



Citation for published version:

Johnston, S & Prochno, J 2022, 'Faà di Bruno's formula and inversion of power series', *Advances in Mathematics*, vol. 395, 108080. <https://doi.org/10.1016/j.aim.2021.108080>

DOI:

[10.1016/j.aim.2021.108080](https://doi.org/10.1016/j.aim.2021.108080)

Publication date:

2022

Document Version

Peer reviewed version

[Link to publication](#)

Publisher Rights

CC BY-NC-ND

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

FAÀ DI BRUNO'S FORMULA AND INVERSION OF POWER SERIES

SAMUEL G.G. JOHNSTON AND JOSCHA PROCHNO

ABSTRACT. Faà di Bruno's formula gives an expression for the derivatives of the composition of two real-valued functions. In this paper we prove a multivariate and synthesized version of Faà di Bruno's formula in higher dimensions, providing a combinatorial expression for the derivatives of chain compositions $F^{(1)} \circ \dots \circ F^{(m)}$ of functions $F^{(l)} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ in terms of sums over labelled trees. We give several applications of this formula, including a new involution formula for the inversion of multivariate power series. We use this framework to outline a combinatorial approach to studying the invertibility of polynomial mappings, giving a purely combinatorial restatement of the Jacobian conjecture. Our methods extend naturally to the non-commutative case, where we prove a free version of Faà di Bruno's formula for multivariate power series in free indeterminates, and use this formula as a tool for obtaining a new inversion formula for free power series.

1. INTRODUCTION

1.1. The one dimensional case. Before discussing our results in full generality, we give an outline of our approach by first considering the one dimensional case. Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$ be smooth functions and consider the derivatives of the product fg . The product rule states that $D[fg] = D[f]g + fD[g]$, and by applying an induction argument it is straightforward to prove the more general Leibniz rule, which states that the n^{th} derivative of fg is given by

$$(1) \quad D^n[fg] = \sum_{S \sqcup T = [n]} D^{\#S}[f]D^{\#T}[g],$$

where $[n] = \{1, \dots, n\}$, and the sum in (1) ranges over all pairs (S, T) of disjoint (possibly-empty) subsets of $[n]$ satisfying $S \cup T = [n]$.

Our somewhat unconventional representation of the Leibniz rule (1) as a sum over pairs of subsets is motivated by a desire to avoid using binomial coefficients. This choice is one we will make throughout the paper: wherever possible we will circumvent combinatorial coefficients by indexing every sum through a sufficiently rich collection of combinatorial objects.

Consider now the derivatives of the *composition* $f \circ g$. Here the chain rule $D[f \circ g] = (D[f] \circ g)D[g]$ plays the role of the product rule, and the analogue of the Leibniz rule (1) is furnished by the lesser known Faà di Bruno formula

$$(2) \quad D^n[f \circ g] = \sum_{\pi \in \mathcal{P}_{[n]}} (D^{\#\pi}[f] \circ g) \prod_{\Gamma \in \pi} D^{\#\Gamma}[g],$$

where the sum ranges over $\mathcal{P}_{[n]}$, the collection of set partitions of $[n]$, $\#\pi$ counts the number of blocks in the partition π , and $\#\Gamma$ counts the number of elements of a block Γ .

Suppose now $f(X) = \sum_{k \geq 1} \frac{f_k}{k!} X^k$ and $g(X) = \sum_{k \geq 1} \frac{g_k}{k!} X^k$ are formal power series with no constant term. Then Faà di Bruno's formula (4.1) tells us that the n^{th} coefficient of the composite power series

2010 *Mathematics Subject Classification.* Primary: 26B05, 05C05, 13F25, 13P99, Secondary: 05A18.
Key words and phrases. Rooted trees, Faà di Bruno formula, power series inversion.

for $(f \circ g)(X) = \sum_{k \geq 1} \frac{(f \circ g)_k}{k!} X^k$ is given by

$$(3) \quad (f \circ g)_n = \sum_{\pi \in \mathcal{P}_{[n]}} f_{\#\pi} \prod_{\Gamma \in \pi} g_{\#\Gamma}.$$

We now show that (3) may be applied to the problem of inverting power series. Indeed, suppose now that f and g are inverses of one another, in the sense that $(f \circ g)(X) = (g \circ f)(X) = X$. Then (3) endows us with the system of equations

$$(4) \quad \sum_{\pi \in \mathcal{P}_{[n]}} f_{\#\pi} \prod_{\Gamma \in \pi} g_{\#\Gamma} = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Using this system of equations it is straightforward to give an expression for each coefficient g_k in terms of the (f_k) and a collection of rooted trees. To this end, let \mathbb{S}_k be the set of finite rooted graph-theoretic trees \mathcal{T} such that leaves are in bijection with $[k]$ and such that every non-leaf vertex has two or more children. Given such a tree \mathcal{T} , and a collection of indeterminates $(h_k : k \geq 2)$, define the h -energy of the tree \mathcal{T} by

$$\mathcal{E}_h(\mathcal{T}) := \prod_{v \in I} h_{\#\text{children of } v},$$

where the product is taken over I , the set of all internal vertices v in the tree (those vertices which are not leaves).

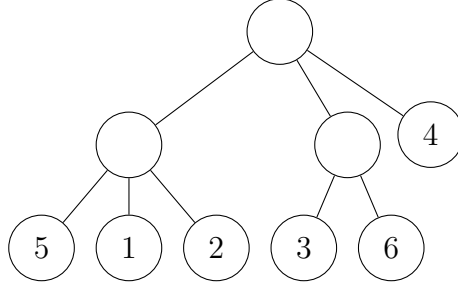


FIGURE 1. An element \mathcal{T} of \mathbb{S}_6 . The h -energy of this tree is $\mathcal{E}_h(\mathcal{T}) = h_2 h_3^2$.

We now use the system of equations (4) to give a sketch proof of the following result, which states that the coefficients of a compositional inverse for a power series with coefficients (h_k) may be written in terms of a sum over a set of trees of h -energies $\mathcal{E}_h(\mathcal{T})$. For the sake of simplicity we assume here that $f'(0) = 1$.

Proposition 1.1. *Let $g(X) = \sum_{k \geq 0} \frac{g_k}{k!} X^k$ be the formal inverse of the power series $f(X) = X - \sum_{k \geq 2} \frac{h_k}{k!} X^k$. Then $g_0 = 0$, $g_1 = 1$, and for $k \geq 2$, the coefficients of g are given by*

$$(5) \quad g_k = \sum_{\mathcal{T} \in \mathbb{S}_k} \mathcal{E}_h(\mathcal{T}).$$

Proof. It is immediate that $g_0 = 0$, and the fact that $g_1 = 1$ follows from the chain rule. The proof of formula (5) for $k \geq 2$ is established by an inductive argument using the system (4). Indeed, plugging $n = k + 1$ in (4), using the fact that $f_1 = 1$ and $f_j = -h_j$ for $j \geq 2$, we obtain

$$g_{k+1} = \sum_{\pi \in \mathcal{P}_{k+1}: \#\pi \geq 2} h_{\#\pi} \prod_{\Gamma \in \pi} g_{\#\Gamma}.$$

Using the inductive hypothesis to expand each term $g_{\#\Gamma}$ in the product as a tree, by thinking of each term in this product as a subtree of a tree in which the root has $\#\pi$ children, each term in the sum on the right-hand side now corresponds to $\mathcal{E}_h(\mathcal{T})$ for a tree \mathcal{T} in \mathbb{S}^{k+1} . \square

While the contents of the discussion above are well known, with Proposition 1.1 following as a fairly straightforward consequence of the single-variable Lagrange inversion formula, the purpose of the present article is to expand this discussion to multivariate functions and power series in both commutative and non-commutative variables, and supply formulas for the composition and inversion of these functions. We now overview both of these cases.

1.2. The commutative case. The main commutative result in this article is Theorem 2.1, a generalisation of Faà di Bruno's formula (4.1) giving a combinatorial expression for the higher derivatives

$$(6) \quad D^\alpha [F^{(1)} \circ \dots \circ F^{(m)}],$$

where $F^{(l)} : \mathbb{R}^N \rightarrow \mathbb{R}^N$ are smooth functions and α is an arbitrary multi-index. The expression for (6) is given in terms of a sum over labelled trees with generations 0 through m : the number of leaves in each of these trees is equal to degree of the multi-index α , and every leaf lies in generation m . Figure 2 gives an example of one of these trees.

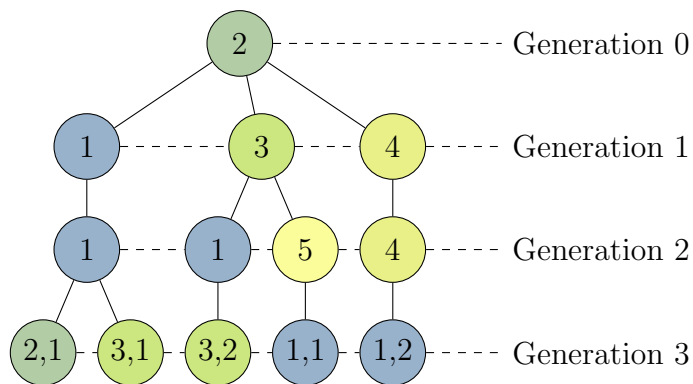


FIGURE 2. The fifth order partial derivatives of a three-fold composition $F \circ G \circ H$ are given in terms of a sum over labelled trees with $3 + 1$ generations and 5 leaves.

We present several combinatorial applications of Theorem 2.1, including most importantly a new inversion formula for formal power series. Indeed, suppose that $F = (F_1, \dots, F_N)$ is an N -dimensional formal power series with i^{th} component

$$F_i(X) = \sum_{\alpha \in \mathbb{Z}_{\geq 0}^N} \frac{F_{i,\alpha}}{\alpha!} X^\alpha,$$

where $F_{i,\alpha}$ take values in some ring. In Theorem 2.5 we show that the coefficients of the compositional formal inverse power series $G = (G_1, \dots, G_N)$ may be given in terms of a sum over labelled trees of energies depending on the coefficients $(F_{i,\alpha})$ — provided the linear term of F is an invertible matrix. The special case of our result where the linear term of F is identical appears implicitly in Haiman and Schmitt [26]. Roughly speaking, we find that when the linear term is non-identical, it interlaces the energy function of the tree. In Section 2.5 we conclude our discussion of commutative power series by applying our framework to outline a combinatorial path to proving the Jacobian conjecture, building heavily on the work of Wright [46–49] and Singer [38–40].

1.3. The non-commutative case. We then turn to looking at the non-commutative case, where the methods above are naturally adapted to studying composition and compositional inversion of formal power series in free variables X_1, \dots, X_N which are not assumed to commute (we use the term ‘free’ throughout to mean ‘non-commutative’). In essence, our results in the commutative case extend to the free case via the rule of thumb “*free variables means planar trees*”.

Our first result in this direction is a free version of the multi-dimensional Faà di Bruno formula, suitable for understanding compositions of formal expressions such as

$$f(X_1, X_2, X_3) = X_1X_2X_1 + X_2X_3 + 4X_3X_2,$$

where X_i are free variables. We permit such expressions to have infinitely many terms, and call such objects *free formal power series*.

In Theorem 3.2 we find that if $F^{(1)}, \dots, F^{(m)}$ are free formal power series in N variables such that each $F^{(l)}$ has N components $(F_1^{(l)}, \dots, F_N^{(l)})$, then the coefficient of a term $X_{i_1} \dots X_{i_k}$ may be given in terms of a sum over labelled *planar* trees with m generations and k leaves. By a *planar* tree, we refer to a rooted tree in which the children of every internal vertex are ranked from left to right.

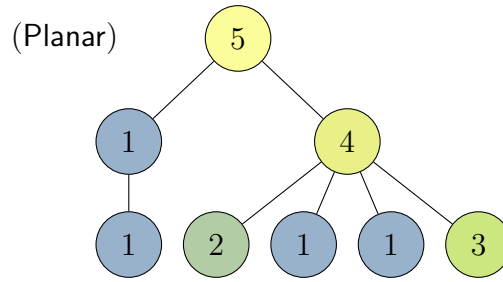


FIGURE 3. The coefficient of $X_1X_2X_1X_1X_3$ in the fifth component of the composition $F \circ G$ may be given in terms of a sum over planar trees with two generations and leaves of types 1, 2, 1, 1, 3 from left to right.

Consider now the compositional inversion of free power series. Our final result, Theorem 3.4, states that like the commutative case, a mapping F in non-commutative variables has a left-inverse G if and only if its linear term is invertible as matrix with entries in the coefficient ring, and that in this case, the mapping G is also a right-inverse. Furthermore, Theorem 3.4 gives an explicit formula for each inverse series G_i in terms of rooted planar trees.

1.4. Structure of the paper. The remainder of the paper is structured as follows:

- In Section 2 we give full statements of our main results in the commutative case. This amounts to our multivariate generalisation of Faà di Bruno’s formula, Theorem 2.1, our inversion formulas for multivariate power series, Theorems 2.3 and 2.5, and a new combinatorial statement for the Jacobian conjecture, Conjecture 2.9.
- In Section 3 we give statements of our results in the non-commutative case. These results consist of Theorem 3.2 — our free Faà di Bruno formula for multivariate power series in N non-commuting indeterminates — as well as an inversion formula for such series, Theorem 3.4.
- In Section 4 we collect together several applications of Faà di Bruno’s formula and the power series inversion formulas, including a representation for the Hermite polynomials in terms of partitions, as well as a new formula for the reciprocal of a power series. Finally, we give a brief overview of recent appearances of Faà di Bruno’s formula in the genealogical structure of a class of random trees known as Galton-Watson trees.

- The remaining sections are dedicated to giving proofs of the results stated in Sections 2 and 3: in Section 5 we prove our generalisation of the classical Faà di Bruno formula, Theorem 2.1, in Section 6 we prove our power series inversion formulas, Theorem 2.3 and Theorem 2.5, and in the final section, Section 7, we prove Theorems 3.2 and 3.4 which give statements about the composition and inversion of power series in non-commuting variables.

2. THE COMMUTATIVE CASE

We shall now present the main results and contributions of this article in the commutative case. In order to begin discussing the higher dimensional setting, we first need to introduce some additional notation.

2.1. Multi-indices and the higher dimensional Leibniz rule. Recall that $[N] := \{1, \dots, N\}$. Whenever S is a finite subset of $[N] \times \mathbb{Z}_{\geq 0}$, let $\#S$ be the multi-index whose i^{th} component counts the number of elements of the form (i, a) in S . Now given a multi-index $\alpha \in \mathbb{Z}_{\geq 0}^N$, define

$$[\alpha] := \{(i, a) \in \mathbb{Z}^2 : i \in [N], 1 \leq a \leq \alpha_i\} \subseteq [N] \times \mathbb{Z}_{\geq 0}.$$

Plainly $\#[\alpha] = \alpha$. Finally, let $(\mathbf{e}_i)_{i \in [N]}$ be the standard basis for $\mathbb{Z}_{>0}^N$.

We are now equipped to provide a high-dimensional version of the Leibniz rule, namely for functions $f^{(1)}, f^{(2)}, \dots, f^{(m)} : \mathbb{R}^N \rightarrow \mathbb{R}$ the higher derivatives of their product $f^{(1)} f^{(2)} \dots f^{(m)}$ are given by

$$(7) \quad D^\alpha [f^{(1)} f^{(2)} \dots f^{(m)}] = \sum_{S_1 \sqcup \dots \sqcup S_m = [\alpha]} \prod_{l=1}^m D^{\#S_l} [f^{(l)}],$$

where the sum ranges over all m -tuples (S_1, \dots, S_m) of disjoint (and possibly empty) subsets of $[\alpha]$ whose union is equal to α . The expression (7) is easily proved by induction.

With a view to stating Theorem 2.1, which gives an expression analogous to (7) for higher derivatives $D^\alpha [F^{(1)} \circ \dots \circ F^{(m)}]$ of chain compositions, in the next section we introduced labelled rooted trees.

2.2. Labelled rooted trees. We say a finite graph (V, E) is a rooted tree if it is a finite tree with a designated root vertex v_0 . For $k \in \mathbb{Z}_{\geq 0}$, let V_k denote the set of vertices whose graph distance from the root is k . We refer to V_k as generation k , and note that $V_0 = \{v_0\}$. Each vertex w in V_{k+1} has a unique parent vertex v in V_k , and in this case we say w is a child of v . We say a vertex is a leaf if it has no children, and we say it is internal if it has at least one child. The vertex set of a rooted tree has a decomposition into a disjoint union $V = L \sqcup I$, where L are the leaves of the tree and the set I consists of the internal vertices of the tree.

For $i \in [N]$ and a finite subset S of $[N] \times \mathbb{Z}_{\geq 0}$ let $\mathbb{T}_{i,S}$ to be the set of labelled trees \mathcal{T} with root type i and leaf types S . More specifically, $\mathbb{T}_{i,S}$ is the set of quadruplets $\mathcal{T} := (V, E, \tau, \phi)$ with the following properties:

- The pair (V, E) denotes a rooted tree.
- The labelling function $\phi : L \rightarrow S$ is a bijection between the leaves of the tree L and the set S , giving each leaf a *label* in S .
- The typing function $\tau : V \rightarrow [N]$ gives each vertex of the tree a *type* in $[N]$ according to certain rules: it gives the root type i , and the type given to each leaf respects the label of that leaf, in the sense that whenever $\phi(v) = (j, a)$ for some $a \in \{1, 2, \dots\}$, we have $\tau(v) = j$. There are no constraints on the types of non-root internal vertices.

