ANNALES SCIENTIFIQUES de L'ÉCOLE NORMALE SUPÉRIEURE

Frédéric FAUVET & Frédéric MENOUS

Ecalle's arborification-coarborification transforms and Connes-Kreimer Hopf algebra

Annales Scientifiques de l'École Normale Supérieure

Publiées avec le concours du Centre National de la Recherche Scientifique

ECALLE'S ARBORIFICATION-COARBORIFICATION TRANSFORMS AND CONNES-KREIMER HOPF ALGEBRA

BY Frédéric FAUVET AND Frédéric MENOUS

ABSTRACT. – We give a natural and complete description of Ecalle's mould-comould formalism within a Hopf-algebraic framework. The arborification transform thus appears as a factorization of characters, involving the shuffle or quasishuffle Hopf algebras, thanks to a universal property satisfied by Connes-Kreimer Hopf algebra. We give a straightforward characterization of the fundamental process of homogeneous coarborification, using the explicit duality between decorated Connes-Kreimer and Grossman-Larson Hopf algebras. Finally, we introduce a new Hopf algebra that systematically underlies the calculations for the normalization of local dynamical systems.

RÉSUMÉ. – Nous donnons une description complète et naturelle du formalisme d'arborification/coarborification d'Ecalle en termes d'algèbres de Hopf. L'arborification apparaît alors comme une factorisation de caractères, impliquant les algèbres shuffle ou quasishuffle, en vertu d'une propriété universelle satisfaite par l'algèbre de Connes-Kreimer. Dans ce cadre, nous obtenons de façon directe le procédé fondamental de coarborification homogène, en utilisant la dualité explicite entre les algèbres de Hopf décorées de Connes-Kreimer et Grossman-Larson. Enfin, nous introduisons une nouvelle algèbre de Hopf qui est sous-jacente aux calculs de normalisation des systèmes dynamiques locaux.

1. Introduction

The local study of dynamical systems, through normalizing transformations, involves calculations in groups, or pseudogroups, of diffeomorphisms (e. g. formal, or analytic) that are tangent to identity. Other situations where these explicit calculations are required are also numerous in key questions of classification of singular geometric structures. Another source of examples is given by the so called mechanism of Birkhoff decomposition ([26]). The group G of formal tangent to Identity diffeomorphisms is the one in which most of the calculations are to be performed.

To tackle problems of this kind, Jean Ecalle has developed a powerful combinatorial environment, named *mould calculus*, that leads to formulas that are surprisingly explicit. This calculus has lately been the object of attention within the algebraic combinatorics community ([6], [7]). However, despite its striking achievements, this formalism has been

doi:

little used in local dynamics, in pending problems that anyway seem out of reach of other approaches. One reason might be that it uses a sophisticated system of notations, in which a number of infinite sums are manipulated, in a way that calls for a number of proofs and explanations, that are to a certain extent still missing in the few existing papers using mould calculus. Beside this, some constructions introduced by Ecalle, though obviously appearing as extraordinarily efficient, might remain a bit mysterious; an example of this is the so called *homogeneous coarborification* ([11]), for which we are now able to give in the present paper a very natural algebraic presentation.

In fact, Ecalle's mould-comould formalism can be very naturally recast in a Hopf-algebraic setting, with the help of a number of Hopf algebras (shuffle, quasi-shuffle, their graded duals, etc) which are now widely used within algebraic combinatorics. In the present text, we show how this can be done, which makes it possible to give simple and quick proofs of important properties regarding mould calculus.

As is now well known ([12], [4]), the Hopf-algebraic formulation of computations on formal diffeomorphisms involves the so called Faà di Bruno Hopf algebra, which encodes the eponymous formula for higher order chain rule. In fact, Hopf algebraic tools and concepts have very recently become pervasive in dynamical systems, see e.g., [22] and the references therein. Now, *an essential point* is the following: the reformulation of a classification problem through the use of Faà di Bruno Hopf algebra (or, more simply, calculations on compositions of diffeomorphisms involving the Faà di Bruno formula), although satisfactory at the formal level will usually be inefficient, in the hard cases, for the question of analyticity of the series. *Indeed, in difficult situations involving resonances and/or small denominators, the formulas obtained through Faà di Bruno are most of the time not explicit enough to obtain satisfactory growth estimates on the coefficients.*

On the other hand, Ecalle's mould-comould expansions often lead to explicit coefficients but, when trying to control the size of these in a straightforward way, we often encounter systematic divergence, which claims for the introduction of something subtler.

So the need was for some sort of *intermediate Hopf algebra*, in which the algebraic calculations would still be tractable, and leading to explicit formulas from which key estimates can be obtained, to eventually get e.g., the analyticity properties we could expect. This is exactly what arborification/coarborification does. Once the original definitions of Ecalle are translated into a Hopf-algebraic setting, with the use of Connes-Kreimer Hopf algebra CK and its graded dual, it is possible to recognize that the arborification transform is nothing else that a property of factorization of characters between Hopf algebras (we perform this at the same time for the shuffle and quasishuffle cases), using the fact that CK is an initial object for Hochschild cohomology for a particular category of cogebras ([8], [13], [14]).

Thus, the universality of the arborification mechanism is directly and naturally connected with a universal property satisfied by Connes-Kreimer Hopf algebra, whose importance is by now widely acknowledged (see e. g. [12]).

The paper is organized as follows. In the next section, we recall a few basic facts concerning normalization in local dynamics, focusing on two basic situations for which it is possible to introduce all the relevant objects in a simple, yet non trivial, context. The following section is devoted to an algebraic study of the group of tangent to identity formal diffeomorphisms, introducing at this stage the Faà di Bruno Hopf algebra $\mathscr{H}_{\mathrm{FdB}}$. This

section doesn't contain new results, yet we have chosen a presentation stressing the role of substitution automorphisms, and adopting a systematic way of looking at normalizing equations as equations on characters of Hopf algebras which are by now classical objects (basic terminology and facts on graded Hopf algebras are included).

Then we are ready to interpret moulds, at least the ones with symmetry properties that are met in practice, as characters or infinitesimal characters on some classical Hopf algebras, namely symmetral (resp. symmetrel) moulds as characters of the shuffle (resp. quasishuffle) Hopf algebra. This is the object of Section 4, where the basic notions regarding moulds, comoulds and their "contractions" are given.

In Section 5 the key dual mechanisms of arborification and coarborifications are introduced, and described through the introduction of CK and its graded dual, known to be isomorphic to the Grossman-Larson Hopf algebra

In fact, we show that the natural isomorphism between these two Hopf algebras leads directly, in the contexts of comoulds, to the process of *homogeneous coarborification*, which was put forward by Ecalle with very little explanation. A cautious handling of the symmetry factors of the trees is crucial, here.

In Section 6 we describe the Hopf algebra CK^+ which is ultimately used in practical calculations of normalizing transformations, for questions of classification of dynamical systems, involving resonances and small denominators. This solves at the same time an algebraic problem and a essential analytic one, regarding the growth estimates of the coefficients of the diffeomorphisms. The point of view which is enhanced in the present paper can be summed up in the following considerations:

- The systematic use of substitution automorphisms, which constitute an alternative – a very profitable one, because it is more *flexible* — to changes of variables, naturally entail a Hopf-algebraic presentation
- Calculations in the Faà di Bruno Hopf algebra are a direct mirror of the traditional approach through normalizing transformations, yet they don't yield results which are explicit enough to tackle difficult cases
- There is a *hierarchy* Sh/Qsh, CK, CK^+ of Hopf algebras, the first ones adapted to the simple formal classification results, the second one necessary for controlling the regularity of the formal constructions, under a strong non resonance condition, and the last one to take care of objects satisfying a weak nonresonance condition

The main results of the text are thus the ones which concern the Hopf algebra CK^+, which is the fundamental one to be used by the practitioner, in difficult problems involving small denominators.

The authors are grateful to the referees for their valuable remarks and suggestions that led to improvements of the text.

2. Normal forms

To study a dynamical system, a standard procedure, since Poincaré, is to try to conjugate the object to another one which is as simple as possible and for which the dynamics is well understood, and which is then called a normal form. The classes of objects that are considered in the present text are vector fields and diffeomorphisms in $\mathbb{C}^\nu$ and we study them near a singularity, namely a vanishing point for the field or a fixed point for the diffeomorphism.

Conjugation is thus obtained through the action of a change of coordinates, performed by a diffeomorphism leaving the singular point invariant. When we consider, say, an analytic germ of vector field X at the origin, the simplest field to which we can hope to conjugate it through an analytic change of coordinates is the linear part X^{lin} of X. When trying to do this, one immediately encounters the possibility of obstructions; indeed, if the eigenvalues of X^{lin} (supposed semi-simple) are $\lambda_1, \ldots, \lambda_\nu$ (with possible multiplicities), then even the *formal* conjugation of X to X^{lin} is not possible when some combinations of the following type do exist:

$$(1) \qquad m_1\lambda_1 + \cdots + m_\nu\lambda_\nu - \lambda_i = 0.$$

In this relation, $i \in \{1, \ldots, \nu\}$; the m_j are nonnegative integers, with $\sum m_j \geqslant 2$.

Such a relation can also be written as $\langle n, \lambda \rangle = 0$ where $\lambda = (\lambda_1, \ldots, \lambda_\nu)$ is the spectrum, n is a ν-uple of integers that belongs to the following set:

$$H = \left\{ (n_1, \ldots, n_\nu); n_i \geqslant -1, \text{at most one being} = -1, \text{and} \sum_1^\nu n_i \geqslant 1 \right\},$$

and $\langle n, \lambda \rangle = \sum_{i=1}^\nu n_i\lambda_i$.

A relation such as (1) is called a *resonance*, and we focus now on the nonresonant case, for which by definition no resonance exists. A standard way to obtain the linearization is then to conjugate with polynomial changes of coordinates our given field to its linear part, up to terms of a given valuation and then composing these transforms in order to obtain in the end only the linear part. The absence of resonance ensures that each step is possible, and the formal convergence of the infinite product is easy. In the next section, an alternative method is described, which directly lead to moulds. In this way we obtain a unique linearizing transformation, if we impose that it is tangent to identity.

Technically, it might happen that under the non resonance condition given above, some of the partial sums might vanish. If we want to avoid this, we have to consider a stronger non resonance condition, namely that the λ_i are independent over $\mathbb{Z}$. We shall below work out the algebraic formulation using the strong condition, and eventually in Section 6 we will be able to cope with the weaker one, once the appropriate Hopf algebra has been defined.

However, the formal transform will not always be convergent: in the process of computation of the linearization transform, whatever the chosen method, we encounter divisions by expressions $m_1\lambda_1 + \cdots + m_\nu\lambda_\nu - \lambda_i$, which, although non zero, might be very small. This problem of occurrence of *small denominators* calls for an extra hypothesis on the spectrum, in order to control the size of the coefficients of the series we are interested in. The original

breakthrough was made by Siegel, and the best (it is known to be optimal in dimension 2) diophantine condition so far, is Brjuno's condition:

$$\sum \frac{1}{2^k} \text{Log}(\frac{1}{\Omega(2^{k+1})}) < \infty,$$

where $\Omega(h) = \min\{|\langle n, \lambda \rangle|, \sum n_i \leqslant h\}$.

Under this condition it can be shown (Brjuno, [2]) that the normalization transform is indeed analytic.

The classification problem for germs of diffeomorphisms goes along the same lines: we consider a diffeomorphism φ at the origin of 0 in $\mathbb{C}^\nu$ and we wish to conjugate it to its linear part φ^{lin}, with $\varphi^{\text{lin}}(x) = (l_1 x_1, \ldots, l_\nu x_\nu)$. In that case, a resonance can be written as

$$l_1^{m_1} \cdots l_\nu^{m_\nu} - 1 = 0,$$

with the exponents (m_i) as above: $m = (m_1, \ldots, m_\nu)$ is in H. In the absence of resonance, such a diffeomorphism is formally conjugate to its linear part, and this can be proved by the same method as for fields.

Here also, we shall have to consider a strong non resonance condition, namely that no relation of the above type vanishes, for any family of coefficients m_i in $\mathbb{Z}$.

Under the following diophantine hypothesis, it is known ([30]) that the linearizing transform is analytic.

$$\sum \frac{1}{2^k} \text{Log}(\frac{1}{\omega(2^{k+1})}) < \infty,$$

where $\omega(h) = \min\{|l_1^{m_1} \cdots l_\nu^{m_\nu} - 1|, \sum m_i \leqslant h\}$.

In dimension one, there is a unique tangent to identity formal diffeomorphism h that conjugates a given diffeomorphism $g : x \longrightarrow \lambda x + \Sigma g_n x^n$ to its linear part g_l, provided g_l is not a periodic rotation (the non resonant case), and its coefficients are given by an explicit but already somewhat complicated recursive expression:

$$h_n = \frac{1}{\lambda^n - \lambda}\left[g_n + \sum_{p=2}^{n-1} g_p \sum_{j_1 + \cdots + j_p = n} h_{j_1} \cdots h_{j_p} \right].$$

These formulas are of little help in directly proving the most delicate analytic linearization results, already in the lowest dimension, let alone in dimension greater than 1.

As a remark, let us mention that the required calculations involving compositions of diffeomorphisms are essentially of the same type when one is interested in classifications of geometric structures with singularities. Consequently, the algebraic formalism developed below can also be used for these problems, in cases where the complexity of the problem tends to make other techniques inoperative (e.g., singular Poisson structures displaying resonances, in a context of small denominators).

Although we consider as examples the cases of non-resonant germs of vector fields or germs of diffeomorphisms, in any dimension at the origin of $\mathbb{C}$, all the algebraic structures, as well as Ecalle's constructions that come into play by following these basic situations as leading thread, are of a universal nature, as notably the Hopf algebra of Section 6.

3. Algebraic structures on the group G of tangent to identity diffeomorphisms

3.1. The Lie algebra g of formal vector fields

We consider now the group G of formal diffeomorphisms that are tangent to Identity at the origin of $\mathbb{C}^\nu$:

$$G = \{\varphi = (\varphi_1, \ldots, \varphi_\nu) : x = (x_1, \ldots, x_\nu) \longmapsto x + \text{h.o.t.}\}.$$

It is well-known that the group G is the Lie group of the Lie algebra g of formal vector fields:

$$g = \left\{ X = \sum_{i=1}^{\nu} X_i(x)\partial_{x_i}, \ X_i(x) \in \mathbb{C}_{\geqslant 2}[[x]] \right\},$$

where $\mathbb{C}_{\geqslant 2}[[x]]$ denotes formal power series in the variables $x = (x_1, \ldots, x_\nu)$ of total valuation greater than 1. Even if we deal with formal power series, the "geometric" interpretation goes as follows: for a given vector field X, consider the differential system:

$$\begin{cases} y_1'(t) = X_1(y_1(t), \ldots, y_\nu(t)) \\ \qquad \vdots \\ y_\nu'(t) = X_\nu(y_1(t), \ldots, y_\nu(t)), \end{cases}$$

with the initial conditions $y(0) = (y_1(0), \ldots, y_\nu(0)) = (x_1, \ldots, x_\nu) = x$. Even formally, the solution at time t is given by $y(t) = \varphi^t(x)$ where $\varphi^t \in G$ and $\varphi^t \circ \varphi^s = \varphi^{t+s}$. Namely, φ^t is the flow of the vector field X, whose exponential is simply $\exp(X) = \varphi^1$.

This correspondence is bijective ($X = \log(\varphi)$) as we shall see in the following section. Note that the computations are not so easy to handle but become clear, once diffeomorphisms are interpreted through their action on formal power series.

3.2. The action of G and substitution automorphisms

From the definition of g it is easy to derive its action on a formal power series f, by the chain-rule formula: $(f(x(t)))' = (X.f)(x(t))$. If $X = \sum_{i=1}^\nu X_i(x)\partial_{x_i}$,

$$X.f = \sum_{i=1}^{\nu} X_i(x)\partial_{x_i} f,$$

as a vector field is a differential operator. Moreover, it is a derivation on $\mathbb{C}[[x]]$ since

$$X.(fg) = (X.f)g + f(X.g).$$

Similarly the natural action of a diffeomorphism φ on a series f is given by

$$(f \lhd \varphi)(x) = (\Theta_\varphi.f)(x) = f \circ \varphi(x).$$

This defines a right action of the group G on the algebra $\mathbb{C}[[x]]$ and Θ_φ is the substitution automorphism associated to φ:

$$\Theta_\varphi.\Theta_\psi.f = (f \lhd \psi) \lhd \varphi = \Theta_{\psi \circ \varphi}.f,$$
$$\Theta_\varphi.(fg) = (fg) \lhd \varphi = (f \lhd \varphi)(g \lhd \varphi) = (\Theta_\varphi.f)(\Theta_\varphi.g).$$

Let us focus on such substitution automorphisms, since they are one of the key ingredients to perform mould calculus.

PROPOSITION 1. – *Let $\tilde{G}$ be the subset of linear endomorphisms Θ, which are continuous wrt the Krull topology of $\mathbb{C}[[x]]$ such that*

1. $\Theta(x) = \Theta.x = \varphi(x) \in G$.
2. *For any series f, g, $\Theta(fg) = \Theta.(fg) = (\Theta.f)(\Theta.g)$.*

then $\tilde{G}$ is a group (the group of substitution automorphisms) and the ("evaluation") map ev *defined by*

$$\mathrm{ev}(\Theta) = \Theta.x \in G$$

is an anti-isomorphism of groups.

