

# PRESERVATION OF ALGEBRAICITY IN FREE PROBABILITY

GREG W. ANDERSON

ABSTRACT. We show that any matrix-polynomial combination of free noncommutative random variables each having an algebraic law has again an algebraic law. Our result answers a question raised by a recent paper of Shlyakhtenko and Skoufranis. The result belongs to a family of results with origins outside free probability theory, including a result of Aomoto asserting algebraicity of the Green function of random walk of quite general type on a free group.

## CONTENTS

1. Statement of the main result and introduction	1
2. Background for the main result and a reduction of the proof	5
3. Hessenberg-Toeplitz matrices and free cumulants	11
4. The linearization step	15
5. Solving the generalized Schwinger-Dyson equation	22
6. Notes on Newton-Puiseux series	26
7. Evaluation of algebraic power series on matrices	29
8. Proof of the main result	35
References	39

## 1. STATEMENT OF THE MAIN RESULT AND INTRODUCTION

Our main result is as follows:

**Theorem 1.** *Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space. Let  $x_1, \dots, x_q \in \mathcal{A}$  be freely independent noncommutative random variables. Let*

$$X \in \text{Mat}_p(\mathbb{C}\langle x_1, \dots, x_q \rangle) \subset \text{Mat}_p(\mathcal{A})$$

*be a matrix. If the laws of  $x_1, \dots, x_q$  are algebraic, then so is the law of  $X$ .*

This paper is devoted to a proof of Theorem 1. We say for short that  $X$  is a *free matrix-polynomial combination* of  $x_1, \dots, x_q$ . See §2 below for notation, terminology, background and references clarifying the theorem statement.

The beginning student of free probability immediately notices that all the important special distributions in the theory are algebraic, e.g., the arcsine law, Wigner's semicircle law, the Marcenko-Pastur law, and so on. Our main result and its proof describes a mechanism at least in part accounting for this ubiquity.

---

*Date:* August 14, 2014.

*2010 Mathematics Subject Classification.* 46L54, 68Q70.

*Key words and phrases.* Free probability, algebraicity, Schwinger-Dyson equation, linearization, realization, formal language.

The phenomenon of preservation of algebraicity established in general here has previously been verified in many special cases, often merely as a byproduct. Here are several important examples. (i) It is implicit in the theory of the  $R$ -transform introduced in [35] that free (additive) convolution preserves algebraicity. We note that this phenomenon has been exploited in a practical way in [13]. Similar remarks apply to free multiplicative convolution. (ii) It is implicit in the theory of commutators of free random variables developed in [27] that formation of free commutators preserves algebraicity. (iii) Algebraicity of the law of a free matrix-polynomial combination of semicircular variables is asserted in [31, Thm. 5.4]. This result forms part of a result of wide scope, namely [31, Thm. 1.1], which gives important constraints on the structure of the law of a free matrix-polynomial combination of noncommutative random variables each of which has a nonatomic law.

Our main result answers a question raised by [31].

One further example of algebraicity, originating outside free probability, deserves special mention as the archetype for Theorem 1.

**Theorem 2** (See, e.g., [33, Cor. 6.7.2, p. 210]). *Let  $G$  be a free group. Let  $\mathbb{C}[G]$  be the group algebra of  $G$  with complex coefficients. Let  $\tau : \mathbb{C}[G] \rightarrow \mathbb{C}$  be the unique  $\mathbb{C}$ -linear map such that  $\tau(g) = \delta_{1g}$  for  $g \in G$ . Then for any  $P \in \mathbb{C}[G]$  and  $g \in G$  the formal power series  $\sum_{n=0}^{\infty} \tau(gP^n)t^n \in \mathbb{C}[[t]]$  is algebraic over the field of rational functions  $\mathbb{C}(t)$ .*

It is remarked in [22] that Theorem 2 has been frequently rediscovered. We do not know the identity of the first discoverer.

Now a unitary noncommutative random variable factors as a product of two free Bernoulli variables. In other plainer words, a group generated by elements  $y_1, \dots, y_{2q}$  subject only to the relations  $y_i^2 = 1$  for  $i = 1, \dots, 2q$  contains a free subgroup on  $q$  generators, for example that generated by  $y_1y_2, y_3y_4, \dots, y_{2q-1}y_{2q}$ . Thus Theorem 2 in the case  $g = 1$  is a consequence of Theorem 1. Given this close relationship, one reasonably looks to the proof of Theorem 2 for clues concerning the proof of Theorem 1.

The proof of Theorem 2 given in [33] is based on [15] which in turn builds upon the theory of algebraic noncommutative formal power series (see e.g., [29] or [33, Chap. 6]), and for the latter, crucial foundations are laid in the seminal paper [11] on context free languages. Tools from the same kit are also used to prove [31, Thm. 5.4].

Somewhat counterintuitively, we prove Theorem 1 by an approach mostly avoiding formal language theory. We rely instead on methods from free probability, random walk on groups, algebraic geometry and commutative algebra. Formal language theory is still involved, but in a different and simpler way. A proof of Theorem 1 parallel to that of [31, Thm. 5.4] might yet be possible and would be very interesting. For now the sticking point seems to be that no suitable generalization of [31, Lemma 5.12] is obviously on offer. It is conceivable that our methods could have in the reverse direction some impact on formal language theory.

Questions quite similar to that answered by Theorem 2 have been treated in the literature of random walk on groups. Here are two particularly important examples. (i) In [4] algebraicity of the Green function of random walk of a fairly general type on a finitely generated free group was proved by explicit calculation. This is the earliest paper of which we are aware which up to some mild and ultimately removable hypotheses proves Theorem 2. (ii) In [40] algebraicity of the Green function of any

finitely supported random walk on a group with a finitely generated free subgroup of finite index was proved by a method based on formal language theory, a method much the same as later used in [33] to prove Theorem 2.

The overall approach to algebraicity used here is adapted not from the literature of formal language theory, but rather from that of random walk on groups and graphs, especially free groups and infinite trees. See [41] for background. From the random walk literature we cite as particular examples [4], [23] and [26]. This short list could be greatly extended because the idea we are adapting is a fundamental trope. There are two main components to this trope. Firstly, one gets relatively simple recursions for Green functions by exploiting finiteness of cone types, or some related principle of self-similarity. Such recursions are often (in effect) consequences of the familiar matrix inversion formula

$$\begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{d}^{-1} \end{bmatrix} + \begin{bmatrix} & \mathbf{1} \\ -\mathbf{d}^{-1}\mathbf{c} & \end{bmatrix} (\mathbf{a} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c})^{-1} \begin{bmatrix} \mathbf{1} & -\mathbf{b}\mathbf{d}^{-1} \end{bmatrix}.$$

Here we put the Boltzmann-Fock space model of free random variables [35] into suitably “arboreal” form in order to gain access to similar recursions. Secondly, there are criteria available for recognizing when recursions have algebraic solutions. Of course formal language theory provides such criteria, but there are also less complicated criteria. For example, [23, Prop. 5.1] is an especially clear and general criterion, and it is applied in the cited paper in an instructive manner. We use a similar but simpler criterion here. See Proposition 2.4.2 below.

We remark that when writing [2], wherein was presented an algebraicity criterion similar to if rather more complicated than [23, Prop. 5.1], namely [2, Thm. 6.1], the authors were unfortunately unaware of [23]. We wish to acknowledge the priority.

We also remark that relations between algebraicity and positivity are highly developed in the random walk literature, e.g., in [23] and [26], leading to local limit theorems. We do not touch those ideas here but we think they could be fruitfully applied in the free probability context.

The paper [8] has been an important influence because, building upon operator-valued  $R$ -transform theory [37], [12] it reveals a rich algebraic and analytic structure to exploit for studying algebraicity. In particular, study of the fixed point equations stated in [8, Thm. 2.2] should in principle lead to a proof of Theorem 1. But because the many-variable setting for these equations is too difficult for us to handle, we work instead with a lightly modified version of the original setup of [35] and we use just the single classical parameter  $z$ . It remains an open and interesting problem to prove algebraicity in the many-variable setting of [8].

The linearization trick which we learned from the papers [16] and [17], which we refined and used in [1], and which in the refined form was also used in [8] plays here an important role as well. But as we have recently learned from [19] and wish to acknowledge here, the trick in its refined form already exists in the literature. A theorem of Schützenberger [30] (see also [7, Thm. 7.1, p. 18]) belonging to formal language theory contains the core idea of the trick sans the self-adjointness-preserving aspect. The whole trick is contained in [19, Lemma 4.1] and is called there *symmetric realization*. (There is a minor issue that the lemma is stated in the real case but the hermitian generalization is routine.) The cited lemma appears in a context perhaps not at first glance closely connected with free probability, but clearly and closely allied with linear systems theory [10], [20]. The author plans with several co-authors to report on these interconnections in a forthcoming paper.

As it happens, because Theorem 1 does not mention self-adjointness, all we need of the linearization/realization technique is Schützenberger’s theorem and of the latter we in fact need only a fragment, which we prove here “from scratch” in a paragraph. See Lemma 4.5.1 below. See also §4.5.2 for an explanation of the connection with Schützenberger’s theorem. Thus formal language theory does enter into our proof of Theorem 1 after all, but in a rather simple way.

In the paper [21], generalizing a result of the earlier paper [22], the following variant of Theorem 2 is proved. In this variant, instead of a group-ring element  $P \in \mathbb{C}[G]$  with complex coefficients, one considers a square matrix  $P \in \text{Mat}_n(\mathbb{Z}[G])$  of group-ring elements with integer coefficients, one considers not the usual generating function but rather the *zeta-function*

$$\exp\left(\sum_{k=1}^{\infty} \tau(\text{tr}(P^k)) \frac{t^k}{k}\right) \in 1 + t\mathbb{C}[[t]],$$

and finally, one concludes strikingly that the latter both has integer coefficients and is algebraic. Thus motivated, we raise the following question. Is “integral-algebraicity of zeta-functions” preserved under “integral-matrix-polynomial combination” of free random variables? Perhaps this question could be answered by combining methods used here with those of the cited papers.

As mentioned above, our formal algebraic setup for proving Theorem 1 is based on the original setup of [35]. From that starting point, we make the rest of our definitions so as to keep our approach to proving Theorem 1 as simple as possible. We make no positivity assumptions—moment sequences of variables can be arbitrarily prescribed sequences of complex numbers. We work over the field  $\mathbb{C}((1/z))$  of formal Laurent series, using simple ideas about Banach algebras over complete ultrametrically normed fields in lieu of operator theory over the complex numbers. Our method reveals nothing about the branch points of the algebraic functions it produces. It is an open problem to recover information about positivity and branch points. Perhaps this is only a matter of unifying features of the several theories mentioned above, but in our opinion some further ingredients from algebraic geometry will be needed. The soliton theory literature, e.g., [25], might provide guidance.

Here is an outline of the paper. In §2, after filling in background in leisurely fashion, and in particular writing down a simple algebraicity criterion, namely Proposition 2.4.2 below, we reformulate Theorem 1 as the conjunction of two propositions both of which concern the generalized Schwinger-Dyson equation. In §3 we introduce the formal algebraic variant of operator theory used in this paper. In §4 we introduce a suitable model of free random variables based on that of [35] and we deploy the linearization/realization technique, a.k.a. Schützenberger’s theorem. In §5 we exhibit the solutions of the generalized Schwinger-Dyson equation needed to prove Theorem 1. In the remainder of the paper we switch to the viewpoint of algebraic geometry and commutative algebra. In §6 we review topics connected with singularities of plane algebraic curves, especially Newton polygons. In §7 we apply the Weierstrass Preparation Theorem in a perhaps unexpected way. Finally, in §8 we complete the proof of Theorem 1 by checking hypotheses in Proposition 2.4.2.

Lastly, we remark that the paper is long only because we have included many explanations and reviews of background to smooth the way for the interested reader who might not be familiar with all the (seemingly) disparate materials collected here.

## 2. BACKGROUND FOR THE MAIN RESULT AND A REDUCTION OF THE PROOF

After recalling principal definitions, fixing notation, and filling in background for Theorem 1, we reduce Theorem 1 to two propositions each treating some aspect of the generalized Schwinger-Dyson equation.

**2.1. Noncommutative probability spaces and free independence.** We present a brief review to fix notation. See, e.g., [3], [28], [32], or [38] for background.

**2.1.1. Algebras.** All algebras in this paper are unital, associative, and have a scalar field containing  $\mathbb{C}$ . The unit of an algebra  $\mathcal{A}$  is denoted by  $1_{\mathcal{A}}$ ; other notation, e.g., simply 1, may be used when context permits. Given elements  $x_1, \dots, x_q \in \mathcal{A}$  of an algebra, let  $\mathbb{C}\langle x_1, \dots, x_q \rangle \subset \mathcal{A}$  denote the subring of  $\mathcal{A}$  generated by forming all finite  $\mathbb{C}$ -linear combinations of monomials in the given elements  $x_1, \dots, x_q$ , including the “empty monomial”  $1_{\mathcal{A}}$ . (But if  $\mathcal{A}$  is commutative, instead of  $\mathbb{C}\langle x_1, \dots, x_q \rangle$ , we write  $\mathbb{C}[x_1, \dots, x_q]$  as is usual in commutative algebra.)

**2.1.2. Noncommutative probability spaces.** A state  $\phi$  on an algebra  $\mathcal{A}$  is simply a  $\mathbb{C}$ -linear functional  $\phi : \mathcal{A} \rightarrow \mathbb{C}$  such that  $\phi(1_{\mathcal{A}}) = 1$ . In our formal algebraic setup no positivity constraints are imposed. A *noncommutative probability space* is a pair  $(\mathcal{A}, \phi)$  consisting of an algebra  $\mathcal{A}$  and a state  $\phi$  on that algebra. Given such a pair  $(\mathcal{A}, \phi)$ , elements of  $\mathcal{A}$  are called *noncommutative random variables*.

**2.1.3. Matrices with algebra entries.** Given an algebra  $\mathcal{A}$  and a positive integer  $n$ , let  $\text{Mat}_n(\mathcal{A})$  denote the algebra of  $n$ -by- $n$  matrices with entries in  $\mathcal{A}$ . More generally let  $\text{Mat}_{k \times \ell}(\mathcal{A})$  denote the space of  $k$ -by- $\ell$  matrices with entries in  $\mathcal{A}$ . For  $A \in \text{Mat}_{k \times \ell}(\mathbb{C})$  and  $a \in \mathcal{A}$  we define  $A \otimes a \in \text{Mat}_{k \times \ell}(\mathcal{A})$  by  $(A \otimes a)(i, j) = A(i, j)a$ . Let  $1 = I_n = I_n \otimes 1_{\mathcal{A}} \in \text{Mat}_n(\mathcal{A})$  denote the identity matrix, as context may permit. Let  $\mathbf{e}_{ij} \in \text{Mat}_{k \times \ell}(\mathbb{C})$  denote the elementary matrix with 1 in position  $(i, j)$  and 0 in every other position. Let  $\text{GL}_n(\mathcal{A})$  denote the group of invertible elements of  $\text{Mat}_n(\mathcal{A})$ . Given a noncommutative probability space  $(\mathcal{A}, \phi)$ , we regard each matrix  $A \in \text{Mat}_n(\mathcal{A})$  as a noncommutative random variable with respect to the state  $\phi_n : \text{Mat}_n(\mathcal{A}) \rightarrow \mathbb{C}$  given by the formula  $\phi_n(A) = \frac{1}{n} \sum_{i=1}^n \phi(A(i, i))$ .

**2.1.4. Free independence.** Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space and let  $\mathcal{A}_1, \dots, \mathcal{A}_q \subset \mathcal{A}$  be subalgebras such that  $1_{\mathcal{A}} \in \cap_{i=1}^q \mathcal{A}_i$ . One says that  $\mathcal{A}_1, \dots, \mathcal{A}_q$  are *freely independent* if for every positive integer  $k$ , sequence  $i_1, \dots, i_k \in \{1, \dots, q\}$  such that  $i_1 \neq i_2, i_2 \neq i_3, \dots, i_{k-1} \neq i_k$  and sequence  $x_1 \in \mathcal{A}_{i_1}, \dots, x_k \in \mathcal{A}_{i_k}$  such that  $\phi(x_1) = \dots = \phi(x_k) = 0$ , one has  $\phi(x_1 \cdots x_k) = 0$ . As a special case of the preceding general definition, one says that noncommutative random variables  $x_1, \dots, x_q \in \mathcal{A}$  are *freely independent* if the subalgebras  $\mathbb{C}\langle x_1 \rangle, \dots, \mathbb{C}\langle x_q \rangle \subset \mathcal{A}$  are freely independent.

**2.1.5. Univariate laws.** Let  $\mathbf{X}$  be a variable. A *univariate law* (or, context permitting, simply a *law*) is by definition a state  $\mu : \mathbb{C}\langle \mathbf{X} \rangle \rightarrow \mathbb{C}$  on the one-variable polynomial algebra  $\mathbb{C}\langle \mathbf{X} \rangle$ . The value  $\mu(\mathbf{X}^n) \in \mathbb{C}$  is called the  $n^{\text{th}}$  moment of  $\mu$ . Note that in our formal algebraic setup the moments of a law are allowed to be arbitrarily prescribed complex numbers. Given a noncommutative probability space  $(\mathcal{A}, \phi)$  and a noncommutative random variable  $x \in \mathcal{A}$ , the *law* of  $x$  is by definition the linear functional  $\mu_x : \mathbb{C}\langle \mathbf{X} \rangle \rightarrow \mathbb{C}$  determined by the formula  $\mu_x(\mathbf{X}^n) = \phi(x^n)$  for integers  $n \geq 0$ .

2.1.6. *Noncommutative joint laws.* Let  $\mathbf{X}_1, \dots, \mathbf{X}_q$  be independent noncommuting algebraic variables and let  $\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle$  be the noncommutative polynomial ring generated by these variables. A  $q$ -variable noncommutative law or, context permitting, simply a *law*, is a state on the algebra  $\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle$ . Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space and let  $x_1, \dots, x_q \in \mathcal{A}$  be noncommutative random variables. The *joint law*  $\mu_{x_1, \dots, x_q} : \mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle \rightarrow \mathbb{C}$  of the  $q$ -tuple  $(x_1, \dots, x_q)$  is by definition the linear functional defined by the rule  $\mu_{x_1, \dots, x_q}(f(\mathbf{X}_1, \dots, \mathbf{X}_q)) = \phi(f(x_1, \dots, x_q))$  for  $f(\mathbf{X}_1, \dots, \mathbf{X}_q) \in \mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle$ . The laws  $\mu_{x_1}, \dots, \mu_{x_q}$  of the individual variables (by analogy with classical probabilistic usage) are called the *marginal laws* for the joint law  $\mu_{x_1, \dots, x_q}$ . A point worth emphasizing is that if  $x_1, \dots, x_q$  are freely independent, then the joint law  $\mu_{x_1, \dots, x_q}$  is uniquely determined by the marginal laws  $\mu_{x_1}, \dots, \mu_{x_q}$ .

