FPSAC 2013 Paris, France DMTCS proc. **AS**, 2013, 361–372

The explicit molecular expansion of the combinatorial logarithm

Gilbert Labelle

LaCIM et Dép. de Mathématiques, UQAM, CP 8888, Succ. Centre-Ville, Montréal, Québec, H3C 3P8 Canada.

Abstract. Just as the power series of $\log(1+X)$ is the analytical substitutional inverse of the series of $\exp(X)-1$, the (virtual) combinatorial species, $\mathrm{Lg}(1+X)$, is the combinatorial substitutional inverse of the combinatorial species, $E(X)-1$, of non-empty finite sets. This *combinatorial logarithm*, $\mathrm{Lg}(1+X)$, has been introduced by A. Joyal in 1986 by making use of an iterative scheme. Given a species $F(X)$ (with $F(0)=1$), one of its main applications is to express the species, $F^{\mathrm{c}}(X)$, of *connected* F-structures through the formula $F^{\mathrm{c}} = \mathrm{Lg}(F) = \mathrm{Lg}(1+F_{+})$ where F_{+} denotes the species of non-empty F-structures. Since its creation, equivalent descriptions of the combinatorial logarithm have been given by other combinatorialists (G. L., I. Gessel, J. Li), but its exact decomposition into irreducible components (molecular expansion) remained unclear. The main goal of the present work is to fill this gap by computing explicitly the molecular expansion of the combinatorial logarithm and of $-\mathrm{Lg}(1-X)$, a “ cousin ” of the tensorial species, $\mathrm{Lie}(X)$, of free Lie algebras.

Résumé. Tout comme la série de puissances de $\log(1+X)$ est l'inverse substitutionnel analytique de la série de $\exp(X)-1$, l'espèce de structures (virtuelle) $\mathrm{Lg}(1+X)$, est l'inverse substitutionnel combinatoire de l'espèce, $E(X)-1$, des ensembles finis non vides. Ce *logarithme combinatoire*, $\mathrm{Lg}(1+X)$, a été introduit par A. Joyal en 1986 en faisant appel à un schéma itératif. Étant donnée une espèce $F(X)$ (telle que $F(0)=1$), l'une de ses principales applications est d'exprimer l'espèce, $F^{\mathrm{c}}(X)$, des F-structures *connexes* par la formule $F^{\mathrm{c}} = \mathrm{Lg}(F) = \mathrm{Lg}(1+F_{+})$ où F_{+} désigne l'espèce des F-structures non vides. Depuis sa création, des descriptions équivalentes du logarithme combinatoire ont été formulées par d'autres combinatoriciens (G. L., I, Gessel, J. Li), mais sa décomposition exacte en composantes irréductibles (développement moléculaire) est demeurée obscure. Le but principal du présent travail est de combler cette lacune en calculant explicitement le développement moléculaire du logarithme combinatoire et de $-\mathrm{Lg}(1-X)$, un “ cousin ” de l'espèce tensorielle, $\mathrm{Lie}(X)$, des algèbres de Lie libres.

Keywords: combinatorial species, combinatorial logarithm, molecular expansion, generating functions

1 Introduction

1.1 Counting connected structures

Since simple graphs are assemblies of connected simple graphs, it is well known that the exponential generating series, $G(x)$, which counts simple graphs, satisfies $G(x) = \exp(G^{\mathrm{c}}(x))$, where $G^{\mathrm{c}}(x)$ is the exponential generating series of connected simple graphs. Now, taking the (analytical) logarithm of both sides of this equation gives,

$$G^{\mathrm{c}}(x) = \log(G(x)) = \log \sum_{n\geq 0} 2^{n(n-1)/2} x^n/n!, \tag{1.1}$$

from which connected simple graphs can be counted exactly, recursively or asymptotically. More generally, the analogous formula,

$$F^{\mathrm{c}}(x) = \log(F(x)), \tag{1.2}$$

holds for the exponential generating series of any species of structures, F and F^{c}, for which F-structures are assemblies of F^{c}-structures. That is, for which the combinatorial equation,

$$F(X) = E \circ F^{\mathrm{c}}(X), \tag{1.3}$$

holds, where $\circ$ denotes the substitution of species, E is the species of finite sets (E stands for *ensembles*, in French) and X is the species of singletons (that is, one-element sets). Taking the cycle index series of both members of (1.3) yields, $Z_F = Z_E \circ Z_{F^{\mathrm{c}}}$, where $\circ$ now denotes the classical plethystic substitution. Since $Z_E = \exp \sum \frac{1}{k} p_k$, this can be written explicitly as(i)

$$Z_F(p_1, p_2, p_3, \ldots) = \exp \sum_{k\geq 1} \frac{1}{k} Z_{F^{\mathrm{c}}}(p_k, p_{2k}, p_{3k}, \cdots). \tag{1.4}$$

Taking the logarithm of both sides of (1.4) and using Möbius inversion gives (see [BLL98]) the following refinement of (1.2),

$$Z_{F^{\mathrm{c}}}(p_1, p_2, p_3, \ldots) = \sum_{k\geq 1} \frac{\mu(k)}{k} \log Z_F(p_k, p_{2k}, p_{3k}, \cdots), \tag{1.5}$$

where μ denotes the classical Möbius function. As a consequence, the (ordinary) generating series, $\widetilde{F^{\mathrm{c}}}(x)$, which counts *unlabeled* F^{c}-structures is obtained via the substitutions, $p_i := x^i, i = 1, 2, \ldots$, in (1.5). All of this is classical in Pólya theory in the context of combinatorial species. For an introduction to species, the reader can consult the basic paper of A. Joyal [Joy81] or the book [BLL98] by Bergeron, Labelle, and Leroux.