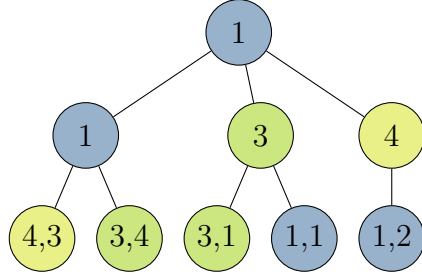


FIGURE 4. An element $\mathbb{F}_{i,S}(2)$ where $i = 1$ and $S = \{(1, 1), (1, 2), (3, 1), (3, 4), (4, 3)\}$. A vertex labelled (k, a) has type k . The colours of the vertices correspond to their types. Note that every leaf lies in the final generation.

We say two labelled trees (V, E, τ, ϕ) and (V', E', τ', ϕ') are *isomorphic* if there is a bijection $\Psi : V \rightarrow V'$ between the underlying vertex sets preserving all of the structure of the tree: $(\Psi(u), \Psi(v)) \in E'$ if and only if $(u, v) \in E$, $\tau'(\Psi(u)) = \tau(u)$ and $\phi'(\Psi(u)) = \phi(u)$. Whenever we speak of a set or collection of trees, technically we mean a set or collection of isomorphism classes according to this equivalence.

So in summary, for subsets S of $[N] \times \{1, 2, \dots\}$, $\mathbb{T}_{i,S}$ is the collection of (isomorphism classes of) labelled trees with root type i and whose leaves are in bijection with S . We will consider several subsets of $\mathbb{T}_{i,S}$:

- The set of *final* trees of length m , $\mathbb{F}_{i,S}(m)$: those trees in $\mathbb{T}_{i,S}$ such that every leaf is contained in generation m . The tree in Figure 4 is a final tree of length 2.
- The set of *proper* trees $\mathbb{S}_{i,S}$: those trees in $\mathbb{T}_{i,S}$ such that every internal vertex has two or more children.
- The set of *alternating* trees $\mathbb{A}_{i,S}$: the subset of $\mathbb{T}_{i,S}$ consisting of trees such that every vertex in some V_{2k} has exactly one child, and every vertex in some V_{2k+1} has no children or two or more children. In particular, every leaf of an alternating tree lies in some V_{2k+1} . See e.g. Figure 5

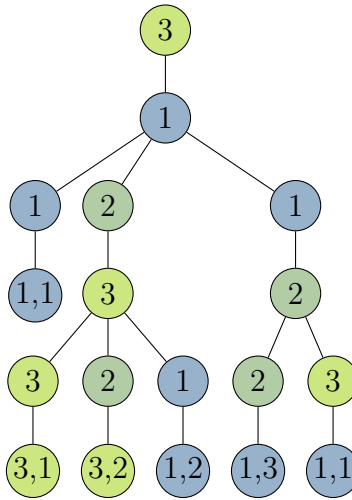


FIGURE 5. An alternating tree in $\mathbb{A}_{i,S}$ with root type $i = 3$, and leaf labelling set $S = \{(1, 1), (1, 2), (1, 3), (3, 1), (3, 2)\} = [3\mathbf{e}_1 + 2\mathbf{e}_3]$. Note that every vertex in an even generation has exactly one child.

To lighten notation we write $\mathbb{T}_{i,\alpha} := \mathbb{T}_{i,[\alpha]}$, and we do similarly with $\mathbb{F}_{i,\alpha}(m)$, $\mathbb{S}_{i,\alpha}$ and $\mathbb{A}_{i,\alpha}$. Final trees $\mathbb{F}_{i,\alpha}(m)$ will appear in our generalisation of Faà di Bruno's formula, specifically for the α^{th} derivative

of the i^{th} component of $F^{(1)} \circ \dots \circ F^{(m)}$. The proper trees $\mathbb{S}_{i,\alpha}$ will appear in our inversion formula for power series with identity linear terms. The unusual looking alternating trees $\mathbb{A}_{i,\alpha}$ are used in our inversion formula for power series with non-identical linear terms.

For an internal vertex v the outdegree $\mu(v)$ of v is defined to be the multi-index whose i^{th} component is given by

$$\mu(v)_i := \#\{w \in V : w \text{ is a child of } v \text{ and } \tau(w) = i\}.$$

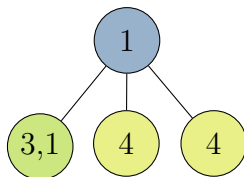


FIGURE 6. An internal vertex v with type $\tau(v) = 1$ and outdegree $\mu(v) = \mathbf{e}_3 + 2\mathbf{e}_4$.

All of our formulas in the commutative case for the composition and inversion of functions take the form $\sum_{\mathcal{T} \in \mathcal{C}} \mathcal{E}(\mathcal{T})$, where \mathcal{C} is some collection of labelled trees, and the energy has the form $\mathcal{E}(\mathcal{T}) = \prod_{v \in I} K(\tau(v), \mu(v))$ for some function K depending on the type of the vertex and its outdegree.

We are now equipped to discuss the specific cases in the next two sections.

2.3. Generalisation of Faà di Bruno's formula. Let $F^* := (F^{(1)}, \dots, F^{(m)})$ be an m -tuple of smooth functions from \mathbb{R}^N to \mathbb{R}^N . We now give a formula for the higher derivatives of their composition in terms of the derivatives of each individual $F^{(l)}$ and a sum over a collection of trees.

Whenever $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$, let $D_i^\alpha[F] : \mathbb{R}^N \rightarrow \mathbb{R}$ denote the α^{th} derivative of the i^{th} component. Given a labelled tree $\mathcal{T} \in \mathbb{F}_{i,\alpha}(m)$ define the energy function $\mathcal{E}_{F^*}(\mathcal{T}) : \mathbb{R}^N \rightarrow \mathbb{R}$ by

$$\mathcal{E}_{F^*}(\mathcal{T}) := \prod_{l=1}^m \prod_{v \in V_{l-1}} D_{\tau(v)}^{\mu(v)}[F^{(l)}] \circ F^{(l+1)} \circ \dots \circ F^{(m)}.$$

Suppose now $F^* = (F, G)$, and consider the tree \mathcal{T} in $\mathbb{F}_{5,\alpha}(2)$ in Figure 4. For this tree we have

$$\mathcal{E}_{F^*}(\mathcal{T}) = (D_5^{\mathbf{e}_1 + \mathbf{e}_3 + \mathbf{e}_4}[F] \circ G) D_1^{\mathbf{e}_3 + \mathbf{e}_4}[G] D_3^{\mathbf{e}_1 + \mathbf{e}_3}[G] D_4^{\mathbf{e}_1}[G].$$

We are now equipped to state our main result, Theorem 2.1, which states that the higher derivatives of chain compositions may be given in terms of a sum over tree energies.

Theorem 2.1. *The function $D_i^\alpha[F^{(1)} \circ \dots \circ F^{(m)}] : \mathbb{R}^N \rightarrow \mathbb{R}$ is given by*

$$D_i^\alpha[F^{(1)} \circ \dots \circ F^{(m)}](s) = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(m)} \mathcal{E}_{F^*}(\mathcal{T})(s).$$

Remark 2.2. A formula for the case where $F^{(l)} : \mathbb{R}^{N_l} \rightarrow \mathbb{R}^{N_{l+1}}$ are maps between Euclidean spaces of different dimensions is easily obtained from Theorem 2.1 by means of an embedding argument. Indeed, choose N large enough so that $N \geq N_l$ for all l , and then canonically associate each $F^{(l)}$ with $\hat{F}^{(l)} : \mathbb{R}^N \rightarrow \mathbb{R}^N$, where $\hat{F}^{(l)}$ is independent of the final $N - N_l$ variables, and the final $N - N_{l+1}$ of $\hat{F}^{(l)}$ components are zero. Applying Theorem 2.1 to the chain $\hat{F}^{(1)} \circ \dots \circ \hat{F}^{(m)}$, this procedure tells us that

$$D_i^\alpha[F^{(1)} \circ \dots \circ F^{(m)}](s) = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(m)} \mathcal{E}_{F^*}(\mathcal{T})(s),$$

where the sum is supported only on trees where the type function τ is restricted so that for each ℓ , the vertices in the ℓ^{th} generation V_ℓ may only have types in $\{1, \dots, N_\ell\} \subset [N]$.

Theorem 2.1 reduces significantly when each $F^{(l)}$ is assumed to be a power series with zero constant term and we evaluate the result at the origin. Namely suppose the i^{th} component of each $F^{(l)}$ is given by

$$F_i^{(l)}(X_1, \dots, X_N) = \sum_{\alpha \neq 0} \frac{F_{i,\alpha}^{(l)}}{\alpha!} X^\alpha.$$

Then by setting $X = 0$ in Theorem 2.1 we see that the coefficient of X^α in the i^{th} component of the composition power series $F^{(1)} \circ \dots \circ F^{(m)}$ is given by $\frac{1}{\alpha!} (F^{(1)} \circ \dots \circ F^{(m)})_{i,\alpha}$, where

$$(8) \quad (F^{(1)} \circ \dots \circ F^{(m)})_{i,\alpha} := \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(m)} \prod_{l=1}^m \prod_{v \in V_{l-1}} F_{\tau(v),\mu(v)}^{(l)}.$$

We now take a brief look at the special case $m = 2$, for which some results have appeared in the literature. Setting $m = 2$ in (8) we obtain

$$(9) \quad (F \circ G)_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} F_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)}.$$

The formula (9) essentially appears in Haiman and Schmitt [26], who give their formula in terms of what they call a *multicolor partition*. Taking a moment to sketch this idea here, each tree $\mathbb{F}_{i,\alpha}(2)$ of length 2 induces a partition π of $[\alpha]$, by letting the blocks of π correspond to the vertices in the middle generation V_1 : we say two elements (i, j) and (i', j') of $[\alpha]$ are in the same block of the partition if the leaves with these labels are children of the same vertex in V_1 . The types of the vertices in V_1 then induces a function $\tau : \pi \rightarrow [N]$. Appealing to this correspondence we may rewrite (9) as

$$(F \circ G)_{i,\alpha} = \sum_{\pi \in \mathcal{P}_{[\alpha]}} \sum_{\tau : \pi \rightarrow [N]} F_{i,\#\tau} \prod_{\Gamma \in \pi} G_{\tau(\Gamma),\#\Gamma},$$

where the former sum ranges over all partitions π of $[\alpha]$, the latter sum ranges over all functions τ labelling the blocks of π , and $\#\tau$ denotes the multi-index whose i^{th} counts the number of blocks of π such that $\tau(\Gamma) = i$.

Before discussing inversion of power series in the next section, we close this section with a brief discussion of the Faà di Bruno formula and its generalisations. While the formula, is as the name indicates, typically attributed to Italian mathematician Faà di Bruno [15–17], it first appeared in the calculus book [2] of Arbogast in the year 1800. We refer the reader to Johnson [28] for a more thorough historical discussion of Faà di Bruno formula.

Several intermediate generalisations of Faà di Bruno's formula have appeared in the literature. Abraham and Robbin [3, Section 1.4] develop a composite mapping formula for sufficiently differentiable functions $f : E \rightarrow F$ and $g : F \rightarrow G$ defined on Banach spaces, only giving the integer coefficients implicitly through a recurrence relation. Their composite mapping formula is mainly used in [3] as a tool to prove the Glaeser rough composition theorem. Constantine and Savits [8] study the case where $f : \mathbb{R}^b \rightarrow \mathbb{R}^c$ and $g : \mathbb{R}^a \rightarrow \mathbb{R}^b$, giving the derivatives for $D^\alpha[f \circ g]$ in terms of various combinatorial coefficients. The book *Vertex Operator Algebras and the Monster* by Frenkel, Lepowsky and Meurman involves a formal calculus including a Faà di Bruno formula for derivations. There are countless related formulas out there, for instance in [24, 25, 27, 31, 51].

2.4. Inversion of power series. In this section we turn to looking at combinatorial inversion of formal power series. We broaden our framework, working with the ring $R := \mathbb{K}[[X_1, \dots, X_N]]$ of formal power series in N determinates with coefficients in a commutative ring \mathbb{K} .

Suppose now that we have a vector $F = (F_1, \dots, F_N)$ in R^N : namely each F_i is an element of R . Each such F may be thought of as a collection $F = (F_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ of elements of the ring \mathbb{K} . The Faà di Bruno formula for power series (8) gives us an associative composition sending a pair (F, G) of elements in R^N to $F \circ G$ in R^N . Moreover, the element $I = (I_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ of R^N defined by $I_{i,\alpha} = 1$ if and only if $\alpha = \mathbf{e}_i$ is an identity for the composition in the sense that $F \circ I = I \circ F = F$. We are interested in the problem of classifying which elements F of R^N are invertible, that is in identifying for which F there exists a G such that $F \circ G = G \circ F = I$, and in obtaining combinatorial expressions for the coefficients of the inverse G in terms of the coefficients of F .

Given an element F of R^N , we write $J(F)$ for its Jacobian matrix, the $N \times N$ matrix given by

$$J(F) := \left(\frac{\partial F_i}{\partial X_j} \right)_{1 \leq i, j \leq N} = \left(\sum_{\alpha \in \mathbb{Z}_{\geq 0}^N} \frac{F_{i,\alpha+\mathbf{e}_j}}{\alpha!} X^\alpha \right)_{1 \leq i, j \leq N}.$$

The Jacobian matrix $J(F)$ is a matrix whose entries take values in the ring $R = \mathbb{K}[[X_1, \dots, X_N]]$. Evaluating the Jacobian matrix at $X = 0$, we obtain the *linear term* of the power series F , the matrix

$$J(F)(0) := (F_{i,\mathbf{e}_j})_{1 \leq i, j \leq N}.$$

The linear term $J(F)(0)$ is a matrix whose entries take values in \mathbb{K} . We will see below that a power series F has a compositional inverse G if and only if its linear term $J(F)(0)$ is invertible as an $N \times N$ matrix with coefficients in the commutative ring \mathbb{K} .

Classifying the coefficients of the power series inverse is more delicate, and we begin with the known case where $F(0) = 0$ and $J(F)(0)$ is the identity matrix, sketching how (9) may be used to identify the derivatives (or equivalently, the coefficients of the power series for) G , much in analogy to how we developed the one-dimensional inversion formula in the introduction

To this end, we introduce some notation. We will consider inverting power series in the following two subsets of R^N :

- Let R_0^N denote the set of power series in R^N such that $F(0) = 0$.
- Let R_1^N denote the set of power series in R^N such that $F(0) = 0$ and $J(F)(0) = I$.

We begin with the simpler case, where F is an element of R_1^N . The i^{th} component of such an F has the form

$$F_i(X) = X_i + \sum_{|\alpha| \geq 2} \frac{F_{i,\alpha}}{\alpha!} X^\alpha.$$

Now suppose G is an inverse of F . Since $F \circ G$ is the identity, we must have $(F \circ G)_{i,\mathbf{e}_i} = 1$ and $(F \circ G)_{i,\alpha} = 0$ whenever $\alpha \neq \mathbf{e}_i$. Using this observation in conjunction with (9) we obtain the relations

$$(10) \quad \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} F_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)} = \begin{cases} 1 & \text{if } \alpha = \mathbf{e}_i, \\ 0 & \text{otherwise,} \end{cases}$$

which hold for all $i \in [N]$ and all multi-indices α . The system of equations (10) gives us a straightforward inductive proof of Theorem 2.3, which states that each $G_{i,\alpha}$ may be written as a sum of weighted trees whose weights are given in terms of the $(F_{i,\alpha})$.

Given a proper labelled tree $\mathcal{T} \in \mathbb{S}_{i,\alpha}$ and a collection $H := (H_{i,\alpha} : 1 \leq i \leq N, \alpha \in \mathbb{Z}_{\geq 0}^N : |\alpha| \geq 2)$ of elements of \mathbb{K} , we define the H -energy of the tree \mathcal{T} to be the element of \mathbb{K} given by

$$(11) \quad \mathcal{E}_H(\mathcal{T}) := \prod_{v \in \mathcal{I}} H_{\tau(v),\mu(v)}$$

The following theorem states that the coefficients of inverse power series may be given in terms of sums of tree energies over proper trees.

Theorem 2.3. *Let F be an element of R_1^N , so that the i^{th} component of F takes the form*

$$F_i(X) = X_i - \sum_{|\alpha| \geq 2} \frac{H_{i,\alpha}}{\alpha!} X^\alpha$$

for some coefficients $(H_{i,\alpha})$. Then the i^{th} -component of the inverse G of F is given by $G_i(X) = X_i + \sum_{|\alpha| \geq 2} \frac{G_{i,\alpha}}{\alpha!} X^\alpha$, where for $|\alpha| \geq 2$,

$$(12) \quad G_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{S}_{i,\alpha}} \mathcal{E}_H(\mathcal{T}).$$

As mentioned in the introduction, Theorem 2.3 is implicit in Corollary 2 of Haiman and Schmitt [26]. In Section 5, we give two proofs of Theorem 2.3, both of which rely on Theorem 2.1 as a key tool.