Proof. – The proof is straightforward: consider a monomial $x^n = x_1^{n_1} \cdots x_\nu^{n_\nu}$. Because of the second property,

$$\Theta.(x^n) = \varphi_1^{n_1} \cdots \varphi_\nu^{n_\nu} \quad (\Theta.x = \varphi(x) = (\varphi_1(x), \ldots, \varphi_\nu(x)).$$

By linearity and continuity,

$$\Theta.f = f \circ \mathrm{ev}(\Theta)$$

and the proposition follows, noticing that $ev(\Theta_1 \Theta_2) = \varphi_2 \circ \varphi_1$ (whence the anti-isomorphism property). $\square$

In the sequel we shall identify G with $\tilde{G}$ when needed, taking advantage of the fact that such substitution automorphisms, as vector fields in g, can be seen as differential operators:

PROPOSITION 2. – *Let $\varphi = x + u(x) \in G$, then*

$$\mathrm{ev}^{-1}(\varphi) = \Theta_\varphi = \mathrm{Id}_{\mathbb{C}[[x]]} + \sum_{\substack{n_1 + \cdots + n_\nu \geqslant 1 \\ n_i \geqslant 0}} \frac{1}{n_1! \cdots n_\nu!} u_1^{n_1}(x) \cdots u_\nu^{n_\nu}(x) \partial_{x_1}^{n_1} \cdots \partial_{x_\nu}^{n_\nu}.$$

This is simply the Taylor formula:

$$\Theta_\varphi.f(x) = f(x + u(x))$$

$$= f(x) + \sum_{\substack{n_1 + \cdots + n_\nu \geqslant 1 \\ n_i \geqslant 0}} \frac{1}{n_1! \cdots n_\nu!} u_1^{n_1}(x) \cdots u_\nu^{n_\nu}(x) \partial_{x_1}^{n_1} \cdots \partial_{x_\nu}^{n_\nu} f(x).$$

There is still some work to do to perform mould calculus, but one can already use this to define explicitly the exponential of a vector field. If $X \in g$, then, for any real number t, the differential operator

$$\Theta^t = \exp(tX) = \mathrm{Id} + \sum_{s \geqslant 1} \frac{t^s}{s!} X^s$$

is a well-defined substitution automorphism and, if $\varphi^t = \mathrm{ev}(\Theta^t)$, it is the flow of the vector field X. Conversely, for a given diffeomorphism φ, if $\Theta = \mathrm{ev}^{-1}(\varphi)$ is its substitution automorphism, then φ is the flow at time $t = 1$ of the vector field

$$X = \log(\Theta) = \log(\mathrm{Id} + (\Theta - \mathrm{Id})) = \sum_{s \geqslant 1} \frac{(-1)^{s-1}}{s} (\Theta - \mathrm{Id})^s.$$

We leave the proof to the reader (see also [20]).

3.3. Degrees and homogeneous components

Vector fields and diffeomorphisms are made of power series, namely series of monomials $x^n = x_1^{n_1} \cdots x_\nu^{n_\nu}$ of degree $n = (n_1, \ldots, n_\nu)$, but what is relevant for such objects in the context of normalization is not the degree, but the notion of *homogeneous components* related to their action on monomials.

More precisely, a formal power series is given by

$$f(x) = \sum_{n \in \mathbb{N}^\nu} f_n x^n,$$

where $n = (n_1, \ldots, n_\nu) \in \mathbb{N}^\nu$ is the degree of $x^n = x_1^{n_1} \cdots x_\nu^{n_\nu}$ and $|n| = n_1 + \cdots + n_\nu$ is its total degree. Such monomial are very well adapted to the product of power series: if $f(x) = \sum_{n \in \mathbb{N}^\nu} f_n x^n$ and $g(x) = \sum_{n \in \mathbb{N}^\nu} g_n x^n$, then their product $h(x) = f(x)g(x) = \sum_{n \in \mathbb{N}^\nu} h_n x^n$ is such that

$$h_n = \sum_{k+l=n} f_k g_l,$$

but, for example, if one considers an elementary vector field $X_{i,n} = x^n \partial_{x_i}$ then

$$X_{i,n}.x^m = m_i x^{n+m-e_i},$$

where e_i is the element of $\mathbb{N}^\nu$ whose i^{th} entry (resp. j^{th} entry with $j \neq i$) is 1 (resp. 0). Regarding its action on monomials such a vector field is "homogeneous" of degree $\eta = n - e_i$ and this will be the right notion of degree for vector fields and diffeomorphisms. This suggests the following notation: for $1 \leqslant i \leqslant \nu$, let

$$H_i = \{\eta = n - e_i, \ n \in \mathbb{N}^\nu, |n| \geqslant 2\}.$$

For any $1 \leqslant i \leqslant \nu$ and $\eta \in H_i$, $|\eta| \geqslant 1$ and any vector field X in $\boldsymbol{g}$ can be written

$$X = \sum_{i=1}^\nu \sum_{\eta \in H_i} b_\eta^i x^\eta x_i \partial_{x_i} = \sum_{\eta \in H} \sum_{i ; \eta \in H_i} b_\eta^i x^\eta x_i \partial_{x_i},$$

where $H = H_1 \cup \cdots \cup H_\nu$. We will note

$$\mathbb{B}_\eta = \sum_{i ; \eta \in H_i} b_\eta^i x^\eta x_i \partial_{x_i} = \sum_i b_\eta^i x^\eta x_i \partial_{x_i},$$

assuming that the sum is restricted to the indices i such that $\eta \in H_i$. With this notation, X is decomposed in "homogeneous" components:

$$X = \sum_{\eta \in H} \mathbb{B}_\eta,$$

where, for any $n \in \mathbb{N}^\nu$ and $\eta \in H$, $\mathbb{B}_\eta.x^n$ is a monomial of degree $n + \eta$ (resp. 0) if $n + \eta \in \mathbb{N}^\nu$ (resp. $n + \eta$ is not in $\mathbb{N}^\nu$).

On the same way, a diffeomorphism $\varphi = (\varphi_1, \ldots, \varphi_\nu)$ is given by ν series

$$\varphi_i(x) = x_i \left(1 + \sum_{\eta \in H_i} \varphi_\eta^i x^\eta \right)$$

and if $\bar{H}$ is the additive semigroup generated by H, then its associated substitution automorphism can be decomposed in homogeneous components

$$\Theta = \mathrm{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta,$$

where

$$\mathbb{D}_\eta = \sum_{s \geqslant 1} \sum_{\substack{1 \leqslant i_1, \ldots, i_s \leqslant v \\ \eta = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{s!} \varphi_{\eta_1}^{i_1} \cdots \varphi_{\eta_s}^{i_s} x^{\eta + e_{i_1} + \cdots + e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}}$$

with finite sums (for any given η), thanks to the fact that $|\eta_k| \geqslant 1$ (thus $s \leqslant |\eta|$). This can be seen using the Taylor expansion of $f \circ \varphi(x)$ and it gives a first flavor of mould calculus:

- If $F = \exp(X)$ with $X = \sum_{\eta \in H} \mathbb{B}_\eta$, then

$$\Theta = \mathrm{Id} + \sum_{s \geqslant 1} \sum_{\eta_1, \ldots, \eta_s \in H} \frac{1}{s!} \mathbb{B}_{\eta_s} \cdots \mathbb{B}_{\eta_1}.$$

- If $X = \log(\varphi)$ with $\Theta_\varphi = \mathrm{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta$, then

$$X = \sum_{s \geqslant 1} \sum_{\eta_1, \ldots, \eta_s \in \bar{H}} \frac{(-1)^{s-1}}{s} \mathbb{D}_{\eta_s} \cdots \mathbb{D}_{\eta_1}.$$

These are in fact two examples of mould-comould expansions. We postpone now the definition and study of mould expansions that will be very useful to deal with linearization equations, namely when we conjugate a given dynamical system to its linear part. But such decompositions can also be used to get the Faà di Bruno formulas.

PROPOSITION 3. – *Let φ and ψ in $\boldsymbol{G}$ and $\phi = \varphi \circ \psi$, then for $1 \leqslant i \leqslant v$,*

$$\phi_i(x) = x_i \left(1 + \sum_{\eta \in H_i} \phi_\eta^i x^\eta \right),$$

with, for $\eta \in H_i$,

$$\phi_\eta^i = \varphi_\eta^i + \psi_\eta^i + \sum_{s \geqslant 2} \sum_{\substack{i_1 = i \\ 1 \leqslant i_2, \ldots, i_s \leqslant v \\ \eta = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{(s-1)!} P_{i_2, \ldots, i_s}^{\eta_1 + e_i} \varphi_{\eta_1}^i \psi_{\eta_2}^{i_2} \cdots \psi_{\eta_s}^{i_s},$$

where $P_{i_2, \ldots, i_s}^{\eta_1 + e_i}$ are integers, independant of φ and ψ.

Proof. – Let φ and ψ in $\boldsymbol{G}$ and $\phi = \varphi \circ \psi$. We can write

$$\Theta_\varphi = \mathrm{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta,$$
$$\Theta_\psi = \mathrm{Id} + \sum_{\eta \in \bar{H}} \mathbb{E}_\eta,$$

where

$$\mathbb{D}_\eta = \sum_{s \geqslant 1} \sum_{\substack{1 \leqslant i_1, \ldots, i_s \leqslant v \\ \eta = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{s!} \varphi_{\eta_1}^{i_1} \cdots \varphi_{\eta_s}^{i_s} x^{\eta + e_{i_1} + \cdots + e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}},$$

$$\mathbb{E}_\eta = \sum_{s \geqslant 1} \sum_{\substack{1 \leqslant i_1, \ldots, i_s \leqslant v \\ \eta = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{s!} \psi_{\eta_1}^{i_1} \cdots \psi_{\eta_s}^{i_s} x^{\eta + e_{i_1} + \cdots + e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}},$$

and $\phi_i(x) = \Theta_\phi.x_i = \Theta_{\varphi \circ \psi}.x_i = \Theta_\psi.\Theta_\varphi.x_i$. We get

$$\Theta_\psi.\Theta_\varphi = \mathrm{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta + \sum_{\eta \in \bar{H}} \mathbb{E}_\eta + \sum_{\eta, \mu \in \bar{H}} \mathbb{E}_\mu \mathbb{D}_\eta.$$

Now, $\mathrm{Id}.x_i = x_i$, $\mathbb{D}_\eta.x_i = \varphi_\eta^i x^\eta x_i$ (resp. 0) if $\eta \in H_i$ (resp. $\eta \notin H_i$) and $\mathbb{E}_\eta.x_i = \psi_\eta^i x^\eta x_i$ (resp. 0) if $\eta \in H_i$ (resp. $\eta \notin H_i$) so it remains to compute $\mathbb{E}_\mu \mathbb{D}_\eta.x_i$. This is zero as soon as $\eta \notin H_i$ and otherwise:

$$\mathbb{E}_\mu \mathbb{D}_\eta.x_i = \qquad\qquad \mathbb{E}_\mu.(\varphi_\eta^i x^{\eta+e_i})$$
$$= \qquad\qquad \varphi_\eta^i \mathbb{E}_\mu.x^{\eta+e_i}$$
$$= \varphi_\eta^i \left(\sum_{s \geqslant 1} \sum_{\substack{1 \leqslant i_1, \ldots, i_s \leqslant \nu \\ \mu = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{s!} \psi_{\eta_1}^{i_1} \cdots \psi_{\eta_s}^{i_s} P_{i_1, \ldots, i_s}^{\eta+e_i} \right) x^{\eta+\mu+e_i},$$

where

$$P_{i_1, \ldots, i_s}^{\eta+e_i} = x^{-\eta-e_i+e_{i_1}+\cdots+e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}} x^{\eta+e_i} \in \mathbb{N}.$$

This gives the announced formula. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We already have in this section the key ingredients to do mould calculus: with the help of these results, we will see that computing a conjugating map will amount to the computation of a character in a quite simple Hopf algebra (shuffle, quasishuffle or, after arborification-coarborification, in the Connes-Kreimer Hopf algebra). But this Hopf algebraic structure is already present when dealing directly with the coefficients of a diffeomorphism and gives rise to the Faà di Bruno Hopf algebra.

3.4. From G to the Faà di Bruno Hopf algebra

3.4.1. *A short reminder on Hopf algebras.* – In the sequel we will deal with graded connected Hopf algebras $\mathscr{H}$. This means first that $\mathscr{H}$ is a graded vector space over $\mathbb{C}$

$$\mathscr{H} = \bigoplus_{n \geqslant 0} \mathscr{H}_n,$$

where $\mathscr{H}_0 \approx \mathbb{C}$. In order to get a graded Hopf algebra, $\mathscr{H}$ has to be a graded algebra with

1. a unit $\eta : \mathbb{C} \to \mathscr{H}_0$,
2. a product $\pi : \mathscr{H} \otimes \mathscr{H} \to \mathscr{H}$,

with the usual commutative diagrams that respectively expresses the unit and associativity properties (see [23]) and such that $\pi(\mathscr{H}_n \otimes \mathscr{H}_m) \subseteq \mathscr{H}_{n+m}$. It also has to be a coalgebra with

1. a counit $\varepsilon : \mathscr{H} \to \mathbb{C}$,
2. a coproduct $\Delta : \mathscr{H} \to \mathscr{H} \otimes \mathscr{H}$,

with the corresponding commutative diagrams for the counit and coassociativity properties respectively ([23]) and such that $\Delta(\mathscr{H}_n) \subseteq \bigoplus_{0 \leqslant k \leqslant n} \mathscr{H}_k \otimes \mathscr{H}_{n-k}$.

With the compatibility relations between the algebra structure and the coalgebra structure, $\mathscr{H}$ becomes a bialgebra and the graded structure (with $\mathscr{H}_0 \approx \mathbb{C}$) ensures that this is a Hopf algebra: there exists an antipode, namely a linear map $S : \mathscr{H} \to \mathscr{H}$ such that:

$$\pi \circ (\mathrm{id} \otimes S) \circ \Delta = \pi \circ (S \otimes \mathrm{id}) \circ \Delta = \eta \circ \varepsilon.$$

Once such a Hopf algebra is given, it induces an algebra structure on $\mathcal{L}(\mathcal{H}, \mathbb{C})$. If u and v are two linear forms, their convolution product is given by

$$u * v = \pi_{\mathbb{C}} \circ (u \otimes v) \circ \Delta,$$

where $\pi_{\mathbb{C}}$ is the usual product on $\mathbb{C}$. Let us remember that among such morphisms, one can distinguish

1. The *characters* (algebra morphisms) that form a group $\mathcal{C}(\mathcal{H})$ for the convolution, with unit ε and the inverse of a character χ is given by $\chi \circ S$.
2. The *infinitesimal characters*, that are the linear morphisms u vanishing on $\mathcal{H}_0$ and such that

$$u \circ \pi = \pi_{\mathbb{C}} \circ (u \otimes \varepsilon + \varepsilon \otimes u).$$

They form a Lie algebra $c(\mathcal{H})$ for the Lie bracket $[u, v] = u * v - v * u$.

As for vector fields and diffeomorphisms $\mathcal{C}(\mathcal{H})$ behaves as the Lie group of the Lie algebra $c(\mathcal{H})$ with the log and exp maps

$$\begin{aligned}
\exp_*(u) &= \varepsilon + \sum_{s \geqslant 1} \tfrac{1}{s!} u^{*s}, \\
\log_*(\chi) &= \sum_{s \geqslant 1} \tfrac{(-1)^{s-1}}{s} (\chi - \varepsilon)^{*s}.
\end{aligned}$$

It is in fact a proalgebraic group, namely an inverse limit of linear algebraic groups, and the exp and log are computed as graded operators at the level of the homogeneous components([9]).

We shall soon see that vector fields and diffeomorphisms can be identified to infinitesimal characters and characters on a Hopf algebra, namely the Faà di Bruno Hopf algebra. But let us first give a concrete example of graded connected Hopf algebra related to power series.

The coalgebra of coordinates of power series can be defined as follows, for $n \in \mathbb{N}^\nu$, let us consider the functional:

$$\begin{aligned}
\alpha_n : \quad & \mathbb{C}[[x]] \quad \rightarrow \quad \mathbb{C} \\
& f(x) = \sum_{n \in \mathbb{N}} f_n x^n \mapsto \alpha_n(f) = f_n.
\end{aligned}$$

The graded vector space $C = \bigoplus_{k \geqslant 0} C_k$, where $C_k = \text{Vect}_{\mathbb{C}}\{\alpha_n, |n| = k\}$ is a graded cocommutative coalgebra for the coproduct induced by the product of series:

$$\Delta \alpha_n = \sum_{k+l=n} \alpha_k \otimes \alpha_l \; (\alpha_n(f.g) = \pi_{\mathbb{C}} \circ (\Delta \alpha_n)(f \otimes g))$$

and the counit is given by $\varepsilon(\alpha_0) = 1$ and $\varepsilon(\alpha_n) = 0$ if $|n| \geqslant 1$. Thanks to this coalgebra structure, the space $\mathcal{L}(C, \mathbb{C})$ is a convolution algebra which is trivially isomorphic to $\mathbb{C}[[x]]$. In order to define a Hopf algebra, let us consider the free commutative algebra generated by $\{\alpha_n, |n| \geqslant 1\}$. By adding a unit $\mathbf{1}$ (which identifies to α_0 in the previous coproduct) and extending the gradation and the previous coproduct to the product of functionals, one gets a Hopf algebra $\mathcal{H}$ whose group of characters can be clearly identified to the group of invertible series:

$$G^{\text{inv}} = \left\{ 1 + \sum_{|n| \geqslant 1} f_n x^n, \; f_n \in \mathbb{C} \right\}.$$

The same idea will govern the contruction of the Faà di Bruno Hopf algebra of coordinates on the group $\boldsymbol{G}$.