2.2. **The Laurent series field  $\mathbb{C}((1/z))$  and related notions.** We recall several definitions together providing a framework in which to discuss algebraicity. See the text [6] by Artin for background on valued fields and algebraic functions.

2.2.1. *Definition of  $\mathbb{C}((1/z))$  and related objects.* Let  $\mathbb{C}((1/z))$  denote the set of *formal Laurent series* in  $z$  of the form

$$(1) \quad f = \sum_{i \in \mathbb{Z}} c_i z^i \quad (c_i \in \mathbb{C} \text{ and } c_i = 0 \text{ for } i \gg 0).$$

(The coefficients  $c_i$  are not subject to any majorization.) Equipped with addition and multiplication in evident fashion, the set  $\mathbb{C}((1/z))$  becomes a field. Note that we have inclusions

$$\mathbb{C}[z] \subset \mathbb{C}(z) \subset \mathbb{C}((1/z)) \text{ and } \mathbb{C}[[1/z]] \subset \mathbb{C}((1/z))$$

where  $\mathbb{C}[z]$  is the ring of polynomials in  $z$ ,  $\mathbb{C}[[1/z]]$  is the ring of formal power series in  $1/z$ , and  $\mathbb{C}(z)$  is the field of rational functions of  $z$ , all with coefficients in  $\mathbb{C}$ . Note also that we have an additive direct sum decomposition

$$\mathbb{C}((1/z)) = \mathbb{C}[z] \oplus (1/z)\mathbb{C}[[1/z]].$$

In our algebraic setup the formal variable  $z$  corresponds to the classical parameter  $z$  in the upper half-plane.

2.2.2. *Algebraic elements of  $\mathbb{C}((1/z))$  and their irreducible equations.* Let  $\mathbb{C}[x, y]$  be the polynomial ring over  $\mathbb{C}$  in two independent (commuting) variables  $x$  and  $y$ . We say that  $f \in \mathbb{C}((1/z))$  is *algebraic* if one and hence all three of the following equivalent conditions hold:

- There exists some  $0 \neq P(x, y) \in \mathbb{C}[x, y]$  such that  $P(z, f) = 0$ .
- There exists some  $0 \neq Q(x, y) \in \mathbb{C}[x, y]$  such that  $Q(1/z, f) = 0$ .
- The field  $\mathbb{C}(z, f)$  generated over  $\mathbb{C}(z)$  by  $f$  is a vector space of finite dimension over  $\mathbb{C}(z)$ .

As is well-known, the algebraic elements form a subfield of  $\mathbb{C}((1/z))$  containing  $\mathbb{C}(z)$ . For algebraic  $f \in \mathbb{C}((1/z))$  there exists irreducible  $F(x, y) \in \mathbb{C}[x, y]$  unique up to a constant multiple such that  $F(1/z, f) = 0$ . (The insertion of  $1/z$  in the preceding definition rather than  $z$  is a technical convenience.) With but slight abuse of language we call any such irreducible polynomial the *irreducible equation* of  $f$ .

2.2.3. *Valuations.* For  $f \in \mathbb{C}((1/z))$  expanded as on line (1) we define

$$\text{val } f = \sup \{i \in \mathbb{Z} \mid c_i \neq 0\} = (\text{the valuation of } f) \in \mathbb{Z} \cup \{-\infty\}.$$

Note that

- (2)  $\text{val } f = -\infty \Leftrightarrow f = 0,$
- (3)  $\text{val}(f_1 f_2) = \text{val } f_1 + \text{val } f_2,$
- (4)  $\text{val}(f_1 + f_2) \leq \max(\text{val } f_1, \text{val } f_2)$  with equality if  $\text{val } f_1 \neq \text{val } f_2.$

Thus  $\text{val}$  is (the logarithm of) a nonarchimedean valuation in the sense of [6]. Thus it becomes possible to use (ultra)metric space ideas to reason about  $\mathbb{C}((1/z))$  and related objects, as in [6], and we will do so throughout this paper. We may speak for example of completeness. It is easy to see that  $\mathbb{C}((1/z))$  is complete with respect to the valuation  $\text{val}$ .

2.2.4. *Banach algebra structure for  $\text{Mat}_n(\mathbb{C}((1/z)))$ .* We now extend the metric space ideas a bit farther. We equip the matrix algebra  $\text{Mat}_n(\mathbb{C}((1/z)))$  with a valuation by the rule  $\text{val } A = \max_{i,j=1}^n \text{val } A(i,j)$ . Then (2) and (4) hold for matrices, (3) holds for multiplication of a matrix by a scalar, and (3) holds for multiplication of two matrices provided that “=” is relaxed to “ $\leq$ .” Thus  $\text{Mat}_n(\mathbb{C}((1/z)))$  becomes a Banach algebra over  $\mathbb{C}((1/z))$ . Later, in §3, a certain infinite-dimensional Banach algebra over  $\mathbb{C}((1/z))$  will be introduced.

2.2.5. *Composition of Laurent series.* The composition  $f \circ g \in \mathbb{C}((1/z))$  of  $f, g \in \mathbb{C}((1/z))$  is defined provided that  $\text{val } g > 0$ . The set  $z + \mathbb{C}[[1/z]]$  forms a group under composition. This group acts on the right side of  $\mathbb{C}((1/z))$  by  $\mathbb{C}$ -linear field automorphisms.

**Lemma 2.2.6.** *If  $f, g \in z + \mathbb{C}[[1/z]]$  satisfy  $f \circ g = z$  and  $f$  is algebraic, then  $g$  is also algebraic.*

*Proof.* For some  $0 \neq P(x, y) \in \mathbb{C}[x, y]$  we have  $0 = P(z, f) \circ g = P(g, z)$ . □

2.3. **Algebraicity of univariate laws.** We recall how to attach to each univariate law a (formal) Stieltjes transform and a (modified formal)  $R$ -transform *à la* Voiculescu. Then we recall how in terms of these transforms one can characterize algebraicity of a law.

2.3.1. *Formal Stieltjes transforms.* For a law  $\mu : \mathbb{C}\langle \mathbf{X} \rangle \rightarrow \mathbb{C}$ , the formal sum

$$S_\mu(z) = \sum_{n=0}^{\infty} \mu(\mathbf{X}^n) / z^{n+1} \in \mathbb{C}((1/z))$$

is by definition the *formal Stieltjes transform* of  $\mu$ . Hereafter we drop the adjective “formal” since no other kind of Stieltjes transform will be considered in this paper.

2.3.2. *Algebraicity of univariate laws.* A law  $\mu$  will be called *algebraic* if its Stieltjes transform  $S_\mu(z) \in \mathbb{C}((1/z))$  is algebraic.

2.3.3. *Free cumulants and  $R$ -transforms.* Given a law  $\mu : \mathbb{C}\langle \mathbf{X} \rangle \rightarrow \mathbb{C}$  one defines in free probability theory for each positive integer  $n$  the  $n^{\text{th}}$  free cumulant  $\kappa_n(\mu) \in \mathbb{C}$ . This can be done various ways, e.g., with generating functions or combinatorially using noncrossing partitions. See, e.g., [3], [28], or [38] for background; the founding document for this theory is [35]. The generating function

$$R_\mu(t) = \sum_{n=1}^{\infty} \kappa_n(\mu) t^{n-1} \in \mathbb{C}[[t]]$$

for the free cumulants is the formal version of the  $R$ -transform of Voiculescu. Hereafter we drop the adjective “formal” since no other kind of  $R$ -transform will be considered in this paper.

2.3.4. *Modified  $R$ -transforms.* To define and make calculations with free cumulants, we will use the generating function method. Consider the *modified  $R$ -transform*

$$\tilde{R}_\mu(z) = z + R_\mu(1/z) = z + \sum_{n=1}^{\infty} \kappa_n(\mu) z^{1-n} \in z + \mathbb{C}[[1/z]],$$

which we will find slightly more convenient. Obviously each of  $R_\mu(t)$  and  $\tilde{R}_\mu(z)$  uniquely determines the other. It is known (see [3], [28], [35], or [38]) that  $\tilde{R}_\mu(z)$  is the unique solution of the equation

$$(5) \quad \left( \frac{1}{S_\mu(z)} \right) \circ \tilde{R}_\mu(z) = z. \quad \left( \text{Equivalently: } \tilde{R}_\mu(z) \circ \left( \frac{1}{S_\mu(z)} \right) = z. \right)$$

Since  $z + \mathbb{C}[[1/z]]$  is a group under composition, the modified  $R$ -transform  $\tilde{R}_\mu(z)$  is well-defined for every law  $\mu$ , hence the sequence  $\{\kappa_n(\mu)\}_{n=1}^{\infty}$  of free cumulants is defined, and it uniquely determines  $\mu$ . Furthermore the free cumulants of a law can be arbitrarily prescribed.

The next lemma expresses algebraicity in terms of free cumulants.

**Lemma 2.3.5.** *Let  $\mu : \mathbb{C}\langle \mathbf{X} \rangle \rightarrow \mathbb{C}$  be a law. Then the following statements are equivalent:*

- (I)  $\mu$  is algebraic.
- (II)  $S_\mu(z)$  is algebraic.
- (III)  $\tilde{R}_\mu(z)$  is algebraic.

*Proof.* The equivalence (I) $\Leftrightarrow$ (II) holds by definition. The equivalence (II) $\Leftrightarrow$ (III) holds by Lemma 2.2.6, statement (5) above, and the fact that  $0 \neq f \in \mathbb{C}((1/z))$  is algebraic if and only if the reciprocal  $1/f$  is algebraic.  $\square$

2.4. **An algebraicity criterion.** We now present the algebraicity criterion which we will use to take the final step of the proof of Theorem 1. To do so we abruptly switch to the optic of commutative algebra.

2.4.1. *Setup for the criterion.* Let  $K/K_0$  be any extension of fields. Let  $x = (x_1, \dots, x_n)$  be an  $n$ -tuple of independent (commuting) variables and let  $K_0[x]$  be the polynomial ring generated over  $K_0$  by these variables. Let  $f = (f_1, \dots, f_n) = f(x) \in K_0[x]^n$  be an  $n$ -tuple of polynomials. Let  $J(x) = \det_{i,j=1}^n \frac{\partial f_i}{\partial x_j} \in K_0[x]$  be the determinant of the Jacobian matrix of  $f$ . Let  $\alpha = (\alpha_1, \dots, \alpha_n) \in K^n$  be an  $n$ -tuple such that  $f(\alpha) = 0$  but  $J(\alpha) \neq 0$ .

**Proposition 2.4.2.** *Notation and assumptions are as above. Every entry of the vector  $\alpha$  is algebraic over  $K_0$ .*

This statement is the same as [24, Prop. 5.3, Chap. VIII, p. 371], and also the same as [33, Lemma 6.6.9, Chap. 6, p. 206].

**2.5. Large-scale organization of the proof of Theorem 1.** We recall the generalized Schwinger-Dyson equation and then we state two technical results about it which together imply Theorem 1.

2.5.1. *The generalized Schwinger-Dyson equation.* In the two technical propositions to be formulated below we consider an instance

$$(6) \quad I_n + a^{(0)}g + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \kappa_j^{(\theta)} (a^{(\theta)}g)^j = 0$$

of the *generalized Schwinger-Dyson equation* for which the data are

$$(7) \quad \left\{ \begin{array}{l} \text{positive integers } q \text{ and } n, \\ \text{matrices } g, a^{(0)} \in \text{Mat}_n(\mathbb{C}((1/z))), \\ \text{matrices } a^{(1)}, \dots, a^{(q)} \in \text{Mat}_n(\mathbb{C}), \text{ and} \\ \text{a family } \left\{ \left\{ \kappa_j^{(\theta)} \right\}_{j=2}^{\infty} \right\}_{\theta=1}^q \text{ of complex numbers.} \end{array} \right.$$

We assume that

$$(8) \quad \lim_{j \rightarrow \infty} \text{val}(a^{(\theta)}g)^j = -\infty \text{ for } \theta = 1, \dots, q$$

in order that the left side of (6) have a well-defined value in  $\text{Mat}_n(\mathbb{C}((1/z)))$ . Furthermore, we impose the following nondegeneracy condition:

$$(9) \quad \begin{array}{l} \text{The linear map} \\ \left( h \mapsto a^{(0)}h + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \sum_{\nu=0}^{j-1} \kappa_j^{(\theta)} (a^{(\theta)}g)^{\nu} (a^{(\theta)}h) (a^{(\theta)}g)^{j-1-\nu} \right) \\ : \text{Mat}_n(\mathbb{C}((1/z))) \rightarrow \text{Mat}_n(\mathbb{C}((1/z))) \text{ is invertible.} \end{array}$$

Note that the map above is well-defined by assumption (8).

We will prove the following two results.

**Proposition 2.5.2.** *Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space. Let  $x_1, \dots, x_q \in \mathcal{A}$  be freely independent noncommutative random variables. Let*

$$X \in \text{Mat}_p(\mathbb{C}\langle x_1, \dots, x_q \rangle) \subset \text{Mat}_p(\mathcal{A})$$

*be a matrix. (Up to this point we have merely repeated the setup for Theorem 1.) For indices  $\theta = 1, \dots, q$  and  $j = 2, 3, 4, \dots$ , let  $\kappa_j^{(\theta)}$  denote the  $j^{\text{th}}$  free cumulant of the law of the noncommutative random variable  $x_{\theta}$ . Then the family  $\left\{ \left\{ \kappa_j^{(\theta)} \right\}_{\theta=1}^q \right\}_{j=2}^{\infty}$  of complex numbers for some integer  $n > p$  can be completed to a family*

$$\left( q, n, g, a^{(0)}, \left\{ a^{(\theta)} \right\}_{\theta=1}^q, \left\{ \left\{ \kappa_j^{(\theta)} \right\}_{j=2}^{\infty} \right\}_{\theta=1}^q \right)$$

of the form (7) satisfying (6), (8), and (9) along with the further conditions

$$(10) \quad a^{(0)} \in \{A + Bz \mid A, B \in \text{Mat}_n(\mathbb{C})\} \text{ and}$$

$$(11) \quad S_{\mu_X} = -\frac{1}{p} \sum_{i=1}^p g(i, i).$$

**Proposition 2.5.3.** *Let data of the form (7) satisfy (6), (8), and (9). Assume furthermore that*

$$(12) \quad a^{(0)} \in \text{Mat}_n(\mathbb{C}(z)) \text{ and}$$

$$(13) \quad \sum_{j=2}^{\infty} \kappa_{j+1}^{(\theta)} z^{-j} \in \mathbb{C}((1/z)) \text{ is algebraic for } \theta = 1, \dots, q.$$

*Then every entry of the matrix  $g$  is algebraic.*

See §4 and §5 below for the proof of Proposition 2.5.2. See §6, §7, and §8 below for the proof of Proposition 2.5.3.

2.5.4. *Reduction of the proof of Theorem 1.* In view of Lemma 2.3.5, it is clear that Propositions 2.5.2 and 2.5.3 together imply Theorem 1.

2.5.5. *Remark.* In the simple case

$$X = x_1 + \dots + x_q \in \mathbb{C}\langle x_1, \dots, x_q \rangle \subset \text{Mat}_1(\mathcal{A}) = \mathcal{A},$$

the instance of generalized Schwinger-Dyson equation emerging from the proof of Proposition 2.5.2 reduces to the standard fact [35] that the  $R$ -transform is additive for the addition of free random variables.

2.5.6. *Remark.* The (un)generalized Schwinger-Dyson equation is familiar in the case that  $\kappa_j^{(\theta)} = \kappa_j(\mu_{x_\theta}) = 0$  for  $j > 2$ . In the latter special case the equation (6) arises naturally in the study of free matrix-polynomial combinations of semicircular variables. See, e.g., [1], [3], [8], [16], [17], [18], [28] and [38].

2.5.7. *Remark.* Proposition 2.5.2 is unsurprising. It is proved here by straightforwardly combining three standard methodologies, namely:

- (i) the Boltzmann-Fock space model of free random variables,
- (ii) the linearization/realization method, and
- (iii) recursions of a type occurring in the study of random walk on infinite trees.

Methodology (i) clearly originates in [35]. Methodology (ii) we learned from the papers [16] and [17], and subsequently we refined it in [1], but as we have recently learned from [19], the essential point apart from the self-adjointness-preserving aspect is contained in Schützenberger's theorem [7, Thm. 7.1, p. 18]. Methodology (iii) has an obscure origin since random walk on many types of graphs has been studied in probability theory for decades and many methods for getting recursions have become commonplace. In this case we point to the examples [4], [23], [26] and [40] as inspirations, and refer the reader to [41] for background.

2.5.8. *Remark.* The form of Proposition 2.5.2 is in key respects quite similar to that of [8, Thm 2.2]. Indeed it could not be essentially different since it has the same origins in operator-valued free probability theory. In particular, (6) can be rewritten as a fixed point equation. But the matrix upper half-plane plays no role either in the statement or the proof of Proposition 2.5.2, greatly simplifying matters.

2.5.9. *Remark.* In the semicircular case  $\kappa_j^{(\theta)} = \kappa_j(\mu_{x_\theta}) = 0$  for  $j > 2$  remarked upon above, Proposition 2.5.3 follows straightforwardly from Propositions 2.4.2. In particular, in this relatively simple case, one can check the nondegeneracy hypothesis of Proposition 2.4.2 directly by using (9). This observation begins to explain the significance of condition (9). A similar approach proves Proposition 2.5.3 in the more general case that  $\kappa_j^{(\theta)} = \kappa_j(\mu_{x_\theta}) = 0$  for all but finitely many pairs  $(j, \theta)$ .

2.5.10. *Remark.* We continue in the setup of the preceding remark. Consider now the remaining case in which  $\kappa_j^{(\theta)} = \kappa_j(\mu_{x_\theta}) \neq 0$  for infinitely many pairs  $(j, \theta)$ . Then one can no longer prove Proposition 2.5.3 by applying Proposition 2.4.2 directly to solutions of the system of equations (6) because this system no longer consists exclusively of polynomial equations. This obstruction is the main difficulty of the proof of Theorem 1. We overcome the obstruction by means of the theory of Newton polygons (see §6 below) and the Weierstrass Preparation Theorem (see §7 below). Using these methods we are able to exhibit a system of  $n^2 + 3qn$  polynomial equations in  $n^2 + 3qn$  unknowns which in a useful sense contains the system (6) and to which Proposition 2.4.2 may be applied directly. (See §8 below.) This enlarged system of equations may be of independent interest. (See (84) below.)

### 3. HESSENBERG-TOEPLITZ MATRICES AND FREE CUMULANTS

We introduce the formal version of operator theory used in this paper and then as an illustration we revisit a key insight of Voiculescu concerning the free cumulants.

3.1. **The algebras  $\mathfrak{M}$  and  $\mathfrak{M}((1/z))$ .** We introduce two algebras of infinite matrices, the first an algebra over  $\mathbb{C}$  and the second a larger algebra over  $\mathbb{C}((1/z))$  possessing Banach algebra structure.

3.1.1. *Notation.* Let  $\mathbb{N}$  denote the set of **nonnegative** integers.