With a view to discussing the Jacobian conjecture below, we would like to highlight a key property of formal inversion of power series. Since every $F \in R_1^N$ may be written

$$F_i(X) = X_i - \sum_{|\alpha| \geq 2} \frac{H_{i,\alpha}}{\alpha!} X^\alpha,$$

we may canonically associate R_1^N with the set

$$R_1^N := (H_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N : |\alpha| \geq 2).$$

(The choice to work with minus the nonlinear coefficients simplifies several formulas we encounter below.) Under this association, consider the map $\Phi : R_1^N \rightarrow R_1^N$ defined by setting

$$\Phi(H)_{i,\alpha} := - \sum_{\mathcal{T} \in \mathbb{S}_{i,\alpha}} \mathcal{E}_H(\mathcal{T}).$$

The following remark is a consequence of the fact that in the setting of Theorem 2.3, the coefficients of F may be recovered from the coefficients of the inverse G . We will use the mapping Φ below in our discussion of the Jacobian conjecture.

Remark 2.4. The mapping Φ is an involution on R_1^N . That is, $\Phi \circ \Phi$ is the identity map on R_1^N .

Next we consider the inversion of power series with non-identity linear terms, showing that the inverse coefficients may be given in terms of a sum of alternating tree energies over alternating trees.

Indeed, given collections

$$(H_{i,\alpha} : i \in [N], |\alpha| \geq 2) \quad \text{and} \quad (Q_{i,e_j} : i, j \in [N])$$

of elements of a ring \mathbb{K} , and an alternating tree $\mathcal{T} \in \mathbb{A}_{i,\alpha}$, define the (H, Q) energy of \mathcal{T} to be the double product

$$\mathcal{E}_{H,Q}^{\text{alt}}(\mathcal{T}) = \prod_{v \text{ even}} Q_{\tau(v),\mu(v)} \prod_{v \text{ odd}} H_{\tau(v),\mu(v)}$$

where the first product is taken over internal vertices whose graph distance from the root is even, and the latter over internal vertices for which it is odd.

Theorem 2.5. *Let F be an element of R_0^N , so that the i^{th} component of F takes the form*

$$F_i(X) = \sum_{j=1}^N P_{i,e_j} X_j - \sum_{|\alpha| \geq 2} \frac{H_{i,\alpha}}{\alpha!} X^\alpha$$

for some matrix $P = (P_{i,e_j})_{1 \leq i,j \leq N}$ and some coefficients $(H_{i,\alpha})$.

If P is invertible, then F has a compositional power series inverse G . Moreover, the i^{th} component of G is given by

$$G_i(s) = \sum_{j=1}^N Q_{i,e_j} X_j + \sum_{|\alpha| \geq 2} \frac{G_{i,\alpha}}{\alpha!} X^\alpha,$$

where Q is the matrix inverse of P , and for $|\alpha| \geq 2$, the coefficients $G_{i,\alpha}$ are given by

$$(13) \quad G_{i,\alpha} := \sum_{\mathcal{T} \in \mathbb{A}_{i,\alpha}} \mathcal{E}_{H,Q}^{\text{alt}}(\mathcal{T}).$$

We remark that thanks to Theorem 2.5, the mapping Φ defined above Remark 2.4 has an extension to the set of all invertible power series which is still an involution.

Moreover, it is straightforward to see that the more general result, Theorem 2.5 implies the special case Theorem 2.3. Indeed, suppose that P in Theorem 2.5 is the identity matrix. Then the inverse Q of P is also the identity, and it follows that each sum in (13) is supported only on trees such that every internal even vertex has outdegree $\mu(v) = \mathbf{e}_{\tau(v)}$, i.e. all vertices in odd generations have the same type as their parent, and in this case we have $\prod_{v \text{ even}} Q_{\tau(v),\mu(v)} = 1$. By collapsing each edge from an even parent to their odd child, we obtain a canonical bijection between those trees in $\mathbb{A}_{i,\alpha}$ such that every odd vertex has the same type as their parent, and trees in $\mathbb{S}_{i,\alpha}$. That this bijection is energy preserving follows from the fact that $\prod_{v \text{ even}} Q_{\tau(v),\mu(v)} = 1$.

Let us take a moment to discuss other formulas for inversion of power series. As mentioned, Theorem 2.3 appears in implicitly in the work of Haiman and Schmitt, where it is used to express the algebra of power series with composition in terms of the powerful *incidence algebra* framework. A less explicit version of Theorem 2.5 appears in Cheng et al. [10].

Bass, Connell and Wright [5] provide inversion formulas for the entire inverse series G in terms of a sum of tree energies over a collection of trees of arbitrary size. (See also the more general Wright [46].) Their analogues for tree energies are themselves polynomials, taking the form

$$\mathcal{E}^{\text{BCW}}(\mathcal{T}) = \prod_{v \in I} H_{\tau(v),\mu(v)} \prod_{v \in L} x_{\tau(v)},$$

where the leaves are included as monomials in the energy products. There are also a large collection of inversion formulas variously referred to as Lagrange inversion formulas, see Gessel [22] for a comprehensive discussion. Finally, let us mention that tree formalisms have been used to tackle problems in countless other related areas. See for instance Carletti [9], who studies a problem in dynamical systems, as well as work by Gentile and coauthors [6, 21].

2.5. The Jacobian conjecture. We now discuss the celebrated *Jacobian conjecture*. Recall that $R^N = \mathbb{K}[[X_1, \dots, X_N]]^N$ is the set of N -tuples (F_1, \dots, F_N) of formal power series in N variables with coefficients in a ring \mathbb{K} . We say an element $F = (F_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ of R^N is a *polynomial mapping* if only finitely many of the values $(F_{i,\alpha})$ are non-zero. A polynomial mapping is said to be a *polynomial automorphism* if there exists another polynomial mapping G such that $F \circ G = G \circ F = I$. The

Jacobian conjecture is concerned with identifying which polynomial mappings are polynomial automorphisms, and has been the subject of a large deal of research over the last fifty years. We refer the reader to the book [45] for the most comprehensive source on the Jacobian conjecture.

For the sake of concreteness, we work over the complex numbers, setting $\mathbb{K} = \mathbb{C}$, though it is known [45, Proposition 1.12] that the statement over \mathbb{C} is equivalent to statements over general fields of characteristic zero. The Jacobian conjecture asserts that in order for a polynomial mapping F to be a polynomial automorphism, it is sufficient to check the global invertibility of its Jacobian matrix:

Conjecture 2.6 (The Jacobian conjecture). *Let $N \geq 1$ be any integer. Let F be a polynomial mapping such that the Jacobian determinant $j(F) : \mathbb{C}^N \rightarrow \mathbb{C}$ defined by*

$$j(F) := \det(J(F)) = \det_{i,j=1}^N \left(\frac{\partial F_i}{\partial X_j} \right)$$

is equal in value on \mathbb{C}^N to a constant in $\mathbb{C} - \{0\}$. Then F is a polynomial automorphism.

Several partial results exist for the Jacobian conjecture, usually studying the invertibility of polynomial mappings in a certain dimension, of a certain degree, or with certain other structural properties. Most notably, building on work by Moh [35], Wang [42] has shown that the Jacobian conjecture is true for all quadratic mappings.

A large amount of literature on the Jacobian conjecture has been concerned with reducing the complexity of the problem. A foundational idea in this direction is the following reduction due to Bass, Connell and Wright [5].

Theorem 2.7 (Bass, Connell and Wright [5]). *In order to prove Conjecture 2.6 it is sufficient to fix any $\delta \geq 3$, and consider maps of the form $F = I - H$, where H is a homogenous polynomial of degree δ . Moreover, we may assume the Jacobian matrix of H is nilpotent.*

We now make a few remarks about Theorem 2.7. It is straightforward to derive a weaker version of Theorem 2.7, namely that in order to prove the Jacobian conjecture one need only consider maps of the form $F = I - H$, where H contains only degree two terms or higher. To see this, first note that we may assume $F(0) = 0$, since translations are clearly polynomial automorphisms. Now suppose F is a polynomial mapping fixing the origin and let L be the (invertible) linear transformation $L := J(F)(0)$ associated with evaluating the Jacobian matrix of F at zero. Then the map $\tilde{F} := F \circ L^{-1}$ has the desired form $\tilde{F} = I - H$, and \tilde{F} is a polynomial automorphism if and only if F is a polynomial automorphism.

The fact that we may further assume that H is homogenous of some degree greater than or equal to three is less elementary, and uses stabilisation ideas from K-theory, which involves treating maps of high degrees in low dimensions as lower degree maps in higher dimensions [5, Section 2].

When H is homogenous, it is fairly straightforward to show that $j(F)$ constant implies that H is nilpotent. Indeed, if $F = I - H$ has constant determinant, then evaluating this determinant at the origin we see that $j(F)(X) = 1$ for all $X \in \mathbb{C}^N$. Define the dilation map $\varphi_t : \mathbb{C}^N \rightarrow \mathbb{C}^N$ by $\varphi_t(X) := tX$ and set $F_t := \varphi_t^{-1} \circ F \circ \varphi_t$. If $F = I - H$ where H is homogenous of degree $\delta \geq 3$, then

$$F_t = I - t^{\delta-1}H.$$

Since φ_t is a polynomial automorphism, F_t is a polynomial automorphism if and only if F is. Moreover, by the chain rule $j(F_t)(X) = j(F)(X) = 1$ for all $X \in \mathbb{C}^N$. In particular, for all $t > 0$,

$$\begin{aligned}
 (14) \quad 1 &= \det (I - t^{\delta-1} J(H)) \\
 &= \exp (\operatorname{Tr} \log (I - t^{\delta-1} J(H))) \\
 &= 1 + \sum_{k \geq 1} a_k t^{k(\delta-1)} \operatorname{Tr} (J(H)^k),
 \end{aligned}$$

for some constants $(a_k)_{k \geq 1}$. For (14) to hold for every t , we must have that every power of $J(H)$ is traceless. In particular, every eigenvalue of $J(H)$ is zero, so that $J(H)$ is nilpotent. It follows that whenever $F = X - H$ has constant Jacobian determinant and H is nilpotent, $J(H)$ is nilpotent.

There are several further reductions to Theorem 2.7. Drużkowski [13] showed that we may assume further that each component of H is the cube of a linear map (such mappings are known as Drużkowski mappings). De Bondt and Van den Essen [11] showed that we may assume that the Jacobian matrix of H is symmetric. Remarkably, if both of these properties are assumed simultaneously for a mapping — i.e. a mapping F is both symmetric and the cube of a linear map — then this mapping is a polynomial automorphism [12].

Several authors have attempted to tackle the Jacobian conjecture from a combinatorial standpoint, most notably Wright [46–49] and Singer [38–40], and we would like to collect together some of their ideas here in our notation to give a purely combinatorial formulation of the Jacobian conjecture that we hope will motivate future research.

To this end, recall the mapping Φ defined above Remark 2.4, which takes an element of R_1^N , i.e. a collection of coefficients $H := (H_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ and outputs a second collection of coefficients $\Phi(H) := (\Phi(H)_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ in such a way that formal power series with components

$$F_i = X_i - \sum_{|\alpha| \geq 2} \frac{H_{i,\alpha}}{\alpha!} X^\alpha \quad \text{and} \quad G_i = X_i - \sum_{|\alpha| \geq 2} \frac{\Phi(H)_{i,\alpha}}{\alpha!} X^\alpha$$

are compositional inverses of one another. (In particular, Φ is an involution.)

In order to pursue this direction further, it is useful to phrase the nilpotency of the non-linear term H in combinatorial terms. We say the Jacobian matrix of H has index of nilpotency m if $J(H)^m$ is equal to zero on \mathbb{C}^N , and m is the smallest integer with this property. Whenever $J(H)$ is an $N \times N$ nilpotent matrix, the index of nilpotency must be at most N .

If H has index of nilpotency m , then certain energy sums over a type of trees we call *ferns* are equal to zero. A fern of length m is a tree with a designated path of vertices $\{v_0, v_1, \dots, v_m\}$ starting from the root such that each v_{i+1} is a child of v_i , and every vertex not equal to some v_i is a leaf. We call $\{v_0, \dots, v_m\}$ the spine of the fern.

We define $\text{Fern}_{i,\alpha,j}(m)$ to be the set of quadruplets $\mathcal{T} = (V, E, \tau, \phi)$ such that

- The underlying graph (V, E) is a fern of length m with spine $\{v_0, \dots, v_m\}$.
- The labelling function ϕ is a bijection between the leaves $L - \{v_m\}$ not equal to v_m and the set $[\alpha]$.
- The typing function $\tau : V \rightarrow [N]$ satisfies $\tau(v_0) = i$, $\tau(v_m) = j$, and for all $v \in L - \{v_m\}$, if $\phi(v) = (k, a)$ for some a then $\tau(v) = k$.

The following lemma states the nilpotency of H in terms of energy sums over sets of ferns.

Lemma 2.8 (The fern lemma). *Let $H : \mathbb{C}^N \rightarrow \mathbb{C}^N$ be a polynomial mapping containing only terms of degree two and higher. The following are equivalent.*

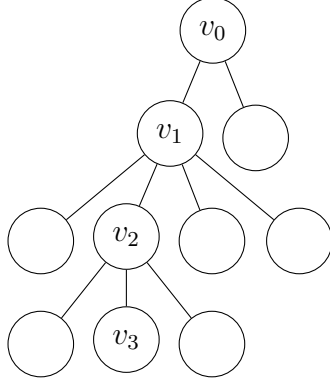


FIGURE 7. A fern of length three.

- (1) The m^{th} power of the Jacobian matrix of H , $J(H)^m$, is equal to the zero matrix on \mathbb{C}^N .
(2) For every i, j, α ,

$$\sum_{\mathcal{T} \in \text{Fern}_{i,\alpha,j}(m)} \mathcal{E}_H(\mathcal{T}) = 0.$$

We remark that an idea similar to Lemma 2.8 appears in the work of Singer [38, Section 3], who studies the special case where H is quadratic. Lemma 2.8 is proved in Section 6.4.

The main result of this section is the following reformulation of the Jacobian conjecture in purely combinatorial terms. This reformulation is based on combining Lemma 2.8, Theorem 2.7, and the definition of the involution Φ described below Theorem 2.3.

Conjecture 2.9 (A combinatorial form of the Jacobian conjecture). *There exists an integer $\delta \geq 3$ with the following property. Suppose $N \geq 1$ and*

$$H := (H_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N, |\alpha| = \delta)$$

is a collection of complex numbers with the property that there exists $m \leq N$ such that for every $i, j \in [N]$ and every $\alpha \in \mathbb{Z}_{\geq 0}^N$,

$$\sum_{\mathcal{T} \in \text{Fern}_{i,\alpha,j}(m)} \mathcal{E}_H(\mathcal{T}) = 0.$$

Now define a second collection of complex numbers

$$\Phi(H)_{i,\alpha} := \sum_{\mathcal{T} \in \mathcal{S}_{i,\alpha}} \mathcal{E}_H(\mathcal{T}).$$

Then only finitely many of the $\Phi(H)_{i,\alpha}$ are non-zero.

To clarify, if Conjecture 2.9 is true for any $\delta \geq 3$, then the Jacobian conjecture holds.

We now make a few further remarks about further assumptions we can make in the setting of Conjecture 2.9 based on the various reductions of the Jacobian conjecture:

- According to Drużkowski's strengthening [13] of the Bass, Connell and Wright reduction, we may assume further that H is the cube of a linear map. This amounts to the existence of a matrix $(L_{i,j})_{1 \leq i,j \leq N}$ such that each $H_{i,\alpha}$ has the form

$$H_{i,\mathbf{e}_a + \mathbf{e}_b + \mathbf{e}_c} = L_{i,a} L_{i,b} L_{i,c}.$$

In fact, in the later work [14], Drużkowski showed that we may further assume $L^2 = 0$.

- If we appeal to the alternative strengthening of the Bass, Connell and Wright result by de Bondt and Van den Essen [11], we may assume that H has a symmetric Jacobian. By the Poincaré lemma, this amounts to the existence of a second collection of constants $(J_\beta : \beta \in \mathbb{Z}_{\geq 0}^N : |\beta| = 4)$ such that

$$H_{i,\alpha} = J_{\alpha+e_i}.$$

See for instance Zhao [52]. Wright has studied analogous statements to Conjecture 2.9 in the symmetric case — though utilising a different power series inversion formula — finding that the symmetry of the Jacobian matrix allows one to reformulate the problem in terms of unrooted trees [48, 49].

It is worth emphasising as mentioned above that the Jacobian conjecture is true for quadratic maps [42]; equivalently, the $\delta = 2$ version of Conjecture 2.9 is true. Unfortunately, to date there are no known combinatorial proofs of this fact, though Singer [38–40] has made some inroads. We hope the formulation we have presented here will yield deeper insight into possible combinatorial approaches to the problem.

3. THE NON-COMMUTATIVE CASE

In this section, we now present the main results of this article concerning power series in N free variables X_1, \dots, X_N such that in general

$$X_i X_j \neq X_j X_i.$$

3.1. Rings of power series in free variables and a free Faà di Bruno formula. Suppose again that \mathbb{K} is a commutative ring, and let $\vec{R} := \mathbb{K}\langle\langle X_1, \dots, X_N \rangle\rangle$ be the ring of formal power series in N free indeterminates with coefficients in \mathbb{K} . Each element f of \vec{R} has the form

$$f(X_1, \dots, X_N) = \sum_{k \geq 0} \sum_{\kappa \in [N]^k} f_\kappa X_\kappa$$

where for $\kappa = (\kappa_1, \dots, \kappa_k) \in [N]^k$, $X_\kappa := X_{\kappa_1} \dots X_{\kappa_k}$, and we use the convention that $[N]^0 := \{\emptyset\}$, where \emptyset is the empty tuple and X_\emptyset refers to the monic constant polynomial in \vec{R} .