1.2 Solving the combinatorial equation $F = E \circ F^{\mathrm{c}}$ for the species F^{c}

In 1986, Joyal [Joy86] went a step further by solving the combinatorial equation (1.3) for the species F^{c} in terms of the species F, thereby refining simultaneously both (1.2) and (1.5). He proceeded along the following lines. Let 1 denote the species of *the* empty set. We have $E = 1 + E_+$, where E_+ is the species of non empty finite sets and $F = 1 + F_+$, where F_+ is the species of F-structures on non empty finite sets(ii). Combinatorial equation (1.3) is then equivalent to

$$F_+(X) = E_+ \circ F^{\mathrm{c}}(X). \tag{1.6}$$

(i) The variables p_k stand for the power sums and are often denoted by x_k in the theory of combinatorial species.
(ii) Since an empty assembly of F^{c}-structures *is* the empty set.

Now, $E_+ = X + E_{\geq 2}$, where $E_{\geq 2}$ is the species of finite sets having at least 2 elements. By his implicit species theorem (see [Joy81]), Joyal concluded that E_+ has a substitutional inverse, $E_+^{<-1>}$, in the realm of *virtual species* (that is, formal differences of species). By adapting the Newton interpolation formula to species (using a special " difference operator " δ), he also gave the following combinatorial formula for this substitutional inverse,

$$E_+^{<-1>}(X) = \sum_{n\geq 0}(-1)^n W_n(X) = \sum_{n:\,\mathrm{even}} W_n(X) - \sum_{n:\,\mathrm{odd}} W_n(X), \tag{1.7}$$

where $W_n = W_n(X)$ are species defined by the recursive scheme,

$$W_0 = X, \qquad W_n = \delta W_{n-1} = W_{n-1}\circ E_+ - W_{n-1}, \quad n \geq 1. \tag{1.8}$$

Hence, a W_n-structure, on a finite set U, is a strictly increasing sequence, $\hat{0} = R_0 < R_1 < \cdots < R_n = \hat{1}$, in the lattice of equivalence relations on U, where $\hat{0}$ and $\hat{1}$ respectively denote the finest and the coarsest equivalence relation on U. Applying $E_+^{<-1>}$ to (1.6) and using (1.7) finally gives,

$$F^{\mathrm{c}}(X) = E_+^{<-1>}\circ F_+(X) = \sum_{n\geq 0}(-1)^n W_n(F_+(X)). \tag{1.9}$$

In the present paper, we use the notation, $\mathrm{Lg}(1+X)$, to denote $E_+^{<-1>}(X)$ and call it the *combinatorial logarithm*[(iii)], by analogy with the fact that, in analysis, the power series of $\log(1+X)$ is the substitutional inverse of that of $\exp(X)-1$. Summarizing, we have,

$$\mathrm{Lg}(1+X) \underset{\mathrm{def}}{=} E_+^{<-1>}(X) \quad \text{and} \quad F^{\mathrm{c}} = \mathrm{Lg}(1+F_+) = \mathrm{Lg}(F). \tag{1.10}$$

Note that (1.10) associates a virtual species F^{c} to *any* species F for which $F(0) = 1$, even in the case where F does not possess connected structures. For this reason, $\mathrm{Lg}(1+X) = (1+X)^{\mathrm{c}}$ is sometimes called, by abuse of language, the virtual species of " connected " $(1+X)$-structures. Although very useful and conceptually elegant, it turns out that the species $\sum_{n:\,\mathrm{even}} W_n$ and $\sum_{n:\,\mathrm{odd}} W_n$ in Joyal's expression for the combinatorial logarithm have plenty of subspecies in common. That is,

$$\mathrm{Lg}(1+X) = \sum_{n:\,\mathrm{even}} W_n - \sum_{n:\,\mathrm{odd}} W_n, \tag{1.11}$$

is *not* a completely reduced expression as a difference of species. Other equivalent $-$ but still *not* completely reduced $-$ expressions for the combinatorial logarithm have been given using special classes of graphs. For example, Gessel and Li in [GL11], found formula (1.12a), where $\mathcal{Q}^{\mathrm{c}}$ is the species of *connected co-point-determining graphs* and $\mathcal{P}^{\mathrm{c}}_{\geq 2}$ is that of *connected point-determining graphs having at least two vertices*. Later, J. Li [Li12] found the further reduced formula (1.12b), where $\mathcal{T}^{\mathrm{c}}$ is the species of *connected co-point-determining cographs* and $\mathcal{S}^{\mathrm{c}}_{\geq 2}$ is that of *connected point-determining cographs having at least two vertices*.

$$a)\quad \mathrm{Lg}(1+X) = \mathcal{Q}^{\mathrm{c}} - \mathcal{P}^{\mathrm{c}}_{\geq 2}, \qquad b)\quad \mathrm{Lg}(1+X) = \mathcal{T}^{\mathrm{c}} - \mathcal{S}^{\mathrm{c}}_{\geq 2}. \tag{1.12}$$

Our main goal is to give a completely reduced expression for the combinatorial logarithm. In Section 2, we describe the irreducible components of $F^{\mathrm{c}} = \mathrm{Lg}(F)$ and, in particular, of $\mathrm{Lg}(1+X)$ and $-\mathrm{Lg}(1-X)$, together with their exact multiplicities. Section 3 contains a compact table for the combinatorial logarithm up to degree 10.

(iii) Joyal uses the notation $\log(1+X)$ for the combinatorial logarithm, but we use it for the *analytical* logarithm in the present text.