3.4.2. *The Faà di Bruno Hopf algebra.* – The group $\boldsymbol{G}$ is associated to a graded connected algebra. Let us first remind that

$$\boldsymbol{G} = \{\varphi(x) = x + u(x), u \in (\mathbb{C}_{\geqslant 2}[[x]])^\nu\},$$

where $x = (x_1, \ldots, x_\nu)$ and $u(x) = (u_1(x), \ldots, u_\nu(x))$ and we can note

$$\varphi(x) = (\varphi_i(x))_{1 \leqslant i \leqslant \nu} = \left(x_i \left(1 + \sum_{\eta \in H_i} \varphi_\eta^i x^\eta \right) \right)_{1 \leqslant i \leqslant \nu}.$$

A diffeomorphism in $\boldsymbol{G}$ is then given by its coefficients φ_η^i, where $i \in \{1, \ldots, \nu\}$ and $\eta \in H_i$, and for any such couple (i, η), one can define functionals on $\boldsymbol{G}$:

$$C_\eta^i : \boldsymbol{G} \to \mathbb{C}$$
$$\varphi \mapsto \varphi_\eta^i.$$

Following the same ideas as for G^{inv}, the Faà di Bruno algebra is the free commutative algebra generated by the functionals C_η^i :

$$\mathscr{H}_{\mathrm{FdB}} = \mathbb{C}[(C_\eta^i)_{i \in \{1, \ldots, \nu\}, \eta \in H_i}].$$

Identifying $\mathbb{C} \subset \mathscr{H}_{\mathrm{FdB}}$ to $\mathbb{C}.1$, where 1 is the functional defined on $\boldsymbol{G}$ by $1(\varphi)=1$, it is clear that $\mathscr{H}_{\mathrm{FdB}}$ acts on $\boldsymbol{G}$, if $P(\ldots, C_\eta^i, \ldots)$ is a polynomial in $\mathscr{H}_{\mathrm{FdB}}$, then

$$P(\ldots, C_\eta^i, \ldots)(\varphi) = P(\ldots, \varphi_\eta^i, \ldots).$$

If we define a gradation by $\mathrm{gr}(1) = 0$ and

$$\mathrm{gr}(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = |\eta_1| + \cdots + |\eta_s|,$$

then $\mathscr{H}_{\mathrm{FdB}}$ is a graded connected commutative algebra. Now, using the Faà di Bruno Formulas 3, it is not difficult to define a coproduct on this algebra by the relation:

$$C_\eta^i(\varphi \circ \psi) = \pi_{\mathbb{C}} \circ (\Delta C_\eta^i)(\varphi \otimes \psi)$$

extended to $\mathbb{C}[(C_\eta^i)_{i \in \{1, \ldots, \nu\}, \eta \in H_i}]$. With this coproduct, $\mathscr{H}_{\mathrm{FdB}}$ is a graded connected Hopf algebra.

Now, any diffeomorphism φ can be identified to the character (also noted φ) on $\mathscr{H}_{\mathrm{FdB}}$ defined by $\varphi(C_\eta^i) = C_\eta^i(\varphi)$ so that $\boldsymbol{G}$ is clearly isomorphic to $\mathcal{C}(\mathscr{H}_{\mathrm{FdB}})$ and, on the same way, the Lie algebra $\boldsymbol{g}$ is isomorphic to $\boldsymbol{c}(\mathscr{H}_{\mathrm{FdB}})$ (taking into account the due order reversal in the formulas) . Note that the log‑exp correspondence between $\boldsymbol{G}$ and $\boldsymbol{g}$ is exactly the $\log_*$‑$\exp_*$ correspondence between $\mathcal{C}(\mathscr{H}_{\mathrm{FdB}})$ and $\boldsymbol{c}(\mathscr{H}_{\mathrm{FdB}})$ and the action of $\boldsymbol{g}$ and $\boldsymbol{G}$ on $\mathbb{C}[[x]]$ corresponds to the coaction $\Phi : C \to C \otimes \mathscr{H}_{\mathrm{FdB}}$ defined by

$$\alpha_n(f \circ \varphi) = \pi_{\mathbb{C}} \circ \Phi(\alpha_n)(f \otimes \varphi),$$

which is such that C is a $\mathscr{H}_{\mathrm{FdB}}$-comodule coalgebra (cf [12]).

This algebraic work on diffeomorphisms and vector fields may not seem to help at the present moment, but it will be useful in the sequel and one can already notice that *linearization equations correspond to equations for characters on* $\mathscr{H}_{\mathrm{FdB}}$.

3.5. Characters and conjugacy equations

3.5.1. *Vector fields.* – Consider a vector field:

$$X(x) = \sum_{i=1}^{\nu} X_i(x) \partial_{x_i}$$

such that

$$X_i(x) = \lambda_i x_i + x_i \sum_{\eta \in H} a_\eta^i x^\eta = \lambda_i x_i + P_i(x) \ (\lambda_i \in \mathbb{C}).$$

This vector field can be seen as a perturbation of it linear part $X^{\mathrm{lin}} = \sum \lambda_i x_i \partial_{x_i}$:

$$X = X^{\mathrm{lin}} + P$$

and one would like to know if, through some formal or analytic change of coordinates $y = \varphi(x)$ $(x = \psi(y))$, the vector field can be conjugated to its linear part:

$$\varphi^*(X) = X^{\mathrm{lin}} \text{ or } \psi^*(X^{\mathrm{lin}}) = X = X^{\mathrm{lin}} + P.$$

If this is the case, the latter equation reads, for $1 \leqslant i \leqslant \nu$,

$$X^{\mathrm{lin}}.\psi_i = X_i \circ \psi = \lambda_i \psi_i + P_i \circ \psi(x),$$

so

$$X^{\mathrm{lin}}.\psi_i - \lambda_i \psi_i = P_i \circ \psi(x).$$

On one hand, if $\psi_i(x) = x_i \left(1 + \sum_{\eta \in H} b_\eta^i x^\eta\right)$, then

$$X^{\mathrm{lin}}.\psi_i - \lambda_i \psi_i = x_i \sum_{\eta \in H} \langle \lambda, \eta \rangle b_\eta^i x^\eta,$$

where, for $\eta = (n_1, \ldots, n_\nu) \in H$, $\langle \lambda, \eta \rangle = \lambda_1 n_1 + \cdots + \lambda_\nu n_\nu$. From the Hopf algebra point of view, let ∇ the derivation on $\mathscr{H}_{\mathrm{FdB}}$ defined by $\nabla C_\eta^i = \langle \lambda, \eta \rangle C_\eta^i$ ($\nabla 1 = 0$), then, if χ is the character associated to ψ, we have

$$X^{\mathrm{lin}}.\psi_i - \lambda_i \psi_i = \sum_{\eta \in H} \chi \circ \nabla(C_\eta^i) x^\eta.$$

On the other hand, If u is the infinitesimal character on $\mathscr{H}_{\mathrm{FdB}}$ defined by $u(C_\eta^i) = a_\eta^i$ ($u(1) = 0$), then the conjugacy equation reads

$$\forall i, \eta, \ \chi \circ \nabla(C_\eta^i) = (u * \chi)(C_\eta^i)$$

and, thanks to the fact that ∇ is a derivation and u infinitesimal, the conjugacy equation reads, on $\mathscr{H}_{\mathrm{FdB}}$,

$$\chi \circ \nabla = u * \chi.$$

Of course, φ is given by the inverse $\chi \circ S$ of χ in $\mathcal{C}(\mathscr{H}_{\mathrm{FdB}}, \mathbb{C})$.

In other words, one can associate to a vector field $X = X^{\mathrm{lin}} + P$ an infinitesimal character u, and this vector field is formally conjugated to X^{lin} if and only if there exists a character χ such that the above equation holds. Moreover we have the very classical ([1]) result:

PROPOSITION 4. – *If, for all $\eta \in H$, $\langle \lambda, \eta \rangle \neq 0$, X is formally conjugated to X^{lin}.*

Proof. – The proof is recursive on the gradation of $\mathscr{H}_{\mathrm{FdB}}$: let $\chi = \sum_{n \geqslant 0} \chi_n$, where χ_n is the restriction to the nth component of $\mathscr{H}_{\mathrm{FdB}}$. Note that, necessarily, $\chi_0 = \varepsilon$. Let us suppose that, for a given $n \geqslant 0$, $\chi_0, \ldots, \chi_n$ are well-defined and such that, for any monomial $C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}$ with $\mathrm{gr}(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = |\eta_1| + \cdots + |\eta_s| = k \leqslant n$,

$$\chi(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = \chi_k(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = \chi(C_{\eta_1}^{i_1}) \cdots \chi(C_{\eta_s}^{i_s}) = \chi_{|\eta_1|}(C_{\eta_1}^{i_1}) \cdots \chi_{|\eta_s|}(C_{\eta_s}^{i_s}).$$

Thanks to the definition of u, if C_η^i is in $\mathscr{H}_{\mathrm{FdB}, n+1}$ ($|\eta| = n + 1$), then the equation reads

$$\langle \lambda, \eta \rangle \chi(C_\eta^i) = \pi \circ (u \otimes \chi)(\Delta(C_\eta^i))$$

$$= u(C_\eta^i) + \sum_{\substack{s \geqslant 2}} \sum_{\substack{i_1 = i \\ 1 \leqslant i_2, \ldots, i_s \leqslant \nu \\ \eta = \eta_1 + \cdots + \eta_s \\ \eta_k \in H_{i_k}}} \frac{1}{s - 1!} P_{i_2, \ldots, i_s}^{\eta_1 + e_i} u(C_{\eta_1}^i) \chi(C_{\eta_2}^{i_2}) \cdots \chi(C_{\eta_s}^{i_s}).$$

Since the right-hand side of this equation is recursively well-defined and $\langle \lambda, \eta \rangle$ is nonzero, $\chi(C_\eta^i)$ is uniquely determined. Now, if $C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}$ is in $\mathscr{H}_{\mathrm{FdB}, n+1}$ with $s \geqslant 2$, then, in order to get a character, one must have

$$\chi(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = \chi(C_{\eta_1}^{i_1}) \cdots \chi(C_{\eta_s}^{i_s}).$$

But, for $1 \leqslant i \leqslant s$, $|\eta_i| \leqslant n$ and one can check that

$$u * \chi(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = \sum_{t=1}^{s} \left(\prod_{r \neq t} \chi(C_{\eta_r}^{i_r}) \right) (u * \chi)(C_{\eta_t}^{i_t})$$

$$= \sum_{t=1}^{s} \left(\prod_{r \neq t} \chi(C_{\eta_r}^{i_r}) \right) \langle \lambda, \eta_t \rangle \chi(C_{\eta_t}^{i_t})$$

$$= \langle \lambda, \eta_1 + \cdots + \eta_t \rangle \chi(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}).$$

Thus, the equation determines a unique character on $\mathscr{H}_{\mathrm{FdB}}$. Note, however, that we don't obtain in this algebra anything near a closed-form solution. $\qquad\square$

This is the non-resonant case. Note that in the above character equation, one only needs to be able to compute the values $\chi(C_\eta^i)$ and assume that this is a character. Moreover, in $\mathscr{H}_{\mathrm{FdB}}$, if we have geometric estimates on the coefficients $\chi(C_\eta^i)$ then it is immediate to conclude on the analyticity of the associated diffeomorphism. But, in this setting the difficulty lies in the explicit computation of these coefficients, since the equation $\chi \circ \nabla = u * \chi$ involves the rather complex coproduct of $\mathscr{H}_{\mathrm{FdB}}$. As we are going to see next, the same work can be done for diffeomorphisms, with the same difficulty in the computation.

3.5.2. *Diffeomorphisms.* – Let $l = (l_1, \ldots, l_\nu) \in (\mathbb{C}^*)^\nu$ and f^{lin} defined by

$$f^{\mathrm{lin}}(x_1, \ldots, x_\nu) = (l_1 x_1, \ldots, l_\nu x_\nu).$$

For a given analytic diffeomorphism f in $\boldsymbol{G}$, the diffeomorphism $f^{\mathrm{lin}} \circ f$ can be seen as a perturbation of f^{lin} and one could ask if, at least formally, this map is conjugate to f^{lin}. In other words, does there exist a diffeomorphism $\varphi \in \boldsymbol{G}$ such that

$$f^{\mathrm{lin}} \circ f \circ \varphi = \varphi \circ f^{\mathrm{lin}}$$

or

$$f \circ \varphi = f^{\mathrm{lin}^{-1}} \circ \varphi \circ f^{\mathrm{lin}}.$$

Now if ξ (resp. χ) is the character associated to f (resp. φ) then the equation reads

$$\xi * \chi = \chi \circ \sigma,$$

where $\sigma(1) = 1$ and $\sigma(C_{\eta_1}^{i_1} \cdots C_{\eta_s}^{i_s}) = l^{\eta_1 + \cdots + \eta_s}$. Once again, it is well-known that if the diffeomorphism $f^{\text{lin}} \circ f$ is non-resonant, i.e.,

$$\forall \eta \in \mathrm{H}, \ l^{\eta} - 1 \neq 0,$$

then there exist a unique solution χ (or φ). Of course, the proof follows the same lines as for vector fields.

In both cases, modulo a condition on the linear part, the conjugacy equation is solvable and can be seen as an equation on characters of $\mathscr{H}_{\text{FdB}}$. Here again, it is *in principle* easy to obtain the analyticity of a diffeomorphism, with the help of the values of the character, provided that one can easily compute this character. But, because of the complexity of the convolution (coproduct in $\mathscr{H}_{\text{FdB}}$), this computation is rather difficult.

The idea of mould calculus amounts to working in much simpler Hopf algebras, whose coproduct is a deconcatenation coproduct. Then the price to pay is that:

1. One has to make some supplementary condition on the linear part, in order that all the coefficients are well defined, namely,

$$\forall \eta \in \bar{H}, \ \langle \lambda, \eta \rangle \neq 0 \text{ or } l^{\eta} - 1 \neq 0.$$

2. Analyticity becomes hidden.

We shall see below how arborification cures both plagues.

4. Mould calculus: a solution to the algebraic complexity of calculations

Moulds are just, in J. Ecalle's own words, functions of a "variable number of variables" (see Definition 2, below); they are endowed with a number of operations and symmetries which yield a very rich and structured algebraic environment and we refer Ecalles' foundational papers, notably to the seminal paper [11] for a general perspective. In the present text, however, it is in a pedestrian way that we introduce them and some of their basic constructions; in this section we shall see how the objects called moulds are ushered in a natural way for the explicit calculations involving normalizing transformations and we right away connect their symmetry properties to some classical combinatorial (Hopf) algebras, moulds being matched in this context with dual objects named comoulds. As we shall see, comould manipulations will amount to calculations with ordinary differential operators in $\mathbb{C}^n$ and the moulds involved will appear as linear forms on some well known algebras; no use of any heavy formalism nor complicated terminology will be necessary along the way.

One of the most fundamental ideas of mould calculus is to consider diffeomorphisms of $\boldsymbol{G}$ as series of "homogeneous" differential operators. Focusing on linearization problems, one starts either with

1. a vector field $X = X^{\text{lin}} + P$ where P, that belong to $\boldsymbol{g}$, can be decomposed in homogeneous components:

$$P = \sum_{\eta \in H} \mathbb{B}_{\eta},$$

2. a diffeomorphism $f^{\text{lin}} \circ f$ with f in $\mathbf{G}$, whose substitution automorphism can be written:

$$\Theta_f = \text{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta,$$

and one has to find a linearization diffeomorphism φ whose substitution automorphism can also decomposed in homogeneous components

$$\Theta_\varphi = \text{Id} + \sum_{\eta \in \bar{H}} \mathbb{F}_\eta.$$

It is then natural to try a priori to express the components $\mathbb{F}_\eta$ as (non commutative) polynomials in the original components delivered by the data of the problem. For example, in the case of vector fields:

$$\Theta_\eta = \sum_{s \geqslant 1} \sum_{\substack{\eta_1 + \cdots + \eta_s = \eta \\ \eta_i \in H}} M^{\eta_1, \dots, \eta_s} \mathbb{B}_{\eta_s} \cdots \mathbb{B}_{\eta_1}.$$

In such an expression, convenient properties of symmetry for the coefficients M will ensure that Θ_φ is a substitution automorphism. Doing so, we have already roughly defined what mould calculus is. Let us focus now on the Hopf algebras underlying this calculus.

4.1. Moulds and the concatenation algebra

In both types of conjugacy equations, one has to compute an element of $\mathbf{G}$, or, equivalently, a substitution automorphism that can be decomposed in homogeneous components. But in both case the initial object already delivers homogeneous components. It seems then reasonable to use them and their composition to compute the conjugating substitution automorphism. This suggests to look at the following concatenation algebra: consider H or $\bar{H}$ as a graded alphabet. Let $\emptyset$ be the empty word. A word will be noted

$$\boldsymbol{\eta} = (\eta_1, \dots, \eta_s).$$

The gradation can be extended to such words (with $|\emptyset| = 0$) and one can also define the length of a word

$$l(\boldsymbol{\eta}) = l((\eta_1, \dots, \eta_s)) = s \ (l(\emptyset) = 0)$$

and its weight

$$\|\boldsymbol{\eta}\| = \eta_1 + \cdots + \eta_s \in \bar{H} \ (\|\emptyset\| = 0).$$

DEFINITION 1. – *Let $\mathbf{H}$ be the set of such words (starting with H or $\bar{H}$), then the linear span of $\mathbf{H}$, noted Conc_H, is a graded unital algebra for the concatenation product:*

$$\forall \boldsymbol{\eta}^1, \boldsymbol{\eta}^2 \in \mathbf{H}, \ \pi(\boldsymbol{\eta}^1 \otimes \boldsymbol{\eta}^2) = \boldsymbol{\eta}^1 \boldsymbol{\eta}^2,$$

where $\boldsymbol{\eta} = \boldsymbol{\eta}^1 \boldsymbol{\eta}^2$ is the usual concatenation of the words $\boldsymbol{\eta}^1$ and $\boldsymbol{\eta}^2$.

Thanks to the gradation, the graded dual of Conc_H is a graded coalgebra Conc_H° whose coproduct is given, on the dual basis (identified to $\mathbf{H}$) by

$$\Delta(\boldsymbol{\eta}) = \sum_{\boldsymbol{\eta}^1 \boldsymbol{\eta}^2 = \boldsymbol{\eta}} \boldsymbol{\eta}^1 \otimes \boldsymbol{\eta}^2$$

and the vector space $\mathcal{L}(\mathrm{Conc}_H^{\circ}, \mathbb{C})$ is an algebra for the convolution product:

$$\forall u, v \in \mathcal{L}(\mathrm{Conc}_H^{\circ}, \mathbb{C}), \ u * v = \pi_{\mathbb{C}} \circ (u \otimes v) \circ \Delta.$$

In fact, we just defined here the algebra of moulds:

DEFINITION 2. – *A mould on H (or $\bar{H}$) with values in $\mathbb{C}$ is a collection $M^{\bullet} = \{M^{\eta}, \ \eta \in H\}$ of complex numbers.*

It is clear that moulds are in one-to-one correspondence with elements of $\mathcal{L}(\mathrm{Conc}_H^{\circ}, \mathbb{C})$: since H is a basis of Conc_H°, a mould represents the values of an element of $\mathcal{L}(\mathrm{Conc}_H^{\circ}, \mathbb{C})$ on the basis H: $M^{\bullet} = \chi(\bullet) = \{M^{\eta} = \chi(\eta), \ \eta \in H\}$.