3.1.2. *The algebra  $\mathfrak{M}$ .* Let  $\mathfrak{M}$  denote the vector space over  $\mathbb{C}$  consisting of  $\mathbb{N}$ -by- $\mathbb{N}$  matrices  $M$  such that for each  $j \in \mathbb{N}$  there exist only finitely many  $i \in \mathbb{N}$  such that  $M(i, j) \neq 0$ . Note that for each  $M \in \mathfrak{M}$  the entry in the upper left corner is denoted by  $M(0, 0)$ , rather than  $M(1, 1)$ , because we are observing the convention that  $0 \in \mathbb{N}$ . Note that every upper-triangular  $\mathbb{N}$ -by- $\mathbb{N}$  matrix with entries in  $\mathbb{C}$  belongs to  $\mathfrak{M}$ . Informally,  $\mathfrak{M}$  consists of the “almost upper-triangular” matrices. It is easy to see that matrix multiplication is well-defined on  $\mathfrak{M}$  and moreover associative, thus making  $\mathfrak{M}$  into a unital associative algebra with scalar field  $\mathbb{C}$ . Indeed, it is clear that  $\mathfrak{M}$  is a copy of the algebra of linear endomorphisms of a complex vector space of countably infinite dimension. We write  $\mathbf{1} = 1_{\mathfrak{M}}$  to abbreviate notation. We equip  $\mathfrak{M}$  with the state  $\phi(M) = M(0, 0)$ , thus defining a noncommutative probability space  $(\mathfrak{M}, \phi)$ .

3.1.3. *The algebra  $\mathfrak{M}((1/z))$ .* Let  $\mathfrak{M}((1/z))$  denote the set of  $\mathbb{N}$ -by- $\mathbb{N}$  matrices  $M$  with entries in  $\mathbb{C}((1/z))$  satisfying one and hence both of the following equivalent conditions:

- There exists a *Laurent expansion*  $M = \sum_{n \in \mathbb{Z}} M_n z^n$  with coefficients  $M_n \in \mathfrak{M}$  such that  $M_n = 0$  for  $n \gg 0$ .
- One has  $\lim_{i \rightarrow \infty} \text{val } M(i, j) = -\infty$  for each  $j \in \mathbb{N}$  (without any requirement of uniformity in  $j$ ) and furthermore one has  $\sup_{i, j \in \mathbb{N}} \text{val } M(i, j) < \infty$ .

From the equivalent points of view described above it is clear that  $\mathfrak{M}((1/z))$  becomes a unital  $\mathbb{C}((1/z))$ -algebra with respect to the usual notion of matrix multiplication. For  $M \in \mathfrak{M}((1/z))$  we define  $\text{val } M = \sup_{i,j \in \mathbb{N}} \text{val } M(i, j)$ . With respect to the valuation function  $\text{val}$  thus extended to  $\mathfrak{M}((1/z))$ , the latter becomes a unital Banach algebra over  $\mathbb{C}((1/z))$ . We write  $\mathbf{1} = 1_{\mathfrak{M}} = 1_{\mathfrak{M}((1/z))}$ .

3.1.4. *Elementary matrices and an abuse of notation.* Let  $\mathbf{e}[i, j] \in \mathfrak{M}$  denote the elementary matrix with entries given by the rule

$$\mathbf{e}[i, j](k, \ell) = \delta_{ik}\delta_{j\ell} \text{ for } i, j, k, \ell \in \mathbb{N}.$$

The notation  $\mathbf{e}[i, j] \in \mathfrak{M}$  introduced here is intended to contrast with the notation  $\mathbf{e}_{ij} \in \text{Mat}_{k \times \ell}(\mathbb{C})$  previously introduced for elementary matrices with finitely many rows and columns. For  $M \in \mathfrak{M}$  supported in a set  $S \subset \mathbb{N} \times \mathbb{N}$  which intersects each column  $\mathbb{N} \times \{j\}$  in a finite set, we abuse notation by writing

$$M = \sum_{(i,j) \in S} M(i, j)\mathbf{e}[i, j]$$

as a shorthand to indicate the placement of entries of  $M$ .

The following simple lemma is a key motivation for the definition of  $\mathfrak{M}((1/z))$ .

**Lemma 3.1.5.** *Fix  $M \in \mathfrak{M}$  arbitrarily and let  $\mu$  denote the law of  $M$ . Then the matrix  $z\mathbf{1} - M \in \mathfrak{M}((1/z))$  is invertible and  $S_{\mu}(z) = (z\mathbf{1} - M)^{-1}(0, 0)$ .*

*Proof.* One has

$$(z\mathbf{1} - M)^{-1} = \frac{1}{z} \sum_{k=0}^{\infty} \frac{M^k}{z^k} \in \mathfrak{M}((1/z)).$$

The geometric series here is convergent because  $\text{val } \frac{M}{z} < 0$ . This noted, it is clear that the series  $(z\mathbf{1} - M)^{-1}(0, 0)$  is the Stieltjes transform of the law of  $M$ .  $\square$

3.1.6. *Remark.* Our setup is inspired by (but is much simpler than) that of [5], and thus belongs to the lineage of [34]. The theory of the  $R$ -transform overlaps in an interesting way with the theory of residues developed in [34], one point of contact being the notion of a Hessenberg-Toeplitz matrix. (See immediately below.) This connection deserves further investigation.

## 3.2. Hessenberg-Toeplitz matrices.

3.2.1. *Basic definitions.* Let  $\{\kappa_j\}_{j=1}^{\infty}$  be any sequence of complex numbers. Consider the infinite matrix

$$(14) \quad H = \begin{bmatrix} \kappa_1 & \kappa_2 & \kappa_3 & \dots & & & & \\ & 1 & \kappa_1 & \kappa_2 & \kappa_3 & \dots & & \\ & & 1 & \kappa_1 & \kappa_2 & \kappa_3 & \dots & \\ & & & 1 & \kappa_1 & \kappa_2 & \kappa_3 & \dots \\ & & & & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \in \mathfrak{M}.$$

Equivalently, in terms of the elementary matrices  $\mathbf{e}[i, j] \in \mathfrak{M}$  we have

$$(15) \quad H = \sum_{k \in \mathbb{N}} \left( \mathbf{e}[1+k, k] + \sum_{j \in \mathbb{N}} \kappa_{j+1} \mathbf{e}[k, j+k] \right).$$

The matrix  $H$  displays the (upper) *Hessenberg* pattern:  $i > j + 1 \Rightarrow H(i, j) = 0$  for  $i, j \in \mathbb{N}$ . The matrix  $H$  also displays the *Toeplitz* pattern:  $H(i + 1, j + 1) = H(i, j)$  for  $i, j \in \mathbb{N}$ . Accordingly we call  $H$  a *Hessenberg-Toeplitz matrix*.

The reason for our interest in the matrix  $H$  is explained by the next lemma.

**Lemma 3.2.2.** *Assumptions and notation are as above. Then for every positive integer  $j$ , the  $j^{\text{th}}$  free cumulant of  $H$  viewed as a noncommutative random variable in the noncommutative probability space  $(\mathfrak{M}, \phi)$  equals  $\kappa_j$ .*

This fact is well-known—it is a key insight for Voiculescu’s theory of the  $R$ -transform [35]. It is therefore not necessary to give a proof. But we nevertheless give a proof in §3.4 below by manipulation of block-decomposed matrices in order to set the pattern for the more difficult calculations undertaken in §5 below.

3.2.3. *Remark.* If one patiently works through the definitions and uses the theory of residues from [34], one can see that Lemma 3.2.2 is equivalent to the one-variable Lagrange Inversion Formula.

**3.3. Inversion of block-decomposed matrices.** We pause to review a method of calculation used repeatedly in the sequel.

3.3.1. *Basic identities.* Let

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

be an invertible square matrix (in practice infinite) decomposed into blocks where  $\mathbf{a}$  and  $\mathbf{d}$  are square and  $\mathbf{d}$  is also invertible. Then we have a factorization

$$(16) \quad \begin{bmatrix} \mathbf{a} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c} & \mathbf{0} \\ \mathbf{0} & \mathbf{d} \end{bmatrix} = \begin{bmatrix} \mathbf{1} & -\mathbf{b}\mathbf{d}^{-1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix} \begin{bmatrix} \mathbf{1} & \mathbf{0} \\ -\mathbf{d}^{-1}\mathbf{c} & \mathbf{1} \end{bmatrix}$$

from which in particular we infer that the *Schur complement*  $\mathbf{a} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c}$  is invertible. Let

$$\mathbf{g} = (\mathbf{a} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c})^{-1}.$$

From (16) one then straightforwardly derives the inversion formula

$$(17) \quad \begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{d}^{-1} \end{bmatrix} + \begin{bmatrix} \mathbf{1} \\ -\mathbf{d}^{-1}\mathbf{c} \end{bmatrix} \mathbf{g} \begin{bmatrix} \mathbf{1} & -\mathbf{b}\mathbf{d}^{-1} \end{bmatrix}.$$

The latter formula also shows that invertibility of  $\mathbf{d}$  and  $\mathbf{a} - \mathbf{b}\mathbf{d}^{-1}\mathbf{c}$  implies invertibility of  $\begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}$ . For convenient application in §5, we restate in abstract form a couple of relations among blocks recorded in formula (17).

**Lemma 3.3.2.** *Let  $\mathcal{A}$  be a unital associative algebra (perhaps not commutative). Let  $\pi, \sigma \in \mathcal{A}$  satisfy  $\pi^2 = \pi \neq 0$ ,  $\sigma^2 = \sigma \neq 0$ ,  $\pi\sigma = \sigma\pi = 0$  and  $1_{\mathcal{A}} = \pi + \sigma$ . Let  $A \in \mathcal{A}$  be invertible. Assume furthermore that  $\sigma A \sigma$  is invertible in the algebra  $\sigma \mathcal{A} \sigma$  and let  $A_{\sigma}^{-1}$  denote the inverse. Then we have*

$$(18) \quad \sigma A^{-1} \pi = -A_{\sigma}^{-1} A \pi A^{-1} \pi \quad \text{and}$$

$$(19) \quad A^{-1} \sigma = (1_{\mathcal{A}} - A^{-1} \pi A \sigma) A_{\sigma}^{-1}.$$

*Proof.* We have

$$\sigma A \sigma A^{-1} \pi = -\sigma A A^{-1} \pi + \sigma A \sigma A^{-1} \pi = -\sigma A \pi A^{-1} \pi.$$

Now left-multiply extreme terms by  $A_\sigma^{-1}$  to recover (18). Similarly, we have

$$\sigma = \sigma A A_\sigma^{-1} = (A - \pi A \sigma) A_\sigma^{-1}.$$

Now left-multiply extreme terms by  $A^{-1}$  to recover (19).  $\square$

**3.4. Proof of Lemma 3.2.2.** Consider the Laurent series

$$f = f(z) = z + \sum_{j=1}^{\infty} \kappa_j z^{j-1} \in z + \mathbb{C}[[1/z]].$$

It will suffice to show that  $f$  is equal to the modified  $R$ -transform of the law of  $H$ . Consider also the Stieltjes transform

$$g = g(z) = S_{\mu_H}(z) \in (1/z) + (1/z^2)\mathbb{C}[[1/z]]$$

of the law of  $H$ . Since  $z + \mathbb{C}[[1/z]]$  forms a group under composition, it will suffice to show that  $z = f \circ \frac{1}{g}$ , equivalently  $z = g^{-1} + \sum_{j=1}^{\infty} \kappa_j g^{j-1}$ , or equivalently

$$1 = (z - \kappa_1)g - \sum_{j=2}^{\infty} \kappa_j g^j.$$

Let

$$A = z\mathbf{1} - H \in \mathfrak{M}((1/z)).$$

By Lemma 3.1.5 the inverse

$$G = A^{-1} \in \mathfrak{M}((1/z))$$

exists and furthermore

$$g = G(0, 0).$$

In view of the relation

$$1 = \sum_{k \in \mathbb{N}} A(0, k)G(k, 0)$$

holding because  $G = A^{-1}$ , it will be enough simply to prove that

$$(20) \quad G(i, 0) = g^{i+1} \text{ for } i \in \mathbb{N}.$$

Now with an eye toward applying (17) above, consider the block decomposition

$$A = \begin{bmatrix} \mathbf{a} & \mathbf{b} \\ \mathbf{c} & \mathbf{d} \end{bmatrix}$$

where

$$\mathbf{a} = z - \kappa_1, \quad \mathbf{b} = -[\kappa_2 \quad \kappa_3 \quad \dots], \quad \mathbf{c} = -\begin{bmatrix} 1 \\ 0 \\ \vdots \end{bmatrix}, \quad \text{and } \mathbf{d} = A.$$

By (17) we have

$$\begin{bmatrix} G(1, 0) \\ G(2, 0) \\ \vdots \end{bmatrix} = -\mathbf{d}^{-1} \mathbf{c} \mathbf{g} = \begin{bmatrix} G(0, 0) \\ G(1, 0) \\ \vdots \end{bmatrix} g,$$

whence (20). The proof is complete.  $\square$

## 4. THE LINEARIZATION STEP

In this section we apply (a fragment of) Schützenberger’s Theorem to a suitable model of free noncommutative random variables with prescribed free cumulants, thus advancing the proof of Proposition 2.5.2.

**4.1. Stars and diamonds.** We build a model for the free unital associative monoid on  $q$  generators for which  $\mathbb{N}$  is the underlying set. Using this monoid structure we will be able to construct and manipulate usefully patterned matrices in  $\mathfrak{M}((1/z))$ .

**4.1.1. Notation.** The parameter  $q$  figuring in Theorem 1 is considered fixed throughout the remainder of the paper. Many constructs below depend on  $q$  but the notation frequently does not show this.

**4.1.2. Improper representations to the base  $q$ .** Suppose at first that  $q > 1$ . In grade school one learns to represent nonnegative integers to the base  $q$  using place notation and digits selected from the set  $\{0, \dots, q-1\}$ . It is not hard to see that using instead digits selected from the set  $\{1, \dots, q\}$  one still gets a unique representation for every member of  $\mathbb{N}$ , it being understood that 0 is represented by the empty digit string  $\emptyset$ . A representation to the base  $q$  of a nonnegative integer using digits  $\{1, \dots, q\}$  will be called *improper*. Improper representations to the base  $q$  make sense also for  $q = 1$ . In the latter extreme case each  $x \in \mathbb{N}$  is represented by a string of 1’s of length  $x$ .

**4.1.3. Example: counting improperly to the base 3.**

$$\emptyset, 1, 2, 3, 11, 12, 13, 21, 22, 23, 31, 32, 33, 111, 112, 113, 121, 122, \dots$$

**4.1.4. The binary operation  $\star_q$ .** We define the binary operation

$$\star = \star_q : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$$

by the rule

$$x \star y = xq^\ell + y \quad \text{where } \ell \in \mathbb{N} \text{ satisfies } \frac{q^\ell - 1}{q - 1} \leq y \leq q \cdot \left( \frac{q^\ell - 1}{q - 1} \right).$$

Informally,  $\ell$  is the number of digits in the improper base  $q$  representation of  $y$ , and  $x \star y$  is the concatenation of the digit strings of  $x$  and of  $y$ . The operation  $\star$  makes  $\mathbb{N}$  into a free associative monoid freely generated by the digits  $1, \dots, q$  with 0 as the identity element.

**Lemma 4.1.5.**  $\mathbb{N} \setminus \{0\}$  is the disjoint union of the sets  $\mathbb{N} \star \theta$  for  $\theta = 1, \dots, q$ .

There is nothing to prove. We record this for convenient reference since however trivial, this is an important point in a later proof.

**4.1.6. The binary operation  $\diamond_q$ .** We define the binary operation

$$\diamond = \diamond_q : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$$

by the formula

$$x^{\diamond y} = \underbrace{x \star \dots \star x}_y.$$

We use exponential-style notation to emphasize the analogy with exponentiation in the usual sense.

**Lemma 4.1.7.** *For  $\theta = 1, \dots, q$ , every  $x \in \mathbb{N}$  has a unique factorization  $x = \theta^{\circ i} \star k$  where  $i \in \mathbb{N}$  and  $k \in \mathbb{N} \setminus \theta \star \mathbb{N}$ .*

Again, there is nothing to prove. We record this for convenient reference.

4.1.8. *Remark.* Consider the graph  $\Gamma = \Gamma_q$  with vertex set  $\mathbb{N}$  and an edge connecting  $x$  to  $\theta \star x$  for each pair  $(\theta, x) \in \{1, \dots, q\} \times \mathbb{N}$ . With  $0 \in \mathbb{N}$  designated as the root, the resulting graph  $\Gamma$  is an infinite rooted planar tree in which every vertex has a “birth-ordered” set of  $q$  children, i.e., a  $q$ -ary rooted tree. We do not explicitly use the  $q$ -ary tree in this paper because we instead rely on the monoid  $(\mathbb{N}, \star)$  to do our bookkeeping. Nonetheless the notion of the  $q$ -ary tree remains a strong guide to intuition.

4.2. **Free random variables with prescribed free cumulants.** The next proposition exhibits a model for  $q$  free noncommutative random variables with prescribed free cumulants. The model is essentially the same as that used in [35], but with the notation designed to make recursions easy to see.

**Proposition 4.2.1.** *Let*

$$\left\{ \left\{ \kappa_j^{(\theta)} \right\}_{j=1}^{\infty} \right\}_{\theta=1}^q$$

*be any family of complex numbers. Then the family*

$$(21) \quad \sum_{k \in \mathbb{N}} \mathbf{e}[\theta \star k, k] + \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{N}} \kappa_{j+1}^{(\theta)} \mathbf{e}[k, \theta^{\circ j} \star k] \in \mathfrak{M} \text{ for } \theta = 1, \dots, q$$

*of noncommutative random variables is freely independent and moreover the  $j^{\text{th}}$  free cumulant of the  $\theta^{\text{th}}$  noncommutative random variable equals  $\kappa_j^{(\theta)}$ .*

The proof requires some preparation and is completed in §4.2.4.

4.2.2. *Self-embeddings of  $\mathfrak{M}$ .* For  $\theta = 1, \dots, q$  and  $A \in \mathfrak{M}$  we define

$$(22) \quad A^{(\theta)} = \sum_{k \in \mathbb{N} \setminus \theta \star \mathbb{N}} \sum_{i, j \in \mathbb{N}} A(i, j) \mathbf{e}[\theta^{\circ i} \star k, \theta^{\circ j} \star k].$$

By Lemma 4.1.7 the matrix  $A^{(\theta)}$  is block-diagonal with copies of  $A$  indexed by  $\mathbb{N} \setminus \theta \star \mathbb{N}$  repeated along the diagonal. Thus the map  $(A \mapsto A^{(\theta)}) : \mathfrak{M} \rightarrow \mathfrak{M}$  is a unital one-to-one homomorphism of algebras. Note that  $A^{(\theta)}(0, 0) = A(0, 0)$  and hence the map  $A \mapsto A^{(\theta)}$  is law-preserving. Let  $\mathfrak{M}^{(\theta)}$  denote the embedded image of  $\mathfrak{M}$  under the map  $A \mapsto A^{(\theta)}$ .

