

On Combinatorial Differential Equations

GILBERT LABELLE*

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

Submitted by G.-C. Rota

We analyse the solution set of first-order initial value differential problems of the form

$$dy/dx = f(x, y), \quad y(0) = 0$$

in the context of combinatorial species in the sense of A. Joyal (*Adv. in Math.* **42** (1981), 1–82). It turns out that the situation is much richer than in the case of formal power series: **many non-isomorphic combinatorial solutions are possible for a given problem, although they all have the same underlying generating series.** We give many examples of this phenomenon and also elaborate a combinatorial Newton–Raphson iterative scheme for the construction of the solutions. The multi-dimensional case is treated explicitly. © 1986 Academic Press, Inc.

1. INTRODUCTION

Using suitable translations, a differential problem of the form

$$\partial y_i / \partial x_j = f_{ij}(x_1, \dots, x_k, y_1, \dots, y_p), \quad 1 \leq i \leq p, \quad 1 \leq j \leq k \quad (1.1)$$

$$y_i(x_1^0, \dots, x_k^0) = y_i^0, \quad 1 \leq i \leq p \quad (1.1a)$$

can always be put in a “normalized” equivalent form in which the initial conditions (1.1a) are reduced to

$$y_i(0, \dots, 0) = 0, \quad 1 \leq i \leq p. \quad (1.2)$$

In the case where the f_{ij} are formal power series

$$f_{ij} = f_{ij}(x_1, \dots, x_k, y_1, \dots, y_p) \in \mathbb{K}[[x_1, \dots, x_k, y_1, \dots, y_p]] \quad (1.3)$$

(where \mathbb{K} is a field of characteristic 0), then such a normalized problem has a solution

$$a = (a_i)_{1 \leq i \leq p}, \quad a_i = a_i(x_1, \dots, x_k) \in \mathbb{K}[[x_1, \dots, x_k]] \quad (1.4)$$

* Grants FCAC EQ1608, Québec and CRSNG A5660, Canada.

if and only if the following compatibility conditions are satisfied:

$$(\forall i)(\forall \mu)(\forall \nu) \mathcal{D}_\mu f_{i,\nu} = \mathcal{D}_\nu f_{i,\mu} \quad (1.5)$$

where

$$\mathcal{D}_\mu := \frac{\partial}{\partial x_\mu} + \sum_{i=1}^p f_{i,\mu} \frac{\partial}{\partial y_i}. \quad (1.6)$$

The solution is then necessarily unique. In fact, the coefficients in the series (1.4) are completely determined recursively by simply taking the successive mixed formal partial derivatives of (1.1) and evaluating each of them at the origin. Conditions (1.5) are equivalent to

$$(\forall \mu)(\forall \nu) \quad \mathcal{D}_\mu \mathcal{D}_\nu = \mathcal{D}_\nu \mathcal{D}_\mu \quad (1.7)$$

and imply that these mixed derivatives can be taken in any order. One then has the following explicit formula

$$\left(\frac{\partial}{\partial x_1} \right)^{n_1} \cdots \left(\frac{\partial}{\partial x_k} \right)^{n_k} a_i(x_1, \dots, x_k) \big|_{\mathbf{x}=0} = \mathcal{D}_1^{n_1} \cdots \mathcal{D}_k^{n_k} y_i \big|_{\mathbf{x}=0, \mathbf{y}=0} \quad (1.8)$$

which is at the basis of Gröbner's theory [11, 12] for differential equations (see also [18, 19]). In the case of one independent variable ($k=1$), the corresponding normalized problem (1.1), (1.2) is automatically compatible and reduces to the form

$$y'_i = f_i(x, y_1, \dots, y_p), \quad 1 \leq i \leq p \quad (1.9)$$

$$y_i(0) = 0, \quad 1 \leq i \leq p \quad (1.10)$$

where $f_i \in \mathbb{K}[[x, y_1, \dots, y_p]]$. Here we always have both existence and unicity of the solution

$$a = (a_i)_{1 \leq i \leq p}, \quad a_i = a_i(x) \in \mathbb{K}[[x]]. \quad (1.11)$$

Via the method of generating series, various differential systems have already appeared in combinatorics to enumerate or define certain families of structures. In this respect we have the following two main directions of research:

- The *first* considers structures that are placed on *arbitrary* finite sets. For example, Goulden, Jackson, and Reilly [10] show that, for fixed $k \leq 4$, the generating series

$$A(x_1, \dots, x_k) = \sum a_{n_1, \dots, n_k} \frac{x_1^{n_1} \cdots x_k^{n_k}}{n_1! \cdots n_k!} \quad (1.12)$$

where a_{n_1, \dots, n_k} is the number of simple graphs whose underlying set consists of n_i vertices of degree i (for $i = 1, \dots, k$), satisfies a certain partial differential system with initial conditions. Moreover, by making use of the symbolic algebra system VAXIMA (based on the system MACSYMA developed by MIT [28]), they have deduced, for example, that the series

$$R_k(x) = \sum_{n \geq 0} r_k(n) x^n / n! \quad (1.13)$$

where $r_k(n)$ denotes the number of k -regular simple graphs on n vertices, satisfies a linear homogeneous differential equation with polynomial coefficients (in x) for each $k \leq 4$. In other words, the sequence $r_k(n)$ is polynomially recursive in the terminology of [31]. They obtained an analogous result concerning the number $q_k(n)$ of k -regular graphs on n vertices. These results solved an open problem stated in Stanley [31].

• The *second* direction of research involves structures which are placed on sets which are, this time, *previously equipped with a linear* (i.e., total) *order*. For example, the theories of alternating, Jacobi's or André's permutations are making use, in an *essential manner*, of a linear order on the underlying sets. Viennot [33] has used differential equations to study these families of permutations. To give a simple example, let us mention that a combinatorial interpretation of the Euler's numbers E_m , $m \geq 0$, can be based on alternating permutations and the differential system

$$y_1' = 1 + y_1^2, \quad y_2' = y_1 y_2, \quad y_1(0) = 0, \quad y_2(0) = 1 \quad (1.14)$$

which is satisfied by $y_1(x) = \operatorname{tg} x$ and $y_2(x) = \sec x$, where

$$\operatorname{tg} x + \sec x = \sum E_m x^m / m! \quad (1.15)$$

More generally, Collins, Goulden, Jackson, and Nierstrasz [7] have shown (among other things) that the enumeration of permutations of $[n] = \{1 < 2 < \dots < n\}$ having a "periodic shape" of "ups" and "downs" is directly connected with certain differential systems of the Riccati type; see also Longtin [24]. Finally, let us mention that the differential equations satisfied by the classical elliptic functions of Jacobi, namely $\operatorname{cn}(u, \alpha)$, $\operatorname{sn}(u, \alpha)$, and $\operatorname{dn}(u, \alpha)$, have given rise to new combinatorial structures on linearly ordered sets; see Dumont [9] and Viennot [32].

By making use of the theory of combinatorial species, in the sense of Joyal [14], one can see that these two main directions of research are fundamentally distinct (see also Bergeron's thesis [2]):

• In the *first case*, one works with *multisorted ordinary species* (we shall simply say *species*): these are, by definition, functors of the form

$$G: \mathbb{B}^m \rightarrow \mathbb{E}, \quad \mathbb{B}^m = \mathbb{B} \times \cdots \times \mathbb{B} \quad (m \text{ factors}) \quad (1.16)$$

where \mathbb{B} is the category of *finite sets* with *bijections*, \mathbb{E} is the category of *finite sets* with *functions* and $m \in \mathbb{N}$ is the number of sorts of points on which the structures are built.

• In the *second case*, one works, instead, with *multisorted \mathbb{L} -species* (we shall simply say *\mathbb{L} -species*): These are, by definition, functors of the form

$$F: \mathbb{L}^m \rightarrow \mathbb{E}, \quad \mathbb{L}^m = \mathbb{L} \times \cdots \times \mathbb{L} \quad (m \text{ factors}) \quad (1.17)$$

where \mathbb{L} is the category of *linearly finite sets* with *increasing bijections*.

It should be noted that if one wants to work with “weighted” structures, the category \mathbb{E} in (1.16) and (1.17) must be replaced by the category $\mathbb{E}_{\mathbb{K}}$ of \mathbb{K} -*weighted finite sets*, where \mathbb{K} is a suitably chosen commutative ring (see [14, 21]).

Two species G_1, G_2 given by (1.16) (resp. \mathbb{L} -species F_1, F_2 given by (1.17)) are said to be *isomorphic* if there exists a *natural isomorphism* between the *functors*

$$G_1 \simeq G_2 \quad (\text{resp. } F_1 \simeq F_2). \quad (1.18)$$

In both the cases of species and \mathbb{L} -species, the structures are counted by using *exponential* (i.e., with factorials) generating series. However, the main distinctions between the two situations can be summarized as follows: For $i = 1, 2$, put

$$\gamma_i = \gamma_i(x_1, \dots, x_k) = \text{generating series of } F_i$$

and

$$g_i = g_i(x_1, \dots, x_k) = \text{generating series of } G_i$$

then it can be easily shown that

$$F_1, F_2 \text{ isomorphic} \Leftrightarrow \gamma_1(x_1, \dots, x_k) = \gamma_2(x_1, \dots, x_k). \quad (1.19)$$

while

$$G_1, G_2 \text{ isomorphic} \not\Leftrightarrow g_1(x_1, \dots, x_k) = g_2(x_1, \dots, x_k). \quad (1.20)$$

In other words, the simple knowledge of the generating series γ of a given \mathbb{L} -species F is sufficient to completely “encode” the isomorphism type of F .

In the case of (ordinary) species, the situation is considerably more complex: the generating series g of a given species G contains only very partial information about the isomorphism type of G .

Moreover, the stabilizer of every Γ -structure is always a trivial group; this is not the case for the stabilizer of a G -structure: it can be isomorphic to *any* finite group.

In the light of these preliminary remarks, it is natural to forget about the underlying generating series and, instead, *lift* the differential problems *themselves* to the level of species (and \mathbb{L} -species). This is indeed possible because of the fact that *combinatorial* operations of derivation, substitution, multiplication, addition, etc., have been defined at this level [14]. Accordingly, given a family

$$(F_{i,j} = F_{i,j}(X_1, \dots, X_k, Y_1, \dots, Y_p))_{1 \leq i \leq p, 1 \leq j \leq k} \quad (1.21)$$

of species (or \mathbb{L} -species) on $k + p$ sorts of points $X_1, \dots, X_k, Y_1, \dots, Y_p$, one may wish to study the families of species (or \mathbb{L} -species)

$$(A_i = A_i(X_1, \dots, X_k))_{1 \leq i \leq p} \quad (1.22)$$

on k sorts, which satisfy the *combinatorial* system

$$\partial Y_i / \partial X_j \simeq F_{i,j}(X_1, \dots, X_k, Y_1, \dots, Y_p) \quad (1.23)$$

$$Y_i[\phi, \dots, \phi] = \phi \quad (1.24)$$

$1 \leq i \leq p, 1 \leq j \leq k$. The arrow in (1.23) denotes a natural isomorphism while the condition (1.24) means that there is no A_i -structure on the empty multiset, for $i = 1, \dots, p$.

In the case of \mathbb{L} -species, the general study of such systems has recently been undertaken by Leroux in a joint work with Viennot [23], when $k = 1$. The solutions they found take the form of special kinds of rooted trees which they call *enriched increasing rooted trees*. They have moreover obtained a unicity theorem for these solutions (up to isomorphism). This result constitutes a kind of combinatorial counterpart for the existence and unicity of the solutions of (1.9 and 1.10) in the case of formal power series (see (1.19)). It is interesting to note that two combinatorial solutions of a given system, although isomorphic, may very well “look” quite different. For example, the \mathbb{L} -species A of “complete odd binary increasing trees” and the \mathbb{L} -species B of “alternating odd permutations starting with a rise” are isomorphic because they both satisfy the combinatorial problem

$$\partial Y / \partial X \simeq 1 + Y^2, \quad Y[\phi] = \phi. \quad (1.25)$$

Leroux and Viennot also show how to adapt the method of “combinatorial bloomings” (“éclosions combinatoires”: see Labelle [18, 19]) to the context of the theory of \mathbb{L} -species. More precisely, they show how to use enriched, increasing, rooted trees to solve the problem of the combinatorial interpretation and proof of the Lie–Gröbner formulas [11, 12] giving the explicit formal power series solution of differential systems such as (1.9) and (1.10).