The set of moulds inherits the structure of algebra and for the product, if $M^{\bullet}$ and $N^{\bullet}$ are two moulds, their product $P^{\bullet} = M^{\bullet} \times N^{\bullet}$ is given by

$$P^{\eta} = \sum_{\eta^1 \eta^2 = \eta} M^{\eta^1} N^{\eta^2},$$

that corresponds to the convolution of the associated morphisms of $\mathcal{L}(\mathrm{Conc}_H^{\circ}, \mathbb{C})$

4.2. The underlying Hopf algebras

4.2.1. *Vector fields and associated shuffle Hopf algebra.* – As we have seen before, a vector field

$$X(x) = \sum_{i=1}^{\nu} X_i(x) \partial_{x_i}$$

such that

$$X_i(x) = \lambda_i x_i + x_i \sum_{\eta \in \mathrm{H}} a_{\eta}^i x^{\eta} = \lambda_i x_i + P_i(x) \ (\lambda_i \in \mathbb{C}),$$

can be decomposed in homogeneous components

$$X = X^{\mathrm{lin}} + \sum_{\eta \in \mathrm{H}} \mathbb{B}_{\eta}$$

with

$$\mathbb{B}_{\eta} = \sum_{i=1}^{\nu} a_{\eta}^i x^{\eta} x_i \partial_{x_i}.$$

Following Ecalle's convention for the composition of operators, with the help of these components, one can associate to any word in H a differential operator in $\mathbb{C}[x, \partial_x]$ acting on $\mathbb{C}[[x]]$ by $\mathbb{B}_{\emptyset} = \mathrm{Id}_{\mathbb{C}[[x]]}$ and

$$\forall \eta = (\eta_1, \ldots, \eta_s) \in H/\{\emptyset\}, \ \mathbb{B}_{\eta} = \mathbb{B}_{\eta_s} \cdots \mathbb{B}_{\eta_1}.$$

The family $\mathbb{B}_{\bullet} = \{\mathbb{B}_{\eta}, \ \eta \in H\}$ is called a *comould* and, from a more algebraic point of view, we have the following:

PROPOSITION 5. – *The map*

$$\rho : \mathrm{Conc}_H \to \mathbb{C}[x, \partial_x]$$

$$\eta \quad \mapsto \quad \mathbb{B}_{\eta}$$

defines an antialgebra morphism (considering the usual composition of differential operators).

Note also that the action of $\mathbb{C}[x, \partial_x]$ on a product in $\mathbb{C}[[x]]$ defines a coproduct Δ : $\mathbb{C}[x, \partial_x] \to \mathbb{C}[x, \partial_x] \otimes \mathbb{C}[x, \partial_x]$ defined by

$$\forall u, v \in \mathbb{C}[[x]], \ \forall D \in \mathbb{C}[x, \partial_x], \ D.(uv) = \pi_{\mathbb{C}[[x]]} \circ \Delta(D).(u \otimes v)$$

and, as we deal here with vector fields,

$$\forall \eta \in \mathrm{H}, \ \Delta(\mathbb{B}_\eta) = \mathrm{Id}_{\mathbb{C}[[x]]} \otimes \mathbb{B}_\eta + \mathbb{B}_\eta \otimes \mathrm{Id}_{\mathbb{C}[[x]]}.$$

This can be extended to the comould and, using

1. the morphism ρ,
2. the operators L_+^η on $\boldsymbol{H}$, defined by

$$L_+^\eta((\eta_1, \ldots, \eta_s)) = (\eta, \eta_1, \ldots, \eta_s) \ (L_+^\eta(\emptyset) = (\eta))$$

and extended by linearity to Conc_H,

we get

THEOREM 1. – *With the coproduct defined on the basis $\boldsymbol{H}$ of Conc_H by $\Delta(\emptyset) = \emptyset \otimes \emptyset$ and*

$$\forall \eta \in \mathrm{H}, \ \forall \boldsymbol{\eta} \in \boldsymbol{H}, \ \Delta(L_+^\eta(\boldsymbol{\eta})) = (\mathrm{Id} \otimes L_+^\eta + L_+^\eta \otimes \mathrm{Id}) \circ \Delta(\boldsymbol{\eta}),$$

the algebra Conc_H becomes a graded, cocommutative bialgebra, and thus a Hopf algebra whose antipode is given by

$$\forall \boldsymbol{\eta} \in \boldsymbol{H}, \ S(\boldsymbol{\eta}) = (-1)^{l(\boldsymbol{\eta})} \mathrm{rev}(\boldsymbol{\eta}),$$

where $\mathrm{rev}(\emptyset) = \emptyset$ and $\mathrm{rev}((\eta_1, \ldots, \eta_s)) = (\eta_s, \ldots, \eta_1)$ otherwise. Moreover the morphism ρ turns to be a coalgebra morphism and, in this case, the comould $\mathbb{B}_\bullet$ is said to be cosymmetral.

The proof is straightforward.

Going back to the given vector field X and the conjugating equation $\varphi^*(X) = X^{\mathrm{lin}}$ becomes, in terms of substitution automorphism,

$$X.\Theta_\varphi = \Theta_\varphi.X^{\mathrm{lin}}.$$

And, since the vector field delivers a family (comould) of differential operators, it seems reasonable to look for a substitution automorphism Θ_φ that can be written as a mould expansion

$$\Theta_\varphi = \sum_{\boldsymbol{\eta} \in \boldsymbol{H}} M^{\boldsymbol{\eta}} \mathbb{B}_{\boldsymbol{\eta}},$$

where $M^\bullet = \{M^{\boldsymbol{\eta}}, \ \boldsymbol{\eta} \in \boldsymbol{H}\}$ is precisely a mould, with the relevant conditions (symmetrality) that ensure that Θ_φ is a substitution automorphism. More precisely, since Conc_H is a graded cocommutative Hopf algebra, its graded dual is a graded commutative Hopf algebra, noted Sh_H (for shuffle Hopf algebra on H) whose product (resp. coproduct) is given by the usual shuffle product (resp. deconcatenation coproduct). But if we consider the group of characters $\mathcal{C}(\mathrm{Sh}_\mathrm{H}, \mathbb{C})$ then

Theorem 2. – *The map*

$$S_\rho : \mathcal{C}(\mathrm{Sh_H}, \mathbb{C}) \to G$$
$$\chi \mapsto \mathrm{ev}\left(\sum_{\eta \in H} \chi(\eta)\rho(\eta)\right)$$

defines a morphism of groups and $\Theta^\chi = \sum_{\eta \in H} \chi(\eta)\rho(\eta)$ *is the substitution automorphism associated to* $S_\rho(\chi)$.

Note that moulds corresponding to such characters are called *symmetral* moulds.

Proof. – Thanks to gradation and homogeneity, it is clear that this defines a diffeomorphism of G. If χ is a character, then for two power series u and v,

$$\Theta^\chi.(uv) = \sum_{\eta \in H} \chi(\eta)\rho(\eta).(uv)$$

$$= \sum_{\eta \in H} \chi(\eta)\pi_{\mathbb{C}[[x]]} \circ (\Delta(\rho(\eta)).(u \otimes v)$$

$$= \pi_{\mathbb{C}[[x]]} \circ \left((\rho \otimes \rho)\left(\sum_{\eta \in H} \chi(\eta)\Delta(\eta)\right)\right).(u \otimes v)$$

$$= \sum_{\eta,\eta^1,\eta^2 \in H} \chi(\eta)\langle\Delta(\eta), \eta^1 \otimes \eta^2\rangle(\rho(\eta^1).u)(\rho(\eta^2)).v)$$

$$= \sum_{\eta,\eta^1,\eta^2 \in H} \chi(\eta)\langle\eta, \pi_{\mathrm{Sh_H}}(\eta^1 \otimes \eta^2)\rangle(\rho(\eta^1).u)(\rho(\eta^2).v)$$

$$= \sum_{\eta^1,\eta^2 \in H} \chi(\pi_{\mathrm{Sh_H}}(\eta^1 \otimes \eta^2))(\rho(\eta^1).u)(\rho(\eta^2).v)$$

$$= \sum_{\eta^1,\eta^2 \in H} (\chi(\eta^1)(\rho(\eta^1).u)\chi(\eta^2)(\rho(\eta^2).v)$$

$$= (\Theta^\chi.u)(\Theta^\chi v),$$

thus Θ^χ is a substitution automorphism. Moreover, if χ^1 and χ^2 are two characters, then

$$\Theta^{\chi^1 * \chi^2} = \Theta^{\chi^2}.\Theta^{\chi^1} = \Theta_{S_\rho(\chi^2)}.\Theta_{S_\rho(\chi^1)} = \Theta_{S_\rho(\chi^1) \circ S_\rho(\chi^2)},$$

thus

$$S_\rho(\chi^1 * \chi^2) = S_\rho(\chi^1) \circ S_\rho(\chi^2). \qquad \square$$

Note that $S_\rho(\mathcal{C}(\mathrm{Sh_H}, \mathbb{C}))$ maybe only be a subgroup of G but, in linearization equations, it is reasonable to look for the change of coordinate in this subgroup. Suppose that $\varphi^*(X) = X^{\mathrm{lin}}$ $(X.\Theta_\varphi = \Theta_\varphi.X^{\mathrm{lin}})$ where $\varphi = S_\rho(\chi)$. If u is the infinitesimal character on $\mathrm{Sh_H}$ defined by

$$u(\eta) = \begin{cases} 1 \text{ if } l(\eta) = 1 \\ 0 \text{ if } l(\eta) \neq 1, \end{cases}$$

then

$$X = X^{\mathrm{lin}} + \sum_{\eta \in H} u(\eta)\rho(\eta)$$

and

$$X.\Theta_\varphi = X^{\mathrm{lin}}.\Theta_\varphi + \sum_{\eta \in H} \chi * u(\eta)\rho(\eta) = \Theta_\varphi.X^{\mathrm{lin}}.$$

Since

$$[X^{\mathrm{lin}}, \mathbb{B}_\eta] = \langle \lambda, \eta \rangle \mathbb{B}_\eta,$$

we have

$$X^{\mathrm{lin}}.\Theta_\varphi - \Theta_\varphi.X^{\mathrm{lin}} = \sum_{\eta \in H} (\nabla\chi)(\eta)\rho(\eta),$$

where

$$(\nabla\chi)(\boldsymbol{\eta}) = \langle \lambda, \|\boldsymbol{\eta}\| \rangle \chi(\boldsymbol{\eta}),$$

so that the conjugacy equation can be turned into a character equation

$$\nabla\chi + \chi * u = 0.$$

For the inverse character ξ, corresponding to the inverse of the diffeomorphism, we get,

$$\nabla\xi = u * \xi.$$

PROPOSITION 6. – *Under the assumption that for any η in $\bar{H}$, $\langle \lambda, \eta \rangle \neq 0$, the above equation determines a unique symmetral mould (character) whose values are*

$$\xi(\eta_1, \ldots, \eta_s) = \frac{1}{\langle \lambda, \eta_1 + \cdots + \eta_s \rangle \langle \lambda, \eta_2 + \cdots + \eta_s \rangle \cdots \langle \lambda, \eta_s \rangle}$$

and its inverse is given by

$$\chi(\eta_1, \ldots, \eta_s) = \frac{(-1)^s}{\langle \lambda, \eta_1 \rangle \langle \lambda, \eta_1 + \eta_2 \rangle \cdots \langle \lambda, \eta_1 + \cdots + \eta_s \rangle}.$$

The proof is straightforward. For example, $\xi(\emptyset) = 1$ and, for $\eta \in H$ and $\boldsymbol{\eta} \in \boldsymbol{H}$, the equation reads

$$\langle \lambda, \eta + \|\boldsymbol{\eta}\| \rangle \xi(L_+^\eta(\boldsymbol{\eta})) = (u * \xi)(L_+^\eta(\boldsymbol{\eta}))) = u(\eta)\xi(\boldsymbol{\eta}) = \xi(\boldsymbol{\eta}).$$

One can check that this is a character (symmetral mould) and χ can be either computed directly or as the inverse of the character ξ. In the latter case, since ξ is a character,

$$\chi = \xi^{*-1} = \xi \circ S,$$

where the antipode S in Sh_H is given by

$$S(\eta_1, \ldots, \eta_s) = (-1)^s(\eta_s, \ldots, \eta_1).$$

Now the same can be done for the linearization of diffeomorphisms.

4.2.2. *Diffeomorphisms and the associated quasishuffle Hopf algebra.* – Once again, let $l = (l_1, \ldots, l_\nu) \in (\mathbb{C}^*)^\nu$ and f^{lin} defined by $f^{\mathrm{lin}}(x_1, \ldots, x_\nu) = (l_1 x_1, \ldots, l_\nu x_\nu)$. For a given analytic diffeomorphism f in $\boldsymbol{G}$, the diffeomorphism $f^{\mathrm{lin}} \circ f$ can be seen as a perturbation of f^{lin} and one could ask if, at least formally, this map is conjugated to f^{lin}. In other words, does there exist a diffeomorphism $\varphi \in \boldsymbol{G}$ or $\varphi \in \boldsymbol{G}_{\mathrm{ana}}$ such that

$$f^{\mathrm{lin}} \circ f \circ \varphi = \varphi \circ f^{\mathrm{lin}}.$$

If we define on $\mathbb{C}[[x]]$ the operator F^{lin} by $F^{\mathrm{lin}}.u = u \circ f^{\mathrm{lin}}$, then the equation becomes

$$F_\varphi.F_f.F^{\mathrm{lin}} = F^{\mathrm{lin}}.F_\varphi.$$

As for vector fields, the substitution automorphims F_f is a series of homogeneous differential operators:

$$F_f = \mathrm{Id}_{\mathbb{C}[[x]]} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta$$

and, as in the previous section, the map

$$
\begin{aligned}
\rho: \quad \mathrm{Conc}_{\bar{H}} \quad &\to \quad \mathbb{C}[x, \partial_x] \\
\boldsymbol{\eta} = (\eta_1, \ldots, \eta_s) &\mapsto \mathbb{D}_{\boldsymbol{\eta}} = \mathbb{D}_{\eta_s} \cdots \mathbb{D}_{\eta_1}
\end{aligned}
$$

defines an algebra antimorphism (with $\rho(\emptyset) = \mathbb{D}_\emptyset = \mathrm{Id}_{\mathbb{C}[[x]]}$). Now the main difference with vector fields is that the definition of ρ is based on homogeneous components of a substitution automorphism, for which we have:

$$\forall \eta \in \bar{H}, \ \Delta(\mathbb{D}_\eta) = \mathrm{Id}_{\mathbb{C}[[x]]} \otimes \mathbb{D}_\eta + \sum_{\eta_1 + \eta_2 = \eta} \mathbb{D}_{\eta_1} \otimes \mathbb{D}_{\eta_2} + \mathbb{D}_\eta \otimes \mathrm{Id}_{\mathbb{C}[[x]]}.$$

But if we define $\Delta(\emptyset) = \emptyset \otimes \emptyset$ and, for $\eta \in \bar{H}$,

$$\Delta((\eta)) = \emptyset \otimes (\eta) + \sum_{\eta_1 + \eta_2 = \eta} (\eta_1) \otimes (\eta_2) + (\eta) \otimes \emptyset,$$

then, extending this coproduct to $\mathrm{Conc}_{\bar{H}}$, we get

THEOREM 3. – *With this coproduct, the algebra $\mathrm{Conc}_{\bar{H}}$ is a graded, cocommutative bialgebra, and thus a Hopf algebra. Moreover the morphism ρ is a coalgebra morphism.*

The proof is quite trivial since this Hopf algebra is the graded dual of a classical quasishuffle Hopf algebra noted $\mathrm{QSh}_{\bar{H}}$ (for quasishuffle Hopf algebra on $\bar{H}$, see [19]) whose product (resp. coproduct) is given by the usual quasishuffle product (resp. deconcatenation coproduct). And, once again,

THEOREM 4. – *The map*

$$
\begin{aligned}
S_\rho: \mathcal{C}(\mathrm{QSh}_{\bar{H}}, \mathbb{C}) &\to \quad \boldsymbol{G} \\
\chi \quad &\mapsto \mathrm{ev}\left(\sum_{\boldsymbol{\eta} \in \boldsymbol{H}} \chi(\boldsymbol{\eta})\rho(\boldsymbol{\eta})\right)
\end{aligned}
$$

defines a morphism of groups and $F^\chi = \sum_{\boldsymbol{\eta} \in \boldsymbol{H}} \chi(\boldsymbol{\eta})\rho(\boldsymbol{\eta})$ is the substitution automorphism associated to $S_\rho(\chi)$.

The proof is the same as above. Once again a mould $M^\bullet = \{M^{\boldsymbol{\eta}}, \ \boldsymbol{\eta} \in \boldsymbol{H}\}$ defines a linear map from $\mathrm{QSh}_{\bar{H}}$ to $\mathbb{C}$ and this mould is

- *symmetrel* if the associated morphism is in $\mathcal{C}(\mathrm{QSh}_{\bar{H}}, \mathbb{C})$,
- *alternel* if the associated morphism is in $c(\mathrm{QSh}_{\bar{H}}, \mathbb{C})$.

