JOURNAL OF COMBINATORIAL THEORY, Series A 50, 235–258 (1989)

On the Generalized Iterates of Yeh's Combinatorial $\mathbb{K}$-Species

GILBERT LABELLE

Département de Mathématiques et d'Informatique, Université du Québec à Montréal, C. P. 8888, Succ. "A," Montréal, Québec, Canada H3C 3P8

Communicated by Gian-Carlo Rota

Received June 9, 1987

Let $f=f(x)=x+a_2x^2+\cdots\in\mathbb{K}[[x]]$ be a "normalized" power series over a (commutative) field $\mathbb{K}$ of characteristic zero. The operator $\Delta_f\colon \mathbb{K}[[x]]\to\mathbb{K}[[x]]$, defined by $\Delta_f g=g\circ f-g$, has been used in (G. Labelle, *European J. Combin.* **1** (1980), 113–138) to obtain formulas for the inverse $f^{\langle -1\rangle}$ and the generalized iterates $f^{\langle t\rangle}$, $t\in\mathbb{K}$, of the series f. A. Joyal (in Lect. Notes in Math. Vol. 1234, pp. 126–159, Springer-Verlag, New York/Berlin, 1986) was the first to realize that Δ_f can be lifted to the combinatorial level. He made use of this fact to obtain a formula for a virtual species $F^{\langle -1\rangle}$ which is the inverse (under substitution) of any given normalized species $F=X+\cdots$. Using the same operator, we show that the concept of $\mathbb{K}$-species in the sense of Y.-N. Yeh (ibid.) (where $\mathbb{K}$ is now only a binomial half-ring) is a good context for the definition of the generalized iterates $F^{\langle t\rangle}$, $t\in\mathbb{K}$, of any normalized species (or $\mathbb{K}$-species). We present a new approach to Yeh's extension of substitution to $\mathbb{K}$-species. We also introduce the notions of "infinitesimal generator," "directional derivatives," and "Lie bracket" of $\mathbb{K}$-species, which turn out to be $\bar{\mathbb{K}}$-species, where $\bar{\mathbb{K}}$ denotes the "rational closure" of $\mathbb{K}$. These concepts give, in return, a better insight into substitution itself. For example, $G\circ F$ can be written in the form $G\circ F=(\exp D_\Phi)G$ for a suitably chosen derivation D_Φ. More generally, $G\circ F^{\langle t\rangle}=(\exp tD_\Phi)G$. Two normalized $\mathbb{K}$-species commute under substitution if and only if the Lie bracket of their infinitesimal generators is zero. Explicitly computed examples are also given.

1. INTRODUCTION

We first recall some basic facts. Let F be any combinatorial species in the sense of A. Joyal [J1] (see also [L4]). For each $n\in\mathbb{N}$ we can extract a subspecies F_n of F by collecting all those F-structures having an "underlying set" of cardinality n. If $F=F_n$, we say that F is *concentrated* on the cardinality n. In the general situation, we obviously have a countable decomposition

$$F=F_0+F_1+F_2+\cdots+F_n+\cdots. \tag{1.1}$$

* Partially supported by grants FCAR (Québec, EQ 1608) and CRSNG (Canada, A5660).

0097-3165/89 $3.00

Further, it is well known [L3, L5, Y] that each F_n can, in turn, be split (up to isomorphism) in a unique way into a finite sum of the form

$$F_n = \sum_{M \in \mathcal{M}_n} f_M M, \qquad f_M \in \mathbb{N}, \tag{1.2}$$

where $\mathcal{M}_n$ denotes a complete set of representatives of all the *molecular species* (i.e., irreducible under addition, and $\neq 0$) that are concentrated on the cardinality n; the equivalence relation being that of isomorphism of species. Each coefficient f_M in (1.2) is called the *multiplicity of M in F.*

Moreover, Y.-N. Yeh has shown [Y] that each molecular species $M \in \mathcal{M}_n$ can itself be factorized (up to isomorphism) in a unique way into a finite product

$$M = \prod_{A \in \mathcal{A}} A^{\gamma_A(M)}, \qquad \gamma_A(M) \in \mathbb{N}, \qquad \sum_{A \in \mathcal{A}} \gamma_A(M) < \infty, \tag{1.3}$$

where $\mathcal{A} \subset \mathcal{M} = \bigcup_{n \geqslant 0} \mathcal{M}_n$ denotes the (countable) set of those molecular species that are *atomic species* (i.e., also irreducible under product and $\neq 1$). Each $\gamma_A(M)$ in (1.3) is called the *exponent of A in M.*

As it is implicit in Eqs. (1.1)–(1.3) above, the equality sign "$=$" is used to denote isomorphic species. This is a standard convention in the theory of species and it will be used throughout the present paper (except when some explicit isomorphisms are needed). Thus every species F has *molecular* and *atomic decompositions*

$$F = \sum_{M \in \mathcal{M}} f_M M = \sum_{M \in \mathcal{M}} f_M \prod_{A \in \mathcal{A}} A^{\gamma_A(M)} \tag{1.4}$$

and the collection of all (isomorphism classes of) species forms a *half-ring* that is isomorphic to the half-ring $\mathbb{N}[[\mathcal{A}]]$ of all formal power series, with non-negative integral coefficients, in the set $\mathcal{A}$ of "atomic" variables.

Complete tables of the finite sets $\mathcal{A}_n$ and $\mathcal{M}_n$ have been made for small values of n (see [L3, L5] for $0 \leqslant n \leqslant 5$). Their construction is based on the fact that a species M, concentrated on cardinality n, is molecular iff the symmetric group $\mathfrak{S}_n$ acts transitively on the set of all M-structures built on $[n] = \{1, 2, ..., n\}$. Of course, infinite families of molecular or atomic species can also be easily exhibited. For example: let X, E, S, C, and L respectively denote the species of *singletons*, *sets*, *permutations*, *circular permutations*, and *linear orders*; then, for $n \geqslant 1$, the species E_n and C_n are all atomic, $L_n = X^n$ are all molecular but only $L_1 = X$ is atomic, while S_n is not molecular for $n \geqslant 2$. Note, moreover, that the species $E_0 = S_0 = L_0 = 1$ of the *empty set* is molecular (and should not be confused with 0, the *empty* species).

There is much more than just a half-ring structure on $\mathbb{N}[[\mathcal{A}]]$ because many other operators (including *substitution* $\circ$, *cartesian product* $\times$, and *derivation* $'$) have been combinatorially added to it [J1]. For example, it can be checked that $E_2 \circ X^2$ is atomic, that $C_3 \times C_3 = 2C_3$, $L'_n = (X^n)' = nX^{n-1}$, $C'_n = L_{n-1}$, $E'_n = E_{n-1}$ and that the species T of all *rooted trees* has an atomic decomposition whose first few terms are given by

$$T = X + X^2 + X^3 + XE_2 + 2X^4 + X^2E_2 + XE_3 + 3X^5 + 3X^3E_2 + X^2E_3 + X \cdot (E_2 \circ X^2) + XE_4 + \cdots. \tag{1.5}$$

Obviously, the combinatorial notion of molecular (or atomic) decomposition (1.4) can be "algebraically" extended by allowing the coefficients f_M to belong to an *arbitrary* half-ring $\mathbb{K}$; this gives rise to the half-ring $\mathbb{K}[[\mathcal{A}]]$. Using "linearity," the operations of cartesian product $\times$ and of derivation $'$ can then easily be extended to $\mathbb{K}[[\mathcal{A}]]$.

However, the problem of finding an analogous "coherent" extension of the operation of substitution $\circ$ is far from being trivial. It has been solved in a satisfactory manner only for special $\mathbb{K}$'s. A. Joyal [J2] solved it in the important case when $\mathbb{K} = \mathbb{Z}$ and, independently, Y. N. Yeh [Y] solved it in the more general situation when $\mathbb{K}$ is an arbitrary *binomial half-ring*; that is, a (commutative) half-ring $\mathbb{K}$ contained in a $\mathbb{Q}$-algebra $\mathbb{L}$ and such that

$$\forall t \in \mathbb{K}, \forall n \in \mathbb{N}, \qquad \binom{t}{n} = t(t-1)(t-2)\cdots(t-n+1)/n! \in \mathbb{K}. \tag{1.6}$$

For example, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{C}$, $\mathbb{Q}[i]$ and $\mathbb{N} + \mathbb{Q}\varepsilon$ (where $\varepsilon^2 = 0$) are all binomial half-rings while $\mathbb{F}_p$ (for prime p) and $\mathbb{Z}[i]$ are not.

From now on, $\mathbb{K}$ will always denote a binomial half-ring and the following terminology will be used: an element $F \in \mathbb{K}[[\mathcal{A}]]$ is a $\mathbb{K}$-*species* (in the sense of Yeh), a $\mathbb{Z}$-species is also called a *virtual species* and, of course, a $\mathbb{N}$-species is, simply, a *species*. The *order* and *degree* of any $F \in \mathbb{K}[[\mathcal{A}]]$ are the following elements of $\mathbb{N} \cup \{\infty\}$, respectively defined by

$$\operatorname{ord}(F) = \min\{n \in \mathbb{N} \mid F_n \neq 0\}, \qquad \deg(F) = \max\{n \in \mathbb{N} \mid F_n \neq 0\}. \tag{1.7}$$

Of course, $0 \leqslant \operatorname{ord}(F) \leqslant \deg(F) \leqslant \infty$ if $F \neq 0$, while $\infty = \operatorname{ord}(0) > \deg(0) = 0$. When $F = 0$ or $\deg(F) < \infty$, we say that F is a *polynomial* $\mathbb{K}$-species; in particular, and $M \in \mathcal{M}_n$ is *monomial* and satisfies $\operatorname{ord}(M) = \deg(M) = n < \infty$.