The goal of the present work is to undertake a combinatorial study of the differential system (1.23) and (1.24) in the setting, this time, of *ordinary species*. Since the generating series of a species encodes *only a part* of its isomorphism type (1.20), we must expect a more complex classification of the solutions of (1.23) and (1.24), *up to isomorphism*, than in the case of the above \mathbb{L} -species. The situation is very rich and we show, in Section 2, that the number of non-isomorphic solutions may vary greatly from problem to problem.

More precisely, we describe (Theorem A) explicit differential problems having *exactly* m non-isomorphic solutions when $m \in \mathbb{N}$ or $m = 2^{\aleph_0}$ (the cardinality of the continuum). We do that by using a notion of “molecular decomposition” of species and a faithful coding of their isomorphism type with the aid of special kinds of multigraphs in which the effect of the operators $\partial/\partial X_j$ can be easily read. Using these notions, Section 3 discusses the “combinatorial approximations” of the solutions of the system (1.23), (1.24). We show (Theorem B) that when two such approximations \mathcal{A} and \mathcal{A}^+ have, respectively, a contact of order N and $2N+2$ with a given solution, then these two approximations are related to each other by a *linear first-order combinatorial differential equation* of a particular form. From that, we deduce efficient iterative schemes (of the Newton–Raphson type) permitting the generation of large families of species of structures (as well as their underlying generating series) starting from combinatorial differential equations, as initial “seeds.” Another consequence of these iterative schemes is the fact that they can be used to prove that the differential system (with $k = 1$)

$$Y'_i = F_i(X, Y_1, \dots, Y_p), \quad Y_i[\phi] = \phi, \quad 1 \leq i \leq p \quad (1.26)$$

always possesses a *canonical combinatorial solution*, among an infinity, in the more general setting of “virtual species.” This follows from remarkable recent *substitution* and *integration* formulas, due to Joyal [16], for virtual species. One of the main interests for this result about virtual species is the fact that the system (1.26) may very well have $m = 0$ combinatorial solution in the context of (ordinary non-virtual) species, because of Theorem A. We mention, at the end of the work, some directions for future exploration.

2. THE NUMBER OF COMBINATORIAL SOLUTIONS

We shall make a constant use of the language, conventions, and basic results of the combinatorial theory of species in the sense of Joyal [14] (see also Labelle [21]). We shall assume, in particular, that the reader knows how to add, multiply, substitute, and derive (multisorted) species; that is, functors of the form (1.16).

Consider a matrix $F = (F_{ij})_{1 \leq i \leq p, 1 \leq j \leq k}$, of species

$$F_{ij} = F_{ij}(X_1, \dots, X_k, Y_1, \dots, Y_p) \quad (2.1)$$

on $k + p$ sorts of points $X_1, \dots, X_k, Y_1, \dots, Y_p$. The normalized differential problem under consideration can be written in the form

$$\partial Y_i / \partial X_j = F_{ij}(X_1, \dots, X_k, Y_1, \dots, Y_p), \quad 1 \leq i \leq p, \quad 1 \leq j \leq k \quad (2.2)$$

$$Y_i[\phi, \dots, \phi] = \phi, \quad 1 \leq i \leq p \quad (2.3)$$

where ϕ denotes the empty set.

DEFINITION 2.1. A combinatorial solution of the differential problem (2.2) and (2.3) is a family of species

$$A = (A_i)_{1 \leq i \leq p} = (A_i(X_1, \dots, X_k))_{1 \leq i \leq p} \quad (2.4)$$

on the k sorts of points X_1, \dots, X_k for which there exists a family of *natural isomorphisms* $\theta = (\theta_{ij})_{1 \leq i \leq p, 1 \leq j \leq k}$ of the form

$$\theta_{ij}: \partial A_i / \partial X_j \simeq F_{ij}(X_1, \dots, X_k, A_1, \dots, A_p). \quad (2.5)$$

Moreover, the family A must satisfy

$$A_i[\phi, \dots, \phi] = \phi, \quad 1 \leq i \leq p. \quad (2.6)$$

That is, there is no A_i -structure on the empty multiset, for $i = 1, \dots, p$.

Remark 2.2. A natural isomorphism between two species is often also called a "combinatorial equality." Thus, by abuse of notation, we shall even write (2.5) in the form

$$\partial A_i / \partial X_j = F_{ij}(X_1, \dots, X_k, A_1, \dots, A_p) \quad (2.7)$$

in cases where the explicit mention of the θ_{ij} is not pertinent. The compatibility conditions (1.5) can be lifted to the combinatorial level by using appropriate isomorphisms.

The concept of a combinatorial solution is illustrated by the following examples:

EXAMPLE 2.3. Unisorted case $k = p = 1$. Consider a single sort X of points and define a species $A = A(X)$ by the combinatorial formula

$$A(X) = XL(E_2(X)) \quad (2.8)$$

where L and E_2 denote the species of *linear orders* and *2-points sets*, respectively. An A -structure thus consists in a singleton followed by a (possibly empty) sequence of disjoint (unordered) pairs of points. It is easy to verify that A is a solution of the differential combinatorial problem

$$Y' = L(E_2(X)) + Y^2, \quad Y[\phi] = \phi. \quad (2.9)$$

EXAMPLE 2.4. Multisorted case $k = p \geq 1$. For $i = 1, \dots, k$ let $A_i = A_i(X_1, \dots, X_k)$ be the species of rooted trees whose roots are points of the sort X_i but whose other vertices are of an arbitrary sort (among X_1, \dots, X_k). It is well known that the family A is recursively characterized by the following system of combinatorial equations

$$A_i = X_i E(A_1 + \dots + A_k), \quad 1 \leq i \leq k \quad (2.10)$$

where $E = E(X) = \exp(X)$ denotes the species of all (finite) sets. A few calculations show that A is a solution of the differential problem (2.2) and (2.3) where the species $F_{i,j}$ are given by

$$F_{i,j} = \delta_{i,j} E(Y_1 + \dots + Y_k) + Y_i L(Y_1 + \dots + Y_k) E(Y_1 + \dots + Y_k) \quad (2.11)$$

where $\delta_{i,j}$ denotes the usual Kronecker symbol, $1 \leq i, j \leq k$. It is interesting to note that the $F_{i,j}$ given by (2.11) do not depend on X_1, \dots, X_k .

Contrarily to the case of formal power series and of \mathbb{L} -species, the number of combinatorial solutions of (2.2) and (2.3) may vary greatly from problem to problem as shown by the following result obtained with the collaboration of J. Labelle and H. Décoste.

THEOREM A. *If m is a finite (possibly null) cardinal number or $m = 2^{\aleph_0}$ (the power of the continuum) then there exists a normalized compatible differential problem having exactly m non-isomorphic combinatorial solutions. Moreover, no differential problem can have exactly $m = \aleph_0$ or $m > 2^{\aleph_0}$ non-isomorphic combinatorial solutions. This non-unicity phenomenon is already present in the case of simple quadratures (with $k = p = 1$) of the form*

$$Y' = F(X), \quad Y[\phi] = \phi \quad (2.12)$$

where $F = F(X)$ is a unisorted species. More explicitly, we have the following examples:

• Case $m = 0$. If $F = E^*$ = the species of pointed sets then (2.12) has no combinatorial solution.

• Case $1 \leq m \in \mathbb{N}$. If $F = 3(m-1) L_2$ where L_2 is the species of linear orders on 2-points sets then (2.12) has exactly m non-isomorphic combinatorial solutions.

• Case $m = 2^{\aleph_0}$. If $F = L$ = the species of linear orders then (2.12) has exactly 2^{\aleph_0} non-isomorphic combinatorial solutions.

Of course, the underlying generating series of the solutions are, in each case, uniquely determined.

We postpone the complete proof of Theorem A to the end of the present section. It requires a preliminary analysis of the decomposition of species into simpler components (cf. Lemma 2.6) as well as a faithful encoding of their isomorphism types in which the combinatorial effect of the operators $\partial/\partial X_j$ can be easily read. The encoding we have chosen consists in a special kind of multigraph having certain symmetry properties (cf. Definition 2.8) in which the operator $\partial/\partial X_j$ is interpreted by erasing certain edges (cf. Lemma 2.10).

The *multicardinality* $|U|$ and the *total cardinality* $\|U\|$ of a multiset $U = (U_1, \dots, U_k)$ are respectively defined by

$$|U| = (|U_1|, \dots, |U_k|) \in \mathbb{N}^k, \quad \|U\| = |U_1| + \dots + |U_k| \in \mathbb{N} \quad (2.13)$$

where $|U_i|$ denotes the cardinality of the set U_i of points of U of sort X_i . Let $n = (n_1, \dots, n_k) \in \mathbb{N}^k$ be a multicardinality and let $H = H(X) = H(X_1, \dots, X_k)$ be a species on k sorts. By convention, the restriction of H to multisets U such that $|U| = n$, is the species $H_n = H_{n_1, \dots, n_k}$ defined by

$$\begin{aligned} H_n[U] &= H[U] & \text{if } |U| = n, \\ &= \phi & \text{otherwise.} \end{aligned} \quad (2.14)$$

The transportation morphisms [14] of H_n are deduced from those of H by restriction. One has the following canonical decomposition

$$H = \sum_n H_n = \sum_n H_{n_1, \dots, n_k} \quad (2.15)$$

where the summation index n varies in the set \mathbb{N}^k of all possible multicardinalities. In ambiguous situations, we shall write H_n in the form $H|_n$ or $H_{,n}$. For example, in the case of a solution $A = (A_i)_{1 \leq i \leq p}$ of (2.2) and (2.3), we write, for each i ,

$$A_i = \sum_n A_i|_{n_1, \dots, n_k} = \sum_n A_{i; n_1, \dots, n_k} \quad (2.16)$$

We shall now see that each of the species H_n entering in the decomposition (2.15) can itself be further decomposed into a *finite sum* of simpler species: the “molecular” species.

DEFINITION 2.5. Let $X = (X_1, \dots, X_k)$ and write $0 = O(X)$ for the empty species. A species $M = M(X)$ is called *molecular* (or *connected*) if and only if $M \neq 0$ and for every species $P = P(X)$ and $Q = Q(X)$,

$$M = P + Q \Rightarrow P = 0 \quad \text{or} \quad Q = 0. \quad (2.17)$$

Because of (2.15), a molecular species must be supported by multisets having a single multicardinality, n say.