Going back to the linearization equation

$$\Theta_\varphi . \Theta_f . \Theta^{\mathrm{lin}} = \Theta^{\mathrm{lin}} . \Theta_\varphi$$

with

$$\Theta_f = \mathrm{Id}_{\mathbb{C}[[x]]} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta.$$

1. $f = S_\rho(\xi)$ where ξ is the character defined by $\xi(\emptyset) = 1$ and, for $s \geqslant 1$,

$$\xi((\eta_1, \ldots, \eta_s)) = \begin{cases} 1 \text{ if } s = 1 \\ 0 \text{ if } s \geqslant 2. \end{cases}$$

2. If there exists a character χ such that

(2) $$\chi \circ \sigma = \xi * \chi \ (\sigma(\boldsymbol{\eta}) = l^{\|\boldsymbol{\eta}\|}\boldsymbol{\eta}),$$

then $\varphi = S_\rho(\chi)$ is a solution to the linearization equation. Finally, we have the following:

PROPOSITION 7. – *Under the assumption that for any η in $\bar{H}$, $l^\eta \neq 1$, the above equation determines a unique symmetral mould (character) whose values are*

$$\chi(\eta_1, \ldots, \eta_s) = \frac{1}{(l^{\eta_1 + \cdots + \eta_s} - 1)(l^{\eta_2 + \cdots + \eta_s} - 1) \cdots (l^{\eta_s} - 1)}$$

and its inverse is given by

$$\chi^{*^{-1}}(\eta_1, \ldots, \eta_s) = \frac{(-1)^s}{(l^{\eta_1} - 1) \cdots (l^{\eta_1 + \cdots + \eta_{s-1}} - 1)(l^{\eta_1 + \cdots + \eta_s} - 1)}.$$

These have been known for a long time, using mould calculus (see [11]). As for vector fields, we have $\chi(\emptyset) = 1$ and, for $\eta \in H$ and $\boldsymbol{\eta} \in \boldsymbol{H}$, the equation for the linearization character reads:

$$l^{\eta + \|\boldsymbol{\eta}\|} \chi(L_+^\eta(\boldsymbol{\eta})) = (\xi * \chi)(L_+^\eta(\boldsymbol{\eta}))) = \xi(\eta)\chi(\boldsymbol{\eta}) + \chi(L_+^\eta(\boldsymbol{\eta})) = \chi(\boldsymbol{\eta}) + \chi(L_+^\eta(\boldsymbol{\eta})).$$

One can check that this is a character (symmetrel mould) and $\chi^{*^{-1}}$ can be either computed directly or as the inverse of the character χ. In the latter case, since χ is a character,

$$\chi^{*^{-1}} = \chi \circ S$$

and the antipode in $\mathrm{QSh}_{\bar{H}}$ is also given by

$$S(\eta_1, \ldots, \eta_s) = (-1)^s \sum_{\boldsymbol{\eta} = (\boldsymbol{\eta}^1 \cdots \boldsymbol{\eta}^t)} (\|\boldsymbol{\eta}^t\|, \ldots, \|\boldsymbol{\eta}^1\|)$$

(the sum involves all the decompositions of the sequence $\boldsymbol{\eta}$ by concatenation of non-empty subsequences $\boldsymbol{\eta}^i$).

4.3. Analyticity and the need for some intermediate Hopf algebras

To sum up the previous sections, under some algebraic condition on λ or l, one can perform the linearization with the help of a formal diffeomorphism, whose substitution automorphism is given by a character χ :

$$\Theta_\varphi = \sum_{\eta \in H} \chi(\eta)\rho(\eta).$$

Under some classical diophantine condition, we shall prove below that such characters have a geometric growth, meaning that an estimate of the following type is satisfied (C being a constant):

$$|\chi(\eta)| \leqslant C^{\mathrm{gr}(\eta)},$$

so that one could hope that the associated diffeomorphism will be analytic. However, this kind of estimates are not sufficient. The reason is the following: if

$$\varphi_i(x) = x_i \left(1 + \sum_{\eta \in \mathrm{H}} a_\eta^i x^\eta \right),$$

then

$$x_i a_\eta^i x^\eta = \sum_{\|\eta\|=\eta} \chi(\eta)\rho(\eta).x_i$$

and the coefficient in $\rho(\eta).x_i$ tends to grow factorially with the length of η, an inevitable feature if we try to bound by brute force the size of the composition of r ordinary differential operators: some $r!$ factors appear. For example, in dimension 1,

$$(t^2\partial_t)^r.t = (r-1)!t^{r+1}.$$

But, on the other hand, this does not mean that the diffeomorphism is divergent: many terms contribute to a same power of x and some *compensations* may arise. Indeed this is the case and, surprisingly, this compensation phenomenon can be taken into account, using the so-called arborification-coarborification process which, algebraically, relies on the use of the Connes-Kreimer Hopf algebra, as we shall see next.

In fact, the situation can profitably be described as such:

– Direct calculations at the level of diffeomorphisms immediately translate into recursive relations in the Faà di Bruno Hopf algebra, which are difficult to solve because the coproduct in $\mathcal{H}_{\mathrm{FdB}}$ is a complicated one, a complexity that mirrors the Faà di Bruno formula for the computation of the nth coefficient of the composition of 2 formal series.

– Mould-comould expansions, on the contrary, lead to simple equations for moulds (be that in the symmetral or the symmetrel case); this simplicity is itself an image of the simplicity of the coproducts of the shuffle or quasishuffle Hopf algebras. These equations yield in fact *closed-form expressions* for the sought moulds, which are *surprisingly explicit*. Yet, when one wants to go beyond the formal level, to eventually get analytic transformations, these expressions are too coarse: although its is usually relatively easy to prove geometric growth estimates based on the explicit mould formulas, the inevitable factorial growth that composing differential operators brings along, is an obstacle to convergence.

There is thus a need for some *intermediate Hopf algebra* for which the calculations are still tractable, yet efficient enough to yield analytic functions when needed. This is exactly what

arborification-coarborification does and, in terms of Hopf algebras, the decorated Connes-Kreimer algebra will then rather naturally enter the stage.

5. Arborification-Coarborification

5.1. Hopf algebras of trees

We use here the results and notations developed in [8], [13] and [14]. A (non-planar) rooted tree T is a connected and simply connected set of oriented edges and vertices such that there is precisely one distinguished vertex (the root) with no incoming edge. An alternative definition can be given in terms of posets containing a smallest element, and for which each element has at most one predecessor. A forest F is a monomial in rooted trees. Let $l(F)$ be the number of vertices in F. Using the set H we can decorate a forest, that is to say that, to each vertex v of F, we associate an element $h(v)$ of H. We note $\mathscr{T}_H$ (resp. $\mathscr{F}_H$) the set of decorated trees (resp. forests) that contains the empty tree noted $\emptyset$. In fact there is a natural equivalence relation for trees, two trees being equivalent iff there is an automorphism of decorated posets that sends one to the other. It is rather the set of equivalent classes of trees that is denoted by $\mathscr{T}_H$, using a traditional abuse of language. As for sequences, if a forest F is decorated by $\eta_1, \ldots, \eta_s$ ($l(F) = s$), we note

$$\|F\| = \eta_1 + \cdots + \eta_s \in \bar{H}, \ \mathrm{gr}(F) = \mathrm{gr}(\eta_1) + \cdots + \mathrm{gr}(\eta_s).$$

For example, if

$$T = \begin{array}{c} \eta_4 \\ | \\ \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array}$$

then $l(T) = 4$ and $\|T\| = \eta_1 + \eta_2 + \eta_3 + \eta_4$.

Let us also recall that, for η in H, the operator B_η^+ associates to a forest of decorated trees the tree with root decorated by η connected to the roots of the forest: $B_\eta^+(\emptyset)$ is the tree with one vertex decorated by η and for example:

$$
(3) \qquad B_\eta^+ \left(\begin{array}{cc} \begin{array}{c} \eta_4 \\ | \\ \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array} & \begin{array}{c} \eta_6 \\ | \\ \eta_5 \end{array} \end{array} \right) = \begin{array}{c} \eta_4 \\ | \\ \eta_2 \quad \eta_3 \eta_6 \\ \backslash \ / \ | \\ \eta_1 \quad \eta_5 \\ \backslash \ / \\ \eta \end{array} \ .
$$

The linear span CK_H of $\mathscr{F}_H$ is a graded commutative algebra for the product

$$\pi(F_1 \otimes F_2) = F_1 F_2$$

and the unit $\emptyset$. Moreover, with the coproduct Δ given by induction by $\Delta(\emptyset) = \emptyset \otimes \emptyset$, $\Delta(T_1 \cdots T_k) = \Delta(T_1) \cdots \Delta(T_k)$ and

$$\Delta(B_\eta^+(F)) = B_\eta^+(F) \otimes \emptyset + (\mathrm{Id} \otimes B_\eta^+) \circ \Delta(F),$$

CK_H is the Connes-Kreimer Hopf algebra of trees decorated by H.

There exists a combinatorial description of this coproduct (see [13]). For a given tree $T \in \mathscr{T}_H$, an admissible cut c is a subset of its vertices such that, on the path from the root to an element of c, no other vertex of c is encountered. For such an admissible cut, $P^c(T)$ is the product of the subtrees of T whose roots are in c and $R^c(T)$ is the remaining tree, once these subtrees have been removed. With these definitions, for any tree T, we have

$$\Delta(T) = \sum_{\mathrm{adm\ cut}} P^c(T) \otimes R^c(T).$$

For example,

$$\Delta\left(\begin{array}{c} \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array} \right) = \begin{array}{c} \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array} \otimes \ \emptyset \ + \quad \eta_3 \ \otimes \ \begin{array}{c} \eta_2 \\ | \\ \eta_1 \end{array} + \eta_2 \otimes \begin{array}{c} \eta_3 \\ | \\ \eta_1 \end{array}$$

$$+ \quad \eta_2\, \eta_3 \ \otimes \ \eta_1 \ + \ \emptyset \ \otimes \ \begin{array}{c} \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array} \ .$$

REMARK 1. – One admissible cut will not yield all the possible subtrees, in the sense of connected subgraphs, of a given tree T. The subtrees t we get are such that
– either t has the same root as T (t will be called a *rooted subtree*),
– or t is a *complete subtree*, that is containing a vertex of T and all its descendants in T.
For instance, the tree $\begin{array}{c} \eta_2 \quad \eta_3 \\ \backslash \ / \\ \eta_1 \end{array}$ is a subtree of the tree on the right of formula (3) which cannot be obtained after one admissible cut.

Once again we can consider the convolution algebra $\mathscr{L}(\mathrm{CK}_H, \mathbb{C})$ and any morphism u of this algebra is given by its values on the basis $\mathscr{F}_H$. The definitions of arborescent moulds can then be rephrased:

DEFINITION 3. – *An arborescent mould $M^{\bullet^<}$ on H with values in $\mathbb{C}$ is a collection of complex numbers $\{M^F \in \mathbb{C}, \ F \in \mathscr{F}_H\}$.*

Such arborescent moulds are in one to one correspondence with the elements of $\mathscr{L}(\mathrm{CK}_H, \mathbb{C})$ and the product of such moulds corresponds to the convolution of the associated linear morphism.
Note that
1. a character on CK_H defines a *separative* mould $M^{\bullet^<}$, i.e.,

$$M^{T_1 \cdots T_s} = M^{T_1} \cdots M^{T_s} \ (\mathrm{and} M^{\emptyset} = 1),$$

2. an infinitesimal character on CK_H defines a *antiseparative* mould $M^{\bullet^<}$, i.e., for $s \geqslant 2$,

$$M^{T_1 \cdots T_s} = 0 \ (\mathrm{and} M^{\emptyset} = 0).$$

Since the coproduct is not as trivial as before, the convolution and inversion of characters are not so easy to handle. Nonetheless, we get partial but useful formulas for "root" characters, namely characters vanishing on forests $T_1 \cdots T_s$ such that at least one of the trees T_i as more than one vertex. For such a character χ, we have

$$\forall u \in \mathscr{L}(\mathrm{CK}_H, \mathbb{C}), \ \forall T = B_\eta^+(F) \in \mathscr{T}_H, \ (u * \chi)(T) = u(T) + u(F)\chi(\bullet_\eta),$$

and one can deduce that for any tree T decorated by $\eta_1, \ldots, \eta_s$ ($l(T) = s$)

$$\chi^{*^{-1}}(T) = (-1)^{l(T)} \chi(\bullet_{\eta_1}) \cdots \chi(\bullet_{\eta_s}).$$

The graded dual of CK_H will play a crucial role in the sequel and is strongly related to the Grossman-Larson Hopf algebra GL_H (see [17], [18], [28], [19] and [33]). The algebra GL_H is the linear span of rooted trees whose vertices (except the root) are decorated by H (see [14]): using 0 to note the absence of decoration, any such tree can be written $B_0^+(F)$ where F is in $\mathcal{F}_H$.

Let $F = T_1 \cdots T_k \in \mathcal{F}_H$ and $T_0 \in B_0^+(\mathcal{F}_H)$, the product of $B_0^+(F)$ and T_0 in GL_H is defined as follows: for any sequence $s = (s_1, \ldots s_k)$ of vertices of T_0 (with possible repetitions), let $(T_1, \ldots, T_k) \circ_s T_0$ be the tree of $B_0^+(F)$ obtained by identifying the root of $B_0^+(T_i)$ with the vertex s_i in T_0. The product π in GL_H is then defined by

$$B_0^+(T_1 \cdots T_k).T_0 = \sum_s (T_1, \ldots, T_k) \circ_s T_0$$

and the unit is $B_0^+(\emptyset)$. The coproduct is given by

$$\Delta(B_0^+(T_1 \cdots T_k)) = \sum_{I \subseteq \{1,\ldots,k\}} B_0^+(T_I) \otimes B_0^+(T_{\{1,\ldots k\}-I}),$$

where I is any subset of $\{1, \ldots, k\}$ and $T_I = \prod_{i \in I} T_i$.

For a forest F in $\mathcal{F}_H$ we remind that the symmetry factor of F is defined by:

1. $s((\eta)) = 1$;
2. $s(B_\eta^+(F)) = s(F)$;
3. $s(T_1^{a_1} \cdots T_k^{a_k}) = s(T_1)^{a_1} \cdots s(T_k)^{a_k} a_1! \cdots a_k!$ if $T_1, \ldots, T_k$ are *distinct* rooted trees.

This factor $s(F)$ is the cardinal of the group of automorphisms of the decorated poset F.

We have the following result, which is by now a classical one, and for which various proofs are available ([14], [21], [28], [19], [33]).

LEMMA 1. – *The map ϕ from GL_H to CK_H° defined by $\phi(B_0^+(F)) = s_F F$ defines an isomorphism of graded Hopf algebras between GL_H and CK_H°.*

5.2. Homogeneous coarborification

In each case (Vector Fields or Diffeomorphisms), the initial object defines a morphism ρ from Sh_H° or Qsh_H° to $\mathbb{C}[x, \partial_x]$ which is a coalgebra morphism and an algebra antimorphism that allows to compute some diffeomorphisms as characters on Sh_H or Qsh_H. We will essentially follow the same lines but with a morphism $\rho^<$ from CK_H° to $\mathbb{C}[x, \partial_x]$. Starting with this map ρ, one can define, using Ecalle's homogeneous coarborification the following linear morphism:

DEFINITION 4. – *The linear morphism $\rho^<$ from CK_H° to $\mathbb{C}[x, \partial_x]$ is defined on its linear basis by the following rules*

1. $\rho^<(\emptyset) = \text{Id}$.
2. *If $T = B_\eta^+(F)$ is a non empty tree, then*

$$\rho^<(T) = \sum_{i=1}^{\nu} (\rho^<(F).(\rho(\eta).x_i)) \partial_{x_i}.$$

3. *If $F = T_1 \cdots T_s$ with $s \geqslant 2$, then*

$$\rho^<(F) = \frac{1}{d_1! \cdots d_k!} \sum_{1 \leqslant i_1, \ldots, i_s \leqslant v} (\rho^<(T_1).x_{i_1}) \cdots (\rho^<(T_s).x_{i_s}) \partial_{x_{i_1}} \cdots \partial_{x_s},$$

where $F = T_1 \cdots T_s$ is the product of k distinct decorated trees, with multiplicities $d_1, \ldots, d_k$ $(d_1 + \cdots + d_k = s)$.