In Section 2 we use a direct combinatorial argument to obtain an alternate definition of Yeh's substitution $G \circ F$ of $\mathbb{K}$-species. Then, taking ideas from [L1, J3], we use the "delta" operator

$$\Delta_F \colon \mathbb{K}[[\mathcal{A}]] \to \mathbb{K}[[\mathcal{A}]], \qquad \text{where} \quad G \mapsto \Delta_F G = G \circ F - G, \tag{1.8}$$

to define the concept of *generalized iterates* $F^{\langle t\rangle}$, $t\in\mathbb{K}$, of any "normalized" $F=X+\cdots\in\mathbb{K}[[\mathcal{Ol}]]$; that is, a one-parameter semigroup $F^{\langle t\rangle}$, $t\in\mathbb{K}$, of $\mathbb{K}$-species, satisfying $F^{\langle 0\rangle}=X$, $F^{\langle 1\rangle}=F$, and

$$F^{\langle s+t\rangle}=F^{\langle s\rangle}\circ F^{\langle t\rangle} \qquad \text{for every} \quad s, t\in\mathbb{K}. \tag{1.9}$$

In Section 3 a more detailed analysis of these generalized iterates leads us, in a very natural manner, to other new notions among which we find:

— the *infinitesimal generator* gen(F) of F, which turns out to be a $\bar{\mathbb{K}}$-species, where $\bar{\mathbb{K}}$ denotes the *rational closure* of $\mathbb{K}$,

— the $\mathbb{Q}$-species $\hat{X}$ of *pseudo-singletons*, which is another kind of "logarithm" of the species E of sets,

— the *directional derivative* $D_H G\in\bar{\mathbb{K}}[[\mathcal{Ol}]]$ of the $\mathbb{K}$-species G in the "direction" of the $\mathbb{K}$-species H and

— the *Lie bracket* $[\Phi, \Gamma]$ of $\mathbb{K}$-species Φ and Γ.

Apart from their intrinsic interest, these notions give, in return, better insights into the concept of a $\mathbb{K}$-species (even in the case of a $\mathbb{N}$-species) and into substitution itself. For example, they imply that any normalized $\mathbb{K}$-species F can be written in the form $F=\exp(D_\Phi)X$ for a suitably chosen $\bar{\mathbb{K}}$-species Φ and that two normalized $\mathbb{K}$-species F, G commute under $\circ$ if and only if the Lie bracket of their infinitesimal generators vanishes.

2. Substitution and Generalized Iteration

Before introducing the "generalized iterates" of a $\mathbb{K}$-species, a preliminary combinatorial analysis of Yeh's substitution will be very useful. In fact, we shall give an alternate definition of Yeh's substitution which turns out to be a direct consequence of a new formula (see Theorem 2.3 below) for the *ordinary* substitution of species (i.e., $\mathbb{N}$-species).

Let $F=\sum_{M\in\mathcal{M}} f_M M$ be a species. Then each coefficient f_M is a natural integer that can be interpreted as a set of f_M "colours." If we write, as in classical set theory, $f_M=\{0, 1, ..., f_M-1\}$, then each such colour $k\in f_M$ can in turn be interpreted as a natural integer satisfying $0\leqslant k\leqslant f_M-1$.

Now, any finite set S can itself be thought of as being a species: a S-structure being simply an element $s\in S$ whose "underlying set" is, by convention, always the empty set $\varnothing$. In particular, each f_M can be interpreted as a species of colours (on the empty set). Using the standard definition of product, we see that a SM-structure is a M-structure that is "coloured" by an element $s\in S$.

With these conventions, an arbitrary $f_M M$-structure can be thought of as a *coloured M-structure*, whose colour is an element $k\in f_M$. Thus, taking the whole sum, a F-structure φ can be canonically represented as a *F-coloured molecular structure* (that is, φ is a M-structure, for some M, that is coloured with colour k, for some $k \in f_M$). Moreover, two F-structures φ, ψ are of the same *isomorphism type* (i.e., are transformable into one another using a suitable bijection between their respective underlying sets) iff they are both of the same colour and belong to the same molecular species.

It is not difficult to see that the subspecies H of F (notation: $H \subsetneq F$) are precisely those species that are of the form

$$H = \sum_{M \in \mathcal{M}} S_M M, \qquad \text{where} \quad S_M \subseteq f_M \tag{2.1}$$

for every $M \in \mathcal{M}$. Collecting these preliminary remarks we can state:

LEMMA 2.1. *Let $F = \sum_{M \in \mathcal{M}} f_M M$ and $P = \sum_{M \in \mathcal{M}} p_M M$ be arbitrary. Then the (possibly infinite) number of distinct subspecies H of F which are isomorphic to P is given by the "generalized binomial coefficient"*

$$\binom{F}{P} := \prod_{M \in \mathcal{M}} \binom{f_M}{p_M}. \tag{2.2}$$

Proof. Subspecies H of F are characterized by (2.1). Now H is isomorphic to P iff card $S_M = p_M$ for each $M \in \mathcal{M}$. Hence the result. ∎

Note that $\binom{F}{P} \neq 0$ iff $p_M \leqslant f_M$ for every $M \in \mathcal{M}$, and that $\binom{F}{P} < \infty$ if P is a polynomial species (i.e., $p_M \neq 0$ only for a finite number of $M \in \mathcal{M}$).

Let now F, G, and H be species such that H is polynomial and $H \subsetneq F$. We can extract a subspecies $G \bullet H$ of $G \circ F$, using the following

DEFINITION 2.2. A given G-assembly γ of F-structures (i.e., a $G \circ F$-structure) is *H-saturated* if and only if the set of types of its members coincides with the set of types of H-structures. The species of all H-saturated $G \circ F$-structures is denoted by $G \bullet H$. The operation $\bullet$ is called *saturated substitution.*

Note. By (2.1), using colourings, an equivalent definition of H-saturation can be stated as follows: The assembly γ is H-saturated iff for each $M \in \mathcal{M}$, S_M coincides with the set of all those colours that are used (at least once) to "paint" the M-structures that appear in the assembly γ.

We are now in a position to state a representation theorem, concerning the substitution of species, which will turn out to be the key tool in our approach to Yeh's extension of substitution.

THEOREM 2.3. *Let G and F be two species such that $F_0 = 0$. Then*

$$G \circ F = \sum_{P \in \mathscr{P}} \binom{F}{P} G \bullet P, \tag{2.3}$$

where $\mathscr{P} = \mathbb{N}[\mathscr{A}]$ denotes the set of all polynomial $\mathbb{N}$-species.

Proof. The condition $F_0 = 0$ is standard and necessary in order to have a finite number of $G \circ F$-structures on any finite set [J1]. Now, write $F = \sum_{M \in \mathscr{M}} f_M M$ with $f_M \in \mathbb{N}$ and consider an arbitrary $G \circ F$-structure γ. For each $M \in \mathscr{M}$ let $S_M \subseteq f_M$ be the set of all those colours that are used (at least once) to "paint" the M-structures that appear in the assembly γ. Of course, $H = \sum_{M \in \mathscr{M}} S_M M$ is a polynomial subspecies of F and γ is a $G \bullet H$-structure (by Definition 2.2). Hence, we can write the decomposition

$$G \circ F = \sum_{H \in \mathscr{H}} G \bullet H, \tag{2.4}$$

where $\mathscr{H} = \{H \subsetneq F \mid H \text{ is a polynomial species}\}$. We then conclude by an application of Lemma 2.1, since each species $G \bullet H$ is isomorphic to $G \bullet P$ if $P = \sum_{M \in \mathscr{M}} p_M M \in \mathscr{P}$, where $p_M = \text{card } S_M$. ∎

In analogy with classical analysis, $G \circ F$ is often written as $G(F)$. The species X of all singletons is the neutral element under substitution. This means that we always have $G(X) = G \circ X = G = X \circ G = X(G)$. The importance of representation formula (2.3) lies in the fact that it immediately leads to explicit *polynomial expressions* for the coefficients of the molecular decomposition of $G \circ F$ in terms of those of F and G:

COROLLARY 2.4. *Let $F = \sum_{M \in \mathscr{M}} f_M M$ and $G = \sum_{N \in \mathscr{M}} g_N N$ be any species such that $F_0 = 0$. Then $G \circ F = \sum_{R \in \mathscr{M}} h_R R$, where each coefficient h_R can be written as a* finite sum *of the form*

$$h_R = \sum_{N \in \mathscr{M}, P \in \mathscr{P}} c_{N,P,R}\, g_N \prod_{M \in \mathscr{M}} \binom{f_M}{p_M} \tag{2.5}$$

in which the $c_{N,P,R} \in \mathbb{N}$ are independent of F and G and are defined by

$$N \bullet P = \sum_{R \in \mathscr{M}} c_{N,P,R} R \tag{2.6}$$

which is also a finite sum.

