

CONTEMPORARY MATHEMATICS

539

Combinatorics and Physics

Mini-Workshop on Renormalization
December 15–16, 2006

Conference on Combinatorics and Physics
March 19–23, 2007
Max-Planck-Institut für Mathematik
Bonn, Germany

Kurusch Ebrahimi-Fard
Matilde Marcolli
Walter D. van Suijlekom
Editors



American Mathematical Society

Combinatorics and Physics

This page intentionally left blank

CONTEMPORARY MATHEMATICS

539

Combinatorics and Physics

Mini-Workshop on Renormalization
December 15–16, 2006

Conference on Combinatorics and Physics
March 19–23, 2007
Max-Planck-Institut für Mathematik
Bonn, Germany

Kurusch Ebrahimi-Fard
Matilde Marcolli
Walter D. van Suijlekom
Editors



American Mathematical Society
Providence, Rhode Island

Editorial Board

Dennis DeTurck, managing editor

George Andrews Abel Klein Martin J. Strauss

2000 *Mathematics Subject Classification*. Primary 81T15, 65D30.

Library of Congress Cataloging-in-Publication Data

Mini-Workshop on Renormalization (2006 : Max-Planck-Institut für Mathematik)

Combinatorics and Physics : Mini-Workshop on Renormalization, December 15–16, 2006, Max-Planck-Institut für Mathematik, Bonn, Germany : Conference on Combinatorics and Physics, March 19–23, 2007, Max-Planck-Institut für Mathematik, Bonn, Germany / Kurusch Ebrahimi-Fard, Matilde Marcolli, Walter D. van Suijlekom, editors.

p. cm. — (Contemporary mathematics ; v. 539)

Includes bibliographical references.

ISBN 978-0-8218-5329-0 (alk. paper)

1. Renormalization group—Congresses. 2. Quantum field theory—Congresses. 3. Numerical integration—Congresses. I. Ebrahimi-Fard, Kurusch, 1973– II. Marcolli, Matilde. III. Suijlekom, Walter D. van., 1978– IV. Max-Planck-Institut für Mathematik. V. Conference on Combinatorics and Physics (2007 : Max-Planck-Institut für Mathematik) VI. Title.

QC20.7.R43M56 2006

530.14'3—dc22

2010048012

Copying and reprinting. Material in this book may be reproduced by any means for educational and scientific purposes without fee or permission with the exception of reproduction by services that collect fees for delivery of documents and provided that the customary acknowledgment of the source is given. This consent does not extend to other kinds of copying for general distribution, for advertising or promotional purposes, or for resale. Requests for permission for commercial use of material should be addressed to the Acquisitions Department, American Mathematical Society, 201 Charles Street, Providence, Rhode Island 02904-2294, USA. Requests can also be made by e-mail to reprint-permission@ams.org.

Excluded from these provisions is material in articles for which the author holds copyright. In such cases, requests for permission to use or reprint should be addressed directly to the author(s). (Copyright ownership is indicated in the notice in the lower right-hand corner of the first page of each article.)

© 2011 by the American Mathematical Society. All rights reserved.

The American Mathematical Society retains all rights
except those granted to the United States Government.

Copyright of individual articles may revert to the public domain 28 years
after publication. Contact the AMS for copyright status of individual articles.

Printed in the United States of America.

⊗ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at <http://www.ams.org/>

10 9 8 7 6 5 4 3 2 1 16 15 14 13 12 11

Contents

| | |
|--|-----|
| Preface | vii |
| List of participants | ix |
| One-particle irreducibility with initial correlations CHRISTIAN BROUDER and FRÉDÉRIC PATRAS | 1 |
| Multiple zeta values and periods: From moduli spaces to Feynman integrals FRANCIS BROWN | 27 |
| From quantum electrodynamics to posets of planar binary trees FRÉDÉRIC CHAPOTON and ALESSANDRA FRABETTI | 53 |
| Sweedler's duals and Schützenberger's calculus GÉRARD H. E. DUCHAMP and CHRISTOPHE TOLLU | 67 |
| Primitive elements of the Hopf algebra of free quasi-symmetric functions L. FOISSY | 79 |
| A Renormalisation Group approach to Stochastic Löwner Evolutions ROLAND FRIEDRICH | 89 |
| On the causal gauge principle JOSÉ M. GRACIA-BONDÍA | 115 |
| Abstract integration, combinatorics of trees and differential equations MASSIMILIANO GUBINELLI | 135 |
| Rooted trees appearing in products and co-products RALF HOLTKAMP | 153 |
| Magnus expansions and beyond ARIEH ISERLES | 171 |
| Wilsonian renormalization, differential equations and Hopf algebras T. KRAJEWSKI and P. MARTINETTI | 187 |
| Algebraic analysis of non-renormalization theorems in supersymmetric field theories E. KRAUS | 237 |
| Not so non-renormalizable gravity DIRK KREIMER | 247 |
| Renormalised multiple zeta values which respect quasi-shuffle relations DOMINIQUE MANCHON | 255 |

| | |
|---|-----|
| Formulas for the Connes–Moscovici Hopf algebra FRÉDÉRIC MENOUS | 269 |
| Hopf algebras and the combinatorics of connected graphs in quantum field theory ÂNGELA MESTRE and ROBERT OECKL | 287 |
| Hopf Algebras of Formal Diffeomorphisms and Numerical Integration on Manifolds ALEXANDER LUNDERVOLD and HANS MUNTHE-KAAS | 295 |
| A combinatorial and field theoretic path to quantum gravity: The new challenges of group field theory D. ORITI | 325 |
| Noncommutative formal Taylor expansions and second quantised regularised traces SYLVIE PAYCHA | 349 |
| Motives: An introductory survey for physicists ABHIJNAN REJ (WITH AN APPENDIX BY MATILDE MARCOLLI) | 377 |
| Combinatorics and Feynman graphs for gauge theories WALTER D. VAN SUIJLEKOM | 417 |
| Multi-scale Analysis and Non-commutative Field Theory FABIEN VIGNES-TOURNERET | 443 |

Preface

These proceedings are based on two meetings held at the Max-Planck-Institut für Mathematik in Bonn, Germany. Both events were centered around algebraic combinatorics in theoretical physics. The short workshop *Renormalization* was held in December 2006 and aimed at bringing together leading specialists to discuss most recent progress in the understanding of mathematical structures—especially algebro-combinatorial ones—in perturbative renormalization of quantum field theory. It included talks given by D. Kreimer, E. Kraus and J. M. Gracia-Bondía. The positive resonance experienced at this meeting led to the follow-up conference *Combinatorics and Physics* held in March 2007. This one week event was broader in its aims and focused on algebraic and combinatorial aspects in modern physics. Besides a rich spectrum of research talks, it included four mini-courses given by leading researchers with, moreover, an acknowledged reputation for being gifted lecturers.