LEMMA 2.6. (Molecular decomposition). *For each multicardinality $n = (n_1, \dots, n_k)$, there exists only a finite number*

$$\mu_n = \mu_{n_1, \dots, n_k} > 0 \quad (2.18)$$

of non-isomorphic molecular species

$$M_n^{(i)} = M_n^{(i)}(X) = M_{n_1, \dots, n_k}^{(i)}(X_1, \dots, X_k), \quad i = 1, \dots, \mu_n \quad (2.19)$$

supported by multisets having multicardinality n . Every species $H = H(X) = H(X_1, \dots, X_k)$ has a unique molecular decomposition of the form

$$H = \sum_{n,i} C_n^{(i)}(H) M_n^{(i)} \quad (2.20)$$

in which the coefficients

$$C_n^{(i)}(H) = C_{n_1, \dots, n_k}^{(i)}(H), \quad n \in \mathbb{N}^k, \quad 1 \leq i \leq \mu_n, \quad (2.21)$$

are integers ≥ 0 . Moreover, for any pair H, K of species we have

$$(H, K \text{ isomorphic}) \Leftrightarrow (\forall n)(\forall i)[C_n^{(i)}(H) = C_n^{(i)}(K)]. \quad (2.22)$$

Proof of Lemma 2.6. Since

$$H \simeq K \Leftrightarrow (\forall n \in \mathbb{N}^k)[H_n \simeq K_n] \quad (2.23)$$

we see that it is sufficient to restrict our attention to the category \mathcal{E}_n of the species of the form $H = H_n$ for fixed n . Write

$$S_n = S_{n_1} \times \cdots \times S_{n_k} \quad (2.24)$$

for the group of multipermutations

$$\sigma = (\sigma_i: n_i \simeq n_i)_{1 \leq i \leq k} \quad (2.25)$$

of n considered as a multiset under the usual convention

$$n_i = \{0, 1, \dots, n_i - 1\} \quad (2.26)$$

for $i = 1, \dots, k$. Because of the transportation morphisms, each $H \in \mathcal{E}_n$ induces an action of S_n

$$\begin{aligned} \alpha_H: S_n \times H[n] &\rightarrow H[n] \\ (\sigma, s) &\mapsto H[\sigma](s) = \sigma_{H \cdot s} \end{aligned} \quad (2.27)$$

on the finite set $H[n]$ of H -structures built on the *multiset* n . We have, in fact, an equivalence of categories (in the sense of [26], p. 91)],

$$\alpha: \mathcal{E}_n \rightarrow A_n, \quad H \mapsto \alpha_H \quad (2.28)$$

where A_n denotes the category of actions of S_n on finite sets. In particular, the *molecular species*, in \mathcal{E}_n , correspond (via α) to *connected* (or *transitive*) actions of S_n . Moreover, since two equivalent categories have isomorphic *skeletons* ([26, p. 93]) we can identify canonically the *isomorphism classes* of species in \mathcal{E}_n with the *isomorphism classes* of actions of S_n . The lemma then easily follows from the following two well-known general properties concerning the actions of an arbitrary group G on sets (i.e., G -sets):

(i) Every G -set can be *uniquely* decomposed as a disjoint union of *connected* (i.e., *transitive*) G -sets.

(ii) There is a bijection between the set of isomorphism classes of connected G -sets and the set of *conjugacy classes* of subgroups of G . More specifically, since here $G = S_n$ is a *finite* group, we deduce that the number of isomorphism classes of molecular species, supported by the multisets U such that $|U| = n$ is *finite* and is given by

$$\mu_n = \mu_{n_1, \dots, n_k} = \text{number of conjugacy classes of subgroups of } S_n. \quad (2.29)$$

Hence, we can choose a finite system of representatives

$$(M_{n_1, \dots, n_k}^{(i)})_{1 \leq i \leq \mu_n} \quad (2.30)$$

for the set of these μ_n classes. Moreover, each coefficient $C_n^{(i)}(H)$ in (2.20) is finite because each of the actions (2.27) is itself finite. Note that bijection (ii) takes the following explicit form: to the isomorphism class of $M_n^{(i)}$ it associates the conjugacy class of the stabilizer of one (any one!) of the $M_n^{(i)}$ -structures on the multiset n . ■

Remark 2.7. If we allow *negative* integral coefficients in (2.20) we obtain the isomorphism classes of the so-called *virtual species*: these are the “formal differences” of ordinary species (where the equality

$H - K = H' - K'$ means $H + K' \simeq H' + K$). The usual product of species induces a commutative ring structure \mathcal{R} on the additive group of the virtual species [15]. It is easy to see that the product of molecular species is always molecular:

$$(\forall m)(\forall n)(\forall i)(\forall j)(\exists ! r) M_m^{(i)} M_n^{(j)} = M_{m+n}^{(r)} \quad (2.31)$$

where $m + n = (m_1 + n_1, \dots, m_k + n_k)$. Yeh [29, 34] has shown that every molecular species M can be factorized in a *unique* way (up to the order of the factors) as a *finite product*

$$M = P_1^{\alpha_1} P_2^{\alpha_2} \dots, \quad \alpha_i \geq 0, \quad \sum \alpha_i < \infty \quad (2.32)$$

of simpler molecular species P_i , $i = 1, 2, \dots$, called *atomic species*. Hence, the ring \mathcal{R} is in fact isomorphic to the *factorial ring* $\mathbb{Z}[[P_1, P_2, \dots]]$ involving an infinity of “variables” P_i .

Now let us pursue our analysis by considering the combinatorial effect of the operators $\partial/\partial X_j$, $j = 1, 2, \dots, k$. The partial derivative $\partial M/\partial X_j$ of a molecular species M is *not necessarily* molecular. To see this, simply take two non-trivial molecular species N, Q , put $M = N \cdot Q$ and use the following well-known *combinatorial* identity [14, p. 49]

$$\frac{\partial}{\partial X_j} M = \frac{\partial}{\partial X_j} (N \cdot Q) = \left(\frac{\partial}{\partial X_j} N \right) \cdot Q + N \cdot \left(\frac{\partial}{\partial X_j} Q \right). \quad (2.33)$$

This phenomenon subsists even in the case when $M = P$ is *atomic* (i.e., without a non-trivial factorization). In the unsorted case ($k = 1$), the simplest example of this is given by the species

$$P = E_2 \circ L_2 \quad (\text{atomic}) \quad (2.34)$$

for which one has

$$\partial P/\partial X = 2L_3 = L_3 + L_3 \quad (\text{non-molecular}) \quad (2.35)$$

which is obviously non-molecular (the calculation will be checked using Fig. 3 below). Thus, in the general situation, all that we can assert a priori is that $(\partial/\partial X_j) M_n^{(i)}$ is isomorphic to *either*

- the null species 0 (if $n_j = 0$) or
- a linear combination (with coefficients ≥ 0) of the species

$$M_{n-1_j}^{(1)}, M_{n-1_j}^{(2)}, \dots, M_{n-1_j}^{(\mu_{n-1_j})}, \quad (\text{if } n_j > 0)$$

where $n - 1_j = (n_1, \dots, n_j - 1, \dots, n_k)$. This immediately follows from Lemma 2.6 and the general definition of the partial derivative of a species as given by Joyal [14].

In view of the specific applications we have in mind, the detailed analysis of an arbitrary species H will be greatly simplified by the introduction of a special diagram $|H|$, called the *type* of H , which faithfully and concisely encodes the isomorphism type (or class) of H (or, which means the same, of the family of actions (2.27) associated with H). In order to facilitate the description of $|H|$, we shall use the following conventions. For every $n = (n_1, \dots, n_k) \in \mathbb{N}^k$, considered as a multiset, write

$$\tau_{i,j}: n \simeq n \quad (0 \leq i \leq n_j - 1) \quad (2.36)$$

for the transposition that interchanges 0 and i within the j th component $n_j = \{0, 1, \dots, n_j - 1\}$ of n . In particular, $\tau_{0,j}$ is interpreted, of course, as the identity transformation on the component n_j . In the *unsorted* case ($k = 1$), we shall simply write τ_i to denote the transposition $\tau_{i,1}$.

DEFINITION 2.8 (The type of a species). Let $H = \Sigma H_n$ be an arbitrary species on k sorts of points X_1, \dots, X_k . The *type* of H is a multigraph $|H|$ defined by

$$|H| = \sum |H_n| \quad (\text{disjoint sum of multigraphs}) \quad (2.37)$$

where, for each $n = (n_1, \dots, n_k)$, $|H_n|$ is constructed as follows:

I. Generate the finite set $H[n]$ of all the H -structures supported by the multiset n .

II. For each $s, s' \in H[n]$ and each $i, j \in \mathbb{N}$ such that $1 \leq i \leq n_j - 1$, $1 \leq j \leq k$, join s and s' by an (unoriented) edge (or loop)

$$s \xrightarrow{i,j} s' \quad (\text{indexed by } i, j) \quad (2.38)$$

if and only if s and s' are interchanged (under transportation) by the action of the transposition $\tau_{i,j}$.

III. For each j such that $n_j = 1$ and $1 \leq j \leq k$, and each $s \in H[n]$, put a loop,

$$0, j \circlearrowleft s \quad (\text{indexed by } 0, j). \quad (2.39)$$

IV. Replace each $s \in H[n]$ by a dot “•” while keeping intact the whole network of edges and loops constructed in steps II and III.

In the unsorted case ($k = 1$), replace (2.38) and (2.39) by

$$s \xrightarrow{i} s' \quad \text{and} \quad 0 \circlearrowleft s \quad (2.39')$$

respectively and $\tau_{i,j}$ by $\tau_{i,1} = \tau_i$.

Remarks 2.9. 1. It is easy to define, in a similar way, the type $|\alpha|$ of an arbitrary action α of S_n . In view of (2.27), we have the equality $|H_n| = |\alpha_{H_n}|$.

2. For each $n \in \mathbb{N}^k$, the set

$$\Gamma = \{\tau_{ij} \mid 1 \leq j \leq k, 1 \leq i \leq n_{j-1}\} \cup \{\tau_{0j} \mid 1 \leq j \leq k, n_j = 1\} \quad (2.40)$$

generates the group S_n . Hence, the steps I, II, and III of the construction of $|H_n|$, constitute an encoding of the actions (2.27). In particular, step III makes a distinction between the cases $n_j = 0$ and $n_j = 1$. Step IV retains only the isomorphism type of each of these actions. Consequently we have for any species H and K

$$H \simeq K \Leftrightarrow |H| = |K| \quad (2.41)$$

and

$$H \text{ is molecular} \Leftrightarrow |H| \text{ is a connected multigraph.} \quad (2.42)$$

It is worth noting that if $H = H_n$ (in particular, if H is molecular) then to recover n from $|H|$, it suffices to take

$$\begin{aligned} n_j &= 0 & \text{if } I_j &= 0 \\ &= 1 + \max I_j & \text{if } I_j &\neq 0 \end{aligned} \quad (2.43)$$

where for $j = 1, \dots, k$ the set I_j is defined by

$$I_j = \{i \mid \overset{ij}{\bullet} \text{ or } i, j \text{ } \bigcirc \text{ appears in } |H|\}. \quad (2.43')$$

3. In our definition of $|H|$ we could have chosen other sets of generators of S_n instead of Γ given by (2.40); for example: those given in Carmichael [6, p. 175]. We have chosen (2.40) because of its symmetry properties and because the resulting multigraphs $|H|$ are easy to manipulate (see below, for example, the proofs of Lemma 2.11 and Theorem A).