Proof. For every $N \in \mathscr{M}$ and every $P \in \mathscr{P}$, the species $N \bullet P$ is of finite degree. Its molecular decomposition is thus finite and can be written in the form (2.6) thereby defining the constants $c_{N,P,R} \in \mathbb{N}$. Formula (2.5) for h_R follows from right distributivity of $\bullet$ over $+$:

$$\begin{aligned} G \circ F &= \sum_{P \in \mathscr{P}} \binom{F}{P} G \bullet P \\ &= \sum_{P \in \mathscr{P}} \binom{F}{P} \sum_{N \in \mathscr{M}} g_N N \bullet P \\ &= \sum_{P \in \mathscr{P}} \binom{F}{P} \sum_{N \in \mathscr{M}} g_N \sum_{R \in \mathscr{M}} c_{N,P,R} R \\ &= \sum_{R \in \mathscr{M}} \left(\sum_{N \in \mathscr{M}, P \in \mathscr{P}} c_{N,P,R}\, g_N \binom{F}{P} \right) R. \end{aligned} \tag{2.7}$$

Now, write $P = \sum_{M \in \mathscr{M}} p_M M$. Then the finitude of sum (2.5) follows from the fact that $F_0 = 0$ and $c_{N,P,R} \neq 0$ implies

$$\deg R \geqslant \deg N \geqslant \sum_{M \in \mathscr{M}} p_M \qquad \text{and} \qquad \deg R \geqslant \sum_{M \in \mathscr{M}} p_M \deg M, \tag{2.8}$$

as is easily verified from the definition of $N \bullet P$. ∎

Note that the finitude of sum (2.5) shows that each h_R is an element of the polynomial half-ring

$$\mathbb{N}\left[(g_N)_{N \in \mathscr{M}}, \binom{f_M}{p}_{M \in \mathscr{M}, p \in \mathbb{N}} \right] \tag{2.9}$$

when each $\binom{f_M}{p}$ and g_N is considered as a formal variable. This fact (first noticed by Yeh [Y], using a different approach) implies that each f_M and g_N can take values in an *arbitrary binomial half-ring* $\mathbb{K}$ and the resulting h_R will also belong to $\mathbb{K}$, for every $R \in \mathscr{M}$. Substitution can thus be extended to $\mathbb{K}[[\mathscr{A}]]$ as follows:

THEOREM 2.5. *Let* $\mathbb{K}$ *be a binomial half-ring and* $G, F \in \mathbb{K}[[\mathscr{A}]]$ *be* $\mathbb{K}$*-species such that* $F_0 = 0$. *Then the operation* $\circ$ *defined by*

$$G \circ F = \sum_{P \in \mathscr{P}} \binom{F}{P} G \bullet P, \qquad G \bullet P = \sum_{N \in \mathscr{M}} g_N N \bullet P \tag{2.10}$$

extends the usual substitution between species and coincides with Yeh's extension.

Proof. This is almost immediate: In Yeh's approach, each h_R (where $R \in \mathscr{M}$) can also be written as a polynomial belonging to the ring (2.9). This polynomial must coincide with polynomial (2.5), since they both take the same value when the variables f_M and g_N take arbitrary non-negative integral values (in the case of $\mathbb{N}$-species). ∎

Of course, the usual identities relating $\circ$, $+$, $\cdot$, $\times$, $'$ (e.g., chain rule, Leibniz rule, etc.) remain valid in $\mathbb{K}[[\mathcal{A}]]$, since the coefficients of their molecular decompositions also involve polynomial expressions (similar to those used in the above proof). Note also that saturated substitution can be expressed in terms of substitution as follows: For $P \in \mathcal{P}$ and $G \in \mathbb{K}[[\mathcal{A}]]$,

$$G \bullet P = \sum_{Q \in \mathcal{P}} (-1)^{|P-Q|} \binom{P}{Q} G \circ Q, \tag{2.11}$$

where $|P-Q| = \sum_{M \in \mathcal{M}} |p_M - q_M|$. This can be seen by adapting the standard technique [R] of (binomial) Möbius inversion to (2.10) with $F = Q \in \mathcal{P}$.

The various infinite families of species (or $\mathbb{K}$-species) appearing in sums (1.1), (1.4), (1.5), (2.1), (2.3), (2.4), (2.7), and (2.10) are all "summable," by construction. These are special instances of

DEFINITION 2.6. Let Λ be an arbitrary set. A family $(F^\lambda)_{\lambda \in \Lambda}$ of $\mathbb{K}$-species is *summable* iff for each finite cardinal $n \in \mathbb{N}$, the set

$$\Lambda_n = \{\lambda \in \Lambda \mid F_n^\lambda \neq 0\} \tag{2.12}$$

is finite. The *sum* $F = \sum_{\lambda \in \Lambda} F^\lambda$ of the family $(F^\lambda)_{\lambda \in \Lambda}$ is the $\mathbb{K}$-species given by

$$F_n = \left(\sum_{\lambda \in \Lambda} F^\lambda\right)_n = \sum_{\lambda \in \Lambda_n} F_n^\lambda, \qquad n = 0, 1, 2, ... \tag{2.13}$$

Other important instances of summable families involve the iterates of the following "delta operators."

DEFINITION 2.7. Let F be a $\mathbb{K}$-species such that $F_0 = 0$. The *delta operator associated to F* is the linear operator

$$\Delta_F\colon \mathbb{K}[[\mathcal{A}]] \to \mathbb{K}[[\mathcal{A}]] : G \mapsto \Delta_F G = G \circ F - G. \tag{2.14}$$

Such operators have been used in [L1] to study the generalized iterates of formal power series and also in [J3] to obtain an explicit virtual species $F^{\langle -1 \rangle} \in \mathbb{Z}[[\mathcal{A}]]$ that is the inverse, under $\circ$, of any species of the form $F = X + \cdots \in \mathbb{N}[[\mathcal{A}]]$. In the present context, they will now be used to define and study the generalized iterates $F^{\langle t \rangle} \in \mathbb{K}[[\mathcal{A}]]$, $t \in \mathbb{K}$, of any $F \in \mathbb{K}[[\mathcal{A}]]$ of the form $F = X + \cdots$. Such F's are called *normalized* $\mathbb{K}$-species and are characterized by the "normalization conditions" $F_0 = 0$ and $F_1 = X$. For example, the above $\mathbb{N}$-species (1.5) of rooted trees is normalized.

LEMMA 2.8. *Let $F = X + \cdots$ be any normalized $\mathbb{K}$-species and $(a_k)_{k \geqslant 0}$ be any sequence of elements of $\mathbb{K}$. Then for any $G \in \mathbb{K}[[\mathscr{A}]]$, the sequence $(a_k \Delta_F^k G)_{k \geqslant 0}$ of $\mathbb{K}$-species is summable (here, $\Delta_F^k G$ is recursively defined by $\Delta_F^0 G = G$ and $\Delta_F^{k+1} G = \Delta_F \Delta_F^k G$, $k = 0, 1, \ldots$).*

Proof. Consider first the particular situation when G, F are $\mathbb{N}$-species and $a_k \in \mathbb{N}$. To establish summability in this case, it is sufficient to show that

$$\operatorname{ord} \Delta_F^k G \geqslant k + 1, \qquad k = 1, 2, 3, \ldots. \tag{2.15}$$

This has been done by A. Joyal who also gave an explicit combinatorial description of the species $\Delta_F^k G$, $k = 0, 1, 2, \ldots$ in terms of "strictly increasing chains of equivalence relations" (see [J3], pp. 143, 144 for more details). In the general situation when G, F are $\mathbb{K}$-species and $a_k \in \mathbb{K}$, inequalities (2.15) are still valid. This can be seen as follows. Fix $k \geqslant 1$ and consider the decomposition

$$a_k \Delta_F^k G = a_k \sum_{M \in \mathscr{M}} \theta_M M, \qquad \theta_M \in \mathbb{K}. \tag{2.16}$$

Because of Corollary 2.4, each θ_M is a polynomial belonging to the ring (2.9). Now, fix $M \in \mathscr{M}$ such that $\operatorname{ord} M \leqslant k$, then (since (2.15) is valid for $\mathbb{N}$-species) the polynomial θ_M takes the value 0 whenever the "variables" f_N, g_N, with $N \in \mathscr{M}$, take values in $\mathbb{N}$. Hence θ_M is identically 0. ∎

Note. In the case of species, (2.15) is a direct consequence of the slightly more general statement

$$H \in \mathbb{N}[[\mathscr{A}]], k \in \mathbb{N}, \qquad \operatorname{ord} H \geqslant k \Rightarrow \operatorname{ord} \Delta_F H \geqslant k + 1, \tag{2.17}$$

which can be seen "geometrically" as follows: Write $F = X + F^+$ where $\operatorname{ord} F^+ \geqslant 2$ and suppose that $\operatorname{ord} H \geqslant k$. Then we have

$$\Delta_F H = H \circ (X + F^+) - H = H(X + Y) - H(X)|_{Y := F^+}, \tag{2.18}$$

where Y is a new sort of points. Now, using the standard graphical conventions concerning species, a $[H(X + Y) - H(X)]$-structure can be described as in Fig. 1, where at least one Y-point appears among at least k X- or Y-points. Hence, because of (2.18), a $\Delta_F H$-structure is an H-assembly of

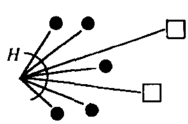

FIG. 1. $[H(X + Y) - H(X)]$-structure, where X: ● and Y: □ (at least one □).

