal{L}^2\big((X_n)^{\mathcal{U}}\big)$ and $(Q_n)^{\mathcal{U}} \in \mathcal{Q}\big((X_n)^{\mathcal{U}}\big)$ are defined according to the rules $((x_n), (y_n)) \mapsto (F_n(x_n, y_n))$ and $(x_n) \mapsto (Q_n(x_n))$, respectively. Moreover,

$$\|(F_n)^{\mathcal{U}}\| = \lim_{\mathcal{U}} \|F_n\| \tag{1.1}$$

and

$$\|(Q_n)^{\mathcal{U}}\| = \lim_{\mathcal{U}} \|Q_n\|. \tag{1.2}$$

We refer to [**Hei80**] for the basics of ultraproducts.

1.2. Stability of commuting maps. Let us recall that a map $T\colon A \to A$ is commuting if

$$[T(a), a] = 0 \qquad (a \in \mathcal{A}). \tag{1.3}$$

1.2.1. *Commuting linear maps.* As mentioned in Section 2.4 every commuting linear map on a prime algebra A is of the form

$$T(a) = \lambda a + \mu(a) \tag{1.4}$$

with λ and $\mathrm{im}(\mu)$ belonging to the extended centroid of A [**BCM07**, Corollary 5.28]. However, in the setting of ultraprime Banach algebras we can neglect the extended centroid.

Remark 1.2.1. If A is an ultraprime Banach algebra then the extended centroid is only the complex field [**Mat89**, Corollary 4.7]. Thus, any commuting linear map $T\colon A \to A$ is of the form (1.4) with $\lambda \in \mathbb{C}$ and $\mu\colon A \to \mathbb{C}\mathbf{1}$ is linear (and clearly μ is continuous whenever T is continuous).

We measure how much a continuous linear operator T from a Banach algebra A into itself satisfies condition (1.3) by considering the constant

$$\mathrm{com}(T) = \sup\big\{\|[T(a), a]\| : a \in A,\ \|a\| = 1\big\}.$$

The subset of $\mathcal{L}(A)$ consisting of commuting maps is denoted by $\mathrm{LCom}(A)$. This is a closed linear subspace of $\mathcal{L}(A)$. Note that the maps $T \mapsto \mathrm{com}(T)$ and $T \mapsto \mathrm{dist}(T, \mathrm{LCom}(A))$ are seminorms on $\mathcal{L}(A)$ which vanish precisely on $\mathrm{LCom}(A)$. Moreover, if $T \in \mathcal{L}(A)$ and $S \in \mathrm{LCom}(A)$, then $\|[T(a), a]\| = \|[(T - S)(a), a]\| \leq 2\|T - S\|\|a\|^2$ for each $a \in A$ and therefore

$$\mathrm{com}(T) \leq 2\,\mathrm{dist}\big(T, \mathrm{LCom}(A)\big)$$

We are interested in whether there is a constant $M > 0$ such that

$$\mathrm{dist}\big(T, \mathrm{LCom}(A)\big) \leq M\,\mathrm{com}(T)$$

It is worth pointing out that $\operatorname{com}(T)$ is nothing but the norm of the quadratic map $a \mapsto [T(a), a]$ on A. Equality (1.2) then gives the following useful property.

Lemma 1.2.2. *Let (A_n) be a sequence of Banach algebras and assume that $T_n \in \mathcal{L}(A_n)$ $(n \in \mathbb{N})$ are given with $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$. Then*

$$\operatorname{com}\left((T_n)^{\mathcal{U}}\right) = \lim_{\mathcal{U}} \operatorname{com}(T_n).$$

Lemma 1.2.3. *Let (A_n) be a sequence of Banach algebras. Then*

$$\kappa\big((A_n)^{\mathcal{U}}\big) = \lim_{\mathcal{U}} \kappa(A_n).$$

Proof. Write $\mathsf{A} = (A_n)^{\mathcal{U}}$. Let $\mathsf{a} = (a_n), \mathsf{b} = (b_n) \in \mathsf{A}$. Then $M_{\mathsf{a},\mathsf{b}} = (M_{a_n,b_n})$ and therefore

$$\|M_{\mathsf{a},\mathsf{b}}\| = \lim_{\mathcal{U}} \|M_{a_n,b_n}\| \geq \lim_{\mathcal{U}} (\kappa(A_n)\|a_n\|\|b_n\|) = \lim_{\mathcal{U}} \kappa(A_n)\|\mathsf{a}\|\|\mathsf{b}\|.$$

This clearly implies that $\kappa(\mathsf{A}) \geq \lim_{\mathcal{U}} \kappa(A_n)$.

In order to prove the converse inequality, for each $n \in \mathbb{N}$, we pick $a_n, b_n \in A_n$ with $\|a_n\| = \|b_n\| = 1$ and $\|M_{a_n,b_n}\| \leq \kappa(A_n) + 1/n$. We then consider $\mathsf{a}, \mathsf{b} \in \mathsf{A}$ given by $\mathsf{a} = (a_n)$ and $\mathsf{b} = (b_n)$. We have

$$\|M_{\mathsf{a},\mathsf{b}}\| = \lim_{\mathcal{U}} \|M_{a_n,b_n}\| \leq \lim_{\mathcal{U}} (\kappa(A_n) + 1/n) = \lim_{\mathcal{U}} \kappa(A_n). \qquad \square$$

Lemma 1.2.4. *Let (A_n) be a sequence of Banach algebras such that $(A_n)^{\mathcal{U}}$ has an identity. Then $\{n \in \mathbb{N} : A_n \text{ has an identity } \} \in \mathcal{U}$.*

Proof. Let $1 = (u_n)$ be the identity of $\mathsf{A} = (A_n)^{\mathcal{U}}$. For every $n \in \mathbb{N}$, let L_{u_n} and R_{u_n} denote the operators of left and right multiplication by u_n on A_n, respectively. Then (L_{u_n}) and (R_{u_n}) are the identity operator on A and therefore $\lim_{\mathcal{U}} \|I_{A_n} - L_{u_n}\| = \lim_{\mathcal{U}} \|I_{A_n} - R_{u_n}\| = 0$. Accordingly,

$$\{n \in \mathbb{N} : \|I_{A_n} - L_{u_n}\|, \|I_{A_n} - R_{u_n}\| < 1\} \in \mathcal{U}.$$

On the other hand, the property $\|I_{A_n} - L_{u_n}\|, \|I_{A_n} - R_{u_n}\| < 1$ implies that both L_{u_n} and R_{u_n} are bijective linear maps from A_n onto itself, which, according to [**CM97**, Proposition 2], implies that A_n has an identity. $\square$

Theorem 1.2.5. *For each $K > 0$ there exists $M > 0$ such that*

$$\operatorname{dist}\big(T, \operatorname{LCom}(A)\big) \leq M \operatorname{com}(T)$$

for each Banach algebra A with $\kappa(A) \geq K$ and $T \in \mathcal{L}(A)$.

Proof. Assume towards a contradiction that the assertion in the theorem is false. Then there exist a constant $K > 0$, a sequence of Banach algebras (A_n) $(n \in \mathbb{N})$, with $\kappa(A_n) \geq K$ $(n \in \mathbb{N})$ and a sequence (F_n) with $F_n \in \mathcal{L}(A_n)$ $(n \in \mathbb{N})$ such that $\operatorname{dist}\big(F_n, \operatorname{LCom}(A_n)\big) > n \operatorname{com}(F_n)$ $(n \in \mathbb{N})$. Set $G_n = \operatorname{dist}\big(F_n, \operatorname{LCom}(A_n)\big)^{-1} F_n$ $(n \in \mathbb{N})$. Then

$$\operatorname{dist}\big(G_n, \operatorname{LCom}(A_n)\big) = 1 \quad (n \in \mathbb{N}) \tag{1.5}$$

and

$$\operatorname{com}(G_n) < 1/n \quad (n \in \mathbb{N}). \tag{1.6}$$

Since the sequence (G_n) is not necessarily bounded, we replace it with a bounded one that still satisfies both (1.5) and (1.6). To this end, on account of (1.5), we can choose a sequence (H_n) with $H_n \in \operatorname{LCom}(A_n)$ $(n \in \mathbb{N})$ and $\lim_{n\to\infty} \|G_n - H_n\| = 1$. We then define $T_n = G_n - H_n$ for each $n \in \mathbb{N}$. It is clear that $\operatorname{dist}\big(T_n, \operatorname{LCom}(A_n)\big) = \operatorname{dist}\big(G_n, \operatorname{LCom}(A_n)\big)$ $(n \in \mathbb{N})$ and then (1.5) gives

$$\operatorname{dist}\big(T_n, \operatorname{LCom}(A_n)\big) = 1 \quad (n \in \mathbb{N}). \tag{1.7}$$

Furthermore, $\operatorname{com}(T_n) = \operatorname{com}(G_n)$ $(n \in \mathbb{N})$ and therefore (1.6) now yields

$$\lim_{n\to\infty} \operatorname{com}(T_n) = 0. \tag{1.8}$$

We now consider $\mathsf{A} = (A_n)^{\mathcal{U}}$ and $\mathsf{T} = (T_n)^{\mathcal{U}} \in \mathcal{L}(\mathsf{A})$. On account of Lemma 1.2.2 and (1.8), $\operatorname{com}(\mathsf{T}) = 0$ and therefore T is commuting. According to Lemma 1.2.3, $\kappa(\mathsf{A}) = \lim_{\mathcal{U}} \kappa(A_n) \geq K$ and hence the Banach algebra A is ultraprime. By Remark 1.2.1 we obtain $\lambda \in \mathbb{C}$ and a continuous linear map $\Phi\colon \mathsf{A} \to Z(\mathsf{A})$ such that

$$\mathsf{T}(\mathsf{a}) = \lambda\mathsf{a} + \Phi(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A}). \tag{1.9}$$

Our next goal is to show that $\Phi = (\mu_n)^{\mathcal{U}}$ for some bounded sequence (μ_n) of continuous linear maps $\mu_n\colon A_n \to Z(A_n)$ $(n \in \mathbb{N})$. We first assume that A does not have an identity. Then $Z(\mathsf{A}) = \{0\}$ so that $\Phi = 0$ and then we can take $\mu_n = 0$ for each $n \in \mathbb{N}$. We now assume that A has an identity. Let $\mathsf{1}$ be the identity of A and let $\varphi\colon \mathsf{A} \to \mathbb{C}$ be a continuous linear functional such that $\Phi(\mathsf{a}) = \varphi(\mathsf{a})\mathsf{1}$ $(\mathsf{a} \in \mathsf{A})$. On account of Lemma 1.2.4, we have $U = \{n \in \mathbb{N} : \ A_n \text{ has an identity}\} \in \mathcal{U}$. For each $n \in U$, let 1_n be the identity of A_n. Then $\kappa(A_n)\|1_n\|^2 \leq \|M_{1_n,1_n}\| = 1$ and so $\|1_n\| \leq K^{-1/2}$ for each $n \in U$. Define $(u_n) \in \mathsf{A}$ by $u_n = 1_n$ for each $n \in U$ and $u_n = 0$ elsewhere. Then $\mathsf{1} = (u_n)$ and (1.9) can be written as

$$\lim_{\mathcal{U}} \|T_n(a_n) - \lambda a_n - \varphi(\mathsf{a})u_n\| = 0 \quad (\mathsf{a} = (a_n) \in \mathsf{A}). \tag{1.10}$$

For each $n \in U$, let $f_n\colon A_n \to \mathbb{C}$ be a continuous linear functional such that $f_n(u_n) = 1$ and $\|f_n\| = \|u_n\|^{-1} \leq 1$. For each $n \in \mathbb{N} \setminus U$ we consider f_n to be the zero functional on A_n. Further, we define a bounded sequence (φ_n) of continuous linear functionals $\varphi_n\colon A_n \to \mathbb{C}$ by

$$\varphi_n(a) = f_n(T_n(a)) - \lambda f_n(a) \quad (a \in A_n,\ n \in \mathbb{N}).$$

Our next objective is to prove that $\varphi = (\varphi_n)^{\mathcal{U}}$, which clearly implies that $\Phi = (\mu_n)^{\mathcal{U}}$, where $\mu_n\colon A_n \to Z(A_n)$ is defined by $\mu_n(a) = \varphi_n(a)u_n$ for all $a \in A_n$ and $n \in \mathbb{N}$. Let $\mathsf{a} = (a_n) \in \mathsf{A}$. Then for each $n \in U$,

$$|\varphi_n(a_n) - \varphi(\mathsf{a})| = \big|f_n\big(T_n(a_n) - \lambda a_n - \varphi(\mathsf{a})u_n\big)\big| \leq \|T_n(a_n) - \lambda a_n - \varphi(\mathsf{a})u_n\|$$

and (1.10) then yields $\varphi(\mathsf{a}) = \lim_{\mathcal{U}} \varphi_n(a_n)$, as claimed.

Finally, (1.9) now reads as follows

$$\lim_{\mathcal{U}} \|T_n - \lambda I_{A_n} - \mu_n\| = 0.$$

Since the map $\lambda I_{A_n} + \mu_n$ lies in $\operatorname{LCom}(A_n)$ for each $n \in \mathbb{N}$, it follows that

$$\lim_{\mathcal{U}} \operatorname{dist}\big(T_n, \operatorname{LCom}(A_n)\big) \leq \lim_{\mathcal{U}} \|T_n - \lambda I_{A_n} - \mu_n\| = 0,$$

which contradicts (1.7). □

1.2.2. *Commuting quadratic maps.* Let us recall the standard form of commuting biadditive maps on a ring R:

$$Q(a) = \lambda a^2 + \mu(a)a + \nu(a) \qquad (a \in R) \tag{1.11}$$

with $\lambda \in Z(R)$, $\mu\colon R \to Z(R)$ an additive map, and $\nu\colon R \to Z(R)$ the trace of a biadditive map (see 4).

The following proposition is a version of [**BCM07**, Theorem 5.32] for ultraprime Banach algebras.

PROPOSITION 1.2.6. *Let A be an ultraprime Banach algebra and let $Q \in \mathcal{Q}(A)$ be a commuting map. Then there exist $\lambda \in \mathbb{C}$, $\mu \in A^*$, and $\nu \in \mathcal{Q}(A)$ with $\nu(A) \subset Z(A)$ such that $Q(a) = \lambda a^2 + \mu(a)a + \nu(a)$ for each $a \in A$.*

PROOF. If A is commutative, then A is isomorphic to $\mathbb{C}$ and the decomposition of Q obviously holds true.

We now turn to the case where A is not commutative. On account Remark 1.2.1 and [**BCM07**, Theorem 5.32], Q is of the form (1.11), where $\lambda \in \mathbb{C}$, $\mu\colon A \to \mathbb{C}$ is a linear functional, and $\nu\colon A \to Z(A)$ is a quadratic map. In order to show that both μ and ν are continuous we pick $u, v \in A$ with $[u, v] \neq 0$ and we now observe that

$$\mu(a)[u,v] = \frac{1}{2}[Q(a+u) - Q(a-u) - 2\lambda(au+ua) - 2\mu(u)a, v] \quad (a \in A),$$

which shows that μ is continuous. Since $\nu(a) = Q(a) - \lambda a^2 - \mu(a)a$ $(a \in A)$, it follows that ν is also continuous. □

Let A be a Banach algebra and $Q \in \mathcal{Q}(A)$. The constant

$$\operatorname{com}(Q) = \sup\bigl\{\|[Q(a),a]\| : \ a \in A, \ \|a\| = 1\bigr\}$$

still makes sense and it gives a measurement of how much Q satisfies condition (1.3). The subset of $\mathcal{Q}(A)$ consisting of commuting maps is denoted by $\operatorname{QCom}(A)$. This is a closed linear subspace of $\mathcal{Q}(A)$. Note that the maps $Q \mapsto \operatorname{com}(Q)$ and $Q \mapsto \operatorname{dist}(Q, \operatorname{QCom}(A))$ are seminorms on $\mathcal{Q}(A)$ which vanish precisely on $\operatorname{QCom}(A)$. Moreover, if $Q \in \mathcal{Q}(A)$ and $Q' \in \operatorname{QCom}(A)$, then $\|[Q(a),a]\| = \|[(Q-Q')(a),a]\| \leq 2\|Q-Q'\|\|a\|^2$ for each $a \in A$ and therefore

$$\operatorname{com}(Q) \leq 2\operatorname{dist}\bigl(Q, \operatorname{QCom}(A)\bigr) \quad (Q \in \mathcal{Q}(A)).$$

We are interested in whether there is a constant $M > 0$ such that

$$\operatorname{dist}\bigl(Q, \operatorname{QCom}(A)\bigr) \leq M \operatorname{com}(Q) \quad (Q \in \mathcal{Q}(A)).$$

In the next theorem we will be required to make a measurement of the commutativity of a given Banach algebra A. To this end, we introduce the following constant

$$\chi(A) = \sup\{\|ab - ba\| : \ a, b \in A, \ \|a\| = \|b\| = 1\}.$$

This constant is the norm of the bilinear map $(a,b) \mapsto [a,b]$ on A and so (1.1) yields the following property.

LEMMA 1.2.7. *Let (A_n) be a sequence of Banach algebras. Then*

$$\chi\bigl((A_n)^{\mathcal{U}}\bigr) = \lim_{\mathcal{U}} \chi(A_n).$$

THEOREM 1.2.8. *For each $K > 0$ there exists $M > 0$ such that*

$$\operatorname{dist}\bigl(Q, \operatorname{QCom}(A)\bigr) \leq M \operatorname{com}(Q)$$

for each Banach algebra A with $\kappa(A) \geq K$ and $Q \in \mathcal{Q}(A)$.

PROOF. This follows by the same method as in Theorem 1.2.5. Suppose the assertion of the theorem is false. Then there exist a constant $K > 0$, a sequence of Banach algebras (A_n) with $\kappa(A_n) \geq K$ $(n \in \mathbb{N})$, and a sequence (Q_n) with $Q_n \in \mathcal{Q}(A_n)$ $(n \in \mathbb{N})$ such that

$$\lim_{n\to\infty} \|Q_n\| = 1,$$

$$\operatorname{dist}\bigl(Q_n, \operatorname{QCom}(A_n)\bigr) = 1 \quad (n \in \mathbb{N}), \tag{1.12}$$

and

$$\lim_{n\to\infty} \operatorname{com}(Q_n) = 0.$$

For each $n \in \mathbb{N}$, let $F_n\colon A_n \times A_n \to A_n$ be the symmetric continuous bilinear map such that $Q_n(a) = F_n(a,a)$ $(a \in A_n)$.

Take $\mathsf{A} = (A_n)^{\mathcal{U}}$, $\mathsf{F} = (F_n)^{\mathcal{U}} \in \mathcal{L}^2(\mathsf{A})$, and $\mathsf{Q} = (Q_n)^{\mathcal{U}} \in \mathcal{Q}(\mathsf{A})$. Of course, F is symmetric and $\mathsf{Q}(\mathsf{a}) = \mathsf{F}(\mathsf{a}, \mathsf{a})$ $(\mathsf{a} \in \mathsf{A})$.

It is a simple matter to check that $\operatorname{com}(\mathsf{Q}) = \lim_{\mathcal{U}} \operatorname{com}(Q_n) = 0$, which implies that Q is commuting. In fact, Lemma 1.2.2 still holds true because the map com is now given by the norm of the trace of a trilinear map and (1.2) works for this case. Next, according to Lemma 1.2.3, A is ultraprime. By Proposition 1.2.6, there exist $\lambda \in \mathbb{C}$, a continuous linear functional $\mathsf{M} \colon \mathsf{A} \to \mathbb{C}$, and a continuous quadratic map $\mathsf{N} \colon \mathsf{A} \to Z(\mathsf{A})$ such that

$$\mathsf{Q}(\mathsf{a}) = \lambda \mathsf{a}^2 + \mathsf{M}(\mathsf{a})\mathsf{a} + \mathsf{N}(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A}). \tag{1.13}$$

Further, let $\mathsf{G} \colon \mathsf{A} \times \mathsf{A} \to Z(\mathsf{A})$ be the symmetric bilinear map associated to N. The linearization of (1.13) gives

$$2\mathsf{F}(\mathsf{a}, \mathsf{b}) = \lambda(\mathsf{ab} + \mathsf{ba}) + (\mathsf{M}(\mathsf{b})\mathsf{a} + \mathsf{M}(\mathsf{a})\mathsf{b}) + 2\mathsf{G}(\mathsf{a}, \mathsf{b}) \quad (\mathsf{a}, \mathsf{b} \in \mathsf{A}). \tag{1.14}$$

Our objective is to prove that $\mathsf{M} = (\mu_n)^{\mathcal{U}}$ and $\mathsf{N} = (\nu_n)^{\mathcal{U}}$ for some bounded sequences (μ_n) of continuous linear functionals $\mu_n \colon A_n \to \mathbb{C}$ and (ν_n) of continuous quadratic maps $\nu_n \colon A_n \to Z(A_n)$.

We first assume that A is commutative. Then A is trivial so that we can take $\mathsf{M} = \mathsf{N} = 0$ in (1.13). Accordingly, the maps $\mu_n = \nu_n = 0$ $(n \in \mathbb{N})$ satisfy our requirement.

We now assume that A is not commutative. Then $\chi(\mathsf{A}) > 0$ and we can pick $0 < \varrho < \chi(\mathsf{A})$. From Lemma 1.2.7, it follows that $V = \{n \in \mathbb{N} : \ \varrho < \chi(A_n)\} \in \mathcal{U}$. For each $n \in V$, we choose $b_n, c_n \in A_n$ with $\|b_n\| = \|c_n\| = 1$ and $\varrho < \|[b_n, c_n]\|$. Furthermore, for each $n \in V$, we take a continuous linear functional $g_n \colon A_n \to \mathbb{C}$ such that $g_n([b_n, c_n]) = 1$ and $\|g_n\| = \|[b_n, c_n]\|^{-1} < \varrho^{-1}$. For each $n \in \mathbb{N} \setminus V$, we consider g_n to be the zero functional on A_n. Let $\mathsf{b}, \mathsf{c} \in \mathsf{A}$ be given by $\mathsf{b} = (b_n)$ and $\mathsf{c} = (c_n)$. On account of (1.14), we have

$$\mathsf{M}(\mathsf{a})[\mathsf{b}, \mathsf{c}] = \big[2\mathsf{F}(\mathsf{a}, \mathsf{b}) - \lambda(\mathsf{ab} + \mathsf{ba}) - \mathsf{M}(\mathsf{b})\mathsf{a}\,,\, \mathsf{c}\big] \quad (\mathsf{a} \in \mathsf{A}).$$

By applying the continuous linear functional $(g_n)^{\mathcal{U}}$ on $(A_n)^{\mathcal{U}}$ we arrive at

$$\mathsf{M}(\mathsf{a}) = \lim_{\mathcal{U}} g_n\big([2F_n(a_n, b_n) - \lambda(a_n b_n + b_n a_n) - \mathsf{M}(\mathsf{b})a_n\,,\, c_n]\big) \quad (\mathsf{a} = (a_n) \in \mathsf{A}),$$

which implies that $\mathsf{M} = (\mu_n)^{\mathcal{U}}$, where (μ_n) is the bounded sequence of continuous linear functionals $\mu_n \colon A_n \to \mathbb{C}$ given by

$$\mu_n(a) = g_n\big([2F_n(a, b_n) - \lambda(a b_n + b_n a) - \mathsf{M}(\mathsf{b})a\,,\, c_n]\big) \quad (a \in A_n,\ n \in \mathbb{N}).$$

In order to show that $\mathsf{N} = (\nu_n)^{\mathcal{U}}$ for some bounded sequence (ν_n) of continuous quadratic maps $\nu_n \colon A_n \to Z(A_n)$ $(n \in \mathbb{N})$, we consider two different cases. First, we assume that A does not have an identity. Then $Z(\mathsf{A}) = \{0\}$, so that $\mathsf{N} = 0$ and then we take $\nu_n = 0$ for each $n \in \mathbb{N}$. We now assume that A has an identity. Let $\psi \colon \mathsf{A} \to \mathbb{C}$ be a quadratic functional such that $\mathsf{N}(\mathsf{a}) = \psi(\mathsf{a})\mathsf{1}$ $(\mathsf{a} \in \mathsf{A})$. On account of Proposition 1.2.4 we have that $U = \{n \in \mathbb{N} : \ A_n \text{ has an identity}\} \in \mathcal{U}$. Define $u_n \in A_n$ and $f_n \in A_n^*$ as in the proof of Theorem 1.2.5. Then (1.13) can be written as

$$\lim_{\mathcal{U}} \|Q_n(a_n) - \lambda a_n^2 - \mu_n(a_n)a_n - \psi(\mathsf{a})u_n\| = 0 \quad (\mathsf{a} = (a_n) \in \mathsf{A}),$$

which implies that

$$\psi(\mathsf{a}) = \lim_{\mathcal{U}} f_n\big(Q_n(a_n) - \lambda a_n^2 - \mu_n(a_n)a_n\big) \quad (\mathsf{a} = (a_n) \in \mathsf{A}).$$

We thus get $\mathsf{N} = (\nu_n)^{\mathcal{U}}$, where $\nu_n \colon A_n \to Z(A_n)$ is defined by

$$\nu_n(a) = f_n\big(Q_n(a) - \lambda a^2 - \mu_n(a)a\big)u_n \quad (a \in A_n,\ n \in \mathbb{N}).$$

Finally, (1.13) reads as $\mathsf{Q} = (P_n)^{\mathcal{U}}$, where $P_n \in \operatorname{QCom}(A_n)$ is defined by

$$P_n(a) = \lambda a^2 + \mu_n(a)a + \nu_n(a) \quad (a \in A_n,\ n \in \mathbb{N}).$$

Hence

$$\lim_{\mathcal{U}} \operatorname{dist}\big(Q_n, \operatorname{QCom}(A_n)\big) \leq \lim_{\mathcal{U}} \|Q_n - P_n\| = 0,$$

which contradicts (1.12). □

1.3. Stability of Lie maps.

1.3.1. *Lie isomorphisms.* The translation of the result [**BCM07**, Corollary 6.5] on the form of Lie isomorphisms to our framework is the following.

Proposition 1.3.1. *Let B be a Banach algebra, let A be an ultraprime Banach algebra, and let $\Phi \in \mathcal{L}(B, A)$ be a Lie isomorphism. Then $\Phi = \Psi + \tau \mathbf{1}$, where $\Psi \in \mathcal{L}(B, A + \mathbb{C}\mathbf{1})$ is either a homomorphism or the negative of an antihomomorphism and $\tau \in B^*$ sends commutators to zero.*

Proof. If A is commutative, then A is isomorphic to $\mathbb{C}$ and the claimed decomposition of Φ obviously holds true.

We now assume that A is not commutative. From [**BCM07**, Corollary 6.5] and Remark 1.2.1 it follows that $\Phi = \Psi + \tau\mathbf{1}$, where $\Psi\colon B \to A + \mathbb{C}\mathbf{1}$ is either a homomorphism or the negative of an antihomomorphism and $\tau\colon B \to \mathbb{C}$ is a linear functional sending commutators to zero. It remains to prove that both Ψ and τ are continuous. Let us consider only the homomorphism case. We can take $v \in B$ and $a \in A$ with $[\Phi(v), a] \neq 0$. It is easily seen that

$$\tau(u)[\Phi(v), a] = [-\Phi(uv) + \Phi(u)\Phi(v) - \tau(v)\Phi(u), a] \quad (u \in B).$$

This implies that τ is continuous. Since $\Psi = \Phi - \tau\mathbf{1}$ we see that Ψ is continuous. □

Our next concern will be the stability of the above result. We introduce the following measurements of the multiplicativity, antimultiplicativity, and Lie multiplicativity of a given map $\Phi \in \mathcal{L}(B, A)$, where A and B are Banach algebras:

$$\operatorname{mult}(\Phi) = \sup\big\{\|\Phi(ab) - \Phi(a)\Phi(b)\| :\ a, b \in B,\ \|a\| = \|b\| = 1\big\},$$
$$\operatorname{amult}(\Phi) = \sup\big\{\|\Phi(ab) - \Phi(b)\Phi(a)\| :\ a, b \in B,\ \|a\| = \|b\| = 1\big\},$$

and

$$\operatorname{lmult}(\Phi) = \sup\big\{\|\Phi([a, b]) - [\Phi(a), \Phi(b)]\| :\ a, b \in B,\ \|a\| = \|b\| = 1\big\}.$$

Further, for $f \in B^*$ we put

$$\|f\|_t = \sup\{|f([a, b])| :\ a, b \in B,\ \|a\| = \|b\| = 1\}$$

It is worth pointing out that the above introduced constants are nothing but the norms of the bilinear maps

$$(a, b) \mapsto \Phi(ab) - \Phi(a)\Phi(b),$$
$$(a, b) \mapsto \Phi(ab) - \Phi(b)\Phi(a),$$
$$(a, b) \mapsto \Phi([a, b]) - [\Phi(a), \Phi(b)],$$

and

$$(a, b) \mapsto f([a, b]),$$

respectively. Consequently, (1.1) yields the following.