4. Table I contains, for the unsorted case ($k = 1$), the *complete list* of all types of the molecular species $M_n^{(i)}$ supported by sets U of cardinality $n \leq 4$. We can see from it that the species 1 (of the empty set), the species X (of all singletons), the species E_r (of all r -sets), and the species C_r (of all circular permutations of length r) are sufficient to generate, via products and substitutions, every $M_n^{(i)}$ ($n \leq 4$) *except* two of them:

- the species $M_4^{(2)} = E_4^\pm$ of all *oriented sets* of 4 elements (2.44)

- the species $M_4^{(6)} = P_4^{\text{bic}}$ of all *bicoloured squares*. (2.45)

THE MOLECULAR SPECIES ON $n \leq 4$ POINTS

$n=0$	\bullet $E_0=1$	$n = \{0, 1, \dots, n-1\}$
$n=1$	$\circ \text{---} \bullet$ $E_1 = X$	$\begin{array}{c} \bullet \\ \\ \bullet \end{array}$: ACTION OF TRANSPOSITION $(0, i)$
$n=2$	$\circ \text{---} \bullet$ E_2	$\bullet \text{---} 1 \text{---} \bullet$ $E_2^\pm = L_2 = X^2$
$n=3$	$\circ \text{---} \bullet$ E_3	$\begin{array}{c} 1 \\ \circ \text{---} \bullet \text{---} 1 \\ 2 \end{array}$ $E_3^\pm = C_3$
	$\bullet \text{---} 2 \text{---} 1 \text{---} \bullet$ $X E_2 = E_3^\bullet$	$\begin{array}{c} 1 \quad 2 \\ \bullet \quad \bullet \\ \quad \\ \bullet \quad \bullet \\ \quad \\ \bullet \quad \bullet \\ 2 \quad 1 \end{array}$ $L_3 = X^3$
$n=4$	$\circ \text{---} \bullet$ E_4	$\begin{array}{c} 1 \\ \circ \text{---} \bullet \text{---} 1 \\ 2 \end{array}$ E_4^\pm
	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $E_2 \circ E_2$	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $X E_3 = E_4^\bullet$
	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $E_2 E_2 = L_2 \circ E_2$	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $X C_3$
	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $L_2^\bullet E_2$	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $E_2 \circ L_2$
	$\begin{array}{c} 1 \quad 2 \quad 3 \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ \quad \quad \\ \bullet \quad \bullet \quad \bullet \\ 2 \quad 3 \quad 1 \end{array}$ $L_4 = X^4$	

TABLE I

By definition, *orienting* an r -set U , amounts to putting the elements of U at the r vertices of an *oriented* $(r - 1)$ -dimensional simplex given in advance (one element of U for each vertex of the simplex). By definition, a *bicoloured square* is an unoriented cycle of length 4 in which two opposite edges have been chosen (i.e., a *linear factor* in the sense of graph theory [13, 20] has been specified in the cycle). Figure 1 shows, for example, how to trace the type $|M_4^{(6)}|$ according to the directives contained in Definition 2.8.

Table I has been obtained in the obvious manner. That is, we listed the whole set of all conjugacy classes of subgroups of every symmetric group

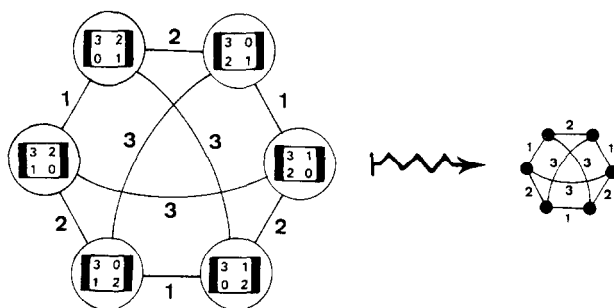


FIGURE 1

S_n (for $n \leq 4$), for each such class we constructed the type of its associated connected action of S_n and, finally, we choose for each such type, a “concrete” representative $M_n^{(i)}$ at the combinatorial level of (unsorted) species. Table I also shows that the sequence μ_n (of Lemma 2.6 with $k = 1$) starts with the values $\mu_0 = 1$, $\mu_1 = 1$, $\mu_2 = 2$, $\mu_3 = 4$, $\mu_4 = 11, \dots$. It is in fact sequence #477 of the book of Sloane [30]. This is a classical sequence in the theory of permutation groups and has been used, for example, in chemistry by Lunn and Senior [25] in a study entitled “Isomerism and Configurations.” For $n = 0$ to 11, the sequence is as follows

$$\mu_n = 1, 1, 2, 4, 11, 19, 56, 96, 296, 554, 1593, 3093, \dots \quad (2.46)$$

In the same general direction, Butler and McKay [3] have used the symbolic algebra system CAYLEY [4, 5] to determine all the transitive permutation groups on k objects, for $k \leq 11$. Using Yeh’s result mentioned above and the values of μ_n given by (2.46), it is not difficult to deduce the values of the numbers π_n defined by

$$\pi_n = \text{the number of types of atomic species supported by } n \text{ points.} \quad (2.47)$$

For $n = 0$ to 11, the sequence runs as follows:

$$\pi_n = 0, 1, 1, 2, 6, 6, 27, 20, 130, 124, 598, 640, \dots \quad (2.48)$$

and is *not included* in Sloane’s book mentioned above. It is interesting to note its *non-monotonicity*. A careful examination of Table I shows that, for $n \leq 4$, the only atomic species are

$$M_1^{(1)}, M_2^{(1)}, M_3^{(1)}, M_3^{(2)}, M_4^{(1)}, M_4^{(2)}, M_4^{(3)}, M_4^{(6)}, M_4^{(7)}, M_4^{(10)} \quad (2.49)$$

in accordance with the first few terms of sequence (2.47). In the multisorted case, it can be shown (see Labelle [22]) that the corresponding families

$$\mu_n = \mu_{n_1, \dots, n_k} \quad \text{and} \quad \pi_n = \pi_{n_1, \dots, n_k} \quad (2.50)$$

where $n \in \mathbb{N}^k$, are related to each other by the remarkable identity

$$\sum_{n_1 \geq 0, \dots, n_k \geq 0} \mu_{n_1, \dots, n_k} t_1^{n_1} \cdots t_k^{n_k} = \prod_{i_1 + \dots + i_k > 0} \left(\frac{1}{1 - t_1^{i_1} \cdots t_k^{i_k}} \right)^{\pi_{i_1, \dots, i_k}}. \quad (2.51)$$

Other results concerning extensions of Table I can be found in [22]. Formula (2.51) can be viewed as a kind of multidimensional extension (to \mathbb{N}^k) of a general decomposition of certain \mathbb{N} -graduated objects into "prime" \mathbb{N} -graduated objects described in Nijenhuis and Wilf [27] under the title "Deus ex Machina."

We shall now describe the relation between the types $|H|$ and $|\partial H / \partial X_j|$.

LEMMA 2.10. *Let $H = H(X_1, \dots, X_k)$ be a species on k sorts of points and let j be a fixed integer, $1 \leq j \leq k$. The type $|\partial H / \partial X_j|$ can be obtained from the type $|H|$ by applying the following algorithm to each of the connected (i.e., molecular) components $|M|$ of $|H|$:*

- I. *If there is no edge (or loop) $\bullet \xrightarrow{\mu} \bullet^{i,j}$ in $|M|$ then replace $|M|$ by the empty multigraph 0;*
- II. *otherwise*
 - IIa. *Put $\mu = \max\{i \mid \bullet \xrightarrow{\mu} \bullet^{i,j} \text{ appears in } |M|\}$;*
 - IIb. *Erase in $|M|$ each edge (or loop) of the form*

$$\bullet \xrightarrow{\mu} \bullet^{i,j} \quad (2.52)$$

while keeping intact the set of other vertices and edges of $|M|$;

- IIc. *If $\mu = 1$ then put a loop $\bullet, i \circlearrowleft$ at each vertex of $|M|$.*

In the unisorted case ($k = 1$), the last conventions of Definition 2.8 are used. For example, (2.52) is written in the form $\bullet \xrightarrow{\mu}$.

Proof of Lemma 2.10. By the linearity of $\partial / \partial X_j$, we need only to consider each individual connected component $|M|$ of $|H|$. Recall that a $\partial M / \partial X_j$ -structure on a multiset $V = (V_1, \dots, V_k)$ is, by definition, an M -structure on

$$U = (U_1, \dots, U_k) = V + \{*_j\} = (V_1, \dots, V_j + \{*_j\}, \dots, V_k) \quad (2.53)$$

where $*_j$ denotes a supplementary point of sort j . Moreover, the transportation of $\partial M / \partial X_j$ -structures coincides with the transportation of M -structures

tures along bijections *leaving invariant* the point $*_j$. Write $M = M_n$ where the $n = (n_1, \dots, n_k)$ is given by (2.43). We have two cases to consider:

• *Case 1.* There is no edge (or loop) $\bullet \xrightarrow{i,j} \bullet$ in $|M|$. In this case we have $n_j = 0$ and we conclude that no $\partial M / \partial M_j$ -structure can exist since we must always have, because of (2.53), $n_j > 0$. In other terms we have $|\partial M / \partial X_j| = 0$ (the empty multigraph).

• *Case 2.* There is at least one edge (or loop) $\bullet \xrightarrow{i,j} \bullet$ in $|M|$. In this case we have

$$\mu = \max \{i \mid \bullet \xrightarrow{i,j} \bullet \text{ appears in } |M|\} = n_j - 1 \geq 0. \quad (2.54)$$

Since we are working up to isomorphism, we can always take the sets V and U in the form

$$V = n - 1_j = (n_1, \dots, n_j - 1, \dots, n_k) \quad \text{and} \quad U = n = (n_1, \dots, n_k) \quad (2.55)$$

and choose $\{*_j\} = \{n_j - 1\}$ (this will make $U_j = n_j = (n_j - 1) \cup \{n_j - 1\} = (n_j - 1) \cup \{*_j\} = V_j + \{*_j\}$ for the given value of j). By definition, the $\partial M / \partial X_j$ -structures on the multiset $n - 1_j$ are identifiable with the M -structures on the multiset n . Hence the vertices of $|\partial M / \partial X_j|$ can be identified with those of $|M|$. Every transposition of the form

$$\tau_{i,v}: n - 1_j \simeq n - 1_j, \quad 1 \leq v \leq k, \quad 0 \leq i \leq n_v - 1 - \delta_{v,j} \quad (2.56)$$

when extended to the whole multiset n , leaves invariant the point $*_j$. Hence $|\partial M / \partial X_j|$ must contain every edge (or loop) $\bullet \xrightarrow{i,v} \bullet$ of $|M|$ *except* those of the form

$$\bullet \xrightarrow{\mu,j} \bullet \quad (\text{i.e., } i = \mu = n_j - 1, v = j). \quad (2.57)$$

Note that in the case when $\mu = n_j - 1 = 1$, we must *add* a loop $0, j$ in accordance with case III of Definition 2.8.

Let us illustrate the derivation algorithm just described with two concrete examples. As a first example, consider the bisorted species

$$H = E_2(E_2(X_1 + X_2)) \quad (2.58)$$

of all partitions in two pairs of points (the sort of each point being taken arbitrarily between X_1 and X_2). Figure 2 shows how to pass from $|H|$ to $|\partial M / \partial X_2|$.

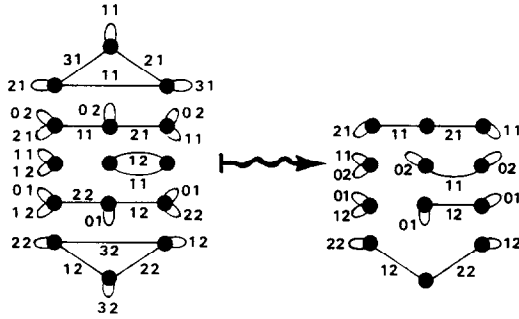


FIGURE 2

The reader can check that the resulting multigraph is the “expected one,” that is, it correspond to the species given by the following calculations:

$$\begin{aligned} \frac{\partial}{\partial X_2} E_2(E_2(X_1 + X_2)) &= E'_2(E_2(X_1 + X_2)) \cdot E'_2(X_1 + X_2) \\ &= (E_2(X_1 + X_2)) \cdot (X_1 + X_2), \end{aligned}$$

which can be identified (up to isomorphism) with the species of all bisorted pointed sets of total cardinality 3.

Another example is given by the unsorted species $P = E_2(L_2(X)) = M_4^{(10)}$ mentioned in (2.54). In view of Table I, the passage from $|P|$ to $|dP/dX| = |P'|$ is described by Fig. 3. This shows, with no calculations, that $P' = 2L_3$ is not molecular. To see that this is the simplest *atomic* species having this property, one has only to check that for each of the nine other atomic species Q of Table I (also given by (2.49)), the multigraph $|Q'|$ is always connected.

The *inverse* process of *combinatorial derivation* of species is *combinatorial integration* of species. It is much more difficult and is not always possible to carry out. We have, in this direction, the following important lemma.