X-points or F^+-structures that contains *at least one* F^+-structure (see Fig. 2). So that $\operatorname{ord} \Delta_F H \geqslant k+1$, since every F^+-structure contains at least two points.

We are now in a position to introduce the generalized iterates $F^{\langle t \rangle}$ of any normalized $\mathbb{K}$-species F.

THEOREM 2.9. *Let $\mathbb{K}$ be a binomial half-ring, $F = X + \cdots \in \mathbb{K}[[\mathcal{A}]$ be a normalized $\mathbb{K}$-species and $\mathcal{M}_+ = \mathcal{M} \backslash \{1\}$. Then there exists a unique family of $\mathbb{K}$-species*

$$F^{\langle t \rangle} = \sum_{M \in \mathcal{M}_+} f_M(t) M, \qquad t \in \mathbb{K}, \tag{2.19}$$

where the coefficients $f_M(t)$ are polynomials in t and are such that

$$n \in \mathbb{N} \Rightarrow F^{\langle n \rangle} = F \circ F \circ \cdots \circ F \qquad (n\text{th } \textit{iterate of } F) \tag{2.20}$$

and

$$s, t \in \mathbb{K} \Rightarrow F^{\langle s+t \rangle} = F^{\langle s \rangle} \circ F^{\langle t \rangle}. \tag{2.21}$$

Moreover, $\deg f_M(t) < \deg M$ *for every* $M \in \mathcal{M}_+$ *and, for any* $G \in \mathbb{K}[[\mathcal{A}]]$, *the* $\mathbb{K}$*-species* $G \circ F^{\langle t \rangle}$ *may be computed by the formula*

$$G \circ F^{\langle t \rangle} = \sum_{k \geqslant 0} \binom{t}{k} \Delta_F^k G. \tag{2.22}$$

Proof. Unicity of the polynomials $f_M(t)$: Let $\varphi_M(t)$ be any other such family of polynomials. Then, because of (2.20), each polynomial $\delta_M(t) = f_M(t) - \varphi_M(t)$ must satisfy $\delta_M(n) = 0$ for every $n \in \mathbb{N}$. Hence, $\delta_M(t)$ must be identically zero. Existence of the polynomials $f_M(t)$: Define $F^{\langle t \rangle}$ by the formula

$$F^{\langle t \rangle} = \sum_{k \geqslant 0} \binom{t}{k} \Delta_F^k X. \tag{2.23}$$

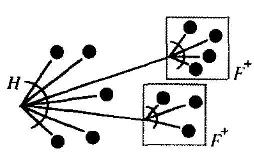

FIG. 2. $\Delta_F H$-structure, where x and y are as in Fig. 1 (at least one F^+-structure).

By Lemma 2.8, this gives a $\mathbb{K}$-species for every $t \in \mathbb{K}$, since the family $\binom{t}{k} \Delta_F^k X$, $k \in \mathbb{N}$, is summable. Moreover,

$$F^{\langle t \rangle} = \sum_{n \geqslant 1} \left(\sum_{k < n} \binom{t}{k} (\Delta_F^k X)_n \right), \tag{2.24}$$

since $(\Delta_F^k X)_n = 0$ if $k \geqslant n$. Hence, for every $M \in \mathcal{M}_+$ with $n = \deg M$, each coefficient $f_M(t)$ of the molecular decomposition of $F^{\langle t \rangle}$ is a linear combination (with integral coefficients) of the binomial coefficients $\binom{t}{k}$ with $k < n$. That is, $f_M(t)$ is a polynomial in t satisfying $\deg f_M(t) < \deg M$. It is easy to check, by induction on n, that (2.20) holds. Finally, (2.21) is a consequence of the classical "Vandermonde formula"

$$\binom{s+t}{k} = \sum_{i+j=k} \binom{s}{i} \binom{t}{j} \tag{2.25}$$

which is valid in any binomial half-ring. The more general formula (2.22) follows by a similar argument. ▌

Of course, when $\mathbb{K} = \mathbb{Z}$, $t = -1$, and $F = X + \cdots \in \mathbb{N}[[\mathcal{A}]]$ then we have $\binom{t}{k} = (-1)^k$ and

$$F^{\langle -1 \rangle} = X - \Delta_F X + \Delta_F^2 X - \Delta_F^3 X + \cdots \tag{2.26}$$

coincides with the virtual species, given by Joyal [J3], which is the inverse of F under the operation of substitution.

EXAMPLE. The reader may check that for the species T of all rooted trees, given by (1.5) above, we have, up to degree 5:

$$\begin{aligned}
\Delta_T X &= X^2 + X^3 + XE_2 + 2X^4 + X^2E_2 + XE_3 + 3X^5 + 3X^3E_2 \\
&\quad + X^2E_3 + X \cdot (E_2 \circ X^2) + XE_4 + \cdots \\
\Delta_T^2 X &= 2X^3 + 7X^4 + 3X^2E_2 + 23X^5 + 12X^3E_2 \\
&\quad + 3X^2E_3 + X \cdot (E_2 \circ X^2) + XE_2^2 + \cdots \\
\Delta_T^3 X &= 6X^4 + 43X^5 + 12X^3E_2 + \cdots \qquad \text{and} \qquad \Delta_T^4 X = 24X^5 + \cdots.
\end{aligned} \tag{2.27}$$

This can be done by noting that the species E_n of all sets of cardinality n satisfies the combinatorial equation

$$E_n \circ (A + B) = \sum_{i+j=n} (E_i \circ A) \cdot (E_j \circ B), \tag{2.28}$$

where A and B are arbitrary species (or $\mathbb{K}$-species), with $A_0 = B_0 = 0$. Hence

$$
\begin{aligned}
T^{\langle t\rangle} = {} & X + \binom{t}{1} X^2 + \left[\binom{t}{1} + 2\binom{t}{2}\right] X^3 + \binom{t}{1} XE_2 \\
& + \left[2\binom{t}{1} + 7\binom{t}{2} + 6\binom{t}{3}\right] X^4 \\
& + \left[\binom{t}{1} + 3\binom{t}{2}\right] X^2E_2 + \binom{t}{1} XE_3 \\
& + \left[3\binom{t}{1} + 23\binom{t}{2} + 43\binom{t}{3} + 24\binom{t}{4}\right] X^5 \\
& + \left[3\binom{t}{1} + 12\binom{t}{2} + 12\binom{t}{3}\right] X^3E_2 + \left[\binom{t}{1} + 3\binom{t}{2}\right] X^2E_3 \\
& + \left[\binom{t}{1} + \binom{t}{2}\right] X\cdot(E_2 \circ X^2) + \binom{t}{1} XE_4 + \binom{t}{2} XE_2^2 + \cdots. \qquad (2.29)
\end{aligned}
$$

It is well known [J1, L2, L4] that the species T is characterized by the combinatorial functional equation $T = X\cdot(E\circ T)$, where E denotes the species of all finite sets. Hence, the virtual species $T^{\langle -1\rangle}$ must satisfy the equation $X = T^{\langle -1\rangle}\cdot E$ and can be decomposed as

$$
\begin{aligned}
T^{\langle -1\rangle} &= X/E = X/(1 + X + E_2 + E_3 + E_4 + \cdots) \\
&= X - X^2 + X^3 - XE_2 - X^4 + 2X^2E_2 - XE_3 + X^5 \\
&\quad - 3X^3E_2 + 2X^2E_3 - XE_4 + XE_2^2 + \cdots. \qquad (2.30)
\end{aligned}
$$

It is easy to check that this last formula coincides with (2.29) when $t = -1$.

Going down to the level of the underlying generating series, formulas (2.29) and (2.30) "collapse" to two formal power series in x

$$
\begin{aligned}
T^{\langle t\rangle}(x) = {} & x + 2\binom{t}{1} x^2/2! + \left[9\binom{t}{1} + 12\binom{t}{2}\right] x^3/3! \\
& + \left[64\binom{t}{1} + 204\binom{t}{2} + 144\binom{t}{3}\right] x^4/4! \\
& + \left[625\binom{t}{1} + 3630\binom{t}{2} + 5880\binom{t}{3} + 2880\binom{t}{4}\right] x^5/5! + \cdots \qquad (2.31)
\end{aligned}
$$

and

$$
T^{\langle -1\rangle}(x) = x - 2x^2/2! + 3x^3/3! - 4x^4/4! + 5x^5/5! + \cdots = xe^{-x} \qquad (2.32)
$$

in accordance with the usual theory (see [L1], for example) of the generalized iterates of normalized power series.