Lemma 1.3.2. *Let (A_n) and (B_n) be sequences of Banach algebras and assume that $f_n \in B_n^*$ and $\Phi_n \in \mathcal{L}(B_n, A_n)$ $(n \in \mathbb{N})$ are given with $\sup_{n\in\mathbb{N}} \|f_n\| < \infty$ and*

$\sup_{n\in\mathbb{N}}\|\Phi_n\| < \infty$. *Then*

$$\begin{aligned}\operatorname{mult}\big((\Phi_n)^{\mathcal{U}}\big) &= \lim_{\mathcal{U}} \operatorname{mult}(\Phi_n),\\ \operatorname{amult}\big((\Phi_n)^{\mathcal{U}}\big) &= \lim_{\mathcal{U}} \operatorname{amult}(\Phi_n),\\ \operatorname{lmult}\big((\Phi_n)^{\mathcal{U}}\big) &= \lim_{\mathcal{U}} \operatorname{lmult}(\Phi_n),\end{aligned}$$

and

$$\big\|(f_n)^{\mathcal{U}}\big\|_t = \lim_{\mathcal{U}} \|f_n\|_t.$$

Note that the maps mult, amult, and lmult vanish exactly on the sets $\operatorname{Hom}(B, A)$ of all homomorphisms, $\operatorname{AHom}(B, A)$ of all antihomomorphisms, and $\operatorname{LHom}(B, A)$ of all Lie homomorphisms from B into A, respectively. If $A = B$ then we write briefly $\operatorname{Hom}(A)$, $\operatorname{AHom}(A)$, and $\operatorname{LHom}(A)$ instead of $\operatorname{Hom}(A, A)$, $\operatorname{AHom}(A, A)$, and $\operatorname{LHom}(A, A)$, respectively. The map $\|\cdot\|_t$ vanishes on the linear space of all continuous linear functionals on B sending commutators to zero.

There are two basic choices of assumptions on a map $\Phi \in \mathcal{L}(B, A)$ that could be required to conclude that Φ is an approximate Lie homomorphism (i.e. $\operatorname{lmult}(\Phi)$ is small). The first choice is to consider $\Phi = \Psi + \tau 1$, where $\Psi \in \mathcal{L}(B, A + \mathbb{C}1)$ and $\tau \in B^*$ are such that $\min\{\operatorname{mult}(\Psi), \operatorname{amult}(-\Psi)\}$ and $\|\tau\|_t$ are small (here we are restricting ourselves to the case when $Z(A)$ is trivial). This pattern of thinking leads to introduce the following constants:

$$\begin{aligned}\operatorname{smult}_+(\Phi) &= \inf\{\operatorname{mult}(\Phi - \tau 1) + \|\tau\|_t : \ \tau \in B^*\},\\ \operatorname{smult}_-(\Phi) &= \inf\{\operatorname{amult}(\tau 1 - \Phi) + \|\tau\|_t : \ \tau \in B^*\},\end{aligned}$$

and

$$\operatorname{smult}(\Phi) = \min\{\operatorname{smult}_+(\Phi), \operatorname{smult}_-(\Phi)\}.$$

The second choice is to assume that $\operatorname{dist}\big(\Phi, \operatorname{LHom}(B, A)\big)$ is small.

PROPOSITION 1.3.3. *Let A be an ultraprime Banach algebra. Then there exists a constant $K > 0$ such that*

$$\operatorname{lmult}(\Phi) \le K \operatorname{smult}(\Phi)$$

and

$$\operatorname{lmult}(\Phi) \le \Big(2 + 2\|\Phi\| + 4\operatorname{dist}\big(\Phi, \operatorname{LHom}(B, A)\big)\Big)\operatorname{dist}\big(\Phi, \operatorname{LHom}(B, A)\big)$$

for each Banach algebra B and $\Phi \in \mathcal{L}(B, A)$.

PROOF. Let $\Phi \in \mathcal{L}(B, A)$. Pick $\tau \in B^*$ and write $\Psi = \Phi - \tau 1$. For all $a, b \in B$, we have

$$\Phi([a, b]) - [\Phi(a), \Phi(b)] = \Psi([a, b]) - [\Psi(a), \Psi(b)] + \tau([a, b])1$$

and so $\|\Phi([a, b]) - [\Phi(a), \Phi(b)]\| \le \|\Psi([a, b]) - [\Psi(a), \Psi(b)]\| + \|\tau\|_t\|1\|\|a\|\|b\|$. We can write $\Psi([a, b]) - [\Psi(a), \Psi(b)]$ in two ways. On the one hand, we have

$$\Psi([a, b]) - [\Psi(a), \Psi(b)] = \big(\Psi(ab) - \Psi(a)\Psi(b)\big) - \big(\Psi(ba) - \Psi(b)\Psi(a)\big),$$

which gives $\|\Psi([a, b]) - [\Psi(a), \Psi(b)]\| \le 2\operatorname{mult}(\Psi)\|a\|\|b\|$. On the other hand, we have

$$\begin{aligned}\Psi([a, b]) - [\Psi(a), \Psi(b)] = \big((-\Psi)(ba) &- (-\Psi)(a)(-\Psi)(b)\big)\\ &- \big((-\Psi)(ab) - (-\Psi)(b)(-\Psi)(a)\big),\end{aligned}$$

which yields $\|\Psi([a, b]) - [\Psi(a), \Psi(b)]\| \le 2\operatorname{amult}(-\Psi)\|a\|\|b\|$. This proves the first inequality in the proposition.

Set $\Psi \in \mathrm{LHom}(B, A)$ and write $\Theta = \Phi - \Psi$. For all $a, b \in A$ we have

$$\begin{aligned}\Phi([a,b]) - [\Phi(a), \Phi(b)] &= \Theta([a,b]) + [\Psi(a), \Psi(b)] - [\Phi(a), \Phi(b)] \\ &= \Theta([a,b]) - [\Psi(a), \Theta(b)] - [\Theta(a), \Phi(b)]\end{aligned}$$

and so

$$\begin{aligned}\|\Phi([a,b]) - [\Phi(a), \Phi(b)]\| &\le \big(2\|\Theta\| + 2\|\Theta\|\big(\|\Psi\| + \|\Phi\|\ \big)\big)\|a\|\|b\| \\ &\le \big(2\|\Theta\| + 2\|\Theta\|(\|\Theta\| + 2\|\Phi\|)\big)\|a\|\|b\|.\end{aligned}$$

This gives $\mathrm{lmult}(\Phi) \le 2\|\Theta\| + 2\|\Theta\|(\|\Theta\| + 2\|\Phi\|)$, which establishes the second inequality in the proposition. □

We are now interested in whether $\mathrm{lmult}(\Phi)$ being small implies that $\mathrm{smult}(\Phi)$ is small.

THEOREM 1.3.4. *For each $K, M, \varepsilon > 0$ there exists $\delta > 0$ such that if A and B are Banach algebras with $\kappa(A) \ge K$ and $\Phi \in \mathcal{L}(B, A)$ is bijective with $\|\Phi\|, \|\Phi^{-1}\| \le M$, and $\mathrm{lmult}(\Phi) < \delta$, then $\mathrm{smult}(\Phi) < \varepsilon$.*

PROOF. Suppose the assertion of the theorem is false. Then there exist $K, M, \varepsilon > 0$, sequences (A_n) and (B_n) of Banach algebras, and a sequence (Φ_n) of bijective continuous linear maps $\Phi_n \colon B_n \to A_n$ $(n \in \mathbb{N})$ with

$$\kappa(A_n) \ge K, \tag{1.15}$$

$$\|\Phi_n\|, \|\Phi_n^{-1}\| \le M, \tag{1.16}$$

$$\mathrm{lmult}(\Phi_n) < 1/n, \tag{1.17}$$

and

$$\mathrm{smult}(\Phi_n) \ge \varepsilon \tag{1.18}$$

for each $n \in \mathbb{N}$.

Set $\mathsf{A} = (A_n)^{\mathcal{U}}$, $\mathsf{B} = (B_n)^{\mathcal{U}}$, and $\mathsf{\Phi} = (\Phi_n)^{\mathcal{U}}$. From (1.16) it follows that $\mathsf{\Phi}$ is bijective with inverse given by $(\Phi_n^{-1})^{\mathcal{U}}$. We claim that $\mathsf{\Phi}$ is a Lie isomorphism. Indeed, if $\mathsf{u} = (u_n), \mathsf{v} = (v_n) \in \mathsf{B}$, then (1.17) yields

$$\begin{aligned}\|\mathsf{\Phi}([\mathsf{u},\mathsf{v}]) - [\mathsf{\Phi}(\mathsf{u}), \mathsf{\Phi}(\mathsf{v})]\| &= \lim_{\mathcal{U}} \|\Phi_n([u_n, v_n]) - [\Phi_n(u_n), \Phi_n(v_n)]\| \\ &\le \lim_{\mathcal{U}} \left(\mathrm{lmult}(\Phi_n)\|u_n\|\|v_n\|\right) = 0.\end{aligned}$$

Furthermore, according to Lemma 1.2.3 and (1.15), $\kappa(\mathsf{A}) \ge K$ and hence A is ultraprime.

On account of Proposition 1.3.1, $\mathsf{\Phi} = \mathsf{\Psi} + \tau 1$, where $\mathsf{\Psi} \in \mathcal{L}(\mathsf{B}, \mathsf{A} + \mathbb{C}1)$ is either a homomorphism or the negative of an antihomomorphism and $\tau \in \mathsf{B}^*$ vanishes on commutators. Our purpose is to show that $\mathsf{\Psi} = (\Psi_n)^{\mathcal{U}}$ and $\tau = (\tau_n)^{\mathcal{U}}$, where $\Psi_n \in \mathcal{L}(B_n, A_n + \mathbb{C}1)$ and $\tau_n \in B_n^*$ for each $n \in \mathbb{N}$. We will consider two cases according to the degree of algebraicity of A.

We first assume that the degree of algebraicity of A is greater than 2. For each $n \in \mathbb{N}$ we define $Q_n \in \mathcal{Q}(A_n)$ by

$$Q_n(a) = \Phi_n\left((\Phi_n^{-1}(a))^2\right) \quad (a \in A_n).$$

According to (1.16), $\|Q_n\| \le M^3$ $(n \in \mathbb{N})$ and therefore we can consider $\mathsf{Q} \in \mathcal{Q}(\mathsf{A})$ given by $\mathsf{Q} = (Q_n)^{\mathcal{U}}$. It is clear that $\mathsf{Q}(\mathsf{a}) = \mathsf{\Phi}\left(\mathsf{\Phi}^{-1}(\mathsf{a})^2\right)$ for each $\mathsf{a} \in \mathsf{A}$. Since $\mathsf{\Phi}$ is a Lie isomorphism, it follows that

$$[\mathsf{Q}(\mathsf{a}), \mathsf{a}] = \mathsf{\Phi}\left([\mathsf{\Phi}^{-1}(\mathsf{a})^2, \mathsf{\Phi}^{-1}(\mathsf{a})]\right) = 0$$

for each $\mathsf{a} \in \mathsf{A}$. Hence Q is commuting. From the proof of Theorem 1.2.8 it may be concluded that

$$\mathsf{Q}(\mathsf{a}) = \lambda \mathsf{a}^2 + \mathsf{M}(\mathsf{a})\mathsf{a} + \mathsf{N}(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A}) \tag{1.19}$$

for some $\lambda \in \mathbb{C}$, $\mathsf{M} = (\mu_n)^{\mathcal{U}}$ with $\mu_n \in A_n^*$ $(n \in \mathbb{N})$, and $\mathsf{N} = (\nu_n)^{\mathcal{U}}$ with $\nu_n \in \mathcal{Q}(A_n)$ and $\nu_n(A_n) \subset Z(A_n)$ $(n \in \mathbb{N})$. Taking $\mathsf{a} = \Phi(\mathsf{u})$ with $\mathsf{u} \in \mathsf{B}$ in (1.19) we arrive at

$$\Phi(\mathsf{u}^2) = \lambda \Phi(\mathsf{u})^2 + \mathsf{M}(\Phi(\mathsf{u}))\Phi(\mathsf{u}) + \mathsf{N}(\Phi(\mathsf{u})). \tag{1.20}$$

On the other hand, we have

$$\begin{aligned} \Phi(\mathsf{u}^2) &= \Psi(\mathsf{u}^2) + \tau(\mathsf{u}^2)1 = \sigma\Psi(\mathsf{u})^2 + \tau(\mathsf{u}^2)1 \\ &= \sigma\Phi(\mathsf{u})^2 - 2\sigma\tau(\mathsf{u})\Phi(\mathsf{u}) + (\sigma\tau(\mathsf{u})^2 + \tau(\mathsf{u}^2))1 \quad (\mathsf{u} \in \mathsf{B}), \end{aligned} \tag{1.21}$$

where $\sigma = 1$ in the case where Ψ is a homomorphism and $\sigma = -1$ in the case where Ψ is the negative of an antihomomorphism. Comparing (1.20) and (1.21) we get

$$(\lambda - \sigma)\Phi(\mathsf{u})^2 + \big(\mathsf{M}(\Phi(\mathsf{u})) + 2\sigma\tau(\mathsf{u})\big)\Phi(\mathsf{u}) + \mathsf{N}(\Phi(\mathsf{u})) - (\sigma\tau(\mathsf{u})^2 + \tau(\mathsf{u}^2))1 = 0$$

for each $\mathsf{u} \in \mathsf{B}$. This can be written as follows

$$\begin{aligned} (\lambda - \sigma)\mathsf{a}^2 + &\big(\mathsf{M}(\mathsf{a}) + 2\sigma\tau(\Phi^{-1}(\mathsf{a}))\big)\mathsf{a} \\ &+ \mathsf{N}(\mathsf{a}) - (\sigma\tau(\Phi^{-1}(\mathsf{a}))^2 + \tau(\Phi^{-1}(\mathsf{a})^2))1 = 0 \end{aligned}$$

for each $\mathsf{a} \in \mathsf{A}$. We claim that

$$\tau\left(\Phi^{-1}(\mathsf{a})\right) = -\frac{1}{2\sigma}\mathsf{M}(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A}). \tag{1.22}$$

Since $\deg(\mathsf{A}) > 2$, it follows that $\lambda = \sigma$ and that

$$\big(\mathsf{M}(\mathsf{a}) + 2\sigma\tau(\Phi^{-1}(\mathsf{a}))\big)\mathsf{a} \in Z(\mathsf{A}) \tag{1.23}$$

for each $\mathsf{a} \in \mathsf{A}$. If $\mathsf{a} \in \mathsf{A} \setminus Z(\mathsf{A})$, then (1.23) yields (1.22). Let $\mathsf{a} \in Z(\mathsf{A})$ and pick $\mathsf{b} \in \mathsf{A} \setminus Z(\mathsf{A})$ (such an element exists because $\dim(Z(\mathsf{A})) \leq 1$). Then $a+b \in \mathsf{A} \setminus Z(\mathsf{A})$ and therefore

$$\begin{aligned} \tau\left(\Phi^{-1}(\mathsf{a})\right) &= \tau\left(\Phi^{-1}(\mathsf{a}+\mathsf{b})\right) - \tau\left(\Phi^{-1}(\mathsf{b})\right) \\ &= -\frac{1}{2\sigma}\mathsf{M}(\mathsf{a}+\mathsf{b}) + \frac{1}{2\sigma}\mathsf{M}(\mathsf{b}) = -\frac{1}{2\sigma}\mathsf{M}(\mathsf{a}), \end{aligned}$$

which gives (1.22), as claimed. From (1.22) it may be concluded that $\tau = (\tau_n)^{\mathcal{U}}$ and $\Psi = (\Psi_n)^{\mathcal{U}}$, where

$$\tau_n = -\frac{1}{2\sigma}\mu_n \circ \Phi_n \colon B_n \to \mathbb{C}$$

and

$$\Psi_n = \Phi_n + \frac{1}{2\sigma}\mu_n \circ \Phi_n \colon B_n \to A_n + \mathbb{C}1$$

for each $n \in \mathbb{N}$. Having established the case where $\deg(\mathsf{A}) > 2$, we now turn to the case where $\deg(\mathsf{A}) \leq 2$. Then A is finite-dimensional (in fact, it is isomorphic either to $\mathbb{C}$ or to $M_2(\mathbb{C})$). By [**BT79**, Theorem 3.1], $W = \{n \in \mathbb{N} : \dim(A_n) = \dim(\mathsf{A})\} \in \mathcal{U}$. Since Φ_n is a bijective linear map, it follows that $\dim(B_n) = \dim(A_n) = \dim(\mathsf{A})$ for each $n \in W$ and [**BT79**, Theorem 3.1] then gives $\dim(\mathsf{B}) = \dim(\mathsf{A})$. From [**Hei80**, Theorem 7.1] it follows that $\tau = (\tau_n)^{\mathcal{U}}$ for some sequence (τ_n) with $\tau_n \in B_n^*$ $(n \in \mathbb{N})$. This clearly implies that $\Psi = (\Psi_n)^{\mathcal{U}}$ where $\Psi_n = \Phi_n - \tau_n \in \mathcal{L}(B_n, A_n)$ $(n \in \mathbb{N})$.

Finally, we are in a position to get a contradiction. Having shown that $\Psi = (\Psi_n)^{\mathcal{U}}$ and $\tau = (\tau_n)^{\mathcal{U}}$ where $\Psi_n \in \mathcal{L}(B_n, A_n + \mathbb{C}1)$ and $\tau_n \in B_n^*$ for each $n \in \mathbb{N}$, we can now apply Lemma 1.3.2 to get

$$\operatorname{mult}(\Psi) = \lim_{\mathcal{U}} \operatorname{mult}(\Phi_n - \tau_n 1) \text{ and } \operatorname{amult}(-\Psi) = \lim_{\mathcal{U}} \operatorname{amult}(\tau_n 1 - \Phi_n).$$

From the definition we see that

$$\operatorname{mult}(\Phi_n - \tau_n \mathbf{1}) \geq \operatorname{smult}_+(\Phi_n) - \|\tau_n\|_t$$

and

$$\operatorname{amult}(\tau_n \mathbf{1} - \Phi_n) \geq \operatorname{smult}_-(\Phi_n) - \|\tau_n\|_t$$

for each $n \in \mathbb{N}$. Since either Ψ is a homomorphism or $-\Psi$ is an antihomomorphism, it may be concluded that

$$\begin{aligned} 0 &= \min\{\operatorname{mult}(\Psi), \operatorname{amult}(-\Psi)\} \\ &= \lim_{\mathcal{U}} \min\{\operatorname{mult}(\Phi_n - \tau_n \mathbf{1}), \operatorname{amult}(\tau_n \mathbf{1} - \Phi_n)\} \\ &\geq \lim_{\mathcal{U}} \big(\operatorname{smult}(\Phi_n) - \|\tau_n\|_t\big) \\ &= \lim_{\mathcal{U}} \operatorname{smult}(\Phi_n) - \lim_{\mathcal{U}} \|\tau_n\|_t. \end{aligned}$$

Since $\lim_{\mathcal{U}} \|\tau_n\|_t = \|\tau\|_t = 0$, it follows that $\lim_{\mathcal{U}} \operatorname{smult}(\Phi_n) = 0$, contrary to (1.18), and the proof is complete. $\square$

1.3.2. *Lie derivations.* Let R be a ring and X be an R-bimodule. Recall that an additive map $\Delta\colon R \to X$ is a *Lie derivation* if

$$\Delta([a,b]) = [\Delta(a), b] + [a, \Delta(b)] \quad (a, b \in R).$$

The typical example of a Lie derivation $\Delta\colon R \to X$ is provided by the map

$$\Phi = D + \tau, \tag{1.24}$$

where $D\colon R \to X$ is a derivation and τ is an additive map from R into the center of X sending commutators to zero. The standard problem consists in determining whether every Lie derivation is in the standard form (1.24). M. Brešar showed that if R is a prime ring with characteristic different from 2, then every Lie derivation $\Delta\colon R \to R$ is of the standard form (1.24) provided that we allow D and τ to map into $RC + C$ and C, respectively, where C is the extended centroid of R [**BCM07**, Corollary 6.9]. In the context of ultraprime Banach algebras this result reads as follows.

Proposition 1.3.5. *Let A be an ultraprime Banach algebra and let $\Delta \in \mathcal{L}(A)$ be a Lie derivation. Then $\Delta = D + \tau\mathbf{1}$, where $D \in \mathcal{L}(A, A + \mathbb{C}\mathbf{1})$ is a derivation and $\tau \in A^*$ sends commutators to zero.*

Proof. If A is commutative, then A is isomorphic to $\mathbb{C}$ and the proposition holds true.

We now proceed with the case where A is not commutative. From [**BCM07**, Corollary 6.9] and Remark 1.2.1 it follows that $\Delta = D + \tau\mathbf{1}$, where $D\colon A \to A + \mathbb{C}\mathbf{1}$ is a derivation and $\tau\colon A \to \mathbb{C}$ is a linear functional sending commutators to zero. The only remaining point concerns the continuity. Since A is not commutative, we can pick $b, c \in A$ with $[b,c] \neq 0$. It is immediate to check that

$$\tau(a)[b,c] = [-\Delta(ab) + \Delta(a)b + a\Delta(b) - \tau(b)a, c] \quad (a \in A),$$

which shows that τ is continuous and finally D is continuous because $D = \Delta - \tau\mathbf{1}$. $\square$

Our next objective consists in analysing the stability of the preceding result. To this end, for a given continuous linear map Δ from a Banach algebra A into a Banach A-bimodule X, we define the constants:

$$\operatorname{der}(\Delta) = \sup\big\{\|\Delta(ab) - \Delta(a)b - a\Delta(b)\| : \ \|a\| = \|b\| = 1\big\}$$

and

$$\operatorname{lder}(\Delta) = \sup\bigl\{\|\Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)]\| : \ \|a\| = \|b\| = 1\bigr\}.$$

It should be mentioned that $\operatorname{der}(\Delta)$ and $\operatorname{lder}(\Delta)$ are nothing but the norms of the bilinear maps

$$(a,b) \mapsto \Delta(ab) - \Delta(a)b - a\Delta(b)$$

and

$$(a,b) \mapsto \Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)],$$

respectively. Consequently, (1.1) yields the following.

LEMMA 1.3.6. *Let (A_n) be a sequence of Banach algebras and assume that, for each $n \in \mathbb{N}$, a Banach A_n-bimodule X_n and $\Delta_n \in \mathcal{L}(A_n, X_n)$ are given with* $\sup_{n\in\mathbb{N}} \|\Delta_n\| < \infty$. *Then*

$$\operatorname{der}\bigl((\Delta_n)^{\mathcal{U}}\bigr) = \lim_{\mathcal{U}} \operatorname{der}(\Delta_n)$$

and

$$\operatorname{lder}\bigl((\Delta_n)^{\mathcal{U}}\bigr) = \lim_{\mathcal{U}} \operatorname{lder}(\Delta_n).$$

The maps $\Delta \mapsto \operatorname{der}(\Delta)$ and $\Delta \mapsto \operatorname{lder}(\Delta)$ define seminorms on $\mathcal{L}(A,X)$ vanishing on the linear subspaces $\operatorname{Der}(A,X)$ consisting of all derivations and $\operatorname{LDer}(A,X)$ consisting of all Lie derivations, respectively. We abbreviate $\operatorname{Der}(A,A)$ to $\operatorname{Der}(A)$ and $\operatorname{LDer}(A,A)$ to $\operatorname{LDer}(A)$. We associate the following constant to $\Delta \in \mathcal{L}(A)$:

$$\operatorname{sder}(\Delta) = \inf\bigl\{\operatorname{der}(\Delta - \tau\mathbf{1}) + \|\tau\|_t : \ \tau \in A^*\bigr\}.$$

The map $\Delta \mapsto \operatorname{sder}(\Delta)$ is a seminorm on $\mathcal{L}(A)$.

PROPOSITION 1.3.7. *Let A be an ultraprime Banach algebra. Then there exists a constant $K > 0$ with*

$$\operatorname{lder}(\Delta) \le K \operatorname{sder}(\Delta) \le 3K \operatorname{dist}\bigl(\Delta, \operatorname{LDer}(A)\bigr) \quad (\Delta \in \mathcal{L}(A)).$$

PROOF. Let $\Delta \in \mathcal{L}(A)$. Pick $\tau \in A^*$ and write $D = \Delta - \tau\mathbf{1}$. For all $a, b \in A$, we have

$$\begin{aligned}
&\Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)] \\
&\qquad = D([a,b]) - [D(a),b] - [a,D(b)] + \tau([a,b])\mathbf{1} \\
&\qquad = \bigl(D(ab) - D(a)b - aD(b)\bigr) - \bigl(D(ba) - D(b)a - bD(a)\bigr) + \tau([a,b])\mathbf{1},
\end{aligned}$$

which implies that

$$\|\Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)]\| \le (2\operatorname{der}(D) + \|\tau\|_t\|\mathbf{1}\|)\|a\|\|b\|$$

and, therefore, $\operatorname{lder}(\Delta) \le 2\operatorname{der}(D) + \|\tau\|_t\|\mathbf{1}\|$. This establishes the first inequality in the proposition.

Pick $D \in \operatorname{LDer}(A)$. Then there exists $\tau \in A^*$ sending commutators to zero such that $D - \tau\mathbf{1}$ is a derivation. Accordingly,

$$\begin{aligned}
\operatorname{sder}(\Delta) &\le \operatorname{der}(\Delta - \tau\mathbf{1}) + \|\tau\|_t = \operatorname{der}(\Delta - \tau\mathbf{1}) \\
&\le \operatorname{der}(\Delta - D) + \operatorname{der}(D - \tau\mathbf{1}) = \operatorname{der}(\Delta - D) \le 3\|\Delta - D\|.
\end{aligned}$$

This gives the second inequality in the proposition. □

We are interested in whether the seminorms $\operatorname{lder}(\cdot)$, $\operatorname{sder}(\cdot)$, and $\operatorname{dist}\bigl(\cdot, \operatorname{LDer}(A)\bigr)$ are actually pairwise equivalent. Our next concern is the analysis of $\operatorname{lder}(\cdot)$ and $\operatorname{sder}(\cdot)$. The next section will be concerned with $\operatorname{dist}\bigl(\cdot, \operatorname{LDer}(A)\bigr)$.

THEOREM 1.3.8. *For each $K, M, \varepsilon > 0$ there exists $\delta > 0$ such that if A is a Banach algebra with $\kappa(A) \geq K$ and $\Delta \in \mathcal{L}(A)$ is such that $\|\Delta\| \leq M$ and* $\mathrm{lder}(\Delta) < \delta$, *then* $\mathrm{sder}(\Delta) < \varepsilon$.

PROOF. This follows by the same method as in Theorem 1.3.4. To obtain a contradiction, suppose the assertion of the theorem is false. Then there exist $K, M, \varepsilon > 0$, a sequence of Banach algebras (A_n) with $\kappa(A_n) \geq K$ $(n \in \mathbb{N})$, and a sequence (Δ_n) with $\Delta_n \in \mathcal{L}(A_n)$, $\|\Delta_n\| \leq M$, $\mathrm{lder}(\Delta_n) < 1/n$, and $\mathrm{sder}(\Delta_n) \geq \varepsilon$ $(n \in \mathbb{N})$.

We then consider $\mathsf{A} = (A_n)^{\mathcal{U}}$ and $\Delta = (\Delta_n)^{\mathcal{U}} \in \mathcal{L}(\mathsf{A})$. We claim that Δ is a Lie derivation on A. Indeed, if $\mathsf{a} = (a_n), \mathsf{b} = (b_n) \in \mathsf{A}$, then

$$\begin{aligned}\|\Delta([\mathsf{a},\mathsf{b}]) - [\Delta(\mathsf{a}),\mathsf{b}] - [\mathsf{a},\Delta(\mathsf{b})]\| \\ &= \lim_{\mathcal{U}} \|\Delta_n([a_n,b_n]) - [\Delta_n(a_n),b_n] - [a_n,\Delta_n(b_n)]\| \\ &\leq \lim_{\mathcal{U}} \big(\mathrm{lder}(\Delta_n)\|a_n\|\|b_n\|\big) = 0.\end{aligned}$$

From Lemma 1.2.3 it follows that A is ultraprime and Proposition 1.3.5 then gives $\Delta = \mathsf{D} + \tau 1$, where $\mathsf{D} \in \mathcal{L}(\mathsf{A}, \mathsf{A} + \mathbb{C}1)$ is a derivation and $\tau \in \mathsf{A}^*$ vanishes on commutators. Our next concern will be to show that $\mathsf{D} = (D_n)^{\mathcal{U}}$ and $\tau = (\tau_n)^{\mathcal{U}}$, where $D_n \in \mathcal{L}(A_n, A_n + \mathbb{C}1)$ and $\tau_n \in A_n^*$ for each $n \in \mathbb{N}$.