We have the following free analogue of the Leibniz rule, namely that if

$$f(X_1, \dots, X_N)g(X_1, \dots, X_N) := \sum_{k \geq 0} \sum_{i \in [N]^k} h_\kappa X_\kappa,$$

then each h_κ is given by

$$h_{\kappa_1, \dots, \kappa_k} = \sum_{0 \leq j \leq k} f_{\kappa_1, \dots, \kappa_j} g_{\kappa_{j+1}, \dots, \kappa_k}.$$

Remark 3.1. In studying the composition of power series in free variables, the object we had in mind were polynomials in several matrix variables with complex coefficients. Indeed, we would like to take a moment to emphasise a few structural aspects of our set up:

- The coefficients f_κ of our power series take values in a commutative ring.
- The variables X_1, \dots, X_N are not assumed to commute with one another.
- The coefficients commute with the variables.

The first point above is where our set up differs from related work by Brouder, Frabetti and Krattenthaler [7] and Anshelevich, Effros and Popa [1]. In both of these works on non-commutative power series, the coefficients themselves take values in a non-commutative algebra, which gives rise to significantly differences which we discuss at the end of this section.

In order to give our free Faà di Bruno formula, we require definitions surrounding *planar trees*. A rooted planar tree is a triplet (V, E, r) where (V, E) is a rooted tree, and r is a function ranking the children of each vertex: so that whenever v is a vertex and the set $\{w : w \text{ is a child of } v\}$ of children of v has cardinality k , r is a bijection between the set of children of v and $\{1, \dots, k\}$. The function r gives rise to a planar embedding of the tree: by drawing the children of a vertices from left to right according to their rankings. Two planar trees (V, E, r) and (V', E', r') are considered isomorphic if there is a rank-preserving graph isomorphism between the underlying vertex sets.

Suppose a rooted planar tree (V, E, r) has k leaves. The ranking induces an ordering on the leaves of the vertex set: so that we label the leaves $\{v_1, \dots, v_k\}$ according to their positions clockwise in the plane; see Figure 8. We refer to v_j as the j^{th} leaf.

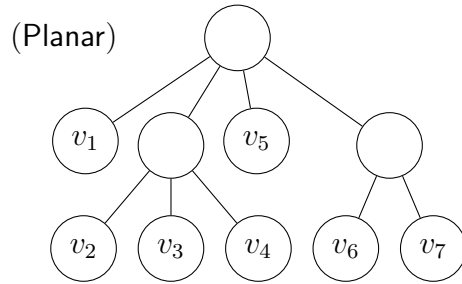


FIGURE 8. A planar tree with seven leaves.

We will be interested in labelled rooted planar trees. Let $\kappa = (\kappa_1, \dots, \kappa_k) \in [N]^k$ be a k -tuple of elements of $[N]$, and let $\vec{\mathbb{T}}_{i,\kappa}$ denote the set of labelled rooted planar trees with root type i and leaf types κ . In other words, $\vec{\mathbb{T}}_{i,\kappa}$ consists of quadruplets $\vec{\mathcal{T}} = (V, E, r, \tau)$ such that

- The triplet (V, E, r) is a rooted planar tree.
- The type function τ is any function $\tau : V \rightarrow [N]$ such that if v_j is the j^{th} leaf of (V, E, r) , then $\tau(v_k) = \kappa_j$.

In analogy to the commutative case, we will consider three subsets of $\vec{\mathbb{T}}_{i,\kappa}$:

- The subset $\vec{\mathbb{S}}_{i,\kappa}$ of $\vec{\mathbb{T}}_{i,\kappa}$ consisting of the proper trees — i.e. those trees where every internal vertex has two or more children.
- The subset $\vec{\mathbb{F}}_{i,\kappa}(m)$ of $\vec{\mathbb{T}}_{i,\kappa}$ consisting of the final trees of length m — i.e. those labelled planar trees where every leaf occurs in generation m .
- The subset $\vec{\mathbb{A}}_{i,\kappa}$ of $\vec{\mathbb{T}}_{i,\kappa}$ consisting of the alternating trees — i.e. those trees where every vertex in an even generation has exactly one child, and everyone vertex in an odd generation is either a leaf or has two or more children.

Our free Faà di Bruno's formula in N non-commutative variables is given in terms of final labelled planar trees, and our inversion formula for power series in N non-commutative variables is given in terms of proper labelled planar trees. See for example Figure 3 above, which depicts an element of $\vec{\mathbb{T}}_{5,\kappa}$, where $\kappa = (1, 2, 1, 3)$.

We now define energies for non-commutative trees. Unsurprisingly perhaps, the energies themselves also depend on the planar structure of the tree. For an internal vertex v of a labelled rooted planar tree with its children listed in ranked order w_1, \dots, w_k , we define the *free outdegree* of v to be the tuple $\vec{\mu}(v) := (\tau(w_1), \dots, \tau(w_k))$. See Figure 9.

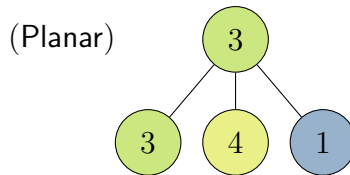


FIGURE 9. An internal vertex v in a labelled planar tree. This vertex v has type $\tau(v) = 3$ and free outdegree $\vec{\mu}(v) = (3, 4, 1)$.

Let \vec{R}^N denote the set of $F = (F_1, \dots, F_N)$, where each component F_i of F is an element of \vec{R} . Suppose $F^* = (F^{(1)}, \dots, F^{(m)})$ is a sequence of elements of \vec{R}^N . We define the non-commutative F^* -energy of a labelled planar tree $\vec{\mathcal{T}} = (V, E, r, \tau)$ in $\vec{\mathbb{F}}_{i,\kappa}(m)$ by

$$\mathcal{E}_{F^*}(\vec{\mathcal{T}}) = \prod_{l=1}^m \prod_{v \in V_{l-1}} F_{\tau(v), \vec{\mu}(v)}^{(l)}.$$

We write \vec{R}_0^N for the subset of \vec{R}^N consisting of power series with no constant term, i.e. F such that $F_i(0, 0, \dots, 0) = 0$ for each $i \in [N]$. We are now equipped to state our free version of Faà di Bruno's formula.

Theorem 3.2. *Suppose that $F^{(1)}, \dots, F^{(m)}$ are elements of \vec{R}_0^N , so that the i^{th} component of $F^{(l)}$ is given by*

$$F_i^{(l)}(X_1, \dots, X_N) := \sum_{k \geq 1} \sum_{\kappa \in [N]^k} F_{i,\kappa} X_\kappa.$$

Then the coefficient of X_κ in the i^{th} component of the composition $F^{(1)} \circ \dots \circ F^{(m)}$ is given by

$$(F^{(1)} \circ \dots \circ F^{(m)})_{i,\kappa} = \sum_{\mathcal{T} \in \vec{\mathbb{F}}_{i,\kappa}(m)} \mathcal{E}_{F^*}(\vec{\mathcal{T}}).$$

Example 3.3. Suppose F and G are elements of \vec{R}^N , and we would like to find the coefficient $(F \circ G)_{2,(3,1,1)}$ of $X_3 X_1^2$ in the 2nd component of the composition $F \circ G$. Then according to Theorem 3.2, $(F \circ G)_{2,(3,1,1)}$ may be written as a sum over trees in $\vec{\mathbb{F}}_{2,(3,1,1)}(2)$. Figure 10 depicts the set of trees in $\vec{\mathbb{F}}_{2,(3,1,1)}$ as the types a, b, c of the internal vertices range over $[N]$.

In particular, if F, G are elements of \vec{R}^N , then

$$\begin{aligned} (F \circ G)_{2,(3,1,1)} &= \sum_{a \in [N]} F_{2,a} G_{a,(3,1,1)} + \sum_{a,b,c \in [N]} F_{2,(a,b,c)} G_{a,3} G_{b,1} G_{c,1} \\ &+ \sum_{a,b \in [N]} G_{2,(a,b)} F_{a,(3,1)} F_{b,1} + \sum_{a,b \in [N]} F_{2,(a,b)} G_{a,3} G_{b,(1,1)}. \end{aligned}$$

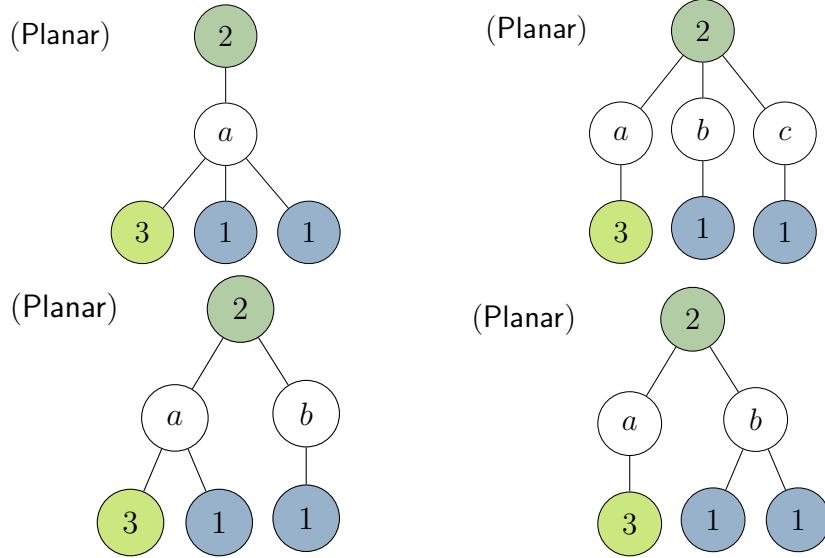


FIGURE 10. The four possible planar topologies for trees in $\vec{\mathbb{F}}_{i,\kappa}(2)$, where $i = 2$ and $\kappa = (3, 1, 1)$. The types a, b, c may take any value in $[N]$.

3.2. Inversion of power series in free variables. We now turn to compositional inversion of elements of \vec{R}^N . Consider first of all the compositional identity element of \vec{R}^N — the element I of \vec{R}^N whose i^{th} component is given by X_i . Namely given an element F of \vec{R}^N , we are interested in identifying whether F has a compositional inverse G satisfying $F \circ G = G \circ F = X$, and in identifying the coefficients of the inverse.

Like in the commutative case, we begin by considering the case where the linear term

$$J(F)(0) := (F_{i,(j)})_{1 \leq i, j \leq N}$$

of F is the identity matrix. Indeed, given a free N -dimensional power series whose i^{th} component has the form $H(X_1, \dots, X_N) := \sum_{k \geq 2} \sum_{\kappa \in [N]^k} H_{i,\kappa} X_\kappa$, define the planar H -energy of a planar tree \vec{T} by

$$\mathcal{E}_H(\vec{T}) := \prod_{v \in I} H_{\tau(v), \vec{\mu}(v)}.$$

Theorem 3.4. *Let F be an element of \vec{R}^N whose i^{th} component has the form*

$$F_i(X_1, \dots, X_N) = X_i - \sum_{k \geq 2} \sum_{\kappa \in [N]^k} H_{i,\kappa} X_\kappa.$$

Then there exists an element G of \vec{R}^N such that G is a compositional inverse of F in that $F \circ G = G \circ F = X$. Moreover, G has the form

$$G_i(X_1, \dots, X_N) = X_i + \sum_{k \geq 2} \sum_{\kappa \in [N]^k} G_{i,\kappa} X_\kappa,$$

where the coefficient $G_{i,\kappa}$ is given in terms of the sum

$$G_{i,\kappa} = \sum_{\vec{T} \in \vec{\mathcal{S}}_{i,\kappa}} \mathcal{E}_H(\vec{T}).$$

Where Theorem 3.4 was concerned with inverting power series with identical linear terms, we now turn to the case where the linear term is non-identical, which we find parallels the commutative case. Given a matrix (Q_{i,e_j}) , and a collection of elements $(H_{i,\kappa} : i \in [N], \kappa \in [N]^k, k \geq 2)$, we define the planar (H, Q) energy of an alternating planar tree $\vec{\mathcal{T}}$ by

$$\mathcal{E}_{H,Q}(\vec{\mathcal{T}}) := \prod_{v \text{ even}} Q_{\tau(v), \vec{\mu}(v)} \prod_{v \text{ odd}} H_{\tau(v), \vec{\mu}(v)}.$$

Theorem 3.5. *Let F be an element of \vec{R}^N whose i^{th} component has the form*

$$F_i(X_1, \dots, X_N) = \sum_{j=1}^N P_{i,j} X_j - \sum_{k \geq 2} \sum_{\kappa \in [N]^k} H_{i,\kappa} X_\kappa.$$

Then if P is an invertible matrix, then G has a composition inverse. Moreover, if Q is the matrix inverse of P , then the i^{th} component of G has the form

$$G_i(X_1, \dots, X_N) = \sum_{j=1}^N Q_{i,j} X_j + \sum_{k \geq 2} \sum_{\kappa \in [N]^k} G_{i,\kappa} X_\kappa$$

where for $k \geq 2$ and $\kappa \in [N]^k$, the coefficients of G are given by

$$G_{i,\kappa} := \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{K}}_{i,\kappa}} \mathcal{E}_{H,Q}^{\text{alt}}(\vec{\mathcal{T}}).$$

We conclude this section by discussing further the aforementioned work of Brouder, Frabetti and Krattenthaler [7] and Anshelevich, Effros and Popa [1], who study the composition and inversion of non-commutative power series in non-commuting variables through the apparatus of incidence algebras. The former paper [7] is concerned with a single-variable and the latter [1] with many variables. As we said above, the set ups of these works differ from ours in that they assume the coefficients themselves of the series take values in a non-commutative algebra — though the coefficients are assumed to commute with the variables. The non-commutativity of the coefficients gives rise to several interesting structural differences. Most notably, the ‘composition’ of power series need not be associative, so that when the coefficients are non-commutative, we have

$$(F \circ G) \circ H \neq F \circ (G \circ H)$$

in general. (To use the example supplied in [1], suppose $N = 1$ and consider the power series $F(X_1) = X_1^2$, $G(X_1) = bX_1$ and $H(X_1) = aX_1$, where the coefficients a and b do not commute.) An interesting further consequence of having non-commutative coefficients is that right and left compositional inverses of a power series need not coincide [1, Corollary 11]. In fact, Anshelevich et al. [1] are hesitant to use the word composition for these reasons, preferring the term *power series substitution*. In summary, Theorem 3.4 is connected with [1, Theorem 8], with various reductions and symmetries arising in our case thanks to the commutativity of the power series coefficients.

More broadly, the study of polynomials and power series in non-commutative variables and with non-commuting coefficients dates back at least to the start of the 20th century (see e.g. Ore [36]). To this day this remains an active research area, with Faà di Bruno’s formula appearing in a multitude of non-commutative settings. This in contrast to our set up — see Remark 3.1. On this front we also mention the recent work of Frabetti and Shestakov [18], who among other things obtain a generalisation of the Lagrange-inversion formula for series with non-commutative coefficients.

3.3. A non-commutative Jacobian conjecture. We conclude the discussion of power series in free variables by exploring possible non-commutative analogues to the Jacobian conjecture.

Recall that in the commutative setting, the Jacobian conjecture asserts that in order for polynomial mapping $F \in R^N$ with coefficients in \mathbb{C} to have a polynomial compositional inverse, it is sufficient to check that the determinant of the Jacobian matrix is equal to a non-zero complex constant. When we attempt to develop a parallel statement in the case where the variables X_1, \dots, X_N do not commute, we find that certain structural aspects of the problem are different.

Indeed, like in the commutative case we say an element $F = (F_{i,\kappa} : i \in [N], \kappa \in [N]^k, k \geq 0)$ of \vec{R}^N is a polynomial mapping if only finitely many of the coefficients $(F_{i,\kappa})$ are non-zero. With a view to defining the Jacobian matrix associated with a polynomial mapping in free variables, we first need a notion of differentiation. To this end, given a multi-index $\kappa = (\kappa_1, \dots, \kappa_k)$, define

$$\frac{\partial}{\partial X_j} X_{\kappa_1} \dots X_{\kappa_k} := \sum_{l=1}^k \mathbb{1}_{\kappa_l=j} X_{i_1} \dots X_{i_{l-1}} X_{i_{l+1}} \dots X_{i_k}.$$

For example,

$$\frac{\partial}{\partial X_2} X_2 X_2 X_1 X_2 X_4 = 2X_2 X_1 X_2 X_4 + X_2 X_2 X_1 X_4.$$

A similar definition for differentiating non-commutative monomials appears in Rota, Sagan and Stein [37], where they refer the operation as the *Hausdorff derivative*. By linearity the operator $\frac{\partial}{\partial X_j}$ extends to the ring of formal power series in N free indeterminates.

Again, for the sake of concreteness, let the coefficient ring $\mathbb{K} = \mathbb{C}$, and suppose we have a power series $F \in \vec{R}^N$ in N free variables with coefficients in \mathbb{C} . We can then define its Jacobian matrix to be the $N \times N$ matrix $\left(\frac{\partial F_i}{\partial X_j} \right)_{1 \leq i, j \leq N}$ with coefficients in the (non-commutative) ring \vec{R} . Here is the point at which the non-commutative case diverges with the commutative case: the invertibility of this Jacobian matrix at a point (X_1, \dots, X_N) may not be expressed straightforwardly in terms of a determinant because the “determinant” of a matrix with non-commuting entries is ill-defined.

