AN EXPONENTIAL FORMULA FOR POLYNOMIAL VECTOR FIELDS (II): LIE SERIES, EXPONENTIAL SUBSTITUITION, AND ROOTED TREES

RUDOLF WINKEL

Abstract. Gröbner's Lie Series [G] and the Exponential Formula [W] provide different explicit formulas for the flow generated by a finite-dimensional polynomial vector field. The present paper gives (1) a generalization of the Lie series in case of non-commuting variables called Exponential Substituition, (2) a structural understanding of the three formulas and their mutual relationships in terms of rooted trees, and (3) as a byproduct new results on the enumeration, coding, and statistics of different kinds of rooted trees.

1. Introduction

Exponential formulas of the type

$$\exp(\Box) = \sum_{n=0}^{\infty} \frac{\Box^n}{n!} \ ,$$

are of widespread use in mathematics. Familiar examples are the ordinary exponential function ($\Box = x$ a real or complex variable), a generating function ($\Box = yf(t)$ with $f(t)$ a formal power series without constant term), the exponential map ($\Box = tX$ with X a tangent vector at the unit of a Lie group G), a unitary one-parameter group ($\Box = tA$ with A a selfadjoint linear operator), and the matrix exponentiation ($\Box = tM$ with M an $N \times N$ matrix over the reals $\mathbb{R}$ or complex numbers $\mathbb{C}$). It is well known that $\exp(tM)$ describes the global flow for the linear sytem of ODE $\dot{x} = Mx$, where $x = (x_1, \ldots, x_N)^*$ ($'*'$ for 'transpose'), and the local power series solution when applied to any initial value $\bar{x} \in \mathbb{R}^N$.

In the present paper the connection between differential operators and rooted trees — as observed already by Cayley [C1,C2] — is used to understand and compare three generalizations of the matrix exponentiation: *Gröbner's Lie series* [G], the *Exponential Formula* for polynomial vector fields [W], and the *exponential substituition* (introduced here for the first time). More specificaly, let R be any commutative ring with unit containing the real numbers $\mathbb{R}$, and for any natural number $N \in \mathbb{N}$ let $f : R^N \longrightarrow R^N$ be the

Date: April 21, 1998.

mapping, which has as components the formal power series $f_1, \ldots, f_N \in R[[x_1, \ldots, x_N]]$. Let ∇ be the "column vector", which contains the formal partial derivatives $\partial_{x_i} \equiv \partial/\partial_{x_i}$ for $i = 1, \ldots, N$ as entries, and set

$$D \equiv D(f) = f \cdot \nabla = f_1 \partial_{x_1} + \cdots + f_N \partial_{x_N} \ . \tag{1.1}$$

Let $D^n = D \circ \cdots \circ D$ be the n-fold composition of D with itself, where of course D^0 is the identity operator. Since the application of D to any formal power series $g \in R[[x_1, \ldots, x_N]]$ is again an element of $R[[x_1, \ldots, x_N]]$, the same is true for its exponentiation:

$$e^{tD} g(x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} D^n g(x) \in R[[x_1, \ldots, x_N, t]] \ . \tag{1.2}$$

If $g = (g_1, \ldots, g_N)^*$ is a vector of formal power series $g_i \in R[[x_1, \ldots, x_N]]$, then the operator $\exp(tD)$ is applied componentwise.— Gröbner has called expressions of the form $\exp(tD)g(x)$ *Lie series*, because they first occured (rather marginally) in the work of Lie [L, 1. Abschnitt, Kap.3, §12]. Gröbner was the first to observe that the Lie series

$$e^{tD} \bar{x} := e^{tD} x|_{x=\bar{x}} \tag{1.3}$$

solves the initial value problem (IVP)

$$\dot{x} = f(x), \ x(0) = \bar{x} \in R^N \tag{1.4}$$

in case of convergent (= analytical) component functions $f_1, \ldots, f_N \in \mathbb{C}\{x_1, \ldots, x_N\}$. In Section 2 we will give a concise account of Gröbner's basic results on Lie series and discuss an application to the inversion of mappings f and the Jacobian Conjecture.

For the second generalization of matrix exponentiation assume that $f : R^N \longrightarrow R^N$ is a *polynomial* mapping, where all components $f_1, \ldots, f_N \in R[x_1, \ldots, x_N]$ are homogeneous of degree m. (This is no restricition of generality, because every non-homogeneous polynomial f can be made homogeneous at the expense of one additional dimension.) In [W] it has been shown that the polynomial IVP (1.4) is solved by the Exponential Formula $\exp(t\mu)$ applied to the initial vector $\bar{x} \in R^N$: let $e_1, \ldots, e_N$ be the canonical basis vectors of R^N, then μ is defined as the R-linear mapping

$$\mu : T_m R^N \equiv \bigotimes^{m} R^N \to R^N, \quad \mu(e_{h_1} \otimes \ldots \otimes e_{h_m}) := \sum_{i=1}^{n} a^i_{h_1 \ldots h_m} e_i \ , \tag{1.5}$$

which is associated to the m-homogeneous polynomial vector field f with components

$$f_i(x_1, \ldots, x_N) := \sum_{h_1, \ldots, h_m = 1}^{N} a^i_{h_1 \ldots h_m} x_{h_1} \cdot \cdots \cdot x_{h_m} \qquad (i = 1, \ldots, N) \ . \tag{1.6}$$

Using the abbreviation

$$[k] := k(m-1)+1 \quad \text{for all } k \in \mathbb{N}_0 \text{ and fixed } m \in \mathbb{N} \tag{1.7}$$

one defines for every $\mu \in L(T_m R^N, R^N)$ the following mappings:

$$d_{\mu,p} : T_p R^N \to T_{p-m+1} R^N$$

$$d_{\mu,p} := \begin{cases} 0 & \text{, if } p < m, \\ \sum_{\nu=1}^{p-m+1} \otimes^{\nu-1} id \otimes \mu \otimes \otimes^{p-m-\nu+1} id & \text{, if } p \geq m \end{cases} , \tag{1.8}$$

$$\delta_p : R^N \longrightarrow T_p R^N, \quad v \mapsto \otimes^p v \quad \text{, if } p > 0, \text{ and} \tag{1.9}$$

$$\mu^n := \begin{cases} id & \text{, if } n = 0 \\ d_{\mu,[1]} \circ \cdots \circ d_{\mu,[n]} \circ \delta_{[n]} & \text{, if } n > 0 \ . \end{cases} \tag{1.10}$$

Therefore the "powers" μ^n can be represented as the concatenation:

$$\mu^n \ : \ R^N \xrightarrow{\delta_{[n]}} T_{[n]} R^N \xrightarrow{d_{\mu,[n]}} T_{[n-1]} R^N \xrightarrow{d_{\mu,[n-1]}} \cdots \xrightarrow{d_{\mu,[2]}} T_{[1]} R^N \ = \ T_m R^N \xrightarrow{d_{\mu,[1]}=\mu} R^N \ ,$$

and the exponentiation of μ as

$$\exp(t\mu) := \sum_{n=0}^{\infty} \frac{t^n}{n!} \mu^n \ . \tag{1.11}$$

It is well known that for (real) analytic f the solution of the IVP (1.4) is (real) analytic, hence it follows immediately from the uniqueness of the local power series solution that for homogeneous polynomial mappings f one has

$$D^n(x) = \mu^n(x) \quad \text{for all } n \geq 0 \ . \tag{1.12}$$

But this identity is not at all obvious from the definitions: in the definition (1.1) of D it is necessary to fix the dimension N, whereas the degree of m of f is secondary; for μ the degree m is primary, whereas the dimension N is irrelevant for the construction of $\exp(t\mu)$. Therefore one of the main tasks of the present paper is to provide a *constructive understanding* of the family of identities (1.12).

So far we have considered the case of commuting variables $x_1, \ldots, x_N$. The third generalization of the matrix exponentiation is concerned with the case of mappings f in non-commuting variables $x_1, \ldots, x_N$: it will be shown that the non-cummutative analog to the linear partial differential operator D of (1.1) is the *substitution operator*

$$\mathcal{D} \equiv \mathcal{D}(f) = (f_1 \downarrow x_1) + \cdots + (f_N \downarrow x_N) \ , \tag{1.13}$$

on monomials defined as

$$(f_k \downarrow x_k)(x_{h_1} \cdot \cdots \cdot x_{h_m}) := \sum_{k=1}^{m} (f_k \overset{k}{\hookrightarrow} x_h) \quad \text{with} \tag{1.14}$$

$$(1.15) \qquad (f_\nu \overset{k}{\hookrightarrow} x_h) := \begin{cases} 0 & \text{, if } k \neq h_k, \\ x_{h_1} \ldots x_{h_{k-1}}\ f_\nu\ x_{h_{k+1}} \ldots x_{h_m} & \text{, if } k = h_k \end{cases}$$

and extended R-linearly to all power series of $R[[x_1, \ldots, x_N]]_{nc}$ (subscript 'nc' for 'non-commuting variables'). In other words: the operator $(f_\nu \downarrow x_k)$ replaces every occurence of x_k in a monomial by f_ν and sums up the resulting expressions. It is not hard to see that for power series in commuting variables the substitution operator $\mathcal{D}$ coincides with the differential operator D, where in particular: $\partial_{x_k} = (1 \downarrow x_k)$. Therefore we conclude that *in the context of formal power series substituition includes differentiation as a special case.* It will be shown in Section 4 that the non-commutative initial value problem (1.4) is solved by the *exponential substitution*

$$(1.16) \qquad e^{t\mathcal{D}}\bar{x} := e^{t\mathcal{D}}x|_{x=\bar{x}} \ .$$

On the other hand the Exponential Formula (1.10) also solves the IVP (1.4), if f is a homogeneous polynomial mapping in non-commuting variables, whence the family of equations (1.12) has the non-commutative analog

$$(1.17) \qquad \mathcal{D}^n(x) = \mu^n(x) \quad \text{for all } n \geq 0,$$

which we will prove constructively. The constructive proofs of (1.12) and (1.17) both have an algebraic part (Sections 3 and 4) and a combinatorial part (Sections 6 and 7), where the combinatorial part is based on a combinatorial description of the structure of the Lie series (1.2), the exponential substitution (1.15), and the Exponential Formula (1.11) in terms of certain *rooted trees* (Section 5 ff.). Different kinds of rooted trees will be seen to provide different perspectives. We describe not only their respective significance for the understanding of the powers D^n, $\mathcal{D}^n$, and μ^n, but also address the following questions:

Counting (the number of rooted trees on n vertices of a given kind), *statistics* (finer counting properties related to levels, branching, leafs, etc.), *enumeration* (the concrete description of all rooted trees of a given kind), and *codes* (linear strings or words of integers). Codes have been systematicaly investigated for the first time by Read [R1,R3] as a tool for the enumeration of trees. We use them mainly as a substitute for a space consuming drawing of a tree and as a description revealing structural features.

Moreover, the relation between different kinds of rooted trees is investigated with emphasis on the properties of *projection classes*: the 'projection' of rooted trees of type A onto rooted trees of type B with less structure induces a partition of A into equivalence classes, where two trees of type A are equivalent iff they have the same image under projection. In Section 7 a new recursive algorithm for the enumeration of rooted trees is described, which works without comparisons and normal forms and which allows the simultaneous computation of certain *projection numbers* (=cardinality of projection classes). In Section 8 we introduce several new statistics for monotonely labeled rooted trees, which extend the well known permutation statistics.

2. Basic theory of Lie series

The R-lineararity of the operators D^n for all $n \geq 0$ implies the R-linearity of the exponential operator e^{tD}. From

$$D(g_1(x)g_2(x)) = D(g_1(x))g_2(x) + g_1(x)D(g_2(x)) \quad \text{for all } g_1, g_2 \in R[[x_1, \dots, x_N]] \tag{2.1}$$

it follows by induction that

$$D^n(g_1(x)g_2(x)) = \sum_{k=0}^{n} \binom{n}{k} D^k(g_1(x))\ D^{n-k}(g_2(x)) \quad \text{for all } n \geq 0. \tag{2.2}$$

An easy calculation [G, p.14] then establishes

$$e^{tD}(g_1(x)g_2(x)) = (e^{tD}g_1(x))\ (e^{tD}g_2(x)), \tag{2.3}$$

which implies [G, (2.11)]

$$e^{tD}P(g_1(x), \dots, g_p(x)) = P(e^{tD}g_1(x), \dots, e^{tD}g_p(x)) \quad \text{for all } P \in R[y_1, \dots, y_p],$$

and in the (formal) limit

$$e^{tD}F(g_1(x), \dots, g_p(x)) = F(e^{tD}g_1(x), \dots, e^{tD}g_p(x)) \quad \text{for all } F \in R[[y_1, \dots, y_p]]. \tag{2.4}$$

Of course the last equation is true also for $F \in \mathbb{C}\{x_1, \dots, x_N\}$ ([G, Satz 6]). Set

$$e^{tD}\bar{x} := e^{tD}x|_{x=\bar{x}} \ \text{ for } \bar{x} = (\bar{x}_1, \dots, \bar{x}_N)^* \in R^N,$$

i.e., $\exp(tD)$ is applied componentwise to the variables x_j, and then evaluated at $\bar{x}$. A usefull notation is

$$X_j = \varphi_j(x;t) := e^{tD}x_j\ , \tag{2.5}$$

whence $\varphi_j(x;0) = x_j$ and as a special case of (2.4) one gets

$$e^{tD}F(x) = F(X)\ . \tag{2.6}$$

Example 2.1. For $D = h_1\partial_1 + \cdots + h_N\partial_N$ with constants $h_1, \dots, h_N \in R$ one has: $D^0x_j = x_j$, $Dx_j = h_j$, and $D^nx_j = 0$ for $n \geq 2$, which yields $e^{tD}x = x + th$. Setting $t = 1$ then gives for all $F \in R[[y_1, \dots, y_N]]$ the (formal) Taylor formula

$$F(x+h) = F(e^Dx) \overset{(2.5)}{=} e^DF(x) \overset{(1.2)}{=} \sum_{n=0}^{\infty} \frac{t^n}{n!}(h \cdot \nabla)^n F(x)\ .$$

Example 2.2. For $f(x) = Ax$ with $A = (a_{ij})$ an $N \times N$-matrix over R one computes that

$$D \equiv D(A) = \sum_{i=1}^{N} \left(\sum_{j=0}^{N} a_{ij} x_j \right) \partial_{x_j} = \nabla^* A x = (\nabla^* A x)^* = x^* A^* \nabla \ .$$

Since $D(f)\ x = f(x)$ for all operators (1.1), one has

$$D(A)x = Ax \ . \tag{2.7}$$

And since ∇x^* is equal to the unit matrix E_N one furthermore computes

$$D^2(A) = x^* A^* \nabla x^* A^* \nabla = x^* (A^*)^2 \nabla = x^* (A^2)^* \nabla = D(A^2)$$

and in general

$$D^n(A) = D(A^n) \quad \text{for all } n \geq 1.$$

Therefore the Lie series for $D \equiv D(A)$ has the usual form

$$e^{tD} x = \sum_{n=0}^{\infty} \frac{t^n}{n!} D(A^n) x \overset{(2.7)}{=} \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} A^n \right) x$$

of a matrix exponentiation. (The Exponential Formula (1.11) for linear μ is immeadiately seen to be the matrix exponentiation and is therefore the genuine generalization of the matrix exponentiation for polynomial vector fields.)