In order to put the two meetings—and this proceedings—into context, let us start with some historical remarks. In the last ten years, a convergence has taken place between seemingly separate fields in mathematics and physics: noncommutative geometry, renormalization in perturbative quantum field theory, and numerical integration of dynamical systems and flows on manifolds. It turned out that these fields share a common (co)algebraic formalism with as key words: combinatorics on trees, Hopf algebras and Lie series. For instance, in the 2007 meeting this is reflected in the talks given by A. Connes, F. Brown, A. Iserles and H. Munthe-Kaas, ranging from noncommutative geometry and multiple zeta values to Lie-Butcher theory in the context of numerical integration methods.

One of the main notions that appear in this context is the one of combinatorial Hopf algebras. Paramount in these developments was the Connes–Kreimer application of unordered rooted trees and the related combinatorial Hopf algebras to renormalization in perturbative quantum field theory, as well as the discovery of truly unexpected connections to the famous Riemann–Hilbert problem, which was eventually addressed in full detail by Connes and Marcolli. In fact, based on the seminal findings of D. Kreimer in the late 1990s, uncovering the fundamental Hopf algebra structure of Feynman rules, the path breaking work of Connes and Kreimer created a new common and fruitful playground for mathematicians and theoretical physicists alike. It reinforced interest in algebro-combinatorial aspects in renormalization theory, culminating into a whole industry of research that lies at the interface of general quantum field theory and advanced algebra and combinatorial methods. Many foundational aspects in this new field have been elaborated, raising new questions which stand at the forefront of current research.

These meetings could not have taken place without the complete financial, logistic and organisational support from the Max-Planck-Institut für Mathematik. We thank Gerd Faltings and Don Zagier for their encouragement and interest in these meetings. We are deeply indebted to Julia Löwenstein and Dagobert Jarisch of the administration of Max-Planck-Institut für Mathematik. Without their professional help and unlimited effort these meetings would not have been possible.

We thank the speakers for their valuable contributions and all participants for their enthusiasm. Together this created an exciting atmosphere resulting in two memorable events. We especially thank Pierre Cartier, Predrag Cvitanović, Philippe di Francesco and Frédéric Patras for giving interesting and stimulating mini-courses.

September 2010

Kurusch Ebrahimi-Fard
Matilde Marcolli
Walter D. van Suijlekom

List of participants

| | |
|------------------------|-------------------------|
| Christian Brouder | Sylvie Paycha |
| Francis Brown | Abhijnan Rej |
| Pierre Cartier | Walter D. van Suijlekom |
| Alain Connes | Christophe Tollu |
| Predrag Cvitanović | Xavier Viennot |
| Gérard H.E. Duchamp | Fabien Vignes-Tourneret |
| Kurusch Ebrahimi-Fard | Raimar Wulkenhaar |
| Loïc Foissy | Don Zagier |
| Alessandra Frabetti | Paul Zinn-Justin |
| Philippe Di Francesco | |
| Roland Friedrich | |
| Hanno Gottschalk | |
| José M. Gracia-Bondía | |
| Massimiliano Gubinelli | |
| Ralf Holtkamp | |
| Arieh Iserles | |
| Thomas Krajewski | |
| Elisabeth Kraus | |
| Dirk Kreimer | |
| Jean-Louis Loday | |
| Dominique Manchon | |
| Matilde Marcolli | |
| Ângela Mestre | |
| Hans Munthe-Kaas | |
| Daniele Oriti | |
| Frédéric Patras | |

This page intentionally left blank

One-particle irreducibility with initial correlations

Christian Brouder and Frédéric Patras

ABSTRACT. In quantum field theory (QFT), the vacuum expectation of a normal product of creation and annihilation operators is always zero. This simple property paves the way to the classical treatment of perturbative QFT. This is no longer the case in the presence of initial correlations, that is if the vacuum is replaced by a general state. As a consequence, the combinatorics of correlated systems such as the ones occurring in many-body physics is more complex than that of quantum field theory and the general theory has made very slow progress. Similar observations hold in statistical physics or quantum probability for the perturbation series arising from the study of non Gaussian measures. In this work, an analysis of the Hopf algebraic aspects of quantum field theory is used to derive the structure of Green functions in terms of connected and one-particle irreducible Green functions for perturbative QFT in the presence of initial correlations.

1. Introduction

In quantum field theory (QFT), the initial state is most often the vacuum. Many quantum field concepts, such as Feynman diagrams, the Dyson equation and the Bethe-Salpeter equation rest on the special properties of the vacuum. These desirable concepts extend to special states called *quasi-free states* [1, 2, 3].

For general initial states, it is not possible to write the Green function in terms of standard Feynman diagrams and the structure of the Green functions is more complex. For example, the Dyson and Bethe-Salpeter equations do not hold. The Dyson equation describes the structure of the two-point interacting Green function. Its extension to non-quasi-free states was discovered by Hall [4] in 1975. However, the equivalent structural equations for $2n$ -point interacting Green functions (with $n > 1$) is not known for general initial states. The determination of this structure is the main purpose of the present article, with a particular emphasis on the notion of one-particle-irreducibility in this framework.

We stress that the problem of the calculation of Green functions for initial states that are not quasi-free has important applications. For instance, many highly-correlated materials contain transition metals where states of the $3d$ shell are degenerate. The consequence of this degeneracy is that a small external perturbation can create a very strong change in the state of the system. For instance, a small external magnetic field induces a strong variation in resistance (giant magnetoresistance), that is used to build high-density storage disks. The knowledge of

the Green functions would enable us to calculate accurately the properties of such materials.