**Lemma 4.2.3.** *The subalgebras  $\mathfrak{M}^{(1)}, \dots, \mathfrak{M}^{(q)} \subset \mathfrak{M}$  are freely independent.*

*Proof.* Fix  $\theta_1, \dots, \theta_k \in \{1, \dots, q\}$  such that

$$\theta_1 \neq \theta_2, \theta_2 \neq \theta_3, \dots, \theta_{k-1} \neq \theta_k.$$

Fix  $A_1, \dots, A_k \in \mathfrak{M}$  such that

$$A_1(0, 0) = \dots = A_k(0, 0) = 0.$$

Our task is to verify that

$$(23) \quad \sum_{(i_1, \dots, i_{k-1}) \in \mathbb{N}^{k-1}} A_1^{(\theta_1)}(0, i_1) A_2^{(\theta_2)}(i_1, i_2) \cdots A_{k-1}^{(\theta_{k-1})}(i_{k-2}, i_{k-1}) A_k^{(\theta_k)}(i_{k-1}, 0) \\ = (A_1^{(\theta_1)} \cdots A_k^{(\theta_k)})(0, 0) = 0.$$

Now by definition, for any matrix  $A \in \mathfrak{M}$  such that  $A(0, 0) = 0$  and  $\theta = 1, \dots, q$ , the matrix entry  $A^{(\theta)}(i, j)$  vanishes unless  $i \in \theta \star \mathbb{N}$  or  $j \in \theta \star \mathbb{N}$ . It follows that for any  $(k-1)$ -tuple  $(i_1, \dots, i_{k-1}) \in \mathbb{N}^{k-1}$  the corresponding term in the sum on the left side of (23) vanishes. Thus, *a fortiori*, (23) holds.  $\square$

4.2.4. *Proof of Proposition 4.2.1.* Let  $H_\theta$  be a copy of the matrix  $H$  defined in (14) and equivalently in (15), with  $\kappa_j$  replaced by  $\kappa_j^{(\theta)}$ . By Lemma 3.2.2 we know already that the  $j^{\text{th}}$  free cumulant of the law of  $H_\theta$  equals  $\kappa_j^{(\theta)}$ . Substituting directly into the definition (22) we have

$$H_\theta^{(\theta)} = \sum_{i \in \mathbb{N}} \sum_{k \in \mathbb{N} \setminus \theta \star \mathbb{N}} \mathbf{e}[\theta^{\circ(i+1)} \star k, \theta^{\circ i} \star k] + \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{N} \setminus \theta \star \mathbb{N}} \kappa_{j+1}^{(\theta)} \mathbf{e}[\theta^{\circ i} \star k, \theta^{\circ(i+j)} \star k].$$

The result follows now via Lemmas 4.1.7 and 4.2.3.  $\square$

4.2.5. *Remark.* Voiculescu [35] introduced the Boltzmann-Fock space model of free random variables using lowering and raising operators for his striking proof of additivity of the  $R$ -transform for addition of free random variables. Also see [3, Cor. 5.3.23] and its proof for a quick review of this material. Proposition 4.2.1 is merely a description of the Boltzmann-Fock space model using notation chosen to make recursions more easily accessible. In the setup of Proposition 4.2.1 the matrices

$$(24) \quad \hat{\lambda}^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[\theta \star k, k] \in \mathfrak{M} \quad \text{and} \quad \lambda^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[k, \theta \star k] \in \mathfrak{M} \quad \text{for } \theta = 1, \dots, q$$

correspond to the lowering and raising operators considered in [35], respectively. Note that using the operators on line (24) we can rewrite the operators on (21) in the more familiar form

$$(25) \quad \hat{\lambda}^{(\theta)} + \sum_{j \in \mathbb{N}} \kappa_{j+1}^{(\theta)} (\lambda^{(\theta)})^j \in \mathfrak{M} \quad \text{for } \theta = 1, \dots, q.$$

The infinite sum here is an abuse of notation but it nonetheless makes sense because the matrices being summed have disjoint supports only finitely many of which meet any given column. Later we will also consider lowering and raising operators

$$(26) \quad \hat{\rho}^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[k \star \theta, k] \in \mathfrak{M} \quad \text{and} \quad \rho^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[k, k \star \theta] \in \mathfrak{M} \quad \text{for } \theta = 1, \dots, q$$

acting (so to speak) on the right rather than the left. The relations

$$(27) \quad \lambda^{(\theta)} \rho^{(\theta')} = \rho^{(\theta')} \lambda^{(\theta)}, \quad \hat{\lambda}^{(\theta)} \hat{\rho}^{(\theta')} = \hat{\rho}^{(\theta')} \hat{\lambda}^{(\theta)},$$

$$(28) \quad \rho^{(\theta')} \hat{\rho}^{(\theta)} = \delta_{\theta\theta'} \mathbf{1}, \quad \text{and} \quad \sum_{\alpha=1}^q \hat{\rho}^{(\alpha)} \rho^{(\alpha)} = \sum_{i=1}^{\infty} \mathbf{e}[i, i]$$

for  $\theta, \theta' = 1, \dots, q$  are easy to verify. See for example [1, Sec. 3.4] where these and further relations are written out as part of an analysis leading (without any reference to noncrossing partitions) to the Schwinger-Dyson equation for semicircular variables.

4.2.6. *Remark.* The interplay of left and right lowering and raising operators is a fundamental feature of the recently introduced bi-free framework of [36].

4.3. **Kronecker products and the isomorphism  $\natural$ .** We introduce notation which is rather tedious to define but convenient to calculate with.

4.3.1. *Classical Kronecker products.* Recall that for matrices of finite size the *Kronecker product*

$$A^{(1)} \otimes A^{(2)} \in \text{Mat}_{k_1 k_2 \times \ell_1 \ell_2}(\mathbb{C}) \quad \left( A^{(\alpha)} \in \text{Mat}_{k_\alpha \times \ell_\alpha}(\mathbb{C}) \text{ for } \alpha = 1, 2 \right)$$

is defined by the rule

$$A^{(1)} \otimes A^{(2)} = \begin{bmatrix} A^{(1)}(1, 1)A^{(2)} & \dots & A^{(1)}(1, \ell_1)A^{(2)} \\ \vdots & & \vdots \\ A^{(1)}(k_1, 1)A^{(2)} & \dots & A^{(1)}(k_1, \ell_1)A^{(2)} \end{bmatrix}$$

or equivalently and more explicitly (if more clumsily)

$$A^{(1)} \otimes A^{(2)}(n(i_1 - 1) + i_2, n(j_1 - 1) + j_2) = A^{(1)}(i_1, j_1)A^{(2)}(i_2, j_2) \\ \text{for } \alpha = 1, 2, i_\alpha = 1, \dots, k_\alpha, \text{ and } j_\alpha = 1, \dots, \ell_\alpha.$$

4.3.2. *Kronecker products involving infinite matrices.* In the mixed infinite/finite case we define the *Kronecker product*

$$x \otimes a \in \mathfrak{M}((1/z)) \quad (x \in \mathfrak{M}((1/z)) \text{ and } a \in \text{Mat}_n(\mathbb{C}((1/z))))$$

by the rule

$$x \otimes a = \begin{bmatrix} x(0, 0)a & x(0, 1)a & \dots \\ x(1, 0)a & x(1, 1)a & \dots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

or equivalently and more explicitly

$$(x \otimes a)(i_1 n + i_2 - 1, j_1 n + j_2 - 1) = x(i_1, j_1)a(i_2, j_2) \\ \text{for } i_1, j_1 \in \mathbb{N} \text{ and } i_2, j_2 = 1, \dots, n.$$

We also define

$$a \otimes x \in \text{Mat}_n(\mathfrak{M}((1/z))) \quad (a \in \text{Mat}_n(\mathbb{C}((1/z))) \text{ and } x \in \mathfrak{M}((1/z)))$$

by the somewhat ungainly iterated index formula

$$((a \otimes x)(i_2, j_2))(i_1, j_1) = a(i_2, j_2)x(i_1, j_1) \\ \text{for } i_1, j_1 \in \mathbb{N} \text{ and } i_2, j_2 = 1, \dots, n.$$

4.3.3. *The operation  $\natural$ .* For  $M \in \text{Mat}_n(\mathfrak{M}((1/z)))$  we define  $M^\natural \in \mathfrak{M}((1/z))$  by the formula

$$M^\natural = \sum_{i_1, j_1 \in \mathbb{N}} \sum_{i_2, j_2=1}^n (M(i_2, j_2)(i_1, j_1))(\mathbf{e}[i_1, j_1] \otimes \mathbf{e}_{i_2 j_2}),$$

thus defining an isometric isomorphism

$$(M \mapsto M^\natural) : \text{Mat}_n(\mathfrak{M}((1/z))) \rightarrow \mathfrak{M}((1/z))$$

of Banach algebras over  $\mathbb{C}((1/z))$ , where the source algebra is given Banach algebra structure by declaring that  $\text{val } A = \max_{i, j=1}^n \text{val } A(i, j)$  for  $A \in \text{Mat}_n(\mathfrak{M}((1/z)))$ . Finally, note that

$$(29) \quad (a \otimes x)^\natural = x \otimes a \quad \text{for } a \in \text{Mat}_n(\mathbb{C}((1/z))) \text{ and } x \in \mathfrak{M}((1/z)).$$

Thus the operation  $\natural$  has a natural interpretation as exchange of tensor factors.

**Lemma 4.3.4.** *Fix  $A \in \text{Mat}_n(\mathfrak{M})$ . Then the following statements hold:*

$zI_n \otimes \mathbf{1} - A \in \text{Mat}_n(\mathfrak{M}((1/z)))$  and  $z\mathbf{1} - A^{\natural} \in \mathfrak{M}((1/z))$  are invertible.

$$\left( (zI_n \otimes \mathbf{1} - A)^{-1} \right)^{\natural} = (z\mathbf{1} - A^{\natural})^{-1}.$$

$$S_{\mu_A}(z) = \frac{1}{n} \sum_{i=0}^{n-1} (z - A^{\natural})^{-1}(i, i) = \left( \frac{1}{n} \sum_{i=1}^n (zI_n \otimes \mathbf{1} - A)^{-1}(i, i) \right) (0, 0).$$

This statement supplements Lemma 3.1.5 only by some minor bookkeeping details. We therefore omit proof.

**4.4. Digital linearization.** Here is the main result in this section.

**Proposition 4.4.1.** *Let  $(\mathcal{A}, \phi)$  be a noncommutative probability space. Let  $x_1, \dots, x_q \in \mathcal{A}$  be freely independent noncommutative random variables. Fix  $X \in \text{Mat}_p(\mathbb{C}\langle x_1, \dots, x_q \rangle) \subset \text{Mat}_p(\mathcal{A})$ . Let  $\kappa_j^{(\theta)} = \kappa_j(\mu_{x_\theta})$  for  $j = 1, 2, \dots$  and  $\theta = 1, \dots, q$ . Then there exist for some  $N > 0$  matrices  $L_0, L_1, \dots, L_q \in \text{Mat}_{p+N}(\mathbb{C})$  with the following properties:*

(30)  $L_0$  vanishes identically in the upper left  $p$ -by- $p$  block.

(31)  $L_1, \dots, L_q$  are supported in the lower right  $N$ -by- $N$  block.

(32) 
$$L = \mathbf{1} \otimes \left( L_0 + \begin{bmatrix} zI_p & 0 \\ 0 & 0 \end{bmatrix} \right) + \sum_{\theta=1}^q \sum_{k \in \mathbb{N}} \mathbf{e}[\theta \star k, k] \otimes L_\theta$$

$$+ \sum_{\theta=1}^q \sum_{j \in \mathbb{N}} \sum_{k \in \mathbb{N}} \kappa_{j+1}^{(\theta)} \mathbf{e}[k, \theta^{\circ j} \star k] \otimes L_\theta \in \mathfrak{M}((1/z))$$
 is invertible.

(33) 
$$S_{\mu_X}(z) = \frac{1}{p} \sum_{i=0}^{p-1} L^{-1}(i, i).$$

We call  $L$  a *digital linearization* of  $X$ . It is worth remarking that this linearization is thoroughgoing in the sense that not only do the variables  $x_1, \dots, x_q$  appear linearly—so also does the variable  $z$ . The proof will be completed in §4.6 below.

**4.4.2. Remark.** Picking up again on the idea mentioned in §4.1.8, and adopting the absurd point of view that probabilities can be square matrices with arbitrary complex number entries, the matrix  $L$  describes a random walk on the  $q$ -ary tree  $\Gamma_q$  such that from a given vertex  $x \in \mathbb{N}$ , one may (i) step one unit back toward the root (if not already at the root), (ii) stay in place, or (iii) step away from the root arbitrarily far along along a geodesic  $\{\theta^{\circ i} \star x \mid i \in \mathbb{N}\}$  for some  $\theta \in \{1, \dots, q\}$ . Whether or not this interpretation of  $L$  is absurd, it does make random walk intuition available to analyze  $L$ . Guided by this intuition we will prove in §5 below that the generalized Schwinger-Dyson equation holds for “upper left corners” of matrices of the form  $L$ , as well as for more general infinite matrices.

**4.5. Schützenberger’s theorem.** The next lemma recalls what we need of the self-adjoint linearization trick, and as we have already noted in the introduction, what we need boils down to Schützenberger’s Theorem [7, Thm. 7.1]. In fact we need only a quite specialized consequence of this theorem, or rather of its proof, simple enough to prove quickly from scratch, as follows.

**Lemma 4.5.1.** *For each  $f \in \text{Mat}_p(\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle)$  there exists a factorization  $f = bd^{-1}c$  (called a linearization of  $f$ ) where*

$$b \in \text{Mat}_{p \times N}(\mathbb{C}), \quad c \in \text{Mat}_{N \times p}(\mathbb{C}), \quad d \in \text{GL}_N(\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle),$$

and each entry of  $d$  belongs to the  $\mathbb{C}$ -linear span of  $1, \mathbf{X}_1, \dots, \mathbf{X}_q$ .

Note that the proof below actually produces  $d$  with the further property that  $d - I_N$  is strictly upper triangular.

*Proof.* If every entry of  $f$  belongs to the  $\mathbb{C}$ -linear span of  $1, \mathbf{X}_1, \dots, \mathbf{X}_q$ , then, say,

$$f = \begin{bmatrix} I_p & 0 \end{bmatrix} \begin{bmatrix} I_p & -f \\ 0 & I_p \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ I_p \end{bmatrix}$$

is a linearization. Thus it will be enough to demonstrate that given linearizable  $f_1, f_2 \in \text{Mat}_p(\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle)$ , again  $f_1 + f_2$  and  $f_1 f_2$  are linearizable. So suppose that  $f_i = b_i d_i^{-1} c_i$  for  $i = 1, 2$  are factorizations of the desired form. We then have

$$f_1 + f_2 = \begin{bmatrix} b_1 & b_2 \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}^{-1} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \quad \text{and}$$

$$f_1 f_2 = \begin{bmatrix} b_1 & 0 & 0 \end{bmatrix} \begin{bmatrix} d_1 & c_1 & 0 \\ 0 & 1 & b_2 \\ 0 & 0 & d_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ c_2 \end{bmatrix}.$$

To assist the reader in checking the second formula, we note that

$$\begin{bmatrix} d_1 & c_1 & 0 \\ 0 & 1 & b_2 \\ 0 & 0 & d_2 \end{bmatrix}^{-1} = \begin{bmatrix} d_1^{-1} & -d_1^{-1}c_1 & d_1^{-1}c_1 b_2 d_2^{-1} \\ 0 & 1 & -b_2 d_2^{-1} \\ 0 & 0 & d_2^{-1} \end{bmatrix}.$$

Thus  $f_1 + f_2$  and  $f_1 f_2$  have linearizations. Consequently the lemma does indeed hold.  $\square$

4.5.2. *Remark.* We refer the reader to the book [7] for a complete discussion of Schützenberger's Theorem and its context in the theory of rational formal noncommutative power series. Nonetheless, we feel that we do owe the reader at least a brief sketch of the interpretation of Lemma 4.5.1 that identifies it as a consequence of Schützenberger's Theorem. For simplicity and with some loss of generality we assume that  $p = 1$ . Without further loss of generality we may assume that

$$d(0) = I_N - \sum_{\theta=1}^q h_\theta \otimes \mathbf{X}_\theta.$$

Let

$$h = (h_1, \dots, h_q) \in \text{Mat}_N^q.$$

Then we have

$$d^{-1} = \sum_{\substack{\text{monomials} \\ \mathbf{M} \in \mathbb{C}\langle \mathbf{X} \rangle}} \mathbf{M}(h) \otimes \mathbf{M}$$

(the sum is actually finite on account of the remark immediately following the statement of Lemma 4.5.1) and hence

$$f = bd^{-1}c = \sum_{\substack{\text{monomials} \\ \mathbf{M} \in \mathbb{C}\langle \mathbf{X} \rangle}} (b\mathbf{M}(h)c)\mathbf{M}.$$

Write

$$f = \sum_{\substack{\text{monomials} \\ \mathbf{M} \in \mathbb{C}\langle \mathbf{X} \rangle}} a_{\mathbf{M}} \mathbf{M} \quad (a_{\mathbf{M}} \in \mathbb{C}).$$

In the spirit of Schützenberger’s theorem we should think of the collection

$$(b, h_1, \dots, h_q, c)$$

as a “linear automaton” which by the rule

$$a_{\mathbf{M}} = b\mathbf{M}(h)c \text{ for monomials } \mathbf{M} \in \mathbb{C}\langle \mathbf{M} \rangle$$

“computes”  $f$  coefficient-by-coefficient. More generally and analogously, Schützenberger’s theorem produces a linear automaton to compute the coefficients of any given rational noncommutative formal power series, and moreover asserts that any noncommutative formal power series so “computable” is rational. We note finally that our application of Lemma 4.5.1 below is not dependent on its “automatic” interpretation. Neither the theory of automata nor the theory of rational formal noncommutative power series are needed in the sequel.