Theorem 2.3. [G, Satz 7] *The Lie series* $e^{tD}\bar{x}$ *with* D *given by* (1.1) *solves the IVP* (1.4) *for commuting variables* $x_1, \ldots, x_N$. *If* f *is (real) analytic, then the Lie series* $e^{tD}x$ *is (real) analytic in* $x_1, \ldots, x_N, t$.

Proof. Gröbner's proof in case of $f_1, \ldots, f_N \in \mathbb{C}\{x_1, \ldots, x_N\}$ generalizes immediately to the formal case:

$$\begin{aligned}
\dot{X} &\overset{(2.5)}{=} \frac{d}{dt} \left(e^{tD} x \right) \\
&\overset{(1.2)}{=} \frac{d}{dt} \left(\sum_{n=0}^{\infty} \frac{t^n}{n!} D^n x \right) = \sum_{n=1}^{\infty} \frac{t^{n-1}}{(n-1)!} D^n x = \sum_{n=0}^{\infty} \frac{t^n}{n!} D^{n+1} x \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} D^n (Dx) \overset{(1.2)}{=} e^{tD} (Dx) = e^{tD} f(x) \\
&\overset{(2.6)}{=} f(X) \ .
\end{aligned}$$

For a proof of analyticity see [G, Satz 2]. $\square$

In case of a polynomial vector field f the convergence of the component power series of $e^{tD}x$ can be derived from the Exponential Formula (1.11), too (cf. [W, Sec.1.5]).

Let

$$D_j = f_{,j} \cdot \nabla = f_{1j}\partial_{x_1} + \cdots + f_{Nj}\partial_{x_N} \quad (j = 1, \ldots, M) \tag{2.8}$$

be M operators of the form (1.1). If $t_1, \ldots, t_M$ are commuting variables, then replacing tD by $t_1D_1 + \cdots + t_MD_M$ in the definition (1.2) of a Lie series leads to a *multi-dimensional Lie series*, for which the analogs of formulas (2.1) – (2.5) remain true. If the operators D_j commute pairwise, then the multi-dimensional Lie operator can be expressed as

$$e^{t_1D_1+\cdots+t_MD_M} = \sum_{n=0}^{\infty} \sum_{\nu \in A(M,n)} \frac{t^\nu}{\nu!} D^\nu \ , \tag{2.9}$$

where ν is a multi-index in

$$A(M,n) = \{\nu = (\nu_1, \ldots, \nu_M) \mid \nu_1, \ldots, \nu_M \geq 0,\ |\nu| = \nu_1 + \cdots + \nu_M = n\} \ , \tag{2.10}$$

and $\nu! = \nu_1! \ldots \nu_M!$, $t^\nu = t_1^{\nu_1} \ldots t_M^{\nu_M}$, $D^\nu = D_1^{\nu_1} \ldots D_M^{\nu_M}$ in multi-index notation. An elementary calculation shows [G, (3.6′)] that the operators D_j commute iff

$$[D_j, D_k] = 0 \quad \Longleftrightarrow \quad \forall i = 1, \ldots, N : \sum_{h=1}^{N} \left(f_{jh} \frac{\partial f_{ki}}{\partial x_h} - f_{kh} \frac{\partial f_{ji}}{\partial x_h} \right) = 0 \ . \tag{2.11}$$

For $x \equiv x(t_1, \ldots, t_M)$ consider an (multi-dimensional) IVP

$$\begin{aligned} \frac{\partial x_i}{\partial t_k} &= f_{ki}(x) \qquad (i = 1, \ldots, N;\ k = 1, \ldots, M) \\ x_i(0) &:= x_i(t)|_{t_1 = \cdots = t_M = 0} = \bar{x}_i \in R \ , \end{aligned} \tag{2.12}$$

where the x_i (or f_{ki}) satisfy the compatibility conditions

$$\frac{\partial^2 x_i}{\partial t_k \partial t_j} = \frac{\partial^2 x_i}{\partial t_j \partial t_k} \ . \tag{2.13}$$

The computation

$$\frac{\partial^2 x_i}{\partial t_k \partial t_j} = \frac{\partial f_{ki}(x)}{\partial t_j} = \sum_{h=1}^{N} \frac{\partial f_{ki}(x)}{\partial x_h} \frac{\partial x_h}{\partial t_j} = \sum_{h=1}^{N} f_{jh} \frac{\partial f_{ki}}{\partial x_h}$$

then shows that (2.13) and (2.11) are equivalent.

Theorem 2.4. [G, Satz 20] *The multi-dimensional Lie series with operators D_j (2.8)*

$$e^{t_1D_1+\cdots+t_MD_M}\ \bar{x} := e^{t_1D_1+\cdots+t_MD_M} x|_{x=\bar{x}}$$

solves the IVP (2.12-13)*. If f is (real) analytic, then the Lie series* $\exp(t_1D_1 + \cdots + t_MD_M)\ x$ *is (real) analytic in $x_1, \ldots, x_N, t_1, \ldots, t_M$.*

Proof. Analogous to the proof of Thm.2.3, where (2.5) is replaced by

$$X_j = \varphi_j(x;t) := e^{t_1 D_1 + \cdots + t_M D_M} \; x_j| . \tag{2.14}$$

□

Note that the multi-dimensional Lie series (2.9) is essentially a Lie series of the form (1.2) with

$$tD = \sum_{i=1}^{N} \left(\sum_{j=1}^{N} t_j f_{ij} \right) \partial_{x_i} \ ,$$

i.e. with coefficients in $R[t_1, \ldots, t_M]$ instead of just R.

We describe next Gröbner's application of multi-dimensional Lie series to the inversion of a mapping of power series.

Corollary 2.5. [G, Satz 23] *Let $F : R^N \longrightarrow R^N$ be a mapping with components*

$$y_i = F_i(x_1, \ldots, x_N) \in R[[x_1, \ldots, x_N]] \quad (i = 1, \ldots, N) \ ,$$

which is invertible in the neighborhood of some point

$$\bar{y} = F(\bar{x}) \in R^N \ ,$$

i.e., the Jacobian J_F has non-vanishing determinant in $\bar{x}$. Then

$$J_F^{-1} = (f_{ij}) \tag{2.15}$$

for certain polynomials $f_{ij} \in R[[x_1, \ldots, x_N]]$ and for operators D_j as in (2.8) *the inversion of the mapping F is given locally by:*

$$x_i = e^{(y_1 - \bar{y}_1)D_1 + \cdots + (y_M - \bar{y}_M)D_M} \bar{x}_i = \sum_{n=0}^{\infty} \sum_{\nu \in A(M,n)} \frac{(y - \bar{y})^\nu}{\nu!} D^\nu \bar{x}_i \ . \tag{2.16}$$

Proof. Observe first that (2.15) is equivalent to the system

$$\frac{\partial x_i}{\partial y_j} = f_{ij}(x) \quad (i, j = 1, \ldots, N)$$

with $x \equiv x(y_1, \ldots, y_N)$ and $x(0) = \bar{x}$. An elementary computation shows that the D_j's commute (cf. [G, p.52]). The result follows immediately from Theorem 2.4 upon setting $t_i = y_i - \bar{y}_i$. □

In the one-dimensional case $N = 1$ Corollary 2.5 spezializes to

Corollary 2.6. [G, Zusatz zu Satz 23] *Let $y = F(x) \in R[[x]]$ be invertible in the neighborhood of some point $\bar{y} = F(\bar{x}) \in R$ with $f = (F')^{-1} \in R[[x]]$. Then the inversion of the mapping F is given locally by the Lie series*

$$x = e^{(y-\bar{y})D}\bar{x} = \sum_{n=0}^{\infty} \sum_{\nu \in A(M,n)} \frac{(y-\bar{y})^{\nu}}{\nu!} D^{\nu}\bar{x} \ , \tag{2.17}$$

whith $D = f(x)\ d/dx$.

As a possible future application of multi-dimensional Lie series and in particular Corollary 2.5 we mention here the *Jacobian Conjecture* of O.H. Keller (1939) (see [BCW] for a historical survey):

> *Let $F : \mathbb{C} \longrightarrow \mathbb{C}$ be a polynomial mapping with* $\det J_F \in \mathbb{C} \setminus \{0\}$*, then the inverse mapping $F^{[-1]}$ is polynomial, too.*

In fact, by Lagrange's inversion formula for matrices the assumption implies that J_F^{-1} has polynomial entries f_{ij}, whence the hard part of the Jacobian Conjecture is to show that (2.16) has only finitely many nonvanishing summands. The approach to the Jacobian Conjecture via labelled rooted trees of [BCW, Sec.III] is an easy consequence of the combinatorial interpretation of Lie series in Section 6. (See also Zeilberger's lovely paper [Z] for another combinatorial approach to the Jacobian Conjecture.)

Gröbner [G] gives further applications of the Lie series to the solution of first order partial differential equations, the parametrization of affine varieties, Abelian integrals, and in particular the n-body problem of celestial mechanics. See also the papers [Sb1,Sb2] (and the references therein) for a wealth of further applications in physics.

3. The algebraic connection between μ and the operators D_j

As stated in the Introduction the constructive proof of the equations (1.12) has a combinatorial part (contained in Section 7 below) and an algebraic part, described in the present section (Lemma 3.1).

Subsequently we use the notations

$$\begin{aligned}
(a)_b :=& a(a-1)\dots(a-b+1) \quad \text{for } a,b \in \mathbb{N}_0 = \mathbb{N} \cup \{0\},\ [(a)_0 := 1], \\
\underline{a} :=& \{1,\dots,a\} \quad \text{for } a \in \mathbb{N}, \text{ and} \\
\underline{a}^{\underline{b}} :=& \{\nu = (\nu_1,\dots,\nu_b) \mid \nu_1,\dots,\nu_b \in \underline{a}\} \quad \text{for } a,b \in \mathbb{N}.
\end{aligned} \tag{3.1}$$

In order to connect the operators D_j (2.8) depending primarily on the dimension N, with the mapping μ depending primarily on the dedree m we first define a bijection

$\Phi : A(N,m) \longrightarrow K(m,N)$ between the sets

$$A(N,m) = \{\alpha = (\alpha_1, \ldots, \alpha_N) \mid \alpha_1, \ldots, \alpha_N \geq 0,\ |\alpha| = \alpha_1 + \cdots + \alpha_N = m\}$$
$$K(m,N) = \{k \in \underline{N}^{\underline{m}} \mid 1 \leq k_1 \leq \cdots \leq k_m \leq N\}$$

by

(3.2) $$\Phi(\alpha) := (\underbrace{1, \ldots, 1}_{\alpha_1}, \underbrace{2, \ldots, 2}_{\alpha_2}, \ldots, \underbrace{N, \ldots, N}_{\alpha_N}) = (k_1, \ldots, k_m) = k\ .$$

We will use $k(\alpha)$ and $\alpha(k)$ as shorthands for $\Phi(\alpha)$ and $\Phi^{-1}(k)$, respectively. Set moreover for any $k \in K(m,N)$:

(3.3) $$V_m(k) := \{\text{ different } m\text{-tuples formed from } k_1, \ldots, k_m\ \}\ .$$

Clearly,

$$|V_m(k)| = \binom{m}{\alpha(k)} := \frac{m!}{\alpha(k)!} = \frac{m!}{\alpha_1(k)! \ldots \alpha_N(k)!}\ ,$$

whence with $x^\alpha = x_1^{\alpha_1} \ldots x_N^{\alpha_N}$:

(3.4) $$\sum_{h \in V_m(k)} x_{h_1} \ldots x_{h_m} = \binom{m}{\alpha(k)} x^{\alpha(k)}\ .$$

For commuting variables $x_1, \ldots, x_N$ one can rewrite any f as given by (1.6) in the form

(3.5) $$f(x) = \sum_{\alpha \in A(N,m)} \binom{m}{\alpha} A_\alpha\ x^\alpha \quad (A_\alpha \in R^N)\ ,$$

since the rationals are contained in R. Using the notation

(3.6) $$\alpha(h) = (\alpha_1(h), \ldots, \alpha_N(h)) \text{ with } \alpha_i(h) := |\{j \mid h_j = i;\ j = 1, \ldots, M\}|$$

for any $h \in \underline{N}^{\underline{M}}$ $(M \leq m)$ then gives

(3.7) $$\begin{aligned} f(x) = &\sum_{k \in K(m,N)} \binom{m}{\alpha(k)} A_{\alpha(k)}\ x^{\alpha(k)} \\ &= \sum_{k \in K(m,N)} A_{\alpha(k)} \sum_{h \in V_m(k)} x_{h_1} \ldots x_{h_m} \overset{(3.4)}{=} \sum_{h \in \underline{N}^{\underline{m}}} A_{\alpha(h)}\ x^{\alpha(h)}\ . \end{aligned}$$

The mapping μ corresponding to this form of f by (1.5-6) is *commutative* in the following sense:

(3.8) $$\mu(e_{h_1} \otimes \ldots \otimes e_{h_m}) = \mu(e_{h_{\pi(1)}} \otimes \ldots \otimes e_{h_{\pi(m)}}) \text{ for all } \pi \in S_m,$$

i.e., for all permutations π of the numbers $1, \ldots, m$ (cf. [W, Example 1.1], the definition of commutativity there does not say what is meant).

Notice that the presentation (3.5) of f, which makes the corresponding μ commutative, is unique.