2 Explicit molecular expansions

2.1 Molecular expansions in general

We first recall the general notions of molecular and atomic species. A species M is *molecular* if $M \neq 0$ and any two M-structures are isomorphic. Equivalently, M is irreducible under the combinatorial sum. A molecular species A is *atomic* if $A \neq 1$ and is irreducible over the combinatorial product. Y. N. Yeh proved in [Yeh86] that every molecular species can be written in a unique way (up to isomorphism) as a commutative finite product of atomic species. The sets $\mathbf{M}$ of all molecular species and $\mathbf{A}$ of all atomic species (up to isomorphism) are countable and we have, up to degree three,

$$\mathbf{M} = \{1, X, E_2, X^2, E_3, C_3, XE_2, X^3, \ldots\}, \quad \mathbf{A} = \{X, E_2, E_3, C_3, \ldots\}, \tag{2.1}$$

where X^n is the species of linear orderings of length n, C_n is the species of oriented n-cycles, and E_n is the species of n-sets. Note that $\mathbf{M}$ is the free commutative monoid (under combinatorial product) generated by $\mathbf{A}$. Moreover, each molecular species, M, is completely determined by the stabilizer $H = \mathrm{Stab}(s) \leq S_n$ of anyone of its structures, say s on $[n]$, where n is the degree of M and $[n] = \{1, 2, \ldots, n\}$. We write $M = X^n/H =$ linear orderings of length n modulo H. In particular, we have $X^n = X^n/\{1\}, \quad E_n = X^n/S_n, \quad C_n = X^n/\langle(1\,2\,\ldots\,n)\rangle$. Two molecular species, X^n/H and X^m/K, are isomorphic (and we write, $X^n/H = X^m/K$) if and only if $n = m$ and H and K are conjugate in S_n. Let now F be any species, not necessarily molecular. Then, one can always write F as a (countable) linear combination with nonnegative integer coefficients of molecular species,

$$F = \sum_{M \in \mathbf{M}} f_M M \in \mathbb{N}[[\mathbf{A}]], \tag{2.2}$$

where f_M denotes the number of subspecies of F that are isomorphic to M. The coefficient f_M is called the *multiplicity* of M in F. Summation (2.2) is unique and is called the *molecular expansion* of F. This expansion is very strong since it is a common refinement of the classical generating series, $F(x), \widetilde{F}(x), Z_F(p_1, p_2, p_3, \ldots)$, associated to the species F. For an example of molecular expansion, consider the well-known species $T = T(X)$ of rooted trees, defined by the combinatorial functional equation $T = XE(T)$. Up to degree 6, we have (see the book [BLL98] by Bergeron, Labelle, and Leroux, for example),

$$\begin{aligned} T = {} & X + X^2 + XE_2 + X^3 + XE_3 + 2X^4 + X^2E_2 + XE_4 + 3X^3E_2 + X\cdot(E_2 \circ X^2) + 3X^5 \\ & + X^2E_3 + X^2E_4 + 6X^4E_2 + 2X^2\cdot(E_2 \circ X^2) + 3X^3E_3 + X^2{E_2}^2 + XE_5 + 6X^6 + \cdots . \end{aligned} \tag{2.3}$$

Note that the species, $E_2 \circ X^2$, which occurs in this expansion is atomic. The usual combinatorial operations (sum, product, composition, differentiation, etc), as well as molecular expansions, have been extended by Joyal and Yeh ([Joy85], [Yeh86]) to *virtual species*, that is formal differences

$$\Phi = F - G, \tag{2.4}$$

of (ordinary) species F and G. The molecular expansion of Φ is defined by,

$$\Phi = \sum_{M \in \mathbf{M}} \phi_M M = \sum_{M \in \mathbf{M}} (f_M - g_M) M \in \mathbb{Z}[[\mathbf{A}]], \tag{2.5}$$

where $\sum_{M\in\mathbf{M}} f_M M$ and $\sum_{M\in\mathbf{M}} g_M M$ are the molecular expansions of F and G, respectively[(iv)]. Every virtual species, Φ, can be represented in the form (2.4) in an infinite number of way, just as $-5 = 0-5 = 1-6 = 2-7 = \cdots$, in the context of the ring, $\mathbb{Z}$, of integers. The less F and G have subspecies in common, the more representation (2.5) is said to be *reduced*. It is always possible to canonically choose F and G in such a way that they have *no* subspecies in common (i.e., are " stranger " species, to use a terminology taken from the theory of signed measures). The corresponding representation is denoted,

$$\Phi = \Phi^+ - \Phi^-, \tag{2.6}$$

and is called the (*completely*) *reduced form* of the virtual species Φ. The species Φ^+ (resp., Φ^-) is called the *positive* (resp., *negative*) part of Φ and the coefficients ϕ_M in (2.5) satisfy $\phi_M > 0$ if M appears in Φ^+ and $\phi_M < 0$ if M appears in Φ^- (otherwise, $\phi_M = 0$). Note that Φ^+ and Φ^- are characterized by the fact that no molecular species appears in both of their molecular expansions.

2.2 The explicit expansions of $\mathrm{Lg}(1+X)$ and of $-\mathrm{Lg}(1-X)$

Let us start by taking a closer look at the Joyal species, W_n, defined by (1.8). It is easy to see that,

$$W_n = \sum_{0\le k\le n} (-1)^k \binom{n}{k} E_+^{<n-k>}, \tag{2.7}$$

where $E_+^{<i>}$ denotes the i-fold iterate of E_+ under $\circ$ (with $E_+^{<0>} = X$). Using the expansion formulas,

$$E_+ = E - 1 = X + E_2 + E_3 + \cdots, \quad E(mA + nB + \cdots) = E(A)^m E(B)^n \cdots, \tag{2.8}$$

massive simplification and cancellation of terms occur in (2.7) and (1.11), and we have, up to degree 6, the molecular expansions,

$$\begin{aligned}\mathrm{Lg}(1+X)^+ = {} & X + XE_2 + XE_3 + E_2\circ E_2 + X^3E_2 + XE_4 + E_2E_3 + X^3E_3 \\ & + 2\,X^2E_2^2 + XE_5 + E_2E_4 + E_3\circ E_2 + E_2\circ E_3 + \cdots,\end{aligned} \tag{2.9}$$

$$\begin{aligned}\mathrm{Lg}(1+X)^- = {} & E_2 + E_3 + X^2E_2 + E_4 + X^2E_3 + XE_2^2 + E_5 + X^4E_2 + X^2E_4 \\ & + 2\,XE_2E_3 + E_2\cdot(E_2\circ E_2) + E_6 + E_2\circ(XE_2) + \cdots,\end{aligned} \tag{2.10}$$