Let us first consider the case when $\deg(\mathsf{A}) > 2$. For each $n \in \mathbb{N}$ we define $Q_n \in \mathcal{Q}(A_n)$ by

$$Q_n(a) = \Delta_n(a^2) - \Delta_n(a)a - a\Delta_n(a) \quad (a \in A_n).$$

Then $\|Q_n\| \leq 3M$ $(n \in \mathbb{N})$ so that we can consider the map $\mathsf{Q} = (Q_n)^{\mathcal{U}} \in \mathcal{Q}(\mathsf{A})$. Of course, $\mathsf{Q}(\mathsf{a}) = \Delta(\mathsf{a}^2) - \Delta(\mathsf{a})\mathsf{a} - \mathsf{a}\Delta(\mathsf{a})$ for each $\mathsf{a} \in \mathsf{A}$. Since Δ is a Lie derivation, we see that

$$0 = \Delta([\mathsf{a}^2,\mathsf{a}]) = [\Delta(\mathsf{a}^2),\mathsf{a}] + [\mathsf{a}^2,\Delta(\mathsf{a})] = [\mathsf{Q}(\mathsf{a}),\mathsf{a}]$$

for each $\mathsf{a} \in \mathsf{A}$. Therefore Q is commuting and, from the proof of Theorem 1.2.8 we deduce that

$$\mathsf{Q}(\mathsf{a}) = \lambda\mathsf{a}^2 + \mathsf{M}(\mathsf{a})\mathsf{a} + \mathsf{N}(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A})$$

for some $\lambda \in \mathbb{C}$, $\mathsf{M} = (\mu_n)^{\mathcal{U}}$ with $\mu_n \in A_n^*$ $(n \in \mathbb{N})$, and $\mathsf{N} = (\nu_n)^{\mathcal{U}}$ with $\nu_n \in \mathcal{Q}(A_n)$ and $\nu_n(A_n) \subset Z(A_n)$ $(n \in \mathbb{N})$. We claim that

$$\tau(\mathsf{a}) = -\frac{1}{2}\mathsf{M}(\mathsf{a}) \quad (\mathsf{a} \in \mathsf{A}). \tag{1.25}$$

To this end, we now compute $\Delta(\mathsf{a}^2)$ in two ways. On the one hand, we have

$$\Delta(\mathsf{a}^2) = \Delta(\mathsf{a})\mathsf{a} + \mathsf{a}\Delta(\mathsf{a}) + \lambda\mathsf{a}^2 + \mathsf{M}(\mathsf{a})\mathsf{a} + \mathsf{N}(\mathsf{a}).$$

On the other hand, we have

$$\begin{aligned}\Delta(\mathsf{a}^2) &= \mathsf{D}(\mathsf{a}^2) + \tau(\mathsf{a}^2)1 = \mathsf{D}(\mathsf{a})\mathsf{a} + \mathsf{a}\mathsf{D}(\mathsf{a}) + \tau(\mathsf{a}^2)1 \\ &= \Delta(\mathsf{a})\mathsf{a} + \mathsf{a}\Delta(\mathsf{a}) - 2\tau(\mathsf{a})\mathsf{a} + \tau(\mathsf{a}^2)1.\end{aligned}$$

We thus get

$$\lambda\mathsf{a}^2 + \big(\mathsf{M}(\mathsf{a}) + 2\tau(\mathsf{a})\big)\mathsf{a} + \big(\mathsf{N}(\mathsf{a}) - \tau(\mathsf{a}^2)1\big) = 0$$

for each $a \in \mathsf{A}$. Since $\deg(\mathsf{A}) > 2$, it follows that $\lambda = 0$ and that

$$\big(\mathsf{M}(\mathsf{a}) + 2\tau(\mathsf{a})\big)\mathsf{a} \in Z(\mathsf{A})$$

for each $\mathsf{a} \in \mathsf{A}$. If $\mathsf{a} \in \mathsf{A} \setminus Z(\mathsf{A})$, then the preceding property obviously gives (1.25). We now fix $\mathsf{b} \in \mathsf{A} \setminus \mathcal{Z}(\mathsf{A})$. If $\mathsf{a} \in Z(\mathsf{A})$, then $\mathsf{a} + \mathsf{b} \in \mathsf{A} \setminus Z(\mathsf{A})$ and therefore

$$\tau(\mathsf{a}) = \tau(\mathsf{a}+\mathsf{b}) - \tau(\mathsf{b}) = -\frac{1}{2}\mathsf{M}(\mathsf{a}+\mathsf{b}) + \frac{1}{2}\mathsf{M}(\mathsf{b}) = -\frac{1}{2}\mathsf{M}(a).$$

From (1.25) it follows that $\mathsf{D} = (D_n)^{\mathcal{U}}$ and $\tau = (\tau_n)^{\mathcal{U}}$ where

$$\tau_n = -\frac{1}{2}\mu_n \in A_n^* \quad (n \in \mathbb{N})$$

and

$$D_n = \Delta_n - \frac{1}{2}\mu_n \mathbf{1} \in \mathcal{L}(A_n, A_n + \mathbb{C}\mathbf{1}) \quad (n \in \mathbb{N}).$$

We now proceed with the case $\deg(\mathsf{A}) \leq 2$. Then A is finite-dimensional and [**Hei80**, Theorem 7.1] shows that $\tau = (\tau_n)^{\mathcal{U}}$ for some sequence (τ_n) with $\tau_n \in A_n^*$ $(n \in \mathbb{N})$. This implies that $\mathsf{D} = (D_n)^{\mathcal{U}}$ where $D_n = \Delta_n - \tau_n \in \mathcal{L}(A_n, A_n + \mathbb{C}\mathbf{1})$ $(n \in \mathbb{N})$.

From the definition we see that

$$\operatorname{sder}(\Delta_n) \leq \operatorname{der}(D_n) + \|\tau_n\|_t \quad (n \in \mathbb{N})$$

and therefore

$$\begin{aligned} \lim_{\mathcal{U}} \operatorname{sder}(\Delta_n) &\leq \lim_{\mathcal{U}} \bigl(\operatorname{der}(D_n) + \|\tau_n\|_t\bigr) = \lim_{\mathcal{U}} \operatorname{der}(D_n) + \lim_{\mathcal{U}} \|\tau_n\|_t \\ &= \lim_{\mathcal{U}} \operatorname{der}(D_n) + \|\tau\|_t = \lim_{\mathcal{U}} \operatorname{der}(D_n), \end{aligned}$$

which finally gives $\varepsilon \leq \lim_{\mathcal{U}} \operatorname{der}(D_n)$. However, since D is a derivation, Lemma 1.3.6 gives $\lim_{\mathcal{U}} \operatorname{der}(D_n) = \operatorname{der}(\mathsf{D}) = 0$, which is impossible. □

1.4. Lie maps on operator algebras. Throughout this subsection we restrict our attention to the Banach algebra $\mathcal{L}(H)$ for a Hilbert space H. Our purpose is to relate lmult and lder to $\operatorname{dist}\bigl(\cdot, \operatorname{LHom}(\mathcal{L}(H))\bigr)$ and $\operatorname{dist}\bigl(\cdot, \operatorname{LDer}(\mathcal{L}(H))\bigr)$. First of all, it should be pointed out that in the case when H is infinite-dimensional the only linear functional on $\mathcal{L}(H)$ sending commutators to zero is the zero functional. This implies that every Lie automorphism of $\mathcal{L}(H)$ is either an automorphism or the negative of an antiautomorphism of $\mathcal{L}(H)$ and every Lie derivation on $\mathcal{L}(H)$ is, in fact, a derivation. We now turn our attention to the stability problem.

Lemma 1.4.1. *Let H be an infinite-dimensional Hilbert space. Then*

$$\|f\| \leq 2\|f\|_t \quad (f \in \mathcal{L}(H)^*).$$

Proof. Let $T \in \mathcal{L}(H)$. On account of [**Hal54**, Corollary of Theorem 8], we have $T = [P, Q] + [R, S]$ with $P, Q, R, S \in \mathcal{L}(H)$. Further, it is straightforward to check that the operators constructed in [**Hal54**, Lemma 2] satisfy the preceding factorisation are such that $\|P\|, \|R\| \leq \|T\|$ and $\|Q\|, \|S\| \leq 1$. If $f \in \mathcal{L}(H)$, then

$$\begin{aligned} \|f(T)\| &\leq \|f([P, Q])\| + \|f([R, S])\| \\ &\leq \|f\|_t \|P\| \|Q\| + \|f\|_t \|R\| \|S\| \leq 2\|f\|_t \|T\|, \end{aligned}$$

which proves the lemma. □

Theorem 1.4.2. *Let H be an infinite-dimensional separable Hilbert space. For each $M, \varepsilon > 0$ there exists $\delta > 0$ such that if $\Phi\colon \mathcal{L}(H) \to \mathcal{L}(H)$ is a bijective continuous linear map with $\|\Phi\|, \|\Phi^{-1}\| \leq M$ and $\operatorname{lmult}(\Phi) < \delta$ then*

$$\min\bigl\{\operatorname{dist}\bigl(\Phi, \operatorname{Hom}(\mathcal{L}(H))\bigr), \operatorname{dist}\bigl(\Phi, -\operatorname{AHom}(\mathcal{L}(H))\bigr)\bigr\} < \varepsilon.$$

Proof. Suppose, contrary to our claim, that there exist M, $\varepsilon > 0$ and a sequence (Φ_n) of bijective continuous linear maps $\Phi_n\colon \mathcal{L}(H) \to \mathcal{L}(H)$ with $\|\Phi_n\|, \|\Phi_n^{-1}\| \leq M$, $\lim_{n\to\infty} \operatorname{lmult}(\Phi_n) = 0$, and

$$\min\bigl\{\operatorname{dist}\bigl(\Phi_n, \operatorname{Hom}(\mathcal{L}(H))\bigr), \operatorname{dist}\bigl(\Phi_n, -\operatorname{AHom}(\mathcal{L}(H))\bigr)\bigr\} \geq \varepsilon \quad (n \in \mathbb{N}). \tag{1.26}$$

From Theorem 1.3.4 we deduce that $\lim_{n\to\infty} \operatorname{smult}(\Phi_n) = 0$. Consequently, we can assume that either

$$\lim_{n\to\infty} \operatorname{smult}_+(\Phi_n) = 0 \tag{1.27}$$

or

$$\lim_{n\to\infty} \operatorname{smult}_-(\Phi_n) = 0. \tag{1.28}$$

We begin by considering the case when (1.27) holds. Then we get a sequence (τ_n) in $\mathcal{L}(H)^*$ such that $\lim\limits_{n\to\infty} \operatorname{mult}(\Phi_n - \tau_n \mathbf{1}) = 0$ and $\lim\limits_{n\to\infty} \|\tau_n\|_t = 0$. From [**Joh88**, Proposition 6.3] we deduce that there is a sequence (Ψ_n) in $\operatorname{Hom}\big(\mathcal{L}(H)\big)$ with $\lim\limits_{n\to\infty} \|\Phi_n - \tau_n \mathbf{1} - \Psi_n\| = 0$. Moreover, from Lemma 1.4.1 we see that $\lim\limits_{n\to\infty} \|\tau_n\| = 0$, and consequently

$$\lim_{n\to\infty} \operatorname{dist}\big(\Phi_n, \operatorname{Hom}(\mathcal{L}(H))\big) \le \lim_{n\to\infty} \|\Phi_n - \Psi_n\| = 0,$$

which contradicts (1.26).

We now turn to the case when (1.28) holds. Then there exists a sequence (σ_n) in $\mathcal{L}(H)^*$ with $\lim\limits_{n\to\infty} \operatorname{amult}(\sigma_n \mathbf{1} - \Phi_n) = 0$ and $\lim\limits_{n\to\infty} \|\sigma_n\|_t = 0$. From [**AEV11**, Proposition 3.8] it may be concluded that there is a sequence (Θ_n) in $\operatorname{AHom}\big(\mathcal{L}(H)\big)$ with $\lim\limits_{n\to\infty} \|\sigma_n \mathbf{1} - \Phi_n - \Theta_n\| = 0$. By Lemma 1.4.1, $\lim\limits_{n\to\infty} \|\sigma_n\| = 0$, and hence

$$\lim_{n\to\infty} \operatorname{dist}\big(\Phi_n, -A\operatorname{Hom}(\mathcal{L}(H))\big) \le \lim_{n\to\infty} \|\Phi_n - (-\Theta_n)\| = 0,$$

contrary to (1.26). ☐

LEMMA 1.4.3. *Let A be an ultraprime Banach algebra, X be a Banach A-bimodule, and $T \in \mathcal{L}(A, X)$. Let $T_I \colon I \to X$ be the restriction of T to a nonzero two-sided ideal I of A. Then $\|T\| \le \kappa(A)^{-1}(\operatorname{der}(T) + 2\|T_I\|)$.*

PROOF. Let $a, b \in A$ and $c \in I$ with $\|a\| = \|b\| = \|c\| = 1$. Then

$$\begin{aligned}\|T(a)bc\| &\le \|T(a)bc + aT(bc) - T(abc)\| + \|T(abc)\| + \|aT(bc)\| \\ &\le \operatorname{der}(T) + 2\|T_I\|.\end{aligned}$$

This implies that $\|M_{T(a),c}\| \le \operatorname{der}(T) + 2\|T_I\|$. Since $\kappa(A)\|T(a)\| \le \|M_{T(a),c}\|$, it follows that $\|T(a)\| \le \kappa(A)^{-1}(\operatorname{der}(T) + 2\|T_I\|)$ and this establishes the result. ☐

LEMMA 1.4.4. *Let H be a Hilbert space. Then there exists $M > 0$ such that*

$$\operatorname{dist}\big(\Delta, \operatorname{Der}(\mathcal{L}(H))\big) \le M \operatorname{der}(\Delta)$$

for each continuous linear operator $\Delta \colon \mathcal{L}(H) \to \mathcal{L}(H)$.

PROOF. Let $\mathcal{K}(H)$ and $\mathcal{N}(H)$ denote the two-sided ideals of $\mathcal{L}(H)$ consisting of the compact linear operators on H and the nuclear operators on H, respectively. Then $\mathcal{K}(H)$ is an amenable Banach algebra, $\mathcal{N}(H)$ is a Banach $\mathcal{K}(H)$-bimodule, and we can identify $\mathcal{L}(H)$ as the dual of that $\mathcal{K}(H)$-bimodule. Consequently, we can apply [**AEV10**, Theorem 3.1] to get a constant $C > 0$ with the property that for every continuous linear map $\Gamma \colon \mathcal{K}(H) \to \mathcal{L}(H)$ there exists $T \in \mathcal{L}(H)$ such that

$$\|\Gamma - \operatorname{ad}_{\mathcal{K}(H)}(T)\| \le C \operatorname{der}(\Gamma).$$

Here, for each $T \in \mathcal{B}(H)$ we write $\operatorname{ad}(T)$ for the inner derivation on $\mathcal{L}(H)$ implemented by T, i.e. $\operatorname{ad}(T)(S) = [T, S]$ $(S \in \mathcal{L}(H))$, and we write $\operatorname{ad}_{\mathcal{K}(H)}(T)$ for the restriction of $\operatorname{ad}(T)$ to $\mathcal{K}(H)$.

Let $\Delta \in \mathcal{L}(\mathcal{L}(H))$. Then we apply the preceding property to the restriction $\Delta_{\mathcal{K}(H)}$ of Δ to $\mathcal{K}(H)$ to get $T \in \mathcal{L}(H)$ such that

$$\|\Delta_{\mathcal{K}(H)} - \operatorname{ad}_{\mathcal{K}(H)}(T)\| \le C \operatorname{der}(\Delta_{\mathcal{K}(H)}) \le C \operatorname{der}(\Delta).$$

Lemma 1.4.3 then yields

$$\|\Delta - \operatorname{ad}(T)\| \le \operatorname{der}(\Delta - \operatorname{ad}(T)) + 2C \operatorname{der}(\Delta) = (2C + 1) \operatorname{der}(\Delta),$$

which implies that $\operatorname{dist}\big(\Delta, \operatorname{Der}(\mathcal{L}(H)\big) \le (2C + 1) \operatorname{der}(\Delta)$, as required. ☐

THEOREM 1.4.5. *Let H be a Hilbert space. Then there exists $M > 0$ such that*
$$\operatorname{dist}\big(\Delta, \operatorname{Der}(\mathcal{L}(H))\big) \leq M \operatorname{lder}(\Delta)$$
for each continuous linear operator $\Delta\colon \mathcal{L}(H) \to \mathcal{L}(H)$.

PROOF. Suppose the assertion of the theorem is false. Then there exists a sequence (Γ_n) in $\mathcal{L}(\mathcal{L}(H))$ such that $\operatorname{dist}\big(\Gamma_n, \operatorname{Der}(\mathcal{L}(H))\big) = 1$ for each $n \in \mathbb{N}$ and $\lim_{n\to\infty} \operatorname{lder}(\Gamma_n) = 0$. For each $n \in \mathbb{N}$ we define $\Delta_n = \Gamma_n - \delta_n$ where $\delta_n \in \operatorname{Der}(\mathcal{L}(H))$ is chosen so that $\lim_{n\to\infty} \|\Delta_n\| = 1$. Moreover, we have

$$\operatorname{dist}\big(\Delta_n, \operatorname{Der}(\mathcal{L}(H))\big) = \operatorname{dist}\big(\Gamma_n, \operatorname{Der}(\mathcal{L}(H))\big) = 1 \quad (n \in \mathbb{N}) \tag{1.29}$$

and $\operatorname{lder}(\Delta_n) = \operatorname{lder}(\Gamma_n)$ $(n \in \mathbb{N})$, which yields $\lim_{n\to\infty} \operatorname{lder}(\Delta_n) = 0$. From Theorem 1.3.8 it follows that $\lim_{n\to\infty} \operatorname{sder}(\Delta_n) = 0$. This implies that there exists a sequence (τ_n) in $\mathcal{L}(H)^*$ such that $\lim_{n\to\infty} \operatorname{der}(\Delta_n - \tau_n \mathbf{1}) = 0$ and $\lim_{n\to\infty} \|\tau_n\|_t = 0$. From Lemma 1.4.4 we deduce that there exists a sequence (D_n) in $\operatorname{Der}(\mathcal{L}(H))$ with $\lim_{n\to\infty} \|\Delta_n - \tau_n \mathbf{1} - D_n\| = 0$. By Lemma 1.4.1, we have $\lim_{n\to\infty} \|\tau_n\| = 0$. Consequently, we have $\lim_{n\to\infty} \|\Delta_n - D_n\| = 0$, which contradicts (1.29). □

In order to give a full picture about the behaviour of the Lie maps on $\mathcal{L}(H)$ for a Hilbert space H, we now complete the information given in Theorems 1.4.2 and 1.4.5 by considering the finite-dimensional case. It is well-known that Lie automorphisms and Lie derivations of the matrix algebra $M_n(\mathbb{C})$ are of the standard form. Actually, every Lie automorphism of $M_n(\mathbb{C})$ is given by $a \mapsto uau^{-1} + \alpha \operatorname{tr}_n(a)\mathbf{1}$ or $a \mapsto -ua^t u^{-1} + \alpha \operatorname{tr}_n(a)\mathbf{1}$ for some invertible $u \in M_n(\mathbb{C})$ and $\alpha \in \mathbb{C}$, where tr_n and $(\cdot)^t$ stand for the trace and transposition on $M_n(\mathbb{C})$, respectively. Every Lie derivation on $M_n(\mathbb{C})$ is given by $\operatorname{ad}(v) + \alpha \operatorname{tr}_n \mathbf{1}$ for some $v \in M_n(\mathbb{C})$ and $\alpha \in \mathbb{C}$. The stability of both $\operatorname{LHom}(M_n(\mathbb{C}))$ and $\operatorname{LDer}(M_n(\mathbb{C}))$ is provided by the following results.

PROPOSITION 1.4.6. *Let A and B be finite-dimensional Banach algebras. For each $M, \varepsilon > 0$ there exists $\delta > 0$ such that if $\Phi \in \mathcal{L}(B, A)$ with $\|\Phi\| \leq M$ and $\operatorname{lmult}(\Phi) < \delta$ then $\operatorname{dist}\big(\Phi, \operatorname{LHom}(B, A)\big) < \varepsilon$.*

PROOF. The proof of [**Joh88**, Proposition 1.3] takes over almost verbatim. Let $M, \varepsilon > 0$. Then the set
$$\mathcal{C}_{M,\varepsilon} = \big\{T \in \mathcal{L}(B, A) : \ \|T\| \leq M, \ \operatorname{dist}\big(T, \operatorname{LHom}(B, A)\big) \geq \varepsilon\big\}$$
is compact. Further, we consider the decreasing net $(\mathcal{G}_\delta)_{\delta>0}$ of open sets given by
$$\mathcal{G}_\delta = \{T \in \mathcal{L}(B, A) : \ \operatorname{lmult}(T) > \delta\} \quad (\delta > 0).$$
Then
$$\mathcal{C}_{M,\varepsilon} \subset \mathcal{L}(B, A) \setminus \operatorname{LHom}(B, A) = \bigcup_{\delta>0} \mathcal{G}_\delta.$$
Consequently, there exists $\delta > 0$ such that $\mathcal{C}_{M,\varepsilon} \subset \mathcal{G}_\delta$, which is the desired conclusion. □

PROPOSITION 1.4.7. *Let A be a finite-dimensional Banach algebra. Then there exists $M > 0$ such that $\operatorname{dist}\big(\Delta, \operatorname{LDer}(A)\big) \leq M \operatorname{lder}(\Delta)$ for each $\Delta \in \mathcal{L}(A)$.*

PROOF. The seminorms $\operatorname{dist}\big(\cdot, \operatorname{LDer}(A)\big)$ and $\operatorname{lder}(\cdot)$ on $\mathcal{L}(A)$ vanish on $\operatorname{LDer}(A)$ so that both of them give rise to norms on the quotient $\mathcal{L}(A)/\operatorname{LDer}(A)$. Since this linear space is finite-dimensional, it follows that both norms are equivalent, which proves the theorem. □

2. Metric versions of Posner's theorems

2.1. First Posner's theorem. The first Posner's theorem states that if R is a prime ring with characteristic different from 2, and D_1, D_2 are derivations on R such that the composition D_1D_2 is also a derivation, then either D_1 or D_2 is zero. The purpose of this section is to give a quantitative estimate of this result. Let A be a Banach algebra and let $T \in \mathcal{L}(A)$. We define a continuous bilinear map $T^\delta : A \times A \to A$ by

$$T^\delta(a,b) = T(ab) - T(a)b - aT(b) \qquad (a,b \in A).$$

Note that the constant $\|T^\delta\|$ equals $\operatorname{der}(T)$.

THEOREM 2.1.1. *Let A be a Banach algebra and let $S,T \in \mathcal{L}(A)$. then*

$$\kappa(A)^2\|S\|\|T\| \le 3\operatorname{der}(ST) + \frac{15}{2}\operatorname{der}(S)\|T\| + \frac{9}{2}\operatorname{der}(T)\|S\|.$$

PROOF. The arguments are similar to those in [**Bre91**].
For all $a,b,c \in A$ we have

$$\begin{aligned} S(a)bT(c) + T(a)bS(c) = &(ST)^\delta(ab,c) - a(ST)^\delta(b,c) \\ &- T^\delta(a,b)S(c) - S^\delta(T(ab),c) \\ &- S^\delta(a,b)T(c) - S^\delta(ab,T(c)) - S(T^\delta(ab,c)) \\ &+ aS^\delta(T(b),c) + aS^\delta(b,T(c)) + aS(T^\delta(b,c)) \end{aligned}$$

and taking norms we arrive at

$$\|S(a)bT(c) + T(a)bS(c)\| \le \big(2\|(ST)^\delta\| + 5\|S^\delta\|\|T\| + 3\|T^\delta\|\|S\|\big)\|a\|\|b\|\|c\|.$$

To shorten notation, we write $\mu = 2\|(ST)^\delta\| + 5\|S^\delta\|\|T\| + 3\|T^\delta\|\|S\|$. On account of [**Bre91**, Observation 2], we have

$$\begin{aligned} 2\,S(a)uT(b)vS(c) = &\big(S(a)uT(b) + T(a)uS(b)\big)vS(c) \\ &+ S(a)u\big(T(b)vS(c) + S(b)vT(c)\big) \\ &- \big(S(a)(uS(b)v)T(c) + T(a)(uS(b)v)S(c)\big), \end{aligned}$$

and hence $2\|S(a)uT(b)vS(c)\| \le 3\mu\|S\|\|a\|\|b\|\|c\|\|u\|\|v\|$ for all a, b, c, u, $v \in A$. This gives $\left\|M_{S(a),T(b)vS(c)}\right\| \le \frac{3}{2}\mu\|S\|\|a\|\|b\|\|c\|\|v\|$ for all $a,b,c,v \in A$. Since $\kappa(A)\|S(a)\|\|T(b)vS(c)\| \le \left\|M_{S(a),T(b)vS(c)}\right\|$, it follows that

$$\kappa(A)\|S(a)\|\|T(b)vS(c)\| \le \frac{3}{2}\mu\|S\|\|a\|\|b\|\|c\|\|v\|$$

for all $a,b,c,v \in A$ and therefore that

$$\kappa(A)\|S(a)\|\left\|M_{T(b),S(c)}\right\| \le \frac{3}{2}\mu\|S\|\|a\|\|b\|\|c\|$$

for all $a,b,c \in A$. From $\kappa(A)\|T(b)\|\|S(c)\| \le \left\|M_{T(b),S(c)}\right\|$ we now deduce that $\kappa(A)^2\|S(a)\|\|T(b)\|\|S(c)\| \le \frac{3}{2}\mu\|S\|\|a\|\|b\|\|c\|$ for all $a,b,c \in A$ and hence that $\kappa(A)^2\|S\|^2\|T\| \le \frac{3}{2}\mu\|S\|$, which clearly establishes the theorem. □

COROLLARY 2.1.2. *Let A be a Banach algebra and let $S,T \in \mathcal{L}(A)$. Then*

$$\kappa(A)^2\min\{\|S\|,\|T\|\} \le \kappa(A)\sqrt{3\operatorname{der}(ST)} + \frac{15}{2}\operatorname{der}(S) + \frac{9}{2}\operatorname{der}(T).$$

PROOF. Of course, we can assume that $\kappa(A), \|S\|, \|T\| \ne 0$.
By applying Theorem 2.1.1 we arrive at

$$1 \le \frac{\alpha}{\|S\|\|T\|} + \frac{\beta}{\|S\|} + \frac{\gamma}{\|T\|},$$

where $\alpha = 3\,\mathrm{der}(ST)\kappa(A)^{-2}$, $\beta = \frac{15}{2}\,\mathrm{der}(S)\kappa(A)^{-2}$, and $\gamma = \frac{9}{2}\,\mathrm{der}(T)\kappa(A)^{-2}$. We now write $\lambda = \min\{\|S\|, \|T\|\}$. Then $1 \leq \frac{\alpha}{\lambda^2} + \frac{\beta}{\lambda} + \frac{\gamma}{\lambda}$ and therefore

$$\lambda^2 - (\beta+\gamma)\lambda - \alpha \leq 0.$$

This implies that

$$\lambda \leq \frac{\beta+\gamma+\sqrt{(\beta+\gamma)^2+4\alpha}}{2} \leq \beta+\gamma+\sqrt{\alpha},$$

which establishes the inequality in the corollary. □

2.2. Second Posner's theorem. A map $T\colon R \to R$ is said to be *centralizing* if

$$[T(a), a] \in Z(R) \qquad (a \in R). \tag{2.1}$$

The second Posner's theorem states that if D is a centralizing derivation on a prime ring R, then either D is zero or R is commutative. Our next concern is to give a quantitative estimate of this result. Our method is motivated by **[Mat92]**. To this end, we measure how much a linear operator T on a Banach algebra A satisfies the commuting, resp. centralizing, condition by considering the constants $\mathrm{com}(T)$ (see Subsubsection 1.2.1) and

$$\mathrm{cen}(T) = \sup\left\{\mathrm{dist}\big([T(a), a], Z(A)\big)\colon\ a \in A,\ \|a\| = 1\right\},$$

respectively. Note that as com also cen is seminorm on $\mathcal{L}(A)$ vanishing precisely on the centralizing maps. Further, we measure the commutativity of A through the constant $\chi(A)$ (see Subsubsection 1.2.2). Let us recall that $Z(A)$ is closed so that the quotient linear space $A/Z(A)$ turns into a Banach space with respect to the norm given by $\|a + Z(A)\| = \mathrm{dist}\big(a, Z(A)\big)$ $(a \in A)$.

LEMMA 2.2.1. *Let A be a Banach algebra. Then*

$$\|[a, b]\| \leq 2\,\|a + Z(A)\|\,\|b + Z(A)\|$$

for all $a, b \in A$.

PROOF. Let $a, b \in A$. For all $u, v \in Z(A)$ we have $[a, b] = [a+u, b+v]$ and so $\|[a, b]\| \leq 2\|a+u\|\|b+v\|$. By taking the infima in u and v we arrive at the claimed inequality. □

LEMMA 2.2.2. *Let A a Banach algebra and let $T \in \mathcal{L}(A)$. Then*

$$\kappa(A)\,\mathrm{com}(T)^2 \leq \big(8\,\mathrm{cen}(T) + \mathrm{der}(T)\big)\|T\|$$

PROOF. For all $a, b \in A$, we have

$$[T(a), b] + [T(b), a] = \frac{1}{2}[T(a+b), a+b] - \frac{1}{2}[T(a-b), a-b].$$

We thus get

$$\|[T(a), b] + [T(b), a] + Z(A)\| \leq 4\,\mathrm{cen}(T) \tag{2.2}$$

for all $a, b \in A$ with $\|a\| = \|b\| = 1$.