Lemma 3.1. *For* $1 \le j \le M \le m$ *let* $f_{,j} = (f_{1,j}, \ldots, f_{N,j})^* : R^N \longrightarrow R^N$ *be any polynomial mappings; let* μ *and* f *be as in* (1.5-6) *and* (3.5), *respectively, with* μ *in its unique commutative form. Then for*

$$\widetilde{D}_M := \sum_{\nu \in \underline{N}^M} f_{\nu_1,1}(x) \ldots f_{\nu_M,M}(x)\ \partial_{\nu_1} \ldots \partial_{\nu_M} \equiv \sum_{\nu \in \underline{N}^M} f_{(\nu)}\ \partial_\nu \tag{3.9}$$

one has

$$\widetilde{D}_M f(x) = (m)_M\ \mu(f_{,1}(x) \otimes \ldots \otimes f_{,M}(x) \otimes \otimes^{m-M} x)\ . \tag{3.10}$$

Proof. The multivariate analog to $(d/dx)^k(x^n) = (n)_k x^{n-k}$ is (with multi-index notation, $\nu \in K(M,N)$, and ∂_ν as in (3.8))

$$\partial_\nu x^\alpha = (\alpha)_{\alpha(\nu)} x^{\alpha - \alpha(\nu)}\ ,$$

Then one computes

$$\widetilde{D}_M x^\alpha \overset{(3.8)}{=} \sum_{\nu \in \underline{N}^M} f_{(\nu)}\ \partial_\nu x^\alpha = \sum_{\nu \in \underline{N}^M} f_{(\nu)}\ (\alpha)_{\alpha(\nu)}\ x^{\alpha - \alpha(\nu)}\ .$$

Application of this last result to $f(x)$ in the form (3.6) yields

$$\begin{aligned} \widetilde{D}_M f(x) &= \sum_{k \in K(m,N)} \binom{m}{\alpha(k)} A_{\alpha(k)}\ \widetilde{D}_M\ (x^{\alpha(k)}) \\ &= (m)_M \sum_{k \in K(m,N)} A_{\alpha(k)} \sum_{\nu \in \underline{N}^M} f_{(\nu)} \binom{m-M}{\alpha(k) - \alpha(\nu)} x^{\alpha(k) - \alpha(\nu)}\ , \end{aligned} \tag{3.11}$$

because

$$\binom{m}{\alpha} (\alpha)_{\alpha(\nu)} = \frac{m!}{\alpha!} \frac{\alpha!}{(\alpha - \alpha(\nu))!} = (m)_M \frac{(m-M)!}{(\alpha - \alpha(\nu))!} \equiv (m)_M \binom{m-M}{\alpha - \alpha(\nu)}\ .$$

By the R-linearity of μ it follows from (1.5) that for any vectors $v_1, \ldots, v_m \in R^N$ with $v_j = (v_j^h)_{h=1,\ldots,N}$ $(j = 1, \ldots, m)$:

$$\begin{aligned} \mu(v_1 \otimes \ldots \otimes v_m) &= \sum_{h_1,\ldots,h_m=1}^{N} v_1^{h_1} \cdot \ldots \cdot v_m^{h_m} \mu(e_{h_1} \otimes \ldots \otimes e_{h_m}) \\ &\overset{(1.5)}{=} \left(\sum_{h_1,\ldots,h_m=1}^{N} a^i_{h_1 \ldots h_m} v_1^{h_1} \cdot \ldots \cdot v_m^{h_m} \right)_{i=1,\ldots,N} = \sum_{h \in \underline{N}^m} a_h v_1^{h_1} \cdot \ldots \cdot v_m^{h_m} \quad (a_h \in R^N). \end{aligned} \tag{3.12}$$

This together with (3.7) implies for the corresponding commutative μ

$$\mu(f_{,1}(x)\otimes\ldots\otimes f_{,M}(x)\otimes\otimes^{m-M}x) = \sum_{k\in K(m,N)} A_{\alpha(k)} \sum_{h\in V_m(k)} f_{h_1,1}(x)\ldots f_{h_M,M}(x)\cdot x_{h_{M+1}}\ldots x_{h_m}\ .$$

Comparison of the last equation with (3.10) and (3.11) shows that the proof is complete, if for fixed $k\in K(m,N)$ the following is true:

(3.13)
$$\sum_{\nu\in \underline{N}^{\underline{M}}} f_{(\nu)}\binom{m-M}{\alpha(k)-\alpha(\nu)} x^{\alpha(k)-\alpha(\nu)} = \sum_{h\in V_m(k)} f_{h_1,1}(x)\ldots f_{h_M,M}(x)\cdot x_{h_{M+1}}\ldots x_{h_m}\ .$$

Take any $h\in V_m(k)$, i.e. $h=(h_1,\ldots,h_m)$ is a permutation of the entries of k, and set $\nu=(h_1,\ldots,h_M)\in \underline{N}^{\underline{M}}$. For the remaining $(m-M)$-tuple $(h_{M+1},\ldots,h_m)$ every one of the $\binom{m-M}{\alpha(k)-\alpha(\nu)}$ different orderings of its entries yields the same factor $x^{\alpha(k)-\alpha(\nu)}$. This shows that for every sumand on the r.h.s. of (3.13) there is exactly one summand on the l.h.s. of (3.13), if one takes into account the multiplicities concealed in the multinomial expression.

Conversely, taking any $\nu\in \underline{N}^{\underline{M}}$, such that $\alpha_i(\nu)\le\alpha_i(k)$ for all $i=1,\ldots,N$, and fixing any order of the variables in $x^{\alpha(k)-\alpha(\nu)}$, yields a unique $h\in V_m(k)$ with $(h_1,\ldots,h_M)=\nu$ and $x_{h_{M+1}}\ldots x_{h_m}=x^{\alpha(k)-\alpha(\nu)}$. This shows that for every sumand on the l.h.s. of (3.13) there is exactly one summand on the r.h.s. of (3.13), which completes the proofs of (3.13) and Lemma 3.1 . □

4. Substitution operators for non-commuting variables

The Exponential Formula $\exp(t\mu)\,\bar{x}$ solves the initial value problem

$$\dot{x}=f(x),\ x(0)=\bar{x}\in R^N\ , \tag{4.1}$$

where $f=(f_1,\ldots,f_N)^*$ and $f_1,\ldots,f_N\in R[[x_1,\ldots,x_N]]_{nc}$ are homogeneous of some degree m in the non-commuting variables $x_1,\ldots,x_N$. This is a consequence of the non-commutativity of the tensor product: since the constuctions (1.5) - (1.10) and the proof of the solution property (see [W, Sec.1.2]) rely only on the tensor product and some elementary algebra, the Exponential Formula is well defined for non-commuting variables and solves the IVP (4.1). There is however one subtle difference between the commutative and the non-commutative formalism of the Exponential Formula: if the arguments of some $\mu\in L(T_mR^N,R^N)$ as in (1.6) involve the non-commuting variables

$x_1, \ldots, x_N$ instead of just the scalars from R, then one *defines*

$$\mu(v_1 \otimes \ldots \otimes v_m) := \sum_{h \in \underline{N}^m} a_h \; v_1^{h_1} \cdot \ldots \cdot v_m^{h_m} \tag{4.2}$$

for vectors $v_j = (v_j^1, \ldots, v_j^N)^*$ instead of deriving it from R-linearity as in (3.12).

In case of non-commuting variables (as in case of commuting variables) the homogenity of the polynomial mapping f is no severe restriction of the applicability of the Exponential Formula. Although homogenization of f through the introduction of an additional variable, say z, is not unique in the case of non-commuting variables, we are interested only in solutions obtained from the homogenized system (– an equation $\dot{z} = 0$ is added to (4.1) –) upon setting $z = 1$. Hence all kinds of homogenizations lead to the same non-homogeneous result.

That homogenity of the polynomial mapping f is not an essential requirement in case of non-commuting variables can be seen also from the non-commutative version of the *convolution formula* described in [W, Sec.2.2], whoose construction we briefly discuss next: Let $x = (x_0, x_1, x_2, \ldots)$ and $y = (y_0, y_1, y_2, \ldots)$ be two sequences of non-commuting variables. Then as usual the convolution of the two is defined by

$$x * y = (x_0 y_0 \ , \ x_0 y_1 + x_1 y_0 \ , \ \ldots \ , \sum_{\substack{i,j \\ i+j=n}} x_i y_j \ , \ \ldots) \equiv (\ (x*y)_{:0}, (x*y)_{:1}, \ldots) \ .$$

This generalizes in the obvious way to the convolution of several sequences $x^{(1)}, \ldots, x^{(N)}$ of non-commuting variables, such that for every polynomial $p \in R[x_1, \ldots, x_N]_{nc}$ there exists a corresponding polynomial $\mathbf{p} \in \mathbf{R}[x^{(1)}, \ldots, x^{(N)}]_{nc}$ different from p only by the interpretation of multiplication as convolution and the interpretation of scalars $r \in R$ as sequences $(r, 0, 0, \ldots) \in \mathbf{R}$. If now the polynomial mapping f is replaced by the corresponding mapping $\mathbf{f}$ and $\mathbf{f}_{:n}$ denotes the n-th component-vector, then the local power series solution

$$\sum_{n=0}^{\infty} x_n t^n \qquad \text{with} \ \ x_n = (x_n^{(1)}, \ldots, x_n^{(N)})^* \tag{4.3}$$

of the non-commutative polynomial IVP (4.1) is given recursively by

$$x_0 = \bar{x} \ , \quad x_{n+1} = \frac{1}{n+1} \, \mathbf{f}_{:n} \ . \tag{4.4}$$

This convolution formula can be shown to be essentially equivalent to the Exponential Formula and can also easily be derived directly with the help of the non-commutative Cauchy product of power series. (The arguments given in [W] work equally well for commuting and non-commuting variables).

It remains therefore to look for a non-commutative analog of the Lie series and as stated already in Section 1 this analog is given by the exponentiation of the substitution operator $\mathcal{D}$. But first the question arises, why the linear partial differential operators D do not work for non-commuting variables $x_1, \dots, x_N$. Re-examination of the arguments leading to the solution property of the Lie series (Thorem 2.3) reveals that the crucial fact is that D is a derivation, i.e., satisfies (2.1). All the formulas (2.2-6) and the proof of Theorem 2.3 are based on this derivation property (and of course the peculiarities of the exponentiation). Now for non-commuting variables $x_1, \dots, x_N$ the partial derivatives ∂_{x_j} are derivations, which commute with each other, but even an operator $D = f(x)\partial_{x_j}$ is not a derivation: for $N > 1$ and $f, g_1, g_2 \in R[[x_1, \dots, x^N]]_{nc}$ one has in general

$$\begin{aligned} D(g_1(x)g_2(x)) = D(g_1(x))\ g_2(x) + f(x)\ g_1(x)\ \partial_{x_j}(g_2(x)) \\ \neq D(g_1(x))\ g_2(x) + g_1(x)\ D(g_2(x))\ . \end{aligned}$$

To the contrary the substitution operators $\mathcal{D}$ defined in (1.13-15) are derivations, and with the same arguments as in Section 2 one derives the non-commutative analog of Theorem 2.3:

Theorem 4.1. *The substitutional series* $e^{t\mathcal{D}}\bar{x} = e^{t\mathcal{D}}x|_{x=\bar{x}}$ *with*

$$\mathcal{D} \equiv \mathcal{D}(f) = (f_1 \downarrow x_1) + \dots + (f_N \downarrow x_N)$$

solves the non-commutative IVP (4.1).

For homogeneous polynomial mappings f the problem remains to prove/understand the equality

$$e^{t\mathcal{D}}x = \exp(t\mu)\ x\ ,$$

respectively, the equalities (1.17) between the powers of $\mathcal{D}$ and μ. This will be done by proving an analog of Lemma 3.1 (Lemma 4.2 below) and by investigating and comparing the tree structures underlying the respective powers (Section 6).

For $1 \leq M \leq m$ and $\nu \in \underline{N}^{\underline{M}}$ let $\alpha(\nu) \in A(N, M)$ be given by

$$\alpha_j(\nu) := |\{q \mid \nu_q = j\}|\ .$$

With

$$\overline{K}(M, m) := \{\rho \in \underline{N}^{\underline{M}} \mid 1 \leq \rho_1 < \dots < \rho_M \leq m\}$$

we define for any $h \in \underline{N}^{\underline{m}}$ the set

$$\begin{gathered} S_M(h) := \{\text{ subwords of } h \text{ of lenght } M\ \}\ ,\ \text{i.e.} \\ k \in S_M(h) :\Longleftrightarrow \exists \rho \in \overline{K}(M, m)\ \forall j \in \underline{M}\ :\ k_j = h_{\rho_j}\ . \end{gathered}$$

For every $h \in \underline{N}^{\underline{m}}$ set

$$S(k,h) := \{\rho \in \overline{K}(M,m) \mid k(h,\rho) := (h_{\rho_1}, \dots, h_{\rho_M}) = k\}$$

and let $s(k,h) := |S(k,h)|$. Clearly:

$$\sum_{k \in S_M(h)} s(k,h) = \binom{m}{M} \quad \text{and} \quad \biguplus_{k \in S_M(h)} S(k,h) = \overline{K}(M,m) \ . \tag{4.5}$$

We are now prepared to define the *higher order substitution operators*:

$$\begin{aligned} (f_\nu \downarrow x_\nu)(x_h) &\equiv (f_{\nu_1}, \dots, f_{\nu_M} \downarrow x_{\nu_1}, \dots, x_{\nu_M})(x_{h_1} \dots x_{h_m}) \\ &:= \alpha(\nu)! \sum_{k \in V_M(\nu)} (f_\nu \overset{k}{\hookrightarrow} x_h) \ , \end{aligned} \tag{4.6}$$

where $\alpha(\nu)!$ and $V_M(\nu)$ are given as in Section 3 and $(f_\nu \overset{k}{\hookrightarrow} x_h)$ means:

> substitute M factors of x_h by the M components of f_ν in accordence to k, i.e., $(f_\nu \overset{k}{\hookrightarrow} x_h) := 0$, if $k \notin S_M(h)$, or otherwise replace for $j = 1, \dots, M$ and all $\rho \in S(k,h)$ the variable x_{k_j} at place ρ_j by $f_{k_j,j}$.

Higher order substitution operators result from repeated first order substitutions, if the latter operators are *not* applied to the already substituted expressions. For example:

$$\begin{aligned} (f_1 \downarrow x_1)[(f_1 \downarrow x_1)(x_1 x_2 x_1)] &= (f_1 \downarrow x_1)(f_1 x_2 x_1 + x_1 x_2 f_1) \\ &= [(f_1 \downarrow x_1)(f_1)] x_2 x_1 + f_1 x_2 f_1 + f_1 x_2 f_1 + x_1 x_2 [(f_1 \downarrow x_1)(f_1)] \ , \end{aligned}$$

but

$$(f_1, f_1 \downarrow x_1, x_1)(x_1 x_2 x_1) = 2 f_1 x_2 f_1 \ .$$

It is not hard to see that for commuting variables

$$(f_\nu \downarrow x_\nu)(x_h) = f_\nu(\partial_\nu \ x_h)$$

and for $M > m$ one has in particular $S_M(h) = \emptyset$ and therefore $(f_\nu \downarrow x_\nu)(x_h) = 0$. Note moreover that $(f_\nu \downarrow x_\nu) = (f_{\pi(\nu)} \downarrow x_{\pi(\nu)})$ for all permutations $\pi \in S_M$ of the entries.