LEMMA 2.11. *Let j be a fixed integer, $1 \leq j \leq k$ and $n = (n_1, \dots, n_k)$ be a given multicardinality. Let $H = H_n = H_n(X_1, \dots, X_k)$ be any species supported*

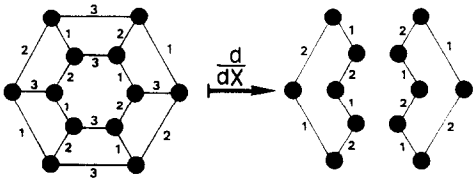


FIGURE 3

by multisets U such that $|U| = n$. Then there exists only a finite number of non-isomorphic species $G = G(X_1, \dots, X_k)$ satisfying the conditions

$$G = G_{n+1_j} \quad \text{and} \quad \frac{\partial G}{\partial X_j} = H \quad (2.59)$$

where $n+1_j = (n_1, \dots, n_j+1, \dots, n_k)$.

Proof of Lemma 2.11. We shall give two proofs of this result. The first proof is based on the fact that to build $|G|$ from $|H|$ we must "reverse" the algorithm contained in Lemma 2.10. More precisely, since $n_j+1 > 0$ and G must be of the form $G = G_{n+1_j}$, the multigraph $|G|$ must necessarily be obtained from $|H|$ by adding new edges (or loops) of the form

$$\bullet \xrightarrow{n_j} \bullet \quad (2.60)$$

Moreover, the following conditions must be satisfied:

- (i) the vertices must not be modified,
- (ii) the new edges (or loops) (2.60) must describe an involution on the totality of the vertices of $|H|$,
- (iii) the resulting multigraph must be the type of a species.

Since the number of vertices of $|H| = |H_n|$ is finite, there is clearly only a finite number of ways to simultaneously satisfy these conditions. In fact, conditions (i) and (ii) are already sufficient to draw this conclusion. The supplementary condition (iii) is more restrictive and shows that it is (at least, a priori) possible that no solutions exist at all.

The second proof is based on the molecular decomposition of species. Let

$$H = \sum_{1 \leq v \leq \mu_n} c_v M_n^{(v)} \quad \text{and} \quad G = \sum_{1 \leq i \leq \mu_{n+1_j}} z_i M_{n+1_j}^{(i)} \quad (2.61)$$

be the molecular decompositions of the given species H and the sought species G . Applying $\partial/\partial X_j$ to each of the molecular species $M_{n+1_j}^{(i)}$ for $i = 1, \dots, \mu_{n+1_j}$, we obtain the decompositions

$$\frac{\partial}{\partial X_j} M_{n+1_j}^{(i)} = \sum_{1 \leq v \leq \mu_n} a_v^{(i)} M_n^{(v)} \quad (2.62)$$

where the coefficients $a_v^{(i)} = a_{v,j}^{(i)}$ are well determined non-negative integers. We must find integral values for the coefficients $z_i \geq 0$ in (2.61) from the equation $\partial M/\partial X_j = H$. Taking into account (2.62), this differential equation amounts to the linear system

$$\sum_{1 \leq i \leq \mu_{n+1_j}} a_v^{(i)} z_i = c_v, \quad 1 \leq v \leq \mu_n \quad (2.63)$$

of μ_n equations in μ_{n+1_j} unknowns. Since $a_v^{(i)}$ and c_v are non-negative integers and (cf. (2.62))

$$(\forall i)(\exists v)(a_v^{(i)} > 0), \quad (2.64)$$

we conclude that the system (2.63) has, at most, a finite number of solutions $z = (z_i)$ such that $z_i \in \mathbb{N}$ for $i = 1, \dots, \mu_{n+1_j}$. ■

Remark 2.12. By further analysis one can show that the general solution of

$$\partial Y / \partial X_j = H_n(X_1, \dots, X_k) \quad (2.65)$$

is (up to isomorphism) of the form

$$G(X_1, \dots, X_k) + K(X_1, \dots, \hat{X}_j, \dots, X_k) \quad (2.66)$$

where G goes through the *finite* set of non-isomorphic species satisfying (2.59) and K goes through (up to isomorphism) the *infinite* set of arbitrary species that are *independent* of the sort X_j . In particular, in the unsorted case ($k = 1$), the species K are necessarily “constant” species (that is, the K -structures live only on the empty set); moreover, if we add the initial condition $Y[\phi] = \phi$ then $K = 0$ (the empty species).

We now have in hand all the tools needed for the complete proof of Theorem A stated above.

Proof of Theorem A. We split it into various cases:

- *Case $m = 0$.* To show that the unsorted quadrature problem (2.12) does not possess a combinatorial solution when $F = E^*$ it suffices to check that even the simpler problem

$$Y' = E_4^*(X), \quad Y[\phi] = \phi, \quad (2.67)$$

restricted to cardinality 4, has no solution $A = A(X)$. Since $E_4^* = M_4^{(4)}$ is a molecular species (see Table I), any eventual solution A must necessarily be isomorphic to one of the 19 above-mentioned molecular species $M_5^{(i)}$ ($1 \leq i \leq 19$). To show that it is impossible, it would suffice, in principle, to make a list of these 19 species and to check, one by one that

$$(\forall i)(1 \leq i \leq 19) \Rightarrow |dM_5^{(i)} / dX| \neq |E_4^*|. \quad (2.68)$$

Happily enough, it is not necessary to make such a list since the following simple *symmetry argument* already suffices to draw the conclusion: The type $|A|$ of a solution A must necessarily be obtained from the type $|E_4^*|$ by *adding* to it new edges (or loops) indexed by the number 4, in such a way that conditions (i), (ii), (iii) of Lemma 2.11 are satisfied. In particular, there

must be *exactly one* new edge (or loop), indexed by the number 4, attached to each of the four vertices of the multigraph $|A|$. Moreover, there must be a total symmetry in the labeling of the edges (and loops) of $|A|$ by the numbers 1, 2, 3, 4 (this is due to the general fact that the transpositions τ_i , interchanging 0 and i , always play a symmetric role relative to each other, when $i \geq 1$). Let us show that it is impossible to satisfy the required conditions by looking at the central vertex of $|E_4^*|$. Two eventualities are, a priori, possible (see Fig. 4):

1. A new loop, labeled 4, is added to the central vertex, or
2. a new edge, labeled 4, joins the central vertex to a peripheral one.

In the first eventuality, each vertex of $|A|$ must have a simple loop, by symmetry. This is impossible since there is already a double loop at each other vertex. In the second eventuality, symmetry would imply the existence, in $|A|$, of double edges of the form

$$\begin{array}{ccc} \text{1} & \text{2} & \text{3} \\ \text{---} & \text{---} & \text{---} \\ \text{2} & \text{3} & \text{1} \end{array} \quad (2.69)$$

This is again impossible since there are no such double edges in $|E_4^*|$ at the beginning. This shows that (2.12) has no solution when $F = E^*$.

• *Case $0 < m \in \mathbb{N}$.* Since the species $3(m-1)L_2$ is supported by 2-sets, Table I shows that every solution A of the differential problem

$$Y' = 3(m-1)L_2, \quad Y[\phi] = \phi \quad (2.70)$$

must necessarily have (up to isomorphism) a molecular decomposition of the form

$$A = z_1 E_3 + z_2 C_3 + z_3 E_3^* + z_4 L_3 \quad (2.71)$$

where the coefficients z_i are non-negative integers to be determined.

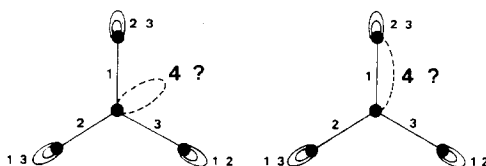


FIGURE 4

Differentiating (2.71) gives

$$\begin{aligned} A' &= z_1 E'_3 + z_2 C'_3 + z_3 E'_3 + z_4 L'_3 \\ &= z_1 E_2 + z_2 L_2 + z_3 (E_2 + L_2) + 3z_4 L_2 \\ &= (z_1 + z_3) E_2 + (z_2 + z_3 + 3z_4) L_2. \end{aligned} \quad (2.72)$$

Hence, the following system must be solved in integers $z_i \geq 0$:

$$z_1 + z_3 = 0, \quad z_2 + z_3 + 3z_4 = 3(m-1). \quad (2.73)$$

The only possible solutions are given by the parametrization

$$z_1 = 0, \quad z_2 = 3r, \quad z_3 = 0, \quad z_4 = m-1-r \quad (2.74)$$

where r runs through the values $0, 1, \dots, m-1$. Hence, problem (2.70) possesses *exactly* the following m non-isomorphic solutions

$$A = 3rC_3 + (m-1-r) L_3, \quad 0 \leq r \leq m-1. \quad (2.75)$$

- *Case $m = 2^{\aleph_0}$.* We have to show that the differential problem

$$Y' = L(X), \quad Y[\phi] = \phi \quad (2.76)$$

possesses exactly 2^{\aleph_0} non-isomorphic solutions (where L denotes the species of linear orders). First note that the number of non-isomorphic solutions is necessarily bounded above by 2^{\aleph_0} , independently of the equation we are dealing with. To see this, one has simply to note that the *total* number of possible molecular decompositions of the form (2.20) is equal to 2^{\aleph_0} . Hence the result will follow if we can build an explicit set of 2^{\aleph_0} of non-isomorphic solutions of (2.76). To achieve this goal, let us generalize the concept of bicoloured squares (see (2.45) above) by the introduction of the species P^{bic} of *bicoloured polygons* defined as follows. A P^{bic} -structure on a finite set U (where $|U| \geq 3$) is a cyclic non-oriented graph having U as vertex-set, together with a linear factor (i.e., a regular subgraph of degree 1). Obviously, there is no P^{bic} -structure on U when $|U|$ is odd and there are exactly $(n-1)!$ such structures when $|U| = n \geq 4$ is even. Accordingly, we have

$$P^{\text{bic}}_{2r-1} = 0, \quad P^{\text{bic}}_{2r} \neq 0, \quad r \geq 2 \quad (2.77)$$

and, in particular, P_4^{bic} is the species of bicoloured squares. Now let us denote by C_n the species of circular permutations of length n . It is easy to

see that, for $r \geq 2$, the species P_{2r}^{bic} and C_{2r} are *molecular* and *non-isomorphic*. Nevertheless, these species have *isomorphic derivatives*, since we have

$$dP_{2r}^{\text{bic}}/dX \simeq L_{2r-1} \simeq dC_{2r}/dX \quad (2.78)$$

as can be seen by examining Fig. 5 (in which $r=3$, $|U| = \{a, b, c, d, e\}$, $|U| = 2r - 1 = 5$).

In the left cycle of this figure, the bold edge (belonging to the linear factor) that is attached to the supplementary point *, indicate, in a *canonical way*, the orientation to be given in the induced chain on U (given by the middle figure). Now, for each

$$S \subseteq \{4, 6, 8, 10, 12, \dots\}, \quad (2.79)$$

let us define a species W_S by the formula

$$W_S = \sum_{n \in S} P_n^{\text{bic}} + \sum_{1 \leq n \notin S} C_n. \quad (2.80)$$

This family (W_S) of species constitutes the 2^{\aleph_0} non-isomorphic solutions of (2.76) that we are looking for. To check this, note first that (because of (2.78) and (2.80))

$$\begin{aligned} \frac{d}{dX} W_S &= \sum_{n \in S} \frac{d}{dX} P_n^{\text{bic}} + \sum_{1 \leq n \notin S} \frac{d}{dX} C_n \\ &= \sum_{1 \leq n} L_{n-1} = L \end{aligned} \quad (2.81)$$

and $W_S[\phi] = \phi$. Second,

$$S, T \subseteq \{4, 6, 8, \dots\}, \quad S \neq T \Rightarrow W_S \neq W_T \quad (2.82)$$

since (2.80) constitutes a molecular decomposition and P_n^{bic} and C_n are non-isomorphic species, for $n \geq 4$.