3. Infinitesimal Generators and Directional Derivatives

The one-parameter semigroup $(F^{\langle t\rangle})_{t\in\mathbb{K}}$ of generalized iterates introduced in Theorem 2.9 above can be interpreted as a "parametrized curve" in the "space" $\mathbb{K}[[\mathcal{A}]]$. For obvious reasons, it is natural to inquire about the family $((d/dt)\,F^{\langle t\rangle})_{t\in\mathbb{K}}$ of all its "tangent vectors." Here, d/dt denotes the operator of formal derivation with respect to t defined, for $k\geqslant 0$, by

$$(d/dt)\binom{t}{k}=(1/k!)\sum_{0\leqslant i\leqslant k} t(t-1)\cdots(t-i+1)(t-i-1)\cdots(t-k+1). \tag{3.1}$$

But there is an obstruction to this, namely: while $\binom{t}{k}\in\mathbb{K}$ for any $t\in\mathbb{K}$, we do not necessarily have $(d/dt)\binom{t}{k}\in\mathbb{K}$. For example, if $\mathbb{K}=\mathbb{N}$, $k=2$, then $(d/dt)\binom{t}{2}=t-\frac{1}{2}\notin\mathbb{N}$, for any $t\in\mathbb{N}$. This motivates the following definition:

Definition 3.1. Let $\mathbb{K}$ be a binomial half-ring. The *rational closure* $\overline{\mathbb{K}}$ of $\mathbb{K}$ is the smallest $\mathbb{Q}$-algebra containing $\mathbb{K}$.

Of course, $\overline{\mathbb{N}}=\mathbb{Q}$, $\overline{\mathbb{K}}$ is a binomial ring, $\overline{\overline{\mathbb{K}}}=\overline{\mathbb{K}}$ and

$$\overline{\mathbb{K}}=\bigcap_{\mathbb{L}\supseteq\mathbb{K}}\mathbb{L}\qquad(\text{where }\mathbb{L}\text{ is a }\mathbb{Q}\text{-algebra}) \tag{3.2}$$

$$=\{r_1t_1+\cdots+r_nt_n \mid n\in\mathbb{N},\, r_i\in\mathbb{Q},\, t_i\in\mathbb{K}, \text{ for } i=1,\ldots,n\}. \tag{3.3}$$

Moreover, the following inclusions always hold

$$\begin{tikzcd}
\mathbb{N}[[X]] \arrow[rr, hook] \arrow[dr, hook] \arrow[dd, hook] & & \mathbb{Q}[[X]] \arrow[dr, hook] & \\
& \mathbb{N}[[\mathcal{A}]] \arrow[rr, hook] \arrow[dd, hook] & & \mathbb{Q}[[\mathcal{A}]] \arrow[dd, hook] \\
\mathbb{K}[[X]] \arrow[rr, hook] \arrow[dr, hook] & & \overline{\mathbb{K}}[[X]] \arrow[dr, hook] & \\
& \mathbb{K}[[\mathcal{A}]] \arrow[rr, hook] & & \overline{\mathbb{K}}[[\mathcal{A}]]
\end{tikzcd} \tag{3.4}$$

and preserve each one of the operations $+$, $\cdot$, $\times$, $\circ$, $'$. An easy computation shows that

$$(d/dt)\binom{t}{k}\bigg|_{t:=0} = (-1)^{k-1}/k \text{ if } k \geqslant 1 \qquad (=0 \text{ if } k=0). \tag{3.5}$$

DEFINITION 3.2. Let $F = X + \cdots \in \mathbb{K}[[\mathcal{A}]]$ be a normalized $\mathbb{K}$-species. The *infinitesimal generator* $\operatorname{gen} F$ of F is the $\bar{\mathbb{K}}$-species

$$\begin{aligned} \operatorname{gen} F &= (d/dt)\, F^{\langle t\rangle}|_{t:=0} \\ &= \Delta_F X - \tfrac{1}{2}\Delta_F^2 X + \tfrac{1}{3}\Delta_F^3 X - \tfrac{1}{4}\Delta_F^4 X + \cdots. \end{aligned} \tag{3.6}$$

EXAMPLES. For the species X of all singletons, we have $\operatorname{gen} X = 0$. The infinitesimal generator of the species $L^* = X + X^2 + X^3 + \cdots$ of all non-empty linear orders is the $\mathbb{N}$-species $\operatorname{gen} L^* = X^2$. The infinitesimal generator of the species T of all rooted trees is the $\mathbb{Q}$-species

$$\begin{aligned} \operatorname{gen} T &= \Delta_T X - \tfrac{1}{2}\Delta_T^2 X + \tfrac{1}{3}\Delta_T^3 X - \tfrac{1}{4}\Delta_T^4 X + \cdots \\ &= X^2 + XE_2 + \tfrac{1}{2}X^4 - \tfrac{1}{2}X^2E_2 + XE_3 - \tfrac{1}{6}X^5 + X^3E_2 \\ &\quad - \tfrac{1}{2}X^2E_3 + \tfrac{1}{2}X\cdot(E_2\circ X^2) + XE_4 - \tfrac{1}{2}XE_2^2 + \cdots. \end{aligned} \tag{3.7}$$

Note that $\operatorname{gen} L^*$ and $\operatorname{gen} T$ are of order 2 and that, in general, $\operatorname{ord} \operatorname{gen} F \geqslant 2$ for any normalized F. It turns out (see Theorem 3.5 below) that F is completely determined by its infinitesimal generator $\Phi = \operatorname{gen} F$. To establish this fact, we need a new notion: the directional derivative of a $\bar{\mathbb{K}}$-species.

DEFINITION 3.3. Let $\bar{\mathbb{K}}$ be the rational closure of the binomial half-ring $\mathbb{K}$ and $G, H \in \bar{\mathbb{K}}[[\mathcal{A}]]$ with $H_0 = 0$. The *directional derivative of* G *in the direction* H is the unique $\bar{\mathbb{K}}$-species $D_H G$ satisfying the condition

$$G(X + tH) = G(X) + tD_H G + O(t^2) \in \bar{\mathbb{K}}[[\mathcal{A}]], \tag{3.8}$$

where $O(t^2)$ denotes a $\bar{\mathbb{K}}$-species in which t^2 can be factored out.

THEOREM 3.4. *For any H with $H_0 = 0$, the operation*

$$D_H\colon \bar{\mathbb{K}}[[\mathcal{A}]] \to \bar{\mathbb{K}}[[\mathcal{A}]]\colon G \mapsto D_H G \tag{3.9}$$

is a derivation which can be computed by

$$\begin{aligned} D_H G &= (d/dt)\, G(X + tH)|_{t:=0} \\ &= \sum_{P\in\mathcal{P}} \delta_{H,P} G \bullet P, \end{aligned} \tag{3.10}$$

where the coefficients $\delta_{H,P} \in \bar{\mathbb{K}}$ *are independent of G and are given by*

$$\delta_{H,P} = (d/dt) \binom{X+tH}{P} \Big|_{t:=0} \in \bar{\mathbb{K}}. \tag{3.11}$$

Moreover, D_H *is also linear in H.*

Proof. This follows from Definition 3.3 and Theorem 2.3. ∎

Note that making use of the "Leibnitz' rule,"

$$D_H(F \cdot G) = (D_H F) \cdot G + F \cdot (D_H G), \tag{3.12}$$

the computation of $D_H G$ can be reduced to the computation of "simpler" directional derivatives of the form $D_M A$, where M and A run through the *molecular* and the *atomic* $\mathbb{N}$-species, respectively.

THEOREM 3.5. *Let* $G, F \in \mathbb{K}[[\mathcal{A}]]$, *F normalized, and* $\Phi = \operatorname{gen} F$. *Then the families* $(F^{\langle t \rangle})_{t \in \mathbb{K}}$ *and* $(G \circ F^{\langle t \rangle})_{t \in \mathbb{K}}$ *satisfy the differential equations in t,*

$$(d/dt)\, F^{\langle t \rangle} = D_\Phi F^{\langle t \rangle} = \Phi \circ F^{\langle t \rangle} \tag{3.13}$$

and

$$(d/dt)(G \circ F^{\langle t \rangle}) = D_\Phi (G \circ F^{\langle t \rangle}) = (D_\Phi G) \circ F^{\langle t \rangle}. \tag{3.14}$$

Moreover,

$$F^{\langle t \rangle} = e^{tD_\Phi} X, \qquad G \circ F^{\langle t \rangle} = e^{tD_\Phi} G. \tag{3.15}$$

Finally, $F \mapsto \operatorname{gen} F$ *defines an injection from the set of all normalized* $\mathbb{K}$*-species into the set of all* $\bar{\mathbb{K}}$*-species of order* $\geqslant 2$.