Of course, the molecular species that appear in $\mathrm{Lg}(1+X)^+$ and $\mathrm{Lg}(1+X)^-$ are all *set-like* (that is, they are build from the E_n's using only products and substitutions), but their exact nature and multiplicities are far from being obvious. In fact, these " surviving " molecular species are *strictly* included in the class of set-like molecular species, since, for example, the set-like molecular species, $X\cdot(E_2\circ E_2)$, neither appears in (2.9) nor in (2.10). We will exhibit their exact form and describe their multiplicities explicitly, using special kinds of integer partitions together with an arithmetical function related to the Möbius function. To do so, we will need to work in the more general setting, $\mathbb{Q}[[\mathbf{A}]]$, of *rational species*, that is, of countably summable linear combinations of molecular species with rational coefficients[(v)]. All usual combinatorial

---

(iv) As a ring, $\mathbb{Z}[[\mathbf{A}]]$ is the completion (under countable summability) of the family, $\mathbf{B}(S_n)_{n\ge 0}$, of the Burnside rings of virtual set-like representations of the symmetric groups, $S_n, n \ge 0$.

(v) More general settings are also possible, for example, $\mathbb{C}[[\mathbf{A}]]$, but $\mathbb{Q}[[\mathbf{A}]]$ is sufficient here.

operations have been extended to this ring by A. Joyal in [Joy85] and Y.-N. Yeh in [Yeh86]. Since the classical ring, $\mathbb{Q}[[X]]$, of power series in X " sits " in $\mathbb{Q}[[\mathbf{A}]]$, both kinds (analytical and combinatorial) of exponentials and logarithms are special cases of rational species, with molecular expansions:

$$\begin{aligned} \exp(X) &= \sum_{n\geq 0} \frac{1}{n!} X^n \in \mathbb{Q}[[\mathbf{A}]], \quad \log(1+X) = \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} X^n \in \mathbb{Q}[[\mathbf{A}]]. \\ E(X) &= \sum_{n\geq 0} E_n(X) \in \mathbb{N}[[\mathbf{A}]], \quad \mathrm{Lg}(1+X) = \sum_{M} \omega_M M(X) \in \mathbb{Z}[[\mathbf{A}]], \end{aligned} \tag{2.11}$$

where the coefficients $\omega_M \in \mathbb{Z}$ are to be determined explicitely. We need to introduce some preliminary definitions, lemmas and notation. For technical reasons, in this paper, a partition of an integer $n \geq 0$ will be a weakly *increasing*(vi) sequence $\lambda = (\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_r)$ of positive integers such that $\lambda_1 + \lambda_2 + \cdots + \lambda_r = n$. We write $\lambda \vdash n$. The number λ_i is called the i-th *part* of λ and $r = \#\lambda$, is the number of parts of λ. As usual, $m_j = m_j(\lambda), j = 1, \ldots, n$ denotes the *multiplicity* of part j in λ; that is, $m_j(\lambda) = \mathrm{card}\{i : \lambda_i = j\}$. The expression $1^{m_1}2^{m_2}3^{m_3}\cdots n^{m_n}$ is called the *type* of λ and $n = |\lambda|$ is called the *size* of λ. It turns out to be useful to freely use the abuse of notation of identifying a partition with its type; so that,

$$\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_r) = 1^{m_1}2^{m_2}\cdots n^{m_n}, \quad m_j = m_j(\lambda), \quad j = 1, \ldots, n. \tag{2.12}$$

Moreover, if $\lambda \vdash n$ and $d \geq 1$, we denote by λ^k the partition of kn defined by

$$\lambda^k = 1^{km_1}2^{km_2}\cdots n^{km_n} \vdash kn, \quad m_j = m_j(\lambda), \quad j = 1, \ldots, n. \tag{2.13}$$

So that $\lambda^k \vdash kn$ have the same parts as λ, but each individual part of λ occurs k times in λ^k. For example, $(1,1,1,4,4,4,4,4,4,5,5,5) = (1,4,4,5)^3 \vdash 42$ and $(2,4,4,7) = (2,4,4,7)^1 \vdash 17$.

Definition 2.1 *A partition $\lambda \vdash n$ is called,*
• primary*, if* $\gcd(m_1(\lambda), m_2(\lambda), \ldots, m_n(\lambda)) = 1$; • fat*, if it has a part $\lambda_i > 1$. Equivalently, $\lambda \neq 1^n$;*
• non-repeating*, if $\lambda \neq m^k$ with $m \geq 1, k > 1$.*

Lemma 2.2 *Every non-empty partition λ can be written in the form $\lambda = \tau^k$, where τ is a primary partition. Moreover, λ is fat* (*resp., non-repeating*) *if and only if τ is fat* (*resp., non-repeating*). □

Borrowing notational conventions from the theory of symmetric functions, we now associate a set-like molecular species, $E_\lambda = E_\lambda(X)$, of degree n, to every partition $\lambda \vdash n$, by the combinatorial products,

$$E_\lambda \underset{\mathrm{def}}{=} E_{\lambda_1}E_{\lambda_2}\cdots E_{\lambda_k} = X^{m_1}E_2^{m_2}\cdots E_n^{m_n}, \quad (E_1 = X). \tag{2.14}$$

Note that, $E_\lambda = X^n/S_\lambda = X^n/S_{\lambda_1,\lambda_2,\ldots,\lambda_k}$, where S_λ denotes the Young subgroup of S_n of type λ. Also, if $\alpha \vdash m, \beta \vdash n$, then $E_\alpha \circ E_\beta = X^m/S_\alpha \circ X^n/S_\beta = X^{mn}/S_\alpha \wr S_\beta$, where $\wr$ is the wreath product.

Lemma 2.3 *Every molecular species, $M \neq 1$, can be written canonically in exactly one of the two forms,*

$$M = P^k \quad \text{or} \quad M = (E_\tau \circ Q)^k, \tag{2.15}$$

where τ is a primary fat partition, P, Q are molecular and $k \geq 1$.