Let $a \in A$ with $\|a\| = 1$. Then

$$\begin{aligned} 4[T(a), a]^2 = 2\big[[T(a), a], T(a)\big]a + 2a\big[[T(a), a], T(a)\big] \\ - \big[[T(a), a^2] + [T(a^2), a], T(a)\big] + \big[[T^\delta(a, a), a], T(a)\big] \end{aligned}$$

and therefore

$$\begin{aligned} 4\left\|[T(a), a]^2\right\| \leq 4\left\|\big[[T(a), a], T(a)\big]\right\| \\ + \left\|\big[[T(a), a^2] + [T(a^2), a], T(a)\big]\right\| + \left\|\big[[T^\delta(a, a), a], T(a)\big]\right\|. \end{aligned}$$

From Lemma 2.2.1 and (2.2) we now deduce that

$$\begin{aligned}\big\|[T(a),a]^2\big\| &\leq 2\big\|[T(a),a]+Z(A)\big\|\|T\| \\ &\quad+\frac{1}{2}\big\|[T(a),a^2]+[T(a^2),a]+Z(A)\big\|\|T\|+\|T^\delta\|\|T\| \\ &\leq \big(4\operatorname{cen}(T)+\operatorname{der}(T)\big)\|T\|.\end{aligned}$$

For each $x\in A$ with $\|x\|=1$, we have

$$[T(a),a]x[T(a),a]=[T(a),a]^2x+[T(a),a]\big[x,[T(a),a]\big]$$

and so

$$\begin{aligned}\|[T(a),a]x[T(a),a]\| &\leq \big\|[T(a),a]^2x\big\|+\big\|[T(a),a]\big[x,[T(a),a]\big]\big\| \\ &\leq \big(4\operatorname{cen}(T)+\operatorname{der}(T)\big)\|T\|+\|[T(a),a]\|\,2\,\big\|[T(a),a)]+Z(A)\big\| \\ &\leq \big(8\operatorname{cen}(T)+\operatorname{der}(T)\big)\|T\|.\end{aligned}$$

We thus get $\big\|M_{[T(a),a],[T(a),a]}\big\|\leq\big(8\operatorname{cen}(T)+\operatorname{der}(T)\big)\|T\|$ and hence

$$\kappa(A)\|[T(a),a]\|^2\leq\big(8\operatorname{cen}(T)+\operatorname{der}(T)\big)\|T\|.$$

Taking the infimum in a we finally obtain the inequality in the lemma. □

THEOREM 2.2.3. *Let A be a Banach algebra and let $T\in\mathcal{L}(A)$. Then*

$$\kappa(A)^2\chi(A)\|T\|\leq 36\ \operatorname{com}(T)+\frac{9}{2}\ \operatorname{der}(T)\chi(A)$$

and

$$\kappa(A)^{5/2}\chi(A)\|T\|\leq 36\big(8\operatorname{cen}(T)+\operatorname{der}(T)\big)^{1/2}\|T\|^{1/2}+\frac{9}{2}\kappa(A)^{1/2}\operatorname{der}(T)\chi(A).$$

PROOF. Let $a,b\in A$ with $\|a\|=\|b\|=1$. We write $\operatorname{ad}(a)$ for the inner derivation on A implemented by a, i.e. $\operatorname{ad}(a)(x)=[a,x]$ for each $x\in A$. Since $\big(-\operatorname{ad}(a)T+\operatorname{ad}(T(a))\big)(b)=\frac{1}{2}[T(a+b),a+b]-\frac{1}{2}[T(a-b),a-b]$, it follows that $\|\operatorname{ad}(a)T-\operatorname{ad}(Ta))\|\leq 4\operatorname{com}(T)$, and consequently $\operatorname{dist}\big(\operatorname{ad}(a)T,\operatorname{Der}(A)\big)\leq 4\operatorname{com}(T)$. On account of [**AEV11**, Proposition 2.2], we have

$$\operatorname{der}\big(\operatorname{ad}(a)T\big)\leq 3\operatorname{dist}\big(\operatorname{ad}(a)T,\operatorname{Der}(A)\big)\leq 12\operatorname{com}(T)$$

and Theorem 2.1.1 now yields

$$\kappa(A)^2\|\operatorname{ad}(a)\|\|T\|\leq 36\operatorname{com}(T)+\frac{9}{2}\operatorname{der}(T)\|\operatorname{ad}(a)\|.$$

Taking the supremum in a we arrive at the first inequality in the theorem. From this inequality together with Lemma 2.2.2 we get the second inequality in the theorem. □

COROLLARY 2.2.4. *Let A be a Banach algebra and let $T\in\mathcal{L}(A)$. Then*

$$\kappa(A)^2\min\{\chi(A),\|T\|\}\leq\frac{9}{2}\operatorname{der}(T)+6\kappa(A)\sqrt{\operatorname{com}(T)}$$

and

$$\kappa(A)^{5/4}\min\{\chi(A),\|T\|^{1/2}\}\leq\sqrt{36\big(8\operatorname{cen}(T)+\operatorname{der}(T)\big)^{1/2}+\frac{9}{2}\kappa(A)^{1/2}\operatorname{der}(T)}.$$

PROOF. Of course, we can assume that $\kappa(A),\chi(A),\|T\|\neq 0$.

By applying the first inequality in Theorem 2.2.3 we arrive at

$$1\leq\frac{\alpha}{\chi(A)\|T\|}+\frac{\beta}{\|T\|},$$

where $\alpha = 36\operatorname{com}(T)\kappa(A)^{-2}$ and $\beta = \frac{9}{2}\operatorname{der}(T)\kappa(A)^{-2}$. Write $\lambda = \min\{\chi(A), \|T\|\}$. Then $1 \leq \frac{\alpha}{\lambda^2} + \frac{\beta}{\lambda}$ and therefore $\lambda^2 - \beta\lambda - \alpha \leq 0$, which implies that

$$\lambda \leq \frac{\beta + \sqrt{\beta^2 + 4\alpha}}{2} \leq \beta + \sqrt{\alpha}$$

and this gives the first inequality in the corollary.

We now apply the second inequality in Theorem 2.2.3 to get

$$1 \leq \frac{\alpha}{\chi(A)\|T\|^{1/2}} + \frac{\beta}{\|T\|}$$

where $\alpha = 36\big(8\operatorname{cen}(T) + \operatorname{der}(T)\big)^{1/2}\kappa(A)^{-5/2}$ and $\beta = \frac{9}{2}\operatorname{der}(T)\kappa(A)^{-2}$. Let $\lambda = \min\{\chi(A), \|T\|^{1/2}\}$. Then $1 \leq \frac{\alpha}{\lambda^2} + \frac{\beta}{\lambda^2}$, which implies $\lambda \leq \sqrt{\alpha + \beta}$ and this proves the second inequality in the corollary. $\square$

CHAPTER 7

f-homomorphisms and f-derivations

In this short chapter we study maps which can be seen as a generalization of Lie maps. Instead of requiring that a map preserves the Lie structure we replace the commutator by an arbitrary multilinear noncommutative polynomial, and study maps which preserve the f-structure of an algebra.

In Section 1 we apply the theory of functional identities to give a description of bijective linear maps ϕ that preserve zeros of f on prime C^*-algebras of big enough dimension for some special polynomials. We prove that they are of the form $\phi = \alpha\theta + \mu 1$ for some $\alpha \in F$, an automorphism or an antiautomorphism θ, and a linear functional μ.

In Section 2 we restrict to matrix algebras $M_n(F)$ and derive the same conclusion for an arbitrary polynomial f under a technical assumption that ϕ preserves scalar matrices and $n \neq 2, 4$ using the Platonov–Đoković theory on linear algebraic groups.

In Section 3 we proceed further with this theory from which the classification of subalgebras of the Lie algebra $\mathfrak{gl}_{n^2}$ that contain all derivations of $M_n(F)$, where $n \geq 5$ and F is an algebraically closed field of characteristic 0, can be easily obtained and apply it to describe f-derivations.

This chapter is based on [**ABŠV12**, **BŠ12b**].

1. Maps preserving zeros of a polynomial on C^*-algebras

Let $\phi : A \to A$ be a linear map, and f a multilinear noncommutative polynomial. We say that ϕ *preserves zeros of* f if $f(\phi(a_1), \ldots, \phi(a_d)) = 0$ for every $(a_1, \ldots, a_d)$ satisfying $f(a_1, \ldots, a_d) = 0$.

In this subsection we will consider some special multilinear polynomials

$$f(x_1, x_2, \ldots, x_d) = \sum_{\sigma \in S_d} \lambda_\sigma x_{\sigma(1)} x_{\sigma(2)} \ldots x_{\sigma(d)} \tag{1.1}$$

for which maps preserving zeros of f can be handled in rather general classes of algebras. Specifically, we will consider polynomials f satisfying one of the following conditions:

$$\frac{\partial^{d-1} f}{\partial x_2 \partial x_3 \ldots \partial x_d} \neq 0, \tag{A}$$

$$\frac{\partial^{d-1} f}{\partial x_2 \partial x_3 \ldots \partial x_d} = 0 \quad \text{and} \quad \frac{\partial^{d-2} f}{\partial x_3 \partial x_4 \ldots \partial x_d} \neq 0, \tag{B}$$

$$\frac{\partial^{d-1} f}{\partial x_2 \partial x_3 \ldots \partial x_d} = 0 \quad \text{and} \quad f(x, x, \ldots, x, y) \neq 0. \tag{C}$$

The conditions (B) and (C) are independent. For example,

$$x_1x_2x_3 - x_1x_3x_2 + x_2x_3x_1 - x_2x_1x_3 + x_3x_1x_2 - x_3x_2x_1$$

satisfies (B) and does not satisfy (C), while

$$x_1(x_2x_3 - x_3x_2) - (x_2x_3 - x_3x_2)x_1$$

satisfies (C) and does not satisfy (B).

1.1. Polynomials satisfying (A)**.** We begin with an elementary lemma.

LEMMA 1.1.1. *Let f be a multilinear polynomial satisfying* (A)*. Suppose A is a unital algebra and $\phi : A \to A$ is a linear map preserving zeros of f and satisfying $\phi(1) \in F^*1$. If $a, b \in A$ are such that $ab = ba = 0$, then $\phi(a)\phi(b) + \phi(b)\phi(a) = 0$.*

PROOF. Without loss of generality we may assume that $\phi(1) = 1$. Namely, if $\phi(1) = \lambda 1$ with $0 \neq \lambda \in F$, then we can replace ϕ by $\lambda^{-1}\phi$ which also preserves zeros of f and does map 1 into 1.

From $ab = ba = 0$ we infer

$$f(a, b, 1, \dots, 1) = f(b, a, 1, \dots, 1) = 0,$$

and hence

$$f(\phi(a), \phi(b), 1, \dots, 1) = f(\phi(b), \phi(a), 1, \dots, 1) = 0.$$

We write f as in (1.1). Note that (A) simply means that

$$\lambda := \sum_{\sigma \in S_d} \lambda_\sigma \neq 0.$$

Since

$$\begin{aligned} &f(\phi(a), \phi(b), 1, \dots, 1) + f(\phi(b), \phi(a), 1, \dots, 1)\\ =&\sum_{\sigma^{-1}(1)<\sigma^{-1}(2)} \lambda_\sigma \phi(a)\phi(b) + \sum_{\sigma^{-1}(2)<\sigma^{-1}(1)} \lambda_\sigma \phi(b)\phi(a)\\ +&\sum_{\sigma^{-1}(1)<\sigma^{-1}(2)} \lambda_\sigma \phi(b)\phi(a) + \sum_{\sigma^{-1}(2)<\sigma^{-1}(1)} \lambda_\sigma \phi(a)\phi(b)\\ =&\lambda\Big((\phi(a)\phi(b) + \phi(b)\phi(a)\Big), \end{aligned}$$

it follows that $\phi(a)\phi(b) + \phi(b)\phi(a) = 0$. $\square$

Recall that a *Jordan epimorphism* on an algebra A is a surjective linear map θ satisfying $\theta(a^2) = \theta(a)^2$ for every $a \in A$.

THEOREM 1.1.2. *Let f be a multilinear polynomial of degree $d \geq 2$ satisfying* (A)*, and let A be a unital C^*-algebra. If a continuous surjective linear map $\phi : A \to A$ preserves zeros of f and satisfies $\phi(1) \in \mathbb{C}^* \cdot 1$, then ϕ is a scalar multiple of a Jordan epimorphism.*

PROOF. The conclusion of Lemma 1.1.1 makes it possible for us to directly apply [**ABEV10**, Theorem 3.3]. The statement of this theorem together with a well-known fact that Jordan epimorphisms preserve unities [**JR50**, Corollary 3, p. 482] immediately gives the desired conclusion. $\square$

COROLLARY 1.1.3. *Assume the conditions of Theorem 1.1.2. If A is a prime algebra, then ϕ is a scalar multiple of either an epimorphism or an antiepimorphism.*

PROOF. If A is prime, then epimorphisms or antiepimorphisms are the only Jordan epimorphisms by Herstein's theorem [**Her56**]. $\square$

1.2. Polynomials satisfying (B)**.** The treatment of (B) is similar to that of (A).

LEMMA 1.2.1. *Let f be a multilinear polynomial satisfying* (B)*. Suppose A is a unital algebra and $\phi : A \to A$ is a linear map preserving zeros of f and satisfying $\phi(1) \in F^*1$. If $a, b \in A$ are such that $ab = ba = 0$, then $\phi(a)\phi(b) = \phi(b)\phi(a)$.*

PROOF. We can reword (B) as

$$\sum_{\sigma\in S_d}\lambda_\sigma = 0 \quad\text{and}\quad \lambda := \sum_{\sigma^{-1}(1)<\sigma^{-1}(2)}\lambda_\sigma \neq 0.$$

Therefore

$$\sum_{\sigma^{-1}(2)<\sigma^{-1}(1)}\lambda_\sigma = -\lambda.$$

We may assume, for the same reason as in the proof of Lemma 1.1.1, that $\phi(1) = 1$. If $a, b \in A$ are such that $ab = ba = 0$, then

$$f(a, b, 1, \dots, 1) = 0,$$

and hence

$$f(\phi(a), \phi(b), 1, \dots, 1) = 0.$$

Since

$$f(\phi(a), \phi(b), 1, \dots, 1) = \sum_{\sigma^{-1}(1)<\sigma^{-1}(2)}\lambda_\sigma\phi(a)\phi(b) + \sum_{\sigma^{-1}(2)<\sigma^{-1}(1)}\lambda_\sigma\phi(b)\phi(a),$$

it follows that $\lambda\Big(\phi(a)\phi(b) - \phi(b)\phi(a)\Big) = 0$, i.e., $\phi(a)$ and $\phi(b)$ commute. □

THEOREM 1.2.2. *Let f be a multilinear polynomial of degree $d \geq 2$ satisfying* (B), *and let A be a unital prime C^*-algebra that is not isomorphic to $M_2(\mathbb{C})$. If a continuous bijective linear map $\phi : A \to A$ preserves zeros of f and satisfies $\phi(1) \in \mathbb{C}^* \cdot 1$, then there exist $\alpha \in \mathbb{C}$, an automorphism or an antiautomorphism θ of A, and a linear functional μ on A such that $\phi(a) = \alpha\theta(a) + \mu(a)1$ for all $a \in A$.*

PROOF. Lemma 1.2.1 makes it possible for us to apply [**ABEV10**, Corollary 3.6], which immediately gives the result. □

1.3. Polynomials satisfying (C)**.** The condition (C) means that there exist $\lambda_1, \dots, \lambda_d \in F$, not all zero, such that

$$\sum_{i=1}^{d}\lambda_i = 0 \quad\text{and}\quad f(x, x, \dots, x, y) = \sum_{i=1}^{d}\lambda_i x^{d-i} y x^{i-1}.$$

The simplest case where $f = x_1x_2 - x_2x_1$ was considered in [**Bre93b**, Theorem 2]. This result was one of the earliest applications of functional identities. Incidentally, [**Bre93b**, Theorem 2] was used in the proof of [**ABEV10**, Corollary 3.6], and therefore indirectly also in the proof of Theorem 1.2.2. What we would now like to show is that using the advanced theory of functional identities one can handle, in a more or less similar fashion, a more general situation where f satisfies (C).

THEOREM 1.3.1. *Let f be a multilinear polynomial of degree $d \geq 2$ satisfying* (C), *let $\operatorname{char}(F) \neq 2, 3$, and let A be a centrally closed prime F-algebra with $\dim_F A > d^2$. If a bijective linear map $\phi : A \to A$ preserves zeros of f, then there exist $\alpha \in F$, an automorphism or an antiautomorphism θ of A, and a linear functional μ on A such that $\phi(a) = \alpha\theta(a) + \mu(a)1$ for all $a \in A$.*

PROOF. As $f(x, x, \dots, x, x^2)$ is obviously 0 if f satisfies (C), we have

$$f(\phi(a), \phi(a), \dots, \phi(a), \phi(a^2)) = 0$$

for all $a \in A$, i.e.,

$$\sum_{i=1}^{d}\lambda_i\phi(a)^{d-i}\phi(a^2)\phi(a)^{i-1} = 0.$$

A complete linearization of this identity leads to a situation where [**BCM07**, Theorem 4.13] is applicable under suitable assumptions on A and ϕ. In view of [**BCM07**, Theorems 5.11 and C.2], these assumptions are fulfilled in our case since ϕ is surjective and $\dim_F A > d^2$. The conclusion is that $\phi(ab + ba)$ is a quasi-polynomial. As $\mathrm{char}(F) \neq 2$, this is equivalent to the existence of $\lambda \in F$ and maps $\mu, \nu : A \to F$ (with μ linear) such that

$$\phi(a^2) = \lambda\phi(a)^2 + \mu(a)\phi(a) + \nu(a)$$

for every $a \in A$. Since ϕ is also injective and $\mathrm{char}(F) \neq 3$, the result now follows from [**Bre93b**, Theorem 2]. $\square$

Let us remind that all prime C^*-algebras are centrally closed [**AM03**, Proposition 2.2.10]. Let us also mention that infinite dimensional algebras are not excluded in Theorem 1.3.1; only algebras of "small" dimension $\leq d^2$ are.

2. Maps preserving zeros of a polynomial on $M_n(F)$

In this section we use the Platonov-Đoković theory to describe maps which preserve zeros of a noncommutative polynomial. We assume that $\mathrm{char}(F) = 0$. First we briefly survey parts of the theory needed in the proof.

2.1. The Platonov-Đoković theory. Let K be an algebraically closed field of characteristic 0. We will write M_n for $M_n(K)$. We have $M_n = M_n^0 \oplus K \cdot 1$, where 1 is the identity matrix, and M_n^0 is the space of all $x \in M_n$ with $\mathrm{tr}(x) = 0$. Let O_{n^2} be the subgroup of GL_{n^2} which preserves the nondegenerate symmetric bilinear form $\mathrm{tr}(xy)$, $x, y \in M_n$. The subgroup of O_{n^2} consisting of operators which fix the identity matrix 1 will be denoted by O_{n^2-1}. The identity components of O_{n^2} and O_{n^2-1}, i.e., subgroups consisting of matrices whose determinant is 1, will be denoted by $S\mathrm{O}_{n^2}$ and SO_{n^2-1}, respectively.

By G we denote the subgroup of GL_{n^2} consisting of all similarity transformations $x \mapsto axa^{-1}$ with $a \in \mathrm{GL}_n$. Next, by P we denote the subgroup of GL_{n^2} which acts trivially on M_n^0 and M_n/M_n^0, and by Q the subgroup of GL_{n^2} which acts trivially on $K1$ and $M_n/K1$. Thus, Q consists of all transformations $x \mapsto x + \mu(x)1$, where μ is a linear functional on M_n such that $\mu(1) = 0$. Let T denote the subgroup of GL_{n^2} which acts by scalar transformations on M_n^0 and $K1$, and set $T_1 = T \cap \mathrm{SL}_{n^2}$.

By τ we denote the transposition map. However, we will write x' for the transpose of x. Note that the group $GQT\langle\tau\rangle$ consists of all invertible linear transformations $\sigma : M_n \to M_n$ that take one of the forms $\sigma(x) = \alpha axa^{-1} + \mu(x)1$ or $\sigma(x) = \alpha ax'a^{-1} + \mu(x)1$, where $\alpha \in K^*$, $a \in \mathrm{GL}_n$, and μ is a linear functional on M_n such that $\mu(1) \neq -\alpha$.

The algebra of all linear transformations on M_n can be identified with the tensor product algebra $M_n \otimes M_n^{opp}$, where M_n^{opp} is the opposite algebra of M_n, via the action $(a \otimes b)(x) = axb$, $a, b, x \in M_n$.

With respect to the notations just introduced, the following theorem can be extracted from [**PĐ95**, Theorems A and B].

THEOREM 2.1.1. **(Platonov-Đoković)** *Let Γ be a proper connected algebraic subgroup of SL_{n^2}, $n \neq 4$, containing G. Then Γ is one of the groups:*

(a) G, GQ, GT_1, GQT_1,
(b) SO_{n^2-1}, $\mathrm{SO}_{n^2-1} T_1$, $\mathrm{SO}_{n^2-1} P$, $\mathrm{SO}_{n^2-1} Q$, $\mathrm{SO}_{n^2-1} PT_1$, $\mathrm{SO}_{n^2-1} QT_1$, SL_{n^2-1}, $\mathrm{SL}_{n^2-1} T_1$, $\mathrm{SL}_{n^2-1} P$, $\mathrm{SL}_{n^2-1} Q$, $\mathrm{SL}_{n^2-1} PT_1$, $\mathrm{SL}_{n^2-1} QT_1$, $tS\mathrm{O}_{n^2}t^{-1}$ *for some* $t \in T_1$,
(c) GP, GPT_1,
(d) $t\Big(\mathrm{SL}_n \otimes \mathrm{SL}_n^{opp}\Big)t^{-1}$ *for some* $t \in T_1$.

Moreover, if Γ *is one of the groups listed in* (a)*, then its normalizer in* GL_{n^2} *is a subgroup of* $GQT\langle\tau\rangle$.

Let us point out that all groups listed in (b) contain SO_{n^2-1}. For $t\mathrm{SO}_{n^2}t^{-1}$, $t \in T_1$, this can be easily checked, while for others this is entirely obvious. Conversely, only the groups from (b) contain SO_{n^2-1}.

2.2. Zeros of f preserving maps and f-homomorphisms. Let $f \in F\langle X\rangle$ be a nonzero multilinear polynomial of degree d. Our goal is to show that under suitable assumptions a linear map $\phi : M_n(F) \to M_n(F)$ that preserves zeros of f is of the standard form; i.e.,

$$\phi(x) = \alpha axa^{-1} + \mu(x)1 \ \text{ or } \ \phi(x) = \alpha ax'a^{-1} + \mu(x)1, \tag{2.1}$$

where $\alpha \in F^*$, $a \in GL(n, F)$, and μ is a linear functional on $M_n(F)$ such that $\mu(1) \neq -\alpha$.

If d was $\geq 2n$, then, by the Amitsur-Levitzki theorem, f could be a polynomial identity, making the assumption that ϕ preserves zeros of f meaningless. We will therefore assume that $d < 2n$. Further, we will assume that $n \neq 2, 4$. It is well-known that the $n = 2$ case must be excluded when dealing with the polynomial $x_1x_2-x_2x_1$. On the other hand, it seems possible that that the exclusion of $n = 4$ is unnecessary. We need it in order to apply Theorem 2.1.1. Another assumption that we have to require is that $\mathrm{char}(F) = 0$. This one is also used because of applying Theorem 2.1.1 and is possibly redundant. Further, we will assume that ϕ is bijective. This is a usual and certainly necessary assumption in this context (cf. [**BŠ06**] that deals with the polynomial $x_1x_2 - x_2x_1$ without assuming bijectivity). Finally, we will assume that $\phi(1) \in F \cdot 1$; the (un)necessity of this assumption will be discussed in the next subsection.

Let us make a few comments and introduce some notations. We are going to consider a bijective linear map ϕ on $M_n(F)$ that preserves the set of zeros of f,

$$S_F = \{(a_1, \dots, a_d) \in M_n(F)^d \ |f(a_1, \dots, a_d) = 0\}.$$

This is an algebraic set. Indeed, considering S_F as a subset of $(F^{n^2})^d$, it is equal to the vanishing set of polynomials $\{f((x^1_{ij}), \dots, (x^d_{ij}))_{st}| \ 1 \leq s, t \leq n\}$. Using [**PĐ95**, Lemma 3] (or [**Dix77**, Lemma 1]) it can be therefore deduced that ϕ^{-1} also preserves S_F. In particular, we have $\phi(S_F) = S_F$. Accordingly, the set of all linear maps satisfying this condition is an algebraic group. The goal of our theorem is to describe those of its elements that also preserve scalar matrices.

By K we now denote an algebraic closure of F. Since $\phi \in \mathrm{GL}(n^2, F) \subseteq \mathrm{GL}_{n^2}$ (= $\mathrm{GL}(n^2, K)$) preserves S_F, it also preserves its Zariski closure S in K^{n^2d}. This is an algebraic set, and therefore, by the same argument as above,

$$\widetilde{G} = \{\psi \in \mathrm{GL}_{n^2} \mid \psi(S) \subseteq S\}$$

is a group (and $\psi(S) = S$ for every $\psi \in \widetilde{G}$). By M we denote the (algebraic) subgroup of GL_{n^2} consisting of all maps that preserve scalar matrices. Thus ϕ is contained in the algebraic group $\widetilde{G} \cap M$.

For every algebraic group L defined over F we denote by L_F the group of F-rational points of L. We have $(GQT)_F = G_FQ_FT_F$ and $(GQT\langle\tau\rangle)_F$ consists of elements in $\mathrm{GL}(n^2, F)$ that are of the form (2.1); cf. [**PĐ95**, p. 176]. Thus, if one can establish that

$$\widetilde{G} \cap M \subseteq GQT\langle\tau\rangle, \tag{2.2}$$

then ϕ, which is defined over F, lies in $G_FQ_FT_F\langle\tau\rangle$ and is therefore of the standard form (2.1).

Note that S_F is invariant under the G_F-action given by

$$g \cdot (a_1, \ldots, a_d) := (g(a_1), \ldots, g(a_d)).$$

Hence its K-closure S is also invariant under G_F, so that $G_F \subseteq \widetilde{G}$. Since $\operatorname{char}(F) = 0$ and G is connected, the rational points G_F are Zariski-dense in G [**Bor91**, Corollary 18.3]. From this one infers that $G = \overline{G_F} \subseteq \widetilde{G}$; moreover, $G \subseteq \widetilde{G} \cap M$. In a similar fashion, by first noticing that S_F is closed under multiplication by nonzero scalars in F we see that $\widetilde{G} \cap M$ is closed under multiplication by nonzero scalars in K; that is, if $a \in \widetilde{G}$ and $\lambda \in K^*$, then $\lambda a \in \widetilde{G} \cap M$.

Let us also mention that if H is an arbitrary algebraic group, then its identity component (i.e., the connected component with respect to Zariski topology that contains the identity) is also an algebraic group, and moreover, it is a normal subgroup of H [**Bor91**, Proposition 1.2].

We now have enough information to describe maps that preserve zeros of a polynomial.

THEOREM 2.2.1. *Let F be a field with* $\operatorname{char}(F) = 0$*, let $f \in F\langle X\rangle$ be a multilinear polynomial of degree $d \geq 2$, and let $\phi : M_n(F) \to M_n(F)$ be a bijective linear map that preserves zeros of f and satisfies $\phi(1) \in F \cdot 1$. Assume that $n \neq 2, 4$ and $d < 2n$. Then ϕ is of the standard form* (2.1).

PROOF. As noticed above, it suffices to establish (2.2). We claim that it is enough to prove

$$\mathrm{SO}_{n^2-1} \not\subseteq \widetilde{G}. \tag{2.3}$$

Indeed, assume (2.3) holds. Consider $H = (\widetilde{G} \cap M) \cap \mathrm{SL}_{n^2}$ and let H_1 be the identity component of H. Then H_1 is an algebraic group, it is connected, and, since $G \subseteq \widetilde{G} \cap M$, it contains G. Therefore H_1 is one of the groups listed in Theorem 2.1.1. As $H_1 \subseteq \widetilde{G} \cap M$ and (2.3) holds, we may exclude the possibilities listed in (b). Furthermore, as the groups from (c) and (d) are not contained in M, H_1 must be one of the groups listed in (a). Theorem 2.1.1 now tells us that the normalizer of H_1 in GL_{n^2} is a subgroup of $GQT\langle\tau\rangle$. Since H_1 is a normal subgroup of H it follows that H is contained in $GQT\langle\tau\rangle$. Now pick $\alpha \in \widetilde{G} \cap M$. As mentioned above, $\widetilde{G} \cap M$ is closed under multiplication by nonzero scalars. Therefore $\det(\alpha)^{\frac{-1}{n}}\alpha \in H \subseteq GQT\langle\tau\rangle$. As $GQT\langle\tau\rangle$ is also closed under multiplication by nonzero scalars it follows that $\alpha = \det(\alpha)^{\frac{1}{n}}\big(\det(\alpha)^{\frac{-1}{n}}\alpha\big) \in GQT\langle\tau\rangle$. This proves (2.2).

Thus, let us prove (2.3). Assume first that d is an even number. Set $k = \frac{d}{2} + 1$ and note that $k \leq n$. Consider the sequence of d matrix units

$$e_{11},\ e_{12},\ e_{22},\ e_{23},\ e_{33},\ e_{34}, \ldots, e_{k-1,k-1},\ e_{k-1,k}. \tag{2.4}$$

The product of these matrices in an arbitrary order except in the given one is equal to zero. Therefore, for an appropriate permutation $(a_1, \ldots, a_d)$ of the matrices (2.4) (corresponding to a nonzero coefficient of f) we have $f(a_1, \ldots, a_d) \neq 0$. Now define a linear transformation θ on $M_n(K)$ according to

$$\theta(e_{12}) = e_{21},\ \theta(e_{21}) = e_{12},\ \theta(e_{11}) = e_{33},\ \theta(e_{33}) = e_{11},$$

and θ fixes all other matrix units. A bit tedious but straightforward verification shows that θ lies in SO_{n^2-1}. Now, θ maps the matrices from (2.4) into the matrices

$$e_{33},\ e_{21},\ e_{22},\ e_{23},\ e_{11},\ e_{34}, \ldots, e_{k-1,k-1},\ e_{k-1,k}.$$

Their product in an arbitrary order is 0, so that $f(\theta(a_1), \ldots, \theta(a_d)) = 0$. This implies that $\theta \notin \widetilde{G}$. Namely, if θ was in $\widetilde{G}$ then θ^{-1} would map S_F into S which is contained in the set of zeros of f. Thus (2.3) is proved in this case.