From this recursive definition, one already sees that the differential operator $\rho^<(F)$ is of order $r(F)$ (number of roots) and of homogeneity $\|F\|$. Thanks to the order of $\rho^<(F)$, this morphism is a coalgebra morphism and we have in fact the following:

THEOREM 5. – $\rho^<$ *is a Hopf morphism.*

Proof. – The proof is based on the following result of Grossman and Larson (see [33], [31]): Let τ the map from GL_H to $\mathbb{C}[x, \partial_x]$ defined by

1. $\tau(B_0^+(\emptyset)) = \mathrm{Id}$.
2. If $T = B_0^+(t)$ where $t = B_\eta^+(t_1 \cdots t_s)$ is a tree of $\mathscr{T}_H$, then

$$\tau(T) = \sum_{i=1}^{v} (\tau(B_0^+(t_1 \cdots t_s)).(\rho(\eta).x_i)) \partial_{x_i}.$$

3. If $T = B_0^+(t_1 \cdots t_s)$ $(s \geq 2)$, then

$$\tau(B_0^+(t_1 \cdots t_s)) = \sum_{1 \leqslant i_1, \ldots, i_s \leqslant v} (\tau(B_0^+(t_1)).x_{i_1}) \cdots (\tau(B_0^+(t_s)).x_{i_s}) \partial_{x_{i_1}} \cdots \partial_{x_{i_s}}.$$

Then τ is a Hopf morphism (the differential operators thus recursively defined are also known as elementary differentials in the literature on B-series, etc). One can convince oneself with the following example where:

$$T_1 = \begin{smallmatrix} \eta_1 \\ | \\ 0 \end{smallmatrix} \quad T_2 = \begin{smallmatrix} \eta_2 \ \eta_3 \\ \backslash \ / \\ 0 \end{smallmatrix} , \quad \pi(T_1 \otimes T_2) = \begin{smallmatrix} \eta_1 \eta_2 \eta_3 \\ \backslash | / \\ 0 \end{smallmatrix} + \begin{smallmatrix} \eta_1 \\ | \\ \eta_3 \ \ \eta_2 \\ \backslash \ / \\ 0 \end{smallmatrix} + \begin{smallmatrix} \eta_1 \\ | \\ \eta_2 \ \ \eta_3 \\ \backslash \ / \\ 0 \end{smallmatrix} .$$

We have:

$$\tau(T_1) = \sum_{i_1=1}^{v} (\rho(\eta_1).x_{i_1}) \partial_{x_{i_1}}, \quad \tau(T_2) = \sum_{i_2,i_3=1}^{v} (\rho(\eta_2).x_{i_2})(\rho(\eta_3).x_{i_3}) \partial_{x_{i_2}} \partial_{x_{i_3}}$$

and, using Leibniz rule,

$$\tau(T_1).\tau(T_2) = \left(\sum_{i_1=1}^{v} (\rho(\eta_1).x_{i_1}) \partial_{x_{i_1}} \right) \left(\sum_{i_2,i_3=1}^{v} (\rho(\eta_2).x_{i_2})(\rho(\eta_3).x_{i_3}) \partial_{x_{i_2}} \partial_{x_{i_3}} \right)$$

$$= \left(\sum_{i_1,i_2,i_3=1}^{v} (\rho(\eta_1).x_{i_1})(\rho(\eta_2).x_{i_2})(\rho(\eta_3).x_{i_3}) \partial_{x_{i_1}} \partial_{x_{i_2}} \partial_{x_{i_3}} \right.$$

$$+ \sum_{i_1,i_2,i_3=1}^{v} \left((\rho(\eta_1).x_{i_1})(\partial_{x_{i_1}}(\rho(\eta_2).x_{i_2})) \right) (\rho(\eta_3).x_{i_3}) \partial_{x_{i_2}} \partial_{x_{i_3}}$$

$$+ \left. \sum_{i_1,i_2,i_3=1}^{v} (\rho(\eta_2).x_{i_2}) \left((\rho(\eta_1).x_{i_1})(\partial_{x_{i_1}}(\rho(\eta_3).x_{i_3})) \right) \partial_{x_{i_2}} \partial_{x_{i_3}} \right)$$

$$= \tau \begin{pmatrix} \eta_1 \eta_2 \eta_3 \\ \backslash | / \\ 0 \end{pmatrix} + \tau \begin{pmatrix} \eta_1 \\ | \\ \eta_3 \ \ \eta_2 \\ \backslash \ / \\ 0 \end{pmatrix} + \tau \begin{pmatrix} \eta_1 \\ | \\ \eta_2 \ \ \eta_3 \\ \backslash \ / \\ 0 \end{pmatrix} .$$

But, thanks to the recursive definition of $\rho^<$, τ and ϕ, we have $\rho^< = \tau \circ \phi^{-1}$ and, since both τ and ϕ^{-1} (see [13, 14]) are Hopf morphisms, so is $\rho^<$. $\qquad\square$

Note that the construction of τ was given by Grossman and Larson only for the case of a family of derivations, which would exactly correspond here to the homogeneous components $\mathbb{B}_\eta$ of a vector field. In the case of the homogeneous components $\mathbb{D}_\eta$ of a diffeomorphism, this corresponds to the construction of Grossman and Larson *for the vector fields*:

$$\mathbb{E}_\eta = \sum_{i=1}^{v} (\mathbb{D}_\eta . x_i) \partial_{x_i}.$$

This means that the construction of the morphism $\rho^<$ only depends on the operators

$$\rho^<(\bullet_\eta) = \sum_{i=1}^{v} (\rho(\eta).x_i)\partial_{x_i} \quad \text{(here, the bullet designates a one vertex tree)}$$

but the origin of ρ (Vector field or diffeomorphism) reappears in the relations between ρ and $\rho^<$:

— In the shuffle case (Vector fields), we have, for $\eta_1 \in H$,

$$\rho((\eta_1)) = \mathbb{B}_{\eta_1} = \rho^<(\bullet^{\eta_1}).$$

— In the Quasishuffle case (Diffeomorphisms), if $f \in G$ is given by

$$f_i(x) = x_i \left(1 + \sum_{\eta \in H} a_\eta^i x^\eta \right)$$

then

$$\Theta_f = \mathrm{Id}_{\mathbb{C}[[x]]} + \sum_{s \geq 1} \sum_{\substack{(\eta_1,\ldots,\eta_s) \in H^s \\ 1 \leqslant i_1,\ldots,i_s \leqslant v}} \frac{1}{s!} a_{\eta_1}^{i_1} \cdots a_{\eta_s}^{i_s} x^{\eta_1 + \cdots + \eta_s} x_{i_1} \cdots x_{i_x} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}}$$

and

$$\rho((\eta)) = \sum_{\substack{(\eta_1,\ldots,\eta_s) \in H^s \\ \eta_1 + \cdots + \eta_s = \eta}} \sum_{1 \leqslant i_1,\ldots,i_s \leqslant v} \frac{1}{s!} a_{\eta_1}^{i_1} \cdots a_{\eta_s}^{i_s} x^{\eta_1 + \cdots + \eta_s} x_{i_1} \cdots x_{i_x} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}},$$

but, for $\eta \in \bar{H}$, one easily sees that

$$\rho((\eta)) = \sum_{\substack{F = \bullet^{\eta_1} \cdots \bullet^{\eta_s} \\ \|F\| = \eta \\ \eta_i \in H}} \rho^<(\bullet^{\eta_1} \cdots \bullet^{\eta_s}).$$

As in Section 3, we have

THEOREM 6. – *The map*

$$S_{\rho^<} : \mathcal{C}(\mathrm{CK}_H, \mathbb{C}) \to \qquad G$$
$$\chi \qquad \mapsto \mathrm{ev}\left(\sum_F \chi(F)\rho^<(F) \right)$$

defines an antimorphism of groups and $\Theta^\chi = \sum_F \chi(F)\rho^<(F)$ *is the substitution automorphism associated to* $S_\rho(\chi)$.

This is a presentation of Ecalle's arborification/coarborification apparatus, within a framework of Hopf algebras.

As we will see below, these series have many advantages in linearization problems:

— Modulo a restriction to a subalgebra of CK_H (and of its graded dual), strong assumptions on the spectrum will become unnecessary.
— There is a very simple criterion on characters χ in $\mathcal{C}(CK_H, \mathbb{C})$ that ensures the analyticity of $S_{\rho^<}(\chi)$.

Moreover, the previous computations of characters on Sh_H or Qsh_H were not useless: in many cases their computation is easier, thanks to the simplicity of the convolution product, and for example, once such a character χ on Sh_H or Qsh_H is given in closed-form expression, one can easily derive a closed-form form expression for the character $\chi^<$ on CK_H such that

$$S_{\rho^<}(\chi^<) = S_\rho(\chi).$$

5.3. Arborification

For the deconcatenation coproduct on Sh_H or Qsh_H, if $L_+^\eta(\eta) = \eta\eta$, then

$$\Delta \circ L_+^\eta = 1 \otimes L_+^\eta + (L_+^\eta \otimes \mathrm{Id}) \circ \Delta.$$

We thus have the cocycle property, and the morphism α such that

$$\alpha \circ B_\eta^+ = L_+^\eta \circ \alpha$$

is a coalgebra *antimorphism* from CK_H to Sh_H or Qsh_H ([13]).

It is the fact that CK_H is an initial object in a category of coalgebras, for a certain cohomology (dual to Hochschild cohomology of algebras) that ensures the existence of the morphism α, which is a morphism of Hopf algebras. We shall not expand on this (as shown by Foissy, the cohomology groups vanish in degree $\geqslant 2$), yet it is satisfactory to have such a simple algebraic characterization of arborification through a universal property of Connes-Kreimer's algebra, which is an important object in its own right.

We shall now see how to recover the same diffeomorphism using α: going back to our conjugacy equations, the change of coordinates, in both cases, is given by a substitution automorphism

$$\Theta = \sum_\eta \chi(\eta)\rho(\eta),$$

and to any such character χ we have associated an arborified character $\chi^< = \chi \circ \alpha$. We should try to use this new character on CK_H to rearrange the above series and finally get some analyticity properties. To do so, let us use the new Hopf algebra morphism $\rho^<$ from CK_H° to $\mathbb{C}[x, \partial_x]$:

$$\Theta = \sum_\eta \chi(\eta)\rho(\eta) = \sum_F \chi^<(F)\rho^<(F).$$

But then

$$\sum_F \chi^<(F)\rho^<(F) = \sum_F \chi(\alpha(F))\rho^<(F)$$

$$= \sum_F \chi\left(\sum_\eta \langle \eta, \alpha(F)\rangle\eta\right)\rho^<(F)$$

$$= \sum_{\eta,F} \chi(\eta)\langle \eta, \alpha(F)\rangle\rho^<(F)$$

$$= \sum_\eta \chi(\eta)\sum_F \langle \eta, \alpha(F)\rangle\rho^<(F)$$

$$= \sum_\eta \chi(\eta)\sum_F \langle \alpha^\circ(\eta), F\rangle\rho^<(F)$$

$$= \sum_\eta \chi(\eta)\rho^<(\alpha^\circ(\eta))$$

so that it appears indeed highly desirable to have such morphisms as $\rho^<$ that fulfills the relation

$$\rho^< \circ \alpha^\circ = \rho.$$

The choice of $\rho^<$ is not unique but the map defined in Section 5.2 works and it is that particular choice which has been called [11] the natural (or homogeneous) coarborification and which is adapted to the analytic study of F.

Theorem 7. – *We have*

$$\rho^< \circ \alpha^\circ = \rho.$$

Proof. – ρ and α° are coalgebra morphisms and algebra antimorphisms and $\rho^<$ is a Hopf morphism, so $\rho^< \circ \alpha^\circ$ and ρ are coalgebra morphisms and algebra antimorphisms.

In the shuffle case (Vector fields), since Sh_H° is freely generated by the words of length 1, it is sufficient to check that both morphisms coincides on these words. But $\alpha^\circ((\eta_1)) = \bullet^{\eta_1}$ thus

$$\rho((\eta_1)) = \mathbb{B}_{\eta_1} = \rho^<(\bullet^{\eta_1}) = \rho^< \circ \alpha^\circ((\eta_1)).$$

The same proof holds in the quasishuffle case: if $f \in G$ is given by

$$f_i(x) = x_i\left(1 + \sum_{\eta\in H} a_\eta^i x^\eta\right),$$

then

$$\Theta_f = \mathrm{Id}_{\mathbb{C}[[x]]} + \sum_{s\geq 1}\sum_{\substack{(\eta_1,\ldots,\eta_s)\in H^s \\ 1\leqslant i_1,\ldots,i_s\leqslant v}} \frac{1}{s!}a_{\eta_1}^{i_1}\cdots a_{\eta_s}^{i_s} x^{\eta_1+\cdots+\eta_s} x_{i_1}\cdots x_{i_x}\partial_{x_{i_1}}\cdots\partial_{x_{i_s}}$$

and

$$\rho((\eta)) = \sum_{\substack{(\eta_1,\ldots,\eta_s)\in H^s \\ \eta_1+\cdots+\eta_s=\eta}}\sum_{1\leqslant i_1,\ldots,i_s\leqslant v} \frac{1}{s!}a_{\eta_1}^{i_1}\cdots a_{\eta_s}^{i_s} x^{\eta_1+\cdots+\eta_s} x_{i_1}\cdots x_{i_x}\partial_{x_{i_1}}\cdots\partial_{x_{i_s}}.$$

But for $\eta \in \bar{H}$, one easily sees that

$$\rho((\eta)) = \sum_{\substack{F = \bullet^{\eta_1} \cdots \bullet^{\eta_s} \\ \|F\| = \eta \\ \eta_i \in H}} \rho^<(\bullet^{\eta_1} \cdots \bullet^{\eta_s})$$

$$= \rho^<(\alpha^\circ((\eta))$$

and this terminates the proof.

$\square$

REMARK 2. – The mechanism of arborification of moulds has in effect been independently rediscovered by Ander Murua in [27], involving Connes-Kreimer Hopf algebra, for efficient calculations involving Lie series in problems of control theory; in that paper, the author is then also lead to coarborification by considering the graded duals, and going thus to the Grossman-Larson algebra.

In the reverse direction, Wenhua Zhao (see [33], [31], [32]) has for his part rediscovered the constructions of coarborification and then obtained in effect the mechanisms of arborification by dualizing and going to CK. Notably, Zhao's results concern in fact both plain and contracting arborification.

More recently, the universal property of CK has also been used (in the non decorated case) in the same way as in our presentation, for a factorization of characters of the quasishuffle algebra in [3].

It must be stressed, however, that the crucial properties for the analyst come *after* these general constructions: namely the existence of closed-form expressions for the arborified moulds, which make it possible to obtain the necessary estimates, as we shall see below.

A very striking instance, though, where an independent approach has exactly lead to arborification, once translated in terms of characters of the relevant Hopf algebras, and includes for the applications a crucial closed-form is [15]. Finally, in several very recent works in the algebraic theory of non-linear control (see [22] and the references therein) some particular characters of the same class of Hopf algebras we are involved with in the present work show up, which translate into moulds of constant use in Ecalle's papers.

5.4. Some examples

5.4.1. *The shuffle case*. – For the tree

$$t = \begin{array}{c} \eta_4 \\ | \\ \eta_2 \quad \eta_3 \\ \backslash \quad / \\ \eta_1 \end{array}$$

we get

$$\alpha(t) = (\eta_1 \eta_2 \eta_3 \eta_4) + (\eta_1 \eta_3 \eta_2 \eta_4) + (\eta_1 \eta_3 \eta_4 \eta_2).$$

Under the strong assumption on the spectrum (the λ_i are independant over the integers), for the character ξ given by

$$\xi(\eta_1, \ldots, \eta_s) = \frac{1}{\langle \lambda, \eta_1 + \cdots + \eta_s \rangle \langle \lambda, \eta_2 + \cdots + \eta_s \rangle \cdots \langle \lambda, \eta_s \rangle},$$

a simple computation yields:

$$\xi^<(t) = \xi(\alpha(t)) = \frac{1}{\langle \lambda, \eta_1 + \eta_2 + \eta_3 + \eta_4 \rangle \langle \lambda, \eta_2 \rangle \langle \lambda, \eta_3 + \eta_4 \rangle \langle \lambda, \eta_4 \rangle}.$$

For this character, even if the evaluation of $\xi^<$ on a tree involves evaluation of ξ on many sequences, there exists finally a surprisingly simple formula for $\xi^<$:

PROPOSITION 8. – *Let f be a tree with s vertices decorated by $\eta_1, \ldots, \eta_s$. For $1 \leqslant i \leqslant s$ if t_i is the (complete) subtree of f whose root is labelled by η_i, then*

$$\xi^<(f) = \prod_{i=1}^s \frac{1}{\langle \lambda, \|t_i\| \rangle}.$$

The reader can check this formula on the previous example where

$$t_1 = \begin{matrix} \eta_2 & \eta_3 \\ \diagdown & \diagup \\ & \eta_1 \end{matrix} \quad , \quad t_2 = \eta_2, \quad t_3 = \begin{matrix} \eta_4 \\ | \\ \eta_3 \end{matrix}, \quad t_4 = \eta_4.$$

Proof. – This result can be proved recursively on the number s of vertices (i.e., the size of the forest). For forests of size 1, this formula is obvious.

If f is a forest of size $s \geqslant 2$ with at least two trees: $f = t_1 \cdots t_n$ ($n \geqslant 2$), then

$$\xi^<(f) = \xi^<(t_1) \cdots \xi^<(t_n),$$

but the size of each tree is less than s and we get by recursion the right formula.

If t is a tree of size $s \geqslant 2$, then $t = B_\eta^+(f)$ and

$$\xi^<(t) = \xi(\alpha(B_\eta^+(f)) = \xi(L_+^\eta(\alpha(f))),$$

but, for any sequence η,

$$\xi(L_+^\eta(\eta)) = \frac{1}{\langle \lambda, \eta + \|\eta\| \rangle} \xi(\eta),$$

thus

$$\xi^<(t) = \xi(L_+^\eta(\alpha(f))) = \frac{1}{\langle \lambda, \|t\| \rangle} \xi(\alpha(f)) = \frac{1}{\langle \lambda, \|t\| \rangle} \xi^<(f)$$

and, once again, we get recursively the right formula.   □

5.4.2. *The quasishuffle case.* – For the tree

$$t = \begin{matrix} & & \eta_4 \\ & & | \\ & \eta_2 & \eta_3 \\ & \diagdown & \diagup \\ & & \eta_1 \end{matrix}$$

we get

$$\alpha(t) = (\eta_1, \eta_2, \eta_3, \eta_4) + (\eta_1, \eta_3, \eta_2, \eta_4) + (\eta_1, \eta_3, \eta_4, \eta_2) + (\eta_1, \eta_2 + \eta_3, \eta_4) + (\eta_1, \eta_3, \eta_2 + \eta_4).$$

Under the strong assumption on the spectrum, for the character χ given by

$$\chi(\eta_1, \ldots, \eta_s) = \frac{1}{(l^{\eta_1 + \cdots + \eta_s} - 1)(l^{\eta_2 + \cdots + \eta_s} - 1) \cdots (l^{\eta_s} - 1)},$$

a simple computation yields:

$$\chi^<(t) = \chi(\alpha(t)) = \frac{1}{(l^{\eta_1+\eta_2+\eta_3+\eta_4} - 1)(l^{\eta_2} - 1)(l^{\eta_3+\eta_4} - 1)(l^{\eta_4} - 1)}.$$

The same proof as before gives

PROPOSITION 9. – *Let f be a tree with s vertices decorated by $\eta_1, \ldots, \eta_s$. For $1 \leqslant i \leqslant s$ if t_i is the (complete) subtree of f whose root is labelled by η_i, then*

$$\chi^<(f) = \prod_{i=1}^{s} \frac{1}{(l^{\|t_i\|} - 1)}.$$

Once again the formula is surprisingly simple and as we shall see in the following section, if we have "geometric" estimates on such an arborified character we will prove the analyticity of the associated diffeomorphism.

But we still have to work with strong assumptions on the spectrum. We will circumvent this difficulty using the following remarks:

1. One can obtain the above formula without arborification by translating directly the linearization equations as character equations on CK_H.
2. We will then prove that, in order to define the corresponding diffeomorphism, it is sufficient to compute a character on a sub-Hopf algebra of CK_H were the sought character is well-defined under the *weak* assumption on the spectrum.