Lemma 4.2. *Let* $\mathrm{Inj}(\underline{M}, \underline{m})$ *be the set of injective mappings from* $\underline{M}$ *to* $\underline{m}$ *given as* M*-tuples* $k \in \underline{m}^{\underline{M}}$*. Let* μ *and* f *be as in* (1.5) *and* (1.6)*, respectively, and* $1 \le j \le M \le m$*. For any collection of polynomial mappings* $f_{,j} = (f_{1,j}, \dots, f_{N,j})^* : R^N \longrightarrow R^N$ *in non-commuting variables* $x_1, \dots, x_N$ *set*

$$\widetilde{\mathcal{D}}_M := \sum_{\nu \in \underline{N}^{\underline{M}}} (f_{\nu_1,1}(x) \dots f_{\nu_M,M}(x) \downarrow x_{\nu_1}, \dots, x_{\nu_M}) \equiv \sum_{\nu \in \underline{N}^{\underline{M}}} (f_{(\nu)} \downarrow x_\nu) \ . \tag{4.7}$$

Then

$$\widetilde{\mathcal{D}}_M f(x) = \sum_{k' \in \mathrm{Inj}(\underline{M}, \underline{m})} \mu(f^{[M]} \overset{k'}{\hookrightarrow} \otimes^m x) \ , \tag{4.8}$$

where $(f^{[M]} \overset{k'}{\hookrightarrow} \otimes^m x)$ *means:*

> *for* $j = 1, \ldots, M$ *substitute the factor* x *at place* k'_j *in* $\otimes^m x$ *by* $f_{,j}$*.*

Proof. One computes

$$\widetilde{\mathcal{D}}_M \overset{(1.6)}{=} \sum_{h \in \underline{N}^{\underline{m}}} a_h \ \widetilde{\mathcal{D}}_M(x_h)$$

and

$$\begin{aligned} \widetilde{\mathcal{D}}_M(x_h) &= \sum_{\nu \in \underline{N}^{\underline{M}}} (f_{(\nu)} \downarrow x_\nu)(x_h) \\ &= \sum_{\nu \in \underline{N}^{\underline{M}}} \alpha(\nu)! \sum_{k \in V_M(\nu)} (f_{(\nu)} \overset{k}{\hookrightarrow} x_h) \\ &= M! \sum_{k \in S_M(h)} (f_{(k)} \overset{k}{\hookrightarrow} x_h) \ , \end{aligned}$$

because for every $k \in V_M(\nu)$ one has: $k \in V_M(\nu')$ for all $\nu' \in V_M(\nu) \subset \underline{N}^{\underline{M}}$ and $|V_M(\nu)| = \binom{M}{\alpha(\nu)} = |V_M(k)|$.

On the other hand from the non-commutative analog (4.2) of (3.12) it follows that

$$\sum_{k' \in \mathrm{Inj}(\underline{M}, \underline{m})} \mu(f^{[M]} \overset{k'}{\hookrightarrow} \otimes^m x) = \sum_{h \in \underline{N}^{\underline{m}}} a_h \sum_{k' \in \mathrm{Inj}(\underline{M}, \underline{m})} \mu(f^{[M]} \overset{k'}{\hookrightarrow} x_h) \ ,$$

where $(f^{[M]} \overset{k'}{\hookrightarrow} x_h)$ means:

> for $j = 1, \ldots, M$ substitute the variable at place k'_j in x_h by the $h_{k'_j}$-th component of $f_{,j}$.

Since $(f^{[M]} \overset{k'}{\hookrightarrow} x_h)$ is invariant under all permutations of the components of k', one concludes that

$$\sum_{k' \in \mathrm{Inj}(\underline{M}, \underline{m})} \mu(f^{[M]} \overset{k'}{\hookrightarrow} x_h) = M! \sum_{k' \in \overline{K}(M,m)} \mu(f^{[M]} \overset{k'}{\hookrightarrow} x_h) \ .$$

Now the assertion of the lemma follows from the identity

$$\sum_{k \in S_M(h)} (f_{(k)} \overset{k}{\hookrightarrow} x_h) = \sum_{k' \in \overline{K}(M,m)} \mu(f^{[M]} \overset{k'}{\hookrightarrow} x_h) \ ,$$

which is true by (4.5) and the definitions. $\square$

5. Rooted trees: generalities

A (finite simple) *graph* G is a tuple $G = (V, E)$, where $V = V(G)$ is a finite set, whose members are called *vertices*, and $E = E(G)$ is a set of two-element subsets of V, whose members are called *egdes*. Two vertices $v_1, v_2 \in V$ are *adjacent*, if $\{v_1, v_2\} \in E$, and a *path* from v_1 to v_r in G is a sequence $(v_1, \dots, v_r)$ of vertices such that v_j and v_{j+1} are adjacent for $j = 1, \dots, r-1$. A path is called a *circuit*, if in addition the vertices $v_1, \dots, v_r$ are all different, r is greater than 2, and v_r and v_1 are adjacent. The *distance* $d(v, v')$ of two vertices v and v' in G is the minimal number r of edges for a path between v and v', where one sets $d(v, v') := \infty$, if there is no path in G between v and v'. G is called *connected*, if $d(v, v')$ is finite for all $v, v' \in V(G)$.

A *tree* T can be defined as a connected, circuit-free graph and a *rooted tree* as a tuple (T, v_0) — or T for short —, where T is a tree and v_0 a distinguished vertex of T called the *root*. The *trivial rooted tree* $T_0 := (\{v_0\}, \emptyset)$ has the root as its only vertex. A vertex v of a rooted tree T has *rank* k or is on *level* k, if $r(v) := d(v, v_0) = k$. Thus a rooted tree T can be identified in a natural way with a *ranked patially ordered set (poset)* (see e.g. [St1]), where $v < v'$ in T iff $d(v, v') = r(v') - r(v) > 0$. Similarly, a rooted tree can be identified in a natural way with a *directed graph* or *digraph*, where the edges are not just two-element subsets $\{v, v'\}$ of the vertex set V, but ordered pairs or *arrows* (v, v'), where v is called the out-vertex and v' the in-vertex of the arrow. Namely, (v, v') is an arrow for a rooted tree T, if $\{v, v'\} \in E(T)$ and $r(v') = r(v) + 1$. $A(T)$ is the set of arrows of the rooted tree T. The *out-degree* of a vertex $v \in V(T)$ then is the number of arrows departing from v:

$$\mathrm{out}(v) := |\{v' \in T \mid (v, v') \in A(T)\}| \ .$$

(The analogously defined in-degree is not very interesting for rooted trees, because every vertex except the root has in-degree 1.) Vertices v with $\mathrm{out}(v) = 0$ are called *leaves*, the set of leaves of some T is denoted by $L(T)$. A *twig* of a tree is a leaf together with the edge leading to it. The *hight of a leaf* is its rank and the *hight of a rooted tree* T is the maximal hight of a leaf:

$$h(v) := r(v) \ \text{ for all } v \in L(T), \ \ h(T) := \max_{v \in V(T)} d(v, v_0) = \max_{v \in L(T)} r(v) \ .$$

The number $|T| := |V(T)|$ of vertices of T will be called the *weight* of T. For any vertex v of a rooted tree T the *subtree*

$$\uparrow(v) \equiv \uparrow(v, T)$$

induced by v is the rooted tree with v as root, vertex set $\{v' \in V(T) \mid v' \geq v\}$ and the edges inherited from T. The *successor set* of a vertex v of a rooted tree T (patriarchaly called 'sons' of v) is the set

$$\mathrm{suc}(v) \equiv \mathrm{suc}(v, T) := \{v' \in T \mid (v, v') \in A(T)\} \ .$$

Clearly, $\mathrm{out}(v) = |\mathrm{suc}(v)|$. The analogously defined notion is that of a *precessor* (or 'father') of a vertex. Every vertex except the root has exactly one precessor and every vertex not a leaf has a successor. The *principal subtrees* $T_1, \ldots, T_s$ of a rooted tree T are the subtrees generated by the successors $v_1, \ldots, v_s$ of the root of T. The relationship between T and its principal subtrees will be written symbolically as

$$T = \lfloor T_1, \ldots, T_s \rfloor \ .$$

As usual we draw the vertices of a graph as points and the edges as lines joining adjacent edges. A rooted tree (T, v_0) is drawn with v_0 at the bottom and vertices of rank $i + 1$ above the vertices of rank i.

The set of all trees on n vertices will be denoted by $\mathbf{T}_n$, such that the set of all finite trees $\mathbf{T}$ is

$$\mathbf{T} = \biguplus_{n \in \mathbb{N}} \mathbf{T}_n \ .$$

Similarly, one has the sets $\mathbf{rT}$ and $\mathbf{rT}_n$ of finite rooted trees and rooted trees on n vertices. The notation $\mathbf{rT}^{\leqslant m}$ will be used for rooted trees, where $\mathrm{out}(v) \leq m$ for all occuring vertices v.

In general we will discuss not just rooted trees, but rooted trees with additional structure, e.g. *ordered* (= *planar*) trees. A rooted tree is ordered, if the successor sets $\mathrm{suc}(v)$ are linearly ordered for every $v \in V(T)$. Ordered rooted trees will be denoted by $\mathbf{OrT}$. For non-ordered trees we will introduce *normal forms*, where all vertices on the same level are ordered according to appropriate rules. Another way to enrich rooted trees is to attach *labels* (from $\mathbb{N}_0$) to their vertices and/or edges. The most important labelings of rooted trees used here are the linear extensions: for $T \in \mathbf{rT}_n$ a *linear extension* is is an injective mapping $e : V(T) \longrightarrow \{0, \ldots, n-1\}$, such that the ordering of the vertices in T is compatible with the linear order on the natural numbers:

$$\forall v, v' \in V(T) : \ v < v' \Longrightarrow e(v) < e(v') \ .$$

In particular: $e(v_0) = 0$ for every linear extension of T. In drawings of linear extensions we simply attach to every vertex v its label $e(v)$. The set of all linear extensions of rooted trees will be denoted by $\mathbf{MrT}$, where $\mathbf{M}$ stands for 'monotonely labeled'. Of course the notions of successor, out-degree, (principal) subtree, etc. apply accordingly for rooted trees with additional structure.

The following picture gives an overview about the most important kinds of trees discussed in the subsequent sections. Double-arrows are bijections, single arrows projections.

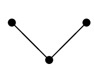
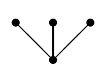
$$
\begin{tikzcd}
 & & & \mathbf{OrT} \arrow[dr] & & \\
\mathbf{S}m\mathbf{rT} \arrow[r, leftrightarrow] & \mathbf{M}m\mathbf{rT} \arrow[r, leftrightarrow] & \mathbf{ME}_m\mathbf{rT} \arrow[ur] \arrow[dr] & & \mathbf{rT} \arrow[r] & \mathbf{T} \\
 & & & \mathbf{MrT} \arrow[ur] & &
\end{tikzcd}
$$

The two leftmost types of rooted trees (discussed in Section 6 below) are m-nary trees, i.e. every non-leaf of such a tree has either out-degree 1 or m, where m is a fixed natural number ≥ 2. In addition there may be labelings.

Subsequently we will discuss the significance of different types of rooted trees for the structural understanding of the Lie series, the exponential substitution and the Exponential Formula and complete the constructive proofs of the identities (1.12) and (1.17). But as already explained in the Introduction we also discuss the counting, enumeration and statistics of trees as well as their codes and their projection properties.

6. m-nary rooted trees, the Exponential Formula and the exponential substitution

For $m \geq 2$ the set $\mathbf{S}m\mathbf{rT}_n$ of *sparse m-nary rooted trees of hight $h(T) = n$* contains all ordered rooted trees T, for which all leafs have rank $h(T) = n$. In addition all vertices on levels $j < n$ have out-degree 1 with the exception of exactly one vertex, which has out-degree m. Alternatively, one can describe the sets $\mathbf{S}m\mathbf{rT}_n$ as follows: Let $\mathbf{S}m\mathbf{rT}_0$ be the set containing only the trivial rooted tree and let $\mathbf{S}m\mathbf{rT}_1$ be the set containing only the *m-bush*, which is the unique rooted tree in $\mathbf{rT}_{m+1}$ of hight 1, i.e. for

$m = 2$:  $m = 3$:  etc.

Assume now that the set $\mathbf{S}m\mathbf{rT}_n$ is already constructed. Then one gets all the elements of $\mathbf{S}m\mathbf{rT}_{n+1}$ by taking any $T \in \mathbf{S}m\mathbf{rT}_n$, selecting a leaf $v \in L(T)$, and attaching an m-bush to v and a twig to every other leaf of T. From this construction it is immediate that

$$|\mathbf{S}m\mathbf{rT}_n| = n!_{m-1} \ , \tag{6.1}$$

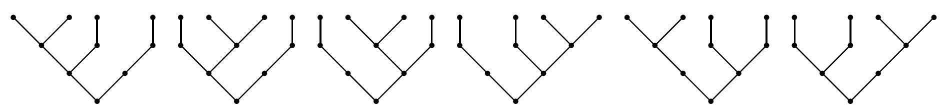
where for every $n \in \mathbb{N}_0$ and $k \in \mathbb{N}$ one defines ([W, Sec.1.5]) the *generalized factorials* $n!_k$ recursively by

$$(n+1)!_k := n!_k \cdot (1+nk) \ , \quad 0!_k := 1 \ . \tag{6.2}$$

Of course: $n!_1 = n!$ and from $(\nu-1)\ (m-1) < (\nu-1)\ (m-1)+1 \leq \nu\ (m-1)$ one gets the estimate

$$(n-1)!\ (m-1)^{n-1} < n!_{m-1} \leq n!\ (m-1)^n \quad \text{for } n > 0 \text{ und } m \geq 2.$$

Example 6.1. For $m=2, n=3$ the 6 elements of $\mathbf{S}2\mathbf{rT}_3$ are



The above recursive construction of the m-nary rooted trees in $\mathbf{S}m\mathbf{rT}_n$ leads not only to formula (6.1), but also to a natural representation of the trees by L-*codes*: there is a natural bijection between the sets $\mathbf{S}m\mathbf{rT}_n$ and

$$L(m,n) := \{l = (l_{n-1}, \ldots, l_0) \mid \nu(m-1) \geq l_\nu \geq 0\} \ ,$$

the latter having cardinality $n!_{m-1}$: Since $T \in \mathbf{S}m\mathbf{rT}_n$ is ordered and has exactly $\nu(m-1)+1 \overset{(1.7)}{=} [\nu]$ vertices on each level ν, one counts from left to right beginning with 0, where on level ν in T the m-bush is attached. This number gives entry l_ν of the L-code $l(T)$ of T. Of course the trees in $\mathbf{S}m\mathbf{rT}_n$ are easily reconstructed from the elements of $L(m,n)$. The L-codes in $L(2,3)$ of the trees in Example 6.1 above are from left to right:

$$(0,0,0),\ (1,0,0),\ (1,1,0),\ (2,1,0),\ (0,1,0),\ (2,0,0) \ .$$

The significance now of the sparse m-nary rooted trees for the understanding of the Exponential Formula is that they reflect exactly the structure underlying the powers of a $\mu \in L(T_m R^N, R^N)$: the m-to-1 mapping μ can be represented pictorially by an m-bush, where the m leafs stand for the m input vectors and the root for the image of the input under μ. The identity mappings correspond to the arrows (v, v') with $\operatorname{out}(v) = 1$, such that the recursive construction of the sets $\mathbf{S}m\mathbf{rT}_n$ corresponds exactly to the building up of the powers μ^n, or, in terms of the sets $L(m,n)$:

Proposition 6.2. *For $\mu \in L(T_m R^N, R^N)$ the powers μ^n for $n > 0$ can be represented as sums over or the L-codes $l \in L(m,n)$:*

$$\mu^n = \sum_{l \in L(m,n)} l(\mu) \ , \quad \textit{where}$$
$$l(\mu) := d^{\,l_0}_{\mu,[1]} \circ \cdots \circ d^{\,l_{n-1}}_{\mu,[n]} \circ \delta_{[n]} \ , \quad \textit{and}$$
$$d^{\,l_\nu}_{\mu,[\nu+1]} := \otimes^{l_\nu} id \otimes \mu \otimes \otimes^{[\nu]-l_\nu-1} id$$

Proof. Immediate from (1.7-10) and the linearity of the tensor product. □

The set $m\mathbf{rT}_n$ of *m-nary rooted trees on n m-bushes* is the set of all ordered rooted trees, such that every vertex has either out-degree m or 0 and that the number of vertices v with $\mathrm{out}(v) = m$ is n. The set $\mathbf{M}m\mathbf{rT}_n$ of *monotonely labeled m-nary (ordered) rooted trees on n m-bushes* is the set of all trees $T \in m\mathbf{rT}_n$, which in addition have a labeling of the n m-bushes by the numbers $\{0, \ldots, n-1\}$, such that the ordering of the m-bushes in T is compatible with the natural linear order on the integers. The sets $\mathbf{M}m\mathbf{rT}_n$ are in bijective correspondence to the sets $\mathbf{S}m\mathbf{rT}_n$: for any $T \in \mathbf{S}m\mathbf{rT}_n$ label the root of an m-bush by its rank and contract all arrows (v, v') with $\mathrm{out}(v) = 1$, such that v and v' are identified. (Clearly, every $T \in \mathbf{M}m\mathbf{rT}_n$ contains all information necessary for the construction of the corresponding tree in $\mathbf{S}m\mathbf{rT}_n$.)