• *Case $m = \aleph_0$.* It remains to be shown that the general differential problem (2.2), (2.3) cannot have exactly \aleph_0 non-isomorphic solutions. We

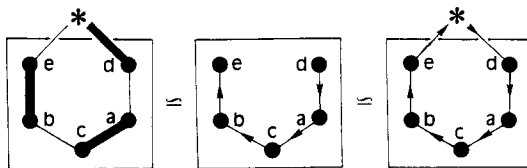


FIGURE 5

shall achieve this by trying to build, in a systematic manner, *all* the solutions $A = (A_i)_{1 \leq i \leq p}$ using *combinatorial successive approximations*. Let $N \geq 0$ be a given integer and suppose that we already have built the canonical components (in the sense of (2.15))

$$A_{i;n} = A_{i;n_1, \dots, n_k}, \quad \text{where} \quad \|n\| = n_1 + \dots + n_k \leq N, \quad 1 \leq i \leq p \quad (2.83)$$

of an "eventual" solution up to total cardinality N . Of course, due to the initial conditions (2.3), we are forced to take, at the beginning

$$A_{i;0, \dots, 0} = 0 \quad (\text{the empty species}), \quad 1 \leq i \leq p. \quad (2.84)$$

We must now only show that there is (at most) a finite number of possible extensions of (2.83) up to multcardinalities v satisfying

$$\|v\| = v_1 + \dots + v_k = N + 1. \quad (2.85)$$

(Indeed, taking, in succession, $N = 0, 1, 2, \dots$ we could then deduce that the total number of solutions is *either* finite (if there exists a value of N from which all the subsequent extensions are unique or inexistent) *or* is equal to 2^{\aleph_0} (otherwise)). Define the approximation $\mathcal{A} = (\mathcal{A}_i)_{1 \leq i \leq p}$, associated to the family (2.83), by

$$\mathcal{A}_i = \sum_{\|n\| \leq N} A_{i;n_1, \dots, n_k}, \quad 1 \leq i \leq p. \quad (2.86)$$

Every v such that $\|v\| = N + 1$ can always be written in the form

$$v = n + 1_j = (n_1, \dots, n_j + 1, \dots, n_k) \quad \text{where} \quad \|n\| = N \quad (2.87)$$

for (at least) one j , $1 \leq j \leq k$, and (at least) one n . Each possible choice for $A_{i;v} = A_{i;n+1_j}$ must be such that $\partial A_{i;v} / \partial X_j$ is supported by the multisets U such that $|U| = n$, where $\|n\| = N$. Hence we must necessarily also have an isomorphism

$$\partial A_{i;v} / \partial X_j \cong [F_{i,j}(X_1, \dots, X_k, \mathcal{A}_1, \dots, \mathcal{A}_p)]|_{n_1, \dots, n_k}. \quad (2.88)$$

This last combinatorial equation is of the general form mentioned in Lemma 2.11, namely:

$$\partial G / \partial X_j \cong H, \quad \text{where} \quad H = H_n(X_1, \dots, X_k) \quad \text{and} \quad G = G_{n+1_j}(X_1, \dots, X_k). \quad (2.89)$$

We conclude, using Lemma 2.11, that there is only a finite number of (non-isomorphic) choices for $A_{i;v}$ (these choices are, a priori, more restricted when j and n in (2.87) are not unique). This concludes the proof of Theorem A. ■

Remark 2.13. In many instances, a symmetry argument, similar to the one used above, can still be used to show that certain differential problems do not possess a solution (among ordinary species). For example, the reader may use such an argument to show that the problem

$$Y' = E_2 \circ E_2(X), \quad Y[\phi] = \phi \quad (2.90)$$

has no solution by trying to add, in a symmetrical manner, an edge, numbered 4, to the type $|M_4^{(3)}| = |E_2 \circ E_2|$ contained in Table I. More generally the following criterion may be useful, in the unsorted case ($k = 1$), to determine or to study the solutions of differential problems. It can be used to decide, for $n > 1$, whether a given non-oriented multigraph A , whose edges are labeled by the numbers 1 to $n - 1$, is or is not the type of a combinatorial species.

CRITERION 2.14. There is a species $H = H_n(X)$ such that $A = |H|$ if, and only if, the following three conditions are satisfied:

(1) For each i , $1 \leq i \leq n - 1$: the set of edges and loops labeled i describes an involution of the vertices of A .

(2) For each i , $1 \leq i \leq n - 1$: each path of the form

$$i(i+1) i(i+1) i(i+1) \quad (2.91)$$

is closed (here we use the convention $i + 1 \equiv 1$ when $i = n - 1$).

(3) For each i, j , $1 \leq i, j \leq n - 1$, $j \neq i$, $j \neq i + 1$: each path of the form

$$i(i+1) jji(i+1) ij \quad (2.92)$$

is closed (here again, we use the convention $i + 1 \equiv 1$ when $i = n - 1$).

The criterion follows directly from the well-known presentation [6] of the symmetric group S_n in terms of "abstract generators" τ_i : for $1 \leq i$, $j \leq n - 1$,

$$\tau_i^2 = 1, \quad (\tau_i \tau_{i+1})^3 = 1, \quad (\tau_i \tau_{i+1} \tau_i \tau_j)^2 = 1 \quad (\tau_n \equiv \tau_1). \quad (2.93)$$

Of course, it is not difficult to formulate similar criteria in the multisorted case ($k > 1$).

Remark 2.15. It is interesting to note that the classical "integration by parts" method can be used to solve, in the ring \mathcal{R} of *virtual species*, differential equations that had no solutions in terms of *ordinary species*. For example, problem (2.12) with $F = E^* = XE$ (the species of pointed sets),

namely $Y' = XE$, $Y[\phi] = \phi$, is unsolvable within ordinary species. However, the “purely formal” calculation

$$E - 1 = \int_0^X E dX = XE - \int_0^X XE dX \quad (2.94)$$

can serve as a “guide” towards the following solution

$$A = E^* - E + 1 \quad (2.95)$$

which belongs to the ring \mathcal{R} . In fact, Joyal has recently shown [16] that, in the unsorted case, a differential equation of the form $Y' = F(X)$, where F is an arbitrary element of the ring \mathcal{R} , is always solvable in the ring \mathcal{R} . Joyal’s solution is a canonical one and is given by the formula

$$Y = E_1 \cdot F - E_2 \cdot F' + E_3 \cdot F'' + \cdots + (-1)^n E_{n+1} \cdot F^{(n)} + \cdots \quad (2.96)$$

where E_r denotes the species of r -sets. Note the the series (2.96) is always summable.

Remark 2.16. Problem (2.12), with $F = L^2 = L \cdot L$, also possesses 2^{\aleph_0} non-isomorphic solutions. These are given by the species

$$Q_S = \sum_{n \in S} L_n + \sum_{1 \leq n \notin S} n C_n \quad (2.97)$$

when S goes through all the subsets of $\mathbb{N} \setminus \{0\}$. Among these, the choice $S = \mathbb{N} \setminus \{0\}$ gives the “classical” solution $Y = L^*$ of all non-empty linear orders. Joyal’s canonical solution (2.96) here takes the particular form

$$E_1 \cdot L^2 - 2!E_2 \cdot L^3 + 3!E_3 \cdot L^4 - 4!E_4 \cdot L^5 + \cdots \quad (2.98)$$

Remark 2.17. The notion of a solution of a combinatorial differential problem, as given by Definition 2.1, can be refined further into the notion of a *strict solution*. By definition, a strict solution of (2.2), (2.3) is an *ordered-pair* (A, θ) satisfying (2.4) through (2.6). Moreover, an *isomorphism*

$$\sigma: (A, \theta) \simeq (B, \omega) \quad (2.99)$$

of strict solutions consists in a natural bijection $\sigma: A \simeq B$ such that the following diagrams commute

$$\begin{array}{ccc} \partial A_i / \partial X_j & \xrightarrow{\theta_{i,j}} & F_{i,j}(X_1, \dots, X_k, A_1, \dots, A_p) \\ \downarrow \partial \sigma_i / \partial X_j & & \downarrow F_{i,j}(X_1, \dots, X_k, \sigma_1, \dots, \sigma_p) \\ \partial B_i / \partial X_j & \xrightarrow{\omega_{i,j}} & F_{i,j}(X_1, \dots, X_k, B_1, \dots, B_p) \end{array} \quad (2.100)$$

for every i, j , $1 \leq i \leq p$, $1 \leq j \leq k$. Any solution A may give rise, a priori, to many non-isomorphic *strict* solutions (A, θ) . It is easy to see that Theorem A remains valid if the word "solution" is replaced by "strict solution" in the cases

$$m = 0, \quad m = 1, \quad m = \aleph_0, \quad m = 2^{\aleph_0}, \quad m > 2^{\aleph_0}. \quad (2.101)$$

However, in the case where $1 < m \in \mathbb{N}$, a detailed analysis shows that example $Y' = 3(m-1)L_2$, $Y[\phi] = \phi$, mentioned above, gives rise to *more than* m non-isomorphic strict solutions. This leaves open the following problem: For which values of $m > 1$ does there exist a combinatorial problem possessing *exactly* m non-isomorphic *strict solutions*?

3. COMBINATORIAL APPROXIMATIONS OF SOLUTIONS

In the proof of Theorem A, we used an iterative scheme enabling us to build, via successive approximations, the solution(s) (if there exists at least one) of the general differential problem

$$\partial Y_i / \partial X_j = F_{ij}(X_1, \dots, X_k, Y_1, \dots, Y_p), \quad Y_i[\phi, \dots, \phi] = \phi \quad (3.1)$$

where $1 \leq i \leq p$, $1 \leq j \leq k$. To go from total cardinalities $\leq N$ to the next one, $N+1$, it was necessary to undertake a systematic search for the solutions of much simpler equations of the form (cf. (2.88), (2.89))

$$\partial Y_i / \partial X_j = P_{ij}(X_1, \dots, X_k) \quad (3.2)$$

in which the species P_{ij} were *independent* of Y_1, \dots, Y_p . Moreover, when N took successively the values $0, 1, 2, \dots$, the P_{ij} had to be recalculated at each step. Borrowing a concept from numerical analysis, we can say that the process is "linearly convergent."

We shall now develop, using purely combinatorial means, another method which is "quadratically convergent." It will make it possible to go directly from total cardinalities $\leq N$ to every total cardinality $\leq 2N+2$. Eq. (3.2) will have to be replaced by equations of the form

$$\partial Y_i / \partial X_j = P_{ij} + \sum_{1 \leq r \leq p} Q_{ij,r} Y_r \quad (3.3)$$

where the right member is, this time, *linear* in Y_1, \dots, Y_p and the species P_{ij} and $Q_{ij,r}$ are independent of Y_1, \dots, Y_p . In the corresponding iterative scheme, N will only have to take the values

$$0, 2, 6, 14, 30, \dots, N, 2N+2, \dots \quad (3.4)$$

and the species $P_{i,j}$ and $Q_{i,j,r}$ will have to be calculated much less often. The essence of the method is contained in Theorem B below. We need first to define certain general notions related to the approximations of species by other species.