Although we will use mainly the language of QFT, let us mention that the situation where the vacuum is the initial state corresponds in statistical physics to the case of Euclidean measures as showing up e.g. in Euclidean Quantum Field Theory. This is because of the joint use of the Wick theorem (in its simplest, Gaussian, form). For general measures, the problem of determining the fine combinatorial structure of perturbation series was addressed in [5] where, in particular, the combinatorics of truncated (or connected) moment functions was studied intensively. We refer to this article, also for further motivations and examples of computations of thermodynamic limits involving the use of generalized Feynman diagrams as we also use them.

Concretely, in the present article, we investigate the structure of Green functions with Hopf algebraic methods. Hopf algebras have been implicitly used for a long time in statistical physics and quantum field theory: Ruelle [6], Borchers [7] and Stora [8] used a product that is called the *convolution product* in the Hopf language. Wightman and Challifour [9] defined a *triple dot product* that was rediscovered only much later in quantum chemistry [10] and was given an algebraic meaning in [11], where it was called *adapted normal product*. The work of Wightman and Challifour was summarized by Stora in ref. [8].

The explicit introduction of Hopf algebras by Kreimer and Connes at the level of trees and Feynman diagrams [12, 13] sparked a reformulation of many quantum field constructions (renormalization [14, 15, 16], Wick's theorem [17], quantization [11, 18], structure of Green functions [19, 20], gauge theory [21]). The result of these efforts is a reasonably complete presentation of QFT in terms of Hopf algebraic concepts [22]. Hopf algebras, which are powerful tools to solve combinatorial problems [23], could be expected to help also in the presence of a general state. Indeed, the use of Hopf algebraic methods resulted in the determination of the equation of motion of the Green functions [24] and the description of the relations between general and connected Green functions in the presence of a general state [22].

The relation between connected and one-particle irreducible (1PI) Green functions, which is the main topic of the present article, is more difficult to understand and depends on the very definition of when a diagram is irreducible. Here, we show that a rather natural definition leads to a complete description of connected Green functions in terms of 1PI Green functions.

We should point out that our main long-term interest is in the study of the electronic structure of highly correlated materials by means of Green functions, also in the framework of classical (non-relativistic) quantum chemistry. However, we restrict here our study of one-particle-irreducibility to the particular case of local potentials. This covers the potentials of quantum field theories -including QED, which is meaningful for our long-term purposes-, but not the Coulomb potential. There are several ways to remedy this problem and extend the constructions in the present article to time-dependent perturbation theory in many-body physics. For example, as the referee pointed out, our results may be extended to the case of general interactions by means of the notion of block truncation as introduced

in [5]. However, for the sake of simplicity -and since the combinatorics of one-particle-irreducibility is already intricate enough for local potentials, we decided to postpone the study of electronic systems (and of non-scalar fields) to further work.

The paper starts with a short introduction to Green functions in the presence of initial correlations and to Hopf algebras, followed by the definition of quantum field forms and their convolution logarithm. Then, the relation between forms and connected forms is made explicit, providing the classical relation between general Green functions and connected Green functions. To discuss 1PI functions, we need to generalize a recent work by Mestre and Oeckl [19, 20]. Then, a rather natural definition of 1PI functions will be proposed and the Mestre-Oeckl approach will be used to write a connected Green function in terms of these 1PI functions. In the process, universal properties of symmetric functions with respect to Hopf algebra derivations are put forward.

2. Green functions with initial correlations

In the present section, we fix the notation and briefly survey the definition of Green functions, emphasizing the role of the initial state.

2.1. Field operators. Quantum fields are operator-valued distributions acting on a Fock space [25]. Here, we describe the construction of Fock space, creation and annihilation operators and the corresponding quantum fields. We start from a self-adjoint operator h acting on a Hilbert space \mathcal{H} and, for notational convenience, we assume that h has a pure point spectrum, so that there is an orthonormal basis $|e_i\rangle$ (with $i \in I$) of \mathcal{H} consisting of eigenvectors of h . In many applications, the Hilbert space \mathcal{H} is a function space and the eigenvectors (written then preferably with the functional notation ϕ_n) are functions of \mathbf{r} (where \mathbf{r} is a point in three-dimensional space).

The tensor product of Hilbert spaces is well defined (see Ref. [26] p. 49) and the *symmetric Fock space* over \mathcal{H} is the Hilbert space $S(\mathcal{H}) = \bigoplus_{N=0}^{\infty} S^N(\mathcal{H})$, where $S^N(\mathcal{H})$ is the N -fold symmetric tensor product of \mathcal{H} . An orthogonal basis of the vector space $S^N(\mathcal{H})$ is provided by the set of vectors

$$(1) \quad |e_{i_1}\rangle \vee \cdots \vee |e_{i_N}\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma} |e_{i_{\sigma(1)}}\rangle \otimes \cdots \otimes |e_{i_{\sigma(N)}}\rangle,$$

where σ runs over the permutations of N elements and where (i_1, \dots, i_N) runs over the subset of I^N such that $i_1 \leq \cdots \leq i_N$. In this formula, the symbol \vee denotes the symmetric product and $1/\sqrt{N!}$ is a normalization factor.

In many-body theory, $S^N(\mathcal{H})$ is called the N -particle space of the system and its elements are the N -particle states. In particular, $S^0(\mathcal{H})$ is a one-dimensional vector space denoted by $\mathbb{C}1$ in the mathematical literature. In many-body physics and quantum field theory, this unit 1 of the tensor product is denoted by $|0\rangle$, this is the vacuum of the theory (i.e. the state without a particle).

The creation operator a_n^\dagger is defined as the linear map from $S(\mathcal{H})$ to itself such that, for any basis vector $|u\rangle$ of $S(\mathcal{H})$, $a_n^\dagger|u\rangle = c_n(u)|e_n\rangle \vee |u\rangle$, where $c_n(u)$ is a normalization factor (see for example [27]). It is called a creation operator because it maps $S^N(\mathcal{H})$ to $S^{N+1}(\mathcal{H})$: it adds a new particle to a N -particle state. Its adjoint a_n is called an annihilation operator. The normalization factor ensures that the commutation relation $a_m \circ a_n^\dagger - a_n^\dagger \circ a_m = \delta_{nm}$ holds, where \circ denotes the composition of operators.