In terms of the powers μ^n the contraction of the arrows with $\mathrm{out}(v) = 1$ to a point makes sense, because exactly these arrows correspond to the identity mappings, which do not alter the input. (But it would be very cumbersome to express algebraically the powers μ^n right from the outset without the use of the identities.)

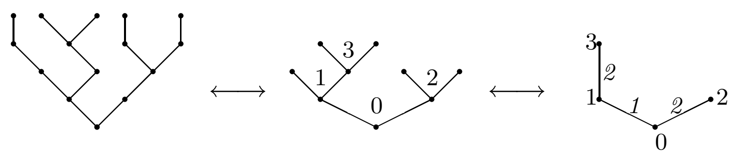
The set $\mathbf{ME}_m\mathbf{rT}_n$ of *monotonely labeled m-nary rooted trees on n vertices with edge-labels in $\underline{m}$* is the set of all (non-ordered) rooted trees, where the vertices are monotonely labeled with $0, \ldots, n-1$ (as explained in Section 5) and the labels of the edges from $\underline{m}$ are restricted only by the condition that the out-going edges of a fixed vertex have different labels. Again the sets $\mathbf{ME}_m\mathbf{rT}_n$ are in bijective correspondence to the sets $\mathbf{M}m\mathbf{rT}_n$: for any $T \in \mathbf{M}m\mathbf{rT}_n$ replace every m-bush by a vertex with the same label and join the new vertices by edges according to the way the m-bushes are stacked upon each other, i.e., if in $\mathbf{M}m\mathbf{rT}_n$ the m-bush b' has as its root a leaf of the m-bush b, then in $\mathbf{ME}_m\mathbf{rT}_n$ the vertex for b' is joined to the vertex for b by an upward leading edge. In addition the number of the leaf of b in $\mathbf{M}m\mathbf{rT}_n$ where b' has its root — the leafs of every m-bush in T are numbered $1, \ldots, m$ from left to right — is recorded by the edge label in $\mathbf{ME}_m\mathbf{rT}_n$. (Clearly, every $T \in \mathbf{ME}_m\mathbf{rT}_n$ contains all information necessary for the construction of the corresponding tree in $\mathbf{M}m\mathbf{rT}_n$.)

Although one has a more compact pictorial representation of the sparse m-nary trees now, the new representation does not add (so far) any new insight on the structure of the powers μ^n. But it provides the information necessary for the proof (below in this section) of the equalities (1.17), and it will be a useful step for the proof of the equalities (1.12) between the powers of μ and D (in the next section). Notice, that for non-commuting variables it is essential to record with the help of the edge labels, on which place μ has a certain input.

Example 6.3. As an illustration for the bijective corresponences

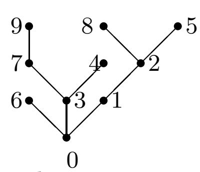
$$\mathbf{S}m\mathbf{rT}_n \quad\longleftrightarrow\quad \mathbf{M}m\mathbf{rT}_n \quad\longleftrightarrow\quad \mathbf{ME}_m\mathbf{rT}_n$$

we give an example for $m = 2$ and $n = 4$:



We describe next the *two-row code* for the trees in $\mathbf{ME}_m\mathbf{rT}_n$. The first row contains a permutation of the vertex labels $1, \ldots, n-1$ and the second row contains the edge labels from $\underline{m}$, namely, the label of the arrow leading up to j, where we identify a vertex v of T by its label j. The permutation of the first row is the *walk-around code* (wa-code) — this idea goes back at least to [dBM], see also [R3] —, which specifies both the information about the underlying rooted tree and the monotone vertex labeling. To determine the wa-code of some tree $T' \in \mathbf{ME}_m\mathbf{rT}_n$ we temporarily remove or ignore the edge labels and investigate only the underlying monotonely labeled tree $T \in \mathbf{MrT}_n$. This tree has to be put in a planar *normal form* first (— the original T' is not planar —) by ordering every set $\mathrm{suc}(v)$ such that the vertex labels decrease from left to right. For this planar normal form of T one writes down in a sequence every newly appearing vertex label while "walking around" the tree in clockwise direction beginning with the root (always having the lines of the tree right hand).

Example 6.4.

$T =$  $\in \mathbf{MrT}_{10}$

is in normal form and has the wa-code 6 3 7 9 4 1 2 8 5.

From the wa-code it is easy to reconstruct the monotonely labeled rooted tree $T \in \mathbf{MrT}_n$ by subdividing the wa-code into *sections* as follows: the first section begins with the first number and ends before the next number smaller than the first. To find the next section repeat the procedure for the remaining wa-code, etc. . In other words: the beginnings of the sections are given by the unique maximal decreasing subsequence of the wa-code, which begins with the first number. The sections of the wa-code 637941285, e.g., are 6, 3794, and 1285. Clearly these sections correspond to the principal subtrees generated by the successors of the root, and a repetition of the sectioning gives the higher order sections (= subtrees with roots on higher levels). Notice, that with the sectioning method one can determine recursively not only the sections (= subtrees) of the wa-code, but also the levels of all vertices (= roots of the subtrees). For example, sectioning and subsectioning of the wa-code 637941285 gives the levels (written below)

$$\begin{matrix} 6 & 3 & 7 & 9 & 4 & 1 & 2 & 8 & 5 \\ 1 & 1 & & & & 1 & & & \end{matrix} \longrightarrow \begin{matrix} 6 & 3 & 7 & 9 & 4 & 1 & 2 & 8 & 5 \\ 1 & 1 & 2 & & 2 & 1 & 2 & & \end{matrix}$$

$$\longrightarrow \begin{matrix} 6 & 3 & 7 & 9 & 4 & 1 & 2 & 8 & 5 \\ 1 & 1 & 2 & 3 & 2 & 1 & 2 & 3 & 3 \end{matrix} .$$

We briefly indicate how the two-row code of a $T' \in \mathbf{ME}_m\mathbf{rT}_n$ can be computed from the L-code of the corresponding tree $T \in \mathbf{S}m\mathbf{rT}_n$ (without sketching the trees):

From the L-code $l = (l_{n-1}, \ldots, l_0) \in L(m, n)$ one constructs first the *parenthesis code* as follows: for l_0 write ${}_0(0, \ldots, 0)$ with m zeros, then replace the l_1-th zero (counted 0, 1, ... from left to right) by ${}_1(0, \ldots, 0)$ (again with m zeros), then replace the l_2-th zero by ${}_2(0, \ldots, 0)$, and so on. Clearly, for the parenthesis code the content of a parenthesis ${}_j(\ \ldots\ )$ is a faithful image of the subtree $\uparrow(v, T)$ of the vertex v with label j. Hence all information necessary for the two-row code of T' can be extracted. For example the tree $T \in \mathbf{S2rT}_4$ of Example 6.3 has L-code $(1, 2, 0, 0)$ and parenthesis code ${}_0({}_1(0,\ {}_3(0, 0)),\ {}_2(0, 0))$.

We are now in a position to complete the constructive proof of the equalities

$$\mathcal{D}^n(x) = \mu^n(x) \qquad (n \geq 0) \tag{6.3}$$

for non-commuting variables $x_1, \ldots, x_N$. In fact we will show an even stronger assertion:

Definition 6.5. *Let $f : R^N \longrightarrow R^N$ be a polynomial mapping with $f = (f_1, \ldots, f_N)^*$ and $f_1, \ldots, f_N \in R[[x_1, \ldots, x^N]]_{nc}$. Then for $T \in \mathbf{MrT}_n$ and $\mathcal{D}$ as in* (1.13) *the substitution operator $\mathcal{D}^T$ is defined recursively as*

$$\begin{aligned}
\mathcal{D}^T &:= \mathcal{D}^T_{\uparrow(0)} \\
\mathcal{D}^T_{\uparrow(j)} &:= \sum_{\nu \in \underline{N}^{\underline{M}}} (f^T_{\nu_1,k_1}, \ldots, f^T_{\nu_M,k_M} \downarrow x_\nu) \quad , \textit{ if } suc(j) = \{k_1, \ldots, k_M\} \neq \emptyset \\
f^T_{p,q}(x) &:= \begin{cases} f_p(x) & , \textit{ if } q \textit{ is a leaf of } T \\ \mathcal{D}^T_{\uparrow(q)} f_p(x) & , \textit{ otherwise.} \end{cases}
\end{aligned}$$

In other words, to every vertex of the tree $T \in \mathbf{MrT}_n$ one associates an N-dimensional vector of polynomials: to every leaf of T the original mapping f and then to every other vertex v in a recursive fashion the result of the componentwise application of the operator $\widetilde{\mathcal{D}}_M$ from Lemma 4.2 to $f(x)$, where the arguments $f_{,1}, \ldots, f_{,M}$ of $\widetilde{\mathcal{D}}_M$ are the M mappings already associated to the M successors of v.

Theorem 6.6. *Let f, T, and $\mathcal{D}^T$ be given as in the definition above, where in addition f is homogeneous of degree m, and let μ be associated to f by* (1.5-6). *Let moreover $\mathbf{E}(T)$ denote the set of all trees in $\mathbf{ME}_m\mathbf{rT}_n$, which reduce to T upon forgetting the edge labels. And finally let $l^{T'}$ be the L-code corresponding to a $T' \in \mathbf{ME}_m\mathbf{rT}_n$. Then with the notation of Proposition* 6.2*:*

$$\mathcal{D}^T(x) = \sum_{T' \in \mathbf{E}(T)} l^{T'}(\mu)\ (x)\ . \tag{6.4}$$

Proof. From the description of $\mathcal{D}^T$ given after its definition it is clear that in every step on can replace $\widetilde{\mathcal{D}}_M f(x)$ by the right hand side

$$\sum_{k \in \mathrm{Inj}(\underline{M},\underline{m})} \mu(f^{[M]} \overset{k}{\hookrightarrow} \otimes^m x)$$

of formula (4.8) of Lemma 4.2. But by the definition of $(f^{[M]} \overset{k}{\hookrightarrow} \otimes^m x)$ this means to replace the factor x at place k_j in $\otimes^m x$ by $f_{,j}$, which in terms of trees means to have the edge label k_j for an arrow leading to the input-vertex $f_{,j}$. Since for every vertex v of T with M successors the possible M-tuples k of edge labels vary over all elements of $\mathrm{Inj}(\underline{M}, \underline{m})$ the assertion (6.4) follows. □

Proof. of (6.3): From Proposition 6.2, the obvious set partition

$$\mathbf{ME}_m\mathbf{rT}_n = \biguplus_{T \in \mathbf{MrT}_n} \mathbf{E}(T)\ ,$$

and the bijection between $\mathbf{ME}_m\mathbf{rT}_n$ and $L(m,n)$ one concludes that

$$\mu^n(x) = \sum_{l \in L(m,n)} l(\mu)(x) = \sum_{T \in \mathbf{MrT}_n} \sum_{T' \in E(T)} l^{T'}(\mu)\ (x) \stackrel{(6.4)}{=} \sum_{T \in \mathbf{MrT}_n} \mathcal{D}^T(x)\ .$$

All that remains to be shown is therefore

$$\mathcal{D}^n(x) = \sum_{T \in \mathbf{MrT}_n} \mathcal{D}^T(x)\ . \tag{6.5}$$

For $n = 1$ one readily sees

$$\mathcal{D}x = (\mathcal{D}x_1, \dots, \mathcal{D}x_N)^* = f(x) = \mathcal{D}^{T_0}(x) = \sum_{T \in \mathbf{MrT}_1} \mathcal{D}^T(x)\ ,$$

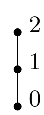
since the trivial tree T_0 on one vertex is the only tree in $\mathbf{MrT}_1$. Similarly one computes for $n = 2$:

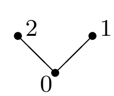
$$\mathcal{D}^2 x = \mathcal{D}f = \sum_{i_1=1}^{N} (f_{i_1} \downarrow x_{i_1})(f) = \sum_{i_1 \in \underline{N}^{\underline{1}}} (f_{i_1} \downarrow x_{i_1})(f)\mathcal{D}^T(x) = \sum_{T \in \mathbf{MrT}_2} \mathcal{D}^T(x)\ ,$$

since again $\mathbf{MrT}_2$ contains only one tree – the root with one successor 1. Now by the definition of higher order substitution operators (4.6) and the derivation property for the operators $(f_i \downarrow x_i)$ one computes for $n = 3$:

$$\begin{aligned}
\mathcal{D}^3 x = \mathcal{D}^2 f &= \sum_{i_2=1}^{N} (f_{i_2} \downarrow x_{i_2}) \left(\sum_{i_1=1}^{N} (f_{i_1} \downarrow x_{i_1})(f) \right) \\
&= \sum_{i_2,i_1=1}^{N} (f_{i_2}, f_{i_1} \downarrow x_{i_2}, x_{i_1})(f) + \sum_{i_2,i_1=1}^{N} ((f_{i_2} \downarrow x_{i_2})(f_{i_1}) \downarrow x_{i_1})(f) \\
&= \sum_{i \in \underline{N}^{\underline{2}}} (f_{i_2}, f_{i_1} \downarrow x_{i_2}, x_{i_1})(f) + \sum_{i_1 \in \underline{N}^{\underline{1}}} (\sum_{i_2 \in \underline{N}^{\underline{1}}} (f_{i_2} \downarrow x_{i_2})(f_{i_1}) \downarrow x_{i_1})(f) \\
&= \mathcal{D}^{T_2}(x) + \mathcal{D}^{T_1}(x) = \sum_{T \in \mathbf{MrT}_3} \mathcal{D}^T(x)
\end{aligned}$$

with

$T_1 =$ 

and $T_2 =$ 

.