The case where d is odd requires only minor modifications. One has to consider the matrix units

$$e_{11},\ e_{12},\ e_{22},\ e_{23},\ \ldots,\ e_{k-1,k},\ e_{k,k},$$

where $k = \frac{d+1}{2} \leq n$, and then follow the above argument. $\square$

Let $f = f(x_1, \ldots, x_d)$ be a noncommutative multilinear polynomial. A linear map ϕ from an algebra A into itself is said to be an *f-homomorphism* if

$$\phi\big(f(a_1, \ldots, a_d)\big) = f(\phi(a_1), \ldots, \phi(a_d))$$

for all $a_1, \ldots, a_d \in A$.

If ϕ is an f-homomorphism, then it clearly preserve zeros of f. If $A = M_n(F)$ and ϕ is bijective then it is very close to an isomorphism by the above theorem.

COROLLARY 2.2.2. *Let $n \neq 2, 4$, $d < 2n$, and let f be a noncommutative polynomial of degree d. If $\phi : M_n(F) \to M_n(F)$ is an f-homomorphism and $\phi(1) \in F1$, then f is of the form* (2.1).

2.3. Preserving scalar matrices. It seems plausible that the assumption from Theorem 2.2.1 that $\phi(1) \in F \cdot 1$ can be removed. To this end one should examine carefully the groups from (c) and (d). Since this would require a detailed and tedious analysis making the proof much lengthier, we omit it.

Let us now rather restrict to polynomials from the subalgebra $F\langle X\rangle_0$ of $F\langle X\rangle$ generated by 1 and all polynomials of the form $[x_{k_1}, [x_{k_2}, \ldots, [x_{k_{r-1}}, x_{k_r}] \ldots]]$. In other words, $F\langle X\rangle_0$ is the subalgebra generated by 1 and all Lie polynomials of degree ≥ 2. It is easy to see that $\frac{\partial f}{\partial x_i}$ is always 0 if $f \in F\langle X\rangle_0$. Moreover, if $\mathrm{char}(F) = 0$, then this property is characteristic for elements from $F\langle X\rangle_0$ [**Ger98**, Proposition 3]. Polynomials from $F\langle X\rangle_0$ might be of special interest in view of [**DD10**].

For these polynomials the argument based on the Platonov-Đoković theory is rather short. However, we will use an alternative approach, based on the following elementary lemma which is perhaps of independent interest.

LEMMA 2.3.1. *Let $f \in F\langle X\rangle$, where F is an arbitrary field, be a multilinear polynomial of degree d. Let $n \geq 2$ be such that $d < 2n$. If $c \in M_n(F)$ satisfies*

$$f(c, a_2, \ldots, a_d) = f(a_1, c, a_3, \ldots, a_d) = \ldots = f(a_1, \ldots, a_{d-1}, c) = 0 \tag{2.5}$$

for all $a_1, \ldots, a_d \in M_n(F)$, then $c \in F \cdot 1$.

PROOF. Pick an arbitrary rank one idempotent $e \in M_n(F)$. Then the algebra $(1-e)M_n(F)(1-e)$ is isomorphic to $M_{n-1}(F)$, so it contains matrix units h_{ij}, $1 \leq i, j \leq n-1$, i.e., elements satisfying $h_{ij}h_{kl} = \delta_{jk}h_{il}$ and $\sum_{k=1}^{n-1} h_{kk} = 1 - e$.

Without loss of generality we may assume that $x_1x_2 \ldots x_d$ is a monomial of f. We set $(s,t) := (\frac{d}{2} - 1, \frac{d}{2})$ if d is even and $(s,t) := (\frac{d-1}{2}, \frac{d-1}{2})$ if d is odd. In any case we have $t \leq n-1$. Examining all possible monomials of f one easily notices that

$$e \cdot f(e, c, h_{11}, h_{12}, h_{22}, h_{23}, \ldots, h_{st}) \cdot h_{t1} = ech_{11}.$$

Since $f(e, c, h_{11}, h_{12}, h_{22}, h_{23}, \ldots, h_{st}) = 0$ by our assumption, we thus have $ech_{11} = 0$. Similarly, by permuting the h_{ij}'s, we see that $ech_{kk} = 0$ for every k. Accordingly, $ec(1-e) = 0$. In a similar fashion, by using $f(a_1, \ldots, a_{d-2}, c, a_d) = 0$, we get $(1-e)ce = 0$. Hence it follows that c commutes with every rank one idempotent e. But then $c \in F \cdot 1$. $\square$

COROLLARY 2.3.2. *Let F be a field with $\mathrm{char}(F) = 0$, let $f \in F\langle X\rangle_0$ be a multilinear polynomial of degree $d \geq 2$, and let $\phi : M_n(F) \to M_n(F)$ be a bijective linear map that preserves zeros of f. Assume that $n \neq 2, 4$ and $d < 2n$. Then ϕ is of the standard form* (2.1).

PROOF. In view of Theorem 2.2.1 it suffices to prove that $c := \phi(1)$ lies in $F\cdot 1$. This is an immediate consequence of Lemma 2.3.1. Namely, since $f \in F\langle X\rangle_0$ we have

$$f(1, b_2, \ldots, b_d) = f(b_1, 1, b_3, \ldots, b_d) = \ldots = f(b_1, \ldots, b_{d-1}, 1) = 0$$

for all $b_i \in M_n(F)$, and hence (2.5) follows. □

3. f-derivations on $M_n(F)$

3.1. Notation. We begin by introducing some notation. In this section we assume that F is algebraically closed and that $n \geq 5$. Let us emphasize that these two assumptions will not be repeated in the statements of our results. They both are connected with the Platonov-Đoković paper [**PĐ95**]. Actually, [**PĐ95**] also deals with the situation where $n < 5$; however, this requires some extra care. For simplicity we will avoid this.

Let us recall that we denote by 1 the identity matrix in $M_n := M_n(F)$, by e_{ij} the standard matrix units in M_n, and by a' the transpose of $a \in M_n$. The Lie algebra $\mathfrak{gl}_{n^2}$ can be identified with $M_n \otimes M_n^{opp}$, where the Lie bracket is given by $[a\otimes b, c\otimes d] = (a\otimes b)(c\otimes d) - (c\otimes d)(a\otimes b) = ac\otimes db - ca\otimes bd$. In this sense $\mathfrak{g} = \{a\otimes 1 - 1\otimes a \mid a \in M_n\}$ is equal to the Lie algebra of all inner derivations of M_n (of course, all derivations on M_n are inner). We are interested in Lie subalgebras of $\mathfrak{gl}_{n^2}$ that contain $\mathfrak{g}$.

We have a direct decomposition $M_n = M_n^0 + F1$, where $M_n^0 = \{a \in M_n \mid \operatorname{tr}(a) = 0\}$. Clearly, every element in $\mathfrak{g}$ can be uniquely written as $a\otimes 1 - 1\otimes a$ with $a \in M_n^0$, and $\mathfrak{g}$ is a simple Lie algebra isomorphic to $\mathfrak{sl}_n$. We denote by $\mathfrak{gl}_{n^2-1}$ the Lie subalgebra of $\mathfrak{gl}_{n^2}$ consisting of elements that preserve the decomposition and send 1 into 0. Next we set

$$\mathfrak{sl}_{n^2-1} = \mathfrak{sl}_{n^2} \cap \mathfrak{gl}_{n^2-1},$$
$$\mathfrak{so}_{n^2} = \{a \in \mathfrak{gl}_{n^2} \mid \operatorname{tr}\big(a(x)y + xa(y)\big) = 0 \text{ for all } x, y \in M_n\},$$
$$\mathfrak{so}_{n^2-1} = \mathfrak{so}_{n^2} \cap \mathfrak{gl}_{n^2-1}.$$

Note that any Lie subalgebra of $\mathfrak{gl}_{n^2}$ containing $\mathfrak{g}$ can be considered as a $\mathfrak{g}$-module. We now state a list of simple $\mathfrak{g}$-submodules of $\mathfrak{sl}_{n^2}$, as given in [**PĐ95**, p. 170]. By ϵ_i we denote the linear functional on diagonal matrices determined by $\epsilon_i(e_{jj}) = \delta_{ij}$.

Module	Highest weight	Dimension	Highest weight vector
$\mathfrak{g}$	$\epsilon_1 - \epsilon_n$	n^2-1	$e_{1n}\otimes 1 - 1\otimes e_{1n}$
V_1	$\epsilon_1+\epsilon_2-2\epsilon_n$	$\frac{1}{4}(n^2-1)(n^2-4)$	$e_{1n}\otimes e_{2n} - e_{2n}\otimes e_{1n}$
V_2	$2\epsilon_1 - \epsilon_{n-1} - \epsilon_n$	$\frac{1}{4}(n^2-1)(n^2-4)$	$e_{1,n-1}\otimes e_{1n} - e_{1n}\otimes e_{1,n-1}$
V_3	$\epsilon_1 - \epsilon_n$	n^2-1	$\sum_i (e_{1i}\otimes e_{in} - e_{in}\otimes e_{1i})$
V_4	$\epsilon_1 - \epsilon_n$	n^2-1	$e_{1n}\otimes 1 + 1\otimes e_{1n}$
V_5	$2\epsilon_1 - 2\epsilon_n$	$\frac{1}{4}n^2(n-1)(n+3)$	$e_{1n}\otimes e_{1n}$
V_6	$\epsilon_1+\epsilon_2-\epsilon_{n-1}-\epsilon_n$	$\frac{1}{4}n^2(n+1)(n-3)$	$e_{2,n-1}\otimes e_{1n} + e_{1n}\otimes e_{2,n-1}$ $-e_{1,n-1}\otimes e_{2n} - e_{2n}\otimes e_{1,n-1}$
V_7	$\epsilon_1 - \epsilon_n$	n^2-1	$\sum_i (e_{1i}\otimes e_{in} + e_{in}\otimes e_{1i})$
V_8	0	1	$1\otimes 1 - n\sum_{i,j} e_{ij}\otimes e_{ji}$
$\mathfrak{p}$	$\epsilon_1 - \epsilon_n$	n^2-1	$\sum_i e_{1i}\otimes e_{in}$
$\mathfrak{q}$	$\epsilon_1 - \epsilon_n$	n^2-1	$\sum_i e_{in}\otimes e_{1i}$
V_4'	$\epsilon_1 - \epsilon_n$	n^2-1	$n(e_{1n}\otimes 1 + 1\otimes e_{1n})$ $-2\sum_i (e_{1i}\otimes e_{in} + e_{in}\otimes e_{1i})$

By T_0 we denote the set of all diagonal matrices of the form

$$\begin{pmatrix} \alpha 1_{n^2-1} & 0 \\ 0 & \beta \end{pmatrix},$$

where $\alpha, \beta \in F$ and 1_{n^2-1} is the identity matrix in M_{n^2-1}. Its subset consisting of all such matrices with $\alpha, \beta \in F^*$ will be denoted by T. Clearly, T is a group.

The notation introduced so far is taken from [**PĐ95**]. Let us introduce another $\mathfrak{g}$-module that will appear in the course of the proof of Proposition 3.2.1 below. As we shall see, it is possible to describe it in terms of V_4, $\mathfrak{p}$, and $\mathfrak{q}$, but first we give a more explicit description. Take $\lambda \in F$, $\lambda \neq -\frac{2}{n}$, and set

$$W(\lambda) = \{x \mapsto ax + xa + \lambda \operatorname{tr}(x)a - \frac{2\lambda}{n\lambda + 2} \operatorname{tr}(xa)1 \mid a \in M_n^0\}.$$

One can verify that $\mathfrak{g} + W(\lambda)$ is a Lie algebra, and implicitly we will in fact make this verification in the proof of Proposition 3.2.1.

Let us point out that we will use both symbols $\subseteq$ and $\subset$. The latter will be of course used to denote a proper subset.

3.2. Lie subalgebras of $\mathfrak{gl}_{n^2}$ containing derivations of $M_n(F)$. This subsection rests heavily on the work by Platonov and Đoković [**PĐ95**]. We will first record several facts that are more or less explicitly stated in the proof of [**PĐ95**, Theorem A].

In addition to the notation introduced above, we let $\mathfrak{l}$ denote a Lie subalgebra of $\mathfrak{sl}_{n^2}$ that contains $\mathfrak{g}$. Therefore $\mathfrak{l}$ can be treated as a $\mathfrak{g}$-module.

(1) $\mathfrak{sl}_{n^2}$, considered as a $\mathfrak{g}$-module, can be directly decomposed into simple $\mathfrak{g}$-modules as follows:

$$\mathfrak{sl}_{n^2} = \mathfrak{g} + \sum_{i=1}^{8} V_i.$$

(2) The $\mathfrak{g}$-modules $\mathfrak{g}, V_3, V_4, V_7, \mathfrak{p}, \mathfrak{q}, V_4'$ are isomorphic, while the others from the above list are pairwise nonisomorphic.

(3) We have the following direct decompositions into simple $\mathfrak{g}$-modules:

$$\begin{aligned} \mathfrak{so}_{n^2-1} &= \mathfrak{g} + V_1 + V_2, \\ \mathfrak{so}_{n^2} &= \mathfrak{so}_{n^2-1} + V_3, \\ \mathfrak{sl}_{n^2-1} &= \mathfrak{so}_{n^2-1} + V_4' + V_5 + V_6, \\ \mathfrak{sl}_{n^2} &= \mathfrak{sl}_{n^2-1} + \mathfrak{p} + \mathfrak{q} + V_8. \end{aligned}$$

(4) $\mathfrak{so}_{n^2}$ is the Lie subalgebra of $\mathfrak{gl}_{n^2}$ consisting of all skew-symmetric tensors.

(5) If $\mathfrak{l}$ is properly contained in $\mathfrak{sl}_{n^2-1}$ and properly contains $\mathfrak{g}$, then $\mathfrak{l} = \mathfrak{so}_{n^2-1}$.

(6) If $V_5 \subset \mathfrak{l}$ or $V_6 \subset \mathfrak{l}$, then $\mathfrak{sl}_{n^2-1} \subseteq \mathfrak{l}$.

(7) If $\mathfrak{p} + \mathfrak{q} \subset \mathfrak{l}$, then $\mathfrak{l} = \mathfrak{sl}_{n^2}$.

(8) $V_4', \mathfrak{p}, \mathfrak{q}$ are pairwise nonisomorphic as $(\mathfrak{g} + V_8)$-modules.

(9) If $V_8 \subset \mathfrak{l}$ and $\mathfrak{p} + \mathfrak{q} \not\subset \mathfrak{l}$, then $\mathfrak{l} \subseteq \mathfrak{sl}_{n^2-1} + \mathfrak{p} + V_8$ or $\mathfrak{l} \subseteq \mathfrak{sl}_{n^2-1} + \mathfrak{q} + V_8$.

(10) If $V_1 \subset \mathfrak{l}$ or $V_2 \subset \mathfrak{l}$, then $\mathfrak{so}_{n^2-1} \subseteq \mathfrak{l}$.

(11) As an $\mathfrak{so}_{n^2-1}$-module, $\mathfrak{sl}_{n^2}$ is a direct sum of five simple modules, namely

$$\mathfrak{so}_{n^2-1},\ V_4' + V_5 + V_6,\ \mathfrak{p},\ \mathfrak{q},\ V_8.$$

Among them only $\mathfrak{p}$ and $\mathfrak{q}$ are isomorphic.

(12) If W is a simple submodule of $\mathfrak{p} + \mathfrak{q}$, different from $\mathfrak{p}$ and $\mathfrak{q}$, then there exists $t \in T$ such that $W = tV_3t^{-1}$.

(13) If W is a $\mathfrak{g}$-submodule of $\mathfrak{p}+\mathfrak{q}$ and $W\neq\mathfrak{p},\mathfrak{q}$, then $\mathfrak{g}+W$ is not a Lie algebra. Next, if $\mathfrak{l}\subset\mathfrak{g}+V_4+\mathfrak{p}+\mathfrak{q}$, then $\mathfrak{l}$ can not be written as a sum of three simple $\mathfrak{g}$-modules.

The $\mathfrak{g}$-modules $\mathfrak{p},\mathfrak{q},V_4$ and V_8 will play particularly prominent roles. Therefore we will now give some further comments about them, which will be used in the proof of Proposition 3.2.1 without reference.

With respect to the decomposition $M_n=M_n^0+F1$ we can represent $\mathfrak{p}$ and $\mathfrak{q}$ with matrices of the form

$$\begin{pmatrix} 0 & x \\ 0 & 0 \end{pmatrix} \text{and} \begin{pmatrix} 0 & 0 \\ x' & 0 \end{pmatrix}, \text{respectively,}$$

where $x\in F^{n^2-1}$ is an arbitrary column vector. Thus, $\mathfrak{p}$ consists of all maps of the form $x\mapsto \mathrm{tr}(x)a$ with $a\in M_n^0$, and $\mathfrak{q}$ consists of all maps of the form $x\mapsto\mathrm{tr}(xa)1$ with $a\in M_n^0$. Next, V_8 consists of scalar multiples of the diagonal matrix

$$\begin{pmatrix} 1_{n^2-1} & 0 \\ 0 & 1-n^2 \end{pmatrix},$$

and $V_4=\{a\otimes 1+1\otimes a\,|\,a\in M_n^0\}$.

As already mentioned, $\mathfrak{g},\mathfrak{p},\mathfrak{q}$, and V_4 are isomorphic simple $\mathfrak{g}$-modules. Since F is algebraically closed, up to scalar multiplication there is exactly one isomorphism from $\mathfrak{g}$ into any of modules $\mathfrak{p},\mathfrak{q}$, and V_4. One can check that these are

$$\begin{aligned} \phi_1:\mathfrak{g}\to\mathfrak{p},\ &\phi_1(a\otimes 1-1\otimes a)(x)=\mathrm{tr}(x)a,\\ \phi_2:\mathfrak{g}\to\mathfrak{q},\ &\phi_2(a\otimes 1-1\otimes a)(x)=\mathrm{tr}(xa),\\ \phi_3:\mathfrak{g}\to V_4,\ &\phi_3(a\otimes 1-1\otimes a)=a\otimes 1+1\otimes a. \end{aligned}$$

Finally, since the highest weight vector of V_4' is a linear combination of the highest weight vectors of $\mathfrak{p}$, $\mathfrak{q}$ and V_4, it easily follows that $\mathfrak{p}+\mathfrak{q}+V_4=\mathfrak{p}+\mathfrak{q}+V_4'$.

Besides the above results we also need the following well-known fact (see e.g. [**Hal03**, Exercise 17, p. 124]). Let $\mathfrak{h}$ be a Lie algebra and let V be a finite dimensional $\mathfrak{h}$-module. Assume that V is a direct sum of simple $\mathfrak{h}$-modules,

$$V=V_1^1+\cdots+V_{n_1}^1+V_1^2+\cdots+V_{n_2}^2+\cdots+V_1^k+\cdots+V_{n_k}^k,$$

where V_i^j is isomorphic to $V_{i'}^j$ for all i,i',j and nonisomorphic to V_l^k for all l and $k\neq j$. Then every simple submodule U of V is contained in $V_1^j+\cdots+V_{n_j}^j$ for some j, and there exist $\lambda_1,\dots,\lambda_{n_j}\in F$, not all zero, such that U consists of all elements of the form

$$\lambda_1 v+\lambda_2\phi_2(v)+\ldots+\lambda_{n_j}\phi_{n_j}(v),\ v\in V_1^j,$$

where $\phi_l:V_1^j\to V_l^j$ is an $\mathfrak{h}$-module isomorphism. We will use this fact frequently and without reference.

Our goal is to give a list of all Lie subalgebras of $\mathfrak{gl}_{n^2}$ that contain $\mathfrak{g}$. In the first and major step we describe those of them that are contained in $\mathfrak{sl}_{n^2}$.

Proposition 3.2.1. *If $\mathfrak{l}$ is a proper Lie subalgebra of $\mathfrak{sl}_{n^2}$ that contains $\mathfrak{g}$, then $\mathfrak{l}$ is one of the following Lie algebras:*

(i) $\mathfrak{sl}_{n^2-1},\mathfrak{sl}_{n^2-1}+\mathfrak{p},\mathfrak{sl}_{n^2-1}+\mathfrak{q},\mathfrak{sl}_{n^2-1}+V_8,\mathfrak{sl}_{n^2-1}+\mathfrak{p}+V_8,\mathfrak{sl}_{n^2-1}+\mathfrak{q}+V_8$,
(ii) $\mathfrak{so}_{n^2-1},\mathfrak{so}_{n^2-1}+\mathfrak{p},\mathfrak{so}_{n^2-1}+\mathfrak{q},\mathfrak{so}_{n^2-1}+V_8,\mathfrak{so}_{n^2-1}+\mathfrak{p}+V_8,\mathfrak{so}_{n^2-1}+\mathfrak{q}+V_8$,
(iii) $\mathfrak{g},\mathfrak{g}+\mathfrak{p},\mathfrak{g}+\mathfrak{q},\mathfrak{g}+V_8,\mathfrak{g}+\mathfrak{p}+V_8,\mathfrak{g}+\mathfrak{q}+V_8$,
(iv) $t\mathfrak{so}_{n^2}t^{-1}$ *for some* $t\in T$,
(v) $\mathfrak{g}+W(\lambda)$ *for some* $\lambda\in F$, $\lambda\neq-\frac{2}{n}$.

Proof. Consider $\mathfrak{l}$ as a $\mathfrak{g}$-module. Since $\mathfrak{g}$ is simple, $\mathfrak{l}$ is a direct sum of simple modules by Weyl's theorem. As $\mathfrak{l}$ is a $\mathfrak{g}$-submodule of $\mathfrak{sl}_{n^2}=\mathfrak{g}+\sum_{i=1}^8 V_i$ (see (1)), it follows from (2) that $\mathfrak{l}$ is equal to a sum of $\mathfrak{g}$, some of the modules V_1,V_2,V_5,V_6,V_8, and a submodule of $V_3 + V_4 + V_7$. We have to figure out when such a sum forms a Lie algebra and can be therefore equal to $\mathfrak{l}$. The proof is divided into several steps. First we consider the case where $\mathfrak{l}$ contains $\mathfrak{sl}_{n^2-1}$, which leads to the list (i). The case where $\mathfrak{sl}_{n^2-1} \not\subseteq \mathfrak{l}$ is more complex. After finding the Lie algebras from (ii) and (iii) we consider separately the situation where $\mathfrak{l}$ contains $\mathfrak{so}_{n^2-1}$ and the situation where it does not. This yields (iv) and (v), respectively.

So first suppose that $\mathfrak{l}$ contains $\mathfrak{sl}_{n^2-1}$. From (3) we see that $\mathfrak{sl}_{n^2} = \mathfrak{sl}_{n^2-1} + \mathfrak{p} + \mathfrak{q} + V_8$. Therefore $\mathfrak{l} = \mathfrak{sl}_{n^2-1} + Z$, where Z is a $\mathfrak{g}$-module contained in $\mathfrak{p} + \mathfrak{q} + V_8$. Since $\mathfrak{p}$ and $\mathfrak{q}$ are isomorphic (2), the only possibilities for Z are

$$0, \mathfrak{p}, \mathfrak{q}, V_8, \mathfrak{p} + V_8, \mathfrak{q} + V_8, \mathfrak{p} + \mathfrak{q}, W, W + V_8,$$

where W is a proper submodule of $\mathfrak{p} + \mathfrak{q}$ different from $0, \mathfrak{p}, \mathfrak{q}$. We have to find out for which of these nine choices $\mathfrak{sl}_{n^2-1} + Z$ is a Lie algebra. It is easy to check that this is true for the first six ones. On the other hand, the last three choices must be ruled out in view of (7), (12), and (9), respectively. Thus, (i) lists all proper Lie subalgebras of $\mathfrak{sl}_{n^2}$ that contain $\mathfrak{sl}_{n^2-1}$.

From now on we assume that $\mathfrak{l}$ does not contain $\mathfrak{sl}_{n^2-1}$. Note that V_5 and V_6 are not contained in $\mathfrak{l}$ due to (6). As

$$\mathfrak{sl}_{n^2} = \mathfrak{so}_{n^2-1} + V_4' + V_5 + V_6 + \mathfrak{p} + \mathfrak{q} + V_8$$

by (3), it follows that $\mathfrak{l} \subset \mathfrak{so}_{n^2-1} + V_4' + \mathfrak{p} + \mathfrak{q} + V_8$ (the strict inclusion holds because of (7)). Moreover, (10) and $\mathfrak{so}_{n^2-1} = \mathfrak{g} + V_1 + V_2$ (see (3)) imply that

$$\mathfrak{l} = \mathfrak{g} + W \quad \text{or} \quad \mathfrak{l} = \mathfrak{so}_{n^2-1} + W,$$

where W is a submodule of $V_4' + \mathfrak{p} + \mathfrak{q} + V_8$. It is easy to check that 0, $\mathfrak{p}$, $\mathfrak{q}$, V_8, $\mathfrak{p} + V_8$, $\mathfrak{q} + V_8$ are appropriate choices for W; that is, $\mathfrak{so}_{n^2-1} + W$ and $\mathfrak{g} + W$ are indeed Lie algebras in each of these cases. From now on we assume that W is different from these modules. In other words, we are assuming that $\mathfrak{l}$ is none of the Lie algebras listed in (ii) and (iii).

Assume that $V_8 \subset \mathfrak{l}$. Then we have $W = V_8 + W'$, where W' is a submodule of $V_4' + p + q$. According to (8) (and (7)), W' can be equal to one of $\mathfrak{p}, \mathfrak{q}, V_4'$ or to a sum of two of them. If $W' \in \{\mathfrak{p}, \mathfrak{q}\}$, then $\mathfrak{l}$ can be found in one of the lists (ii) or (iii). On account of (7), the only additional examples can be obtained if $V_4' \subseteq W'$. Thus, either $\mathfrak{so}_{n^2-1} + V_4' \subset \mathfrak{l}$ or $\mathfrak{g} + V_4' \subset \mathfrak{l}$. However, from (5) (and (3)) it follows that in each of these two cases $\mathfrak{l}$ contains $\mathfrak{sl}_{n^2-1}$, contrary to our assumption. Therefore $V_8 \not\subset \mathfrak{l}$.

Assume now that $\mathfrak{l}$ contains $\mathfrak{so}_{n^2-1}$, so that $\mathfrak{l} = \mathfrak{so}_{n^2-1} + W$. Since the Lie algebras from (ii) and (iii) have already been excluded, we see from (11) that W is a submodule of $\mathfrak{p} + \mathfrak{q}$ different from $0, \mathfrak{p}, \mathfrak{q}$ and isomorphic to $\mathfrak{p} \cong \mathfrak{q}$. By (12), $W = tV_3t^{-1}$ for some $t \in T$. Since $\mathfrak{so}_{n^2} = \mathfrak{so}_{n^2-1} + V_3$ (see (3)), it follows that $\mathfrak{l} = t\mathfrak{so}_{n^2}t^{-1}$. On the other hand, $t\mathfrak{so}_{n^2}t^{-1}$ indeed is a Lie algebra for every $t \in T$. We have thus arrived at the case (iv).

Finally we consider the case where $\mathfrak{l} = \mathfrak{g} + W$ with $W \subset V_4' + \mathfrak{p} + \mathfrak{q}$. As mentioned above, this is equivalent to $W \subset V_4 + \mathfrak{p} + \mathfrak{q}$. In view of (13) and (7) we may assume that W is a simple module and $W \not\subset \mathfrak{p} + \mathfrak{q}$. Accordingly, there exist $\lambda, \mu \in F$ such that W consists of all maps of the form

$$w_a : x \mapsto ax + xa + \lambda \operatorname{tr}(x)a + \mu \operatorname{tr}(xa)1, \ \text{ where } a \in M_n^0.$$

The fact that $\mathfrak{l}$ is a Lie algebra gives rise to some restrictions on λ and μ. Indeed, we have

$$[w_a, w_b](x) = [[a,b],x] + (n\lambda\mu + 2\lambda + 2\mu)\big(\operatorname{tr}(xb)a - \operatorname{tr}(xa)b\big) \in \mathfrak{l} = \mathfrak{g} + W$$

Since $x \mapsto [[a,b],x]$ lies in $\mathfrak{g} \subset \mathfrak{l}$, it follows that either $n\lambda\mu + 2\lambda + 2\mu = 0$ or $x \mapsto \operatorname{tr}(xb)a - \operatorname{tr}(xa)b$ lies in $\mathfrak{l}$ for all $a, b \in M_n^0$. The rank of this operator is at most 2. On the other hand, operators in $\mathfrak{l}$ are of the form

$$h : x \mapsto [c, x] + dx + xd + \lambda \operatorname{tr}(x)d + \mu \operatorname{tr}(xd)1 = ex + xf + \lambda \operatorname{tr}(x)d + \mu \operatorname{tr}(xd)1,$$

where $e = d + c$ and $f = d - c$. An elementary argument (see, e.g., [**Kuč74**]) shows that the rank of $x \mapsto ex + xf$ is at least n, unless $e = -f$ is a scalar matrix (and hence $h = 0$). Accordingly, nonzero elements in $\mathfrak{l}$ have rank at least $n - 2$. Since $n \geq 5$, $\mathfrak{l}$ cannot contain operators $x \mapsto \operatorname{tr}(xb)a - \operatorname{tr}(xa)b$. Therefore $n\lambda\mu + 2\lambda + 2\mu = 0$, showing that $\lambda \neq -\frac{2}{n}$ and $W = W(\lambda)$. That is, $\mathfrak{l}$ is of the form described in (v). $\square$

Now we can easily complete the last proposition to give a list of Lie subalgebras of $\mathfrak{gl}_{n^2}$ containing $\mathfrak{g}$.