In functional notation, the corresponding quantum field is the (self-adjoint) operator-valued distribution on the three-dimensional space:

$$\varphi_S(\mathbf{r}) = \sum_{n \in I} \phi_n(\mathbf{r}) a_n + \phi_n^*(\mathbf{r}) a_n^\dagger.$$

This formalism is used to describe scalar particles or photons (up to an additional vector index in that case; recall that a self-adjoint field operator describes a neutral particle, charged scalar or fermion field operators are not self-adjoint). We remind that we focus in the present article on scalar particles (without a charge).

2.2. Adiabatic limit. The adiabatic limit is a very general way of solving the Schrödinger equation for a system described by the Hamiltonian $H = H_0 + V$ where the eigenstates of H_0 are known but not those of H . The general idea behind the technique is that, for a particle evolving in a potential V , the effect of the potential on the motion can be treated (or can be expected to be treated) perturbatively. As we mentioned, besides the perturbative expansions of QFT, the same general idea shows up in the perturbation series of statistical physics [5].

As far as adiabatic limits are concerned, the basic idea is quite simple. We define a time-dependent Hamiltonian $H(t) = H_0 + e^{-\epsilon|t|}V$. When ϵ is small, the interaction $H(t)$ is very slowly switched on from $t = -\infty$ where $H(-\infty) = H_0$ to $t = 0$ where $H(0) = H$. It is hoped that, if ϵ is small enough, then an eigenstate of H_0 is transformed into an eigenstate of H .

To implement this picture, the time-dependent Schrödinger equation given by $i\partial|\Psi_S(t)\rangle/\partial t = H(t)|\Psi_S(t)\rangle$ is solved. However, the solution $|\Psi_S(t)\rangle$ is not convenient because it has no limit when $t \rightarrow -\infty$. Therefore, we define $|\Psi(t)\rangle = e^{iH_0 t}|\Psi_S(t)\rangle$ that satisfies $i\partial|\Psi(t)\rangle/\partial t = H_{\text{int}}(t)|\Psi(t)\rangle$ with respect to $H_{\text{int}}(t) = e^{iH_0 t}V e^{-iH_0 t}e^{-\epsilon|t|}$. Now $H_{\text{int}}(-\infty) = 0$ and $|\Psi(-\infty)\rangle$ makes sense. Using H_{int} , we can start from the ground state $|\Phi_0\rangle$ of H_0 and solve the time-dependent Schrödinger equation with the boundary condition $|\Psi(-\infty)\rangle = |\Phi_0\rangle$. When no eigenvalue crossing takes place, $|\Phi_0\rangle$ should be transformed into the ground state $|\Psi(0)\rangle$ of H .

Instead of calculating directly $|\Psi(t)\rangle$ it is convenient to define the unitary operator $U(t)$ as the solution of $i\partial U(t)/\partial t = H_{\text{int}}(t)U(t)$, with the boundary condition $U(-\infty) = 1$. Thus, $|\Psi(t)\rangle = U(t)|\Phi_0\rangle$. Note that $U(t)$ depends on ϵ , as $H_{\text{int}}(t)$. But is $\lim_{\epsilon \rightarrow 0} U(0)|\Phi_0\rangle$ an eigenstate of H ? It would if the limit existed, but it does not [30]. However, Gell-Mann and Low [28] discovered in 1951 that

$$|\Psi_{\text{GL}}\rangle = \lim_{\epsilon \rightarrow 0} \frac{U(0)|\Phi_0\rangle}{\langle\Phi_0|U(0)|\Phi_0\rangle}$$

exists and is an eigenstate of H . A mathematical proof of this fact for reasonable Hamiltonians came much later [29]. Notice that the above scheme works when the ground state of H_0 is non degenerate. When it is degenerate, the problem is more subtle [33, 34] and the limit $\epsilon \rightarrow 0$ only exists when $|\Phi_0\rangle$ is properly chosen [31, 32].

2.3. Green functions. We now come to the heart of QFT: the calculation of Green functions (or moment functions). Green functions are important because they allow for the calculation of practically all relevant physical observables: energy, charge density, transport coefficients, current density, dielectric constants, etc. In particular, they show up naturally in the perturbative expansions arising from

adiabatic limits. If we could calculate Green functions exactly, we would know all interesting properties of matter. Of course, as far as many-body theory is concerned, we cannot calculate exact Green functions for realistic materials, but non-perturbative approximations are now used with great success [35].

When the dynamics of the particles is described by a one-body Hamiltonian H_0 , the n -point Green function for scalar particles is defined by

$$G_n^0(x_1, \dots, x_n) = \langle \Phi_0 | T(\varphi(x_1) \dots \varphi(x_n)) | \Phi_0 \rangle,$$

where $x = (t, \mathbf{r})$, T is the time-ordering operator and $\varphi(x)$ is related to $\varphi_S(\mathbf{r})$ by

$$\varphi(x) = e^{iH_0 t} \varphi_S(\mathbf{r}) e^{-iH_0 t} = \sum_{n \in I} e^{-i\epsilon_n t} \phi_n(\mathbf{r}) a_n + e^{i\epsilon_n t} \phi_n^*(\mathbf{r}) a_n^\dagger,$$

where the $\phi_n(\mathbf{r})$ are eigenvectors of H_0 with associated eigenvalues ϵ_n . The time-ordering operator orders the quantum fields $\varphi(x_1), \dots, \varphi(x_n)$ so that the field $\varphi(x_i)$ is on the left of $\varphi(x_j)$ if t_i is greater (i.e. later) than t_j . For example $T(\varphi(x_1)\varphi(x_2)) = \varphi(x_1)\varphi(x_2)$ if $t_1 > t_2$ and $T(\varphi(x_1)\varphi(x_2)) = \varphi(x_2)\varphi(x_1)$ if $t_1 < t_2$.