**4.6. Proof of Proposition 4.4.1.** Without loss of generality we may assume that  $(\mathcal{A}, \phi)$  is the noncommutative probability space  $(\mathfrak{M}, \phi)$  and we may take  $\{x_\theta\}_{\theta=1}^q$  to be the family constructed in Proposition 4.2.1. Fix

$$f = f(\mathbf{X}_1, \dots, \mathbf{X}_q) \in \text{Mat}_p(\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle)$$

such that  $f(x_1, \dots, x_q) = X$ , write  $f = bd^{-1}c$  as in Lemma 4.5.1, and then write

$$\begin{aligned} \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} &= L_0 \otimes 1_{\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle} + L_1 \otimes \mathbf{X}_1 + \dots + L_q \otimes \mathbf{X}_q \\ &\in \text{Mat}_{p+N}(\mathbb{C}\langle \mathbf{X}_1, \dots, \mathbf{X}_q \rangle) \quad (L_0, L_1, \dots, L_q \in \text{Mat}_{p+N}(\mathbb{C})) \end{aligned}$$

in the unique possible way. Finally, let

$$L = \mathbf{1} \otimes \left( L_0 + \begin{bmatrix} zI_p & 0 \\ 0 & 0 \end{bmatrix} \right) + x_1 \otimes L_1 + \dots + x_q \otimes L_q \in \mathfrak{M}((1/z))$$

noting that this expression when expanded in terms of elementary matrices takes by (21) the desired form (32). Let

$$B = b \otimes \mathbf{1} \in \text{Mat}_{p \times N}(\mathfrak{M}) \quad \text{and} \quad C = c \otimes \mathbf{1} \in \text{Mat}_{N \times p}(\mathfrak{M}).$$

Let

$$D \in \text{GL}_N(\mathfrak{M}) \quad \text{and} \quad F \in \text{Mat}_p(\mathfrak{M})$$

be the evaluations of  $d$  and  $f$ , respectively, at  $\mathbf{X}_\theta = x_\theta$  for  $\theta = 1, \dots, q$ . Since  $D$  is the image of an invertible matrix under a unital algebra homomorphism,  $D$  is indeed invertible. Furthermore  $zI_p \otimes \mathbf{1} - F$  is invertible by Lemma 4.3.4. It follows by the discussion after formula (17) that the matrix

$$\begin{bmatrix} zI_p \otimes \mathbf{1} & B \\ C & D \end{bmatrix} \in \text{Mat}_{p+N}(\mathfrak{M}((1/z)))$$

is invertible, and from (17) itself it follows that

$$\begin{aligned} &\begin{bmatrix} zI_p \otimes \mathbf{1} & B \\ C & D \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & D^{-1} \end{bmatrix} + \begin{bmatrix} I_p \otimes \mathbf{1} \\ -D^{-1}C \end{bmatrix} (zI_p \otimes \mathbf{1} - F)^{-1} \begin{bmatrix} I_p \otimes \mathbf{1} & -BD^{-1} \end{bmatrix}. \end{aligned}$$

In turn, by (29) we have

$$L = \begin{bmatrix} zI_p \otimes \mathbf{1} & B \\ C & D \end{bmatrix}^{\natural},$$

hence  $L$  is invertible and moreover (33) holds by Lemma 4.3.4. The proof of Proposition 4.4.1 is complete.  $\square$

## 5. SOLVING THE GENERALIZED SCHWINGER-DYSON EQUATION

We finish the proof of Proposition 2.5.2 by constructing sufficiently many solutions of the generalized Schwinger-Dyson equation.

**5.1. Statement of the construction.** Here is our main result in this section.

**Proposition 5.1.1.** *Fix data of the form (7). Consider the matrix*

$$(34) \quad A = -\mathbf{1} \otimes a^{(0)} - \sum_{\theta=1}^q \sum_{k \in \mathbb{N}} \left( \mathbf{e}[\theta \star k, k] + \sum_{j=1}^{\infty} \kappa_{j+1}^{(\theta)} \mathbf{e}[k, \theta^{\circ j} \star k] \right) \otimes a^{(\theta)}.$$

constructed by using these data. Assume that

$$(35) \quad G = A^{-1} \in \mathfrak{M}((1/z)) \text{ exists, and}$$

$$(36) \quad g(i, j) = G(i-1, j-1) \text{ for } i, j = 1, \dots, n.$$

Then (6), (8), and (9) hold, i.e., the data (7) constitute a solution of the generalized Schwinger-Dyson equation.

We complete the proof below in §5.2 after deducing Proposition 2.5.2 from Proposition 5.1.1 and proving a lemma excusing us from having to verify (9) when proving Proposition 5.1.1. The minus signs inserted in definition (34) turn out to save us from writing many minus signs later.

**5.1.2. Remark.** We can rewrite (34) as

$$(37) \quad A = -\mathbf{1} \otimes a^{(0)} - \sum_{\theta=1}^q \left( \hat{\lambda}^{(\theta)} + \sum_{j=1}^{\infty} \kappa_{j+1}^{(\theta)} \lambda^{(\theta)} \right) \otimes a^{(\theta)}$$

in terms of the lowering and raising operators considered in Remark 4.2.5, thus making the relationship of the definition of  $A$  to the setup of [35] more transparent.

**5.1.3. Completion of the proof of Proposition 2.5.2 with Proposition 5.1.1 granted.**

We identify  $X$  in Proposition 2.5.2 with  $X$  in Proposition 4.4.1. We complete the choice of positive integer  $q$  and the given family  $\left\{ \left\{ \kappa_j^{(\theta)} \right\}_{\theta=1}^q \right\}_{j=2}^{\infty}$  to a family

$$(38) \quad \left( q, n, g, a^{(0)}, \left\{ a^{(\theta)} \right\}_{\theta=1}^q, \left\{ \left\{ \kappa_j^{(\theta)} \right\}_{j=2}^{\infty} \right\}_{\theta=1}^q \right)$$

of the form (7) where

$$\begin{aligned} n &= p + N > p, \\ a^{(0)} &= - \left( L_0 + \begin{bmatrix} zI_p & 0 \\ 0 & 0 \end{bmatrix} - \sum_{\theta=1}^q L_{\theta} \right) \in \text{Mat}_n(\mathbb{C}[z]), \\ a^{(\theta)} &= -L^{(\theta)} \in \text{Mat}_n(\mathbb{C}) \text{ for } \theta = 1, \dots, q \text{ and} \\ g(i, j) &= -L^{-1}(i-1, j-1) \text{ for } i, j = 1, \dots, n. \end{aligned}$$

For the family (38) the hypotheses (35) and (36) of Proposition 5.1.1 are fulfilled by (32) and the definition of  $g$ , respectively. Thus (38) is a solution of the generalized Schwinger-Dyson equation, i.e., (6), (8), and (9) hold. Property (10) holds by construction and property (11) holds by (33). The proof of Proposition 2.5.2 is complete modulo the proof of Proposition 5.1.1.  $\square$

**Lemma 5.1.4.** *To prove Proposition 5.1.1 it is necessary only to verify statements (6) and (8) for data (7) satisfying hypotheses (35) and (36).*

In the proof below we are reusing elements of the “secondary trick” used in [1] to obtain certain correction terms.

*Proof.* The weakened version of Proposition 5.1.1 delivering only conclusions (6) and (8) for data (7) satisfying (35) and (36) we will call Proposition 5.1.1- $\epsilon$ . Our task is to derive Proposition 5.1.1 from Proposition 5.1.1- $\epsilon$ . To that end fix  $b \in \text{Mat}_n(\mathbb{C}((1/z)))$  arbitrarily and consider new data consisting of

$$(39) \quad \left\{ \begin{array}{l} \text{a positive integer } \hat{n} = 2n \text{ (but } q \text{ the same as before),} \\ \text{a matrix } \hat{a}^{(0)} = \begin{bmatrix} a^{(0)} & b \\ 0 & a^{(0)} \end{bmatrix} \in \text{Mat}_{\hat{n}}(\mathbb{C}((1/z))), \\ \text{matrices } \hat{a}^{(\theta)} = \begin{bmatrix} a^{(\theta)} & 0 \\ 0 & a^{(\theta)} \end{bmatrix} \in \text{Mat}_{\hat{n}}(\mathbb{C}) \text{ for } \theta = 1, \dots, q, \\ \text{a matrix } \hat{g} = \begin{bmatrix} g & h \\ 0 & g \end{bmatrix} \in \text{Mat}_{\hat{n}}(\mathbb{C}((1/z))) \text{ (} h \text{ to be determined), and} \\ \text{a family } \{\{\kappa_j^{(\theta)}\}_{j=2}^{\infty}\}_{\theta=1}^q \text{ of complex numbers (same as before).} \end{array} \right.$$

We will apply Proposition 5.1.1- $\epsilon$  to the new data (39) thereby deriving (9) for the old data (7). To apply Proposition 5.1.1- $\epsilon$  we need first to verify invertibility of the matrix

$$\hat{A} = -\hat{a}^{(0)} - \sum_{\theta=1}^q \sum_{k \in \mathbb{N}} \left( \mathbf{e}[\theta \star k, k] + \sum_{j=1}^{\infty} \kappa_{j+1}^{(\theta)} \mathbf{e}[k, \theta^{\circ j} \star k] \right) \otimes \hat{a}^{(\theta)}.$$

Because  $\natural$  is an isomorphism, there exists unique  $\tilde{A} \in \text{Mat}_n(\mathfrak{M}((1/z)))$  such that  $(\tilde{A})^{\natural} = A$ . Using (29) we obtain the relation

$$\hat{A} = \begin{bmatrix} \tilde{A} & -b \otimes \mathbf{1} \\ 0 & \tilde{A} \end{bmatrix}^{\natural} \in \mathfrak{M}((1/z)),$$

and we have explicitly

$$\hat{G} = \left( \left[ \begin{array}{cc} \tilde{A} & -b \otimes \mathbf{1} \\ 0 & \tilde{A} \end{array} \right]^{-1} \right)^{\natural} = \left[ \begin{array}{cc} \tilde{A}^{-1} & \tilde{A}^{-1}(b \otimes \mathbf{1})\tilde{A}^{-1} \\ 0 & \tilde{A}^{-1} \end{array} \right]^{\natural}.$$

Thus the new data (39) satisfy (35), and moreover there is a choice of  $h$  we can in principle read off from the last displayed line above so that hypothesis (36) is satisfied. (An explicit formula for  $h$  is not needed.) Statement (8) of Proposition 5.1.1- $\epsilon$  applied to the new data asserts that

$$\lim_{j \rightarrow \infty} \text{val} \left[ \begin{array}{cc} a^{(\theta)}g & a^{(\theta)}h \\ 0 & a^{(\theta)}g \end{array} \right]^j = -\infty \text{ for } \theta = 1, \dots, q.$$

This can also be deduced directly from (8) as it pertains to the old data (7). Finally, the key point is that by statement (6) as it pertains to the new data (39) we have

$$I_{\hat{n}} + \begin{bmatrix} a^{(0)} & b \\ 0 & a^{(0)} \end{bmatrix} \begin{bmatrix} g & h \\ 0 & g \end{bmatrix} + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \kappa_j^{(\theta)} \left( \begin{bmatrix} a^{(\theta)} & 0 \\ 0 & a^{(\theta)} \end{bmatrix} \begin{bmatrix} g & h \\ 0 & g \end{bmatrix} \right)^j = 0.$$

Looking in the upper left corners, we obtain an identity

$$a^{(0)}h + bg + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \sum_{\nu=0}^{j-1} \kappa_j^{(\theta)} (a^{(\theta)}g)^\nu (a^{(0)}h) (a^{(\theta)}g)^{j-1-\nu} = 0.$$

The latter equation, because  $b$  is arbitrary and  $g$  is invertible by (6) as it pertains to the old data (7), proves that (9) holds for the old data. In other words, Proposition 5.1.1- $\epsilon$  does indeed imply Proposition 5.1.1.  $\square$

**5.2. Proof of Proposition 5.1.1.** In broad outline the proof is similar to the proof we previously gave for Lemma 3.2.2. But more machinery is needed.

**5.2.1. Block decompositions.** Throughout the proof it will be convenient to work with the block decompositions defined by the formulas

$$A = \sum_{i,j \in \mathbb{N}} \mathbf{e}[i, j] \otimes A\langle i, j \rangle \quad \text{and} \quad G = \sum_{i,j \in \mathbb{N}} \mathbf{e}[i, j] \otimes G\langle i, j \rangle$$

where

$$A\langle i, j \rangle, G\langle i, j \rangle \in \text{Mat}_n(\mathbb{C}((1/z))) \quad \text{and in particular} \quad G\langle 0, 0 \rangle = g.$$

The key point is contained in the following result which says that the first column of blocks in  $G$  has a relatively simple structure.

**Lemma 5.2.2.** *We have*

$$(40) \quad G\langle \theta_1 \star \dots \star \theta_k, 0 \rangle = ga^{(\theta_1)}ga^{(\theta_2)} \dots ga^{(\theta_k)}g$$

for  $k \in \mathbb{N}$  and  $\theta_1, \dots, \theta_k \in \{1, \dots, q\}$ .

This relation generalizes formula (20) above.

*Proof.* For the proof we will use Lemma 3.3.2 in the case

$$A = \mathfrak{M}((1/z)), \quad A \in \mathcal{A} : \text{ as on line (34),}$$

$$\pi = \mathbf{e}[0, 0] \otimes I_n \in \mathcal{A} \quad \text{and} \quad \sigma = \sum_{i=1}^{\infty} \mathbf{e}[i, i] \otimes I_n \in \mathcal{A}.$$

We will also use the matrices

$$R^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[k \star \theta, k] \otimes I_n \in \mathfrak{M} \quad \text{and} \quad \hat{R}^{(\theta)} = \sum_{k \in \mathbb{N}} \mathbf{e}[k, k \star \theta] \otimes I_n \in \mathfrak{M}$$

for  $\theta = 1, \dots, q$ .

These matrices satisfy

$$(41) \quad \hat{R}^{(\theta)} R^{(\theta')} = \delta_{\theta\theta'} \mathbf{1} \quad \text{for } \theta, \theta' = 1, \dots, q,$$

$$(42) \quad \hat{R}^{(\theta)} A R^{(\theta')} = \delta_{\theta\theta'} A \quad \text{for } \theta, \theta' = 1, \dots, q, \text{ and}$$

$$(43) \quad \sum_{\theta=1}^q R^{(\theta)} \hat{R}^{(\theta)} = \sigma,$$

as can either be verified directly by straightforward if tedious calculation or by using the definitions and relations (24), (25), (26), (27), and (28) listed above in Remark 4.2.5, along with the rewrite (37) of the definition of  $A$ . It follows by (42) and (43) that

$$\sigma A \sigma = \left( \sum_{\theta=1}^q R^{(\theta)} \hat{R}^{(\theta)} \right) A \left( \sum_{\theta=1}^q R^{(\theta')} \hat{R}^{(\theta')} \right) = \sum_{\theta=1}^q R^{(\theta)} A \hat{R}^{(\theta)}.$$

It follows in turn that  $A_\sigma^{-1}$  exists and more precisely that

$$(44) \quad A_\sigma^{-1} = \sum_{\theta=1}^q R^{(\theta)} G \hat{R}^{(\theta)},$$

as one verifies using (41). We furthermore have

$$\sigma A \pi A^{-1} \pi = - \sum_{\theta=1}^q (\mathbf{e}[\theta, 0] \otimes a^{(\theta)}) (\mathbf{e}[0, 0] \otimes g) = - \sum_{\theta=1}^q R^{(\theta)} (\mathbf{e}[0, 0] \otimes a^{(\theta)} g),$$

as one can immediately check. Finally, we have the following chain of equalities:

$$\begin{aligned} \sum_{\theta=1}^q \sum_{i \in \mathbb{N}} \mathbf{e}[i \star \theta, 0] \otimes G \langle i \star \theta, 0 \rangle &= \sigma A^{-1} \pi = -A_\sigma^{-1} A \pi A^{-1} \pi \\ &= \sum_{\theta=1}^q \sum_{\theta'=1}^q R^{(\theta)} G \hat{R}^{(\theta)} R^{(\theta')} (\mathbf{e}[0, 0] \otimes a^{(\theta')} g) \\ &= \sum_{\theta=1}^q \sum_{i \in \mathbb{N}} \mathbf{e}[i \star \theta, 0] \otimes G \langle i, 0 \rangle a^{(\theta)} g. \end{aligned}$$

At the first step we used Lemma 4.1.5, at the second equation (18) of Lemma 3.3.2, at the third (44), and at the last (41). Thus (40) holds.  $\square$

5.2.3. *Proof of (8).* By definition of  $\mathfrak{M}((1/z))$  we have

$$\lim_{i \rightarrow \infty} \text{val } G \langle i, 0 \rangle = -\infty.$$

Thus by Lemma 5.2.2 we have for  $\theta = 1, \dots, q$  that

$$\lim_{i \rightarrow \infty} \text{val } (a^{(\theta)} g)^i = \lim_{i \rightarrow \infty} \text{val } a^{(\theta)} G \langle \theta^{\circ i}, 0 \rangle = 0,$$

which proves statement (8).

5.2.4. *Proof of (6).* Consider the following calculation:

$$\begin{aligned} I_n &= \sum_{j \in \mathbb{N}} A \langle 0, j \rangle G \langle j, 0 \rangle = -a^{(0)} G \langle 0, 0 \rangle - \sum_{\theta=1}^q \sum_{j=1}^{\infty} \kappa_{j+1}^{(\theta)} a^{(\theta)} G \langle \theta^{\circ j}, 0 \rangle \\ &= -a^{(0)} g - \sum_{\theta=1}^q \sum_{j=1}^{\infty} \kappa_{j+1}^{(\theta)} (a^{(\theta)} g)^j. \end{aligned}$$

The first step holds by definition of  $G$ , the second by definition of  $A$ , and the last by Lemma 5.2.2. This calculation proves statement (6).

5.2.5. *Completion of the proof.* By Lemma 5.1.4 it is necessarily the case that statement (9) holds. The proof of Proposition 5.1.1 is complete and in turn the proof of Proposition 2.5.2 is complete.  $\square$

### 5.3. Miscellaneous remarks.

5.3.1. In Lemma 5.2.2 and its proof we are exploiting the type of recursion used, for example, in [26], and used more generally in many investigations of random walk on infinite trees.

5.3.2. By exploiting the identities (41), (42), and (43), we are reusing some features of the proof of [1, Prop. 9].

5.3.3. In the recently introduced bi-free probability setup [36] the left and right variants of lowering and raising operators play an equal role. Our approach here is potentially useful for calculating the laws of matrix-polynomial combinations of bi-free collections of noncommutative random variables.

5.3.4. By continuing the line of argument in the proof of Lemma 5.2.2 and using (19) of Lemma 3.3.2, it is not difficult to prove that

$$(45) \quad GR^{(\theta)} = (R^{(\theta)} - G\pi AR^{(\theta)})G \text{ for } \theta = 1, \dots, q.$$

Since statement (45) is not needed for the proof of Theorem 1, we omit its proof. It is easy to see that Lemma 5.2.2 and (45) together allow one to make every block  $G\langle i, j \rangle$  explicit in terms of  $g$  and  $A$ . Doing so in a systematic way one would obtain a generalization of Theorem 1 having the full statement of Theorem 2 as a consequence.

## 6. NOTES ON NEWTON-PUISEUX SERIES

At this point in the paper we switch from the noncommutative viewpoint previously stressed to the viewpoint of commutative algebra and algebraic geometry. The latter is maintained throughout the rest of the paper.

**6.1. Newton-Puiseux series and Newton polygons.** We review basic devices for understanding singularities of plane algebraic curves in characteristic zero and make some definitions needed for later calculations.