Proof. We simply use Theorem 2.3, Definition 3.3, and the analogy with classical analysis. Let s, t be formal variables taking values in $\bar{\mathbb{K}}$. Since $\Phi = (d/ds)\, F^{\langle s \rangle}|_{s:=0}$, we can write $F^{\langle s \rangle}$ in the form $F^{\langle s \rangle} = X + s\Phi + O(s^2)$. To establish (3.14), we need only to compute the following three equivalent expressions for $G \circ F^{\langle t+s \rangle}$ (which are valid in $\bar{\mathbb{K}}[[\mathcal{A}]]$):

$$G \circ F^{\langle t+s \rangle} = G \circ F^{\langle t \rangle} + s(d/dt)(G \circ F^{\langle t \rangle}) + O(s^2), \tag{3.16}$$

$$\begin{aligned} G \circ F^{\langle t+s \rangle} &= (G \circ F^{\langle t \rangle}) \circ F^{\langle s \rangle} \\ &= (G \circ F^{\langle t \rangle}) \circ (X + s\Phi + \cdots) \\ &= G \circ F^{\langle t \rangle} + sD_\Phi (G \circ F^{\langle t \rangle}) + O(s^2), \end{aligned} \tag{3.17}$$

and

$$\begin{aligned} G\circ F^{\langle t+s\rangle} &= (G\circ F^{\langle s\rangle})\circ F^{\langle t\rangle} \\ &= (G\circ(X+s\Phi+\cdots))\circ F^{\langle t\rangle} \\ &= (G+sD_\Phi G+\cdots)\circ F^{\langle t\rangle} \\ &= G\circ F^{\langle t\rangle}+s(D_\Phi G)\circ F^{\langle t\rangle}+O(s^2). \end{aligned} \tag{3.18}$$

Equations (3.13) follow from (3.14) with $G=X$. Formulas (3.15) are the Taylor expansions, in t, of the solutions of the differential equations (3.13) and (3.14). The injectivity of $F\mapsto \operatorname{gen} F$ follows from the sustitution $t:=1$ in (3.15) which expresses F (and $G\circ F$) in terms of Φ (and G) as

$$F=e^{D_\Phi}X, \qquad G\circ F=e^{D_\Phi}G. \quad \blacksquare \tag{3.19}$$

Remarks. Let $\Phi=\operatorname{gen} F$ then, because of (3.19), each operator Δ_F and D_Φ may be expressed in terms of the other,

$$\Delta_F=e^{D_\Phi}-I, \qquad D_\Phi=\ln(I+\Delta_F), \tag{3.20}$$

where $I\colon \mathbb{K}[[\mathcal{C}]]\to\mathbb{K}[[\mathcal{C}]]$ denotes the identity operator. There is an obvious analogy here with the following classical identities taken from the calculus of finite differences,

$$\Delta=e^D-I, \qquad D=\ln(I+\Delta), \tag{3.21}$$

where

$$\Delta G(X)=G(X+1)-G(X), \qquad DG(X)=(d/dX)\,G(X)=G'(X). \tag{3.22}$$

Indeed, take the *polynomial* $F=X+1\in\mathbb{K}[X]$ then $F^{\langle t\rangle}=X+t\in\mathbb{K}[X]$ and $\Phi=\operatorname{gen} F=1\in\mathbb{K}[X]$. This gives

$$\Delta_{X+1}=\Delta\colon \mathbb{K}[X]\to\mathbb{K}[X], \qquad D_1=D\colon \mathbb{K}[X]\to\mathbb{K}[X]. \tag{3.23}$$

However, (3.20) *is a strict extension of* (3.21) and one must take great care when making such analogies. For example, the classical identity

$$(D_HG)(X)=G'(X)\cdot H(X) \qquad \text{for} \quad G,H\in\mathbb{K}[[X]],\ H_0=0, \tag{3.24}$$

is no longer valid in $\mathbb{K}[[\mathcal{C}]]$. This means that, in general,

$$(D_HG)(X)\neq G'(X)\cdot H(X) \qquad \text{for} \quad G,H\in\mathbb{K}[[\mathcal{C}]],\ H_0=0 \tag{3.25}$$

even though D_HG and $G'\cdot H$ always have the *same* underlying generating series.

The simplest illustration of this phenomenon is obtained by taking $H = X$, the species of all singletons and $G = E_2$, the species of all 2-elements sets. From Theorem 3.4, we get the $\mathbb{Q}$-species

$$D_H G = D_X E_2 = \tfrac{1}{2}X^2 + E_2 \tag{3.26}$$

which is distinct from

$$G' \cdot H = E_2' \cdot X = X^2. \tag{3.27}$$

The common underlying generating series for (3.26) and (3.27) is $2x^2/2!$.

Apart from (3.10), there are other interesting explicit expressions for the directional derivatives $D_H G$ (see Theorem 3.8 below). One of these expressions makes use of a new $\mathbb{Q}$-species, $\hat{X}$, which is defined as

DEFINITION 3.6. Let $E^* = E^*(X)$ be the species of all non-empty (finite) sets. The $\mathbb{Q}$-species $\hat{X}$ of *pseudo-singletons* (*of sort* X) is the infinite summable series

$$\hat{X} = E^* - \tfrac{1}{2}(E^*)^2 + \tfrac{1}{3}(E^*)^3 - \tfrac{1}{4}(E^*)^4 + \cdots \tag{3.28}$$

$$\begin{aligned} &= X + (E_2 - \tfrac{1}{2}E_1^2) + (E_3 - E_1E_2 + \tfrac{1}{3}E_1^3) \\ &\quad + (E_4 - \tfrac{1}{2}E_2^2 - E_1E_3 + E_1^2E_2 - \tfrac{1}{4}E_1^4) + \cdots \\ &\quad + \left(\sum (-1)^{\nu_1+\nu_2+\cdots-1}((\nu_1+\nu_2+\cdots-1)!/\nu_1!\,\nu_2!\cdots)\,E_1^{\nu_1}E_2^{\nu_2}\cdots\right) + \cdots, \end{aligned} \tag{3.29}$$

where, for each $n \in \mathbb{N}$, the sum is extended in the general term over all the ν_i's satisfying $\nu_1 + 2\nu_2 + 3\nu_3 + \cdots = n$.

This species is obtained by substituting $E^* = E_1 + E_2 + E_3 + \cdots$ for X in the classical formal *power* series

$$\ln(1+X) = X - \tfrac{1}{2}X^2 + \tfrac{1}{3}X^3 - \tfrac{1}{4}X^4 + \cdots \in \mathbb{Q}[[X]] \subseteq \mathbb{K}[[\mathscr{A}]]. \tag{3.30}$$

Hence, we may write

$$\hat{X} = \ln(1 + E^*) = \ln E. \tag{3.31}$$

The $\mathbb{Q}$-species (3.30) should not be confused with the "combinatorial" logarithm, defined by A. Joyal [J3] and Y.-N. Yeh [Y], which is a $\mathbb{Z}$-species $\Lambda(X) = \mathrm{Lg}(1+X)$ satisfying $\Lambda \circ E^* = X$. This distinction being made, $\hat{X}$ may be called the "analytical" logarithm of the species of all sets. Its properties are very similar to those of the species X of all singletons.

LEMMA 3.7. *The generating series of $\hat{X}$ is x. The cycle indicator series $Z_{\hat{X}}$ of $\hat{X}$ is given by*

$$Z_{\hat{X}} = x_1 + \tfrac{1}{2}x_2 + \tfrac{1}{3}x_3 + \tfrac{1}{4}x_4 + \cdots. \tag{3.32}$$

Let X and Y be two species of singletons. The following equations are valid

$$\widehat{X+Y} = \hat{X} + \hat{Y} \tag{3.33}$$

$$\widehat{tX} = t\hat{X} \qquad \textit{for every } t \in \mathbb{K}. \tag{3.34}$$

The species E of all (finite) sets is the "analytical" exponential of $\hat{X}$,

$$E(X) = e^{\hat{X}}, \tag{3.35}$$

where e^X is the $\mathbb{Q}$-species defined by the usual formal power series

$$e^X = \sum_{k \geqslant 0} (1/k!)\, X^k \in \mathbb{Q}[[X]] \subseteq \mathbb{K}[[\mathscr{A}]]. \tag{3.36}$$

Proof. We establish (3.35) first. It is well known [C1, p. 29] that the formal power series (3.30) and (3.36) satisfy, in $\mathbb{Q}[[X]]$,

$$1 + X = e^{\ln(1+X)}. \tag{3.37}$$

Since the operation $\circ$ has been extended in a compatible way to the whole of $\mathbb{K}[[\mathscr{A}]]$, we may, in particular, substitute E^* for X in (3.37). This gives

$$E = 1 + E^* = e^{\ln(1+E^*)} = e^{\hat{X}} \tag{3.38}$$

as desired. Now taking the generating series of both sides of (3.35) gives $e^x = e^{\hat{X}(x)}$. Hence $\hat{X}(x) = x$. Similarly, taking the cycle indicator series (see [J1] or Definition 3.11 below) of (3.35) gives $\exp(x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3 + \frac{1}{4}x_4 + \cdots) = \exp(Z_{\hat{X}})$. Hence (3.32) holds. To establish (3.33) we must work in the ring of $\mathbb{K}$-species on two sorts of points X, Y (see [Y] for the definition of this ring); substituting $X+Y$ for X in (3.35) gives

$$\exp(\widehat{X+Y}) = E(X+Y) = E(X) \cdot E(Y) = \exp(\hat{X}) \cdot \exp(\hat{Y}) = \exp(\hat{X} + \hat{Y}). \tag{3.39}$$

Hence (3.33) holds. Iterating (3.33) shows that (3.34) is valid whenever $t = n \in \mathbb{N}$. Since the coefficients of the molecular decompositions of both sides of (3.34) are polynomials in t, taking the same values for $t \in \mathbb{N}$, it follows that (3.34) holds for every $t \in \mathbb{K}$. ∎

Remarks. There is a current intentional "abuse of notation" in the standard literature on the theory of species; namely, $E(X)=e^X$. In fact, $E(X)$ is a ℕ-species while e^X, defined by (3.36), is a ℚ-species which is *not* a ℕ-species. So that, strictly speaking,

$$E(X) \neq e^X. \tag{3.40}$$

The correct equation relating the species of sets and the (analytical) exponential series is, of course, given by (3.35). It is interesting to note that substituting tX for X in (3.35) and using (3.34) yields

$$E(tX) = \exp(\widehat{tX}) = \exp(t\hat{X}) = (\exp(\hat{X}))^t = (E(X))^t. \tag{3.41}$$

The equations

$$E(tX) = (E(X))^t \qquad \text{and} \qquad G(tX) = G(X) \times E(tX) \tag{3.42}$$

are, in fact, the starting points of Yeh's extension of substitution [Y].