When the dynamics of the particles is described by a Hamiltonian $H = H_0 + V$, where H_0 is one-body, the expression for the Green function becomes [36, 27]

$$G_n(x_1, \dots, x_n) = \frac{\langle \Phi_0 | T(\varphi(x_1) \dots \varphi(x_n) e^{-i \int H_{\text{int}}(t) dt}) | \Phi_0 \rangle}{\langle \Phi_0 | T(e^{-i \int H_{\text{int}}(t) dt}) | \Phi_0 \rangle},$$

where $H_{\text{int}}(t) = e^{iH_0 t} V e^{-iH_0 t} e^{-\epsilon|t|}$ and the limit $\epsilon \rightarrow 0$ is implicitly taken. This generalizes to non-scalar particles and, in the example of the non-relativistic electrons,

$$H_{\text{int}}(t) = e^{-\epsilon|t|} \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \psi^\dagger(t, \mathbf{r}) \psi^\dagger(t, \mathbf{r}') V_{ee}(\mathbf{r} - \mathbf{r}') \psi(t, \mathbf{r}') \psi(t, \mathbf{r}).$$

We are now ready to enter correlated systems. Assuming that the initial state $|\Phi_0\rangle$ is the free field vacuum $|0\rangle$ implies the classical expansion of Green functions in terms of Feynman propagators and, ultimately, of (usual) Feynman diagrams. In many-body theory, the use of this decomposition of Green function into sums of (usual) Feynman diagrams is restricted to very specific states $|\Phi_0\rangle$ called *quasi-free states*. For the other states (or, equivalently, for the study of general functional measures in statistical physics), the structure of Green functions is more complex. Let us give a simple example. We can define the quantity $D_4(x_1, x_2, x_3, x_4)$ by

$$\begin{aligned} G_4^0(x_1, x_2, x_3, x_4) = & G_2^0(x_1, x_2) G_2^0(x_3, x_4) + G_2^0(x_1, x_3) G_2^0(x_2, x_4) \\ & + G_2^0(x_1, x_4) G_2^0(x_2, x_3) + D_4(x_1, x_2, x_3, x_4). \end{aligned}$$

When the initial state is the vacuum or a quasi-free state, the term D_4 is zero. For a general initial state, it is not.

For a fermionic system, a term D_4 can be defined similarly. It is absent when the ground state of H_0 can be written as a Slater determinant. It is present when the ground state of H_0 is degenerate, as in open shell systems¹. In that case, $|\Phi_0\rangle$ can be written as a linear combination of Slater determinants and D_4 describes the correlation between these determinants. The presence of several Slater determinants in the initial state is rather catastrophic for many-body theory. Yaris and

¹For non-equilibrium systems, additional complications come from the fact that time-ordered products must be defined over a closed time path [37]. However, this does not change the combinatorial aspects of the problem.

Taylor summarized the situation [38]: “The inability to handle open-shell systems is a ubiquitous problem in many-body theory. It basically arises when one cannot find a single-determinant unperturbed ground state which connects to the exact ground state when the residual interaction is adiabatically switched on. When this situation holds, one cannot properly define occupied and unoccupied single-particle states, Wick’s theorem does not hold, and Dyson equations, Bethe-Salpeter equations, etc. do not exist.” To this list one can add that ordinary Feynman diagrams and the Gell-Mann and Low formula are lost. In other words, most of the tools of quantum field theory break down. Since the seminal work by Bloch and Horowitz in 1958 [39], many works were devoted to the rebuilding of these tools. Morita discovered a modified version of the Gell-Mann and Low theorem [40], Fujita defined generalized Feynman diagrams [41], Hall derived a Dyson equation for degenerate systems [4]. Since then, progress has been quite slow because of the combinatorial complexity of the problem.

To illustrate this complexity, we first describe the generalized Feynman diagrams introduced by Fujita [41], Hall [4] and Djah et al. [5]. For bosonic and fermionic systems, D_4 can be thought of as a sort of 4-point Feynman propagator, as $D_2(x, y) = G_2^0(x, y)$ is the 2-point Feynman propagator. We shall see that D_4 plays the role of a cumulant, as in the decomposition of a distribution function. Higher order Green functions G_{2n}^0 give rise to higher order propagators D_{2n} and the precise relation between them will be described in the following. In standard quantum field theory, the Green function of the interacting system can be written by adding all possible Feynman diagrams involving the two-point propagator D_2 . When the initial state is not quasi-free, the Green function is written as the sum of all possible Feynman diagrams involving 2-point, 4-point, and $2n$ -point propagators for arbitrary n . An example will be given in figure 2 of this paper.

3. Hopf algebra

We do not provide the general definition of a Hopf algebra (see e.g. [42]) and consider only the special case of the symmetric Hopf algebra $S(V) = \bigoplus_n S^n(V) = \bigoplus_n V^{\otimes n} / \mathcal{S}_n$, where V is a complex vector space and where \mathcal{S}_n , the symmetric group of order n , acts by permutation on the components of the tensor power $V^{\otimes n}$. The commutative product of $S(V)$ is denoted by concatenation. The counit is the linear map $\varepsilon : S(V) \rightarrow \mathbb{C}$ defined by $\varepsilon(1) = 1$, $\varepsilon(u) = 0$ if $u \in S^n(V)$ with $n > 0$. The coproduct is the linear map $\Delta : S(V) \rightarrow S(V) \otimes S(V)$ determined by $\Delta 1 = 1 \otimes 1$, $\Delta a = 1 \otimes a + a \otimes 1$ for $a \in V$ and $\Delta(uv) = (\Delta u)(\Delta v)$, for u and v in $S(V)$. We employ the strengthened Sweedler notation for the coproduct [43]: $\Delta u = u_{(1)} \otimes u_{(2)}$. Recall that there is an implicit summation in the notation, which does not lead to ambiguities when handled correctly: the right hand side should be understood not as the mere tensor product of two elements in $S(V)$ but as a sum of such elements (so that e.g. $a_{(1)} \otimes a_{(2)} = 1 \otimes a + a \otimes 1$ for $a \in V$). More generally, an expression such as $u_{(1)}v_{(1)} \otimes u_{(2)}v_{(2)}$, which stands for $\Delta(u)\Delta(v) = \Delta(uv)$, contains an implicit double summation and should be understood as: $(u_{(1)} \otimes u_{(2)})(v_{(1)} \otimes v_{(2)})$, and similarly for expressions of higher orders.