6.1.1. *The algebraic closure of  $\mathbb{C}((1/z))$ .* Let  $\mathbb{K}$  denote the union of the tower of fields  $\{\mathbb{C}((1/z^{1/n!}))\}_{n=1}^{\infty}$ . In other words,  $\mathbb{K}$  arises by adjoining roots of  $z$  of all orders to  $\mathbb{C}((1/z))$ . We call an element of  $\mathbb{K}$  a *Newton-Puiseux series*. When discussing  $\mathbb{K}$  below we often use the more apposite abbreviated notation  $\mathbb{K}_0 = \mathbb{C}((1/z))$ . It is well-known that  $\mathbb{K}$  is the algebraic closure of  $\mathbb{K}_0$ . See, e.g., [14, Cor. 13.15]. The original insight is due to Newton.

6.1.2. *Extension of the valuation function  $\text{val}$  to  $\mathbb{K}$ .* Each element  $f \in \mathbb{K}$  has by definition a unique *Newton-Puiseux expansion*  $f = \sum_{u \in \mathbb{Q}} c_u z^u$  with coefficients  $c_u \in \mathbb{C}$  such that for some positive integer  $N$  depending on  $f$  one has  $c_u = 0$  unless  $u \leq N$  and  $Nu \in \mathbb{Z}$ . To extend to  $\mathbb{K}$  the valuation defined on  $\mathbb{K}_0$ , we define

$$\text{val } f = \sup\{u \in \mathbb{Q} \mid c_u \neq 0\} \in \mathbb{Q} \cup \{-\infty\}.$$

The properties (2), (3), and (4) continue to hold for the extension of  $\text{val}$  to  $\mathbb{K}$ .

**Proposition 6.1.3.** *Let  $P(y) \in \mathbb{K}_0[y]$  be a polynomial of degree  $n$  in a variable  $y$  with coefficients in the field  $\mathbb{K}_0$ . Write*

$$P(y) = \sum_{i=0}^n a_i y^{n-i} = a_0 \prod_{i=1}^n (y - r_i) \quad (a_i \in \mathbb{K}_0 \text{ and } r_i \in \mathbb{K}),$$

enumerating the roots  $r_i$  so that  $\text{val } r_1 \geq \dots \geq \text{val } r_n$ . (i) Then we have

$$(46) \quad \text{val } a_i \leq \text{val } a_0 + \text{val } r_1 \cdots r_i \text{ for } i = 0, \dots, n,$$

$$(47) \quad \text{with equality for } i = 0, n \text{ and furthermore} \\ \text{for } i = 1, \dots, n-1 \text{ s.t. } \text{val } r_i > \text{val } r_{i+1}.$$

(ii) Let  $\psi : [0, n] \rightarrow \mathbb{R} \cup \{-\infty\}$  be the infimum of all affine linear functions  $\lambda : [0, n] \rightarrow \mathbb{R}$  satisfying  $\lambda(i) \geq \text{val } a_i$  for  $i = 0, \dots, n$ . Then we have the integral formula  $\psi(s) = \text{val } a_0 + \int_0^s \text{val } r_{\lceil u \rceil} du$ .

Here  $\lceil u \rceil$  denotes the least integer not less than  $u$ . The function  $\psi$  is concave by construction. The function  $\psi$  (or rather, its graph) is the *Newton polygon* associated with  $P(y)$ , up to reflections in and translations parallel to the horizontal and vertical axes (conventions vary from author to author). Similarly, a Newton polygon is attached to any one-variable polynomial with coefficients in a discretely valued field. For background on Newton polygons see [6, Chap. 2, Sec. 5] or [9, Part III, Chap. 8, Sec. 3].

*Proof.* (i) Since  $(-1)^i a_i / a_0$  for  $i > 0$  is the  $i^{\text{th}}$  symmetric function of the roots  $r_1, \dots, r_n$ , the result follows straightforwardly from (2), (3) and (4). (ii) The deduction of this statement from the preceding one is standard.  $\square$

6.1.4. *Extension of the valuation to  $\text{Mat}_n(\mathbb{K})$ .* We extend  $\text{val}$  from  $\text{Mat}_n(\mathbb{K}_0)$  to  $\text{Mat}_n(\mathbb{K})$  by the rule  $\text{val } A = \max_{i,j=1}^n \text{val } A(i, j)$ . Then  $\text{Mat}_n(\mathbb{K})$  satisfies all the axioms of a Banach algebra over  $\mathbb{K}$  except completeness. Lack of completeness will not be an issue.

6.2. **Applications.** We present several applications of the preceding machinery needed for the proof of Proposition 2.5.3.

6.2.1. *Specialized matrix notation.* Given  $A \in \text{Mat}_n(\mathbb{C}((1/z)))$ , we define

$$e(A) = (e_1(A), \dots, e_n(A)) \in \mathbb{C}((1/z))^n$$

by the formula

$$\det(tI_n - A) = t^n + \sum_{i=1}^n e_i(A) (-1)^i t^{n-i} \in \mathbb{C}((1/z))[t].$$

The Cayley-Hamilton Theorem takes then the form

$$(48) \quad A^n + \sum_{i=1}^n (-1)^i e_i(A) A^{n-i} = 0.$$

**Proposition 6.2.2.** *For  $A \in \text{Mat}_n(\mathbb{C}((1/z)))$  the following are equivalent:*

- (I)  $e(A) \in (1/z)\mathbb{C}[[1/z]]^n$ .
- (II)  $\lim_{k \rightarrow \infty} \text{val } A^k = -\infty$ .

*Proof.* We add one further statement to the list above:

- (III) Every eigenvalue of  $A$  in  $\mathbb{K}$  has (strictly) negative valuation.

Statements (I) and (III) are equivalent by Proposition 6.1.3. It remains only to prove the equivalence (II) $\Leftrightarrow$ (III). It is actually easier to prove more. We will prove the equivalence (II) $\Leftrightarrow$ (III) for  $A \in \text{Mat}_n(\mathbb{K})$ . Supposing at first that  $A$  consists of a single Jordan block, one verifies the equivalence by inspection. In general we can write  $A = WJW^{-1}$  where  $W \in \text{GL}_n(\mathbb{K})$  and  $J \in \text{Mat}_n(\mathbb{K})$  is block-diagonal with diagonal blocks of the Jordan form and we have a bound

$$|\text{val } A^k - \text{val } J^k| \leq \text{val } W^{-1} + \text{val } W$$

which establishes the equivalence (II) $\Leftrightarrow$ (III) in general.  $\square$

**6.2.3. Negative spectral valuation.** We say that  $A \in \text{Mat}_n(\mathbb{K}_0)$  has *negative spectral valuation* if the equivalent conditions (I) and (II) above hold.

**6.2.4. Algebraic and nonsingular algebraic elements of  $\mathbb{C}[[t]]$ .** In some situations  $1/z$  rather than  $z$  is the natural parameter to work with. We therefore make the following definitions which are in principle redundant but in practice convenient. We say that  $f(t) \in \mathbb{C}[[t]]$  is *algebraic* if  $F(t, f(t)) = 0$  for some  $0 \neq F(x, y) \in \mathbb{C}[x, y]$ . Of course  $f(t) \in \mathbb{C}[[t]]$  is algebraic if and only if  $f(1/z) \in \mathbb{C}((1/z))$  is algebraic in the sense defined in §2.2.2. We call  $f(t)$  *nonsingular algebraic* if there exists  $F(x, y) \in \mathbb{C}[x, y]$  satisfying  $\frac{\partial F}{\partial y}(0, f(0)) \neq 0$  and  $F(t, f(t)) = 0$ .

The next statement is the key to desingularization.

**Lemma 6.2.5.** *Let  $f(t) \in t\mathbb{C}[[t]]$  be a power series with vanishing constant term. Let  $F(x, y) \in \mathbb{C}[x, y]$  be a polynomial not divisible by  $x$  such that  $F(t, f(t)) = 0$ . Then the following conditions are equivalent:*

- (I)  $F(1/z, y) \in \mathbb{K}_0[y]$  has exactly one root in  $\mathbb{K}$  of negative valuation.
- (II)  $\frac{\partial F}{\partial y}(0, 0) \neq 0$ .

*Proof.* Write  $F(x, y) = \sum_{i=0}^n p_i(x)y^{n-i}$  where  $n > 0$ ,  $p_i(x) \in \mathbb{C}[x]$  and  $p_0(x) \neq 0$ . For  $p(x) \in \mathbb{C}[x]$ , let  $\text{ord } p(x)$  denote the exponent of the highest power of  $x$  dividing  $p(x)$ . Let  $\phi : [0, n] \rightarrow \mathbb{R} \cup \{+\infty\}$  denote the supremum of all affine linear functions  $\lambda : [0, n] \rightarrow \mathbb{R}$  such that  $\lambda(i) \leq \text{ord } p_i(x)$  for  $i = 0, \dots, n$ . The function  $\phi$  is convex. Since  $F(0, 0) = 0$  we have  $\phi(n) > 0$ . Since  $x$  does not divide  $F(x, y)$ , we have  $\min_{i=0}^n \text{ord } p_i(x) = \min_{i=0}^n \phi(i) = 0$ . Let  $i_0$  be the maximum of  $i = 0, \dots, n-1$  such that  $\phi(i_0) = 0$ . Statement (ii) is equivalent to the assertion that  $i_0 = n-1$ . Now  $F(1/z, y)$  has  $n - i_0$  roots of negative valuation by Proposition 6.1.3 and the observation that  $\text{ord } p(x) = -\text{val } p(1/z)$ . Thus statement (I) is equivalent to the assertion that  $i_0 = n-1$ .  $\square$

The next statement summarizes just enough of the theory of resolution of singularities of plane algebraic curves in characteristic zero for our purposes.

**Proposition 6.2.6.** *Let  $\sum_{i=0}^{\infty} c_i t^i \in t\mathbb{C}[[t]]$  ( $c_i \in \mathbb{C}$ ) be an algebraic power series. Then  $\sum_{i=N}^{\infty} c_{i+N} t^i \in t^N \mathbb{C}[[t]]$  is a nonsingular algebraic power series for all  $N \gg 0$ .*

*Proof.* Let  $f = \sum_{i=1}^{\infty} c_i z^{-i} \in \mathbb{C}((1/z))$  and  $f_N = \sum_{i=N}^{\infty} c_{i+N} z^{-i} \in \mathbb{C}((1/z))$ . Let  $F(x, y) \in \mathbb{C}[x, y]$  (resp.,  $F_N(x, y) \in \mathbb{C}[x, y]$ ) denote the irreducible equation (see §2.2.2) of  $f$  (resp.,  $f_N$ ). It will be enough to show that  $F_N(1/z, y) \in \mathbb{K}_0[y]$  has exactly one root in  $\mathbb{K}$  of negative valuation for  $N \gg 0$ . If  $f_{N_0} = 0$  for some  $N_0$  then  $f_N = 0$  for all  $N \geq N_0$  and there is nothing to prove. Thus we may assume without loss of generality that  $f_N \neq 0$  for all  $N \geq 0$ . Let  $n$  denote the dimension of  $\mathbb{C}(z, f)$  over  $\mathbb{C}(z)$ . It is clear that  $f$  and  $f_N$  generate the same extension of  $\mathbb{C}(z)$ . Thus  $F(x, y)$  and  $F_N(x, y)$  have the same degree in  $y$ , namely

$n$ . Let  $r_1, \dots, r_n$  denote the roots in  $\mathbb{K}$  of the polynomial  $F(1/z, y) \in \mathbb{K}_0[y]$ , enumerated so that  $r_n = f$ . Let  $h_N = \sum_{i=0}^{2N-1} a_i z^{-i} = f - f_N/z^N \in \mathbb{C}(z)$ , in which case necessarily  $f_N = z^N(f - h_N)$ . Then for a suitable enumeration  $r_{1,N}, \dots, r_{n,N}$  of the roots in  $\mathbb{K}$  of the polynomial  $F_N(1/z, y) \in \mathbb{K}_0[y]$ , we have  $r_{i,N} = z^N(r_i - h_N)$  for  $i = 1, \dots, n$  and  $r_{n,N} = f_N$ . Now the roots  $r_1, \dots, r_n$  are distinct due to irreducibility of  $F(x, y)$ , and clearly  $\text{val}(f - h_N) \leq -2N$ . Because  $h_N \rightarrow_{N \rightarrow \infty} f$  with respect to the valuation  $\text{val}$ , it follows that for some integer  $N_0 > 0$  depending only on  $f$ , and all integers  $N \geq N_0$ , we have

$$\min_{i=1}^{n-1} \text{val } r_{i,N} = N + \min_{i=1}^{n-1} \text{val}(r_i - f) \geq 0 > -N \geq \text{val } f_N = \text{val } r_{n,N},$$

whence the result via Lemma 6.2.5.  $\square$

## 7. EVALUATION OF ALGEBRAIC POWER SERIES ON MATRICES

The main result of this section is Proposition 7.3.4 below. The main tools used in this section are the Cayley-Hamilton Theorem, the Weierstrass Preparation Theorem, and Proposition 6.2.6 above.

**7.1. Motivation.** Let  $A$  be an  $n$ -by- $n$  matrix with complex entries. We take our inspiration from the undergraduate level approach in [39] to computing the matrix exponential  $\exp(tA)$ . The approach is lengthy as presented for sophomores but it can be summarized quickly at graduate level as follows. Perform Weierstrass division (possible globally in this case) in order to obtain an identity relating two-variable entire functions of complex variables  $t$  and  $X$ , namely

$$(49) \quad \exp(tX) = \sum_{k=0}^{n-1} y_k(t) X^k + Q(X, t) \det(XI_n - A),$$

for suitable and unique remainder  $\sum_{k=0}^{n-1} y_k(t) X^k$  and quotient  $Q(X, t)$ . Differentiation of (49) on both sides with respect to  $t$  yields a first order homogeneous linear differential equation for the vector function  $y(t) = (y_1(t), \dots, y_n(t))$  which together with the evident initial value data uniquely determines  $y(t)$ . One can then go on to solve explicitly for the functions  $y_k(t)$  in closed form. By plugging in  $X = A$  on both sides of (49) and using the Cayley-Hamilton Theorem one then has finally

$$(50) \quad \exp(tA) = \sum_{k=0}^{n-1} y_k(t) A^k.$$

Formula (50) makes no reference to the Jordan canonical form of  $A$ . Indeed, by construction, the coefficients  $y_k(t)$  are uniquely determined by the characteristic polynomial of  $A$  alone. We will make roughly analogous use of Weierstrass division below to evaluate algebraic power series on matrices with entries in  $\mathbb{C}((1/z))$  of negative spectral valuation.

**7.2.  $I$ -adic convergence, power series, and Weierstrass division.** We pause to review generalities connected with the formal power series version of the Weierstrass Preparation Theorem.

7.2.1. *I-adic convergence.* Given a commutative ring  $R$  with unit, an ideal  $I$ , and a sequence  $\{a\} \cup \{a_i\}_{i=1}^{\infty}$  in  $R$ , one says  $\lim_{i \rightarrow \infty} a_i = a$  holds *I-adically* if for every positive integer  $k$  there exists a positive integer  $i_0 = i_0(k)$  such that  $a - a_i \in I^k$  for all  $i \geq i_0$ . Similarly, one can speak of *I-adic Cauchy sequences* and *I-adic completeness*. Consider, e.g., the ring  $\mathbb{C}[[u_1, \dots, u_n]] = \mathbb{C}[[u]]$  and the maximal ideal  $I = (u_1, \dots, u_n) \subset \mathbb{C}[[u]]$ . Then  $f_i \in \mathbb{C}[[u]]$  converges *I-adically* to  $f \in \mathbb{C}[[u]]$  if and only if for every  $n$ -tuple  $(\nu_1, \dots, \nu_n)$  of nonnegative integers and every sufficiently large index  $i$  depending on  $(\nu_1, \dots, \nu_n)$ , the Taylor coefficient  $\frac{1}{\nu_1! \cdots \nu_n!} \frac{\partial^{\nu_1 + \cdots + \nu_n} f_i}{\partial u_1^{\nu_1} \cdots \partial u_n^{\nu_n}}(0)$  equals the Taylor coefficient  $\frac{1}{\nu_1! \cdots \nu_n!} \frac{\partial^{\nu_1 + \cdots + \nu_n} f}{\partial u_1^{\nu_1} \cdots \partial u_n^{\nu_n}}(0)$ . It is easy to see that the ring  $\mathbb{C}[[u]]$  is *I-adically complete*.

7.2.2. *Weierstrass division.* We now briefly recall the *Weierstrass Preparation Theorem* from a more active point of view emphasizing the algorithm of Weierstrass division. See, e.g., [42, Thm. 5, p. 139, Chap. VII, §1] for background and proof. The theorem concerns an  $(n+1)$ -variable power series ring over a field, with one of the variables singled out for special treatment. For definiteness we take the coefficient field to be  $\mathbb{C}$ . Consider the ring  $\mathbb{C}[[u_1, \dots, u_n, t]] = \mathbb{C}[[u, t]]$ , with  $t$  distinguished. One is given a *divisand*  $F(u, t) \in \mathbb{C}[[u, t]]$  and a *divisor*  $D(u, t) \in \mathbb{C}[[u, t]]$ . Of the latter it is assumed that there exists a positive integer  $m$  (called the *multiplicity* of the divisor) such that  $D(0, t) = t^m U(t)$  for some  $U(t) \in \mathbb{C}[[t]]$  such that  $U(0) \neq 0$ . The Weierstrass division process delivers a *quotient*  $Q(u, t) \in \mathbb{C}[[u, t]]$  and a *remainder*  $R(u, t) \in \mathbb{C}[[u]][t]$ . The pair  $(Q(u, t), R(u, t))$  is uniquely determined by two requirements. Firstly, the division equation  $F(u, t) = Q(u, t)D(u, t) + R(u, t)$  must hold. Secondly,  $R(u, t)$  must be a polynomial in  $t$  of degree  $< m$ . It bears emphasis that if  $D(u, t) \in \mathbb{C}[[u]][t]$  is monic of degree  $m$  such that  $D(0, t) = t^m$ , and  $F(u, t) \in \mathbb{C}[[u]][t]$ , then the Euclidean (i.e., high school) and Weierstrass division processes deliver the same quotient and remainder.

**Lemma 7.2.3.** *We continue in the setting of the preceding paragraph. However, for simplicity we assume now that  $D(u, t)$  is monic of degree  $m$  such that  $D(0, t) = t^m$ . Consider the ideal  $I = (u_1, \dots, u_n) \subset \mathbb{C}[[u]][t]$ . Let  $k$  be a positive integer. If  $t^k$  divides  $F(u, t)$ , then  $R(u, t)$  belongs to the ideal  $I^{\lfloor k/m \rfloor}$ . (Here  $\lfloor c \rfloor$  denotes the greatest integer not exceeding  $c$ .)*

It follows that formation of Weierstrass remainder upon division by  $D(u, t)$  viewed as a function from  $\mathbb{C}[[u, t]]$  to  $\mathbb{C}[[u]][t]$  is continuous with respect to the  $(t)$ -adic topology on the source and the  $(u_1, \dots, u_n)$ -adic topology on the target.