For $n = 4$ the application of $(f_{i_3} \downarrow x_{i_3})$ to the first sum yields three sums over

$$(f_{i_3}, f_{i_2}, f_{i_1} \downarrow x_{i_3}, x_{i_2}, x_{i_1})(f)\ ,\ ((f_{i_3} \downarrow x_{i_3})(f_{i_2}), f_{i_1} \downarrow x_{i_2}, x_{i_1})(f)\ ,$$
$$\text{and}\ \ (f_{i_2}, (f_{i_3} \downarrow x_{i_3})(f_{i_1}) \downarrow x_{i_2}, x_{i_1})(f)$$

and to the second sum it yields three sums over

$$(((f_{i_3} \downarrow x_{i_3})(f_{i_2}) \downarrow x_{i_2})(f_{i_1}) \downarrow x_{i_1})(f) \ , \ ((f_{i_3}, f_{i_2} \downarrow x_{i_3}, x_{i_2})(f_{i_1}) \downarrow x_{i_1})(f) \ ,$$
$$\text{and} \ \ ((f_{i_2} \downarrow x_{i_2})(f_{i_1}), f_{i_3} \downarrow x_{i_1}, x_{i_3})(f) \ .$$

The general pattern is emergent now: by the derivation property the application of the sum of operators $(f_{i_{n+1}} \downarrow x_{i_{n+1}})$ to the summands of $\mathcal{D}^n x$ corresponds on the level of trees to adding a twig with vertex label $n+1$ to every vertex of every $T \in \mathbf{MrT}_n$, which gives exactly the sum over the trees of $\mathbf{MrT}_{n+1}$. This completes the proof of (6.5) and (6.3). □

Remark 6.7. Binary ordered rooted trees (in $2\mathbf{rT}$) with labeled leafs are natural representatives for repeated brackets in free Lie algebras. Namely, the labeled leafs represent the (numbered) generators, and the two successors of a non-leaf the two factors of a bracket. For a comprehensive exposition of the algebra and combinatorics of the free Lie algebras see the Reutenauer's book [Re].

7. Monotonely labeled rooted trees, the Exponential Formula and the Lie series

The sets $\mathbf{MrT}_n$ of monotonely labeled rooted trees and their walk-around codes (wa-codes), which are permutations in S_{n-1}, have been introduced already in Section 6. The natural recursive construction of the sets $\mathbf{MrT}_n$ proceeds as follows: the trees in $\mathbf{MrT}_{n+1}$ can be derived by taking any tree $T \in \mathbf{MrT}_n$, selecting a vertex v, and adjoining a twig with label $n+1$ to it. This immediately shows

$$|\mathbf{MrT}_{n+1}| = n!$$

and simultaneously sugests how to build up an L-code for this trees. Namely, for every l in

$$L_n := L(2,n) = \{l = (l_{n-1}, \ldots, l_0) \mid \nu \geq l_\nu \geq 0\}$$

one constructs recursively a tree $T \in \mathbf{MrT}_{n+1}$ by attaching a twig with label $\nu+1$ to the vertex l_ν (= the vertex with label l_ν). In other words: the vertex $\nu > 0$ of T covers the vertex $l_{\nu-1}$. The L-code, e.g., for the tree from Example 6.4 is $(7,2,3,0,2,3,0,1,0) \in L_9$. From the L-code of a $T \in \mathbf{MrT}_n$ one immediately reads off the out-degree of a vertex j: it is simply the number of occurences of j in $L(T)$.

We briefly discuss (a) how the L-code of some $T \in \mathbf{MrT}_n$ can be computed from the L-code of some $T' \in \mathbf{S}m\mathbf{rT}_n$, if T is the projection of T', i.e. T originates from the bijective image of T' in $\mathbf{ME}_m\mathbf{rT}_n$ by forgetting the edge labels; and (b) how the L-code of some $T \in \mathbf{MrT}_n$ can be computed from its wa-code.

For (a) compute the parenthesis code for T' as described in Section 6. Then the vertex j of T covers i, if ${}_i($ is the next unmatched parenthesis to the left of ${}_j($, where 'unmatched' means that the matching right parenthesis for ${}_i($ is not between ${}_i($ and ${}_j($.

For (b) compute the levels of the wa-code of T as described in Section 6. Then all vertices j on level 1 cover the root 0, and every j on a level greater 1 covers the next vertex i to the left, whoose level is smaller than the level of j. For example, let $T \in \mathbf{MrT}_{10}$ have

wa-code:	6	3	7	9	4	1	2	8	5
vertex on level:	1	1	2	3	2	1	2	3	3
covers vertex:	0	0	3	7	3	0	1	2	2

whence the L-code of T is $(7, 2, 3, 0, 2, 3, 0, 1, 0)$.

Proposition 7.1. *For every natural number $m \geq 2$ and every tree $T \in \mathbf{MrT}_n$ let $p_T(m)$ be the number of $T' \in \mathbf{ME}_m\mathbf{rT}_n$, which project onto T (upon forgetting the edge labels). Then p_T is a polynomial in $\mathbb{Z}[m]$ and (with notation* (3.1)*)*

$$p_T(m) = \prod_{j=0}^{n-1} (m)_{\beta(j)} \quad \textit{with } \beta(j) := |\{\nu \mid l_\nu = j\}| \ , \tag{7.1}$$

where of course $l = (l_{n-1}, \ldots, l_0)$ is the L-code of T.

Proof. Let j be a vertex of T with out-degree $\mathrm{out}(j)$. Then for the given m there are exactly $(m)_{\mathrm{out}(j)}$ possibilities to label the outgoing edges with different labels from $\underline{m}$. Since in terms of the L-code $\mathrm{out}(j)$ equals $\beta(j)$ the wanted projection number for T is the product over all factors $(m)_{\beta(j)}$. □

For the rest of this section assume that f is a polynomial mapping, which is homogeneous of degree m in the commuting variables $x_1, \ldots, x_N$, and that the mapping μ associated to f is in commutative form (see Section 3). We are now in a position to complete the constructive proof of the equalities

$$D^n(x) = \mu^n(x) \qquad (n \geq 0) \tag{7.2}$$

Despite the fact that (6.3) in case of commuting variables implies (7.2), because then the substitution operator $\mathcal{D}$ specializes to the differential operator D, there are some new phenomena in case of commuting variables, which make worthwile a fresh start.

Proposition 7.2. *For commutative $\mu \in L(T_m R^N, R^N)$ the powers $\mu^n(x)$ for $n > 0$ can be represented as*

$$\begin{aligned} \mu^n(x) &= \sum_{T \in \mathbf{MrT}_n} p_T(m)\ \mu^T(x)\ , \quad \textit{where recursively} \\ \mu^T(x) &:= \mu^T_{\uparrow(0)}(x) \quad \textit{and} \\ \mu^T_{\uparrow(j)}(x) &:= \mu(\mu^T_{\uparrow(j_1)}(x) \otimes \ldots \otimes \mu^T_{\uparrow(j_M)}(x) \otimes \otimes^{m-M} x)\ \ , \textit{ if } suc(j) = \{j_1, \ldots, j_M\}\ . \end{aligned}$$

In particular: $\mu^T_{\uparrow(j)}(x) = \mu^1(x) = \mu(\otimes^m x)$, if j is a leaf.

Proof. Recall from the proof of (6.3) that

$$\mu^n(x) = \sum_{T \in \mathbf{MrT}_n} \sum_{T' \in \mathbf{E}(T)} l^{T'}(\mu)(x)\ .$$

Since for commutative μ the ordering of the input vectors is not relevant, it follows from Proposition 7.1 that for all $T \in \mathbf{MrT}_n$

$$\sum_{T' \in \mathbf{E}(T)} l^{T'}(\mu)(x) = p_T(m)\ l^{T''}(\mu)(x)\ ,$$

where T'' is any fixed tree in $\mathbf{E}(T)$. But commutativity of μ means for the parenthesis code of the $T'' \in \mathbf{ME}_m\mathbf{rT}_n$ that all entries of every parenthesis ${}_j(\ldots)$ commute. Therefore all subparethesis of all ${}_j(\ldots)$ can be assumed to be in leftmost position, which gives exactly $\mu^T(x)$ upon recursive evaluation. Hence

$$l^{T''}(\mu)(x) = \mu^T(x)$$

and the proof is complete. □

Definition 7.3. *Let $f : R^N \longrightarrow R^N$ be a polynomial mapping with $f = (f_1, \ldots, f_N)^*$ and $f_1, \ldots, f_N \in R[[x_1, \ldots, x^N]]$. Then for $T \in \mathbf{MrT}_n$ and D as in* (1.1) *the linear partial differential operator D^T is defined recursively as*

$$\begin{aligned} D^T &:= D^T_{\uparrow(0)} \\ D^T_{\uparrow(j)} &:= \sum_{\nu \in \underline{N}^{\underline{M}}} f^T_{\nu_1, j_1} \ldots f^T_{\nu_M, j_M} \partial_\nu \ \ , \textit{ if } suc(j) = \{j_1, \ldots, j_M\} \neq \emptyset \\ f^T_{p,q}(x) &:= \begin{cases} f_p(x) & , \textit{ if } q \textit{ is a leaf of } T \\ D^T_{\uparrow(q)} f_p(x) & , \textit{ otherwise.} \end{cases} \end{aligned}$$

In other words, if the M polynomial mappings $f_{,1}, \ldots, f_{,M}$ for the M successors of a non-leaf j have been computed already, then they build up $D^T_{\uparrow(j)}$ the same way as the operator $\widetilde{D}_M$ of Lemma 3.1. We can now proof the following refinement of (7.2):

Theorem 7.4. *Let f, T, and D^T be given as in Definition 7.3, where in addition f is homogeneous of degree m, and let μ be the unique commutative mapping in $L(T_m R^N, R^N)$ associated to f. Then with the notations of Propositions 7.1 and 7.2:*

$$D^T(x) = p_T(m)\ \mu^T(x)\ . \tag{7.3}$$

Proof. From the description of the operator D^T given after its definition it is clear that in every recursive step on can replace $\widetilde{D}_M f(x)$ by the right hand side of formula (3.9) of Lemma 3.1:

$$(m)_M\ \mu(f_{,1}(x) \otimes \ldots \otimes f_{,M}(x) \otimes \otimes^{m-M} x)$$

Then the result follows from formula (7.1) of Proposition 7.1. □

Proof. of (7.2): Since by Proposition 7.2 and (7.3) one has

$$\mu^n(x) = \sum_{T \in \mathbf{MrT}_n} p_T(m)\ \mu^T(x) = \sum_{T \in \mathbf{MrT}_n} D^T(x)\ ,$$

it remains to be shown that

$$D^n(x) = \sum_{T \in \mathbf{MrT}_n} D^T(x)\ . \tag{7.4}$$

From

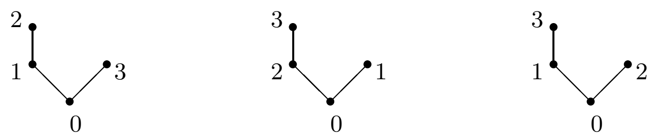
$$\begin{aligned}
Dx &= f(x) \\
D^2 x &= \sum_{i_1=1}^{N} f_{i_1}(\partial_{i_1} f) \\
D^3 x &= \sum_{i_2,i_1=1}^{N} [\ f_{i_2}(\partial_{i_2} f_{i_1})(\partial_{i_1} f) + f_{i_2} f_{i_1}(\partial_{i_2}\partial_{i_1} f)\ ] \\
D^4 x &= \sum_{i_3,i_2,i_1=1}^{N} [\ f_{i_3}(\partial_{i_3} f_{i_2})(\partial_{i_2} f_{i_1})(\partial_{i_1} f) + f_{i_3} f_{i_2}(\partial_{i_3 i_2} f_{i_1})(\partial_{i_1} f) \\
&\qquad + f_{i_3} f_{i_2}(\partial_{i_2} f_{i_1})(\partial_{i_3 i_1} f) + f_{i_3}(\partial_{i_3} f_{i_2}) f_{i_1}(\partial_{i_2 i_1} f) \\
&\qquad + f_{i_3} f_{i_2}(\partial_{i_3} f_{i_1})(\partial_{i_2 i_1} f) + f_{i_3} f_{i_2} f_{i_1}(\partial_{i_3 i_2 i_1} f)\ ]
\end{aligned}$$

it is easy to verify (7.4) directly for $n = 1$ and $n = 2$. For $n \geq 3$ the general pattern is clear: for given $T \in \mathbf{MrT}_n$ the application of the operator $f_{i_{n+1}} \partial_{i_{n+1}}$ to the summand $D^T(x)$ of $\mathcal{D}^n x$ corresponds to an addition of a twig with vertex label $n+1$ to every vertex of T. Summing up now proves (7.4) and (7.2). □

Notice, that the sums over the expressions

$$f_{i_3} f_{i_2}(\partial_{i_2} f_{i_1})(\partial_{i_3 i_1} f), \quad f_{i_3}(\partial_{i_3} f_{i_2}) f_{i_1}(\partial_{i_2 i_1} f), \quad f_{i_3} f_{i_2}(\partial_{i_3} f_{i_1})(\partial_{i_2 i_1} f)$$

correspond respectively to the trees



in $\mathbf{MrT}_4$, where

$$(\partial_{i_M \ldots j_1} f_j) \qquad \text{corresponds to} \qquad \operatorname{suc}(j) = \{j_1, \ldots, j_M\}.$$

Notice further, that all three sums, which are the expressions $D^T(x)$ for the three depicted trees, are equal. More generally one sees: if $\mathbf{M}(T) \subset \mathbf{MrT}_n$ is the set of all monotone labelings or all linear extensions of a rooted tree $T \in \mathbf{rT}_n$, then

$$\forall\ T_1, T_2 \in \mathbf{M}(T): \quad D^{T_1}(x) = D^{T_2}(x)\ . \tag{7.5}$$

This is true, because the result depends only on the kind of composition of the operators D and not on their absolute order in the composition or, in terms of the parenthesis code: the result depends on how the parenthesis are set and not on their subscript numbers. Therefore the expression $D^T(x)$ is well defined for every $T \in \mathbf{rT}_n$: use any linear extension $T' \in \mathbf{M}(T)$ and compute $D^{T'}(x)$ according to Definition 7.3.

Similarly, on sees for every $T \in \mathbf{rT}_n$ and every commutative μ that $\mu^T(x)$ is well defined, because

$$\forall\ T_1, T_2 \in \mathbf{M}(T): \quad \mu^{T_1}(x) = \mu^{T_2}(x)\ . \tag{7.6}$$

This leads naturally to the study of rooted trees $\mathbf{rT}$ and the projection numbers

$$e_T := |\mathbf{M}(T)| \quad \text{for all } T \in \mathbf{rT}, \tag{7.7}$$

because, e.g., in terms of μ one has:

Proposition 7.5. *For commutative $\mu \in L(T_m R^N, R^N)$ the powers $\mu^n(x)$ for $n > 0$ can be represented as*

$$\mu^n(x) = \sum_{T \in \mathbf{rT}_n} e_T\ p_T(m)\ \mu^T(x)$$

where $\mu^T(x) := \mu^{T'}(x)$ for some $T' \in \mathbf{M}(T)$ according to Proposition 7.2.