With the aid of the ℚ-species of pseudo-singletons, the directional derivatives $D_H G$ may also be computed as follows.

THEOREM 3.8. *Let $G, H \in \bar{\mathbb{K}}[[\mathcal{A}]]$ with $H_0 = 0$, then*

$$D_H G = G(X+Y) \times (E(X) \cdot \hat{Y})|_{Y:=H} \tag{3.43}$$

$$\begin{aligned} &= (G(X+H) - G(X)) - \tfrac{1}{2}(G(X+2H) - 2G(X+H) + G(X)) \\ &\quad + \tfrac{1}{3}(G(X+3H) - 3G(X+2H) + 3G(X+H) - G(X)) + \cdots \\ &\quad + ((-1)^{k-1}/k)\left(\sum_{0 \leqslant \nu \leqslant k} (-1)^\nu \binom{k}{\nu} G(X+(k-\nu)H)\right) + \cdots. \end{aligned} \tag{3.44}$$

Proof. Taking (3.41) and (3.42) into account yields, successively,

$$\begin{aligned} G(X+tH) &= G(X+tY)|_{Y:=H} \\ &= G(X+Y) \times (E(X) \cdot E(tY))|_{Y:=H} \\ &= G(X+Y) \times (E(X) \cdot e^{t\hat{Y}})|_{Y:=H} \\ &= G(X+Y) \times (E(X) \cdot (1 + t\hat{Y} + O(t^2)))|_{Y:=H} \\ &= G(X) + tG(X+Y) \times (E(X) \cdot \hat{Y})|_{Y:=H} + O(t^2). \end{aligned} \tag{3.45}$$

Hence (3.43) holds. Formula (3.44) follows immediately by expanding $\hat{Y}$ as in (3.28) and using, for $k = 0, 1, 2, \ldots$, the identities

$$(E^*(Y))^k = (E(Y)-1)^k = \sum_{0 \leqslant \nu \leqslant k} (-1)^\nu \binom{k}{\nu} E((k-\nu)Y). \quad \blacksquare \tag{3.46}$$

Remarks. An alternate way to obtain (3.44) is to compute the function $t \mapsto G(X+tH)$ using the classical (forward) Newton's expansion from the theory of finite-differences. The "more refined" molecular decomposition (3.29), when applied to $\hat{Y}$, may also be used to compute the molecular decomposition of $D_H G$ via (3.43), once the molecular decomposition of $G(X+Y)$ is known. Note that (3.44) is always summable. In particular, when G and H are molecular, the corresponding sum is *finite* and of degree equal to the product $(\deg G)\cdot(\deg H)$.

To illustrate Theorem 3.8, one may check that the directional derivative $D_C E$ of the species E of all sets in the direction of the species C of all circular permutations is the following $\mathbb{Q}$-species

$$D_C E = E\cdot(S^* - \tfrac{1}{2}S^{*2} + \tfrac{1}{3}S^{*3} - \cdots), \tag{3.47}$$

where $S^* = E^*(C)$ is the species of all nonempty permutations. More generally,

$$D_H E = E\cdot\hat{H}, \qquad \text{where} \quad \hat{H} = \hat{X}\circ H. \tag{3.48}$$

It is well known in classical analysis that the notion of infinitesimal generator provides a better understanding of the notion of substitution of power series. For example, it can be shown (see [L1] for a proof), that for any two normalized formal *power series* $F, G \in \mathbb{K}[[X]]$, with infinitesimal generators $\Phi = \operatorname{gen} F$, $\Gamma = \operatorname{gen} G$:

$$G\circ F = F\circ G \qquad \text{iff} \quad \Gamma\Phi' = \Phi\Gamma' \tag{3.49}$$

and

$$G\circ F = F\circ G \qquad \text{iff} \quad \exists s, t\in\mathbb{K}: G^{\langle s\rangle} = F^{\langle t\rangle},\ s\neq 0 \text{ or } t\neq 0. \tag{3.50}$$

As a simple illustration, take for example, $F = X/(1-2X)$, $G = X/(1-3X)$. Then $\operatorname{gen} F = 2X^2$, $\operatorname{gen} G = 3X^2$, and $s=2$, $t=3$ ($\mathbb{K}=\mathbb{N}$, here).

Such kinds of results can be lifted (at least partially) to combinatorial situations:

DEFINITION 3.9. Let Φ and Γ be $\mathbb{K}$-species with $\Phi_0 = \Gamma_0 = 0$. The *Lie bracket* of D_Φ and D_Γ is the operator

$$[D_\Phi, D_\Gamma] = D_\Phi D_\Gamma - D_\Gamma D_\Phi. \tag{3.51}$$

The *Lie bracket* of Φ and Γ is the $\bar{\mathbb{K}}$-species

$$[\Phi, \Gamma] = D_\Phi \Gamma - D_\Gamma \Phi. \tag{3.52}$$

THEOREM 3.10. *The Lie brackets are related by the equation*

$$[D_\Phi, D_\Gamma] = D_{[\Phi, \Gamma]}. \tag{3.53}$$

Moreover, for any normalized $F, G \in \mathbb{K}[[\mathcal{X}]]$, with infinitesimal generators $\Phi, \Gamma \in \mathbb{K}[[\mathcal{X}]]$, we have

$$G \circ F = F \circ G \qquad \textit{iff} \quad [\Phi, \Gamma] = 0. \tag{3.54}$$

Proof. We follow the general guidelines taken from classical analysis. A straightforward computation shows that

$$(X + s\Gamma) \circ (X + t\Phi) = (X + st[\Phi, \Gamma] + \cdots) \circ (X + t\Phi) \circ (X + s\Gamma). \tag{3.55}$$

This gives, for any $W \in \mathbb{K}[[\mathcal{X}]]$,

$$\begin{aligned} &W \circ [(X + s\Gamma) \circ (X + t\Phi)] - W \circ [(X + t\Phi) \circ (X + s\Gamma)] \\ &\quad = [st\, D_{[\Phi, \Gamma]} W + \cdots] \circ (X + t\Phi) \circ (X + s\Gamma) \\ &\quad = st\, D_{[\Phi, \Gamma]} W + \text{higher terms in } s, t. \end{aligned} \tag{3.56}$$

Similarly,

$$\begin{aligned} &[W \circ (X + s\Gamma)] \circ (X + t\Phi) - [W \circ (X + t\Phi)] \circ (X + s\Gamma) \\ &\quad = st(D_\Phi D_\Gamma - D_\Gamma D_\Phi) W + \text{higher terms in } s, t. \end{aligned} \tag{3.57}$$

Hence (3.53) follows by equating (3.56) and (3.57).

We now prove (3.54). It is easily seen by induction that $G \circ F = F \circ G$ if and only if

$$\forall m, n \in \mathbb{N}, \forall W \in \mathbb{K}[[\mathcal{X}]]: W \circ G^{\langle m \rangle} \circ F^{\langle n \rangle} = W \circ F^{\langle n \rangle} \circ G^{\langle m \rangle}. \tag{3.58}$$

Since the coefficients of the molecular decomposition of $W \circ G^{\langle m \rangle} \circ F^{\langle n \rangle}$ and of $W \circ F^{\langle n \rangle} \circ G^{\langle m \rangle}$ are polynomials in m, n we obtain that $G \circ F = F \circ G$ if and only if

$$\forall s, t \in \mathbb{K}, \forall W \in \mathbb{K}[[\mathcal{X}]]: W \circ G^{\langle s \rangle} \circ F^{\langle t \rangle} = W \circ F^{\langle t \rangle} \circ G^{\langle s \rangle}. \tag{3.59}$$

Using Theorem 3.5 we see that (3.59) is equivalent to

$$\forall s, t \in \mathbb{K}, \forall W \in \mathbb{K}[[\mathcal{X}]]: (\exp tD_\Phi)(\exp sD_\Gamma) W = (\exp sD_\Gamma)(\exp tD_\Phi) W \tag{3.60}$$

which is equivalent to $D_\Phi D_\Gamma = D_\Gamma D_\Phi$ by equating the coefficients of st. ▮

Remarks. Of course, (3.49) is a particular case of (3.54). This can be seen by taking the *power series* $F, G \in \mathbb{K}[[X]] \subseteq \mathbb{K}[[\mathcal{X}]]$ and noting that, in this case, $[\Phi, \Gamma] = \Phi\Gamma' - \Gamma\Phi'$. Moreover, (3.50) follows at once from (3.49), since

$$\Phi\Gamma' = \Gamma\Phi' \quad \text{iff} \quad \exists s, t \in \mathbb{K}: s\Gamma = t\Phi, s \neq 0 \text{ or } t \neq 0. \tag{3.61}$$

However, the situation is, once again, strictly richer at the combinatorial level: (3.50) is no longer valid in the general context of $\mathbb{K}$-species as the following example shows. Take $\mathbb{K} = \mathbb{Z}$ and

$$F = X + X^2 - 2E_2, \qquad G = X - X^2 + 2E_2. \tag{3.62}$$

After a few computations, one gets

$$G \circ F = F \circ G = X - X^4 + 2E_2^2 + 2E_2 \circ X^2 - 4E_2 \circ E_2, \tag{3.63}$$

$$F^{\langle t\rangle} = X + t(X^2 - 2E_2) + \binom{t}{2}(X^4 - 2E_2^2 - 2E_2 \circ X^2 + 4E_2 \circ E_2) + \cdots \tag{3.64}$$

$$G^{\langle s\rangle} = X + s(2E_2 - X^2) + \binom{s}{2}(X^4 - 2E_2^2 - 2E_2 \circ X^2 + 4E_2 \circ E_2) + \cdots. \tag{3.65}$$

Hence, for this choice of F and G,

$$G \circ F = F \circ G \quad \text{but} \quad G^{\langle s\rangle} = F^{\langle t\rangle} \quad \text{iff} \quad s = t = 0. \tag{3.66}$$

We shall show below that (3.50) is not valid even at the level of cycle indicator series in the sense of [J1] (see also [B, D, L5, NR]). These are special series that lie between the "combinatorial level" (i.e., molecular and atomic series) and the "analytical level" (i.e., generating power series), in the terminology of [L2]. They constitute a good context for the generalization, to species, of the classical Pólya counting theory [C2, D, HP, P]. They are defined as follows in the context of $\mathbb{K}$-species.