THEOREM 3.2.2. *If $\mathfrak{h}$ is a proper Lie subalgebra of $\mathfrak{gl}_{n^2}$ that contains $\mathfrak{g}$, then either $\mathfrak{h} = \mathfrak{sl}_{n^2}$ or $\mathfrak{h} = \mathfrak{l} + Ft$, where $\mathfrak{l}$ is a Lie algebra from Proposition 3.2.1 and $t \in T_0$. Moreover, if $\mathfrak{l}$ is a Lie algebra from* (iv) *or* (v)*, then $t = 0$ or $t = 1$.*

PROOF. In view of Proposition 3.2.1 we may assume that $\mathfrak{h}$ is not contained in $\mathfrak{sl}_{n^2}$. Thus, $\mathfrak{l} = \mathfrak{h} \cap \mathfrak{sl}_{n^2}$ is a proper Lie subalgebra of $\mathfrak{h}$. Since $\mathfrak{sl}_{n^2}$ has codimension 1 in $\mathfrak{gl}_{n^2}$, $\mathfrak{l}$ has codimension 1 in $\mathfrak{h}$. Also, $\mathfrak{l}$ is a $\mathfrak{g}$-module, and so $\mathfrak{h} = \mathfrak{l} + U$ for some 1-dimensional $\mathfrak{g}$-submodule U of $\mathfrak{gl}_{n^2}$. In the decomposition (1) there is only one 1-dimensional module, namely V_8, so U is a submodule of $V_8 + F1$. Therefore $U = Ft$, where t is a matrix of the form $\begin{pmatrix} \alpha 1_{n^2-1} & 0 \\ 0 & \beta \end{pmatrix} \in T_0$ for some $\alpha, \beta \in F$. It is easy to see that for any choice of α and β, $\mathfrak{h} = \mathfrak{l} + Ft$ is a Lie algebra if $\mathfrak{l}$ is listed in (i), (ii), or (iii). On the other hand, a brief examination shows that α must be equal to β if $\mathfrak{l}$ is listed in (iv) or (v). $\square$

Let f be a multilinear noncommutative polynomial of degree l, $l \geq 2$. A linear map d from an algebra A into itself is said to be an *f-derivation* if

$$d\big(f(a_1, \ldots, a_l)\big) = \sum_{i=1}^{l} f(a_1, \ldots, a_{i-1}, d(a_i), a_{i+1}, \ldots, a_l) \tag{3.1}$$

for all $a_1, \ldots, a_l \in A$. Derivations are obvious examples, and the question is whether they are basically also the only possible examples. An affirmative answer has been obtained for rather general algebras (see, e.g., [**BCM07**, Section 6.5]), but, surprisingly, the case where $A = M_n$ still has not been completely settled. Under certain technical restrictions we are now in a position to handle it.

COROLLARY 3.2.3. *Let f be a multilinear polynomial of degree $l < 2n$. If $d : M_n \to M_n$ is an f-derivation such that $d(1) = 0$, then d is a derivation.*

PROOF. Note that the set $\mathfrak{h}$ of all f-derivations of M_n is a Lie subalgebra of $\mathfrak{gl}_{n^2}$ that contains $\mathfrak{g}$. Hence $\mathfrak{h}$ is one of the Lie algebras listed in Theorem 3.2.2. It suffices to show that $\mathfrak{h}$ does not contain $\mathfrak{so}_{n^2-1}$, $\mathfrak{p}$, $\mathfrak{q}$, $W(\lambda)$ and $\mathfrak{g} + Ft$ with $t \in T_0 \setminus \{0\}$, as all Lie algebras from the list, except $\mathfrak{g}$, contain at least one of these sets. The second and the fourth possibility can be cancelled out due to the initial assumption $d(1) = 0$.

We claim that $\mathfrak{so}_{n^2-1} \not\subseteq \mathfrak{h}$. Without loss of generality we may assume that $x_1 \ldots x_l$ is one of the monomials of f. Choose $e_{13} \otimes e_{22} - e_{22} \otimes e_{13} \in \mathfrak{so}_{n^2-1}$ (cf. (4)). If $l = 2k-2$ (resp. $l = 2k-1$), take $(x_1, \ldots, x_n) = (e_{11}, e_{32}, e_{22}, e_{23}, \ldots, e_{k-1,k})$ (resp. $(e_{11}, e_{32}, e_{22}, e_{23}, \ldots, e_{k,k})$). Then observe that in this case the left-hand side of (3.1) differs from its right-hand side. This proves our claim.

The task now is to exclude the case $\mathfrak{q} \subset \mathfrak{h}$. Note that $\sum_{i=1}^{n} e_{i1} \otimes e_{ki} \in \mathfrak{q}$ with $(x_1, \ldots, x_n) = (e_{11}, e_{12}, e_{22}, e_{23}, \ldots, e_{k-1,k})$ (resp. $(e_{11}, e_{12}, e_{22}, e_{23}, \ldots, e_{k,k})$), depending on the parity of l, does the trick.

We are reduced to proving that $t \in T_0 \setminus \{0\}$ cannot belong to $\mathfrak{h}$. We can restrict ourselves to the case where t acts as a scalar multiple of the identity on M_n^0 and sends 1 to 0. Now choose a maximal subset S of $\mathbb{N}_l = \{1, \ldots, l\}$ such that the polynomial $f(y_1, \ldots, y_l)$, where $y_i = 1$ if $i \in S$ and $y_i = x_i$ if $i \notin S$, is not zero (the case where $S = \emptyset$ is not excluded). Since the degree of $f(y_1, \ldots, y_l)$ is less than $2n$, this polynomial is not an identity of M_n. Therefore there exist $a_1, \ldots, a_l \in M_n$ such that $a_i = 1$ if $i \in S$ and $f(a_1, \ldots, a_l) \neq 0$. Moreover, because of the maximality assumption we may assume that $a_i \in M_n^0$ whenever $i \notin S$. Note that (3.1) yields

$$f(a_1, \ldots, a_l) - \frac{1}{n}\operatorname{tr}\big(f(a_1, \ldots, a_l)\big)1 = (l - s)f(a_1, \ldots, a_l),$$

where $s = |S|$. This is possible only when $l - s = 0$ or $l - s = 1$. Actually, from the definition of S it is clear that the last possibility cannot occur. Therefore $l = s$. Considering $f(a, \ldots, a)$ for an arbitrary $a \in M_n^0$ we easily derive a contradiction. $\square$

Some of the Lie algebras from Theorem 3.2.2 (and Proposition 3.2.1) are also closed under the associative product, and are therefore associative algebras. In the next corollary we will list all of them. Although the symbols such as $\mathfrak{gl}$ etc. are traditionally reserved for Lie algebras, we will slightly abuse the notation and consider them as associative algebras.

Corollary 3.2.4. *If A is a proper associative subalgebra of $\mathfrak{gl}_{n^2}$ that contains $\mathfrak{g}$, then A is either*

$$\mathfrak{gl}_{n^2-1},\ \mathfrak{gl}_{n^2-1} + \mathfrak{p},\ \mathfrak{gl}_{n^2-1} + \mathfrak{q},$$
$$\mathfrak{gl}_{n^2-1} + Ft,\ \mathfrak{gl}_{n^2-1} + \mathfrak{p} + Ft,\ \text{ or } \mathfrak{gl}_{n^2-1} + \mathfrak{q} + Ft$$

for some $t \in T$.

Proof. All we have to do is to find out which of the Lie algebras from Theorem 3.2.2 are closed under the associative product. Take elements $e_{12} \otimes 1 - 1 \otimes e_{12}, e_{34} \otimes 1 - 1 \otimes e_{34} \in \mathfrak{g}$. Their product $u = -e_{12} \otimes e_{34} - e_{34} \otimes e_{12}$ preserves the decomposition $M_n = M_n^0 + F1$ and has zero trace, thus it lies in $A \cap \mathfrak{gl}_{n^2-1} \cap \mathfrak{sl}_{n^2}$. Note that $\mathfrak{l} \cap \mathfrak{gl}_{n^2-1} \cap \mathfrak{sl}_{n^2} = \mathfrak{l} \cap \mathfrak{sl}_{n^2-1}$ for $\mathfrak{l}$ listed in Proposition 3.2.1 (i), (ii), (iii), (iv), (v) is equal to $\mathfrak{sl}_{n^2-1}, \mathfrak{so}_{n^2-1}, \mathfrak{g}, \mathfrak{so}_{n^2-1}, \mathfrak{g}$, respectively. But u lies neither in $\mathfrak{so}_{n^2-1}$ nor in $\mathfrak{g}$. Hence $\mathfrak{sl}_{n^2-1} \subset A$. Therefore A also contains $\mathfrak{gl}_{n^2-1}$, which is the associative algebra generated by $\mathfrak{sl}_{n^2-1}$. All Lie algebras from Theorem 3.2.2 that contain $\mathfrak{gl}_{n^2-1}$ are indeed associative algebras. These are the algebras listed in the statement of the corollary. $\square$

Bibliography

[ABE+13] J. Alaminos, M. Brešar, J. Extremera, Š. Špenko, and A. R. Villena, *Determining elements in C^*-algebras through spectral properties*, J. Math. Anal. Appl. **405** (2013), no. 1, 214–219.

[ABE+14] ———, *Derivations preserving quasinilpotent elements*, Bull. Lond. Math. Soc. **46** (2014), no. 2, 379–384.

[ABEV09] J. Alaminos, M. Brešar, J. Extremera, and A. R. Villena, *Maps preserving zero products*, Studia Math. **193** (2009), no. 2, 131–159.

[ABEV10] J. Alaminos, J. M. Brešar, J. Extremera, and A. R. Villena, *Characterizing Jordan maps on C^*-algebras through zero products*, Proc. Edinb. Math. Soc. (2) **53** (2010), no. 3, 543–555.

[ABŠV12] J. Alaminos, M. Brešar, Š. Špenko, and A. R. Villena, *Maps preserving zeros of a polynomial*, Linear Algebra Appl. **436** (2012), no. 7, 2504–2512.

[AD03] D. Alpay and C. Dubi, *A realization theorem for rational functions of several complex variables*, Systems Control Lett. **49** (2003), no. 3, 225–229.

[AEŠV12a] J. Alaminos, J. Extremera, Š. Špenko, and A. R. Villena, *Metric versions of Posner's theorems*, Taiwanese J. Math. **16** (2012), no. 6, 1951–1957.

[AEŠV12b] ———, *Stability of commuting maps and Lie maps*, Studia Math. **213** (2012), no. 1, 25–48.

[AEV10] J. Alaminos, J. Extremera, and A. R. Villena, *Hyperreflexivity of the derivation space of some group algebras*, Math. Z. **266** (2010), no. 3, 571–582.

[AEV11] ———, *Approximately spectrum-preserving maps*, J. Funct. Anal. **261** (2011), no. 1, 233–266.

[AG60a] M. Auslander and O. Goldman, *The Brauer group of a commutative ring*, Trans. Amer. Math. Soc. **97** (1960), 367–409.

[AG60b] ———, *Maximal orders*, Trans. Amer. Math. Soc. **97** (1960), 1–24.

[AKV13] G. Abduvalieva and D. S. Kaliuzhnyi-Verbovetskyi, *Fixed point theorems for noncommutative functions*, J. Math. Anal. Appl. **401** (2013), no. 1, 436–446.

[AM57] A. A. Albert and B. Muckenhoupt, *On matrices of trace zeros*, Michigan Math. J. **4** (1957), 1–3.

[AM03] P. Ara and M. Mathieu, *Local multipliers of C^*-algebras*, Springer Monographs in Mathematics, Springer-Verlag London, Ltd., London, 2003.

[AM14] J. Agler and J. E. McCarthy, *The implicit function theorem and free algebraic sets*, 2014, arXiv:1404.6032.

[AM15] ———, *Global Holomorphic Functions in Several Noncommuting Variables*, Canad. J. Math. **67** (2015), no. 2, 241–285.

[Ami57] S. A. Amitsur, *A generalization of Hilbert's Nullstellensatz*, Proc. Amer. Math. Soc. **8** (1957), 649–656.

[Art69] M. Artin, *On Azumaya algebras and finite dimensional representations of rings.*, J. Algebra **11** (1969), 532–563.

[Aup91] B. Aupetit, *A primer on spectral theory*, Universitext, Springer-Verlag, New York, 1991.

[AY14] J. Agler and N. J. Young, *Symmetric functions of two noncommuting variables*, J. Funct. Anal. **266** (2014), no. 9, 5709–5732.

[BB01] K. I. Beidar and M. Brešar, *Extended Jacobson density theorem for rings with derivations and automorphisms*, Israel J. Math. **122** (2001), 317–346.

[BB09] Y. Bahturin and M. Brešar, *Lie gradings on associative algebras*, J. Algebra **321** (2009), no. 1, 264–283.

[BBCM00] K. I. Beidar, M. Brešar, M. A. Chebotar, and W. S. Martindale, III, *On functional identities in prime rings with involution. II*, Comm. Algebra **28** (2000), no. 7, 3169–3183.

[BBR09] G. Braatvedt, R. Brits, and H. Raubenheimer, *Spectral characterizations of scalars in a Banach algebra*, Bull. Lond. Math. Soc. **41** (2009), no. 6, 1095–1104.

[BBS11] Y. Bahturin, M. Brešar, and I. Shestakov, *Jordan gradings on associative algebras*, Algebr. Represent. Theory **14** (2011), no. 1, 113–129.

[BC00a] K. I. Beidar and M. A. Chebotar, *On functional identities and d-free subsets of rings. I, II*, Comm. Algebra **28** (2000), no. 8, 3925–3951, 3953–3972.

[BC00b] ——, *On Lie-admissible algebras whose commutator Lie algebras are Lie subalgebras of prime associative algebras*, J. Algebra **233** (2000), no. 2, 675–703.

[BCM07] M. Brešar, M. A. Chebotar, and W. S. Martindale, III, *Functional identities*, Frontiers in Mathematics, Birkhäuser Verlag, Basel, 2007.

[BD73] F. F. Bonsall and J. Duncan, *Complete normed algebras. Ergebnisse der mathematik und ihrer grenzgebiete, band 80*, Springer-Verlag, New York-Heidelberg, 1973.

[Bei98] K. I. Beidar, *On functional identities and commuting additive mappings*, Comm. Algebra **26** (1998), no. 6, 1819–1850.

[Ber06] G. M. Bergman, *Can one factor the classical adjoint of a generic matrix?*, Transform. Groups **11** (2006), no. 1, 7–15.

[BF99] K. I. Beidar and Y. Fong, *On additive isomorphisms of prime rings preserving polynomials*, J. Algebra **217** (1999), no. 2, 650–667.

[BFLW97] K. I. Beidar, Y. Fong, P.-H. Lee, and T.-L. Wong, *On additive maps of prime rings satisfying the Engel condition*, Comm. Algebra **25** (1997), no. 12, 3889–3902.

[BGM06] J. A. Ball, G. Groenewald, and T. Malakorn, *Bounded real lemma for structured noncommutative multidimensional linear systems and robust control*, Multidimens. Syst. Signal Process. **17** (2006), no. 2-3, 119–150.

[BK09] M. Brešar and I. Klep, *Values of noncommutative polynomials, Lie skew-ideals and tracial Nullstellensätze*, Math. Res. Lett. **16** (2009), no. 4, 605–626.

[BK11] ——, *Tracial Nullstellensätze*, Notions of positivity and the geometry of polynomials, Trends Math., Birkhäuser/Springer Basel AG, Basel, 2011, pp. 79–101.

[BK13] ——, *A local-global principle for linear dependence of noncommutative polynomials*, Israel J. Math. **193** (2013), no. 1, 71–82.

[BLVdB10] R.-O. Buchweitz, G. J. Leuschke, and M. Van den Bergh, *Non-commutative desingularization of determinantal varieties I*, Invent. Math. **182** (2010), no. 1, 47–115.

[BLVdB11] ——, *Non-commutative desingularization of determinantal varieties, II: Arbitrary minors*, 2011, arXiv:1106.1833.

[BM95a] M. Brešar and M. Mathieu, *Derivations mapping into the radical. III*, J. Funct. Anal. **133** (1995), no. 1, 21–29.

[BM95b] M. Brešar and C. R. Miers, *Commuting maps on Lie ideals*, Comm. Algebra **23** (1995), no. 14, 5539–5553.

[BM04] N. Boudi and M. Mathieu, *Commutators with finite spectrum*, Illinois J. Math. **48** (2004), no. 2, 687–699.

[BMM96] K. I. Beidar, W. S. Martindale, III, and A. V. Mikhalev, *Rings with generalized identities*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 196, Marcel Dekker, Inc., New York, 1996.

[BMŠ12] M. Brešar, B. Magajna, and Š. Špenko, *Identifying derivations through the spectra of their values*, Integral Equations Operator Theory **73** (2012), no. 3, 395–411.

[Bor91] A. Borel, *Linear algebraic groups*, second ed., Graduate Texts in Mathematics, vol. 126, Springer-Verlag, New York, 1991.

[BPŠ15] Matej Brešar, Claudio Procesi, and Špela Špenko, *Quasi-identities on matrices and the Cayley–Hamilton polynomial*, Adv. Math. **280** (2015), 439–471.

[Bre91] M. Brešar, *On the distance of the composition of two derivations to the generalized derivations*, Glasgow Math. J. **33** (1991), no. 1, 89–93.

[Bre93a] ——, *Centralizing mappings and derivations in prime rings*, J. Algebra **156** (1993), no. 2, 385–394.

[Bre93b] ——, *Commuting traces of biadditive mappings, commutativity-preserving mappings and Lie mappings*, Trans. Amer. Math. Soc. **335** (1993), no. 2, 525–546.

[Bre96] ——, *Applying the theorem on functional identities*, Nova J. Math. Game Theory Algebra **4** (1996), no. 1, 43–54.

[Bre00] ——, *Functional identities: a survey*, Algebra and its applications (Athens, OH, 1999), Contemp. Math., vol. 259, Amer. Math. Soc., Providence, RI, 2000, pp. 93–109.

[Bre04] ——, *Commuting maps: a survey*, Taiwanese J. Math. **8** (2004), no. 3, 361–397.

[Bri93] M. Brion, *Sur les modules de covariants*, Ann. Sci. École Norm. Sup. (4) **26** (1993), no. 1, 1–21.

[BŠ96] M. Brešar and P. Šemrl, *Derivations mapping into the socle*, Math. Proc. Cambridge Philos. Soc. **120** (1996), no. 2, 339–346.

[BŠ99] ——, *On locally linearly dependent operators and derivations*, Trans. Amer. Math. Soc. **351** (1999), no. 3, 1257–1275.

[BŠ03] ———, *Commuting traces of biadditive maps revisited*, Comm. Algebra **31** (2003), no. 1, 381–388.

[BŠ06] ———, *On bilinear maps on matrices with applications to commutativity preservers*, J. Algebra **301** (2006), no. 2, 803–837.

[BŠ08] ———, *An extension of the Gleason-Kahane-Żelazko theorem: a possible approach to Kaplansky's problem*, Expo. Math. **26** (2008), no. 3, 269–277.

[BŠ10] N. Boudi and P. Šemrl, *Derivations mapping into the socle. III*, Studia Math. **197** (2010), no. 2, 141–155.

[BŠ12a] M. Brešar and Š. Špenko, *Determining elements in Banach algebras through spectral properties*, J. Math. Anal. Appl. **393** (2012), no. 1, 144–150.

[BŠ12b] ———, *On Lie and associative algebras containing inner derivations*, Linear Algebra Appl. **437** (2012), no. 2, 648–658.

[BŠ14a] ———, *Functional identities in one variable*, J. Algebra **401** (2014), 234–244.

[BŠ14b] ———, *Functional identities on matrix algebras*, 2014, arXiv:1403.5454, to appear in Algebras Repr. Theory.

[BT79] M. W. Boyd and W. R. Transue, *Properties of ultraproducts*, Rend. Circ. Mat. Palermo (2) **28** (1979), no. 3, 387–397.

[Bur11] S. Burgdorf, *Trace-positive polynomials, sums of hermitian squares and the tracial moment problem*, Ph.D. thesis, Universität Konstanz and Université de Rennes 1, 2011.

[BV03] J. A. Ball and V. Vinnikov, *Formal reproducing kernel Hilbert spaces: the commutative and noncommutative settings*, Reproducing kernel spaces and applications, Oper. Theory Adv. Appl., vol. 143, Birkhäuser, Basel, 2003, pp. 77–134.

[BW13] D. Buzinski and R. Winstanley, *On multilinear polynomials in four variables evaluated on matrices*, Linear Algebra Appl. **439** (2013), no. 9, 2712–2719.

[CD08] B. Collins and K. Dykema, *A linearization of Connes' embedding problem*, New York J. Math. **14** (2008), 617–641.

[CFL05] M. A. Chebotar, Y. Fong, and P.-H. Lee, *On maps preserving zeros of the polynomial* $xy - yx$, Linear Algebra Appl. **408** (2005), 230–243.

[CGMS09] B. Collins, A. Guionnet, and E. Maurel-Segala, *Asymptotics of unitary and orthogonal matrix integrals*, Adv. Math. **222** (2009), no. 1, 172–215.

[CH97] Z. Chen and Z. Huan, *On the continuity of the mth root of a continuous nonnegative definite matrix-valued function*, J. Math. Anal. Appl. **209** (1997), no. 1, 60–66.

[Che52] C. Chevalley, *The Betti numbers of the exceptional simple Lie groups*, Proceedings of the International Congress of Mathematicians, Cambridge, Mass., 1950, vol. 2, Amer. Math. Soc., Providence, R. I., 1952, pp. 21–24.

[CHMN13] J. Cimprič, J. W. Helton, S. McCullough, and C. Nelson, *A noncommutative real nullstellensatz corresponds to a noncommutative real ideal: algorithms*, Proc. Lond. Math. Soc. (3) **106** (2013), no. 5, 1060–1086.

[CHSY03] J. F. Camino, J. W. Helton, R. E. Skelton, and J. Ye, *Matrix inequalities: a symbolic procedure to determine convexity automatically*, Integral Equations Operator Theory **46** (2003), no. 4, 399–454.

[Chu90] C.-L. Chuang, *On ranges of polynomials in finite matrix rings*, Proc. Amer. Math. Soc. **110** (1990), no. 2, 293–302.

[CKL06] M. A. Chebotar, W.-F Ke, and P.-H. Lee, *On a Brešar-Šemrl conjecture and derivations of Banach algebras*, Q. J. Math. **57** (2006), no. 4, 469–478.

[CL10] C.-L. Chuang and T.-K. Lee, *Density of polynomial maps*, Canad. Math. Bull. **53** (2010), no. 2, 223–229.

[CLO07] D. Cox, J. Little, and D. O'Shea, *Ideals, varieties, and algorithms. An introduction to computational algebraic geometry and commutative algebra*, third ed., Undergraduate Texts in Mathematics, Springer, New York, 2007.

[CM97] M. Cabrera and J. Martínez, *Inner derivations on ultraprime normed algebras*, Proc. Amer. Math. Soc. **125** (1997), no. 7, 2033–2039.

[Coh95] P. M. Cohn, *Skew fields. Theory of general division ring*, Encyclopedia of Mathematics and its Applications, vol. 57, Cambridge University Press, Cambridge, 1995.

[Con90] J. B. Conway, *A course in functional analysis*, second ed., Graduate Texts in Mathematics, vol. 96, Springer-Verlag, New York, 1990.

[Dal00] H. G. Dales, *Banach algebras and automatic continuity*, London Mathematical Society Monographs. New Series, vol. 24, The Clarendon Press, Oxford University Press, New York, 2000.

[DCPP13] C. De Concini, P. Papi, and C. Procesi, *The adjoint representation inside the exterior algebra of a simple Lie algebra*, 2013, arXiv:1311.4338.

[DD10] T. D. Dinh and M. Donzella, *On maps preserving zeros of Lie polynomials of small degrees*, Linear Algebra Appl. **432** (2010), no. 2-3, 493–498.

[Dix77] J. D. Dixon, *Rigid embedding of simple groups in the general linear group*, Canad. J. Math. **29** (1977), no. 2, 384–391.

[DT75] J. Duncan and A. W. Tullo, *Finite dimensionality, nilpotents and quasinilpotents in Banach algebras*, Proc. Edinburgh Math. Soc. (2) **19** (1974/75), 45–49.

[Dyn59] E. B. Dynkin, *Homologies of compact Lie groups*, Amer. Math. Soc. Transl. (2) **12** (1959), 251–300.

[Eis95] D. Eisenbud, *Commutative algebra. With a view toward algebraic geometry*, Graduate Texts in Mathematics, vol. 150, Springer-Verlag, New York, 1995.

[EKM08] R. Elman, N. Karpenko, and A. Merkurjev, *The algebraic and geometric theory of quadratic forms*, American Mathematical Society Colloquium Publications, vol. 56, American Mathematical Society, Providence, RI, 2008.

[FH91] W. Fulton and J. Harris, *Representation theory. A first course*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991.

[Fon84] C. K. Fong, *Range inclusion for normal derivations*, Glasgow Math. J. **25** (1984), no. 2, 255–262.

[For72] E. Formanek, *Central polynomials for matrix rings*, J. Algebra **23** (1972), 129–132.

[For91] ———, *The polynomial identities and invariants of $n \times n$ matrices*, CBMS Regional Conference Series in Mathematics, vol. 78, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1991.

[Ger98] L. Gerritzen, *Taylor expansion of noncommutative polynomials*, Arch. Math. (Basel) **71** (1998), no. 4, 279–290.

[GK09] A. E. Guterman and B. Kuzma, *Preserving zeros of a polynomial*, Comm. Algebra **37** (2009), no. 11, 4038–4064.

[Gli60] J. Glimm, *A Stone-Weierstrass theorem for C^*-algebras*, Ann. of Math. (2) **72** (1960), 216–244.

[Hal54] P. R. Halmos, *Commutators of operators. II*, Amer. J. Math. **76** (1954), 191–198.

[Hal69] H. Halpern, *Irreducible module homomorphisms of a von Neumann algebra into its center*, Trans. Amer. Math. Soc. **140** (1969), 195–221.

[Hal03] B. C. Hall, *Lie groups, Lie algebras, and representations. An elementary introduction*, Graduate Texts in Mathematics, vol. 222, Springer-Verlag, New York, 2003.

[Har77] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 52, Springer-Verlag, New York-Heidelberg, 1977.

[HBJP87] J. W. Helton, J. A. Ball, C. R. Johnson, and J. N. Palmer, *Operator theory, analytic functions, matrices, and electrical engineering*, CBMS Regional Conference Series in Mathematics, vol. 68, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1987.

[Hei80] S. Heinrich, *Ultraproducts in Banach space theory*, J. Reine Angew. Math. **313** (1980), 72–104.

[Her56] I. N. Herstein, *Jordan homomorphisms*, Trans. Amer. Math. Soc. **81** (1956), 331–341.

[Hig08] N. J. Higham, *Functions of matrices. Theory and computation*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[HKM11] J. W. Helton, I. Klep, and S. McCullough, *Proper analytic free maps*, J. Funct. Anal. **260** (2011), no. 5, 1476–1490.

[HKM12] ———, *Free analysis, convexity and LMI domains*, Mathematical methods in systems, optimization, and control, Oper. Theory Adv. Appl., vol. 222, Birkhäuser/Springer Basel AG, Basel, 2012, pp. 195–219.

[HLW08] J. Hou, C.-K. Li, and N.-C. Wong, *Jordan isomorphisms and maps preserving spectra of certain operator products*, Studia Math. **184** (2008), no. 1, 31–47.

[HM04] J. W. Helton and S. A. McCullough, *A Positivstellensatz for non-commutative polynomials*, Trans. Amer. Math. Soc. **356** (2004), no. 9, 3721–3737.

[HMV06] J. W. Helton, S. A. McCullough, and V. Vinnikov, *Noncommutative convexity arises from linear matrix inequalities*, J. Funct. Anal. **240** (2006), no. 1, 105–191.

[IW14a] O. Iyama and M. Wemyss, *Maximal modifications and Auslander-Reiten duality for non-isolated singularities*, Invent. Math. **197** (2014), no. 3, 521–586.

[IW14b] ———, *Singular derived categories of $\mathbb{Q}$-factorial terminalizations and maximal modification algebras*, Adv. Math. **261** (2014), 85–121.

[Jac35] N. Jacobson, *Rational methods in the theory of Lie algebras*, Ann. of Math. (2) **36** (1935), no. 4, 875–881.

[Jel05] Z. Jelonek, *On the effective Nullstellensatz*, Invent. Math. **162** (2005), no. 1, 1–17.

[Joh88] B. E. Johnson, *Approximately multiplicative maps between Banach algebras*, J. London Math. Soc. (2) **37** (1988), no. 2, 294–316.

[JR50] N. Jacobson and C. E. Rickart, *Jordan homomorphisms of rings*, Trans. Amer. Math. Soc. **69** (1950), 479–502.

[JW75] B. E. Johnson and J. P. Williams, *The range of a normal derivation*, Pacific J. Math. **58** (1975), no. 1, 105–122.

[Kap70] I. Kaplansky, *Algebraic and analytic aspects of operator algebras*, American Mathematical Society, Providence, R.I., 1970, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 1.