6. Back to linearization

6.1. Equations for characters of CK_H

As in Section 3, if

$$X = X^{\text{lin}} + \sum_{\eta \in H} \mathbb{B}_\eta = X^{\text{lin}} + P,$$

with $X^{\text{lin}} = \sum_{1 \leqslant i \leqslant v} \lambda_i x_i \partial_{x_i}$, the vector field P is given by the infinitesimal character u on CK_H:

$$u(f) = \begin{cases} 1 \text{ if } f = \bullet_\eta \\ 0 \text{ otherwise.} \end{cases}$$

That is to say:

$$X = X^{\text{lin}} + \sum_{\eta \in H} \mathbb{B}_\eta = X^{\text{lin}} + \sum_{f \in \mathscr{F}_H} u(f)\rho^<(f).$$

The diffeomorphism φ that linearizes X ($X^{\text{lin}}.F_\varphi = F_\varphi.X$) can be obtained as $\varphi = S_{\rho^<}(\xi)$ where ξ is a character on CK_H such that

$$\nabla \xi = \xi * u,$$

where

$$(\nabla \chi)(f) = \langle \lambda, \|f\| \rangle \chi(f).$$

It is then easy to check directly on this equation that if $\langle \lambda, \eta \rangle \neq 0$ for any η in $\bar{H}$, this character is uniquely defined and is given by proposition 8.

On the same way, for a diffeomorphism $f^{\text{lin}} \circ f$ where

$$F_f = \text{Id} + \sum_{\eta \in \bar{H}} \mathbb{D}_\eta,$$

the character ξ on CK_H given by

$$\xi(f) = \begin{cases} 1 \text{ if } f = \bullet_{\eta_1} \cdots \bullet_{\eta_s} \\ 0 \text{ otherwise} \end{cases}$$

is such that

$$F_f = \sum_{f \in \mathscr{F}_H} \xi(f) \rho^<(f),$$

and if χ is a character such that

$$\chi \circ \sigma = \chi * \xi \; (\sigma(f) = l^{\|f\|} f),$$

then $\varphi = S_{\rho^<}(\chi)$ is such that

$$f^{\text{lin}} \circ f \circ \varphi = \varphi \circ f^{\text{lin}}.$$

Once again, if, for any $\eta \in \bar{H}$, $l^\eta \neq 1$, then χ is well-defined and is given by proposition **9**.

We still have the strong condition because, in order to compute such characters on a forest f, one has to divide by $\langle \lambda, \|f\| \rangle$ or $l^{\|f\|} - 1$ and $\|f\|$ runs over $\bar{H}$. But, as we shall see, when considering a substitution automorphism

$$F = \sum_{f \in \mathscr{F}_H} \chi(f) \rho^<(f),$$

there are many forests f such that $\rho^<(f) = 0$. Omitting these terms in the series defining F, one can consider that f runs over a subset $\mathscr{F}_H^+$ of $\mathscr{F}_H$ which is the linear basis of a sub-Hopf algebra CK_H^+ of CK_H. We will thus be able to consider the previous character equations on CK_H^+ and there will exist a unique solution as soon as $\langle \lambda, \eta \rangle \neq 0$ or $l^\eta - 1 \neq 0$ for all η in H.

6.2. The non-resonance condition and the subalgebras of CK_H

The attempted sub-Hopf algebra CK_H^+, is indeed a polynomial algebra over trees in a subset $\mathscr{T}_H^+ \subset \mathscr{T}_H$ which is defined recursively on the number of vertices.

DEFINITION 5. – *Let $\mathscr{T}_H^{+,1}$ be the set of trees T with one vertex ($l(T) = 1$) such that $\|T\| \in H$. For $l \geq 2$, the sets $\mathscr{T}_H^{+,l}$ are defined recursively as follows. A tree T is in $\mathscr{T}_H^{+,l}$ if and only if:*

1. *T has l vertices,*
2. *$\|T\| \in H$,*
3. *for any nontrivial any admissible cut $(R^c(T), P^c(T))$, that is $R^c(T) = T_0 \neq \emptyset$ and $P^c(T)) = T_1 \cdots T_s \neq \emptyset$, the trees $T_0, T_1, \ldots T_s$ are in $\bigcup_{k=1}^{l-1} \mathscr{T}_H^{+,k}$.*

We note $\mathscr{T}_H^+ = \bigcup_{k \geq 1} \mathscr{T}_H^{+,k}$ and $\mathscr{F}_H^+$ the set of forests of such trees.

It is readily checked that CK_H^+, as a polynomial algebra, is a subalgebra of CK_H but, thanks to point 3 in the recursive definition of $\mathscr{T}_H^+$, for any tree in CK_H^+,

$$\Delta(T) = \sum_{\text{adm cut}} P^c(T) \otimes R^c(T) \in CK_H^+ \otimes CK_H^+,$$

thus CK_H^+ is a sub-Hopf algebra of CK_H. In order to reach the weak condition in linearization problems, one can also notice that any (rooted or complete) subtree t of a tree $T \in \mathscr{T}_H^+$ is also in $\mathscr{T}_H^+$ (and consequently $\|t\| \in H$) since such a subtree is obtained through an admissible cut of T. Moreover:

THEOREM 8. – *If a forest F in $\mathscr{F}_H$ does not belong to $\mathscr{F}_H^+$, then*

$$\rho^<(F) = 0.$$

Proof. – Starting with a diffeomorphism or a vector field, it is clear that for $\eta \in H$, the image of the one node tree decorated by η is:

$$\rho^<(\eta) = \sum_{i=1}^{\nu} u_\eta^i x^{\eta + e_i} \partial_{x_i},$$

where $u_\eta^i \in \mathbb{C}$. For any forest $F = T_1 \cdots T_s$ in $\mathscr{F}_H$, $\rho^<(F)$ is an endomorphism of $\mathbb{C}[x]$ such that

$$\rho^<(T_1 \cdots T_s) = \sum_{1 \leqslant i_1,\ldots,i_s \leqslant \nu} P_F^{i_1,\ldots i_s}(u) x^{\|F\| + e_{i_1} + \cdots + e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}},$$

where the coefficients $P_F^{i_1,\ldots i_s}(u)$ are polynomials in the variables $u = \{u_\eta^i\}$ with coefficients in $\mathbb{Q}^+$ ($P_F^{i_1,\ldots i_s}(u) \in \mathbb{Q}^+[u]$). Thanks to Definition 4, if $s \geq 2$ and $F = T_1 \cdots T_s$ is the product of k distinct decorated trees, with multiplicities $d_1, \ldots, d_k$ ($d_1 + \cdots + d_k = s$), we get

$$(4) \qquad P_F^{i_1,\ldots i_s}(u) = \frac{1}{d_1! \cdots d_k!} P_{T_1}^{i_1}(u) \ldots P_{T_s}^{i_s}(u).$$

With this remark, we are ready to prove the theorem recursively on the number of vertices of $F \in \mathscr{F}_H \setminus \mathscr{F}_H^+$. For $l \geq 1$, let

$$E_l = \{F \in \mathscr{F}_H \setminus \mathscr{F}_H^+ ; l(F) \leq l\}.$$

Since E_1 is empty, the theorem is true for any forest in E_1. Let us suppose now that, for a given $l \geq 1$, the theorem is true for any element of E_l and consider a forest with $l+1$ vertices in $\mathscr{F}_H \setminus \mathscr{F}_H^+$. There are three cases to consider:

1. Either $F = T_1 \ldots T_s \notin \mathscr{F}_H^+$ with $s \geq 2$,
2. or F is a tree T and $\|T\| \notin H$,
3. or F is a tree T, $\|T\| \in H$ but there exists a nontrivial admissible cut $(R^c(T), P^c(T))$ such that $R^c(T) \notin \mathscr{F}_H^+$ or $P^c(T) \notin \mathscr{F}_H^+$.

In the first case, at least one of the trees, say T_1 (with $l(T_1) \leq l$) is not in $\mathscr{F}_H^+$. This means that $T_1 \in E_l$ thus

$$\rho^<(T_1) = \sum_{i=1}^{\nu} P_{T_1}^i(u) x^{\|T\| + e_i} \partial_{x_i} = 0.$$

In other words, $P_{T_1}^i(u) = 0$ for any $1 \leqslant i \leqslant \nu$ and, thanks to Equation 4, $\rho^<(F) = 0$.

In the second case, F is a tree T in $\mathscr{T}_H$ such that $\|T\| \notin H$. This means that, for any $1 \leqslant i \leqslant \nu$, $\|T\| + e_i \notin \mathbb{N}^\nu$. But

$$\rho^<(T) = \sum_{i=1}^\nu P_T^i(u) x^{\|T\|+e_i} \partial_{x_i},$$

with

$$P_T^i(u) x^{\|T\|+e_i} = \rho^<(T).x_i \in \mathbb{C}[x]$$

and, since $x^{\|T\|+e_i}$ is not in $\mathbb{C}[x]$, for $1 \leqslant i \leqslant \nu$, $P_T^i(u) = 0$ and $\rho^<(T) = 0$.

It remains to look at the third case. There exists a nontrivial admissible cut $(R^c(T), P^c(T))$ with $R^c(T)$ or $P^c(T)$ in E_l: One can deduce that either $\rho^<(R^c(T)) = 0$ or $\rho^<(P^c(T)) = 0$ thus,

$$\rho^<(P^c(T)).\rho^<(R^c(T)) = 0 = \rho^< \circ \pi(P^c(T) \otimes R^c(T)),$$

where π is the product in CK_H°, dual to the coproduct of CK_H. But thanks to the definition of this coproduct

$$\pi(P^c(T) \otimes R^c(T)) = cT + Q,$$

where $c \in \mathbb{N}^*$ and Q is a combination of forests with coefficients in $\mathbb{N}$. For $1 \leqslant i \leqslant \nu$,

$$\rho^<(P^c(T)).\rho^<(R^c(T)).x_i = c P_T^i(u) x^{\|T\|+e_i} + \rho^<(Q).x_i = 0$$
$$= (c P_T^i(u) + Q^i(u)) x^{\|T\|+e_i} = 0.$$

This means that the polynomial $c P_T^i(u) + Q^i(u)$ is zero but, since it is a linear combination (with positive coefficients) of polynomials in $\mathbb{Q}^+[[u]]$,

$$P_T^i(u) = Q^i(u) = 0$$

and finally $\rho^<(T) = 0$.

This ends the recursive proof since any forest F in E_{l+1} is such that $\rho^<(F) = 0$. $\qquad\square$

This means that, if p is the projection of CK_H on CK_H^+ or of CK_H° on $\mathrm{CK}_H^{+\circ}$ defined by

$$\forall F \in \mathscr{F}_H, \ p(F) = \begin{cases} 0 \ \text{if } F \notin \mathscr{F}_H^+ \\ F \ \text{if } F \in \mathscr{F}_H^+, \end{cases}$$

then $\rho^< \circ p$ is still a Hopf algebra morphism. Moreover, for any character χ on CK_H or $\mathrm{CK}_{\bar{H}}$, $\chi \circ p$ is a character on CK_H^+ and

$$\sum_F \chi(F)\rho^<(F) = \sum_F \chi(p(F))\rho^<(p(F)) = \sum_{F \in \mathrm{CK}_H^+} \chi(F)\rho^<(F).$$

In other words, in the linearization equation, one can look for a substitution morphism given by a character on CK_H^+ and this one is well-defined as soon as we have the *weak* non-resonance condition.

It is easy to see that *any* subtree of a given tree T can be obtained from T through at most *two* admissible cuts. The recursive definition of $\mathscr{T}_H^+$ then ensures that any subtree t of a given tree $T \in \mathscr{T}_H^+$ is also in $\mathscr{T}_H^+$ and thus has its weight $\|t\|$ in H. We can state the following proposition (suggested by one referee of this paper) that can be used as an alternative definition of $\mathscr{T}_H^+$:

PROPOSITION 10. – *The set $\mathcal{T}_H^+$ is the set of trees T such that, for **any** subtree T' of T, $\|T'\| \in H$.*

As observed above, any subtree of a tree $T \in \mathcal{T}_H^+$ has this property and the converse is proved by induction on the number of vertices: for trees with one vertex, this property coincides with Definition 5. Otherwise, consider a tree T such that, for **any** subtree T' of T, $\|T'\| \in H$. In particular, $\|T\| \in H$ and for any non trivial admissible cut c of T, with $P^c(T) = T_0$ and $R^c = T_1, \ldots, T_s$, as any subtree t of $T_0, \ldots, T_s$ is also a subtree of T, we have $\|t\| \in H$. Using the induction hypothesis, $T_0, \ldots, T_s$ belong to $\mathcal{T}_H^+$ and thus T is in $\mathcal{T}_H^+$.

The Hopf algebra CK_H^+, which does not appear in the literature, is the relevant object one has to use, in order to recover the usual results on formal linearization:

1. It works with the classical conditions on the spectrum; no extra assumption is needed.
2. The diffeomorphism is expressed by a character which is given without ambiguity.

It remains to prove that CK_H^+ is also extremely well-suited to consider the analyticity of such a diffeomorphism. In other words, the Hopf algebra CK_H^+ is the right algebra to deal with questions of convergence in linearization problems (and in fact also in more general normalization problems, in situations involving resonances).

6.3. Majorant series and analyticity

Using majorant series, it is easy to see that

$$\boldsymbol{G}_{\mathrm{ana}} = \{\varphi = (\varphi_1, \ldots, \varphi_\nu) \in \boldsymbol{G} \; ; \; \varphi_i(x) \in \mathbb{C}\{x\}\}$$

is a subgroup of $\boldsymbol{G}$ and this still holds for many subsets of diffeomorphisms whose coefficients satisfy some particular estimates (see [26]).

THEOREM 9. – *Let $B = \{B_\eta \in \mathbb{R}^+, \ \eta \in H\}$ be a set of submultiplicative estimates: for all η, η_1, η_2 in H such that $\eta = \eta_1 + \eta_2$, $B_{\eta_1} B_{\eta_2} \leqslant B_\eta = B_{\eta_1 + \eta_2}$. Let $\boldsymbol{G}_B$ be the subset of $\boldsymbol{G}$ of diffeomorphisms φ such that there exists $A > 0$ and*

$$\forall 1 \leqslant i \leqslant \nu, \forall \eta \in H_i, \ |\varphi_\eta^i| \leqslant B_\eta A^{|\eta|}.$$

Then $\boldsymbol{G}_B$ is a subgroup of $\boldsymbol{G}$.

The complete proof can be found in [26]. It relies on majoring series: let $\varphi(x) = x + u(x)$ in $\boldsymbol{G}$, we say that $\psi(x) = x + v(x)$ is a majorant series of φ ($\varphi \prec \psi$) if,

$$\forall 1 \leqslant i \leqslant \nu, \forall \eta \in H_i, \ |\varphi_\eta^i| \leqslant \psi_\eta^i.$$

For a given set B and $A > 0$, let $\psi_{B,A}$ be the diffeomorphism such that $C_\eta^i(\psi_{B,A}) = B_\eta A^{|\eta|}$. It is clear that $\psi_{B,A}$ is in $\boldsymbol{G}_B$ and φ belongs to $\boldsymbol{G}_B$ if and only if there exists $A > 0$ such that

$$\varphi \prec \psi_{B,A}.$$

Now the proof of the theorem relies on classic estimates that gives:

1. If $\varphi_1 \prec \psi_{B,A_1}$ and $\varphi_2 \prec \psi_{B,A_2}$ then there exists $A_3 > 0$ such that $\varphi_1 \circ \varphi_2 \prec \psi_{B,A_3}$. In other words, $\boldsymbol{G}_B$ is stable under the composition of diffeomorphisms.
2. If $\varphi_1 \prec \psi_{B,A_1}$ then there exists $A_2 > 0$ such that $\varphi_1^{\circ^{-1}} \prec \psi_{B,A_2}$ and this finally proves that $\boldsymbol{G}_B$ is a subgroup.

Note that the analytic subgroup corresponds to G_B with,

$$\forall \eta \in H, \; B_\eta = 1.$$

Now, using the same ideas as in [26], one easily gets that

THEOREM 10. – *Suppose that, the map $\rho^<$, restricted to $\mathrm{CK}_H^{+\circ}$ is such that:*

$$\rho^<(\bullet_\eta) = \sum_{1 \leqslant i \leqslant \nu} u_\eta^i x^\eta x_i \partial_{x_i},$$

with $|u_\eta^i| \leqslant B_\eta A^{|\eta|}$ for some $A > 0$. If χ is a character on CK_H^+ such that, for all forests $F \in \mathrm{CK}_H^+$,

$$|\chi(F)| \leqslant C^{\mathrm{gr}(F)},$$

then the diffeomorphism φ such that

$$\Theta_\varphi = \sum_{F \in \mathrm{CK}_H^+} \chi(F) \rho^<(F)$$

is in G_B.