(vi) Instead of weakly *decreasing*, contrarily to the usual practice in the theory of partitions.

Proof: Let $M = A_1^{\alpha_1} A_2^{\alpha_2} \cdots$ be the atomic factorization of M, take $k = \gcd(\alpha_1, \alpha_2, \ldots)$ and consider the molecular species $P = A_1^{\alpha_1/k} A_2^{\alpha_2/k} \cdots$. Clearly, we canonically have $M = P^k$. Now, if P is not of the form $E_\tau \circ Q$, with $|\tau| > 1$, then we are done. On the contrary, if $P = E_\tau \circ Q$, with $|\tau| > 1$, then τ must be fat since if $\tau = 1^s, s > 1$, then $M = P^k = (E_{1^s} \circ Q)^k = (Q^s)^k = Q^{sk}$, contradicting the fact that $k = \gcd(\alpha_1, \alpha_2, \ldots)$, since $sk > k$. Moreover, τ must be primary, since if $\tau = \phi^s, s > 1$, then $M = P^k = (E_{\phi^s} \circ Q)^k = ((E_\phi \circ Q)^s)^k = (E_\phi \circ Q)^{sk}$, which is again a contradiction. $\square$

Finally, we need the following special function π defined on the set, $\mathbb{P}^+$, of non empty partitions and the " Möbius-like " arithmetical function, ν, defined on the set, $\mathbb{N}^+$, of positive integers.

Definition 2.4 *The function, $\pi : \mathbb{P}^+ \to \mathbb{Q}$, is defined, for $\lambda = 1^{m_1} 2^{m_2} \cdots n^{m_n} \vdash n$ by,*

$$\pi(\lambda) = \frac{(-1)^{\#\lambda - 1}}{\#\lambda} \binom{\#\lambda}{m_1, \ldots, m_n}. \tag{2.16}$$

In particular, for $\lambda = 1^n$, $\pi(1^n)$ reduces to the number-theoretic multiplicative function,

$$\theta : \mathbb{N}^+ \to \mathbb{Q},\; n \mapsto \frac{(-1)^{n-1}}{n}. \tag{2.17}$$

Definition 2.5 *The function, $\nu : \mathbb{N}^+ \to \mathbb{Q}$, is the inverse of θ under the Dirichlet $(\star)$ convolution*(vii)*. Explicitly,*

$$\nu(n) = \begin{cases} \frac{1}{2}\mu(i)/i, & if\ n = 2^k i,\ i\ odd,\ k \geq 1, \\ \mu(i)/i, & otherwise. \end{cases} \tag{2.18}$$

We are now ready to state and prove our main result from which each individual coefficient of the molecular expansion of the *combinatorial* logarithm, $\mathrm{Lg}(F)$, of a species, F, can be computed from the coefficients of the molecular expansion of its *analytical* logarithm, $\log(F)$. This analytical logarithm is very easy to expand, since we have, in view of (2.11),

$$\log(F) = \log(1 + F_+) = \sum_{n \geq 1} \frac{(-1)^n}{n} (F_+)^n. \tag{2.19}$$

The expansions of $\mathrm{Lg}(1 + X)$ and $-\mathrm{Lg}(1 - X)$ will then follow as special cases (Corollaries 2.7 – 2.8).

Theorem 2.6 *Consider a species, $F = 1 + F_+$, with molecular expansion $F = 1 + \sum_{M \neq 1} f_M M$, together with the molecular expansions of its two kinds of logarithms,*

$$\mathrm{Lg}(F) = \sum_{M \neq 1} g_M M, \qquad \log(F) = \sum_{M \neq 1} h_M M. \tag{2.20}$$

Then the coefficients g_M can be computed from the coefficients h_M via the recursive scheme,

$$g_M = h_M - \sum_{\substack{E_\lambda \circ N = M \\ |\lambda| > 1}} \pi(\lambda) g_N. \tag{2.21}$$

(vii) $(f \star g)(n) = \sum_{d|n} f(d) g(n/d)$.

More precisely, if M is written in the canonical form (2.15)*, then*

$$g_M = \begin{cases} \nu(k) \star h_{P^k}, & if\ M = P^k, \\ \nu(k) \star \left(h_{(E_\tau \circ Q)^k} - \pi(\tau^k) g_Q\right), & if\ M = (E_\tau \circ Q)^k, \end{cases} \tag{2.22}$$

where $(\star)$ *denotes Dirichlet convolution.*

Proof: Consider the special rational species, $\widehat{X}$, of *pseudo-singletons*, that we introduced in [Lab90] as the *analytical* logarithm of the species, E, of finite sets. Expanding, we have explicitly (see [Lab08], for more detail),

$$\begin{aligned} \widehat{X} \underset{\mathrm{def}}{=} \log(E) = \log(1 + E_+) &= \sum_{k \geq 1} \frac{(-1)^{k-1}}{k} (E_1 + E_2 + E_3 + \cdots)^k \\ &= P_1 + \frac{1}{2} P_2 + \frac{1}{3} P_3 + \cdots + \frac{1}{n} P_n + \cdots \in \mathbb{Q}[[\mathbf{A}]], \end{aligned} \tag{2.23}$$

where $P_n = P_n(X)$ are *virtual* species that are " combinatorial liftings " of the classical power sums symmetric functions[(viii)], p_n, and can be computed by the " Newton like " combinatorial recursive scheme,

$$P_1 = X, \qquad P_n = nE_n - E_1 P_{n-1} - E_2 P_{n-2} - \cdots - E_{n-1} P_1, \quad n \geq 2. \tag{2.24}$$

Taking *analytical* exponential, $\exp$, of (2.23) gives alternate expressions for the species of finite sets:

$$E = \exp(\widehat{X}) = e^{\widehat{X}} = \exp\left(\sum_{n \geq 1} \frac{1}{n} P_n \right). \tag{2.25}$$