By Theorem 3.5, we know that a polynomial mapping F in \vec{R}^N is guaranteed to have a compositional inverse G in \vec{R}^N if and only if its Jacobian matrix is invertible at the origin. However, it is by no means clear to the authors under what conditions the inverse G is itself a polynomial. This discussion leads us pose the following question.

Question 3.6. Let F be a polynomial mapping in \vec{R}^N in N non-commuting variables whose Jacobian matrix is invertible at the origin so that F has a power series inverse G . When is the inverse free power series G itself a polynomial?

One possible way forward in terms of formulating a statement analogous to the Jacobian conjecture in the free case would be in terms of *quasi-determinants*, which serve as analogues to the classical determinant for matrices with non-commutative entries. In their present form, quasi-determinants were introduced by Gel'fand and Retakh in [20]. We refer the reader to Gel'fand, Gel'fand, Retakh and Wilson [19] for a survey.

4. A FEW APPLICATIONS

Before proving our main results in Sections 5 and 6, in this section we discuss a few applications of Faà di Bruno's formula and our inversion formula to various areas of mathematics, including the Stirling numbers, the Hermite polynomials, the cumulants of random variables, reciprocation of power series, enumeration of trees, and the genealogy of Galton-Watson trees.

4.1. Stirling numbers. Let $B(k, j)$ denote the number of set partitions the set $\{1, \dots, k\}$ into j blocks. (So that in particular, $B(k, j)$ is zero when $j > k$.) The numbers $B(k, j)$ are known as the Stirling numbers of the second kind. We may use the one-dimensional Faà di Bruno to obtain a quick derivation of the joint generating function of the sequence $(B(k, j))_{k, j \geq 1}$. Indeed, consider the function $h(x) := \exp(ae^x - a)$, which we may regard as the composition $h = f \circ g$ of the functions

$$f(x) = e^{ax} \quad \text{and} \quad g(x) := e^x - 1.$$

By using the Faà di Bruno formula, we immediately see that

$$(15) \quad \frac{d^k}{dx^k} \exp(ae^x - 1) \Big|_{x=0} = \sum_{\pi \in \mathcal{P}_k} a^{\#\pi}.$$

In particular, by summing (15) over k , we have

$$\exp(ae^{bx} - 1) = \sum_{k=0}^{\infty} \frac{b^k}{k!} \sum_{\pi \in \mathcal{P}_k} a^{\#\pi} = \sum_{k, j \geq 0} B(k, j) a^j \frac{b^k}{k!},$$

which gives a joint generating function of the Stirling numbers of the second kind $(B(k, j))_{k, j \geq 1}$.

4.2. Hermite polynomials. The Hermite polynomials $(H_n(x))_{n \geq 0}$ are a sequence of classical orthogonal polynomials arising in numerous applications across mathematics. The Hermite polynomials satisfy the orthogonality property

$$\int_{-\infty}^{\infty} H_i(s) H_j(s) \frac{e^{-s^2/2} ds}{\sqrt{2\pi}} = \delta_{i,j} i!,$$

and may be defined through the higher derivatives

$$H_k(x) := (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}.$$

We now use the one dimensional Faà di Bruno formula to obtain a new representation for the Hermite polynomials in terms of partitions of $\{1, \dots, k\}$ into sets of size at most 2. Indeed, setting $f(x) = e^x$ and $g(x) = -x^2/2$ in the Faà di Bruno formula we obtain

$$(16) \quad H_k(x) = \sum_{\pi \in \mathcal{P}_k} (-1)^{n+\#\pi} \prod_{\Gamma \in \pi} (\mathbb{1}_{\#\Gamma=1} x + \mathbb{1}_{\#\Gamma=2}).$$

The formula (16) states that the Hermite polynomials have a representation as a *matching polynomial*; see for instance Godsil [23].

4.3. Cumulants. Let $Y = (Y_1, \dots, Y_N)$ be an \mathbb{R}^N -valued random variable, and consider the moment and cumulant generating functions

$$M(s_1, \dots, s_N) := \mathbb{E}[e^{s_1 Y_1 + \dots + s_N Y_N}] \quad \text{and} \quad K(s_1, \dots, s_N) := \log M(s_1, \dots, s_N).$$

For simplicity, suppose that Y has moments of all orders, so that M and K may be written as a power series

$$M(s) = 1 + \sum_{\alpha \neq 0} \frac{m_\alpha}{\alpha!} s^\alpha \quad \text{and} \quad K(s) = \sum_{\alpha \neq 0} \frac{\kappa_\alpha}{\alpha!} s^\alpha.$$

Of course, each coefficient m_α is equal to the moment $\mathbb{E}[Y^\alpha]$ of Y . The coefficients (κ_α) of K on the other hand are known as the *cumulants* of X . Though the moments are more natural than the cumulants, there are many occasions in probability where it is more convenient to work with the cumulants. Take for instance the fact that the standard one-dimensional Gaussian distribution is characterised by the

succinct property that it is the unique random variable with cumulants $\kappa_n := \mathbb{1}_{n=2}$. We now address the problem of computing the cumulants in terms of the moments, and vice versa.

By the special case $m = 2$ of Theorem 2.1, and the following Remark 2.2, if $f : \mathbb{R} \rightarrow \mathbb{R}$ and $G : \mathbb{R}^N \rightarrow \mathbb{R}$, the derivatives of $(f \circ G) : \mathbb{R}^N \rightarrow \mathbb{R}$ are given by

$$(17) \quad D^\alpha[f \circ G](s) = \sum_{\mathcal{T} \in \mathbb{F}_{1,\alpha}(2)} D^{\mu(v_0)}[f] \circ G(s) \prod_{v \in V_1} D^{\mu(v)}[G](s).$$

Now using the simple facts

$$\left(\frac{d}{dx}\right)^j \log(x) \Big|_{x=1} = (-1)^{j-1} (j-1)! \quad \text{and} \quad \left(\frac{d}{dx}\right)^j \exp(x) \Big|_{x=0} = 1,$$

in conjunction with respectively setting $(f, G) = (\log, M)$ and (\exp, K) and $s = 0$ in (17), we obtain the following formulas allowing us to pass between the multivariate moments and cumulants:

$$(18) \quad \kappa_\alpha := \sum_{\pi \in \mathcal{P}_{[\alpha]}} (-1)^{\#\pi-1} (\#\pi - 1)! \prod_{\Gamma \in \pi} m_{\#\Gamma} \quad \text{and} \quad m_\alpha := \sum_{\pi \in \mathcal{P}_{[\alpha]}} \prod_{\Gamma \in \pi} \kappa_{\#\Gamma}.$$

The derivation of (18) using Faà di Bruno's formula is a slick alternative to the direct combinatorial approach often used in the literature (see e.g. Exercise 4 of Section 1.1 in Mingo and Speicher [34]).

4.4. Reciprocals of power series. Suppose we have a formal power series

$$f(X) = 1 + \sum_{|\alpha| \neq 0} \frac{f_\alpha}{\alpha!} X^\alpha.$$

It is straightforward to show using (17) that $\frac{1}{f(X)} := 1 + \sum_{\alpha \neq 0} \frac{h_\alpha}{\alpha!} X^\alpha$ where for $\alpha \neq 0$,

$$h_\alpha = \sum_{\pi \in \mathcal{P}_{[\alpha]}} (-1)^{\#\pi} \#\pi! \prod_{\Gamma \in \pi} f_{\#\Gamma}.$$

4.5. Enumeration of trees. For positive integers $k \geq 1$, recall the set \mathbb{S}_k of rooted proper trees with leaves in bijection with $[k]$ introduced in Section 1. Theorem 1.1 supplies us with a quick way to enumerate \mathbb{S}_k .

Indeed, in the context of Theorem 1.1 consider the power series h such that $\mathcal{E}_h(\mathcal{T}) = 1$ for every tree \mathcal{T} . We see that the unique power series with no constant or linear term giving this property is $h(X) = e^X - 1 - X$. It is then an immediate consequence of Theorem 1.1 that

$$\#\mathbb{S}_k = \text{Coefficient of } \frac{X^k}{k!} \text{ in power series inverse of } 2X + 1 - e^X.$$

In other words, the exponential generating function of \mathbb{S}_k is given by the functional inverse of $f(X) = 2X + 1 - e^X$. This result appears in Stanley [41].

We may now treat the energy functional $\mathcal{E}_h(\mathcal{T})$ more generally as a partition function counting the number of vertices of certain degrees with trees of \mathbb{S}_k . For instance, suppose we want to count the size of the subset $\mathbb{S}_k^{\text{even}}$ of \mathbb{S}_k consisting only of trees in which vertices have an even number of children. Then we look at the power series

$$h(X) = \sum_{k \geq 2} \frac{\mathbb{1}_{k \text{ even}}}{k!} X^k = \cosh(X) - 1,$$

As a result, we have

$$\#\mathbb{S}_k^{\text{even}} = \text{Coefficient of } \frac{X^k}{k!} \text{ in functional inverse of } 2X + 1 - \cosh(X).$$

4.6. The genealogical structure of Galton-Watson trees. Faà di Bruno's formula is required to understand the genealogical structure of Galton-Watson trees, which are stochastic processes modelling population growth which we now define. Before setting this up, we note that whenever f, g are smooth functions with non-negative derivatives of all orders, and $D^n(f \circ g)(x) > 0$, the function $P : \mathcal{P}_n \rightarrow \mathbb{R}$ given by

$$(19) \quad P(\pi) := \frac{D^{\#\pi}[f](g(x)) \prod_{\Gamma \in \pi} D^{\#\Gamma}[g](x)}{D^n[f \circ g](x)}$$

defines a probability measure on the set \mathcal{P}_n of partitions of $\{1, \dots, n\}$. It turns out that random partitions with probability laws of the form (19) occur naturally in the study of Galton-Watson trees.

Let $(p_i)_{i \geq 0}$ be a collection of non-negative reals satisfying $\sum_{i \geq 0} p_i = 1$. The (continuous-time) Galton-Watson tree with offspring distribution $(p_i)_{i \geq 0}$ is the stochastic process defined as follows. We start with a single particle at time zero who lives for a random period of time of length τ , where τ is standard exponential so that

$$\mathbb{P}(\tau > t) = e^{-t}.$$

Upon death this particle is replaced by a random number of children L , where

$$\mathbb{P}(L = i) = p_i.$$

Each one of these children then independently repeats the behaviour of their parent: living for a standard exponential amount of time and then being replaced by a random number of children upon death with probabilities $(p_i)_{i \geq 0}$. We make the simple observation that for every pair of times $t < T$, that each individual living at some time T is descended from a unique ancestor living in the time t population.

Galton-Watson trees are best studied through their generating functions. For $s \in [0, 1]$ let $f(s) := \sum_{i \geq 0} p_i s^i$, and suppose that N_t is the number of particles in the process at time t . Then the process generating function $F_t(s) := \mathbb{E}[s^{N_t}]$ satisfies the partial differential equation

$$\frac{\partial F}{\partial t} = f(F) - F, \quad F_0(s) = s.$$

(See e.g. Athreya and Ney [4].) Moreover, F_t is a semigroup, in that $F_{t_1} \circ F_{t_2} = F_{t_1+t_2}$.

Fix $T > 0$, and consider now conditioning on the event that $\{N_T \geq k\}$ and picking k distinct individuals u_1, \dots, u_k uniformly from the population at time T . We may create a random partition π_t of $\{1, \dots, k\}$ by declaring i and j to be in the same block of π_t if individuals u_i and u_j are descended from the same ancestor in the time t -population. According to Theorem 3.1 of [29], the law of π_t is given by the integral formula

$$(20) \quad \mathbb{P}(\pi_t = \pi) = \int_0^1 \Lambda_T^k(s) \frac{F_t^{\#\pi}(F_{T-t}(s)) \prod_{\Gamma \in \pi} F_{T-t}^{\#\Gamma}(s)}{F_T^k(s)} ds,$$

where $F_t^j(s) = \left(\frac{\partial}{\partial s}\right)^j F_t(s)$, and

$$\Lambda_T^k(s) ds := \frac{(1-s)^{k-1} F_T^k(s)}{(k-1)! \mathbb{P}(N_T \geq k)} ds$$

is a probability measure for $s \in [0, 1]$. We remark that the fact that the right-hand side of (20) constitutes a probability measure on \mathcal{P}_n is a consequence of the fact that $\Lambda_T^k(s)$ is a probability measure on $[0, 1]$,

and that the internal quotients are themselves probability measures by virtue of setting $f = F_t, g = F_{T-t}$ in the representation (19) and using the semigroup property $F_t \circ F_{T-t} = F_T$.

We refer the reader to other appearances of Faà di Bruno's formula on work in the genealogical structure of branching processes: in the setting of Galton-Watson trees by Vatutin and co-authors [32,50], as well as continuous-state branching processes [30].

That completes the section on applications. In the next section we work towards proving our generalisation of Faà di Bruno's formula in the commutative case, Theorem 2.1.

5. PROOF OF THE MULTIVARIATE FAÀ DI BRUNO FORMULA

In this section we work towards proving our main result in the commutative case, Theorem 2.1, which is a generalisation of Faà di Bruno's formula for a chain composition of functions in N commutative variables. In Section 5.1 we begin by making some first observations. In Section 5.2 we prove Theorem 2.1 in the special case $m = 2$, and in the following Section 5.3 the result is proved for general m .

First, a word on notation: we will use the convention that in any expression, square brackets precede composition, which proceeds multiplication. For instance

$$D[f] \circ g \ h \text{ refers to } ((D[f]) \circ g) \ h.$$

Our proof is based on a doubly inductive argument, first we prove the special case $m = 2$ via an induction argument on α . Then we prove the general case by induction on m . In the next section we investigate how the differentiation and projection operators interact with multiplication and composition of functions.

5.1. Differentiation and combining functions. Let \mathcal{A} be the set of smooth scalar-valued functions $f : \mathbb{R}^N \rightarrow \mathbb{R}$, and \mathcal{A}^N be the set of smooth vector-valued functions $F : \mathbb{R}^N \rightarrow \mathbb{R}^N$. Finally, let $\pi_j : \mathcal{A}^N \rightarrow \mathcal{A}$ be the projection of a function $F = (F_1, \dots, F_N)$ onto its j^{th} component F_j .

We have several ways of combining functions in \mathcal{A} or \mathcal{A}^N to form a new function in one of these sets. First of all, we have addition. Namely given any pair of functions f and g in \mathcal{A} (resp. \mathcal{A}^N), we may define a new function $f + g$ in \mathcal{A} (resp. \mathcal{A}^N) by setting $(f + g)(s) := f(s) + g(s)$. Secondly, we have multiplication. Given any pair of functions f and g both in \mathcal{A} , we may define a function fg in \mathcal{A} defined by setting $fg(s) := f(s)g(s)$. Finally, we have composition. Namely, whenever f is an element of \mathcal{A} (resp. \mathcal{A}^N) and g is an element of \mathcal{A}^N , we may define a function $f \circ g$ in \mathcal{A} (resp. \mathcal{A}^N) by setting $f \circ g(x) := f(g(x))$.

We now have a look at how first order differentiation $D^{\mathbf{e}_i} := \frac{\partial}{\partial x_i}$ interacts with each of these ways of combining functions. First of all, we have linearity. Namely, for f_1, \dots, f_p in \mathcal{A} we have

$$(21) \quad D^{\mathbf{e}_i} \left[\sum_{r=1}^p f_r \right] = \sum_{r=1}^p D^{\mathbf{e}_i} [f_r].$$

As for multiplication, we have the product rule. For any f, g in \mathcal{A} , we have $D^{\mathbf{e}_i} [fg] = D^{\mathbf{e}_i} [f]g + fD^{\mathbf{e}_i} [g]$ and more generally, it follows by induction that for f_1, \dots, f_p in \mathcal{A} ,

$$(22) \quad D^{\mathbf{e}_i} \left[\prod_{r=1}^p f_r \right] = \sum_{r=1}^p D^{\mathbf{e}_i} [f_r] \prod_{s \neq r} f_s.$$

Finally, differentiation interacts with composition according to the chain rule. Namely whenever $f \in \mathcal{A}$ and G is an element of \mathcal{A}^N , the first order derivatives of their composition satisfy

$$(23) \quad D^{\mathbf{e}_i}[f \circ G] = \sum_{j=1}^N (D^{\mathbf{e}_j}[f] \circ G) D_j^{\mathbf{e}_i}[G].$$

We will see that these three rules (21), (22) and (23) are the building blocks in a proof of Theorem 2.1. In the next section we start by proving the case $m = 2$ of Theorem 2.1.

5.2. The $m = 2$ case: derivatives of $F \circ G$. On our way to proving Theorem 2.1, first we prove the special case $m = 2$, which we state as a lemma.

Recall that $\mathbb{F}_{i,\alpha}(2)$ is the set of labelled trees whose root v_0 has type i , the leaves are in bijection with $[\alpha]$, and every leaf of the tree lies in the second generation.

Lemma 5.1. *For smooth functions $F, G : \mathbb{R}^N \rightarrow \mathbb{R}^N$, we have*

$$(24) \quad D_i^\alpha[F \circ G] = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} \mathcal{E}_{F,G}(\mathcal{T}),$$

where

$$\mathcal{E}_{F,G}(\mathcal{T}) := \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G].$$

Proof. We proceed by induction on the multi-index α .