Proof. – If we consider

$$u(x) = (u_1(x), \ldots, u_\nu(x)),$$

where

$$u_i(x) = x_i + \sum_{\eta \in H} \rho^<(\bullet_\eta).x_i = x_i + \sum_{\eta \in H_i} u_\eta^i x^\eta x_i,$$

then $u \prec \psi_{B,A} = v$. We note $\rho_u = \rho^<$ and ρ_v the similar morphism such that

$$\rho_v(\bullet_\eta) = \sum_{1 \leqslant i \leqslant \nu} v_\eta^i x^\eta x_i \partial_{x_i}.$$

For any forest $F = T_1 \cdots T_s$ in $\mathscr{F}_H^+$, we have once again

$$\rho_u(T_1 \cdots T_s) = \sum_{1 \leqslant i_1, \ldots, i_s \leqslant \nu} P_F^{i_1, \ldots, i_s}(u) x^{\|F\| + e_{i_1} + \cdots + e_{i_s}} \partial_{x_{i_1}} \cdots \partial_{x_{i_s}},$$

where the coefficients $P_F^{i_1, \ldots, i_s}(u)$ are polynomials in the variables $u = \{u_\eta^i\}$ with coefficients in $\mathbb{Q}^+$ ($P_F^{i_1, \ldots, i_s}(u) \in \mathbb{Q}^+[u]$). Since the coefficients of such polynomials are non-negative, it is clear that

$$|P_F^{i_1, \ldots, i_s}(u)| \leqslant P_F^{i_1, \ldots, i_s}(v),$$

and, if

$$\varphi(x) = \Theta_\varphi.x = \sum_{F \in \mathrm{CK}_H^+} \chi(F) \rho_u(F).x,$$

we have

$$\varphi(x) \prec \sum_{F \in \mathrm{CK}_H^+} |\chi(F)| \rho_v(F).x \prec \sum_{F \in \mathrm{CK}_H^+} C^{\mathrm{gr}(F)} \rho_v(F).x = \phi(x) \in G.$$

The map ξ defined on CK_H^+ by $\xi(F) = C^{\mathrm{gr}(F)}$ is a character and it is easy to check that its inverse is defined by

$$\xi^{*^{-1}}(F) = \begin{cases} 1 & \text{if } F = \emptyset \\ (-1)^s C^{\mathrm{gr}(F)} & \text{if } F = \bullet_{\eta_1} \cdots \bullet_{\eta_s} \\ 0 & \text{otherwise.} \end{cases}$$

This means that

$$\begin{aligned} \phi^{\circ^{-1}}(x) &= \sum_{F \in \mathrm{CK}_H^+} \xi^{*^{-1}}(F)\rho_u(F).x \\ &= x + \sum_{\eta \in H} \xi^{*^{-1}}(\bullet_\eta)\rho_v(\bullet_\eta).x \\ &= x - \sum_{\eta \in H} C^{\mathrm{gr}(\eta)}\rho_v(\bullet_\eta).x \\ &= 2x - \left(x + \sum_{\eta \in H} C^{\mathrm{gr}(\eta)}\rho_v(\bullet_\eta).x \right) \\ &= 2x - \frac{1}{C}v(Cx). \end{aligned}$$

As v is in G_B, so is $\phi^{\circ^{-1}}$ and, since G_B is a group, we have

$$\varphi(x) \prec \phi(x) \in G_B$$

and φ is in G_B. $\qquad\square$

The previous argument is a systematization of a process that was introduced by one of us (FM), and implemented in 2 previous papers ([25], [26]), regarding respectively non-linear q-difference equations and "Birkhoff decomposition" in spaces of Gevrey series.

6.4. Growth estimates for the arborified moulds

In order to give a nontrivial application, we show how Brjuno's classical result on linearization for non resonant fields can be obtained, once we match the previous estimate on the comould side with another one, regarding the geometric growth of the arborified mould.

We start with the vector field case and we denote by $M^{\bullet^<}$ the arborescent mould corresponding to the character ξ, for which a closed-form expression was obtained above.

In order to avoid technicalities in diophantine approximation, in the present work which is focussed on algebraic constructions, we shall consider vector fields which satisfy the following strong version of Brjuno's diophantine condition:

$$\sum \frac{1}{2^k}\mathrm{Log}(\frac{1}{\Omega(2^{k+1})}) < \infty,$$

where $\Omega(h) = \Omega(h) = \min\{|\langle n, \lambda \rangle|, n_i \in \mathbb{Z}, |\langle n, \lambda \rangle| > 0 \text{ and } \sum n_i \leqslant h\}$.

Note that, since Ω is decreasing, the above condition is equivalent to the condition:

$$S = \sum \frac{1}{2^k}\left|\mathrm{Log}(\frac{1}{\Omega(2^{k+1})})\right| < \infty$$

PROPOSITION 11. – *The arborified mould $M^{\bullet^{<}}$ has a geometric growth: there exists a constant K such that for any decorated forest F, we have $|M^F| \leqslant K^{\mathrm{gr}(F)}$.*

All proofs of normalization results under Brjuno's arithmetical condition rely at some point on an key estimate, usually known as a "Brjuno's counting lemma" (see e.g., the classical paper by J. Poeschel [29] for a particularly clear exposition of this, in the case of diffeomorphisms). The proof of the previous proposition will unsurprisingly also crucially depend as well on a version of a counting lemma which we give below. In the form that we use here, the estimate is proved in the paper [5], for the version of the lemma that is relevant in the case of diffeomorphisms). The paper [16], for example, explains the way trees appear naturally in this context; it is then straightforward to translate the version of the counting lemma for fields which is contained in Brjuno's seminal paper in the language of trees.

Note however that our presentation is different and totally independent of the one used in [16] and [5] but it is the very same counting argument that is crucial, as in any other proofs of results involving Brjuno's condition. The proof of the proposition itself will simply consist in regrouping subtrees in "slices" that are determined by a total weight comprised between two successive values of $\Omega(2^l)$.

LEMMA 2. – *(Tree version of "Brjuno's counting lemma")*
Let F a decorated forest with r vertices and let $s = \mathrm{gr}(F)$ (we consider here only forests such that $\langle \lambda, \|F\| \rangle \neq 0$). If, for any nonnegative integer k, $N_k(F)$ is the number of complete subtrees t of F that satisfy the following inequality:

$$\frac{1}{2}\Omega(2^{k+1}) \leqslant \langle \lambda, \|t\| \rangle < \frac{1}{2}\Omega(2^k),$$

then

$$N_k(F) \leqslant \begin{cases} 0 \ \text{if} \ s < 2^k \\ 2\nu\frac{s}{2^k} - 1 \ \text{if} \ 2^k \leqslant s. \end{cases}$$

Let us now consider a forest F with r vertices decorated by $\eta_1, \ldots, \eta_r$. Let $s = \mathrm{gr}(F)$ and let l be the integer such that $2^l \leqslant s < 2^{l+1}$. The closed-form expression of the mould (see Proposition 8) is given by:

$$M^F = \prod_{i=1}^{r} \frac{1}{\langle \lambda, \|t_i\| \rangle},$$

where, for $1 \leqslant i \leqslant r$ if t_i is the complete subtree of F whose root is labelled by η_i.

We immediately obtain:

$$|M^F| \leqslant \prod_{k=0}^{l} \left(\frac{2}{\Omega(2^{k+1})} \right)^{N_k(F)} = 2^r \prod_{k=0}^{l} \left(\frac{1}{\Omega(2^{k+1})} \right)^{N_k(F)}.$$

Thanks to the previous lemma, for indices $k \leqslant l$, $N_k(F) \leqslant 2\nu\frac{s}{2^k}$ thus

$$|M^F| \leqslant 2^r \exp\left(\sum_{k=0}^{l} N_k(F) \mathrm{Log}(\frac{1}{\Omega(2^{k+1})}) \right)$$
$$\leqslant 2^r \exp\left(\sum_{k=0}^{l} 2\nu\frac{s}{2^k} \left| \mathrm{Log}(\frac{1}{\Omega(2^{k+1})}) \right| \right)$$
$$\leqslant 2^r \exp(2\nu s S) \leqslant C^{\mathrm{gr}(F)}$$

with $C = 2\exp(2\nu S)$.

The case of diffeomorphisms is settled in exactly the same way. We denote by $N^{\bullet^<}$ the linearizing mould that corresponds to the character χ, and we have the following:

PROPOSITION 12. – *The arborified mould $N^{\bullet^<}$ has a geometric growth: there exists a constant D such that $|N^F| \leqslant D^{\mathrm{gr}(F)}$.*

The proof goes along the same lines as for vector fields, using instead the following closed form:

$$N^F = \prod_{i=1}^{r} \frac{1}{e^{2\pi i \langle \lambda, \|t_i\| \rangle} - 1},$$

where, for $1 \leqslant i \leqslant r$, t_i is the complete subtree of F whose root is labelled by η_i, and applying the relevant counting lemma as in [5].

6.5. The analytic normalization scheme with CK_H^+

Let us recollect now the scheme for linearizing a non resonant dynamical system, using the Hopf algebra CK_H^+:

1. We express the equation regarding the normalizing substitution automorphism as an equation on characters of CK_H^+.
2. We solve this equation, obtaining this way a well-defined character, even for forests displaying "fake resonances" for some of their subtrees.
3. We prove some geometrical growth estimate for this character, by using the Diophantine hypothesis on the spectrum.
4. We match this with the geometric growth for the comould part in the expansion.
5. We obtain a convergent series of operators, which makes it possible to conclude to the analyticity of the transformation thus constructed.

So in fact, strictly speaking, we don't need to arborify moulds, we can work from the outset at the arborescent level, and directly at the level of the algebra CK^+, which is the one that underlies all the computations, and for which no fake obstruction remains. However, plain (i.e., non arborescent) moulds are nevertheless very useful because it is usually easier to guess a closed-form expression for them, before proving that their arborescent counterparts also have a closed-form of the same kind (and this is a very general phenomenon for the use of arborification, cf [25] and [26]).

To dispel any idea that the scheme we have described in the present text is too special and only limited to giving a new proof of already well known results achieved by common methods, let us indicate 2 directions:

– A natural question is the linearization of nonresonant dynamical systems for data of various classes of regularity. In [5], the author proved new results of linearization for diffeomorphisms or vector fields which are formal series with Gevrey growth estimates, under a Brjuno condition. It is straightforward to get the same results with the mould apparatus, using the approach detailed in the present text. The algebraic constructions are *exactly* the same, all the results on the mould side can be used unchanged, the only supplementary thing is to show the geometric growth for the comould part, adapted to spaces of Gevrey series, instead of analytic ones, which is easy. Now, the point is that in order to go beyond such results performed on *formal* spaces of series with some growth conditions, to tackle the

same question for *functional spaces*, e. g. data which are summable in one variable, or multi-summable, or resurgent, the same scheme remains valid in the mould/comould formalism, whereas under other approaches would require ad hoc estimates that would be quite difficult to prove.

– Next we can consider the question of normalization of resonant local dynamical systems; there, linearization is generically not possible using formal series, but there are simple normal forms and the normalizing series are generically divergent ([24], [11]). The substitution automorphisms for the normalizing transformations can be expressed by mould/comould expansions, where the moulds take their values in some algebra $\mathscr{R}$ of resurgent functions [11]. In the presence of diophantine small denominators, the arborification/coarborification machinery is used in the same way as in the present paper; in this Hopf-algebraic presentation, arborification is a factorization of characters from CK to the (commutative) algebra $\mathscr{R}$ and the comould constructions are exactly the same (and CK$_+$ plays an important role, there, too). All the constructions and theorems are already in Ecalle's foundational papers, but with arguments that are very concise; the presentation we give yield easy proofs of algebraic properties of the arborification formalism and makes it possible to connect it to some very recent work in algebraic combinatorics. Applications of arborescent moulds go much further than its original domain of application, namely irregular singularities of local dynamical systems: Stochastic Processes, in particular the theory of rough paths is one striking example (see in particular Section 4.2 of [15], where the concept of extension is *exactly* the factorization of characters as we have formulated it; see also [10]); the fast expanding algebraic theory of non-linear control theory, with Hopf algebraic formulations of (Lie–)Butcher series is another one ([22]).

7. Conclusion

Ecalle's mould-comould formalism has been in the present text given a presentation in terms of some Hopf algebras (Faà di Bruno, shuffle, quasishuffle, Connes-Kreimer, Grossman-Larson...) which are by now standard objects in algebraic combinatorics. In this way, symmetral moulds appear as characters of a decorated shuffle Hopf algebra, and symmetrel ones as characters of a quasishuffle one. Next we have shown that arborification (resp. contracting arborification) of moulds is the outcome of a factorization of characters, by using a universal property satisfied by Connes-Kreimer Hopf algebra.

Then, going to the graded duals, we have been able to *characterize* the fundamental process of *homogeneous coarborification* in a simple way, and consequently easily obtaining justifications of its properties, by building on known facts regarding Grossman-Larson Hopf algebra.

We have introduced a *subalgebra of the decorated Connes-Kreimer algebra* which underlies the calculations of normalization of analytic dynamical systems at singularities. Namely, computing a normalizing transformation will amount to finding a character of this algebra, which satisfies a particular equation that directly comes from the normalization relation itself. In the present paper, we have illustrated the method by the well-known problem of linearization of non-resonant dynamical systems in any dimension, in the presence of small denominators. In problems involving resonances together with small denominators, the same

Hopf-algebraic apparatus governs the calculations and the only thing that changes is that the characters are not scalar any more but take their values in relevant algebras of resurgent functions.

BIBLIOGRAPHY

[1] V. I. Arnol′d, *Geometrical methods in the theory of ordinary differential equations*, Grundl. math. Wiss. **250**, Springer, New York-Berlin, 1983.

[2] A. D. Brjuno, Analytic form of differential equations. I, II, *Trudy Moskov. Mat. Obšč.* **25** (1971), 131–288, **26** (1972), 199–239.

[3] D. Calaque, K. Ebrahimi-Fard, D. Manchon, Two interacting Hopf algebras of trees: a Hopf-algebraic approach to composition and substitution of B-series, *Adv. in Appl. Math.* **47** (2011), 282–308.

[4] J. F. Cariñena, K. Ebrahimi-Fard, H. Figueroa, J. M. Gracia-Bondía, Hopf algebras in dynamical systems theory, *Int. J. Geom. Methods Mod. Phys.* **4** (2007), 577–646.

[5] T. Carletti, The Lagrange inversion formula on non-Archimedean fields. Non-analytical form of differential and finite difference equations, *Discrete Contin. Dyn. Syst.* **9** (2003), 835–858.

[6] F. Chapoton, The anticyclic operad of moulds, *Int. Math. Res. Not.* **2007** (2007), Art. ID rnm078.

[7] F. Chapoton, F. Hivert, J.-C. Novelli, J.-Y. Thibon, An operational calculus for the mould operad, *Int. Math. Res. Not.* **2008** (2008), Art. ID rnn018.

[8] A. Connes, D. Kreimer, Hopf algebras, renormalization and noncommutative geometry, *Comm. Math. Phys.* **199** (1998), 203–242.

[9] K. Ebrahimi-Fard, D. Manchon, The combinatorics of Bogoliubov's recursion in renormalization, in *Renormalization and Galois Theories* (A. Connes, F. Fauvet, J.-P. Ramis, eds.), EMS Publ. House, 2009.

[10] K. Ebrahimi-Fard, A. Lundervold, S. J. A. Malham, H. Munthe-Kaas, A. Wiese, Algebraic structure of stochastic expansions and efficient simulation, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **468** (2012), 2361–2382.

[11] J. Écalle, Singularités non abordables par la géométrie, *Ann. Inst. Fourier (Grenoble)* **42** (1992), 73–164.

[12] H. Figueroa, J. M. Gracia-Bondía, Combinatorial Hopf algebras in quantum field theory. I, *Rev. Math. Phys.* **17** (2005), 881–976.

[13] L. Foissy, Les algèbres de Hopf des arbres enracinés décorés. I, *Bull. Sci. Math.* **126** (2002), 193–239.

[14] L. Foissy, Les algèbres de Hopf des arbres enracinés décorés. II, *Bull. Sci. Math.* **126** (2002), 249–288.

[15] L. Foissy, J. Unterberger, Ordered forests, permutations, and iterated integrals, *Int. Math. Res. Not.* **2013** (2013), 846–885.

[16] G. Galavotti, Exact renormalization group, IHP Séminaire Bourbaphy, CNRS, 2002.

[17] R. Grossman, R. G. Larson, Hopf-algebraic structure of families of trees, *J. Algebra* **126** (1989), 184–210.

[18] R. Grossman, R. G. Larson, Hopf-algebraic structure of combinatorial objects and differential operators, *Israel J. Math.* **72** (1990), 109–117.

[19] M. E. Hoffman, (Non)commutative Hopf algebras of trees and (quasi)symmetric functions, in *Renormalization and Galois theories*, IRMA Lect. Math. Theor. Phys. **15**, Eur. Math. Soc., Zürich, 2009, 209–227.

[20] Y. Ilyashenko, S. Yakovenko, *Lectures on analytic differential equations*, Graduate Studies in Math. **86**, Amer. Math. Soc., Providence, RI, 2008.

[21] D. Kreimer, Chen's iterated integral represents the operator product expansion, *Adv. Theor. Math. Phys.* **3** (1999), 627–670.

[22] A. Lundervold, H. Z. Munthe-Kaas, On algebraic structures of numerical integration on vector spaces and manifolds, in *Faà di Bruno Hopf algebras, Dyson-Schwinger equations, and Lie-Butcher series*, IRMA Lect. Math. Theor. Phys. **21**, Eur. Math. Soc., Zürich, 2015, 219–263.

[23] D. Manchon, Hopf algebras, from basics to applications to renormalization, in *Comptes-rendus des rencontres mathématiques de Glanon*, 2003.

[24] J. Martinet, J.-P. Ramis, Problèmes de modules pour des équations différentielles non linéaires du premier ordre, *Publ. Math. IHÉS* **55** (1982), 63–164.

[25] F. Menous, An example of local analytic q-difference equation: analytic classification, *Ann. Fac. Sci. Toulouse Math.* **15** (2006), 773–814.

[26] F. Menous, On the stability of some groups of formal diffeomorphisms by the Birkhoff decomposition, *Adv. Math.* **216** (2007), 1–28.

[27] A. Murua, The Hopf algebra of rooted trees, free Lie algebras, and Lie series, *Found. Comput. Math.* **6** (2006), 387–426.

[28] F. Panaite, Relating the Connes-Kreimer and Grossman-Larson Hopf algebras built on rooted trees, *Lett. Math. Phys.* **51** (2000), 211–219.

[29] J. Pöschel, On invariant manifolds of complex analytic mappings near fixed points, *Exposition. Math.* **4** (1986), 97–109.

[30] H. Rüssmann, On the convergence of power series transformations of analytic mappings near a fixed point into normal form, preprint IHÉS, 1977.

[31] W. Zhao, Differential operator specializations of noncommutative symmetric functions, *Adv. Math.* **214** (2007), 639–665.

[32] W. Zhao, Noncommutative symmetric systems over associative algebras, *J. Pure Appl. Algebra* **210** (2007), 363–382.

[33] W. Zнао, *A noncommutative symmetric system over the Grossman-Larson Hopf algebra of labeled rooted trees*, *J. Algebraic Combin.* **28** (2008), 235–260.

Frédéric Fauvet
Université de Strasbourg
IRMA
UMR 7501 du CNRS
7 rue René-Descartes
67084 Strasbourg Cedex, France
E-mail: `frederic.fauvet@gmail.com`

Frédéric Menous
Laboratoire de Mathématiques d'Orsay
Université Paris-Sud
CNRS
Université Paris-Saclay
91405 Orsay, France
E-mail: `frederic.menous@math.u-psud.fr`