[KBMR12] A. Kanel-Belov, S. Malev, and L. Rowen, *The images of non-commutative polynomials evaluated on* 2×2 *matrices*, Proc. Amer. Math. Soc. **140** (2012), no. 2, 465–478.

[KBMR13] ———, *The images of multilinear polynomials evaluated on* 3×3 *matrices*, 2013, arXiv:1306.4389.

[Kem87] A. R. Kemer, *Finite basability of identities of associative algebras*, Algebra i Logika **26** (1987), no. 5, 597–641, 650.

[KK83] L. Kaup and B. Kaup, *Holomorphic functions of several variables*, de Gruyter Studies in Mathematics, vol. 3, Walter de Gruyter & Co., Berlin, 1983.

[KMRT98] M.-A. Knus, A. Merkurjev, M. Rost, and J.-P. Tignol, *The book of involutions. With a preface in French by J. Tits*, American Mathematical Society Colloquium Publications, vol. 44, American Mathematical Society, Providence, RI, 1998.

[Kno86] F. Knop, *Über die Glattheit von Quotientenabbildungen*, Manuscripta Math. **56** (1986), no. 4, 419–427.

[Kol88] J. Kollár, *Sharp effective Nullstellensatz*, J. Amer. Math. Soc. **1** (1988), no. 4, 963–975.

[Kos97] B. Kostant, *Clifford algebra analogue of the Hopf-Koszul-Samelson theorem, the* ρ*-decomposition* $C(\mathfrak{g}) = \operatorname{End} V_\rho \otimes C(P)$*, and the* $\mathfrak{g}$*-module structure of* $\bigwedge \mathfrak{g}$, Adv. Math. **125** (1997), no. 2, 275–350.

[KP96] H. Kraft and C. Procesi, *Classical invariant theory*, 1996, www.math.unibas.ch/ kraft/Papers/KP-Primer.pdf.

[KP02] S. G. Krantz and H. R. Parks, *A primer of real analytic functions*, second ed., Birkhäuser Advanced Texts: Basler Lehrbücher., Birkhäuser Boston, Inc., Boston, MA, 2002.

[KR86] R. V. Kadison and J. R. Ringrose, *Fundamentals of the theory of operator algebras. Vol. II*, Pure and Applied Mathematics, vol. 100, Academic Press, Inc., Orlando, FL, 1986.

[KS01] E. Kissin and V. S. Shulman, *On the range inclusion of normal derivations: variations on a theme by Johnson, Williams and Fong*, Proc. London Math. Soc. (3) **83** (2001), no. 1, 176–198.

[KS08] I. Klep and M. Schweighofer, *Connes' embedding conjecture and sums of Hermitian squares*, Adv. Math. **217** (2008), no. 4, 1816–1837.

[KŠ14a] I. Klep and Š. Špenko, *Free function theory through matrix invariants*, 2014, arXiv:1407.7551.

[KŠ14b] ———, *A tracial Nullstellensatz*, Bull. Lond. Math. Soc. **46** (2014), no. 5, 1091–1098.

[KŠ15] ———, *Sweeping words and the length of a generic vector space*, in preparation.

[Kuč74] V. Kučera, *The matrix equation* $AX + XB = C$, SIAM J. Appl. Math. **26** (1974), 15–25.

[Kuz08] Alexander Kuznetsov, *Lefschetz decompositions and categorical resolutions of singularities*, Selecta Math. (N.S.) **13** (2008), no. 4, 661–696.

[KVV12] D. S. Kaliuzhnyi-Verbovetskyi and V. Vinnikov, *Noncommutative rational functions, their difference-differential calculus and realizations*, Multidimens. Syst. Signal Process. **23** (2012), no. 1-2, 49–77.

[KVV14] ———, *Foundations of free noncommutative function theory*, Mathematical Surveys and Monographs, vol. 199, American Mathematical Society, Providence, RI, 2014.

[Lan93a] S. Lang, *Real and functional analysis*, third ed., Graduate Texts in Mathematics, vol. 142, Springer-Verlag, New York, 1993.

[Lan93b] C. Lanski, *An Engel condition with derivation*, Proc. Amer. Math. Soc. **118** (1993), no. 3, 731–734.

[LB89] L. Le Bruyn, *Quiver concomitants are often reflexive Azumaya*, Proc. Amer. Math. Soc. **105** (1989), no. 1, 10–16.

[LBP87] L. Le Bruyn and C. Procesi, *Étale local structure of matrix invariants and concomitants*, Algebraic groups Utrecht 1986, Lecture Notes in Math., vol. 1271, Springer, Berlin, 1987, pp. 143–175.

[LBVdB88] L. Le Bruyn and M. Van den Bergh, *Regularity of trace rings of generic matrices*, J. Algebra **117** (1988), no. 1, 19–29.

[Lee05] T.-K. Lee, *Derivations on noncommutative Banach algebras*, Studia Math. **167** (2005), no. 2, 153–160.

[Ler75] U. Leron, *Nil and power-central polynomials in rings*, Trans. Amer. Math. Soc. **202** (1975), 97–103.

[Lit31] D. E. Littlewood, *Identical Relations Satisfied in an Algebra*, Proc. London Math. Soc. **S2-32** (1931), no. 1, 312–320.

[LP67] C. Le Page, *Sur quelques conditions entraînant la commutativité dans les algèbres de Banach*, C. R. Acad. Sci. Paris Sér. A-B **265** (1967), A235–A237.

[LT07] A. Luttman and T. Tonev, *Uniform algebra isomorphisms and peripheral multiplicativity*, Proc. Amer. Math. Soc. **135** (2007), no. 11, 3589–3598.

[Lun73] D. Luna, *Slices étales*, Sur les groupes algébriques, Soc. Math. France, Paris, 1973, pp. 81–105. Bull. Soc. Math. France, Paris, Mémoire 33.

[LWLW97] P.-H. Lee, T.-L. Wong, J.-S. Lin, and R.-J. Wang, *Commuting traces of multiadditive mappings*, J. Algebra **193** (1997), no. 2, 709–723.

[LZ09] T.-K. Lee and Y. Zhou, *Right ideals generated by an idempotent of finite rank*, Linear Algebra Appl. **431** (2009), no. 11, 2118–2126.

[Mag93] B. Magajna, *On the distance to finite-dimensional subspaces in operator algebras*, J. London Math. Soc. (2) **47** (1993), no. 3, 516–532.

[Mar69a] W. S. Martindale, III, *Prime rings satisfying a generalized polynomial identity*, J. Algebra **12** (1969), 576–584.

[Mar69b] ———, *When are multiplicative mappings additive?*, Proc. Amer. Math. Soc. **21** (1969), 695–698.

[Mat88] M. Mathieu, *Elementary operators on prime C^*-algebras. II*, Glasgow Math. J. **30** (1988), no. 3, 275–284.

[Mat89] ———, *Rings of quotients of ultraprime Banach algebras, with applications to elementary operators*, Conference on Automatic Continuity and Banach Algebras (Canberra, 1989), Proc. Centre Math. Anal. Austral. Nat. Univ., vol. 21, Austral. Nat. Univ., Canberra, 1989, pp. 297–317.

[Mat92] ———, *Posner's second theorem deduced from the first*, Proc. Amer. Math. Soc. **114** (1992), no. 3, 601–602.

[MF82] D. Mumford and J. Fogarty, *Geometric invariant theory*, second ed., Ergebnisse der Mathematik und ihrer Grenzgebiete [Results in Mathematics and Related Areas], vol. 34, Springer-Verlag, Berlin, 1982.

[MM91] M. Mathieu and G. J. Murphy, *Derivations mapping into the radical*, Arch. Math. (Basel) **57** (1991), no. 5, 469–474.

[Mol02] L. Molnár, *Some characterizations of the automorphisms of $B(H)$ and $C(X)$*, Proc. Amer. Math. Soc. **130** (2002), no. 1, 111–120.

[MS11] P. S. Muhly and B. Solel, *Progress in noncommutative function theory*, Sci. China Math. **54** (2011), no. 11, 2275–2294.

[Mur90] G. J. Murphy, *C^*-algebras and operator theory*, Academic Press, Inc., Boston, MA, 1990.

[Onn94] S. Onn, *Hilbert series of group representations and Gröbner bases for generic modules*, J. Algebraic Combin. **3** (1994), no. 2, 187–206.

[PĐ95] V. P. Platonov and D. Ž. Đoković, *Linear preserver problems and algebraic groups*, Math. Ann. **303** (1995), no. 1, 165–184.

[Pap97] C. J. Pappacena, *An upper bound for the length of a finite-dimensional algebra*, J. Algebra **197** (1997), no. 2, 535–545.

[Pas14] J. E. Pascoe, *The inverse function theorem and the Jacobian conjecture for free analysis*, Math. Z. **278** (2014), no. 3-4, 987–994.

[Paz84] A. Paz, *An application of the Cayley-Hamilton theorem to matrix polynomials in several variables*, Linear and Multilinear Algebra **15** (1984), no. 2, 161–170.

[Pop06] G. Popescu, *Free holomorphic functions on the unit ball of $B(\mathcal{H})^n$*, J. Funct. Anal. **241** (2006), no. 1, 268–333.

[Pop10] ———, *Free holomorphic automorphisms of the unit ball of $B(\mathcal{H})^n$*, J. Reine Angew. Math. **638** (2010), 119–168.

[Pos57] E. C. Posner, *Derivations in prime rings*, Proc. Amer. Math. Soc. **8** (1957), 1093–1100.

[Pro76] C. Procesi, *The invariant theory of $n \times n$ matrices*, Advances in Math. **19** (1976), no. 3, 306–381.

[Pro87] ———, *A formal inverse to the Cayley-Hamilton theorem*, J. Algebra **107** (1987), no. 1, 63–74.

[Pro07] ———, *Lie groups. An approach through invariants and representations*, Universitext, Springer, New York, 2007.

[Pro13] ———, *On the theorem of Amitsur–Levitzki*, 2013, arXiv:1308.2421.

[Ptá78] V. Pták, *Derivations, commutators and the radical*, Manuscripta Math. **23** (1977/78), no. 4, 355–362.

[PTD13] J. E. Pascoe and R. Tully-Doyle, *Free Pick functions: representations, asymptotic behavior and matrix monotonicity in several noncommuting variables*, 2013, arXiv:1309.1791.

[Raz73] Ju. P. Razmyslov, *A certain problem of Kaplansky*, Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 483–501.

[Raz74] ———, *Identities with trace in full matrix algebras over a field of characteristic zero*, Izv. Akad. Nauk SSSR Ser. Mat. **38** (1974), 723–756.

[Ree97] M. Reeder, *Exterior powers of the adjoint representation*, Canad. J. Math. **49** (1997), no. 1, 133–159.

[Ric48] C. E. Rickart, *One-to-one mappings of rings and lattices*, Bull. Amer. Math. Soc. **54** (1948), 758–764.

[Ric60] ———, *General theory of Banach algebras*, The University Series in Higher Mathematics, D. van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960.

[Ros76] S. Rosset, *A new proof of the Amitsur-Levitski identity*, Israel J. Math. **23** (1976), no. 2, 187–188.

[Row74] L. H. Rowen, *Universal pi-algebras and algebras of generic matrices*, Israel J. Math. **18** (1974), 65–74.

[Row80] ———, *Polynomial identities in ring theory*, Pure and Applied Mathematics, vol. 84, Academic Press, Inc., New York-London, 1980.

[Sal99] D. J. Saltman, *Lectures on division algebras*, CBMS Regional Conference Series in Mathematics, vol. 94, Published by American Mathematical Society, Providence, RI; on behalf of Conference Board of the Mathematical Sciences, Washington, DC, 1999.

[Sho37] K. Shoda, *Einige Sätze über Matrizen*, Jpn. J. Math. **13** (1937), no. 3, 361–365.

[Smi76] J. H. Smith, *Commutators of nilpotent matrices*, Linear and Multilinear Algebra **4** (1976), no. 1, 17–19.

[Som99] M. Sombra, *A sparse effective Nullstellensatz*, Adv. in Appl. Math. **22** (1999), no. 2, 271–295.

[Sta79] R. P. Stanley, *Combinatorics and invariant theory*, Relations between combinatorics and other parts of mathematics, Proc. Sympos. Pure Math., XXXIV, Amer. Math. Soc., Providence, R.I., 1979, pp. 345–355.

[SVdB08] J. T. Stafford and M. Van den Bergh, *Noncommutative resolutions and rational singularities*, Michigan Math. J. **57** (2008), 659–674.

[SW55] I. M. Singer and J. Wermer, *Derivations on commutative normed algebras*, Math. Ann. **129** (1955), 260–264.

[Špe13] Š. Špenko, *On the image of a noncommutative polynomial*, J. Algebra **377** (2013), 298–311.

[ŠVdB15] Š. Špenko and M. Van den Bergh, *Non-commutative resolutions of quotient singularities*, 2015, arXiv:1502.05240.

[Tay73] J. L. Taylor, *Functions of several noncommuting variables*, Bull. Amer. Math. Soc. **79** (1973), 1–34.

[Tho88] M. P. Thomas, *The image of a derivation is contained in the radical*, Ann. of Math. (2) **128** (1988), no. 3, 435–460.

[TL09] T. Tonev and A. Luttman, *Algebra isomorphisms between standard operator algebras*, Studia Math. **191** (2009), no. 2, 163–170.

[TS87] Y. V. Turovskiĭ and V. S. Shulman, *Conditions for the massiveness of the range of a derivation of a Banach algebra and of associated differential operators*, Mat. Zametki **42** (1987), no. 2, 305–314, 346.

[VdB91] M. Van den Bergh, *Cohen-Macaulayness of modules of covariants*, Invent. Math. **106** (1991), no. 2, 389–409.

[VdB93] ———, *Cohen-Macaulayness of semi-invariants for tori*, Trans. Amer. Math. Soc. **336** (1993), no. 2, 557–580.

[VdB99] ———, *Local cohomology of modules of covariants*, Adv. Math. **144** (1999), no. 2, 161–220.

[VdB04] ———, *Non-commutative crepant resolutions*, The legacy of Niels Henrik Abel, Springer, Berlin, 2004, pp. 749–770.

[Voi04] D.-V. Voiculescu, *Free analysis questions I. Duality transform for the coalgebra of* $\partial_{X\,:\,B}$, Int. Math. Res. Not. (2004), no. 16, 793–822.

[Voi10] ———, *Free analysis questions II: the Grassmannian completion and the series expansions at the origin*, J. Reine Angew. Math. **645** (2010), 155–236.

[VV10] M. Varagnolo and E. Vasserot, *Double affine Hecke algebras and affine flag manifolds, I*, Affine flag manifolds and principal bundles, Trends Math., Birkhäuser/Springer Basel AG, Basel, 2010, pp. 233–289.

[Wem14] M. Wemyss, *Aspects of the homological minimal model program*, 2014, arXiv:1411.7189.

Povzetek

V algebri lahko del strukture pozabimo. Zanima nas, kako izluščeni del odraža prvotno strukturo. Realno bi to lahko ponazorili z gibanjem gazele, pri katerem poskušamo le iz sledi gazele kar najbolje razumeti celotno gibanje.

Prvotni sta linearna in multiplikativna struktura. Prvo obdržimo, drugo pa nekoliko okrnemo. Vnaprej dani produkt lahko zamenjamo z Liejevim, porojenim s komutatorjem, jordanskim, porojenim z jordanskim polinomom, ali bolj splošnim produktom, porojenim s poljubnim multinearnim nekomutativnim polinomom. Zanima nas, kakšne so linearne preslikave med algebrami, ki ohranjajo tako množenje, predvsem njih povezava s prvotnimi preslikavami med algebrami. Učinkovita metoda za reševanje tovrstnih problemov je teorija funkcijskih identitet. Praznino je pustila le v končno razsežnih algebrah. V disertaciji poskušamo zakrpati to vrzel.

Z druge strani pristopimo k Banachovim algebram. Teh prvotna struktura je bogatejša, dodatno so opremljene z normo. Množenje v Banachovih algebrah lahko nadomestimo s spektralno funkcijo. Z normo porojena metrika odpre še prostor približnosti, ki omogoči raziskovanje stabilnosti omenjenih preslikav.

Identitete na matrikah in Cayley-Hamiltonov polinom

Funkcijska identiteta je identična relacija v kolobarju, ki poleg poljubnih elementov, kateri se pojavijo na podoben način kot v polinomski identiteti, vsebuje tudi poljubne neznane funkcije. Splošna teorija funkcijskih identitet poskuša opisati te funkcije. Kot plod rešitve dolgo odprtega Hersteinovega problema o Liejevih izomorfizmih [**Bre93b**] se je teorija funkcijskih identitet razvila v uporabno metodo za obravnavanje raznovrstnih problemov v nekomutativni algebri, neasocativni algebri, teoriji operatorjev in funkcionalni analizi.

V dani funkcijski identiteti najprej poiščemo očitne rešitve; funkcije, ki zadoščajo tej identiteti iz formalnih razlogov, ne glede na strukturo opazovanega kolobarja. Take rešitve imenujemo *standardne rešitve*. Značilen rezultat pravi, da so bodisi standardne rešitve edine možne rešitve bodisi je opazovani kolobar poseben. Obstoječa teorija funkcijskih identitet, povzeta v [**BCM07**], tako poda dokončne zaključke za širok razred nekomutativnih kolobarjev, paradoksno pa ne seže do tal osnovnega primera nekomutativnega kolobarja, matrične algebre $M_n(F)$. To se odraža tudi v uporabi. V kopici rezultatov, katerih dokazi temeljijo na splošni teoriji funkcijskih identitet, je potrebno izvzeti $M_n(F)$ (za majhne n), iz narave rezultatov pa lahko slutimo, da so izvzetki nepotrebni. Nekaj tipičnih primerov lahko najdemo v [**BB09**, **BBS11**, **BBCM00**, **BC00b**, **BF99**]. Težava pri matrični algebri $M_n(F)$ sloni na nestandardnih rešitvah, katerih opis se je zdel mnogo težji kot opis standardnih rešitev. Koordinatni pristop ter teorija generičnih matrik in kolobarjev s sledjo končno omogočita napredek še v tej smeri.

Začnemo z raziskovanjem podrazreda funkcijskih identitet. Bodi A končno razsežna algebra in $m \in \mathbb{Z} \cup \{\infty\}$, s $\mathcal{C}$ označimo komutativni kolobar polinomskih funkcij na m kopijah A, $\mathcal{C}\langle X\rangle$ stoji za prosto algebro v m spremenljivkah $X = \{x_1, \ldots, x_m\}$. Zadnjo imenujemo algebra *kvazi-polinomov* na A. Elemente $\mathcal{C}\langle X\rangle$

lahko vrednotimo na A in *kvazi-identitete* algebre A so tisti kvazi-polinomi, katerih vse evalvacije so ničelne.

Kvazi-polinome, v nekaterih člankih imenovane Beidarjeve polinome, sta vpeljala Beidar in Chebotar leta 2000 [**BC00a**], od takrat pa igrajo pomembno vlogo v teoriji funkcijskih identitet in njenih uporabah. Standardne rešitve kvazi-identitet lahko zelo preprosto opišemo: vse koeficientne funkcije morajo biti enake 0 (cf. [**BCM07**, Lemma 4.4]), Cayley-Hamiltonov izrek pa vodi do osnovnega primera kvazi-identitete na matrični algebri $M_n(F)$ z nestandardnimi rešitvami. Imenujemo jo Cayley-Hamiltonova identiteta.

Tako se naravno zastavi vprašanje, če vsaka kvazi-identiteta na $M_n(F)$ izhaja iz Cayley-Hamiltonove identitete. Vprašanje izvira pri sorodnih identitetah s sledjo. Procesi [**Pro76**] in Razmyslov [**Raz74**] sta neodvisno pokazala, da so le-te posledica Cayley-Hamiltonove identitete.

Pokažemo, da za kvazi-identitete to drži, če dopuščamo centralne imenovalce.

Izrek. *Bodi P kvazi-identiteta na $M_n(F)$. Za vsak centralen polinom c na $M_n(F)$ z ničelnim konstantnim koeficientom obstaja $m \in \mathbb{N}$, da je $c^m P$ posledica Cayley-Hamiltonove identitete.*

Globalno vse identitete ne izhajajo iz Cayley-Hamiltonove identitete. Kvocientni prostor $\mathfrak{I}_n/(Q_n)$ kvazi-identitet po prostoru tistih, ki so posledica Cayley-Hamiltonove identitete, je invarianta kvocientne preslikave delovanja projektivne linearne grupe s hkratnim konjugiranjem na prostoru m-teric matrik. Ta prostor se pojavi kot modul na kvocientni raznoterosti in prejšnji izrek pove, da je podprt na singularni množici. Z omejitvijo na podmodul pokažemo, da je neničeln. Antisimetrične kvazi-identitete na $M_n(F)$ stopnje n^2, ki niso posledica Cayley-Hamiltonove identitete, poiščemo med kvazi-identitetami, ki jih splošna linearna grupa transformira kot adjungirano upodobitev.

Izrek. *Obstaja razcep*

$$\mathbb{G}_n[n^2] = \mathbb{G}_n[n^2]_{CH} \oplus \bigwedge^{n^2-2} N_n^* X^2.$$

V formuli $G_n[n^2]$ označuje izotipično komponento antisimetričnih kvazi-identitet stopnje n^2, ki ustreza adjungirani upodobitvi GL_n na prostoru N_n matrik s sledjo 0, $G_n[n^2]_{CH}$ pa označuje podmodul tistih, ki sledijo iz Cayley-Hamiltonove identitete.

Struktura modula $\mathfrak{I}_n/(Q_n)$ ostaja precej neraziskana, dokažemo le, da je modul končno generiran. Tako ima Spechtov problem za kvazi-identitete pozitivno rešitev.

Preidemo k splošnim dvostranskim kvazi-identitetam. To so identitete oblike

$$\sum_{k\in K} F_k(\overline{x}_m^k)x_k = \sum_{l\in L} x_l G_l(\overline{x}_m^l) \quad \text{za vse } x_1 \ldots, x_m \in M_n(F),$$

kjer $\overline{x}_m^k = (x_1, \ldots, \hat{x}_k, \ldots, x_m)$, K in L sta podmnožici $\{1, \ldots, m\}$, F_k, G_l pa poljubne funkcije iz $M_n(F)^{m-1}$ v $M_n(F)$.

Cayley-Hamiltonova identiteta je ponovno osnovni primer funkcijske identitete na $M_n(F)$, katere rešitve niso standardne. V nasprotju s kvazi-identitetami, vse nestandardne rešitve dvostranskih funkcijskih identitet sledijo iz najosnovnejše.

Izrek. *Vsaka dvostranska kvazi-identiteta je posledica Cayley-Hamiltonove identitete.*

Posebej obravnavamo še funkcijske identitete v eni spremenljivki, s poudarkom na komutirajočih preslikavah.

Bodi R kolobar. Preslikava $q : R \to R$ je *komutirajoča*, če velja $[q(x), x] = 0$ za vsak $x \in R$. Preslikava q med Abelovima grupama (oz. vektorskima prostoroma) A in B je sled d-aditivne (oz. d-linearne) preslikave, če obstaja taka

d-aditivna (oz. d-linearna) preslikava $M : A^d \to B$, da je $q(x) = M(x, \ldots, x)$ za vsak $x \in A$. Standardne so oblike preslikav $q(x) = \sum_{i=0}^{d} \mu_i(x)x^i$, kjer je μ_i sled $(d-i)$-aditivne preslikave, ki slika iz R v razširjeni centroid C kolobarja R. Ali je komutirajoča sled d-aditivne preslikave an prakolobarju (oz. praalgebri) nujno standardne oblike? Vprašanje si je leta 1993 zastavil Brešar in poiskal pritrdilen odgovor za $d = 1$ [**Bre93a**], pri $d = 2$ [**Bre93b**] pa je bilo potrebno le izvzeti kolobarje karakteristike 2 in take, ki zadoščajo standardni polinomski identiteti stopnje 4. Za $d = 3$ so problem obravnavali Beidar, Martindale, in Mikhalev [**BMM96**]. Omenjeni trije rezultati so se izkazali za uporabne pri številnih problemih, posebno v teoriji Liejevih algeber, in igrali nenadomestljivo vlogo pri razvoju teorije funkcijskih identitet. Za splošen d so leta 1997 na vprašanje pritrdilno odgovorili Lee, Lin, Wang in Wong [**LWLW97**] na kolobarjih karakteristike 0 ali $> d$, ki ne zadoščajo standardni polinomski identiteti stopnje $2d$. Te predpostavke narekuje metoda dokaza. Za $d = 2$ se je izkazalo, da so nepotrebne [**BŠ03**]. V splošnem ostaja vprašanje brez odgovora. Pritrdilno mu odgovorimo na razredu centralno zaprtih praalgeber.

IZREK. *Če je A centralno zaprta praalgebra in $q : A \to A$ komutirajoča sled multilinearne preslikave, potem je q standardne oblike.*

Obravnavamo tudi splošnejše funkcijske identitete ene spremenljivke. Podrobno si ogledamo funkcijske identitete, za katere je $\sum_{i=0}^{m} x^i q_i(x) x^{m-i}$, kjer so q_i sledi multilinearni preslikav, vedno enak 0 ali pa leži v centru. V končno razsežnem primeru pokažemo, da so nestandardne rešitve spet posledica Cayley-Hamiltonove identitete.

Kolobarji s sledjo

Neločljivo povezana z identitetami na matrikah sta generična matrična algebra, generirana z m generičnimi matrikami, in kolobar s sledjo, ki ga dobimo, ko generični matrični algebri dodamo sledi produktov generičnih matrik. Objekta sta univerzalna v kategoriji algeber (oz. algeber s sledjo), ki zadoščajo polinomskim identitetam (oz. identitetam s sledjo) $n \times n$ matrik. Tako ustrezata polinomskemu kolobarju z nekomutativno geometričnega zornega kota. Zanimajo nas njune geometrijske lastnosti.

Hilbertov Nullstellensatz je klasični rezultat v algebraični geometriji. Nad algebraično zaprtim poljem opiše polinome, ki so ničelni na množici ničel vnaprej danih polinomov. Pomembnost izreka je vodila do številnih posplošitev in razširitev v različne smeri, tudi do prostih algeber. Amitsurjev Nullstellensatz [**Ami57**] opiše polinome, ki so ničelni na skupni množici ničel končne množice nekomutativnih polinomov ovrednotenih na matrični algebri $M_n(F)$. V drugi smeri Bergmanov Nullstellensatz [**Ber06**] študira šibkejšo, smerno ničelnost neodvisno od razsežnosti, kar omogoči močnejše zaključke. V nasprotju s Hilbertovim in Amitsurjevim Nullstellensatzom se lahko izogne potencam v končnem algebraičnem kriteriju. V [**BK11**] so predstavljeni prosti Nullstellensatzi.

Poskušamo napraviti korak k izrekom o ničlah sledi.

IZREK. *Naj bodo $f_1, \ldots, f_r, f \in F\langle X \rangle$. Potem je*

$$\operatorname{tr}(f_1(A)) = \cdots = \operatorname{tr}(f_r(A)) = 0 \qquad \Longrightarrow \qquad \operatorname{tr}(f(A)) = 0$$

za vse $n \in \mathbb{N}$ in vse $A \in M_n(F)^g$ natanko tedaj, ko je f ciklično ekvivalenten linearni kombinaciji polinomov f_i, $1 \le i \le r$ ali pa je linearna kombinacija f_i, $1 \le i \le r$, ciklično ekvivalentna neničelnemu skalarju.

Pri dokazu se močno naslonimo na Kollarjeve efektivne omejitve stopenj v Hilbertovem Nullstellensatzu [**Kol88**], podobne omejitve stopenj najdemo tudi v

[**Som99**] in [**Jel05**], ter na teorijo polinomskih identitet in identitet s sledjo. Kot posledica sledi momentni problem s sledjo.

Medtem ko so ničle osnovni objekt (klasične) algebraične geometrije, je slika enega polinoma nad algebraično zaprtim poljem precej nezanimiva. Slika elementov kolobarja s sledjo ali generične matrične algebre na matrični algebri $M_n(F)$ pa postane zanimiv objekt. Po [**KBMR12**] ga je najprej opazil Kaplansky. Podmnožica $M_n(F)$, ki želi biti slika nekomutativnega polinoma, mora biti očitno zaprta za konjugiranje z obrnljivimi matrikami. Chuang [**Chu90**] je pokazal, da je za končno polje F ta pogoj že zadosten. V neskončnem primeru pa množica nilpotentnih matrik kljub zaprtosti za konjugiranje ni slika nekomutativnega polinoma.