The special case where $\alpha = \mathbf{e}_j$ follows immediately from the chain rule. Indeed, a tree with two generations and one leaf is simply a line with three vertices. The root has type i and the sole leaf has type j , leaving N choices of labelling for the internal vertex.

We now fix $\beta \in \mathbb{Z}_{\geq 0}^N$, and show that if (24) holds for $\alpha = \beta$, then it holds for $\alpha = \beta + \mathbf{e}_j$. Indeed, using the inductive hypothesis in the second equality below, and the linearity of differentiation in the third, we have

$$(25) \quad \begin{aligned} D_i^{\beta + \mathbf{e}_j}[F \circ G] &= D^{\mathbf{e}_j} \left[D_i^\beta[F \circ G] \right] \\ &= D^{\mathbf{e}_j} \left[\sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} \mathcal{E}_{F,G}(\mathcal{T}) \right] \\ &= \sum_{\mathcal{T} \in \mathbb{F}^{[\beta]}(2)} D^{\mathbf{e}_j} [\mathcal{E}_{F,G}(\mathcal{T})]. \end{aligned}$$

We now expand the summands using the product rule and the chain rule (23). Indeed, using the product rule to obtain the second equality below, and the chain rule to obtain the third, we have

$$(26) \quad \begin{aligned} D^{\mathbf{e}_j} [\mathcal{E}_{F,G}(\mathcal{T})] &:= D^{\mathbf{e}_j} \left[\left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G] \right] \\ &= D^{\mathbf{e}_j} \left[\left(D_i^{\mu(v_0)}[F] \circ G \right) \right] \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G] \\ &+ \sum_{v \in V_1} \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{w \in V_1} G_{\tau(w), \mu(w) + \mathbb{1}_{v=w} \mathbf{e}_j}. \end{aligned}$$

Now note that by the chain rule, $D^{\mathbf{e}_j} \left[\left(D_i^{\mu(v_0)}[F] \circ G \right) \right] = \sum_{k=1}^N \left(D_i^{\mu(v_0) + \mathbf{e}_k}[F] \circ G \right) D_k^{\mathbf{e}_j}[G]$, so that

$$(27) \quad D^{\mathbf{e}_j} [\mathcal{E}_{F,G}(\mathcal{T})] = \sum_{k=1}^N A_k(\mathcal{T}) + \sum_{v \in V_1} B_v(\mathcal{T}),$$

where

$$A_k(\mathcal{T}) := \left(D_i^{\mu(v_0) + \mathbf{e}_k}[F] \circ G \right) D_k^{\mathbf{e}_j}[G] \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G]$$

and

$$B_v(\mathcal{T}) := \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{w \in V_1} G_{\tau(w), \mu(w) + \mathbb{1}_{v=w} \mathbf{e}_j}.$$

First, we observe that for each $k \in [N]$, $\mathcal{A}_k(\mathcal{T}) = \mathcal{E}_{F,G}(\mathcal{T}_k)$, where \mathcal{T}_k is the tree in $\mathbb{F}_{i, \alpha + \mathbf{e}_j}(2)$ obtained from \mathcal{T} by adding a vertex of type k to V_1 , and letting this vertex have the sole leaf $(j, \beta_j + 1)$ as a child. (Note that $\{(j, \beta_j + 1)\} = [\beta + \mathbf{e}_j] - [\beta]$.)

Next, we note that $B_v(\mathcal{T}) = \mathcal{E}_{F,G}(\mathcal{T}_v)$ for each $v \in V_1$, where \mathcal{T}_v is the tree in $\mathbb{F}_{i, \alpha + \mathbf{e}_j}(2)$ obtained from \mathcal{T} by making the new leaf $(j, \beta_j + 1)$ a child of v .

We now note that every tree \mathcal{T}' in $\mathbb{F}_{i, \beta + \mathbf{e}_j}(2)$ is obtained in this way by exactly one tree in $\mathbb{F}_{i, \beta}(2)$, that is, for each $\mathcal{T}' \in \mathbb{F}_{i, \beta + \mathbf{e}_j}(2)$ there is a unique \mathcal{T} in $\mathbb{F}_{i, \beta}(2)$ such that either $\mathcal{T}' = \mathcal{T}_k$ or $\mathcal{T}' = \mathcal{T}_v$.

Using (25), (27) and this observation in the second equality below we obtain

$$\begin{aligned} D_i^{\beta + \mathbf{e}_j}[F \circ G] &= \sum_{\mathcal{T} \in \mathbb{F}_{i, \beta}(2)} \left\{ \sum_{k=1}^N A_k(\mathcal{T}) + \sum_{v \in V_1} B_v(\mathcal{T}) \right\} \\ &= \sum_{\mathcal{T}' \in \mathbb{F}_{i, \beta + \mathbf{e}_j}(2)} \mathcal{E}_{F,G}(\mathcal{T}'), \end{aligned}$$

which proves the lemma. □

5.3. The general case.

Proof of Theorem 2.1. We now prove Theorem 2.1 by induction on m .

Now suppose the result holds for any m -fold composition of functions. We now show it holds for $(m+1)$ -fold compositions. Indeed, let $F^{(1)} \circ \dots \circ F^{(m+1)}$ be the composition of $m+1$ smooth functions from $\mathbb{R}^N \rightarrow \mathbb{R}^N$. By the special case $m=2$, with $F = F^{(1)}$ and $G = F^{(2)} \circ \dots \circ F^{(m+1)}$, we have

$$(28) \quad D_i^\alpha [F^{(1)} \circ \dots \circ F^{(m+1)}](s) = \sum_{\mathcal{T} \in \mathbb{F}_{i, \alpha}(2)} \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G].$$

We may now use the inductive hypothesis to compute $D_{\tau(v)}^{\mu(v)}[G]$. For each $v \in V_1$, by the inductive hypothesis we have

$$D_{\tau(v)}^{\mu(v)}[G] = \sum_{\mathcal{T} \in \mathbb{F}_{\tau(v), \mu(v)}(2)} \mathcal{E}_{G^*}(\mathcal{T}),$$

where G^* is the m -tuple $(F^{(2)}, \dots, F^{(m+1)})$. In particular, expanding (28) we have

$$\begin{aligned} D_i^\alpha[F^{(1)} \circ \dots \circ F^{(m+1)}](s) &= \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} \sum_{\mathcal{T}_v \in \mathbb{F}_{\mu(v), \tau(v)}(m)} \mathcal{E}_{G^*}(\mathcal{T}_v) \\ &= \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} \sum_{\mathcal{T}_v \in \mathbb{F}_{\mu(v), \tau(v)}(m): v \in V_1} \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} \mathcal{E}_{G^*}(\mathcal{T}_v). \end{aligned}$$

We now note that each term in the sum — that is, each combination of $\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)$ and a set of m -trees $(\mathcal{T}_v : v \in V_1)$ — gives rise to a tree \mathcal{T}' in $\mathbb{F}_{i,\alpha}(m+1)$ by ‘glueing’ the tree \mathcal{T}_v to the vertex v in V_1 , and that

$$\mathcal{E}_{F^*}(\mathcal{T}) = \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} \mathcal{E}_{G^*}(\mathcal{T}_v).$$

We note that each \mathcal{T}' in $\mathbb{F}_{i,\alpha}(m+1)$ is obtained uniquely in this way. It follows that

$$D_i^\alpha[F^{(1)} \circ \dots \circ F^{(m+1)}](s) = \sum_{\mathcal{T}' \in \mathbb{F}_{i,\alpha}(m+1)} \mathcal{E}_{F^*}(\mathcal{T}'),$$

proving the result. □

6. PROOF OF THE POWER SERIES INVERSION FORMULAS

This section is dedicated towards proving the power series inversion formulas given in Theorem 2.3 and Theorem 2.5, as well as to proving Lemma 2.8, which was stated in our discussion of the Jacobian conjecture in Section 2.5.

We begin in the next section by making some first remarks about the ring of power series in N commuting variables.

6.1. The ring of formal power series. Let \mathbb{K} be a commutative ring, and consider the ring

$$R := \mathbb{K}[[X_1, \dots, X_N]]$$

of formal power series in N commutative indeterminates with coefficients in \mathbb{K} . Namely, each element f of R is simply a collection $(f_\alpha : \alpha \in \mathbb{Z}_{\geq 0}^N)$ of elements of the underlying ring \mathbb{K} , though we think of f as the formal expression

$$f(X) := \sum_{\alpha \in \mathbb{Z}_{\geq 0}^N} \frac{f_\alpha}{\alpha!} X^\alpha.$$

There are naturally defined notions of addition and multiplication of elements of R . Namely, $(f+g)_\alpha := f_\alpha + g_\alpha$, and by the Leibniz rule (7), the α^{th} coefficient of the product fg is given by

$$(fg)_\alpha = \sum_{S \sqcup T = \alpha} f_{\#S} g_{\#T}.$$

Now define the set R^N of objects of the form $F = (F_i)_{i=1, \dots, N}$, where each F_i is an element of R . In other words, R^N is simply the set of tuples $F = (F_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$.

For each pair of elements F, G in R^N , we would like to define the composition $F \circ G$ in R^N . To this end recall the special case $m = 2$ of the Faà di Bruno formula Lemma 5.1, which states that

$$(29) \quad D_i^\alpha[F \circ G] = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} \mathcal{E}_{F,G}(\mathcal{T}) \quad \text{where} \quad \mathcal{E}_{F,G}(\mathcal{T}) := \left(D_i^{\mu(v_0)}[F] \circ G \right) \prod_{v \in V_1} D_{\tau(v)}^{\mu(v)}[G].$$

With (29) in mind, we define the composition $F \circ G$ to be the element of R^N whose coefficients are given by

$$(30) \quad (F \circ G)_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} F_{i,\mu(v_0)} \prod_{v \in V_1} D_{\tau(v),\mu(v)}.$$

Composition is clearly associative in that $(F \circ G) \circ H = F \circ (G \circ H)$, with the coefficients of multiple compositions afforded by the power series variant (8) of Theorem 2.1.

Define the element I of R^N by

$$I_{i,\alpha} = \begin{cases} 1 & \text{if } \alpha = \mathbf{e}_i, \\ 0 & \text{otherwise.} \end{cases}$$

We note that I is both a left and right identity for composition of functions in R^N , in that for every F in R^N we have $I \circ F = F \circ I = F$.

We say an element F of R^N is invertible if there exists a G in R^N such that $F \circ G = G \circ F = I$. In Section 6.2 we prove Theorem 2.3, which states that all elements F of R^N whose linear term is the identity matrix are invertible, and gives an explicit expression for the coefficients of the inverse. In the following section, Section 6.3, we prove Theorem 2.5, which states that an element F of R^N is invertible if and only if its linear term is invertible, and again gives an expression for the inverse.

6.2. Inverses of power series with identity linear terms. In this section we now give two proofs of Theorem 2.3 concerning the inverse of power series whose linear term is the identity. The first proof is short, and is simply a matter of verifying that the formula works. The second proof provides more insight, giving an inductive construction of the trees that helps motivate the formula.

Proof 1 of Theorem 2.3. A verification proof. Let G be the power series whose coefficients are given by (12). We now show directly that $F \circ G = I$. Indeed, separating (30) into trees $\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)$ in which the root has one child or more than one child, and using the fact that $F_{i,\mathbf{e}_j} = \delta_{i,j}$, we have

$$(F \circ G)_{i,\alpha} = G_{i,\alpha} - \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2): |\mu(v_0)| \geq 2} H_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)}.$$

We now claim that

$$\sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}: |\mu(v_0)| \geq 2} H_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)} = G_{i,\alpha} := \sum_{\mathcal{U} \in \mathbb{S}_{i,\alpha}} \mathcal{E}_H(\mathcal{U}).$$

Indeed, let for each $\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)$, and each v in the first generation $V_1(\mathcal{T})$ of this tree, expand the term $G_{\tau(v),\mu(v)}$ in terms of a sum over trees (\mathcal{T}_v) . Each possible combination of $\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)$ and $(\mathcal{T}_v : v \in V_1(\mathcal{T}))$ creates a tree \mathcal{U} in $\mathbb{S}_{i,\alpha}$ in such a way that $\mathcal{E}_H(\mathcal{U}) = H_{i,\mu(v_0)} \prod_{v \in V_1(\mathcal{T})} \mathcal{E}_H(\mathcal{T}_v)$. Moreover, every tree \mathcal{U} in $\mathbb{S}_{i,\alpha}$ is created in this way by exactly one possible combination. \square

Proof 2 of Theorem 2.3. A constructive proof. The proof follows by an induction argument on the degree $|\alpha|$ of α . By definition, if F and G are formal inverses of one another, we must have

$$(31) \quad (F \circ G)_{i,\alpha} = \begin{cases} 1 & \text{if } \alpha = \mathbf{e}_i, \\ 0 & \text{otherwise.} \end{cases}$$

On the other hand, by (30) and the definition $(F \circ G)_{i,\alpha} := (F_i \circ G)_{\alpha}$, we have

$$(32) \quad (F \circ G)_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} F_{i,\mu(v_0)} \prod_{v \in V_1} D_{\tau(v),\mu(v)}.$$

In particular, combining (31) and (32), we obtain the following system of equations in i and α :

$$(33) \quad \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2)} F_{i,\mu(v_0)} \prod_{v \in V_1} D_{\tau(v),\mu(v)} = \begin{cases} 1 & \text{if } \alpha = \mathbf{e}_i, \\ 0 & \text{otherwise.} \end{cases}$$

First we show that the linear term $J(G)(0)$ of the inverse G is the identity map. This follows from noting that for $|\alpha| = 1$, the system of equations (33) reads as saying

$$\sum_{k=1}^N F_{i,\mathbf{e}_k} G_{k,\mathbf{e}_j} = \delta_{i,j},$$

or in other words, the matrix composition of $J(F)(0)$ and $J(G)(0)$ is the identity matrix. Since $J(F)(0)$ is the identity, so is $J(G)(0)$.

Now when $|\alpha| = 2$, since every tree in $\mathbb{F}_{i,\alpha}(2)$ has two leaves, the generation V_1 either contains one or contains two vertices. In particular, setting $|\alpha| = 2$ in (33) and using the fact that $F_{i,\mathbf{e}_j} = G_{i,\mathbf{e}_j} = \delta_{i,j}$, we obtain

$$F_{i,\alpha} + G_{i,\alpha} = 0 \quad \text{for } |\alpha| = 2,$$

which establishes (12) for $|\alpha| = 2$, since there is exactly one tree in $\mathbb{S}_{i,\alpha}$. The case $|\alpha| = 2$ will form a base case in our inductive proof.

Suppose now that $\beta \in \mathbb{Z}_{\geq 0}^d$, and that the formula (12) holds for every j in $[N]$ and every $\alpha \in \mathbb{Z}_{\geq 0}^d$ satisfying $|\alpha| < |\beta|$. We now show it holds for β . Indeed, again using (33), and separating $F_i^\alpha(2)$ into those trees in which the root has one child or more children, and then using the fact that $F_{i,\mathbf{e}_j} = \delta_{i,j}$, we obtain

$$0 = G_{i,\alpha} + \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2): |\mu(v_0)| \geq 2} F_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)}.$$

Equivalently, by using the definition of H we have

$$G_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{F}_{i,\alpha}(2): |\mu(v_0)| \geq 2} H_{i,\mu(v_0)} \prod_{v \in V_1} G_{\tau(v),\mu(v)}.$$

Since $G_{i,\mathbf{e}_j} = \delta_{i,j}$, every vertex in V_1 either has more than one child, or they have a single child of the same type as themselves. With this in mind, define $\tilde{\mathbb{F}}_{i,\alpha}(2)$ to be the subset of trees in $\mathbb{F}_{i,\alpha}(2)$ such that the root has two or more children, and such that every $v \in V_1$ with only one child, we have $\mu(v) = \mathbf{e}_{\tau(v)}$.

$$(34) \quad G_{i,\alpha} = \sum_{\mathcal{T} \in \tilde{\mathbb{F}}_{i,\alpha}(2)} H_{i,\mu(v_0)} \prod_{v \in V_1: |\mu(v)| \geq 2} G_{\tau(v),\mu(v)}.$$

Now take a tree \mathcal{T} in $\mathbb{F}_{i,\alpha}(2)$. Since $|\mu(v_0)| \geq 2$, we have $|\mu(v)| < |\alpha|$ for every $v \in V_1$. In particular we may use the inductive hypothesis to expand each $G_{\tau(v),\mu(v)}$ in terms of a tree. In particular, by using the inductive hypothesis in (34) we have

$$\begin{aligned} G_{i,\alpha} &= \sum_{\mathcal{T} \in \tilde{\mathbb{F}}_{i,\alpha}(2)} H_{i,\mu(v_0)} \prod_{v \in V_1: |\mu(v)| \geq 2} \sum_{\mathcal{T}_v \in \mathbb{S}_{\tau(v),\mu(v)}} \mathcal{E}_H(\mathcal{T}_v) \\ &= \sum_{\mathcal{T} \in \tilde{\mathbb{F}}_{i,\alpha}(2)} \sum_{\{(\mathcal{T}_v)_{v \in V_1}: \mathcal{T}_v \in \mathbb{S}_{\tau(v),\mu(v)}: v \in V_1\}} H_{i,\mu(v_0)} \prod_{v \in V_1: |\mu(v)| \geq 2} \mathcal{E}_H(\mathcal{T}_v). \end{aligned}$$

Let $\mathcal{T} \in \widetilde{\mathbb{F}}_{i,\alpha}(2)$, and for each $v \in V_1$ let $\mathcal{T}_v \in \mathbb{S}_{\tau(v),\mu(v)}$. Then it is easily seen that

$$H_{i,\mu(v_0)} \prod_{v \in V_1: |\mu(v)| \geq 2} \mathcal{E}_H(\mathcal{T}_v) = \mathcal{E}(\mathcal{T}'),$$

where \mathcal{T}' is the element of $\mathbb{S}_{i,\alpha}$ such that the subtree of \mathcal{T}' of descendants of each vertex $v \in V_1$ is given by \mathcal{T}_v . Noting that every tree \mathcal{T}' in $\mathbb{S}_{i,\alpha}$ is obtained uniquely in this way, we see that

$$G_{i,\alpha} = \sum_{\mathcal{T}' \in \mathbb{S}_{i,\alpha}} \mathcal{E}_H(\mathcal{T}'),$$

which proves the result. □

6.3. Inverses of power series with non-identical linear terms. In this section we prove the more general result, Theorem 2.5, concerning inversion of power series with non-identical linear terms. We will see that this result may be obtained as a corollary of Theorem 2.3, where the linear term is assumed to be the identity matrix.