Expanding (2.23) and using (2.16), we get the molecular expansion,

$$\frac{1}{n} P_n = \sum_{\lambda \vdash n} \pi(\lambda) E_\lambda. \tag{2.26}$$

Now, a basic property of P_n is that it behaves *linearly* (see [Lab08]) under substitution[(ix)]:

$$P_n \circ (aA + bB + \cdots) = aP_n \circ A + bP_n \circ B + \cdots, \quad a, b, \ldots \in \mathbb{Q}, \quad A, B, \ldots \in \mathbb{Q}[[\mathbf{A}]]. \tag{2.27}$$

Of course, such a linear behavior is *far* from being true in general. In particular, it is far from being true for E_λ. Nevertheless, thanks to $(2.26) - (2.27)$, we have,

$$\left(\sum_{\lambda \vdash n} \pi(\lambda) E_\lambda \right) \circ (aA + bB + \cdots) = \sum_{\lambda \vdash n} \pi(\lambda)(aE_\lambda \circ A + bE_\lambda \circ B + \cdots). \tag{2.28}$$

(viii) In fact, at the level of linear representations, the cycle index series of P_n is p_n. That is, $Z_{P_n} = p_n$.

(ix) They even behave *plethystic linearly* under substitution: for weight variables, $s, t, \ldots$, we have, $P_n(asF + btG + \cdots) = as^n P_n(F) + bt^n P_n(G) + \cdots$.

This last equation is *crucial* in the next steps of the present proof. By (2.25), we have,

$$\exp\left(\left(\sum_n \frac{1}{k}P_k\right)\circ\left(\sum_{N\neq 1} g_N N\right)\right) = E(\sum_{N\neq 1} g_N N) = E(\mathrm{Lg}(F)) = F = 1+\sum_{M\neq 1} f_M M. \quad (2.29)$$

Taking $\log$, using (2.26) and linearity property (2.28), gives,

$$\sum_{\substack{k,N\\ \lambda\vdash k}} \pi(\lambda) g_N E_\lambda \circ N = \log(F) = \sum_{M\neq 1} h_M M. \quad (2.30)$$

Extracting the coefficient of M on the leftmost and rightmost sides of (2.30), we can write,

$$\sum_{E_\lambda \circ N = M} \pi(\lambda) g_N = h_M, \quad (2.31)$$

which is equivalent to the recursive scheme (2.21), since $\pi(1) = 1$ and $E_1 \circ N = X \circ N = N$. Finally, consider the canonical form (2.15) of M. If $M = P^k$, then $E_\lambda \circ N = M$, with $|\lambda| > 1$, if and only if $\lambda = 1^d, N = P^{k/d}$, with $1 < d|k$. So that, (2.21) takes the form,

$$g_{P^k} = h_{P^k} - \sum_{1<d|k} \pi(1^d) g_{P^{k/d}} = h_{P^k} + \pi(1) g_{P^k} - \theta(k) \star g_{P^k}, \quad (2.32)$$

which reduces to $\theta(k) \star g_{P^k} = h_{P^k}$. This is equivalent to $g_{P^k} = \nu(k) \star h_{P^k}$. On the other hand, if $M = (E_\tau \circ Q)^k$, where τ is a primary fat partition, then $E_\lambda \circ N = M$, with $|\lambda| > 1$, if and only if $\lambda = 1^d, N = (E_\tau \circ Q)^{k/d}$, with $1 < d|k$, **or** $\lambda = \tau^k, N = Q$. This time, (2.21) takes the form,

$$\begin{aligned} g_{(E_\tau\circ Q)^k} &= h_{(E_\tau\circ Q)^k} - \pi(\tau^k) g_Q - \sum_{1<d|k} \pi(1^d) g_{(E_\tau\circ Q)^{k/d}} \\ &= h_{(E_\tau\circ Q)^k} - \pi(\tau^k) g_Q + g_{(E_\tau\circ Q)^k} - \theta(k) \star g_{(E_\tau\circ Q)^k}. \end{aligned} \quad (2.33)$$

This reduces to $\theta(k) \star g_{(E_\tau\circ Q)^k} = h_{(E_\tau\circ Q)^k} - \pi(\tau^k) g_Q$, which proves (2.22). □

Corollary 2.7 *The molecular expansion of the combinatorial logarithm is explicitly given by,*

$$\mathrm{Lg}(1+X) = \sum_M \omega_M M, \quad (2.34)$$

where each molecular component, M, is of the form of a finite composition,

$$M = E_{\phi^{(1)}} \circ E_{\phi^{(2)}} \circ \cdots \circ E_{\phi^{(s)}}, \qquad s \geq 0, \quad (2.35)$$

in which each $\phi^{(i)}$ is a non-repeating fat *partition. The coefficients, $\omega_M \in \mathbb{Z}\backslash\{0\}$, and their sign,* $\mathrm{sgn}(\omega_M)$*, are given by,*

$$\omega_M = c(\phi^{(1)}) \cdots c(\phi^{(s)}), \qquad \mathrm{sgn}(\omega_M) = (-1)^{\#\phi^{(1)}+\cdots+\#\phi^{(s)}}, \quad (2.36)$$

where each factor, for non-repeating fat $\phi = \tau^k$ with primary non-repeating fat τ, is given by,

$$c(\phi) = c(\tau^k) = -\nu(k) \star \pi(\tau^k) = -\sum_{d|k} \nu(k/d)\pi(\tau^d). \quad (2.37)$$