The *iterated coproducts* Δ^k are defined by $\Delta^0 = \text{id}$, $\Delta^1 = \Delta$ and $\Delta^{k+1} = (\text{id}^{\otimes k} \otimes \Delta)\Delta^k$. Their action on an element u of $S(V)$ is denoted by $\Delta^k u = u_{(1)} \otimes \cdots \otimes u_{(k+1)}$. For any $u \in S(V)$, the *reduced coproduct* is the map $\underline{\Delta} : S(V) \rightarrow$

$S(V) \otimes S(V)$ such that $\underline{\Delta}u = \Delta u - 1 \otimes u - u \otimes 1$. The iterated reduced coproducts $\underline{\Delta}^k$ are defined by $\underline{\Delta}^0 = \text{id}$, $\underline{\Delta}^1 = \underline{\Delta}$ and $\underline{\Delta}^{k+1} = (\text{id}^{\otimes k} \otimes \underline{\Delta})\underline{\Delta}^k$. Their action on an element u of $S(V)$ is denoted by $\underline{\Delta}^k u = u_{(1)} \otimes \cdots \otimes u_{(k+1)}$. The coproduct and the reduced coproduct are cocommutative, that is:

$$\Delta(u) = u_{(1)} \otimes u_{(2)} = u_{(2)} \otimes u_{(1)}, \quad \underline{\Delta}(u) = u_{(1)} \otimes u_{(2)} = u_{(2)} \otimes u_{(1)}.$$

The coproduct is an algebra morphism, but the reduced coproduct is not. Its relation with the product is described by the following simple and useful lemma.

LEMMA 3.1. *If $a \in V$ and $u \in S(V)$, then*

$$\underline{\Delta}(au) = a \otimes u + u \otimes a + au_{(1)} \otimes u_{(2)} + u_{(1)} \otimes au_{(2)},$$

and, for $k > 1$,

$$\underline{\Delta}^k(au) = a \otimes \underline{\Delta}^{k-1}u + au_{(1)} \otimes \underline{\Delta}^{k-1}u_{(2)} + u_{(1)} \otimes \underline{\Delta}^{k-1}(au_{(2)}).$$

More explicitly, for $k > 0$,

$$\begin{aligned} \underline{\Delta}^k(au) &= \sum_{i=1}^{k+1} u_{(1)} \otimes \cdots \otimes u_{(i-1)} \otimes a \otimes u_{(i)} \otimes \cdots \otimes u_{(k)} \\ &\quad + \sum_{i=1}^{k+1} u_{(1)} \otimes \cdots \otimes u_{(i-1)} \otimes au_{(i)} \otimes u_{(i+1)} \otimes \cdots \otimes u_{(k+1)}, \end{aligned}$$

where the terms $i = 1$ and $i = k + 1$ are $a \otimes \underline{\Delta}^k u$ and $\underline{\Delta}^k u \otimes a$ in the first sum and $(a \otimes 1^{\otimes k})\underline{\Delta}^k u$ and $(1^{\otimes k} \otimes a)\underline{\Delta}^k u$ in the second term.

For an arbitrary $u \in S^n(V)$, $n > 0$ and $v \in S(V)$, we also have:

$$\begin{aligned} \Delta^k(uv) &= u_{(1)}v_{(1)} \otimes \cdots \otimes u_{(k)}v_{(k)} \\ &= \sum_{1 \leq p \leq k} \sum_{1 \leq i_1 < \cdots < i_p \leq k} v_{(1)} \otimes \cdots \otimes u_{(1)}v_{(i_1)} \otimes \cdots \otimes u_{(p)}v_{(i_p)} \otimes \cdots \otimes v_{(k)}. \end{aligned}$$

4. Green functions for quasi-free states

Let V be the vector space generated by the symbols $\varphi(x)$, where x runs over points of \mathbb{R}^d . In physical terms, φ should be thought of as a free bosonic field operator, that is, as an operator-valued distribution (think of the quantum fields $\phi_{\mathbf{S}}(\mathbf{r})$). Our forthcoming developments can be adapted easily to fermionic systems, the adaptation amounting mathematically to replacing the symmetric algebra $S(V)$ by the exterior (or Grassmann) algebra $\Lambda(V)$, see [11].

Defining a time-ordered product of fields at the same point gives rise to major difficulties and is the subject of renormalization [44]. Here, we take advantage of the fact that the combinatorics of Green functions is in many respects a self-contained topic and leave aside these questions (renormalization, operator product expansion). We will therefore treat powers of fields such as $\varphi^4(x)$ as formal expressions, that is as monomials belonging to the symmetric Hopf algebra $S(V) \supset S^4(V)$. Note that $\varphi^0(x) = 1$ is the unit of the algebra $S(V)$.²

²In another paper [22], an algebra different from $S(V)$ was used, where $\varphi^0(x)$ was not the unit of the algebra, in order to obtain some desirable coalgebraic properties. That alternative construction considers the field products $\varphi^n(x)$ as the basis of a Hopf algebraic fiber at x . However such a point of view is not required in the present paper.

4.1. Convolution. In this section we survey some Hopf algebraic concepts that provide a startling simplification of the decomposition of the expectation value of time-ordered products in terms of Feynman diagrams. We define a *form* as a linear map from $S(V)$ to \mathbb{C} . A *unital form* is a form ρ such that $\rho(1) = 1$. In our context, that is when $\varphi(x)$ is the quantum field of QFT or many-body theory, unital forms are defined from states of H_0 : if $|\Phi\rangle$ is a normalized state and $u \in S(V)$, then $\rho(u) = \langle\Phi|T(u)|\Phi\rangle$ is a unital form because it is obviously linear and $\rho(1) = \langle\Phi|1|\Phi\rangle = 1$. The unital form corresponding to the vacuum is denoted by ρ_0 , so that $\rho_0(u) = \langle 0|T(u)|0\rangle$.

To express $\rho_0(u)$ in Hopf algebraic terms, we first need a few definitions. The *convolution* product of two forms ρ and σ is the form $\rho * \sigma$ defined by $(\rho * \sigma)(u) = \rho(u_{(1)})\sigma(u_{(2)})$. Notice that, because of the commutativity and cocommutativity of $S(V)$, $\sigma * \rho = \rho * \sigma$. The space of unital forms equipped with the convolution product is a commutative group, denoted by \mathcal{S} , whose unit is the counit ε .

The n -th *convolution power* of a form ρ is the form ρ^{*n} defined recursively by $\rho^{*0} = \varepsilon$, $\rho^{*1} = \rho$ and $\rho^{*(n+1)} = \rho^{*n} * \rho$. The *convolution exponential* of a form ρ is the form $e^{*\rho}$ defined by

$$e^{*\rho} = \sum_{n=0}^{\infty} \frac{\rho^{*n}}{n!}.$$

The convolution logarithm $\log^* \rho$ of the form ρ is the form defined by

$$\log^* \rho = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\rho - \varepsilon)^{*n}.$$

Note that, if ρ is a unital form, $\log^* \rho$ satisfies $\log^* \rho(1) = 0$. A form σ such that $\sigma(1) = 0$ is called an *infinitesimal form* because it is the logarithm of a unital form. In the present paper, the convolution exponential is always applied to infinitesimal forms. Note that, if $\sigma = \log^* \rho$, then $e^{*\sigma} = \rho$. In other words, convolution exponential and convolution logarithm are inverse functions of each other. At last, note that, if α and β are two unital forms with convolution logarithms a and b , then $\alpha * \beta = e^{*a} * e^{*b} = e^{*(a+b)}$.