DEFINITION 3.1. Let $S \subseteq \mathbb{N}^k$ be a set of multicharacteristics and $H = H(X_1, \dots, X_k)$ be an arbitrary species on k sorts. The *restriction of H to the multicharacteristics of S* is the species H_S defined by

$$H_S = H|_S = \sum_{n \in S} H_{n_1, \dots, n_k} \quad (3.5)$$

where the species H_{n_1, \dots, n_k} are given by (2.14). If $M, N \in \mathbb{N}$, the interval $[M, N]$ is the following set

$$[M, N] = \{n \in \mathbb{N}^k \mid M \leq \|n\| = n_1 + \dots + n_k \leq N\} \subseteq \mathbb{N}^k. \quad (3.6)$$

In particular, when $M = 0$, we say that

$$H_{[0, N]} = H|_{[0, N]} = \sum_{\|n\| \leq N} H_{n_1, \dots, n_k} \quad (3.7)$$

is an *approximation to H of order N* . A *contact of order N* between two species G and H is a natural isomorphism

$$\alpha: G|_{[0, N]} \xrightarrow{\sim} H|_{[0, N]}. \quad (3.8)$$

When α comes from the restriction of a natural transformation $\omega: G \rightarrow H$ to multisets U such that $\|U\| \leq N$ we say that ω *induces a contact of order N* between G and H , and we write

$$G \stackrel{\omega}{\underset{(N)}{=}} H \quad (\text{or } G \underset{(N)}{=} H \text{ if } \omega \text{ is not necessary}). \quad (3.9)$$

Note that every approximation of order N gives rise, in a trivial manner, to a contact of order N :

$$K = H|_{[0, N]} \Rightarrow K \underset{(N)}{=} H. \quad (3.10)$$

THEOREM B (Combinatorial Newton–Raphson scheme for the solutions). *Let $N \geq 0$ be an integer and let (A, θ) be a strict solution of the differential problem (2.2), (2.3) where θ is a family of natural isomorphisms*

$$\theta_{i,j}: \partial A_i / \partial X_j \xrightarrow{\sim} F_{i,j}(X_1, \dots, X_k, A_1, \dots, A_p) \quad (3.11)$$

with $1 \leq i \leq p$, $1 \leq j \leq k$. For each i , consider the following two approximations of A_i :

$$\alpha_i = A_i|_{[0, N]} = \sum_{\|n\| \leq N} A_{in_1, \dots, n_k} \quad (3.12)$$

and

$$\alpha_i^+ = A_i|_{[0, 2N+2]} = \sum_{\|n\| \leq 2N+2} A_{in_1, \dots, n_k} \quad (3.13)$$

of order N and $2N+2$, respectively. Then each species α_i^+ can be split as a sum

$$\alpha_i^+ = \alpha_i + \mathcal{B}_i, \quad \mathcal{B}_i = A_i|_{[N+1, 2N+2]} \quad (3.14)$$

where the species \mathcal{B}_i are related to the α_j as follows:

• By restriction to the multisets V such that $N \leq \|V\| \leq 2N+1$, each isomorphism θ_{ij} induces a contact of order $2N+1$ of the form

$$\frac{\partial \mathcal{B}_i}{\partial X_j} = P_{ij} + \sum_{1 \leq r \leq p}^{(2N+1)} Q_{ij,r} \mathcal{B}_r \quad (3.15)$$

where

$$P_{ij} = P_{ij}(X_1, \dots, X_k) \quad \text{and} \quad Q_{ij,r} = Q_{ij,r}(X_1, \dots, X_k) \quad (3.16)$$

are species defined by

$$P_{ij} = F_{ij}(X_1, \dots, X_k, \alpha_1, \dots, \alpha_p)|_{[N, 2N+1]} \quad (3.17)$$

and

$$Q_{ij,r} = \frac{\partial F_{ij}}{\partial Y_r}(X_1, \dots, X_k, \alpha_1, \dots, \alpha_p)|_{[0, N]}. \quad (3.18)$$

Proof of Theorem B. Because of the isomorphisms (3.11), each $(\partial A_i / \partial X_j)$ -structure can be canonically identified with a F_{ij} -assembly of X_ρ -structures (i.e., singletons of sort X_ρ) together with A_τ -structures (where ρ and τ take arbitrary values, with possible repetitions, in the ranges $1 \leq \rho \leq k$, $1 \leq \tau \leq p$). In accordance with the standard graphical conventions used in the theory of species [8, 14, 17–19, 21, 22], the left-hand side of Fig. 6 exhibits the generic $(\partial A_i / \partial X_j)$ -structure while its right-hand side shows the corresponding F_{ij} -assembly. The star $*$ _j denotes the supplementary point, of sort X_j , within the $(\partial A_i / \partial X_j)$ -structure.

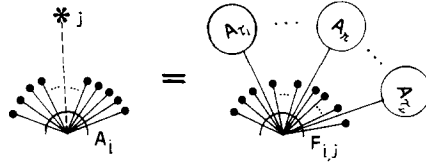


FIGURE 6

By definition, an \mathcal{A}_i -structure is an A_i -structure supported by a multiset U such that $\|U\| \leq N$. Equations (3.14) are equivalent to

$$A_i \mid_{[0,2N+2]} = A_i \mid_{[0,N]} + A_i \mid_{[N+1,2N+2]} \quad (3.19)$$

and define uniquely the species \mathcal{B}_i : a \mathcal{B}_i -structure is an \mathcal{A}_i -structure supported by a multiset such that $N+1 \leq \|U\| \leq 2N+2$. To establish (3.15) combinatorially, we must analyse the internal constituents of an arbitrary $(\partial \mathcal{B}_i / \partial X_j)$ -structure s supported by a multiset V such that

$$N \leq \|V\| \leq 2N+1. \quad (3.20)$$

By analogy with [8], let us use the following terminology:

- An \mathcal{A}_i -structure is called a *light* A_i -structure.
- A \mathcal{B}_i -structure is called a *heavy* A_i -structure.

Being, in particular, a $(\partial \mathcal{A}_i / \partial X_j)$ -structure, the structure s decomposes into a $F_{i,j}$ -assembly as in Figure 6 and we fall into one of the following exclusive cases:

- Case 1.* Every A_i -structure of the assembly is *light*.
- Case 2.* Every one is light, except *at most one*, which is *heavy*.

Indeed, if at least two of these A_i -structures were heavy then the total cardinality $\|V\|$ of the underlying multiset V would satisfy

$$\|V\| \geq (N+1) + (N+1) = (2N+2) > 2N+1 \quad (3.21)$$

contradicting (3.20). Figure 7 summarizes the situation.

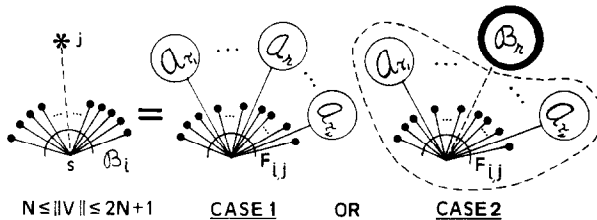


FIGURE 7

Taking (3.20) into account, Fig. 7 contains, in case 1, a $P_{i,j}$ -structure. In case 2, the dotted closed line contains a $Q_{i,j,r}$ -structure. This establishes the contact (3.15) since r can take any value in the range $1 \leq r \leq p$. By examining the underlying total cardinalities in Fig. 7, the reader may check that the intervals $[N, 2N+1]$ and $[0, N]$ entering in (3.17) and (3.18) are the minimal intervals which are compatible with constraint (3.20). ■

Remark 3.2. It is important to fully realize that (3.14) through (3.18) give rise to a quite efficient iterative scheme for a systematic search of the solutions of (3.1). More explicitly, suppose that we already know the canonical components

$$A_{i;n} = A_{i;n_1, \dots, n_k}, \quad \|n\| = n_1 + \dots + n_k \leq N, \quad 1 \leq i \leq p \quad (3.22)$$

of an eventual solution $A = (A_i)_{1 \leq i \leq p}$ up to order N and that we want to find *all its possible extensions*,

$$A_{i;v} = A_{i;v_1, \dots, v_k}, \quad N+1 \leq \|v\| \leq 2N+2, \quad 1 \leq i \leq p \quad (3.23)$$

up to order $2N+2$. We can always write

$$v = n + 1_j, \quad \text{where } N \leq \|n\| \leq 2N+1. \quad (3.24)$$

Taking into account that

$$\begin{aligned} \mathcal{B}_{i;v_1, \dots, v_k} &= A_{i;v_1, \dots, v_k} & \text{if } N+1 \leq \|v\| \leq 2N+2, \\ &= 0 & \text{otherwise,} \end{aligned} \quad (3.25)$$

we obtain that contacts (3.15), when restricted to multicardinality n , give rise to the more explicit equations

$$\frac{\partial}{\partial X_j} A_{i;n+1_j} = P_{i,j;n} + \sum_{\substack{1 \leq r \leq p \\ (u,v) \in T}} Q_{i,j,r;u} A_{r;v} \quad (3.26)$$

where

$$T = T_{N,n} = \{(u, v) \mid u + v = n, \|u\| \leq N, \|v\| \geq N+1\} \quad (3.27)$$

and $N \leq \|n\| \leq 2N+1$. Moreover,

$$(u, v) \in T_{N,n} \Rightarrow N+1 \leq \|v\| \leq \|n\|. \quad (3.28)$$

Hence, the algorithm is clear: Using (3.17) and (3.18), calculate first the various species $P_{i,j}$ and $Q_{i,j,r}$ and then generate the families (3.23) as successive “strata” of the form

$$(A_{i;v_1, \dots, v_k})_{1 \leq i \leq p}, \quad \|v\| = c = \text{const} \quad (3.29)$$

where c takes, one by one, the values $c = N + 1, N + 2, \dots, 2N + 2$. Each component $A_{i,v} = A_{i,n+1_j}$ of each stratum (3.29) *must be chosen in the finite set of solutions of equations* (3.26) whose right-hand sides are known species depending only on previously calculated strata. Each of these choices, if it exists, can be made using the methods described in the proof of Lemma 2.11. Note that, exceptionally, the *first stratum* (3.29), corresponding to $c = N + 1$, is calculated from the simpler equations

$$\frac{\partial}{\partial X_j} A_{i,n+1_j} = P_{i,j;n} \quad (3.30)$$

since the set of previous strata is then empty (indeed, $T_{N,N}$ is empty because of (3.28)). In this particular case, Eq. (3.30) is the same as Eq. (2.88) met earlier.

Starting with $N = 0$, the initial approximation $\alpha = 0$ (i.e., $\alpha_i = A_{i,0,\dots,0} =$ the empty species, $i = 1, \dots, p$) and applying the algorithm for $N = 0, 2, 6, 14, 30, \dots$, in all possible ways, we see that the *set of all possible strata* of the form

$$(A_{i,v_1,\dots,v_k})_{\|v\| = c}, \quad c = 0, 1, 2, \dots, \quad (3.31)$$

constitute, in fact, the vertex-set of a rooted tree \mathcal{S} that we will call *the rooted tree generating the eventual solutions* of the given differential problem. More precisely, each *elementary path* \mathcal{C} in \mathcal{S} joining the root 0 to a stratum can be identified with an approximation α of an eventual solution in the following way:

- α is the (combinatorial) sum of all the stratums in the path \mathcal{C} .
- The order of contact of α (with a corresponding solution) is the length of the path \mathcal{C} .
- Any (existing) solution of (3.1) can be identified an *infinite elementary path* in \mathcal{S} starting from the root 0.

Note that even when a given approximation α (of order c) does not possess an extension, we can still talk about α -structures living on multisets having a “small” (i.e., $\leq c$) total multicardinality. Indeed, it gives rise to a contact of the following form

$$\partial \alpha_i / \partial X_j \underset{(c-1)}{=} F_{i,j}(X_1, \dots, X_k, \alpha_1, \dots, \alpha_p), \quad (3.32)$$

$1 \leq i \leq p, 1 \leq j \leq k$. Of course, the quadratic algorithm, just described, becomes more efficient than the linear algorithm (of the previous section) as N becomes large. For small values of N the two algorithms are comparable.

As an application of these algorithms, we have calculated, in the unsorted case ($k=p=1$), the first few vertices of the rooted tree \mathcal{S} corresponding to the very simple differential problem

$$Y' = (1 + Y)^2, \quad Y[\phi] = \phi. \quad (3.33)$$

Here, each stratum (3.31) reduces to a single species $A_c = A|_c$ and \mathcal{S} is given, for $c = 0, 1, 2, 3, \dots$, by Fig. 8.