Definition 3.11. Let $x_1, x_2, \ldots$ be a countable sequence of indeterminates and let $F = \sum_{M \in \mathcal{M}} f_M M \in \mathbb{K}[[\mathcal{A}]]$. The *cycle indicator series* of F is the series $Z_F = \sum_{M \in \mathcal{M}} f_M Z_M$ with

$$Z_M = Z_M(x_1, x_2, \ldots) = \sum \operatorname{fix} M[\sigma_1, \sigma_2, \ldots]\, x_1^{\sigma_1} x_2^{\sigma_2} \cdots / 1^{\sigma_1}\sigma_1!\, 2^{\sigma_2}\sigma_2! \cdots, \tag{3.67}$$

where the last sum is extended over all (finite) sequences $(\sigma_1, \sigma_2, \ldots)$ of integers such that $\sigma_1 + 2\sigma_2 + \cdots + \nu\sigma_\nu + \cdots = \deg M$ and $\operatorname{fix} M[\sigma_1, \sigma_2, \ldots]$ denotes the number of M-structures, on a fixed set U, which are invariant under the action of any given permutation σ of U of type $\sigma_1, \sigma_2, \ldots, \sigma_\nu, \ldots$ (here σ_ν denotes the number of cycles of σ of length ν).

The set of all (formal) indicator series constitutes a differential ring $\mathbb{K}\{\{x_1, x_2, ...\}\}$ which is further equipped with the so-called *plethystic substitution*,

$$g \circ f = g(f_1, f_2, ..., f_k, ...), \qquad f_k = f(x_k, x_{2k}, ..., x_{\nu k}, ...), \tag{3.68}$$

if $f, g \in \mathbb{K}\{\{x_1, x_2, ...\}\}$ with $f(0, 0, ...) = 0$. Every identity between $\mathbb{K}$-species in this paper leads to a corresponding identity between cycle indicator series, simply by applying the operator Z on both sides of the identity.

In particular, the generalized iterates of any normalized $f = x_1 + \cdots \in \mathbb{K}\{\{x_1, x_2, ...\}\}$ are given by

$$f^{\langle t\rangle} = f^{\langle t\rangle}(x_1, x_2, ...) = \sum_{k \geqslant 0} \binom{t}{k} \Delta_f^k x_1, \tag{3.69}$$

where $\Delta_f g = g \circ f - g$ for every $g \in \mathbb{K}\{\{x_1, x_2, ...\}\}$. The infinitesimal generator φ of f is

$$\varphi = \operatorname{gen} f = \Delta_f x_1 - \tfrac{1}{2}\Delta_f^2 x_1 + \tfrac{1}{3}\Delta_f^3 x_1 - \cdots \in \bar{\mathbb{K}}\{\{x_1, x_2, ...\}\} \tag{3.70}$$

and the directional derivative $D_\varphi\, g$ of g is

$$D_\varphi\, g = \sum_{k \geqslant 0} \varphi_k\, \partial g/\partial x_k, \qquad \varphi_k = \varphi(x_k, x_{2k}, ...). \tag{3.71}$$

For normalized $f, g \in \mathbb{K}\{\{x_1, x_2, ...\}\}$, with $\varphi = \operatorname{gen} f$ and $\gamma = \operatorname{gen} g$:

$$g \circ f = f \circ g \qquad \text{iff} \quad D_\varphi \gamma - D_\gamma \varphi = [\varphi, \gamma] = 0 \tag{3.72}$$

but

$$g \circ f = f \circ g \not\Rightarrow \exists s, t \in \mathbb{K}\colon g^{\langle s\rangle} = f^{\langle t\rangle}, \qquad s \neq 0 \text{ or } t \neq 0 \tag{3.73}$$

as the following example shows: Take F, G as in (3.62), then

$$f = Z_F = x_1 - x_2, \qquad g = Z_G = x_1 + x_2, \qquad g \circ f = f \circ g = x_1 - x_4, \tag{3.74}$$

$$f^{\langle t\rangle} = x_1 - \binom{t}{1} x_2 + \binom{t}{2} x_4 - \binom{t}{3} x_8 + \cdots, \tag{3.75}$$

$$g^{\langle s\rangle} = x_1 + \binom{s}{1} x_2 + \binom{s}{2} x_4 + \binom{s}{3} x_8 + \cdots. \tag{3.76}$$

Hence,

$$g \circ f = f \circ g \qquad \text{but} \qquad g^{\langle s\rangle} = f^{\langle t\rangle} \qquad \text{iff} \quad s = t = 0. \tag{3.78}$$

These counterexamples to (3.50) lead to the problem of investigating the additional conditions to be put on normalized $F, G \in \mathbb{K}[[\mathcal{A}]]$ in order that (3.50) hold. There is an analogous problem corresponding to the cycle indicator series. We leave these open.

Acknowledgments

I would like to thank Jacques Labelle and Pierre Leroux for their helpful comments, Hélène Décoste who MACDREW the figures, and Yeong Nan Yeh who carefully checked all the formulas, examples, and counterexamples.

References

[B] F. Bergeron, Une combinatoire du pléthysme, *J. Combin. Theory Ser. A* **46**, No. 2 (1987), 291–305.

[C1] H. Cartan, "Théorie Élémentaire des Fonctions Analytiques d'une ou Plusieurs Variables Complexes, Hermann, Paris, 1961.

[C2] L. Comtet, "Analyse Combinatoire," (2 vols.), Collection SUP, Presses Univ. de France, 1970.

[D] H. Décoste, "Séries indicatrices d'espèces pondérées et q-analogues," Ph.D. thesis, Université de Montréal–UQAM, in preparation.

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

[J1] A. Joyal, Une théorie combinatoire des séries formelles, *Adv. in Math.* **42** (1981), 1–82.

[J2] A. Joyal, Règle des signes en algèbre combinatoire, *C. R. Math. Rep. Acad. Sci. Canada* **7**, No. 5 (1985), 285–290.

[J3] A. Joyal, Foncteurs analytiques et espèces de structures, *in* "Combinatoire énumérative, Proceedings, Montréal, Québec 1985," (G. Labelle and P. Leroux, Ed.), Lecture Notes in Math. Vol. 1234, pp. 126–159, Springer-Verlag, New York/Berlin, 1986.

[L1] G. Labelle, Sur l'inversion et l'itération continue des séries formelles, *European J. Combin.* **1** (1980), 113–138.

[L2] G. Labelle, Some New Computational Methods in the Theory of Species, *in* "Combinatoire énumérative, Proceedings, Montréal, Québec 1985," (G. Labelle and P. Leroux, Ed.), Lecture Notes in Math. Vol. 1234, pp. 192–209, Springer-Verlag, New York/Berlin, 1986.

[L3] G. Labelle, On Combinatorial Differential Equations, *J. Math. Anal. Appl.* **113**, No. 2 (1986), 344–381.

[L4] J. Labelle, Applications diverses de la théorie combinatoire des espèces de structures, *Ann. Sci. Math. Québec* **7**, No. 1 (1983), 59–94.

[L5] J. Labelle, Quelques espèces sur les ensembles de petite cardinalité, *Ann. Sci. Math. Québec* **9** No. 1 (1985), 31–58.

[NR] O. Nava and G. C. Rota, Plethysm, categories and combinatorics, *Adv. in Math.* **58** (1985), 61–88.

[P] G. Pólya, Kombinatorische Anzahbestimmungen fur Gruppen, Graphen und chemische Verbindungen, *Acta Math.* **68** (1937), 145–254.

[R] G. C. Rota, On the foundations of combinatorial theory, I: Theory of Möbius functions, *Z. Wahrsch. Verw. Gebiete* **2** (1964), 340–368.

[Y] Y. N. Yeh, The calculus of virtual species and $\mathbb{K}$-species, in "Combinatoire énumérative, Proceedings, Montréal, Québec 1985," (G. Labelle and P. Leroux, Ed.), Lecture Notes in Math. Vol. 1234, pp. 351–369, Springer-Verlag, New York/Berlin, 1986.