4.2. Expansion in Feynman diagrams. In standard quantum field theory, Wick's theorem states that, if $u = \varphi^{k_1}(x_1) \dots \varphi^{k_n}(x_n)$, $\langle 0|T(u)|0\rangle$ is calculated as the sum of all pairings of k_1 times the point x_1 , \dots , k_n times the point x_n . A pairing is the choice of a pair of different points represented graphically as a line and analytically as a Feynman propagator. Graphically, $\rho_0(u) = \langle 0|T(u)|0\rangle$ is therefore represented by the sum of all the graphs with n vertices labeled by x_1, \dots, x_n such that k_i edges are incident to the vertex labeled by x_i , for $i = 1, \dots, n$. Each graph is weighted by a proper combinatorial factor.

To express $\rho_0(u)$ in Hopf algebraic terms, we define the infinitesimal form τ by

$$\tau(\varphi(x_1)\varphi(x_2)) := D_F(x_2 - x_1) \text{ if } x_1 \neq x_2,$$

and

$$\tau(\varphi(x_1) \dots \varphi(x_n)) := 0, \text{ if } n \neq 2 \text{ or } n = 2 \text{ and } x_1 = x_2.$$

The form τ is called the *Feynman form*. The function D_F is defined³ by

$$D_F(x) = \int \frac{d^4 p}{(2\pi)^4} \frac{i}{p^2 - m^2 + i\varepsilon} e^{-i(p \cdot x)}.$$

We can now restate Wick's theorem algebraically:

THEOREM 4.1. ([6, 7, 11, 22]) *The unital form ρ_0 is the convolution exponential of the Feynman form :*

$$\rho_0 = e^{*\tau}.$$

This theorem extends to the case of quasi-free states [45], the only change is that τ is now defined by

$$\tau(\varphi(x_1)\varphi(x_2)) := D_2(x_1, x_2) \text{ if } x_1 \neq x_2,$$

and

$$\tau(\varphi(x_1)\dots\varphi(x_n)) := 0, \text{ if } n \neq 2 \text{ or } n = 2 \text{ and } x_1 = x_2,$$

where $D_2(x_1, x_2) = \langle \Phi | T(\varphi(x_1)\varphi(x_2)) | \Phi \rangle$.

5. Green functions for general states

Most often the relevant object to deal with in perturbative expansions is actually not the unital form ρ built from the ground state $|\Phi\rangle$ (or, abstractly, the group \mathcal{S}) but its convolution logarithm r (resp. the corresponding commutative Lie algebra \mathcal{L}). The infinitesimal form r is called the *cumulant form*. As we shall see in section 6.2, this is exactly what we need to calculate the Green functions in the presence of initial correlations.

The theorem (4.1) of the previous section trivially generalizes to arbitrary states and unital forms:

THEOREM 5.1. ([6, 7, 5]) *The unital form ρ is the convolution exponential of its cumulant form :*

$$\rho = e^{*r}.$$

Although our approach is not the usual one, writing ρ as e^{*r} is in fact quite common in physics. The notion of a cumulant form is related to the cumulant expansion, and expresses the generalized Wick theorem used for solving the Anderson model [46]. Moreover, it is a way to isolate the singularities of the forms because a natural property of a quantum field is that $r(\varphi(x_1)\dots\varphi(x_n))$ is a smooth function of x_1, \dots, x_n , except possibly for $n = 2$ (see [47], as well as [48] for a related result in many-body theory).

Finally, an observation that will prove essential in our forthcoming developments: all our previous reasonings suggest that a unital form $\rho = e^{*r}$ should be dealt with by means of generalized propagators in the same way as vacuum expectations of time-ordered products of free fields are dealt with by means of 2-point Green functions and Feynman propagators in the usual picture of QFT. However, whereas the Feynman propagator, which is associated to the unique non trivial component of τ (recall that $\tau = 0$ on $S^i(V)$ for any $i \neq 2$), is described graphically by a line linking two vertices, we may have now $r(\varphi(x_1)\dots\varphi(x_n)) \neq 0$ with $n \neq 2$. Accordingly, we shall represent graphically the “ n -point propagator”

³When x_1 and x_2 are separated by a light-like interval, the definition of τ does not make sense and D_F should be replaced by a smooth regularization. We do not enter into these details here since we consider only the combinatorial aspects of the problem.

$D_n(x_1, \dots, x_n) = r(\varphi(x_1) \dots \varphi(x_n))$ ⁴ by a white dot with n edges linked to the n vertices x_1, \dots, x_n , as shown in figure 1 (a similar convention was used by Djah et al. [5]).

$$D_3(x, y, z) = \begin{array}{c} \bullet \\ \diagup \quad \diagdown \\ x \quad y \end{array}$$

FIGURE 1. The generalized propagator $D_3(x, y, z) = r(\varphi(x)\varphi(y)\varphi(z))$

As we already mentioned, we can also consider $\rho = e^{*r}$ as a generalization of Wick's theorem when the latter is stated algebraically. The same observation holds for graphical statements of the theorem: we saw that, in standard quantum field theory, Wick's theorem states that, if $u = \varphi^{k_1}(x_1) \dots \varphi^{k_n}(x_n)$, $\langle 0|T(u)|0 \rangle$ is calculated as the sum of all pairings of k_1 times the point x_1, \dots, k_n times the point x_n . In the many-body context $e^{*r}(u)$ is replaced by $e^{*r}(u)$. This amounts to say that we write $e^{*r}(u)$ as the sum of all ways to partition the multiset made of k_1 times point x_1, \dots, k_n times point x_n into sub-multisets of any multiplicity (i.e. not only pairs and not only different points). See figure 2 for an example. The n -point propagators are then a convenient way to represent these sub-multisets.