Proof: Take $F = 1 + X$ in Theorem 2.6. Then $g_M = \omega_M$ and, from (2.11), $h_M = \theta(k) = \frac{(-1)^{k-1}}{k}$, if $M = X^k$ and 0, otherwise. Hence, by (2.22), $\omega_{X^k} = \nu(k) \star \theta(k) = 1$, if $k = 1$; 0, otherwise. If $M = (E_\tau \circ Q)^k = E_{\tau^k} \circ Q$, with τ primary fat, then, by (2.22), $\omega_{E_{\tau^k} \circ Q} = \nu(k) \star (0 - \pi(\tau^k)\omega_Q) = -\nu(k) \star \pi(\tau^k)\omega_Q = c(\tau^k)\omega_Q$. In particular, if $\tau^k = m^k$, with $1 < m \in \mathbb{N}$, then $\omega_{E_{m^k} \circ Q} = \nu(k) \star (0 - \pi(m^k)\omega_Q) = -\nu(k) \star \theta(k)\omega_Q = -\omega_Q$ if $k = 1$; 0, otherwise. Summarizing, let $\phi = \tau^k$, be fat, then, $\omega_{E_\phi \circ Q} = c(\phi)\omega_Q$, where $c(\phi)$ is defined by (2.37). Moreover, if $\phi = m^k$, with $k > 1$, i.e., when ϕ is repeating, then $\omega_{E_\phi \circ Q} = 0$. This means that the molecular species that can contribute to $\mathrm{Lg}(1 + X)$ are of the form X or of the form $E_\phi \circ Q$, where ϕ is non-repeating fat and Q also contribute to $\mathrm{Lg}(1 + X)$. This implies that formula (2.36) for ω_M holds, since, in this case, $\omega_{E_\phi \circ Q} = c(\phi)\omega_Q$. The sign of ω_M follows from the fact that the leading term in (2.37) corresponds to $d = k$ and its sign is $(-1)^{\#\phi}$. □

The virtual species, $-\mathrm{Lg}(1-X) = \mathrm{Lg}\frac{1}{(1-X)}$, of " connected " linear orders, is a " cousin" of the tensorial species, $\mathrm{Lie}(X)$, of free Lie algebras (see A. Joyal in [Joy86] and C. Reutenauer in [Reu86]).

Corollary 2.8 *The following molecular expansion holds,*

$$-\mathrm{Lg}(1 - X) = \sum_M \ell_M M, \tag{2.38}$$

where each molecular component, M, is of the form of a finite composition,

$$M = E_{\phi^{(1)}} \circ E_{\phi^{(2)}} \circ \cdots \circ E_{\phi^{(s)}} \circ X^{2^j}, \quad s \geq 0, \quad j \geq 0, \tag{2.39}$$

in which each $\phi^{(i)}$ is a non-repeating fat *partition. The coefficients, $\ell_M \in \mathbb{Z}\backslash\{0\}$, and their sign, $\mathrm{sgn}(\ell_M)$, are given by,*

$$\ell_M = c(\phi^{(1)}) \cdots c(\phi^{(s)}), \qquad \mathrm{sgn}(\ell_M) = (-1)^{\sum_{i=1}^{s} \#\phi^{(i)}}. \tag{2.40}$$

Proof: Take $F = 1 - X$ in Theorem 2.6 and argue as in the proof of Corollary 2.7. A simpler proof is to use $\frac{1}{1-X} = (1 + X)(1 + X^2)(1 + X^4) \cdots (1 + X^{2^j}) \cdots$ and apply Corollary 2.7, to obtain,

$$-\mathrm{Lg}(1 - X) = \mathrm{Lg}\left(\frac{1}{1 - X}\right) = \sum_{j \geq 0} \mathrm{Lg}(1 + X^{2^j}) = \sum_{M \neq 1, j \geq 0} \omega_M M(X^{2^j}). \tag{2.41}$$

□

Of course, from (1.5), we have the underlying cycle index series,

$$Z_{\mathrm{Lg}(1+X)} = \sum_{k \geq 1} \frac{\mu(k)}{k} \log(1 + p_k), \qquad Z_{-\mathrm{Lg}(1-X)} = -\sum_{k \geq 1} \frac{\mu(k)}{k} \log(1 - p_k). \tag{2.42}$$

Many other applications of Theorem 2.6 are possible. But, due to lack of space, we conclude with a Table, obtained using Maple, which gives the explicit molecular expansion of the combinatorial logarithm up to degree 10. Much more extended tables are easily obtained.

3 Compact table for the combinatorial logarithm up to degree 10

Note that for $s = 0$ in (2.35), the sequence, $\phi^{(1)}, \dots, \phi^{(s)}$, of non-repeating fat partitions is empty. So that the corresponding s-fold composition is a 0-fold composition, hence is equal to X, which is the neutral element under composition. Moreover, the corresponding product (2.36) being empty, is equal to 1. This is coherent with the fact that the molecular expansion of $\mathrm{Lg}(1+X)$ starts with X (see (2.9)). Moreover, if $M = E_{\phi^{(1)}} \circ \cdots \circ E_{\phi^{(s)}}$, is a molecular component in (2.34) then, for any permutation, $\sigma \in S_s$, $M^\sigma = E_{\phi^{(\sigma(1))}} \circ \cdots \circ E_{\phi^{(\sigma(s))}}$, is also a molecular component and the coefficients are equal: $\omega_M = \omega_{M^\sigma}$. Table 1, below[(x)], gives the molecular expansion (2.34) up to degree 10 and uses this fact to compact its size. The following convention is used, for $s > 1$ and $M = E_{\phi^{(1)}} \circ \cdots \circ E_{\phi^{(s)}}$:

$$\overline{M} \underset{\mathrm{def}}{=} \sum_{N \in \Lambda} N, \qquad \Lambda = \{M^\sigma : \sigma \in S_s\}. \tag{3.1}$$

Acknowledgements. The authors would like to thank Jérôme Tremblay for his LaTeXand Maple help.