Notice, that $p_T(m) = 0$, if $\operatorname{out}(v) > m$ for some vertex v of T. Therefore:

$$\mu^n(x) = \sum_{T \in \mathbf{rT}_n^{\leqslant m}} e_T\ p_T(m)\ \mu^T(x) \quad \text{for fixed } m.$$

Proposition 7.6. *For $T \in \mathbf{rT}$ let $\operatorname{Aut}(T)$ denote the group of graph automorphisms of T, i.e. $\varphi \in \operatorname{Aut}(T)$ is a bijection on $V(T)$, which preserves adjacency. Let*

$$\gamma(T) := \prod_{v \in V(T)} |\uparrow(v)| \tag{7.8}$$

the product of weights of T. *Then*

$$e_T = \frac{|T|!}{\gamma(T)\ |\mathrm{Aut}(T)|}\ . \tag{7.9}$$

Proof. (7.9) is trivial for $|T| = 1$. Assume that T with $|T| \geq 2$ has s non-isomorphic types $T_1, \ldots, T_s$ of principal subtrees with multiplicities $m_1, \ldots, m_s$, respectively. Then

$$|\mathrm{Aut}(T)| = \prod_{i=1}^{s} m_i!\ |\mathrm{Aut}(T_i)|^{m_i}\ ,$$

because $\mathrm{Aut}(T)$ is the direct product of the groups $S_{m_i} \wr \mathrm{Aut}(T_i)$. Since the principal subtrees have all together $|T| - 1$ vertices, there are

$$\frac{(|T|-1)!}{\prod_{i=1}^{s} m_i!\ (|T_i|!)^{m_i}}$$

possibilities to distribute the labels $1, \ldots, |T| - 1$ to these vertices, where the factors $m_i!$ are due to the fact that the m_i principal subtrees of type T_i are indistinguishable. One then computes inductively:

$$\begin{aligned}
e_T &= \frac{(|T|-1)!}{\prod_{i=1}^{s} m_i!\ (|T_i|!)^{m_i}}\ \prod_{i=1}^{s} (e_{T_i})^{m_i} \\
&= \frac{(|T|-1)!}{\prod_{i=1}^{s} m_i!\ (|T_i|!)^{m_i}}\ \prod_{i=1}^{s} \left(\frac{|T_i|!}{\gamma(T_i)\ |\mathrm{Aut}(T_i)|} \right)^{m_i} \\
&= \frac{(|T|-1)!}{\prod_{i=1}^{s} \gamma(T_i)^{m_i}\ m_i!\ |\mathrm{Aut}(T_i)|^{m_i}} = \frac{(|T|-1)!}{|\mathrm{Aut}(T)|\ \prod_{i=1}^{s} \gamma(T_i)^{m_i}} \\
&= \frac{(|T|-1)!}{|\mathrm{Aut}(T)|\ \gamma(T)/|T|} = \frac{|T|!}{\gamma(T)\ |\mathrm{Aut}(T)|}\ .
\end{aligned}$$

□

Directly from the proof of Proposition 7.6 and from (7.1) one infers the following

Corollary 7.7. *Let T be a rooted tree, which has s non-isomorphic types $T_1, \ldots, T_s$ of principal subtrees with multiplicities $m_1, \ldots, m_s$, respectively. Then the number of linear extensions of T can be computed recursively by*

$$e_T = \frac{(|T|-1)!}{\prod_{i=1}^{s} m_i!}\ \prod_{i=1}^{s} \left(\frac{e_{T_i}}{|T_i|!} \right)^{m_i}\ . \tag{7.10}$$

Moreover:

$$p_T(m) = (m)_{\beta(0)} \prod_{i=1}^{s} (p_{T_i}(m))^{m_i} \quad \textit{with} \quad \beta(0) := \sum_{i=1}^{s} m_i\ . \tag{7.11}$$

Since Corollary 7.7 allows the recursive calculation of the correct multiplicities for the expressions $\mu^T(x)$ occuring in the sum $\mu^n(x)$ (cf. Proposition 7.5), the problem arises how to compute recursively the sets $\mathbf{rT}_n$. Ideally such a recursion will generate all representatives of all isomorphism classes of ordered trees in linear order without the need of comparisons, normal forms, etc. (see [R1,R2] for general considerations on the orderly generation of tables of graphs). Such a recursion will be described now.

The set $\mathbf{rT}_1$ contains only the trivial tree. Assume that the rooted trees of $\mathbf{rT}_j$ for $j = 1, \dots, n$ are described as (finite) lists of the form

$$\mathbf{rT}_j = (T_1^j, T_2^j, T_3^j, \dots) ,$$

i.e. the superscript j indicates the number of vertices and the subscript enumerates the elements of $\mathbf{rT}_j$ in a certain linear order. Let

$$\Lambda_n := \{\lambda = (\lambda_1, \dots, \lambda_s) \mid \lambda_1 \geq \dots \geq \lambda_s \geq 1 , |\lambda| := \lambda_1 + \dots + \lambda_s = n\}$$

be the set of all *partitions* of the number n and let $\Lambda_n^{\succ}$ be the set Λ_n ordered lexicographically with respect to the linear order $0 \succ 1 \succ 2 \succ 3 \dots$. Then the weights of the principal subtrees of every $T \in \mathbf{rT}_{n+1}$ can be assumed to form an element of Λ_n. Since for every $T \in \mathbf{rT}_{n+1}$ the weights of the principal subtrees can be assumed to form an element of Λ_n and since every principal subtree of T is an element of some $\mathbf{rT}_j$ with $1 \leq j \leq n$, one concludes that

$$T = T_i^{n+1} = \lfloor T_{i_1}^{\lambda_1}, \dots, T_{i_s}^{\lambda_s} \rfloor \quad \text{with} \ \ 1 \leq i_\nu \leq |\mathbf{rT}_{\lambda_\nu}| , \tag{7.12}$$

where the position i of T in the list of $\mathbf{rT}_{n+1}$ is not yet specified. To get this linear order take first the upper indices in the linear order of $\Lambda_n^{\succ}$ and take second the admissable s-tuples $(i_1, \dots, i_s)$ for every partition λ in the lexicographic order, which is induced by the usual linear order on $\mathbb{N}$. A last point to be observed is that for a m_ν-fold occurence of a part λ_ν all permutations of the entries of the corresponding m_ν-tuples of admissable subindices describe the same rooted tree. Therefore one chooses only the lexicographically smallest representative of all permutations of such a tuple. The result up to $n = 5$ together with the multiplicities $e_T \, p_T(m)$ is :

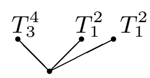
$$\begin{array}{llcll}
\mathbf{MrT}_1 : & T_1^1 & = & \text{trivial tree} & \\
 & & & & \\
\mathbf{MrT}_2 : & T_1^2 & = & \lfloor T_1^1 \rfloor & m \\
 & & & & \\
\mathbf{MrT}_3 : & T_1^3 & = & \lfloor T_1^2 \rfloor & m^2 \\
 & T_2^3 & = & \lfloor T_1^1, T_1^1 \rfloor & m(m-1)
\end{array}$$

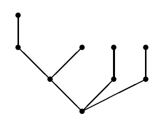
$$\begin{array}{rrcll} \mathbf{MrT}_4: & T_1^4 & = & \lfloor T_1^3 \rfloor & m^3 \\ & T_2^4 & = & \lfloor T_2^3 \rfloor & m^2(m-1) \\ & T_3^4 & = & \lfloor T_1^2, T_1^1 \rfloor & 3m^2(m-1) \\ & T_4^4 & = & \lfloor T_1^1, T_1^1, T_1^1 \rfloor & m(m-1)(m-2) \end{array}$$

$$\begin{array}{rrcll} \mathbf{MrT}_5: & T_1^5 & = & \lfloor T_1^4 \rfloor & m^4 \\ & T_2^5 & = & \lfloor T_2^4 \rfloor & m^3(m-1) \\ & T_3^5 & = & \lfloor T_3^4 \rfloor & 3m^3(m-1) \\ & T_4^5 & = & \lfloor T_4^4 \rfloor & m^2(m-1)(m-2) \\ & T_5^5 & = & \lfloor T_1^3, T_1^1 \rfloor & 4m^3(m-1) \\ & T_6^5 & = & \lfloor T_2^3, T_1^1 \rfloor & 4m^2(m-1)^2 \\ & T_7^5 & = & \lfloor T_1^2, T_1^2 \rfloor & 3m^3(m-1) \\ & T_8^5 & = & \lfloor T_1^2, T_1^1, T_1^1 \rfloor & 6m^2(m-1)(m-2) \\ & T_9^5 & = & \lfloor T_1^1, T_1^1, T_1^1, T_1^1 \rfloor & m(m-1)(m-2)(m-3) \end{array}$$

With the help of these tables it is easy to draw a rooted tree T_j^j and to compute its multiplicities:

Example 7.8. Given $T = T_{\dots}^9 = \lfloor T_3^4, T_1^2, T_1^2 \rfloor$. Then

$T =$ 

$=$  .

Since T_3^4 and T_1^2 have multiplicities $3m^2(m-1)$ and m, respectively, one computes with (7.10-11): $p_T(m) = (m)_3 \cdot m^2(m-1) \cdot m \cdot m = m^5(m-1)^2(m-2)$ and

$$e_T = \frac{8!}{1!\ 2!} \cdot \frac{3}{4!} \cdot \left(\frac{1}{2!}\right)^2 = 7 \cdot 6 \cdot 5 \cdot 3 = 630 \ .$$

If $c^*(p,q)$ denotes the number of q-element multisets, which can be choosen from p elements, and if

$$I_{m,n} := \{ i = (i_1, \dots, i_m) \mid i_1 \geq \dots \geq i_m \geq 0 \ , |i| := i_1 + \dots + i_m = n \} \ ,$$

and $\alpha_i(k) := |\{\nu \mid i_\nu = k\}|$, then with a little contemplation of the above recursive procedure for the recursive generation of ordered rooted trees one sees that the numbers

$$C_m(n) := |\mathbf{rT}_n^{\leqslant m}|$$

satisfy the recursion

$$C_m(n+1) = \sum_{i \in I_{m,n}} \prod_{k=0}^{n} c^*(C_m(k), \alpha_i(k)) \quad \text{with } C_m(0) := 1 \ . \tag{7.13}$$

Remark 7.9. Formula (7.13) has been observed first by Melzak [Mz], but his statement and proof of the result are much more complicated. The asymptotic behaviour for $n \longrightarrow \infty$ of the numbers $C_n(n)$, $C_m(n)$, and $|\mathbf{T}_n|$ has been determined by Polya [P,PR] and Otter [Ot] with the help of functional equations for the respective generating functions — in [HRS] these classical methods are crystallized into a twenty step algorithm. The numbers $C_2(n)$ are known as "Wedderburn-Etherington numbers" [Sl, No.298], because Wedderburn [W] found them while investigating the number of parenthesations of a non-assoziative product of n factors — and in fact every commutative $\mu \in L(T_2R^N, R^N)$ can be viewed as a commutative, but in general non-assoziative multiplication law on R^N.

More (asymptotic) counting results for other types of trees are described and reviewed in the paper [HPr] and the book [HP]. Counting and coding of trees in connection with the chemical theory of isomers is investigated further in [GK]. And Labelle [Lb] studies in depth the counting problems for different kinds of asymmetric trees, i.e. trees with trivial automorphism group.

Remark 7.10. If derivations and elementary algebraic operations on power series are represented by certain combinatorial operations on a *linear species* M (i.e. M is a functor from the category of finite linearly ordered sets with order preserving bijections to the category of finite sets with functions), then a solution of the IVP

$$\frac{dY}{dT} = M(T,Y), \quad Y(0) = Z$$

can be interpreted combinatorially in terms of "M-enriched arborescences" [LV]. A similar combinatorial approach based on the interaction between ordinary differential equations, substitutions ("grammars"), and trees in some special cases is the theme of [DR].

Remark 7.11. Grossman and Larson [GL1-3] have introduced a Hopf-algebraic structure on rooted trees (with possible additional structure), which is compatible with the evaluation of trees by differential operators as in Definition 7.3. Namely, the product of two trees $T \cdot T'$ is defined as sum over all possible attachments of the principal subtrees of T to the vertices of T'. This is the extension of the procedure of attaching a twig in all possible ways to a tree as used in the proof of (7.2). The coproduct of some T is the sum over all possible products of the two trees, which can be formed from the set of principal subtrees of T. Clearly, the multiplication is non-commutative and the comultiplication cocommutative. The obvious unit and counit then complete the Hopf-algebraic structure, which Grossman and Larson subsequently apply to the investigation and simplification of the computations with higher order derivations and Lie brackets of differential operators.

8. Combinatorial supplements

Tree statistics count finer structural properties of different kinds of trees. They describe the distribution of levels, branching (= out-degrees), hights, leaves, (principal) subtrees, distances, chains and antichains, etc. . The most important two-parameter statistics are the following ones, where $\Box$ stands for any one of
$\mathbf{rT}$, $\mathbf{OrT}$, $\mathbf{MrT}$, $\mathbf{ME}_m\mathbf{rT}$, $\ldots$:

The *principal tree statistics* counts the out-degrees of the roots:

$$p_\Box(n,k) := |\{T \in \Box_n \mid \mathrm{out}(v_0) = k\}| \ .$$

The *branching statistics* counts the out-degrees of the vertices:

$$b_\Box(n,k) := \sum_{T\in\Box_n} |\{v \in V(T) \mid \mathrm{out}(v) = k\}| \ .$$

The *level statistics* counts the number of vertices on a fixed level:

$$l_\Box(n,k) := \sum_{T\in\Box_n} |\{v \in V(T) \mid r(v) = k\}| \ .$$

The *hight statistics* counts the hight of trees:

$$h_\Box(n,k) := |\{T \in \Box_n \mid h(T) = k\}| \ .$$

The *leaf statistics* counts the number of trees having a fixed number of leafs:

$$f_\Box(n,k) := |\{T \in \Box_n \mid |L(T)| = k\}| \ .$$

The *leaf-level statistics* counts the number of leafs on a fixed level:

$$ll_\Box(n,k) := \sum_{T\in\Box_n} |\{v \in L(T) \mid r(v) = k\}| \ .$$

As a three-parameter statistic we mention the *level-branching statistics*, which counts the out-degrees in relation to the levels:

$$lb_\Box(n,k,l) := \sum_{T\in\Box_n} |\{v \in V(T) \mid \mathrm{out}(v) = k,\ r(v) = l\ \}| \ .$$

In view of the results of Sections 6 and 7 the case $\Box = \mathbf{MrT}$ appears to be especially interesting. Since the wa- and the L-code have been seen to induce bijections between the symmetric groups S_n of permutations and the sets $\mathbf{MrT}_{n+1}$, one can use trees to refine and extend the extensively investigated permutation statistics (cf. [St1,BW,Rw]). It is well known [St1, Prop.1.3.16] that (suppressing subscripts $\mathbf{MrT}$)

$$p(n+1,k) = c(n,k) \quad \text{(signless Stirling numbers of the first kind),}$$
$$f(n+1,k) = A(n,k) \quad \text{(Eulerian numbers).}$$

Propositions 8.1 - 3 below characterize the other two-parameter statistics in case of **MrT**, except for the hight statistics, which seems to be more complicated. First of all, the principal tree statistics and the level statistics appear to be closely related:

Proposition 8.1. *The level statistics for* **MrT** *is*

$$l(n,k) = c(n,k+1) \ . \tag{8.1}$$

Proof. It is obvious that $l(n+1,0) = n!$ and $l(1,k) = \delta_{0,k}$. Therefore it is enough to prove

$$l(n+1,k) = n\ l(n,k) + l(n,k-1) \ , \tag{8.2}$$

which is the recursion relation for the signless Stirling numbers of the first kind. Recall form the beginning of Section 7 the recursive generation of the set $\mathbf{MrT}_{n+1}$ through the attachment of a twig with label $n+1$ to every vertex v of every $T \in \mathbf{MrT}_n$. Since T has n vertices this produces n copies of every T, where in addition for every vertex v of rank $r(v)$ there is a new vertex on level $r(v)+1$. This proves (8.2) and therefore (8.1). □

The statistics described in Propositions 8.2-3 below and the hight statistics seem to be new (they are not contained in [Sl]).