*Proof.* Let  $F_0(u, t) = F(u, t)/t^k$ . Let  $R_0(u, t)$  denote the remainder of  $F_0(u, t)$  upon Weierstrass division by  $D(u, t)$ . Then  $R(u, t)$  is the remainder of  $t^k R_0(u, t)$  upon high school division by  $D(u, t)$ . This noted, there is no loss of generality in assuming that  $F(u, t) = t^k$ . Write  $D(u, t) = t^m + \sum_{i=0}^{m-1} a_i t^i$  with coefficients  $a_i = a_i(u) \in \mathbb{C}[[u]]$  such that  $a_i(0) = 0$ . Write  $R(u, t) = \sum_{i=0}^{m-1} b_i t^i$  with coefficients  $b_i = b_i(u) \in \mathbb{C}[[u]]$ . Then we have

$$\begin{bmatrix} & & & -a_0 \\ & & & \vdots \\ & & & \vdots \\ 1 & & & \vdots \\ & \ddots & & \vdots \\ & & 1 & -a_{m-1} \end{bmatrix}^k \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} b_0 \\ \vdots \\ b_{m-1} \end{bmatrix},$$

where the matrix on the left is the so-called *companion matrix* for  $D(u, t)$ . Clearly every entry of the  $m^{\text{th}}$  power of the companion matrix belongs to the ideal  $I$ , and hence every entry of the  $k^{\text{th}}$  power belongs to the ideal  $I^{\lfloor k/m \rfloor}$ .  $\square$

**7.3. Formulation of the main result.** We state a technical result needed to make the final arguments of the proof of Proposition 2.5.3.

**7.3.1. Variables and rings.** Throughout the remainder of §7 we fix a positive integer  $n$  and we work with the family of independent (commuting) algebraic variables

$$\{u_i\}_{i=1}^n \cup \{v_i\}_{i=1}^{2n} \cup \{t, x, y\}.$$

Let

$$u = (u_1, \dots, u_n) \text{ and } v = (v_1, \dots, v_{2n}).$$

We write

$$\begin{aligned} \mathbb{C}[u] &= \mathbb{C}[u_1, \dots, u_n], \quad \mathbb{C}[[u]] = \mathbb{C}[[u_1, \dots, u_n]], \\ \mathbb{C}[u, v] &= \mathbb{C}[u_1, \dots, u_n, v_1, \dots, v_{2n}], \end{aligned}$$

and so on. We use similar notation below for building up rings from the given variables without further comment. Given, for example  $P(u, v) \in \mathbb{C}[u, v]^{2n}$ , we denote by  $\frac{\partial P}{\partial u}(u, v)$  the  $2n$ -by- $n$  matrix with entries  $\frac{\partial P_i}{\partial u_j}(u, v)$ . We use similar notation for derivatives of vector functions below without further comment.

**7.3.2. Specialized matrix notation.** Given  $A \in \text{Mat}_n(\mathbb{C}((1/z)))$ , let

$$\begin{aligned} A^{\flat} &= \begin{bmatrix} A(1,1) & \dots & A(n,1) & \dots & A(1,n) & \dots & A(n,n) \end{bmatrix}^{\text{T}} \\ &\in \text{Mat}_{n^2 \times 1}(\mathbb{C}((1/z))). \end{aligned}$$

Note that

$$(51) \quad (BA)^{\flat} = (I_n \otimes B)A^{\flat}$$

for  $B \in \text{Mat}_n(\mathbb{C}((1/z)))$ .

**7.3.3. Setup for the main result.** We are given an algebraic power series

$$(52) \quad f(t) = \sum_{i=0}^{\infty} c_i t^i \in \mathbb{C}[[t]] \quad (c_i \in \mathbb{C})$$

and a matrix

$$A \in \text{Mat}_n(\mathbb{C}((1/z)))$$

of negative spectral valuation, i.e., a matrix satisfying the conditions

$$\lim_{k \rightarrow \infty} \text{val} A^k = -\infty \text{ and } e(A) \in (1/z)\mathbb{C}[[1/z]]^{2n}$$

which by Proposition 6.2.2 are equivalent.

**Proposition 7.3.4.** *Notation and assumptions are as above. For every  $N \geq 0$  such that  $\sum_{i=N}^{\infty} c_i + N t^i \in \mathbb{C}[[t]]$  is nonsingular algebraic and divisible by  $t^{2n}$ , there exist*

$$\gamma \in (1/z)\mathbb{C}[[1/z]]^{2n} \text{ and } P(u, v) \in \mathbb{C}[u, v]^{2n}$$

such that the following statements hold:

$$(53) \quad P(e(A), \gamma) = 0.$$

$$(54) \quad \det \frac{\partial P}{\partial v}(e(A), \gamma) \neq 0.$$

$$(55) \quad [ (A^0)^{\flat} \quad \dots \quad (A^{2n-1})^{\flat} ] \left( \frac{\partial P}{\partial v}(e(A), \gamma) \right)^{-1} \frac{\partial P}{\partial u}(e(A), \gamma) = 0.$$

$$(56) \quad \sum_{i=N}^{\infty} c_{i+N} \begin{bmatrix} A \otimes I_n & I_n \otimes I_n \\ 0 & I_n \otimes A \end{bmatrix}^i = \sum_{i=1}^{2n} \gamma_i \begin{bmatrix} A \otimes I_n & I_n \otimes I_n \\ 0 & I_n \otimes A \end{bmatrix}^{i-1}.$$

The proof takes up the rest of §7 and is completed in §7.6. Note that every  $N \geq 2n$  sufficiently large depending on  $f(t)$  satisfies the hypotheses of Proposition 7.3.4 by Proposition 6.2.6.

7.3.5. *Remark.* In the application to the proof of Proposition 2.5.3 we will need to use several consequences of the conclusions of Proposition 7.3.4 which are easy to check once written down but might otherwise be obscure. For the reader's convenience we write these consequences down. Firstly, we observe that the statement

$$(57) \quad [ (A^N)^{\flat} \quad \dots \quad (A^{N+2n-1})^{\flat} ] \left( \frac{\partial P}{\partial v}(e(A), \gamma) \right)^{-1} \frac{\partial P}{\partial u}(e(A), \gamma) = 0$$

follows from statement (55) via statement (51). Secondly, we observe that (56) implies

$$(58) \quad \sum_{i=N}^{\infty} c_{i+N} \begin{bmatrix} A & B \\ 0 & A \end{bmatrix}^i = \sum_{i=1}^{2n} \gamma_i \begin{bmatrix} A & B \\ 0 & A \end{bmatrix}^{i-1} \quad \text{for } B \in \text{Mat}_n(\mathbb{C}((1/z)))$$

by (so to speak) substituting  $B$  for  $\otimes$ . Finally, we observe that

$$(59) \quad \sum_{i=2N}^{\infty} c_i A^i = \sum_{i=1}^{2n} \gamma_i A^{N+i-1} \quad \text{and}$$

$$(60) \quad \sum_{i=2N}^{\infty} \sum_{\nu=0}^{i-1} c_i A^{\nu} B A^{i-1-\nu} = \sum_{i=1}^{2n} \sum_{\nu=0}^{N+i-2} \gamma_i A^{\nu} B A^{i-2-\nu} \quad \text{for } B \in \text{Mat}_n(\mathbb{C}((1/z))).$$

These last two statements are obtained by right-multiplying statement (58) on both sides by  $\begin{bmatrix} A & B \\ 0 & A \end{bmatrix}^N$  and expanding the matrix powers.

7.3.6. *Reduction of the proof of Proposition 7.3.4.* After replacing  $f(t)$  by  $\sum_{i=N}^{\infty} c_{i+N} t^i$ , we may assume without loss of generality that  $N = 0$ . Thus we are making two further special assumptions concerning  $f(t)$  which for the sake of clarity and convenient reference we write out explicitly. Firstly, we are assuming that

$$(61) \quad f(t) \in t^{2n} \mathbb{C}[[t]].$$

Secondly, we are assuming that there exists

$$F(x, y) \in \mathbb{C}[x, y]$$

such that

$$(62) \quad F(t, f(t)) = 0 \quad \text{and} \quad \frac{\partial F}{\partial y}(0, 0) \neq 0.$$

Note that the formula

$$(63) \quad F(0, 0) = \frac{\partial F}{\partial x}(0, 0) = \dots = \frac{\partial^{2n-1} F}{\partial x^{2n-1}}(0, 0) = 0$$

follows straightforwardly from (61) and (62).

#### 7.4. A candidate for $\gamma$ .

7.4.1. *A special polynomial.* Let

$$(64) \quad D(u, t) = t^n + \sum_{i=1}^n (-1)^i u_i t^{n-i} \in \mathbb{C}[u, t].$$

Note that the left side of (48) equals  $D(e(A), A)$ . This is the motivation for the definition of  $D(u, t)$ .

7.4.2. *Construction of  $\varphi(u)$ .* Perform Weierstrass division by  $D(u, t)^2$  to obtain an identity

$$(65) \quad f(t) = \sum_{i=1}^{2n} \varphi_i(u) t^{i-1} + Q_1(u, t) D(u, t)^2,$$

where

$$\varphi(u) = (\varphi_1(u), \dots, \varphi_{2n}(u)) \in u\mathbb{C}[[u]]^{2n} \quad \text{and} \quad Q_1(u, t) \in \mathbb{C}[[u, t]].$$

Note that we indeed have  $\varphi(0) = 0$  as one verifies by substituting  $u = 0$  on both sides of (65) and using hypothesis (61).

**Lemma 7.4.3.** *We have*

$$(66) \quad \sum_{j=0}^{\infty} c_j \begin{bmatrix} A \otimes I_n & I_n \otimes I_n \\ 0 & I_n \otimes A \end{bmatrix}^j = \sum_{j=1}^{2n} \varphi_j(e(A)) \begin{bmatrix} A \otimes I_n & I_n \otimes I_n \\ 0 & I_n \otimes A \end{bmatrix}^{j-1} \quad \text{and}$$

$$(67) \quad \left[ (A^0)^\flat \quad \dots \quad (A^{2n-1})^\flat \right] \frac{\partial \varphi}{\partial u}(e(A)) = 0.$$

Thus the reasonable candidate for  $\gamma$  is  $\varphi(e(A))$ .

*Proof.* Note that (67) can be rewritten

$$(68) \quad \sum_{i=1}^{2n} \frac{\partial \varphi_i}{\partial u_j}(e(A)) A^{i-1} = 0 \quad \text{for } j = 1, \dots, n.$$

Note also that by differentiation we deduce from (65) that

$$(69) \quad \sum_{i=1}^{2n} \frac{\partial \varphi_i}{\partial u_j}(u) t^{i-1} \\ = - \left( \frac{\partial Q_1}{\partial u_j}(u, t) D(u, t) + 2Q_1(u, t) \frac{\partial D}{\partial u_j}(u, t) \right) D(u, t) \quad \text{for } j = 1, \dots, n.$$

Now suppose temporarily that  $f(t) \in \mathbb{C}[t]$ . Then we have  $\varphi(u) \in u\mathbb{C}[u]^{2n}$  and  $Q_1(u, t) \in \mathbb{C}[u, t]$  since high school division in this case gives the same result as Weierstrass division. Substituting

$$(u, t) = \left( e(A), \begin{bmatrix} A \otimes I_n & I_n \otimes I_n \\ 0 & I_n \otimes A \end{bmatrix} \right)$$

into (65) and using the Cayley-Hamilton Theorem (48), we obtain (66). Substituting  $(u, t) = (e(A), A)$  into (69) and using the Cayley-Hamilton Theorem (48) again, we obtain (68). The general case follows by a routine approximation argument based on Lemma 7.2.3 and the remark immediately following.  $\square$

### 7.5. A candidate for $P(u, v)$ .

7.5.1. *Construction of the candidate.* Perform high school division of

$$F \left( t, \sum_{i=1}^{2n} v_i t^{i-1} \right) \in \mathbb{C}[u, v, t]$$

by  $D(u, t)^2$  to obtain an identity

$$(70) \quad F \left( t, \sum_{i=1}^{2n} v_i t^{i-1} \right) = \sum_{i=1}^{2n} P_i(u, v) t^{i-1} + Q_2(u, v, t) D(u, t)^2$$

where

$$Q_2(u, v, t) \in \mathbb{C}[u, v, t]$$

and

$$P(u, v) = (P_1(u, v), \dots, P_{2n}(u, v)) \in \mathbb{C}[u, v]^{2n}.$$

The latter is our candidate for  $P(u, v)$ .

**Lemma 7.5.2.** *Assumptions and notation are as above. Then the following statements hold:*

$$(71) \quad P(u, \varphi(u)) = 0.$$

$$(72) \quad \det \frac{\partial P}{\partial v}(0, 0) \neq 0 \text{ and hence } \frac{\partial P}{\partial v}(u, \varphi(u)) \in \text{GL}_{2n}(\mathbb{C}[[u]]).$$

$$(73) \quad \frac{\partial \varphi}{\partial u}(u) = \left( \frac{\partial P}{\partial v}(u, \varphi(u)) \right)^{-1} \frac{\partial P}{\partial u}(u, \varphi(u)).$$

7.5.3. *Proof of (71).* Perform Weierstrass division of  $F(x, y)$  by  $y - f(x)$  in the formal power series ring  $\mathbb{C}[[x, y]]$  to obtain the identity

$$(74) \quad F(x, y) = (y - f(x))U(x, y)$$

for some  $U(x, y) \in \mathbb{C}[[x, y]]$ . *A priori* one should add a remainder term  $r(x) \in \mathbb{C}[[x]]$  to the right side but substitution of  $y = f(x)$  on both sides and the hypothesis (62)

shows that  $r(x) = 0$ . We then have the following chain of equalities:

$$\begin{aligned} & \sum_{i=1}^{2n} P_i(u, \varphi(u))t^{i-1} + Q_2(u, \varphi(u), t)D(u, t)^2 = F\left(t, \sum_{i=1}^{2n} \varphi(u)t^{i-1}\right) \\ &= \left(\sum_{i=1}^{2n} \varphi(u)t^{i-1} - f(t)\right)U\left(t, \sum_{i=1}^{2n} \varphi_i(u)t^{i-1}\right) \\ &= -Q_1(u, t)D(u, t)^2U\left(t, \sum_{i=1}^n \varphi(u)t^{i-1}\right). \end{aligned}$$

Justifications for the steps are as follows. The first equality we obtain by substituting  $v = \varphi(u)$  into (70). The second equality we obtain by substituting into the factorization given in (74). The third equality we obtain by rearrangement of (65). The equality between the extreme terms of the chain of equalities above forces (71) to hold by the uniqueness of the remainder produced by Weierstrass division.

7.5.4. *Proof of (72)*. Differentiation on both sides of (70) with respect to  $v_j$  followed by evaluation at  $u = v = 0$  yields the relation

$$(75) \quad \frac{\partial F}{\partial y}(t, 0)t^{j-1} = \sum_{i=1}^{2n} \frac{\partial P_i}{\partial v_j}(0, 0)t^{i-1} + \frac{\partial Q_2}{\partial v_j}(0, 0, t)t^{2n}.$$

Now write

$$\frac{\partial F}{\partial y}(x, 0) = \sum_{i=0}^{\infty} b_i x^i \quad (b_i \in \mathbb{C}).$$

By (75) we have

$$\frac{\partial P_i}{\partial v_j}(0, 0) = \begin{cases} b_{i-j} & \text{if } j \leq i, \\ 0 & \text{if } j > i \end{cases}$$

for  $i, j = 1, \dots, 2n$ . Thus we have

$$\det_{i,j=1}^{2n} \frac{\partial P_i}{\partial v_j}(0, 0) = \left(\frac{\partial F}{\partial y}(0, 0)\right)^{2n}.$$

The right side does not vanish by assumption (62). Thus (72) holds.

7.5.5. *Proof of (73)*. Formula (73) holds by implicit differentiation of formula (71). The proof of Lemma 7.5.2 is complete.  $\square$

7.6. **Completion of the proof of Proposition 7.3.4.** As noted above, we may assume that  $N = 0$  and hence that assumptions (61), (62) and (63) are in force. Property (53) follows from formula (71). Property (54) follows from formula (72). Property (55) follows from formulas (67) and (73). Property (56) follows from formula (66). The proof of Proposition 7.3.4 is complete.  $\square$

## 8. PROOF OF THE MAIN RESULT

We finish the proof of Proposition 2.5.3 by checking hypotheses in Proposition 2.4.2, thereby completing the proof of Theorem 1.

**8.1. Review of the setup for Proposition 2.5.3.** Let us start simply by repeating statements (6) and (9) here for the reader's convenience:

$$(76) \quad I_n + a^{(0)}g + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \kappa_j^{(\theta)} (a^{(\theta)}g)^j = 0.$$

(77) The linear map

$$\left( h \mapsto a^{(0)}h + \sum_{\theta=1}^q \sum_{j=2}^{\infty} \sum_{\nu=0}^{j-1} \kappa_j^{(\theta)} (a^{(\theta)}g)^{\nu} (a^{(\theta)}h) (a^{(\theta)}g)^{j-1-\nu} \right)$$

:  $\text{Mat}_n(\mathbb{C}((1/z))) \rightarrow \text{Mat}_n(\mathbb{C}((1/z)))$  is invertible.

Concerning the data appearing above, we have by (7) and (12) that

$$(78) \quad a^{(0)} \in \text{Mat}_n(\mathbb{C}(z)), \quad a^{(1)}, \dots, a^{(q)} \in \text{Mat}_n(\mathbb{C}) \quad \text{and} \quad g \in \text{Mat}_n(\mathbb{C}((1/z))).$$

**8.2. Application of Proposition 7.3.4.** By (7) and (13) we have that

$$\sum_{j=2}^{\infty} \kappa_j^{(\theta)} t^j \in \mathbb{C}[[t]] \text{ is algebraic for } \theta = 1, \dots, q.$$

By assumption (8) and Proposition 6.2.2 we have that

$$\lim_{k \rightarrow \infty} \max_{\theta=1}^q \text{val}(a^{(\theta)}g)^k = 0 \quad \text{and} \quad e(a^{(1)}g), \dots, e(a^{(q)}g) \in (1/z)\mathbb{C}[[1/z]]^n.$$

Thus by Propositions 6.2.6 and 7.3.4 along with the remarks immediately following the latter, there exist an integer  $N \geq 2$  along with

$$\gamma^{(1)}, \dots, \gamma^{(q)} \in (1/z)\mathbb{C}[[1/z]]^{2n} \quad \text{and} \quad P^{(1)}(u, v), \dots, P^{(q)}(u, v) \in \mathbb{C}[u, v]^{2n}$$

such that

$$(79) \quad P^{(\theta)}(e(a^{(\theta)}g), \gamma^{(\theta)}) = 0,$$

$$(80) \quad \det \frac{\partial P^{(\theta)}}{\partial v}(e(a^{(\theta)}g), \gamma^{(\theta)}) \neq 0,$$

$$(81) \quad \sum_{j=2N}^{\infty} \kappa_j^{(\theta)} (a^{(\theta)}g)^j = \sum_{j=1}^{2n} \gamma_j^{(\theta)} (a^{(\theta)}g)^{N+j-1},$$

$$(82) \quad \left[ ((a^{(\theta)}g)^N)^{\flat} \quad \dots \quad ((a^{(\theta)}g)^{N+2n-1})^{\flat} \right] \\ \times \left( \frac{\partial P^{(\theta)}}{\partial v}(e(a^{(\theta)}g), \gamma^{(\theta)}) \right)^{-1} \frac{\partial P^{(\theta)}}{\partial u}(e(a^{(\theta)}g), \gamma^{(\theta)}) = 0, \quad \text{and}$$

$$(83) \quad \sum_{j=2N}^{\infty} \sum_{\nu=0}^{j-1} \kappa_j^{(\theta)} (a^{(\theta)}g)^{\nu} (a^{(\theta)}h) (a^{(\theta)}g)^{j-1-\nu} \\ = \sum_{j=1}^{2n} \sum_{\nu=0}^{N+j-2} \gamma_j^{(\theta)} (a^{(\theta)}g)^{\nu} (a^{(\theta)}h) (a^{(\theta)}g)^{N+j-2-\nu}$$

for  $\theta = 1, \dots, q$  and any  $h \in \text{Mat}_n(\mathbb{C}((1/z)))$ .