n	Compact form for the terms of degree n in the combinatorial logarithm $\mathrm{Lg}(1+X)$
0	0
1	X
2	$-E_2$
3	$-E_3 + E_{1,2}$
4	$-E_4 + E_{1,3} - E_{1^2,2} + E_2 \circ E_2$
5	$-E_5 + E_{1,4} + E_{2,3} - E_{1^2,3} - E_{1,2^2} + E_{1^3,2}$
6	$-E_6 + E_{1,5} + E_{2,4} - E_{1^2,4} - 2E_{1,2,3} + E_{1^3,3} + 2E_{1^2,2^2} - E_{1^4,2} + \overline{E_2 \circ E_3} - \overline{E_2 \circ E_{1,2}}$
7	$-E_7 + E_{1,6} + E_{2,5} + E_{3,4} - E_{1^2,5} - 2E_{1,2,4} - E_{1,3^2} - E_{2^2,3} + E_{1^3,4} + 3E_{1^2,2,3} + E_{1,2^3}$ $-E_{1^4,3} - 2E_{1^3,2^2} + E_{1^5,2}$
8	$-E_8 + E_{1,7} + E_{2,6} + E_{3,5} - E_{1^2,6} - 2E_{1,2,5} - 2E_{1,3,4} - E_{2^2,4} - E_{2,3^2} + E_{1^3,5} + 3E_{1^2,2,4}$ $+2E_{1^2,3^2} + 3E_{1,2^2,3} - E_{1^4,4} - 4E_{1^3,2,3} - 2E_{1^2,2^3} + E_{1^5,3} + 2E_{1^4,2^2} - E_{1^6,2} + \overline{E_2 \circ E_4}$ $-\overline{E_2 \circ E_{1,3}} + \overline{E_2 \circ E_{1^2,2}} - E_2 \circ E_2 \circ E_2$
9	$-E_9 + E_{1,8} + E_{2,7} + E_{3,6} + E_{4,5} - E_{1^2,7} - 2E_{1,2,6} - 2E_{1,3,5} - E_{1,4^2} - E_{2^2,5} - 2E_{2,3,4}$ $+E_{1^3,6} + 3E_{1^2,2,5} + 3E_{1^2,3,4} + 3E_{1,2^2,4} + 3E_{1,2,3^2} + E_{2^3,3} - E_{1^4,5} - 4E_{1^3,2,4} - 2E_{1^3,3^2}$ $-6E_{1^2,2^2,3} - E_{1,2^4} + E_{1^5,4} + 5E_{1^4,2,3} + 3E_{1^3,2^3} - E_{1^6,3} - 3E_{1^5,2^2} + E_{1^7,2} + E_3 \circ E_3$ $-\overline{E_3 \circ E_{1,2}} + E_{1,2} \circ E_{1,2}$
10	$-E_{10} + E_{1,9} + E_{2,8} + E_{3,7} + E_{4,6} - E_{1^2,8} - 2E_{1,2,7} - 2E_{1,3,6} - 2E_{1,4,5} - E_{2^2,6} - 2E_{2,3,5}$ $-E_{2,4^2} - E_{3^2,4} + E_{1^3,7} + 3E_{1^2,2,6} + 3E_{1^2,3,5} + 2E_{1^2,4^2} + 3E_{1,2^2,5} + 6E_{1,2,3,4} + E_{1,3^3}$ $+E_{2^3,4} + 2E_{2^2,3^2} - E_{1^4,6} - 4E_{1^3,2,5} - 4E_{1^3,3,4} - 6E_{1^2,2^2,4} - 6E_{1^2,2,3^2} - 4E_{1,2^3,3} + E_{1^5,5}$ $+5E_{1^4,2,4} + 2E_{1^4,3^2} + 10E_{1^3,2^2,3} + 2E_{1^2,2^4} - E_{1^6,4} - 6E_{1^5,2,3} - 5E_{1^4,2^3} + E_{1^7,3} + 4E_{1^6,2^2}$ $-E_{1^8,2} + \overline{E_2 \circ E_5} - \overline{E_2 \circ E_{1,4}} - \overline{E_2 \circ E_{2,3}} + \overline{E_2 \circ E_{1^2,3}} + \overline{E_2 \circ E_{1,2^2}} - \overline{E_2 \circ E_{1^3,2}}$

Tab. 1: Compact form for the terms of degree n in the combinatorial logarithm, for $0 \le n \le 10$.

[(x)] Made using the Maple package `combinat` together with the `define('linear')` command.

References

[BLL98] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial species and tree-like structures*, Encyclopedia of Mathematics and its Applications, vol. 67, Cambridge U. Press, Cambridge, 1998.

[GL11] I. M. Gessel and J. Li, *Enumeration of point-determining graphs*, J. Combinatorial Theory Ser. A **118** (2011), 591–612.

[Joy81] A. Joyal, *Une théorie combinatoire des séries formelles*, Adv. Math. **42** (1981), no. 1, 1–82.

[Joy85] ______, *Règle des signes en algèbre combinatoire*, C.R. Math. Rep. Acad. Sci. Canada **7** (1985), no. 5, 285–290.

[Joy86] ______, *Foncteurs analytiques et espèces de structures*, Combinatoire énumérative (Montreal, Que., 1985) (P. Leroux G. Labelle, ed.), Lecture Notes in Mathematics, vol. 1234, Springer, Berlin, 1986, p. 126–159.

[Lab90] G. Labelle, *Dérivées directionnelles et développements de taylor combinatoires*, Discrete Mathematics **79** (1990), no. 3, 279 –297.

[Lab08] ______, *New combinatorial computational methods arising from pseudo-singletons*, Discrete Mathematics and Theoretical Computer Science (2008), 247–258.

[Li12] J. Li, *Combinatorial logarithm and point-determining cographs*, The electronic journal of combinatorics **19** (2012), no. 3, P8.

[Reu86] C. Reutenauer, *Theorem of Poincaré-Birkhoff-Witt, the logarithm and symmetric group representations of degrees equal to Stirling numbers*, Combinatoire énumérative (Montreal, Que., 1985) (P. Leroux G. Labelle, ed.), Lecture Notes in Mathematics, vol. 1234, Springer, Berlin, 1986, p. 267–284.

[Yeh86] Y. N. Yeh, *The calculus of virtual species and k-species*, Combinatoire énumérative (Montreal, Que., 1985) (P. Leroux G. Labelle, ed.), Lecture Notes in Mathematics, vol. 1234, Springer, Berlin, 1986, p. 351–369.