Proposition 8.2. *The branching statistics for* **MrT** *obeys the recursion*

$$\begin{aligned} b(n+1,k) = (n-1)\ b(n,k) + b(n,k-1) \ , \\ b(n,0) = n!/2 \ \ (\text{for } n \geq 2), \qquad b(1,k) = \delta_{1,k} \ . \end{aligned} \tag{8.3}$$

Proof. For the proof of (8.3) we show first $b(n,0) = n!/2$, which is the total number of leafs in $\mathbf{MrT}_n$. As in the proof of Proposition 8.1 above an addition of twigs (in the recursive generation of $\mathbf{MrT}_{n+1}$ from $\mathbf{MrT}_n$) generates n copies of every $T \in \mathbf{MrT}_n$, whence the new number of leafs is $n\ b(n,0)$ plus the $n(n-1)!$ leafs from the new twigs minus the $b(n,0)$ cases, where the twig is attached to a leaf. Since $b(1,0) = b(2,0) = 1$ the assertion $b(n,0) = n!/2$ now follows. Similarly, for $k \geq 1$ the number $b(n+1,k)$ equals $n\ b(n,k)$ minus the $b(n,k)$ cases, where the out-degree is enlarged from k to $k+1$, and plus the $b(n,k-1)$ cases, where the out-degree is enlarged from $k-1$ to k. This proves the assertion. □

Proposition 8.3. *The leaf-level statistics for the case* **MrT** *obeys the recursion*

$$\begin{aligned} ll(n+1,k) = n\ ll(n,k) + ll(n,k-1) \ \ (\text{for } n \geq 2) \ , \\ ll(n,1) = (n-1)! \ (\text{for } n \geq 2), \qquad ll(2,k) = \delta_{1,k} \ . \end{aligned} \tag{8.4}$$

Moreover, the level statistics and the leaf-level statistics are connected by the surprising formula

$$l(n,k) = ll(n,k+1) + ll(n,k) \ . \tag{8.5}$$

Proof. Again we use the recursion step leading from $\mathbf{MrT}_n$ to $\mathbf{MrT}_{n+1}$ and we show first (8.4) in case of $k = 1$: the number $ll(n+1,1)$ of leafs on level 1 generated from the trees in $\mathbf{MrT}_n$ is $n \ ll(n,1)$ minus the $ll(n,1)$ cases, where a twig is added to a leaf on level 1, and plus the number of roots ($= (n-1)!$), which yields the desired conclusion. For $k > 1$ one concludes similarly that

$$ll(n+1,k) = (n-1) \ ll(n,k) + l(n,k-1) \ , \tag{8.6}$$

where $l(n,k-1)$ is the number of (arbitrary) vertices on level $k-1$ in $\mathbf{MrT}_n$. But the recursion (8.6) implies (8.4), if (8.5) is correct. The latter can be seen as follows: For $k = 0$ we know already that $l(n,0) = (n-1)! = ll(n,1)$ and for $k \geq 1$ the number $l(n,k) - ll(n,k)$ is the cardinality of the set

$$R(n,k) := \{v \in \biguplus_{T \in \mathbf{MrT}_n} V(T) \mid r(v) = k, \ |\uparrow(v)| > 1\} \ .$$

For a fixed vertex $u \in R(n,k)$ of some tree T' let $C(u) := T' \setminus \uparrow(u)$ be the subtree of T', which is complementary to $\uparrow(u)$. Let $\overline{C}(u)$ be the class of all $T \in \mathbf{MrT}_n$, which contain $C(u)$ as a subtree. Then by the symmetry of monotone labelings the set of all complements of $C(u)$ for the trees in $\overline{C}(u)$ is a complete set of monotonely labeled trees on $|\uparrow(v)| > 1$ vertices, where the set of labels is given by the labels of $\uparrow(u)$. But from the case $k = 0$ we know already that $|\mathbf{MrT}_n| = ll(n,1)$ for all $n > 1$, whence the number of twigs attached to the vertices corresponding to the set $\overline{C}(u)$ is $|\overline{C}(u)|$. Since all the sets $\overline{C}(u)$ induce a partition of $R(n,k)$ into equivalence classes, formula (8.5) follows. □

Notice that the monotone labelings in case of $\mathbf{MrT}$ allow additional refinements of statistics. For example Moon [Mo] computes the mean and variance of the distances $d(i,j)$ of the vertices $0 \leq i < j \leq n-1$ in $\mathbf{MrT}_n$.

We turn attention now to ordered rooted trees $\mathbf{OrT}$. These trees are counted by the celebrated *Catalan numbers*

$$|\mathbf{OrT}_{n+1}| = C_n = \frac{1}{n+1}\binom{2n}{n} \ ,$$

which form the solution of the recurrence

$$C_{n+1} = \sum_{i=0}^{n} C_i C_{n-i} \ , \quad C_0 = 1 \ .$$

The book [St2] describes 65 families of discrete structures, which are counted by the Catalan numbers. That **OrT** belongs to the Catalanian structures is easily seen by considering the operation $\rfloor$ on **OrT**: $T \rfloor T'$ is defined as the ordered rooted tree T, which has T' adjoined as its rightmost principal subtree. The assertion then follows from

$$\mathbf{OrT}_{n+1} = \biguplus_{i=1}^{n} \quad \mathbf{OrT}_i \rfloor \mathbf{OrT}_{n+1-i} \ .$$

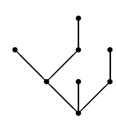
About the statistics of **OrT** the following is known: [DZ] containes explicit formulas for the statistics of leafs, principal trees, branching, and level-branching. [dBKR] gives a recursion for the hight statistics and computes the asymtotic average hight.

Ordered rooted trees are coded much easier than rooted trees, where the idea of 'walking around' appears in many variations:

(1) *Parenthesis code*: While walking around one writes a left parenthesis for every step up and a right parenthesis for every step down.
(2) *Binary code*: Same as the parenthesis code, but with 0 instead of a left parenthesis and 1 instead of a right parenthesis.
(3) *Walk around valency code* (WAV): Attach the out-degree to every vertex, and then walk around.
(4) *Walk around level code* (WAL): Attach the level to every vertex.
(5) *Walk around weight code* (WAW): Attach the 'weight' $|\uparrow(v)|$ to every vertex. WAW is closest in spirit to our recursive enumeration for **rT**.

The idea of the *Bottom up valency code* (BUV) is a bit different from the WAV: to every vertex attach its out-degree, but record it from bottom to top and on each level from left to right instead of walking around.

Example 8.4.

$T =$  has

parenthesis code : (()(()))()(()), binary code : 00100111010011, WAV : 32010010, WAL : 1223112, WAW : 84121121, BUV : 32010100 .

The idea of parenthesis code and binary code goes back in principle to the work of Cayley [C1,C2] and reappears in [dBM]. BUV is introduced in [R3], wheras WAL and WAW seem to be new.

For the sets **rT** and **T** coding ideas similar to WAV and WAL apply, where the WAV for **T** is explained — together with the more sophisticated Smolenski code — in [R3, Sec.12]. For the WAL of a tree T first label all leafs with 0, then remove or ignore them and label the new leafs with 1, etc. . The 'shelling' of T thus recorded by the labels stops when the *center* of T consisting of one or two vertices is reached. Next to nothing seems to be known about the statistics of $\mathbf{rT}$ and $\mathbf{T}$ and about the projection numbers from $\mathbf{OrT}$ to $\mathbf{T}$ or $\mathbf{rT}$ to $\mathbf{T}$.

References

[BCW] H. Bass, E.H. Connell, D. Wright, The Jacobian Conjecture: Reduction of degree and formal expansion of the inverse, *Bull. Amer. Math. Soc.* **7** (1982), 287 - 330.

[BW] A. Björner, M.L. Wachs, Permutation Statistics and Linear Extensions of Posets, *J. Comb. Theory A* **58** (1991), 85 - 114.

[C1] A. Cayley, On the theory of analytical forms called trees, *in:* "Collected Mathematical papers of Arthur Cayley", Vol.3, 242 - 246, Cambridge University Press, 1890.

[C2] A. Cayley, On the theory of analytical forms called trees, second part, *in:* "Collected Mathematical papers of Arthur Cayley", Vol.4, 112 - 115, Cambridge University Press, 1891.

[dBM] N.G. de Bruijn, B.J.M. Morselt, A note on plane trees, *J. Comb. Theory* **2** (1967), 27 - 34.

[dBKR] N.G. de Bruijn, D.E. Knuth, S.O. Rice, The average hight of planted plane trees, *in:* R.C. Read (ed.), "Graph Theory and Computing", Academic Press, New York, 1972.

[DR] D. Dumont, A. Ramamonjisoa, Grammaire de Ramanujan et Arbres de Cayley, *Electronic Journal of Combinatorics* **3 (2)** (1996) # R17, 18pp.

[DZ] N. Derschowitz, S. Zaks, Enumeration of ordered trees, *Discrete Math.* **31** (1980), 9 - 28.

[G] W. Gröbner, "Die Lie-Reihen und ihre Anwendungen", VEB Deutscher Verlag der Wissenschaften, Berlin, 1960.

[GL1] R. Grossman, R.G. Larson, Hopf–Algebraic Structure of Families of Trees, *J. Algebra* **126** (1989), 184 - 210.

[GL2] R. Grossman, R.G. Larson, Solving Nonlinear Equations from Higher Order Derivations in Linear Stages, *Advances in Math.* **82** (1990), 180 - 202.

[GL3] R. Grossman, R.G. Larson, Symbolic Computation of Derivations Using Labelled Trees, *J. Symb. Comp.* **13** (1992), 511 - 523.

[Gr] R. Grossman, Using trees to compute approximate solutions to ordinary differential equations exactly, *in:* "Differential Equations and Computer Algebra", Academic Press, 1991.

[GK] M. Gordon, J.W. Kennedy, The counting and coding of trees of fixed diameter, *SIAM J. Appl. Math.* **28** (1975), 376 - 398.

[HP] F. Harary, E.M. Palmer, "Graphical Enumeration", Academic Press, New York, 1973.

[HPr] F. Harary, G. Prins, The number of homeomorphically irreducible trees, and other species, *Acta Math.* **101** (1959), 141 - 162.

[HRS] F. Harary, R.W. Robinson, A.J. Schwenk, Twenty-step algorithm for determining the asymptotic number of trees of various species, *J. Austral. Math. Soc.* **20** (1975), 483 - 503.

[L] S. Lie, "Theorie der Transformationsgruppen", Teubner, Leipzig, 1888.

[Lb] G. Labelle, Counting Asymmetric Enriched Trees, *J. Symbolic Computation* **14** (1992), 211 - 242.

[LV] P. Leroux, X.G. Viennot, Combinatorial Resolution of Systems of Differential Equations, I. Ordinary Differential Equations, *Springer Lecture Notes Math.* **1234** (1986), 210 - 245.

[Mo] J.W. Moon, The distance between nodes in recursive trees, *in:* "Combinatorics", *London Math. Soc. Lecture Notes* **13**, Cambridge Univ. Press, Cambridge, 1974.

[Mz] Z.A. Melzak, A note on homogeneous dendrits, *Canad. Math. Bull.* **11** (1968), 85 - 93.

[Ot] R. Otter, The number of trees, *Ann. Math.* **49** (1948), 583 - 599.

[PR] G. Polya, R.C. Read, "Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds", Springer, New York, 1987.

[P] G. Polya, Kombinatorische Anzahlbestimmung für Graphen und chemische Verbindungen, *Acta Math.* **68** (1937), 145 - 254.

[R1] R.C. Read, Every one a winner or How to avoid isomorphism search when cataloguing combinatorial configurations, *Ann. Discrete Math.* **2** (1978), 107 - 120.

[R2] R.C. Read, A Survey of Graph Generation Thechniques, *in:* "Combinatorial Mathematics VIII", *Springer Lecture Notes Math.* **884** (1981), 77 - 89.

[R3] R.C. Read, The coding of various kinds of unlabeled trees, *in:* R.C. Read (ed.), "Graph Theory and Computing", Academic Press, New York, 1972.

[Re] C. Reutenauer, "Free Lie Algebras", Clarendon Press, Oxford, 1993.

[Rw] D. Rawlings, A Binary Tree Decomposition Space of Permutation Statistics, *J. Comb. Theory A* **59** (1992), 111 - 124.

[Sb1] S. Steinberg, Lie Series and Nonlinear Ordinary Differential Equations, *J. Math. Anal. Appl.* **101** (1984), 39 - 63.

[Sb2] S. Steinberg, Lie series, Lie transforms, and their applications, *in:* "Lie Methods in Optics", *Springer Lecture Notes Physics* **250** (1986), 45 - 103.

[Sl] N.J. Sloane, "Sloane's On-Line Encyclopedia of Integer Sequences", *at:* `http://www.research.att.com/~njas/sequences`

[St1] R.P. Stanley, "Enumerative Combinatorics Volume I", Wadsworth & BrooksCole, Montery California, 1986.

[St2] R.P. Stanley, "Enumerative Combinatorics Volume II", Wadsworth & BrooksCole, Montery California, 1998.

[Wd] J.H.M. Wedderburn, The functional equation $g(x^2) = 2\alpha x + [g(x)]^2$, *Ann. Math.* **24** (1922), 121 - 140.

[W] R. Winkel An Exponential Formula for Polynomial Vector Fields, *Advances in Math.* **128** (1997), 190 - 216.

[Z] D. Zeilberger, Toward a Combinatorial Proof of the Jacobian Conjecture, *Springer Lecture Notes Math.* **1234** (1986), 370 - 380.

Institut für Reine und Angewandte Mathematik, RWTH Aachen, D-52056 Aachen, Germany, *papers/preprints*: http://www.iram.rwth-aachen.de/ ~winkel/

E-mail address: `winkel@iram.rwth-aachen.de`