Proof of Theorem 2.5. Namely, suppose we have a general element $F = (F_{i,\alpha} : i \in [N], \alpha \in \mathbb{Z}_{\geq 0}^N)$ of R^N whose i^{th} component is given by

$$F_i(X) = \sum_{j=1}^N P_{i,e_j} X_j + \sum_{|\alpha| \geq 2} \frac{F_{i,\alpha}}{\alpha!} X^\alpha,$$

and consider the formal inverse G of F . We now note that if the matrix (P_{i,e_j}) is not invertible, then neither is F , since the linear term of a functional inverse G is necessarily a matrix inverse of F .

Now suppose that (Q_{i,e_j}) is a matrix inverse of (P_{i,e_j}) , so that Q may be thought of as an element of R^N with i^{th} component $Q_i(s) = \sum_{j=1}^N Q_{i,e_j} X_j$. Let $F = P - H$, and define the power series

$$\widetilde{F} := F \circ Q.$$

It is easily verified that the i^{th} component of \widetilde{F} has the form

$$\widetilde{F}_i(s) = X_i - \widetilde{H}_i(s),$$

where $\widetilde{H} = H \circ Q$ contains only terms of degree two or higher. By Theorem 2.3 we have an expression for the coefficients of the inverse \widetilde{G} of \widetilde{F} . Indeed,

$$\widetilde{G}_i(X) = X_i + \sum_{|\alpha| \geq 2} \frac{\widetilde{G}_{i,\alpha}}{\alpha!} X^\alpha,$$

where

$$(35) \quad \widetilde{G}_{i,\alpha} = \sum_{\mathcal{T} \in \mathbb{S}_{i,\alpha}} \mathcal{E}_{\widetilde{H}}(\mathcal{T}).$$

Now it is easily verified that the element G of R^N given by

$$G = Q \circ \widetilde{G}$$

is the inverse of F , and we may use Faà di Bruno's formula to compute the coefficients of G . Indeed, using (29) and the fact that $Q_{i,\alpha}$ is zero whenever α has degree two or higher to obtain the first equality below, and (35) to obtain the second, we have

$$\begin{aligned}
 G_{i,\alpha} &= \sum_{j \in [N]} Q_{i,e_j} \tilde{G}_{j,\alpha} \\
 &= \sum_{j \in [N]} Q_{i,e_j} \sum_{\mathcal{T} \in \mathbb{S}_{j,\alpha}} \mathcal{E}_{\tilde{H}}(\mathcal{T}) \\
 (36) \quad &= \sum_{j \in [N], \mathcal{T} \in \mathbb{S}_{j,\alpha}} Q_{i,e_j} \mathcal{E}_{\tilde{H}}(\mathcal{T}).
 \end{aligned}$$

Now recall that $\mathbb{A}_{i,\alpha}$ is the set of alternating trees: those vertices with an even distance from the root have exactly one child, and vertices lying an odd distance from the root have two or more children.

Note that there is a natural projection $\mathbf{P} : \mathbb{A}_{i,\alpha} \rightarrow \cup_{j \in [N]} \mathbb{S}_{j,\alpha}$ defined as follows. Let \mathcal{A} be an alternating tree in $\mathbb{A}_{i,\alpha}$, and define an element $\mathcal{T}' = (V', E', \tau', \phi') = \mathbf{P}(\mathcal{A})$ of $\mathbb{S}_{i,\alpha}$ as follows:

- The vertex set of \mathcal{T}' is given by

$$V' := \bigcup_{k \geq 0} V_{2k+1}(\mathcal{A}),$$

- The edge set of \mathcal{T}' is given by

$$E' := \{(v_0, v_1)\} \cup \{(v, w) \in V' : v \text{ is a grandchild of } w \text{ in } \mathcal{A} \text{ or vice versa}\},$$

where (v_0, v_1) is the edge between the root and the unique vertex in V_1 .

- Finally, we let τ' and ϕ' be the restrictions of τ and ϕ to V' . (We note that every leaf of the alternating tree \mathcal{T} is in an odd generation, and hence also an element of V' .)

Given \mathcal{T}^* in $\mathbb{S}_{i,\alpha}$, define $\mathbf{P}^{-1}(\mathcal{T}^*)$ to be the subset of $\mathbb{A}_{i,\alpha}$ consisting of alternating trees \mathcal{A} such that $\mathbf{P}(\mathcal{A}) = \mathcal{T}^*$. The result now follows from the simple observation that

$$(37) \quad Q_{i,e_j} \mathcal{E}_{H \circ Q}(\mathcal{T}) = \sum_{\mathcal{A} \in \mathbf{P}^{-1}(\mathcal{T})} \mathcal{E}_{H,Q}^{\text{alt}}(\mathcal{A}).$$

Indeed, plugging (37) into (36) we obtain

$$G_{i,\alpha} = \sum_{j \in [N], \mathcal{T} \in \mathbb{S}_{j,\alpha}} \sum_{\mathcal{A} \in \mathbf{P}^{-1}(\mathcal{T})} \mathcal{E}_{H,Q}^{\text{alt}}(\mathcal{A}).$$

The result follows by noting that we may write $\mathbb{A}_{i,\alpha}$ as a disjoint union

$$\mathbb{A}_{i,\alpha} = \bigcup_{j \in [N]} \bigcup_{\mathcal{T} \in \mathbb{S}_{j,\alpha}} \mathbf{P}^{-1}(\mathcal{T}).$$

□

6.4. Proof of Lemma 2.8. In this section we prove the *fern lemma*, Lemma 2.8, which was used in our combinatorial restatement of the Jacobian conjecture, Conjecture 2.9.

Proof of Lemma 2.8. Let H be a polynomial mapping containing only terms of degree two or higher. The $(i, j)^{\text{th}}$ entry of the matrix $J(H)^m$ may be written

$$G_{i,j}(X) := \sum_{k_0, \dots, k_m \in [N]} \mathbb{1}\{k_0 = i, k_m = j\} \prod_{l=1}^m D_{k_{l-1}}^{e_{k_l}}[H](X).$$

Clearly, since H contains only degree two terms or higher, $G_{i,j}(X)$ contains only terms of degree m or higher, and hence for $G_{i,j}(0) = 0$ whenever $m \geq 1$.

The property that $J(H)^m$ is equal to the zero matrix on \mathbb{C}^N is equivalent to $G_{i,j}(X) = 0$ for all $1 \leq i, j \leq N$ and all $X \in \mathbb{C}^N$. Equivalently, all derivatives of $G_{i,j}$ are also equal to zero, so that for every $\alpha \in \mathbb{Z}_{\geq 0}^d$,

$$(38) \quad 0 = D^\alpha[G_{i,j}](X) = \sum_{k_0, \dots, k_m \in [N]} \mathbb{1}\{k_0 = i, k_m = j\} D^\alpha \left[\prod_{l=1}^m D_{k_{l-1}}^{e_{k_l}}[H] \right] (X).$$

Using the multivariate Leibniz rule (7), we see that

$$(39) \quad D^\alpha \left[\prod_{l=1}^m D_{k_{l-1}}^{k_l}[H] \right] = \sum_{S_1 \sqcup \dots \sqcup S_m = [\alpha]} \prod_{l=1}^m D_{k_{l-1}}^{e_{k_l} + \#S_l}[H].$$

We now note that every tuple of numbers $(k_0, \dots, k_m) \in [N]$ with $k_0 = i$ and $k_m = j$, together with any sequence of sets (S_1, \dots, S_m) , gives rise to a fern \mathcal{T} in $\text{Fern}_{i,\alpha,j}(m)$ by letting the leaves with labels in S_i be children of v_{i-1} , and letting the type of v_i be k_i . Moreover, the energy of such a fern is given by

$$(40) \quad \mathcal{E}_H(\mathcal{T}) = \prod_{l=1}^m D_{k_{l-1}}^{e_{k_l} + \#S_l}[H].$$

In particular, by plugging (39) into (38), and using (40), we see that (38) reads as saying

$$\sum_{\mathcal{T} \in \text{Fern}_{i,\alpha,j}(m)} \mathcal{E}_H(\mathcal{T}) = 0.$$

□

7. THE NON-COMMUTATIVE CASE

In this section, we shall present the non-commutative/free versions of our main results and the reader should recall the notation introduced in Section 3. In Section 7.1 we prove the non-commutative generalisation of Faà di Bruno's formula, Theorem 3.2. In Section 7.2 we prove the non-commutative inversion formulas, Theorems 3.4 and 3.5.

7.1. Proof of the free Faà di Bruno formula. In analogy with proving the commutative version of Faà di Bruno's formula, in working towards a proof of Theorem 3.2, we begin by considering the case $m = 2$.

The case $m = 2$ of Theorem 3.2 reads as saying that whenever F, G are power series in \vec{R}_0^N with zero constant term, then the coefficient of $X_{\kappa_1} \dots X_{\kappa_k}$ is given by

$$(41) \quad (F \circ G)_{i,\kappa} = \sum_{\vec{\tau} \in \vec{\mathbb{F}}_{i,\kappa}(2)} F_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v),\vec{\mu}(v)}.$$

We now prove this equation.

Proof of equation (41). Note that if F and F' are elements of \vec{R}^N , then

$$(F + F') \circ G = F \circ G + F' \circ G.$$

In particular, $(F \circ G)_{i,\kappa}$ depends linearly on F . Since the right-hand side of (41) also clearly depends linearly on the coefficients of F , it is sufficient to prove (41) for F_i of the form

$$(42) \quad F_i(X_1, \dots, X_N) = X_{\gamma_1} \dots X_{\gamma_j}.$$

For F of the form (42), $F_i \circ G$ is given by

$$F_i \circ G(X_1, \dots, X_N) = G_{\gamma_1}(X_1, \dots, X_N) \dots G_{\gamma_j}(X_1, \dots, X_N).$$

Clearly then the coefficient of $X_{\kappa_1} \dots X_{\kappa_k}$ in $(F_i \circ G)$ is given by

$$(43) \quad (F \circ G)_{i, \kappa} = \sum_{1 \leq l_1 < l_2 < \dots < l_{j-1} < k} G_{\gamma_1, (\kappa_1, \dots, \kappa_{l_1})} G_{\gamma_2, (\kappa_{l_1+1}, \dots, \kappa_{l_2})} \dots G_{\gamma_j, (\kappa_{l_{j-1}+1}, \dots, \kappa_k)}.$$

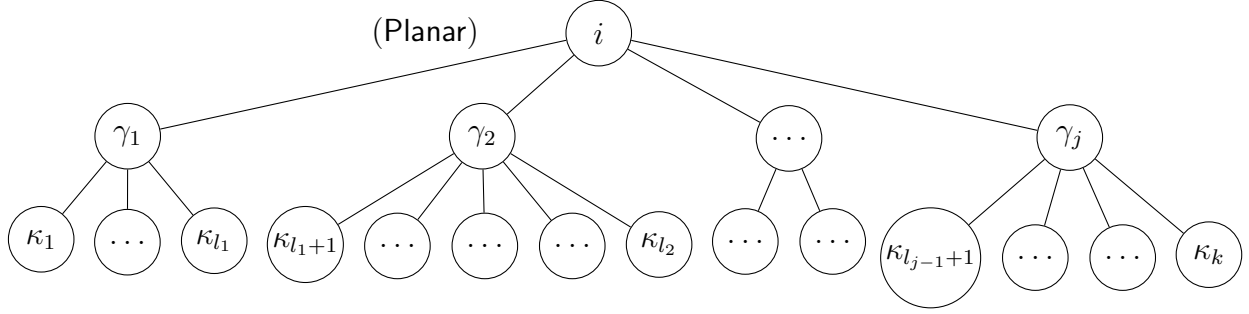


FIGURE 11. The coefficient of $X_{\kappa_1} \dots X_{\kappa_k}$ in the i^{th} component of $(F \circ G)$ may be given in terms of a sum over planar trees with two generations.

Now there is a bijection between every tuple $1 \leq l_1 < l_2 < \dots < l_{j-1} < k$ the subset of $\overrightarrow{\mathbb{F}}_{i, \kappa}(2)$ consisting of the planar trees with j vertices $\{v_1, \dots, v_j\}$ in generation 1 (listed in planar order) by letting the leaves with labels $\{l_{j-1} + 1, \dots, l_j\}$ be descended from v_j . Write $\mathcal{T}(l_1, \dots, l_{j-1})$ for the unique tree with this property. It follows that (43) may be written

$$(44) \quad (F_i \circ G)_{\kappa} = \sum_{\vec{\tau} \in \overrightarrow{\mathbb{F}}_{i, \kappa}(2)_{\gamma}} \prod_{v \in V_1} G_{\tau(v), \vec{\mu}(v)},$$

where $\overrightarrow{\mathbb{F}}_{i, \kappa}(2)_{\gamma}$ is the subset of $\overrightarrow{\mathbb{F}}_{i, \kappa}(2)$ consisting of those trees with vertices $\{v_1, \dots, v_j\}$ in generation 1 with types $\gamma_1, \dots, \gamma_j$.

In other words, if F_i has the form (42), then $F_{i, \gamma} = 1$ and $F_{i, \gamma'} = 0$ for all tuples $\gamma' \neq \gamma$, so that we may write (44) as

$$(F \circ G)_{i, \kappa} = \sum_{\vec{\tau} \in \overrightarrow{\mathbb{F}}_{i, \kappa}(2)} F_{i, \vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v), \vec{\mu}(v)},$$

which amounts to (41) for $F_i(X_1, \dots, X_k) = X_{\gamma}$, which by linearity, implies (41) for general F . \square

We now prove Theorem 3.2 by induction on m . Since the proof bears a great deal of similarity with our proof in the analogous commutative case, we will provide less detail here than we did in the commutative case.

Proof of Theorem 3.2. We proceed by induction on m by using the $m = 2$ case proved above to prove the inductive step. Suppose we have a sequence $(F^{(1)}, \dots, F^{(m+1)})$ of elements of \overrightarrow{R}_0^N . Then by letting $F = F^{(1)}$ and $G = (F^{(2)} \circ \dots \circ F^{(m+1)})$ in (41), we obtain

$$(45) \quad (F^{(1)} \circ \dots \circ F^{(m+1)})_{i, \kappa} = \sum_{\mathcal{T} \in \overrightarrow{\mathbb{F}}_{i, \kappa}(2)} F_{i, \vec{\mu}(v_0)}^{(1)} \prod_{v \in V_1} (F^{(2)} \circ \dots \circ F^{(m+1)})_{\tau(v), \vec{\mu}(v)}.$$

By the inductive hypothesis, we may expand each term

$$(46) \quad (F^{(2)} \circ \dots \circ F^{(m+1)})_{\tau(v), \vec{\mu}(v)} = \sum_{\mathcal{T}_v \in \mathbb{F}_{\tau(v)}^{\vec{\mu}(v)}(m)} \mathcal{E}_{(F^{(2)}, \dots, F^{(m+1)})}(\vec{\mathcal{T}}_v).$$

Suppose now that $\vec{\mathcal{T}}$ is a tree in $\vec{\mathbb{F}}_{i, \kappa}(2)$, such that the vertices in generation 1 have types and free outdegrees $\{(\tau(v), \vec{\mu}(v)) : v \in V_1\}$. Given a collection of trees $(\vec{\mathcal{T}}_v : v \in V_1)$ such that each $\vec{\mathcal{T}}_v$ is an element of $\vec{\mathbb{F}}_{\tau(v), \vec{\mu}(v)}(m)$, we may create a tree $\mathcal{S} = \text{Tree}(\vec{\mathcal{T}}; \vec{\mathcal{T}}_v : v \in V_1(\vec{\mathcal{T}}))$ in $\vec{\mathbb{F}}_{i, \kappa}(m+1)$ by “glueing” each \mathcal{T}_v to the vertex v in $\vec{\mathcal{T}}$. Moreover, the $F^* = (F^{(1)}, \dots, F^{(m+1)})$ -energy of the resulting tree is given by

$$\mathcal{E}_{F^*}(\mathcal{S}) = \prod_{l=1}^{m+1} \prod_{v \in V_{l-1}} F_{\tau(v), \vec{\mu}(v)}^{(l)} = F_{i, \vec{\mu}(v_0)} \prod_{v \in V_1} \mathcal{E}_{(F^{(2)}, \dots, F^{(m+1)})}(\vec{\mathcal{T}}_v).$$

In particular, since every tree \mathcal{S} has a unique decomposition into trees $(\vec{\mathcal{T}}_v : v \in V_1(\mathcal{S}))$ it follows that by plugging (46) into (45) that

$$\begin{aligned} (F^{(1)} \circ \dots \circ F^{(m+1)})_{i, \kappa} &= \sum_{\mathcal{T} \in \vec{\mathbb{F}}_{i, \kappa}(2)} F_{i, \vec{\mu}(v_0)}^{(1)} \prod_{v \in V_1} (F^{(2)} \circ \dots \circ F^{(m+1)})_{\tau(v), \vec{\mu}(v)} \\ &= \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i, \kappa}(2)} F_{i, \vec{\mu}(v_0)}^{(1)} \prod_{v \in V_1} \sum_{\vec{\mathcal{T}}_v \in \vec{\mathbb{F}}_{\tau(v), \vec{\mu}(v)}(m)} \mathcal{E}_{(F^{(2)}, \dots, F^{(m+1)})}(\vec{\mathcal{T}}_v) \\ &= \sum_{\vec{\mathcal{U}} \in \vec{\mathbb{F}}_{i, \kappa}(m+1)} \mathcal{E}_{(F^{(1)}, \dots, F^{(m+1)})}(\vec{\mathcal{U}}). \end{aligned}$$

This shows that the formula also holds for $m+1$. □

7.2. Proof of the free inversion theorem. In this section we prove Theorem 3.4, which gives a formula for the compositional inverse of a formal power series in non-commutative indeterminates. Recall that in Section 6.2 we gave two proofs of the commutative analogue Theorem 2.3 of Theorem 3.4 — the first constructive and the latter a simple verification.