**8.3. Polynomial version of (76).** We embed (76) into a system of  $3nq + n^2$  polynomial equations in  $3nq + n^2$  variables with coefficients in  $\mathbb{C}(z)$ .

8.3.1. *Variables.* We employ the family of variables

$$\{U_i\}_{i=1}^{qn} \cup \{V_i\}_{i=1}^{2qn} \cup \{\xi_i\}_{i=1}^{n^2}.$$

Let

$$U = (U_1, \dots, U_{qn}), \quad V = (V_1, \dots, V_{2qn}), \quad \text{and} \quad \xi = (\xi_1, \dots, \xi_{n^2}).$$

Let

$$\Xi = \begin{bmatrix} \xi_1 & \cdots & \xi_{n^2-n+1} \\ \vdots & & \vdots \\ \xi_n & \cdots & \xi_{n^2} \end{bmatrix} \in \text{Mat}_n(\mathbb{C}[\xi]).$$

We break the  $U$ 's and  $V$ 's down into groups by introducing the following notation:

$$\begin{aligned} U_i^{(\theta)} &= U_{i+(\theta-1)q} \quad \text{for } i = 1, \dots, n \text{ and } \theta = 1, \dots, q. \\ V_i^{(\theta)} &= V_{i+(\theta-1)q} \quad \text{for } i = 1, \dots, 2n \text{ and } \theta = 1, \dots, q. \\ U^{(\theta)} &= (U_1^{(\theta)}, \dots, U_n^{(\theta)}) \quad \text{and} \quad V^{(\theta)} = (V_1^{(\theta)}, \dots, V_{2n}^{(\theta)}) \quad \text{for } \theta = 1, \dots, q. \end{aligned}$$

8.3.2. *A special matrix.* We define

$$\begin{aligned} &\mathcal{H}(V, \xi) \\ &= I_n + a^{(0)}\Xi + \sum_{\theta=1}^q \left( \sum_{k=2}^{2N-1} \kappa_k^{(\theta)} (a^{(\theta)}\Xi)^k + \sum_{k=1}^{2n} V_k^{(\theta)} (a^{(\theta)}\Xi)^{N+k-1} \right) \\ &\in \text{Mat}_n(\mathbb{C}(z)[V, \xi]). \end{aligned}$$

8.3.3. *Polynomials.* We define  $3qn + n^2$  polynomials belonging to  $\mathbb{C}(z)[U, V, \xi]$  as follows.

$$\begin{aligned} &F_{i+n(\theta-1)}(U, \xi) \\ &= F_i^{(\theta)}(U^{(\theta)}, \xi) = U_i^{(\theta)} - e_i(a^{(\theta)}\Xi) \quad \text{for } i = 1, \dots, n \text{ and } \theta = 1, \dots, q, \\ &G_{i+2n(\theta-1)}(U, V) \\ &= P^{(\theta)}(U^{(\theta)}, V^{(\theta)}) \quad \text{for } i = 1, \dots, 2n \text{ and } \theta = 1, \dots, q, \text{ and} \\ &H_{i+n(j-1)}(V, \Xi) \\ &= \mathcal{H}(V, \Xi)(i, j) \quad \text{for } i, j = 1, \dots, n. \end{aligned}$$

8.3.4. *Presentation of the system of equations.* Let

$$\begin{aligned} F(U, \xi) &= [ F_1(U, \xi) \quad \cdots \quad F_{qn}(U, \xi) ]^T, \\ G(U, V) &= [ G_1(U, V) \quad \cdots \quad G_{2qn}(U, V) ]^T, \\ H(V, \xi) &= [ H_1(V, \xi) \quad \cdots \quad H_{n^2}(V, \xi) ]^T. \end{aligned}$$

Then our system of polynomial equations takes the form

$$(84) \quad F(U, \xi) = 0, \quad G(U, V) = 0, \quad H(V, \xi) = 0.$$

Note that this system has all coefficients in  $\mathbb{C}(z)$  by (78) and the definitions.

8.3.5. *The solution  $\Upsilon_0$ .* We claim that the following formulas specify a solution over  $\mathbb{C}((1/z))$  of the system of equations (84):

$$(85) \quad U_i^{(\theta)} = e_i(a^{(\theta)}g) \text{ for } i = 1, \dots, n \text{ and } \theta = 1, \dots, q.$$

$$(86) \quad V_i^{(\theta)} = \gamma_i^{(\theta)} \text{ for } i = 1, \dots, 2n \text{ and } \theta = 1, \dots, q.$$

$$(87) \quad \xi_{i+n(j-1)} = \Xi(i, j) = g(i, j) \text{ for } i, j = 1, \dots, n.$$

The equation  $F(U, \xi) = 0$  is obviously satisfied. The equation  $G(U, V) = 0$  is satisfied because it merely restates the system of equations (79). Finally, one verifies that  $H(U, \xi) = 0$  is satisfied by using (76), (81), (85), and (86). The claim is proved. The solution of (84) specified by (85), (86), and (87) will be denoted by  $\Upsilon_0$ .

8.4. **Analysis of the Jacobian determinant.** Now we study the Jacobian matrix

$$(88) \quad \begin{bmatrix} \frac{\partial F}{\partial U}(U, \xi) & 0 & \frac{\partial F}{\partial \xi}(U, \xi) \\ \frac{\partial G}{\partial U}(U, V) & \frac{\partial G}{\partial V}(U, V) & 0 \\ 0 & \frac{\partial H}{\partial V}(V, \xi) & \frac{\partial H}{\partial \xi}(V, \xi) \end{bmatrix} \in \text{Mat}_{n^2+3qn}(\mathbb{C}(z)[U, V, \xi])$$

for the system of equations (84). Let

$$(89) \quad \begin{bmatrix} I_n & 0 & b_{13} \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \in \text{Mat}_{n^2+3qn}(\mathbb{C}((1/z)))$$

be the result of evaluating (88) at the point  $\Upsilon_0$ . To prove Proposition 2.5.3 and thereby to complete the proof of Theorem 1, we have by Proposition 2.4.2 only to prove that the determinant of the matrix (89) does not vanish. Now provided that  $\det b_{22} \neq 0$ , we have a matrix identity

$$\begin{aligned} & \begin{bmatrix} I_n & 0 & b_{13} \\ b_{21} & b_{22} & 0 \\ 0 & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} I_n & 0 & 0 \\ -b_{22}^{-1}b_{21} & I_{2n} & 0 \\ 0 & 0 & I_{n^2} \end{bmatrix} \begin{bmatrix} I_n & 0 & -b_{13} \\ 0 & I_{2n} & 0 \\ 0 & 0 & I_{n^2} \end{bmatrix} \\ &= \begin{bmatrix} I_n & 0 & 0 \\ 0 & b_{22} & 0 \\ -b_{32}b_{22}^{-1}b_{21} & b_{32} & b_{32}b_{22}^{-1}b_{21}b_{13} + b_{33} \end{bmatrix}. \end{aligned}$$

Thus it will be enough to prove that

$$(90) \quad \det b_{22} \neq 0,$$

$$(91) \quad b_{32}b_{22}^{-1}b_{21} = 0, \text{ and}$$

$$(92) \quad \det b_{33} \neq 0.$$

8.5. **Proof of (90).** We have by the definitions

$$(93) \quad b_{22} = \sum_{\theta=1}^q \mathbf{e}_{\theta\theta} \otimes \frac{\partial P^{(\theta)}}{\partial v}(e(a^{(\theta)}g), \gamma^{(\theta)}).$$

Thus (90) holds by (80).

8.6. **Proof of (91).** For  $b_{21}$  and  $b_{32}$  we have formulas similar to (93), namely

$$\begin{aligned} b_{21} &= \sum_{\theta=1}^q \mathbf{e}_{\theta\theta} \otimes \frac{\partial P^{(\theta)}}{\partial u}(e(a^{(\theta)}g), \gamma^{(\theta)}) \text{ and} \\ b_{32} &= \sum_{\theta=1}^q \mathbf{e}_{\theta\theta} \otimes [ ((a^{(\theta)}g)^N)^{\flat} \quad \dots \quad ((a^{(\theta)}g)^{N+2n-1})^{\flat} ]. \end{aligned}$$

Thus (91) holds by (82).

8.7. **Proof of (92).** We have for  $i, j = 1, \dots, n$  that

$$\begin{aligned} &\frac{\partial \mathcal{H}(V, \xi)}{\partial \xi_{i+(j-1)n}} \\ &= a^{(0)} \mathbf{e}_{ij} + \sum_{\theta=1}^q \sum_{k=2}^{2N-1} \sum_{\nu=0}^{k-1} \kappa_k^{(\theta)} (a^{(\theta)} \Xi)^{\nu} (a^{(\theta)} \mathbf{e}_{ij}) (a^{(\theta)} \Xi)^{k-1-\nu} \\ &\quad + \sum_{\theta=1}^q \sum_{k=1}^{2n} \sum_{\nu=0}^{N+k-2} V_k^{(\theta)} (a^{(\theta)} \Xi)^{\nu} (a^{(\theta)} \mathbf{e}_{ij}) (a^{(\theta)} \Xi)^{N+k-2-\nu} \end{aligned}$$

and hence after evaluating both sides at  $\Upsilon_0$  and using (83), we find that  $b_{33}$  is a matrix describing with respect to the basis

$$\mathbf{e}_{11}, \dots, \mathbf{e}_{n1}, \dots, \mathbf{e}_{1n}, \dots, \mathbf{e}_{nn} \in \text{Mat}_n(\mathbb{C}((1/z)))$$

the invertible linear map considered in (77). Thus (92) holds. Thus the proof of Proposition 2.5.3 is complete and with it the proof of Theorem 1.

**Acknowledgements:** I thank Serban Belinschi, J. William Helton, Tobias Mai, and Roland Speicher for communications concerning the self-adjoint linearization trick and symmetric realization. I thank Steven Lalley for communications concerning algebraicity in relation to random walk. I thank Christine Berkesch Zamaere for communications concerning algebraicity criteria and commutative algebra.

#### REFERENCES

- [1] Anderson, G., *Convergence of the largest singular value of a polynomial in independent Wigner matrices*. Annals of Prob. **41**(2013), 2103–2181. MR3098069
- [2] Anderson, G., Zeitouni, O.: *A law of large numbers for finite-range dependent random matrices*. Comm. Pure Appl. Math. **61**(2008), 1118–1154. MR2417889
- [3] Anderson, G., Guionnet, A., and Zeitouni, O., *An Introduction to Random Matrices*. Cambridge Studies in Advanced Mathematics **118**. Cambridge University Press, Cambridge 2010. MR2760897
- [4] Aomoto, K. *Spectral theory on a free group and algebraic curves*. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **31**(1984), no. 2, 297–318. MR0763424
- [5] Arbarello, E., De Concini, C., Kac, V. G. *The infinite wedge representation and the reciprocity law for algebraic curves*. Theta functions–Bowdoin 1987, Part 1 (Brunswick, ME, 1987), 171–190, Proc. Sympos. Pure Math., **49**, Part 1, Amer. Math. Soc., Providence, RI, 1989. MR1013132
- [6] Artin, E., *Algebraic numbers and algebraic functions*. Reprint of the 1967 original. AMS Chelsea Publishing, Providence, RI, 2006. MR2218376
- [7] Berstel, J., Reutenauer, C., *Noncommutative rational series with applications*. Encyclopedia of Mathematics and its Applications, **137**. Cambridge University Press, Cambridge, 2011. MR2760561

- [8] Belinschi, S., Mai, T., and Speicher, R., *Analytic subordination theory of operator-valued free additive convolution and the solution of a general random matrix problem*. [arXiv:1303.3196](#)
- [9] Brieskorn, E., Knörrer, H.: *Plane algebraic curves*. Translated from the German original by John Stillwell. [2012] reprint of the 1986 edition. Modern Birkhser Classics. Birkhäuser/Springer Basel AG, Basel, 1986. MR2975988
- [10] Brockett, R., *Finite dimensional linear systems*. John Wiley and Sons, New York, 1970.
- [11] Chomsky, N., Schützenberger, M., *The algebraic theory of context-free languages*. 1963 Computer programming and formal systems, pp. 118–161, North-Holland, Amsterdam MR0152391
- [12] Dykema, K., *On the  $S$ -transform over a Banach algebra*. J. Funct. Anal. **231**(2006), no. 1, 90–110. MR2190164
- [13] Rao, N. Raj, Edelman, A. *The polynomial method for random matrices*. Found. Comput. Math. **8**(2008), no. 6, 647–702. MR2461243
- [14] Eisenbud, D. *Commutative algebra with a view toward algebraic geometry*. Graduate Texts in Mathematics, **150**, Springer-Verlag, New York, 1995. MR1322960
- [15] Haiman, M., *Noncommutative rational power series and algebraic generating functions*. European J. Combin. **14** (1993), no. 4, 335–339. MR1226580
- [16] Haagerup, U., Schultz, H. and Thorbjørnsen, S. *A random matrix approach to the lack of projections in  $C_{\text{red}}^*(\mathbb{F}_2)$* . Adv. Math. **204**(2006), 1–83. MR2233126
- [17] Haagerup, U. and Thorbjørnsen, S. *A new application of random matrices:  $\text{Ext}(C^*(\mathbb{F}_2))$  is not a group*. Annals Math. **162**(2005), 711–775. MR2183281
- [18] Helton, J., Rashidi Far, R., Speicher, R., *Operator-valued semicircular elements: solving a quadratic matrix equation with positivity constraints*. Int. Math. Res. Not. IMRN **2007**, no. 22, Art. ID rnm086. MR2376207
- [19] Helton, J., McCullough, S., Vinnikov, V. *Noncommutative convexity arises from linear matrix inequalities*. J. Funct. Anal. **240**(2006), no. 1, 105–191. MR2259894
- [20] Kalman, R., Falb, P., Arbib, M. A., *Topics in mathematical system theory*. McGraw-Hill Book Co., New York-Toronto, Ont.-London 1969, MR0255260
- [21] Kassel, C. and Reutenauer, C., *Algebraicity of the zeta function associated to a matrix over a free group algebra* [arXiv:1303.3481](#)
- [22] Kontsevich, M. *Noncommutative identities* [arXiv:1109.2469](#)
- [23] Lalley, S. *Random walks on regular languages and algebraic systems of generating functions*. Algebraic methods in statistics and probability (Notre Dame, IN, 2000), 201–230, Contemp. Math. **287**, Amer. Math. Soc., Providence, RI, 2001. MR1873677
- [24] Lang, S. *Algebra*. Revised third edition. Graduate Texts in Mathematics **211**. Springer-Verlag, New York, 2002. MR1878556
- [25] Mumford, D. *An algebro-geometric construction of commuting operators and of solutions to the Toda lattice equation, Korteweg de Vries equation and related nonlinear equation*. Proceedings of the International Symposium on Algebraic Geometry (Kyoto Univ., Kyoto, 1977), pp. 115–153, Kinokuniya Book Store, Tokyo, 1978. MR0578857
- [26] Nagnibeda, T., Woess, W. *Random walks on trees with finitely many cone types*. J. Theoret. Probab. **15**(2002), no. 2, 383–422. MR1898814
- [27] Nica, A., Speicher, R., *Commutators of free random variables*. Duke Math. J. **92**(1998), no. 3, 553–592. MR1620518
- [28] Nica, A. and Speicher, R., *Lectures on the combinatorics of free probability*. London Mathematical Society Lecture Note Series **335**. Cambridge University Press, Cambridge 2006. MR2266879
- [29] Salomaa, A., Soittola, M., *Automata-theoretic aspects of formal power series*. Texts and Monographs in Computer Science. Springer-Verlag, New York-Heidelberg, 1978. MR0483721
- [30] Schützenberger, M. P., *On the definition of a family of automata*. Information and Control **4** 1961 245270. MR0135680
- [31] Shlyakhtenko, D., Skoufranis, P. *Freely Independent Random Variables with Non-Atomic Distributions*. [arXiv:1305.1920](#)
- [32] Speicher, R., *Free probability and random matrices*. To appear in the proceeding of ICM 2014. [arXiv:1404.3393](#)

- [33] Stanley, R., *Enumerative combinatorics. Vol. 2*. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge Studies in Advanced Mathematics, **62**. Cambridge University Press, Cambridge, 1999, MR1676282
- [34] Tate, J. *Residues of differentials on curves*. Ann. Sci. cole Norm. Sup. (4) **1** 1968 149–159. MR0227171
- [35] Voiculescu, D., *Addition of certain noncommuting random variables*. J. Funct. Anal. **66**(1986), no. 3, 323346. MR0839105
- [36] Voiculescu, D., *Free Probability for Pairs of Faces I* [arXiv:1306.6082](https://arxiv.org/abs/1306.6082).
- [37] Voiculescu, D. *Operations on certain non-commutative operator-valued random variables*. Recent advances in operator algebras (Orléans, 1992). Astérisque No. **232** (1995), 243–275. MR1372537
- [38] Voiculescu, D. V., Dykema, K. J. and Nica, A., *Free random variables. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups*. CRM Monograph Series **1**. American Mathematical Society, Providence, RI, 1992. MR1217253
- [39] Williamson, R., Trotter, H. *Multivariable mathematics*. Fourth edition. Pearson, 2003.
- [40] Woess, W. *Context-free languages and random walks on groups*. Discrete Math. **67**(1987), no. 1, 81–87. MR0908187
- [41] Woess, W., *Random walks on infinite graphs and groups*. Cambridge Tracts in Mathematics, **138**. Cambridge University Press, Cambridge, 2000. MR1743100
- [42] Zariski, O. and Samuel, P., *Commutative algebra. Vol. II*. Reprint of the 1960 edition. Graduate Texts in Mathematics, Vol. **29**. Springer-Verlag, New York-Heidelberg, 1975. MR0389876

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MN 55455  
E-mail address: [gwanders@umn.edu](mailto:gwanders@umn.edu)