Zožimo se na algebraično zaprta polja karakteristike 0. Če je f polinomska identiteta, potem je $\operatorname{im}(f) = \{0\}$. Centralni polinomi imajo v svoji sliki le skalarne matrike. Katere slike smatrati za majhne? Upoštevati moramo, da je slika zaprta za konjugiranje, slike mnogih polinomov pa so zaprte tudi za množenje s skalarji. Zato označimo $a^\sim = \{\lambda t a t^{-1} \,|\, t \in GL_n, \lambda \in F\}$. Je $\operatorname{im}(f)$ lahko vsebovan v $a^\sim$? V primeru $n = 2$ odgovorimo hitro: $\operatorname{im}(x_1x_2 - x_2x_1)^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^\sim$. Zaradi centralnosti polinoma $(x_1x_2 - x_2x_1)^2$ lahko $\operatorname{im}((x_1x_2 - x_2x_1)^3)$ sestoji le iz tistih matrik s sledjo 0 z neničelno determinanto. Omenimo še, da ima tudi $x_1x_2 - x_2x_1$ razmeroma majhno sliko $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}^\sim \bigcup \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}^\sim$. V splošnem polinom imenujemo *končni na* $M_n(F)$, če obstajajo $a_1, \dots, a_k \in M_n(F)$ z lastnostjo $\{0\} \neq \operatorname{im}(f) \subseteq a_1^\sim \bigcup \cdots \bigcup a_k^\sim$. Bodi j naravno število, ki deli n, in izberimo primitiven j-ti koren enote μ_j, označimo z $\mathbf{1}_r$, $r = \frac{n}{j}$, identično matriko v $M_r(F)$, z $\mathbf{w}_j$ pa diagonalno matriko v $M_n(F)$ z $\mathbf{1}_r, \mu_j\mathbf{1}_r, \dots, \mu_j^{j-1}\mathbf{1}_r$ na diagonali.

Izrek. *Nekomutativen polinom f je končen na $M_n(F)$ natanko tedaj, ko obstaja naravno število j, ki deli n, da je f^j centralen na $M_n(F)$ in f^i ni centralen za $1 \leq i < j$. V tem primeru je $\operatorname{im}(f) \subseteq \mathbf{w}_j^\sim \bigcup n_2^\sim \bigcup \cdots \bigcup n_k^\sim$ za nilpotentne matrike n_i. Dodatno $\operatorname{im}(f^{j+1}) \subseteq \mathbf{w}_j^\sim$.*

Spomnino, da je podmnožica F^{n^2} standardna odprta, če je komplement množice ničel enega polinoma. Primer take množica je GL_n, ki jo lahko preprosto predstavimo kot sliko nekomutativnega polinoma. Enako velja za poljubne standardne odprte množice.

Izrek. *Če je U standardna odprta podmnožica F^{n^2}, ki je zaprta za množenje z neničelnimi skalarji in konjugiranje, potem je $U \cup \{0\} = \operatorname{im}(f)$ za nek nekomutativen polinom f.*

Ošinemo še gostost $\operatorname{im}(f)$ v topologiji Zariskega. Hitro vidimo, da je gostost $\operatorname{im}(f)$ v $M_n(F)$ (oz. $M_n(F)^0$) ekvivalentna algebraični neodvisnosti komutativnih polinomov $\operatorname{tr}(f), \operatorname{tr}(f^2), \dots, \operatorname{tr}(f^n)$ (oz. $\operatorname{tr}(f^2), \dots, \operatorname{tr}(f^n)$). V ta kot smo se zazrli zaradi Lvovovega vprašanja o sliki multilinearnega polinoma. Je slika multilinearnega polinoma vektorski prostor? Pri $n = 2$ in z nekaj izpuščenimi primeri pri $n = 3$ je odgovor pritrdilen [**KBMR12, KBMR13**]. V splošnem ostaja brez odgovora. Če bi bil pritrdilen, potem bi veljalo $\operatorname{im}(f) = M_n(F)$ ali $\operatorname{im}(f) = M_n(F)^0$ za poljuben nekomutativen multilinearen polinom, ki ni identiteta ali centralni polinom [**BK09**]. Gostost bi lahko bila pomemben polkorak k dokazu teh identitet.

V senci Lvovega vprašanja si zastavimo naslednji vprašanji o sliki multilinearnega polinoma. Je f centralen na $M_n(F)$, če obstaja $k \leq 2$, da je f^k centralen na $M_n(F)$, $n \neq 2$? Je f identiteta, če obstaja $k \geq 2$, da velja $\operatorname{tr}(f^k) = 0$ na $M_n(F)$, $n \neq 2$? Prvo se je pojavilo že v [**Ler75, Row74**]. Pritrdilni odgovor na Lvovovo vprašanje bi narekoval, da sta tudi odgovora na zastavljeni vprašanji pritrdilna.

Da bi pritrdili zadnjemu, bi bilo dovolj pokazati, da je $\mathrm{im}(f) \cap M_n(F)^0$ gost v $M_n(F)^0$. Ker ima $\mathbf{w}_j$ sled 0, lahko iz izreka zgoraj hitro ugotovimo, da bi pritrdilni odgovor na drugo vprašanje podal enak odgovor na prvo. Odgovori so nam žal neznani. Pokažemo le razsežnosti prosto različico; če je f neničelen multilinearen polinom, potem f^k, $k \geq 2$, ni vsota komutatorjev. Lvovo domnevo pa potrdimo za multilinearne Liejeve polinome stopnje največ 4.

Nadaljujemo s posebnimi polinomski preslikavami, podanimi z n^2-tericami besed stopnje d na $M_n(F)$. Pokažemo, da v sliki obstaja n^2-terica linearno neodvisnih matrik za skoraj najmanjši možen d.

Izrek. *Bodita $g \geq 2$ in $d = \lceil \log_g n \rceil$. Obstajajo $w_1, \ldots, w_{n^2} \in \langle x_1, \ldots, x_g \rangle_{2d}$, ki so $M_n(F)$-lokalno linearno neodvisne.*

K obravnavi tovrstnih preslikav nas napelje Pazova domneva. Bodi V vektorski podprostor $M_n(F)$. Z V^k označimo vektorski prostor razpet na besede dolžine največ k ovrednotene na V. *Dolžina* V je naravno število ℓ, ki določi dolžino verige

$$V \subsetneq V^2 \subsetneq \cdots \subsetneq V^\ell = V^{\ell+1}.$$

Paz je pokazal, da je $\ell(V) \leq \lceil (n^2+2)/3 \rceil$ [**Paz84**], kasneje je Pappacena mejo znatno izboljšal do reda $O(n^{3/2})$ [**Pap97**]. Pazova domneva trdi, da je optimalna meja $2n-2$. Ni težko poiskati primerov, ki to mejo dosežejo. Za generični vektorski podprostor V prostora $M_n(F)$ pa bi pričakovali dolžino reda $O(\log n)$, kar imenujemo generična Pazova domneva. Enostavna posledica zgornjega izreka potrdi ugibanje.

Kolobarji s sledjo koreninijo v klasični invariantni teoriji. Bodi V n-razsežen vektorski prostor. Grupa $G = \mathrm{GL}_n(\mathbb{C})$ s hkratnim konjugiranjem deluje na prostor $W = \mathrm{End}(V)^{\oplus m}$. Procesi je pokazal, da je kolobar invariant $(SW)^G$ izomorfen centru $\mathcal{T}_{m,n}$ kolobarja s sledjo $\mathcal{T}_{m,n}\langle \xi_k \rangle$, $\mathcal{T}_{m,n}\langle \xi_k \rangle = (SW \otimes \mathrm{End}(V))^G$ pa je algebra kovariant [**Pro76**]. Medtem ko je $\mathcal{T}_{m,n}$ singularen za vse (m,n) z izjemo $(m,n) = (2,2)$ [**LBP87**, Proposition II.3.1], ima $\mathcal{T}_{m,n}\langle \xi_k \rangle$ za odtenek svetlejše homološke lastnosti. Končne globalne razsežnosti je za natanko $(m,n) = (2,2), (3,2), (2,3)$ [**LBVdB88**, **LBP87**].

Nedaven razvoj singularne teorije je pospešil Van den Bergh z vpeljavo "nekomutativnih odprav singularnosti".

Definicija. Bodi S normalen noetherski Gorensteinov kolobar brez deliteljev niča. *Nekomutativna krepantna odprava singularnosti* (NCCR) kolobarja S je S-algebra končne globalne razsežnosti oblike $\Lambda = \mathrm{End}_S(M)$, kjer je M neničeln končno generiran refleksiven S-modul, Λ pa Cohen-Macaulayev S-modul.

Za razlago definicije napotimo na izvirni članek [**VdB04**]. V splošnem NCCR v vedenju precej natančno posnema svojo komutativno različico, krepantno odpravo iz algebraične geometrije. Koncepta sta tesno povezana, v razsežnosti 3 celo ekvivalentna [**VdB04**] in to je del motivacije za algebraični pristop Iyame in Wemyssa [**IW14a**] k programu trirazsežnega minimalnega modela.

Nekomutativne odprave so dobro poznane za kvocientne singularnosti končnih grup. Bodi G končna grupa, z $\hat{G}$ označimo množico izomorfnih razredov nerazcepnih G-upodobitev. Za končno razsežno G-upodobitev U označimo z $M(U) = (U \otimes F[X])^G$ pripadajoči $F[X]^G$-*modul kovariant*. Postavimo $U = \oplus_{V \in \hat{G}} V$. Če je W upodobitev G in je $G \subset \mathrm{SL}(W)$, potem je $\Lambda = \mathrm{End}_{SW^G}(M(U))$ NCCR kolobarja SW^G.

Za kvocientne singularnosti neskončnih grup so bile nekomutativne krepantne resolucije sestavljene za determinante raznoterosti [**BLVdB10**, **BLVdB11**] in raznoterosti $n \times n$ poševno-simetričnih matrik ranga < 4 za lihe n [**Kuz08**].

Pokažemo, da imajo kvocientne singularnosti reduktivnih grup vedno nekomutativne resolucije v primernem smislu. Kot pri končnih grupah pri grajenju resolucij preučimo lastnosti kategorije $\mathrm{mod}(G, SW)$. Vendar je v tem primeru potrebna globlja analiza zaradi dveh netrivialnih zapletov:

(1) kategorija $\mathrm{mod}(G, SW)$ nima projektivnega generatorja, ker ima G neskončno neizomorfnih nerazepnih upodobitev,
(2) moduli kovariant običajno niso Cohen-Macaulayevi.

Prvega odpletemo s konstrukcijo kompleksov, ki povežejo različne projektivne module v $\mathrm{mod}(G, SW)$. Drugega rešijo kriteriji v [**VdB91**, **VdB93**, **VdB99**].

Predstavimo tudi razmeroma širok razred takih singularnosti, ki imajo (vite) nekomutativne krepantne resolucije. Z razvitimi metodami lahko pristopimo k veliki množici primerov, tako starih kot novih. V posebnem NC(C)R obstajajo v prej nepoznanih primerih determinantnih raznoterosti simetričnih in poševnosimetričnih matrik. Omejimo se na kolobarje s sledjo. Vita NCCR je podobna običajni NCCR, le generično je centralno enostavna algebra in ne nujno matrična.

Izrek. *Bodita $m \geq 2$, $n \geq 2$. Potem $\mathcal{T}_{m,n}$ ima vito NCCR.*

Prosta funkcijska teorija skozi matrične invariante

Z identitetami in njihovimi pripadajočimi univerzalnimi objekti zaključimo s prikazom uporabe v polju proste analize.

Proste preslikave so preslikave na g-tericah matrik poljubne velikosti, ki ohranjajo hkratno podobnost in direktne vsote. Naravno se pojavijo v prosti verjetnosti, študiju nekomutativnih racionalnih funkcij [**AD03**, **BGM06**, **HMV06**] in teoriji sistemov [**HBJP87**, **KVV12**]. S študijem teh preslikav se ukvarja prosta analiza [**AM15**, **AM14**, **AY14**, **AKV13**, **BV03**, **KVV14**, **HKM12**, **MS11**, **Pas14**, **PTD13**, **Pop06**, **Pop10**, **Tay73**, **Voi04**, **Voi10**].

K prosti analizi pristopimo algebraično in vanjo vpeljemo močne metode invariantne teorije [**Pro76**]. Čeprav je sicer glavni tok usmerjen k prostim preslikavam brez involucije, sorodnim analitičnim funkcijam v več kompleksnih spremenljivkah, ki se izkažejo za zelo rigidne, se usmerimo k prostim preslikavam z involucijo, med katerimi najdemo nekomutativne polinome, racionalne funkcije in potenčne vrste v x, x^t. Metode, ki jih predstavimo, delujejo v obeh primerih. Tako ponovno odkrijemo nekaj obstoječih lastnosti prostih preslikav brez involucije ([**AM14**, **KVV14**, **Pas14**]).

Pokažemo, da je prosta preslikava z involucijo f polinom v x, x^t natanko tedaj, ko obstaja tak $d \in \mathbb{N}$, da so na $M_n(F)$ zožene funkcije $f[n]$ polinomske stopnje $\leq d$ za vsak $n \in \mathbb{N}$. S tem zgradimo

Izrek. *Analitična prosta preslikava z involucijo ima razvoj v potenčno vrsto okoli skalarnih točk.*

Pokažemo, da enako velja za razvoj okoli neskalarnih točk, homogene komponente so v tem primeru posplošeni polinomi. Dokažemo prosti različici izrekov o inverzni in implicitni funkciji za diferenciabilne preslikave z involucijo.

S primeri opomnimo na šibkejšo rigidnost prostih preslikav z involucijo glede na običajne proste preslikave. Med drugim obstaja omejena gladka prosta preslikava z involucijo, ki ni analitična.

Spekter kot invarianta Banachovih algeber

Spekter elementa a Banachove algebre A označimo s $\sigma(a)$. Z $r(a)$ označimo njegov spektralni radij. Za $Z(A)$ stoji center A. Banachova algebra je polenostavna natanko tedaj, ko je edini element z lastnostjo $\sigma(ax) = \{0\}$ za vsak $x \in A$ ničelni element. V polenostavni Banachovi algebri je tako $a = 0$ edini element z ničelno spektralno funkcijo $x \mapsto \sigma(ax)$. Poskušamo ugotoviti, če lahko tudi druge elemente prepoznamo po spektralni funkciji.

VPRAŠANJE. Bodi A polenostavna Banachova algebra. Naj $a, b \in A$ zadoščata

$$\sigma(ax) = \sigma(bx) \quad \text{za vsak } x \in A.$$

Sledi $a = b$?

Odgovora v splošnem ne poznamo. V različnih posebnih primerih pokažemo, da je pritrdilen. To na primer velja, če a lahko zapišemo kot produkt idempotenta in obrnljivega elementa. Dokaz temelji na spektralni karakterizaciji centralnih idempotentov. Odgovor podamo še v primeru komutativnih Banachovih algeber in C^*-algeber.

IZREK. *Če v C^*-algebri A elementa $a, b \in A$ zadoščata $\sigma(ax) = \sigma(bx)$ za vsak $x \in A$, potem velja $a = b$.*

Obravnavamo tudi splošnješi pogoj, kjer spekter nadomestimo s spektralnim radijem. Bodi A polenostavna Banachova algebra. Naj za $a, b \in A$ velja $r(ax) \leq r(bx)$ za vse $x \in A$. Kakšen je odnos med a in b? Odgovor je odvisen od opazovane algebre in izbranih elementov. V posebnem primeru $b = 1$ je Ptak [**Ptá78**] pokazal, da mora a ležati v $Z(A)$. V C^*-praalgebri ugotovimo, da morata biti a in b linearno odvisna.

Bodita B in A Banachovi algebri in $\phi : B \to A$ surjektivna linearna preslikava z lastnostjo

$$\sigma\big(\varphi(x)\big) = \sigma(x) \quad \text{za vsak } x \in B.$$

Kdaj je ϕ jordanski homomorfizem? To je klasični problem teorije Banachovih algeber, z začetkom pri Kaplanskyem [**Kap70**], ki napelje na študij prepoznavanja elementov preko spektralne funkcije. Pritrdilen odgovor se pričakuje za C^*-algebre, morda celo za splošne polenostavne Banachove algebre. Reštev problema se zdi še zelo oddaljena, pot zgodovine je popisana v [**BŠ08**]. Poskusimo z močnejšim pogojem, ki morda drobno osvetli klasičnega.

V [**Mol02**] je Molnar opisal ne nujno linearne surjektivne preslikave ϕ, ki zadoščajo $\sigma\big(\varphi(x)\varphi(y)\big) = \sigma(xy)$ za vsaka $x, y \in B$ v primeru, ko je $B = A = B(H)$ ali $B = A = C(K)$. Rezultati so se širili v veliko smeri [**HLW08**, **LT07**, **TL09**], vendar samo na posebnih primerih algeber. Obravnavamo podobna, a lažje dostopna pogoja

$$\rho\big(\varphi(x)\varphi(y)\varphi(z)\big) = \rho(xyz) \quad \text{za vse } x, y, z \in B,$$

kjer $\rho \in \{\sigma, r\}$, z uporabo prej omenjenih rezultatov.

Naprej raziskujemo druge pare linearnih preslikav na A z enako spektralno funkcijo. Obravnavamo par odvajanj d in g. Namesto enakosti študiramo nekoliko splošnejšo vsebovanost spektrov:

$$\sigma(g(x)) \subseteq \sigma(d(x)) \quad \text{za vsak } x \in A.$$

Spektralne lastnosti vrednosti odvajanj so obravnavali za odvajanja in produkte odvajanj s kvazinilpotentnimi vrednostmi v [**CKL06**, **Lee05**, **TS87**], spektralno omejena odvajanja v [**BM95a**], in odvajanja, katerih vse vrednosti imajo končen spekter, v [**BM04**, **BŠ10**, **BŠ96**].

Trivialno se vsebovanost pojavi v primeru $g = d$. Če zaloga vrednosti d sestoji iz neobrnljivih elementov, na primer za notranje odvajanje d porojeno z elementom iz pravega ideala, potem je g lahko tudi 0. Če je d notranje odvajanje, porojeno z algebraičnim elementom stopnje 2, pa se lahko pojavi tudi $g = -d$. Pokažemo, da so omenjene možnosti edine, če je A primitivna Banachova algebra z neničelnim podstavkom. S tem sklepom podobno zaključimo v razredu splošnih polenostavnih

Banachovih algeber s predpostavko, da je spekter $d(x)$ končen za vsak $x \in A$. Na von Neumannovih algebrah izpostavljeni pogoj vodi do razcepa.

IZREK. *Bodi A von Neumannova algebra in $a, b \in A$. Če velja $\sigma([b,x]) \subseteq \sigma([a,x]) \cup \{0\}$ za vsak $x \in A$, potem $b = p_1a - p_2a + z$ za ortogonalne centralne projekcije p_1, p_2 in $z \in Z(A)$.*

Omejimo se še na C^*-algebro A in notranji odvajanji $d : x \mapsto [a,x]$, $g : x \mapsto [b,x]$. Pokažemo, da b leži v $\{a\}''$, (relativnem) bikomutantu $\{a\}$, če je a normalen. Iz vsebovanosti spektrov očitno sledi pogoj $r([b,x]) \leq Mr([a,x])$ z $M = 1$. Če sta a in b sebiadjungirana, potem sta komutatorja $[a,x]$ in $[b,x]$ antisebiadjungirana, če je x sebiadjungiran. V tem primeru pogoj lahko zapišemo kot

$$||[b,x]|| \leq M||[a,x]|| \quad \text{za vse sebiadjungirane } x \in A.$$

Po [**JW75**, Lemma 1.1] lahko zadnji pogoj smatramo za dualni problem vsebovanosti slik. Po sledi [**JW75**] in [**Fon84**, **KS01**] se nanj osredotočimo za normalen element a v C^*-algebri A. Najbolj zadovoljivo ga sicer razumemo z enakostjo za $M = 1$ v primeru sebiadjungiranih elementov a, b.

Zadržimo se še nekoliko pri odvajanjih, katerih spektralna funkcija je ničelna na elementih s trivialnim spektrom. S $Q = Q_A$ označimo množico kvazinilpotentnih elementov v A, torej $Q = \{q \in A \mid \sigma(q) = \{0\}\}$, z $\mathrm{rad}(A)$ pa (Jacobsonov) radikal A. Spomnimo, $\mathrm{rad}(A) = \{q \in A \mid qA \subseteq Q\}$.

Dobro je znano, da je $d(A) \subseteq \mathrm{rad}(A)$, če je A komutativna. S predpostavko zveznosti sta to dokazala Singer in Wermer [**SW55**], odstranil jo je Thomas [**Tho88**]. Razšitiev na nekomutativne algebre je stekla v več smeri. Le Page [**LP67**] je dokazal, da $d(A) \subseteq Q$ povzroči $d(A) \subseteq \mathrm{rad}(A)$, če je d notranje odvajanje. Za splošno odvajanje sta to pokazala Turovskii in Shulman [**TS87**] (in neodvisno [**MM91**]). V [**BM95b**] je dokazano, da $d(A) \subseteq \mathrm{rad}(A)$, če obstaja $M > 0$ z lastnostjo $r(d(x)) \leq Mr(x)$ za vsak $x \in A$. Katavolos in Stamatopoulos [**KS08**] sta pokazala, da za notranje odvajanje d, porojeno s kvazinilpotentnim elementom, $d(Q) \subseteq Q$ narekuje $d(A) \subseteq \mathrm{rad}(A)$.

Vprašanje, če $d(Q) \subseteq Q$ povzroči $d(A) \subseteq \mathrm{rad}(A)$ za poljubno odvajanje d, se zdi naravno, saj pogoj $d(Q) \subseteq Q$ za poljuben d zaobjame vse pogoje prejšnjega odstavka. V splošnem je odgovor sicer negativen, saj je Q lahko $\{0\}$ tudi v nekomutativni algebri A [**DT75**], in v tem primeru vsako neničelno notranje odvajanje ponudi protiprimer. Zato se omejimo na poseben razred Banachovih algeber.

IZREK. *Bodi A Banachova algebra z lastnostjo β, $Q \subseteq A$ množica kvazinilpotentnih elementov. Če odvajanje d zadošča $d(Q) \subseteq Q$, potem $d(A) \subseteq \mathrm{rad}(A)$.*

Razred algeber z nekoliko tehnično lastnostjo β je precej širok, vsebuje C^*-algebre, grupne algebre poljubnih lokalno kompaktnih grup, z idempotenti generirane Banachove algebre.

Analitične Liejeve preslikave

Zaradi dodatne metrične strukture Banachovih algeber se pojavijo analitične različice odvajanj, komutirajočih in Liejevih preslikav. Hiter račun pokaže, da za poljuben zvezen linearen (oz. kvadratičen) operator T na Banachovi algebri A velja

$$\sup\{\|T(a)a - aT(a)\| : \ a \in A, \|a\| = 1\} \leq 2\|T - S\|$$

za vsak komutirajoč zvezen linearen (oz. kvadratičen) operator S na A. Pokažemo, da iz majhnosti $\sup_{a\in A,\|a\|=1} \|T(a)a - aT(a)\|$ sledi, da je T blizu neki komutirajoči preslikavi. Primeren okvir za obravnavo je razred ultrapra Banachovih algeber, v katerih imajo algebraični opisi komutirajočih in podobnih preslikav posebej lično obliko.

V tem okviru študiramo tudi aproksimativne Liejeve izomorfizme in aproksimativna Liejeva odvajanja. Bodita A in B Banachovi algebri, $\Phi\colon B \to A$ in $\Delta\colon A \to A$ pa zvezni linearni preslikavi. Liejevo multiplikativnost Φ in Liejevo odvedljivost Δ merimo s konstantami

$$\mathrm{lmult}(\Phi) = \sup\{\|\Phi([a,b]) - [\Phi(a),\Phi(b)]\| :\ a,b \in B,\ \|a\| = \|b\| = 1\}$$

in

$$\mathrm{lder}(\Delta) = \sup\{\|\Delta([a,b]) - [\Delta(a),b] - [a,\Delta(b)]\| :\ a,b \in A,\ \|a\| = \|b\| = 1\}.$$

Zanima nas, če majhnost $\mathrm{lmult}(\Phi)$ (oz. $\mathrm{lder}(\Delta)$) povzroči, da sta Φ (oz. Δ) blizu pravemu Liejevemu homomorfizmu (oz. Liejevemu odvajanju). Dokončne rezultate, sloneče na [**Joh88**] in [**AEV10**], dobimo za $\mathcal{L}(H)$, kjer je H Hilbert prostor.

IZREK. *Bodi H separabilen Hilbertov prostor. Za poljubna $M, \varepsilon > 0$ obstaja tak $\delta > 0$, da za bijektivno zvezno linearno preslikavo $\Phi\colon \mathcal{L}(H) \to \mathcal{L}(H)$ z $\|\Phi\|, \|\Phi^{-1}\| \le M$ in $\mathrm{lmult}(\Phi) < \delta$ velja*

$$\min\left\{\mathrm{dist}\big(\Phi, \mathrm{Hom}(\mathcal{L}(H))\big), \mathrm{dist}\big(\Phi, -\,\mathrm{AHom}(\mathcal{L}(H))\big)\right\} < \varepsilon.$$

Povezan problem aproksimativnih jordanskih izomorfizmov se je pojavil pri študiju aproksimativnih preslikav, ki ohranjajo spekter, v [**AEV11**]. Klasični Hersteinov problem jordanskih epimorfizmov je stabilen v smislu, da so aproksimativni jordanski epimorfizmi aproksimativni epimorfizmi ali aproksimativni antiepimorfizmi [**AEV11**]. Uporabimo podobne tehnike, vendar so komutirajoče in Liejeve preslikave tehnično zahtevnejše zaradi prisotnosti centralnih preslikav.

Na odvajanja pogledamo še z analitične strani. Podamo metrični različici klasičnih Posnerjevih izrekov o odvajanjih [**Pos57**]. Metrično različico prvega Posnerjevega izreka z oceno razdalje kompozituma D_1D_2 odvajanj D_1 in D_2 na ultrapra Banachovi algebri do posplošenih odvajanj na A najdemo v [**Bre91**]. Odvedljivost danega zveznega linearnega operatorja T na ultrapra Banachovi algebri A merimo s konstanto

$$\mathrm{der}(T) = \sup\{\|T(ab) - T(a)b - aT(b)\|\colon \|a\| = \|b\| = 1\}.$$

Ocenimo $\|S\|\|T\|$ glede na $\mathrm{der}(S)$, $\mathrm{der}(T)$, in $\mathrm{der}(ST)$ za poljubna zvezna linearna operatorja S in T na A. Metrično različico drugega Posnerjevega izreka pa podamo z oceno $\|T\|\sup\{\|ab - ba\|\colon \|a\| = \|b\| = 1\}$ glede na $\mathrm{der}(T)$ in $\sup\left\{\mathrm{dist}\big([T(a),a], Z(A)\big)\colon \|a\| = 1\right\}$.

f-homomorfizmi in f-odvajanja

Bodi f nekomutativen multilinearen polinom. Pravimo, da preslikava $\phi : A \to A$ *ohranja ničle* f, če za poljubne $a_1, \dots, a_d \in A$ velja

$$f(a_1,\dots,a_d) = 0 \implies f(\phi(a_1),\dots,\phi(a_d)) = 0.$$

Seznam preslikav na A, ki ohranjajo ničle f, mora seveda vsebovati skalarne večkratnike avtomorfizmov, za nekatere polinome tudi skalarne večkratnike antiavtomorfizmov (na primer za $f = x_1x_2 + x_2x_1$), za nekatere pa celo vse preslikave oblike

$$\phi(x) = \alpha\theta(x) + \mu(x),$$

pri čemer je $\alpha \in F$, $\theta : A \to A$ avtomorfizem ali antiavtomorfizem, $\mu : A \to Z(A)$ pa linearna preslikava (na primer za $f = x_1x_2 - x_2x_1$). Take preslikave imenujemo standardne.

Za nekaj enostavnih polinomov, predvsem za $f = x_1x_2$ in $f = x_1x_2 - x_2x_1$, ima problem ohranjanja ničel dolgo zgodovino, v [**ABEV09**] in [**BCM07**] se najdejo njen popis ter dodatne reference. Za splošne polinome je bil problem eksplicitno zastavljen v [**CFL05**] na matrični algebri $A = M_n(F)$. Delne rešitve so se pokazale v dveh nedavnih člankih: [**GK09**] se ukvarja s primerom, ko je vsota koeficientov f neničelna brez predpostavke linearnosti ϕ, [**DD10**] pa obravnava Liejeve polinome stopnje največ 4.

Sorodni in na prvi pogled bližji homomorfizmom so f-homomorfizmi; linearne preslikave, ki ohranjajo vse vrednosti f, $\phi(f(a_1, \ldots, a_d)) = f(\phi(a_1), \ldots, \phi(a_d))$ za vse $a_i \in A$. S pomočjo funkcijskih identitet so bili opisani v precejšnji splošnosti, le za končno razsežne algebre, vključno z $M_n(F)$, dobljeni opisi niso optimalni, [**BF99**] in [**BCM07**, Section 6.5].

Najprej pokažemo, da so preslikave, ki ohranjajo ničle polinoma za posebne polinome v splošnih razredih praalgeber in C^*-algeber, standardne oblike. Na matrični algebri $A = M_n(F)$ opišemo vse take preslikave za poljubne polinome z milimi tehničnimi omejitvami.

Izrek. *Bodi F polje s $\mathrm{char}(F) = 0$, $f \in F\langle X\rangle$ multilinearen polinom stopnje $d \geq 2$, in bodi $\phi : M_n(F) \to M_n(F)$ bijektivna linearna preslikava, ki ohranja ničle f in zadošča $\phi(1) \in F \cdot 1$. Če sta $n \neq 2, 4$ in $d < 2n$, potem je ϕ standardne oblike.*

Zadošča opaziti, da je množica tovrstnih preslikav algebraična podgrupa GL_{n^2}, ki vsebuje (notranje) avtomorfizme $M_n(F)$. Te sta opisala Platonov in Đoković v [**PĐ95**]. Na seznamu hitro izločimo neustrezne. Z natančno analizo njunega dokaza lahko napravimo tudi seznam vseh algebraičnih Liejevih podalgeber $\mathfrak{gl}_{n^2}$, ki vsebujejo Liejevo algebro vseh (notranjih) odvajanj. Tako opišemo še f-odvajanja.