In terms of proving Theorem 3.4 here, we only give a verification proof, since a constructive argument may be intimated from paralleling the constructive argument in the commutative case.

Proof of Theorem 3.4. Suppose we are in the context of Theorem 3.4. Namely, suppose F and G are elements of \vec{R}_0^N such that each component F_i of F has the form

$$F_i(X_1, \dots, X_N) = X_i - \sum_{k \geq 2} \sum_{\kappa \in [N]^k} H_{i, \kappa} X_\kappa,$$

and each component of G has the form

$$G_i(X_1, \dots, X_N) = X_i + \sum_{k \geq 2} \sum_{\kappa \in [N]^k} G_{i, \kappa} X_\kappa,$$

where

$$G_{i, \kappa} = \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{S}}_{i, \kappa}} \mathcal{E}_H(\vec{\mathcal{T}}).$$

We now use the special case $m = 2$ of the free Faà di Bruno formula (41) to verify that both $F \circ G$ and $G \circ F$ are equal to the identity power series I .

We begin with $F \circ G$. For $j \in [N]$, by letting κ be the 1-tuple $\kappa = (j)$ in (41), we see that the coefficient of X_j in the i^{th} component of $(F \circ G)$ is given by

$$(F \circ G)_{i,j} = \sum_{k \in [N]} F_{i,k} G_{k,j} = \mathbb{1}_{i=j}.$$

We now show that whenever $k \geq 2$, for any $\kappa \in [N]^k$, the coefficient of X_κ in the i^{th} component of $(F \circ G)$ is zero. Indeed, by (41)

$$(47) \quad (F \circ G)_{i,\kappa} = \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2)} F_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v),\vec{\mu}(v)} = G_{i,\kappa} - \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2):|V_1|>2} H_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v),\vec{\mu}(v)}.$$

We now note that each $G_{\tau(v),\vec{\mu}(v)}$ may be expanded in terms of a sum over trees in $\vec{\mathbb{S}}_{\tau(v),\vec{\mu}(v)}$. In particular,

$$(48) \quad \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2):|V_1|>2} H_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v),\vec{\mu}(v)} = \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2):|V_1|>2} \sum_{(\vec{\mathcal{T}}_v)_{v \in V_1}, \vec{\mathcal{T}}_v \in \vec{\mathbb{S}}_{\tau(v),\vec{\mu}(v)}} H_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} \mathcal{E}_H(\vec{\mathcal{T}}_v).$$

However, for each term in the summand on the right-hand side, we may write

$$\mathcal{E}(\vec{\mathcal{T}}^*) = H_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} \mathcal{E}_H(\vec{\mathcal{T}}_v),$$

where $\vec{\mathcal{T}}^*$ is the tree in $\vec{\mathbb{S}}_{i,\kappa}$ obtained by “glueing” each $\vec{\mathcal{T}}_v$ to v . Moreover, each tree $\vec{\mathcal{T}}^*$ in $\vec{\mathbb{S}}_{i,\kappa}$ is obtained uniquely in this from a tree $\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2)$, and a collection $(\vec{\mathcal{T}}_v)$ of trees such that each $\vec{\mathcal{T}}_v$ is contained in $\vec{\mathbb{S}}_{\tau(v),\vec{\mu}(v)}$.

In particular, (48) reads as saying

$$(49) \quad \sum_{\vec{\mathcal{T}} \in \vec{\mathbb{F}}_{i,\kappa}(2):|V_1|>2} H_{i,\vec{\mu}(v_0)} \prod_{v \in V_1} G_{\tau(v),\vec{\mu}(v)} = \sum_{\vec{\mathcal{T}}^* \in \vec{\mathbb{S}}_{i,\kappa}} \mathcal{E}_H(\vec{\mathcal{T}}^*),$$

so that by (47), $(F \circ G)_{i,\kappa} = 0$. The proof that $G \circ F$ is equal to the identity is similar, and left to the reader. \square

We now turn to proving Theorem 3.5, concerning inverses of power series with non-identical linear terms in N free indeterminates. Since the proof of Theorem 3.5 is extremely similar to our proof of its commutative counterpart, Theorem 2.5, we shall only sketch the main details.

Proof of Theorem 3.5. Let F be a free power series in \vec{R}^N with components of the form

$$F_i(X_1, \dots, X_N) = \sum_{j \in [N]} P_{i,j} X_j - \sum_{k \geq 2} \sum_{\kappa \in [N]^k} H_{i,\kappa} X_\kappa, \quad i \in [N].$$

Now consider the element \tilde{F} of \vec{R}^N defined by setting $\tilde{F} = F \circ Q$. It is easily verified using (41) that \tilde{F} has components of the form

$$\tilde{F}_i(X_1, \dots, X_N) = X_i - \sum_{k \geq 2} \sum_{\kappa \in [N]^k} \tilde{H}_{i,\kappa} X_\kappa, \quad i \in [N].$$

where $\tilde{H} = H \circ Q$. In particular, by Theorem 3.4 the inverse \tilde{G} exists and has components of the form

$$\tilde{G}_i(X_1, \dots, X_N) = X_i + \sum_{k \geq 2} \sum_{\kappa \in [N]^k} G_{i,\kappa} X_\kappa,$$

where

$$\tilde{G}_{i,\kappa} = \sum_{\vec{\mathcal{T}} \in \vec{\mathcal{S}}_{i,\kappa}} \mathcal{E}_{\tilde{H}}(\vec{\mathcal{T}}).$$

It is easily verified now that $G := Q \circ \tilde{F}$ is inverse to F . Moreover, in a planar expansion identical to its non-planar counterpart in the proof of Theorem 2.5, we see that the coefficients of G are given by

$$G_{i,\kappa} = \sum_{\vec{\mathcal{T}} \in \vec{\mathcal{A}}_{i,\kappa}} \mathcal{E}_{H,Q}^{\text{alt}}(\vec{\mathcal{T}}),$$

as required. □

Acknowledgments. The authors are extremely grateful to an anonymous referee whose suggestions have greatly improved this article. We would also like to thank Michael Anshelevich for directing us towards several useful references.

SJ and JP have been supported by the Austrian Science Fund (FWF) Project P32405 “Asymptotic Geometric Analysis and Applications” of which JP is principal investigator. JP has also been supported by a *Visiting International Professor Fellowship* from the Ruhr University Bochum and its Research School PLUS.

Finally, we gratefully acknowledge the support of the Oberwolfach Research Institute for Mathematics, where several discussions were held during the workshop “New Perspectives and Computational Challenges in High Dimensions” (Workshop ID 2006b).

REFERENCES

- [1] ANSHELEVICH, M., EFFROS, E.G. AND POPA, M. (2006). Zimmermann type cancellation in the free Faà di Bruno algebra, *J. Funct. Anal.* **237** 76–104.
- [2] ARBOGAST, L.F.A. (1800). *Du Calcul des Dérivations*, Levrault, Strasbourg.
- [3] ABRAHAM, R. AND ROBBIN, J. (1967). *Transversal Mappings and Flows*. New York/Amsterdam. W. A. Benjamin, Inc.
- [4] Athreya, K.B. and Ney. P.E. (1972). *Branching Processes*, Springer-Verlag, New York Heidelberg Berlin.
- [5] BASS, H., CONNELL, E. H. AND WRIGHT, D. (1982). The Jacobian conjecture: reduction of degree and formal expansion of the inverse, *Bull. Amer. Math. Soc.* **7**(2), 287–330.
- [6] BERETTI, A. AND GENTILE, G. (1998.) Bryuno function and the standard map. *Comm. Math. Phys.* **220** 623–656.
- [7] BROUDER, C., FRABETTI, A. AND KRATTENTHALER, C. (2006) Non-commutative Hopf algebra of formal diffeomorphisms, *Adv. Math.*, **200** 479–524.
- [8] CONSTANTINE, G. M. AND SAVITS, T.H. (1996). A Multivariate Faà di Bruno Formula with Applications, *T. Am. Math. Soc.*, **348** 503-520.
- [9] CARLETTI, T. (2003). The Lagrange inversion formula on non-Archimedean fields. Non-analytical form of differential and finite difference equations. *Discrete Contin. Dyn. Sys.* **9**(4) 835–858.
- [10] CHENG, C.C., MCKAY J. H., TOWBER J., WANG S. S.-S., AND WRIGHT D. L. (Reversion of power series and the extended Raney coefficients. *Tran. Am. Math. Soc* **349**(5) 1769–1782.
- [11] DE BONDT, M. AND VAN DEN ESSEN, A. (2005). A reduction of the Jacobian conjecture to the symmetric case. *P. Am. Math. Soc.* **133**(8) 2201–2205.
- [12] DE BONDT, M. AND VAN DEN ESSEN, A. (2005). The Jacobian Conjecture for symmetric Drużkowski mappings. *Ann. Polonici Mathematici* **86**(1) 43–46.
- [13] DRUŻKOWSKI, L.M. (1983). An effective approach to Keller’s Jacobian conjecture, *Math. Ann.*, **264** 303–313.

- [14] DRUŻKOWSKI, L.M. (2001). New reductions in the Jacobian conjecture. *Universitatis Iagellonicae Acta Mathematica*, Fasciculus XXXIX.
- [15] FAÀ DI BRUNO, F. (1855). Sullo Sviluppo delle Funzioni, *Annali di Scienze Matematiche e Fisiche* **6**, 479–480.
- [16] FAÀ DI BRUNO, F. (1857). Note sur un nouvelle formule de calcul différentiel, *Quarterly Journal of Pure and Applied Mathematics*, **1**, 359–360.
- [17] FAÀ DI BRUNO, F. (1885). Sullo Sviluppo delle Funzioni, *Annali di Scienze Matematiche e Fisiche* **6**, 479–480.
- [18] FRABETTI, A. AND SHESTAKOV, I.P. (2019). Loop of formal diffeomorphisms and Faà di Bruno coloop bialgebra, *Adv. Math.*, **351**, 495–569.
- [19] GEL'FAND, I. M., GEL'FAND, S., RETAKH, V. AND WILSON, R. L. (2005). Quasideterminants, *Adv. Math.*, **193**(1) 56–141.
- [20] GEL'FAND, I.M. AND RETAKH, V.S. (1991). Determinants of matrices over noncommutative rings, *Funktsional. Anal. i Prilozhen.* **25**(2), 13–25, 96.
- [21] GENTILE, G (2009). Diagrammatic techniques in perturbation theory, In: *Encyclopedia of Complexity and Systems Science*.
- [22] GESSEL, I.M. (1987). A combinatorial proof of the multivariable Lagrange inversion formula, *J. Comb. Theory. A* **45** 178–195.
- [23] GODSIL, C.D. (1981). Hermite polynomials and a duality relation for matchings polynomials, *Combinatorica* **1**(3) 257–262.
- [24] GZYL, H. (1986). Multidimensional Extension of Faa di Bruno's formula, *J. Math. Anal. Appl.*, **116** 450–455.
- [25] HARDY, M. (2006). The combinatorics of partial derivatives, *Electron. J. Comb.* **11**(1).
- [26] HAIMAN, M. AND SCHMITT, W. (1989). Incidence algebra antipodes and Lagrange inversion in one and several variables. *J. Comb. Theory A*, **50** 172–185.
- [27] HERNÁNDEZ ENCINAS, L. AND MUÑOZ MASQUÉ, J. (2003) A short proof of the generalized Faà di Bruno's formula. *Appl. Math. Lett.* **16** 975–979.
- [28] JOHNSON, W. P. (2002). The curious history of Faà di Bruno's formula, *Am. Math. Mon.*, **109** 217–234.
- [29] JOHNSTON, S.G.G. (2019). The genealogy of Bienaymé-Galton-Watson trees, *Electron. J. Probab.*, **24**, Paper no. 94.
- [30] JOHNSTON, S.G.G. AND LAMBERT, A. (2019). The coalescent structure of branching processes: a unifying Poissonization approach. arXiv: 1912.00198.
- [31] LEIPNIK, R. AND REID, T. (1996). Multivariable Faa di Bruno Formulas, *Electronic Proceedings of the Ninth Annual International Conference on Technology in Collegiate Mathematics*.
- [32] LIU, M. AND VATUTIN, V.A. (2019). Reduced critical branching processes for small populations. *Theory Probab. Appl.* **63**(4) 648–656.
- [33] MEURMAN, A., FRENKEL, I. AND LEPOWSKY, J. (1988). *Vertex operator algebras and the Monster*, Pure and Applied Mathematics, **134**, Boston, MA.
- [34] MINGO, J.A. AND SPEICHER, R. (2017). *Free Probability and Random Matrices*. Fields Institute Monographs, **35**.
- [35] MOH, T.-T. (1983). On the Jacobian conjecture and the configuration of roots, *J. Reine Angew. Math.*, **340** 140–212.
- [36] ORE, O. (1933). Theory of Non-Commutative Polynomials, *Ann. Math.* **34**(3), 480–508.
- [37] ROTA, G.-C., SAGAN, B. AND STEIN, P.R. (1980). A cyclic derivative in non-commutative algebra. *J. Algebra* **64**, 54 – 75.
- [38] SINGER. D (2001). On Catalan trees and the Jacobian conjecture. *Electron. J. Combinatorics.* **8**(1).
- [39] SINGER. D (2009). A graph-theoretic method for choosing a spanning set for a finite-dimensional vector space, with applications to the Grossman-Larson-Wright module and the Jacobian conjecture *Electron. J. Combinatorics.* **16**.
- [40] SINGER. D (2011). Towards a combinatorial proof of the Jacobian conjecture. *Electron. J. Combinatorics.* **18**(2).
- [41] STANLEY, R.P. (1999). *Enumerative Combinatorics, Volume 2*. Cambridge University Press.
- [42] WANG, L.-C. (2005). On the Jacobian conjecture. *Taiwan. J. Math.* **9**(3) 421–431.
- [43] WANG, S. (1980). A jacobian criterion for separability, *J. Algebra* **65**, 453–494.
- [44] TAYLOR, M. E. (1999). *Partial differential equations, Volume I, Basic Theory*, Springer-Verlag New York, Inc.
- [45] VAN DEN ESSEN, A. (2000). *Polynomial automorphisms and the Jacobian conjecture*, Birkhäuser Basel.
- [46] WRIGHT, D. (1987). Formal inverse expansion and the Jacobian conjecture. *J. Pure. Appl. Alg.* **48** 199–219.
- [47] WRIGHT, D. (1989). The tree formulas for reversion of power series. *J. Pure. Appl. Algebra*, **57** 191–211. In: *Affine Algebraic Geometry, in honor of Masayoshi Miyanishi*, edited by Takayuki Hibi, published by Osaka University Press 2007.
- [48] WRIGHT, D. (2005). The Jacobian conjecture: ideal membership questions and recent advances. *Contemp. Math.* **369**.

- [49] WRIGHT, D. (2007). The Jacobian conjecture as a problem in combinatorics. In: *Affine Algebraic Geometry, in honor of Masayoshi Miyanishi*, edited by Takayuki Hibi, published by Osaka University Press 2007.
- [50] VATUTIN, V.A., HONG, W., AND JI, Y. (2018). Reduced critical Bellman-Harris branching processes for small populations. *Discrete Math. Appl.* **28**(5) 319–330.
- [51] YANG, W.C. (2000). Derivatives are essentially integer partitions. *Discrete Math.* **222** 235–245.
- [52] ZHAO, W. (2005). Inversion problem, Legendre transform and inviscid Burgers' equations. *J. Pure. Appl. Algebra* **199** 299–317.

SAMUEL G. G. JOHNSTON: DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF BATH, UNITED KINGDOM.
Email address: samuel.johnston@uni-graz.at

JOSCHA PROCHNO: FACULTY OF COMPUTER SCIENCE AND MATHEMATICS, UNIVERSITY OF PASSAU, GERMANY.
Email address: joscha.prochno@uni-passau.de