The upper (bold) path in \mathcal{S} is extendable to infinity and corresponds to the "usual" solution

$$L^*(X) = X/(1 - X) \quad (3.34)$$

of (3.33) which is the species of non-empty linear orders. Using Table I, we calculated that the 4 strata at level $c = 3$ generate a total of

$$10 + 4 + 8 + 16 = 38 \quad (3.35)$$

strata at level $c = 4$. Hence, it is easy to conjecture that problem (3.33) has an infinity of non-isomorphic solutions! Nevertheless, we do not know which are the paths (excepting the upper one) that are extendable to infinity.

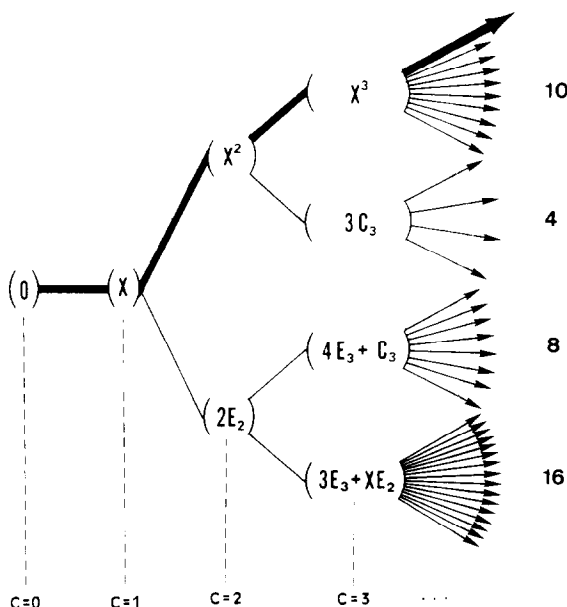


FIGURE 8

Remark 3.4. Going down to the level of the *underlying generating series*, Theorem B gives also rise to an efficient algorithm for the calculation of the solution of the general (normalized and compatible) differential problem (1.1), (1.2) in the “numerical” context of multidimensional formal power series. Here, the rooted tree \mathcal{S} “degenerates” into a single infinite chain since we always have both existence and unicity of the solution. The algorithm thus obtained is in fact a slightly modified version of the classical Newton–Raphson iterative scheme (the “tangent” method) adapted to the context of power series. Let us briefly illustrate this statement in the unisorted case: Let $f(x, y) \in \mathbb{K}[[x, y]]$ and let $\alpha = \alpha(x) \in \mathbb{K}[[x]]$ be an approximation to the solution $a = a(x) \in \mathbb{K}[[x]]$ of the differential problem $y' = f(x, y)$, $y(0) = 0$. The classical Newton–Raphson scheme, applied to the equation $\Phi(y) = 0$, where $\Phi(y) = y' - f(x, y)$, replaces the approximation $\alpha = \alpha(x)$ by a better one $\alpha^+ = \alpha^+(x)$ given by the formula

$$\alpha^+ = \alpha - [\Phi'(\alpha)]^{-1}(\Phi(\alpha)) \quad (3.36)$$

where the derivative $\Phi'(\alpha)$ is explicitly given by

$$\Phi'(\alpha)(h) = h' - \frac{\partial f}{\partial y}(x, \alpha) h. \quad (3.37)$$

As for the numerical algorithm following from Theorem B, it can be written in the form

$$\alpha^+ = \alpha + \beta, \quad \beta' = f(x, \alpha) - \alpha' + \frac{\partial f}{\partial y}(x, \alpha) \beta \quad (3.38)$$

and corresponds to the slightly different formula

$$\Phi'(\alpha)(\alpha^+) = \Phi'(\alpha)(\alpha) - \Phi(\alpha) \quad (3.39)$$

which can also be obtained, incidentally, by an application of $\Phi'(\alpha)$ to the two sides of (3.36). This has the advantage of avoiding a preliminary calculation of the inverse $[\Phi'(\alpha)]^{-1}$ of the linear operator described by (3.37). Thus, Theorem B (see also [8]) appears as a kind of „lifting” of the multidimensional Newton–Raphson iterative scheme to the context of differential equations where the unknowns (and the equations themselves) are combinatorial species (not “mere” power series,...).

To conclude, let us mention some directions for future exploration:

1. It would be interesting to obtain appreciable extensions of Table I and of tables contained in [3 and 22], in both the unidimensional and multidimensional context. Symbolic manipulation systems such as CAYLEY [4, 5], DARWIN [1], and MACSYMA [28] should be very

useful in this respect, helping, for instance, to make efficient calculations within the differential ring \mathcal{R} of virtual species.

2. In the context of recent studies connecting combinatorics to algebraic topology, Joyal [15, 16] has defined, as we said above, remarkable substitution and integration (2.96) formulas in the context of the ring \mathcal{R} . This permits the reformulation of any general differential problem of the form (2.2), (2.3) to the context of virtual species. The iterative schemes developed in the present paper can still be applied. Thus, because of (2.88), (2.89), and (3.15)–(3.18), the problem of solving (2.2), (2.3) in \mathcal{R} can be reduced to searching “virtual integrals” of simpler equations of the form

$$\partial Y_i / \partial X_j = H_{ij}(X_1, \dots, X_k) \quad (3.40)$$

$1 \leq i \leq p$, $1 \leq j \leq k$. Because of the integration formula (2.96), we thus see that, in the *unsorted case* ($k = 1$), every differential problem of the form

$$Y'_i = F_i(X, Y_1, \dots, Y_p), \quad Y_i[\phi] = \phi, \quad F_i \in \mathcal{R}, \quad 1 \leq i \leq p, \quad (3.41)$$

(having only one independent variable) is solvable in \mathcal{R} and thus always have, in fact an infinity of non-isomorphic solutions of the form $A = A(X) = B(X) - C(X)$, where B and C are ordinary (i.e., non-virtual) species. Do there exist extensions of (2.96) to the *multisorted case*? How can we adequately “classify” these solutions? Which conditions can we add to general differential problems to obtain *unicity* of solutions in \mathcal{R} ?

ACKNOWLEDGMENTS

The author wishes to thank Robert V. Anderson, François Bergeron, Pierre Bouchard, Hélène Décoste, Jacques Hurtubise, André Joyal, Jacques Labelle, Pierre Leroux, Gian-Carlo Rota, Stephen Schanuel, Volker Strehl and Yeong Nan Yeh for their encouragements and suggestions. Table I has been made with the help of Hélène Décoste.

REFERENCES

1. F. BERGERON AND C. PICHET, DARWIN: A Discovery System for the Study of Combinatorics,” Univ. du Québec à Montréal, Canada, 1984.
2. F. BERGERON, Une systématique de la combinatoire algébrique,” Ph.D. thesis, Univ. du Québec à Montréal, Canada, 1986.
3. G. BUTLER AND J. MACKAY, The transitive groups of degree up to eleven, *Comm. Algebra* **11**, No. 8, (1983), 863–911.
4. J. J. CANNON, An introduction to the group theory language CAYLEY, in “Proc. L.M.S. Symp. Computational Group Theory, Durham, Aug. 1982” (M. Atkinson, Ed.) Academic Press, New York, 1984.

5. J. J. CANNON, "A Language for Group Theory," Lecture Notes in Computer Science, Springer-Verlag, Berlin, to appear.
6. R. D. CARMICHAEL, Introduction to the Theory of Groups of Finite Order," Dover, New York, 1956.
7. C. B. COLLINS, I. P. GOULDEN, D. M. JACKSON, AND O. M. NIERSTRASZ, A combinatorial application of matrix Riccati equations and their q -analogues, *Discrete Math.* **36** (1981), 139–153.
8. H. DÉCOSTE, G. LABELLE, AND P. LEROUX, Une approche combinatoire pour l'itération de Newton-Raphson, *Adv. in Appl. Math.* **3** (1982), 407–416.
9. D. DUMONT, Une approche combinatoire des fonctions elliptiques de Jacobi, *Adv. in Math.* **41**, (1981), 1–39.
10. I. P. GOULDEN, D. M. JACKSON, AND J. W. REILLY, The Hammond series of a symmetric function and its applications to P -recursiveness, *SIAM J. Algebraic Discrete Methods* **4**, No. 2, (1983), 179–193.
11. W. GRÖBNER, "Die Lie-Reihen und ihre Anwendungen," Deut. Verlag Wissenschaften, Berlin, 1967.
12. W. GRÖBNER AND H. KNAPP, "Contributions to the Theory of Lie Series," Bibliographisches Institut-Mannheim, Hochschultaschenbücher-Verlag, Mannheim, 1967.
13. F. HARARY, "Graph Theory," Addison-Wesley, Reading, Mass., 1968.
14. A. JOYAL, Une théorie combinatoire des séries formelles, *Adv. in Math.* **42** (1981), 1–82.
15. A. JOYAL, Règle des signes en algèbre combinatoire, *C. R. Math. Acad. Sci. Canada*, Vol. VII, No. 5, Oct. 1985, 285–290.
16. A. JOYAL, Calcul intégral combinatoire et homologie du groupe symétrique, *C. R. Math. Acad. Sci. Canada*, Vol. VII, No. 6, Dec. 1985, 337–342.
17. G. LABELLE, Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange, *Adv. in Math.* **42** (1981), 217–247.
18. G. LABELLE, Éclotions combinatoires appliquées à l'inversion multidimensionnelle des séries formelles, *J. Combin. Theory Ser. A*, Vol. 39, No. 1, (1985), 52–82.
19. G. LABELLE, Une combinatoire sous-jacente au théorème des fonctions implicites, *J. Combin. Theory Ser. A*, (in press).
20. J. LABELLE, "Théorie des graphes," Modulo Éditeur, Outremont, Québec, Canada, 1981.
21. J. LABELLE, Applications diverses de la théorie combinatoire des espèces de structures, *Ann. Sci. Math. Québec*, **7**, No. 1 (1983), 58–94.
22. J. LABELLE, Quelques espèces sur les ensembles de petites cardinalités, *Ann. Sci. Math. Québec*, **9**, No. 1, (1985), 31–58.
23. P. LEROUX, AND G. VIENNOT, Combinatorial resolution of systems of differential equations, I. Ordinary differential equations, in "Colloque de combinatoire énumérative," Univ. du Québec à Montréal, 1985 (Proceedings to appear in the Springer Lecture Notes in Math.).
24. A. LONGTIN, Une combinatoire non-commutative pour l'étude des permutations alternantes et des fonctions trigonométriques généralisées, lecture, in "Séminaire de Combinatoire," Univ. du Québec à Montréal, May 1984.
25. A. C. LUNN, AND J. K. SENIOR, Isomerism and Configurations, *J. Phys. Chem.* **33** (1929), 1026–1079.
26. S. MAC LANE, Categories for the Working Mathematician," *Graduate Texts in Math.*, Vol. 5, Springer-Verlag, New York, 1971.
27. A. NIJENHUIS AND H. S. WILF, "Combinatorial Algorithms," Comput. Sci. Appl. Math., Academic Press, New York 1975.
28. R. H. RAND, "Computer Algebra in Applied Mathematics: An Introduction to MAC-SYMA," Research Notes in Math. Vol. 94, Pitman, New York, 1984.
29. S. SCHANUEL, Yeh's proof of the factoriality of Joyal's ring of species, lecture, in "Séminaire de Combinatoire," Univ. du Québec à Montréal, April 1984.

30. N. J. A. SLOANE, "A Handbook of Integer Sequences," Academic Press, New York 1973.
31. R. P. STANLEY, Differentiably finite power series, *European J. Combin.* **1**, (1980), 175–188.
32. G. VIENNOT, Une interprétation combinatoire des coefficients des développements en séries entières des fonctions elliptiques de Jacobi, *J. Combin. Theory, Ser. A* **29** (1980), 121–133.
33. G. VIENNOT, Interprétations combinatoires des nombres d'Euler et de Genocchi, in "Séminaire de Théorie des Nombres," Exposé No. 11, U.E.R. de Math. et d'Info., Univ. de Bordeaux I, 1981.
34. Y. N. YEH, "On the Combinatorial Species of Joyal," Ph.D. thesis, State Univ. of New York at Buffalo, 1985.