To conclude this section, we state three easy but important lemmas

LEMMA 5.1. *If $a \in V$ and $E = e^a$, then for any form ρ with logarithm r , we have $\rho(E) = e^{r(E)}$.*

LEMMA 5.2. *If $a \in V$ and $u \in S(V)$, then, for any linear map $r : S(V) \rightarrow \mathbb{C}$ such that $r(1) = 0$,*

$$e^{*r}(au) = \sum r(au_{(1)})e^{*r}(u_{(2)}).$$

More generally,

LEMMA 5.3. *For any $u \in \ker \epsilon$ and any v in $S(V)$,*

$$(2) \quad e^{*r}(uv) = \sum_{k=1}^{\infty} \frac{1}{k!} r(u_{(1)}v_{(1)}) \dots r(u_{(k)}v_{(k)}) e^{*r}(v_{(k+1)}).$$

PROOF. The first lemma is a simple consequence of the fact that E is group-like (that is, $\Delta(E) = E \otimes E$). The second lemma was shown in [22], it follows from the cocommutativity of the coproduct and the fact that a is a primitive element (that is, $\Delta(a) = 1 \otimes a + a \otimes 1$). The third lemma follows from the last identity in Lemma 3.1, from the properties of the binomial coefficients, and from the cocommutativity of the coproduct. \square

The first lemma is often used with $a = \int j(x)\varphi(x)dx$ (up to a suitable extension of the definition of V when the function $j(x)$ has not a discrete support). In that case, it relates the generating function of the moments of ρ to that of the moments

⁴The definition of D_n generalizes the definition of D_4 in section 2.3 –this should be clear from our forthcoming developments.

of r . The second and third lemmas provide powerful tools for the recursive proof of the properties of e^{*r} . Notice in particular that, using the last Lemma with $v = 1$:

$$(3) \quad \rho(a^n) = \sum_{k=1}^n \frac{1}{k!} \sum_{i_1 + \dots + i_k = n} \frac{n!}{i_1! \dots i_k!} r(a^{i_1}) \dots r(a^{i_k}),$$

where, for $p = 1, \dots, k$, $i_p > 0$. A formula with less terms can be given using the Faà di Bruno coefficients:

$$(4) \quad \rho(a^n) = \sum_{\alpha} \frac{n! r(a^1)^{\alpha_1} \dots r(a^n)^{\alpha_n}}{\alpha_1! (1!)^{\alpha_1} \alpha_2! (2!)^{\alpha_2} \dots \alpha_n! (n!)^{\alpha_n}},$$

where $(\alpha_1, \dots, \alpha_n)$ are nonnegative integers such that $\sum_i i \alpha_i = n$. For the partition represented by α , n is cut into $k = \sum_i \alpha_i$ parts. For example, $\rho(a) = r(a)$, $\rho(a^2) = r(a^2) + r(a)^2$, $\rho(a^3) = r(a^3) + 3r(a)r(a^2) + r(a)^3$.

6. Connected forms

In quantum field theory, an important simplification comes from the fact that all physical quantities can be expressed in terms of connected diagrams. We first define the notion of a *connected form* by analogy with that of a connected diagram. A monomial of $S(V)$ can always be written $u = \varphi^{n_1}(x_1) \dots \varphi^{n_k}(x_k)$, where all points x_i are distinct. Now, for any form ρ with convolution logarithm r , we use eq. (4) to expand $\rho(u)$ in terms of r .

PROPOSITION 6.1. *We have*

$$\begin{aligned} \rho(u) = e^{*r}(u) &= \sum_{l \in \mathbb{N}} \frac{1}{l!} \sum_{\substack{n_1^1 + \dots + n_i^l = n_i \\ i=1 \dots k}} \prod_{i=1}^k \frac{n_i!}{n_i^1! \dots n_i^l!} \\ &\quad \times r(\varphi^{n_1^1}(x_1) \dots \varphi^{n_k^1}(x_k)) \dots r(\varphi^{n_1^l}(x_1) \dots \varphi^{n_k^l}(x_k)), \end{aligned}$$

where, for $i = 1, \dots, k$, the sum is over all the l -tuples of nonnegative integers (n_i^1, \dots, n_i^l) such that $n_i^1 + \dots + n_i^l = n_i$.

Although a straightforward application of the Hopf algebra formalism, the result is important since it allows us to compute the multiplicity of a graph –or symmetry factor– in the Feynman diagrammatic perturbative expansion of amplitudes. We refer for example to the expansion of the connected Green functions for φ^3 theory with an arbitrary ground state in the present section of the article.

Let us consider a term $t := r(\varphi^{n_1^1}(x_1) \dots \varphi^{n_k^1}(x_k)) \dots r(\varphi^{n_1^l}(x_1) \dots \varphi^{n_k^l}(x_k))$ of $e^{*r}(u)$. We say that $x_i \cong_t x_j$, $1 \leq i, j \leq k$ if there exists $m \leq l$ with $n_i^m n_j^m \neq 0$. The transitive closure \equiv_t of the binary relation \cong_t defines the connectedness of t : t is said to have n connected components if there are n equivalence classes associated to the equivalence relation \equiv_t . The connected component of x_i in t is defined similarly as the product of all the $r(\varphi^{n_1^m}(x_1) \dots \varphi^{n_k^m}(x_k))$ with $n_j^m \neq 0$ for at least one coefficient j with $x_i \equiv_t x_j$. When $n = 1$ (resp. $n \neq 1$), we also say that the term t is connected (resp. disconnected). Let us take a simple example. For $u = \varphi(x)\varphi^2(y)$, we have $\rho(u) = r(\varphi(x))r(\varphi^2(y)) + r(\varphi(x))r(\varphi(y))^2 + 2r(\varphi(x)\varphi(y))r(\varphi(y)) + r(\varphi(x)\varphi^2(y))$, where the first two terms are disconnected (they actually have two connected components). The connected components of y in the four terms are respectively

$r(\varphi^2(y)), r(\varphi(y))^2, r(\varphi(x)\varphi(y))r(\varphi(y))$ and $r(\varphi(x)\varphi^2(y))$. The definition of connected form is actually best formulated in algebraic terms: this is the purpose of the next section.

6.1. Another coproduct on $S(V)$. As we have just seen, a pedestrian definition of connectedness makes an essential use of the fact that some points x_i are equal or distinct. In order to reflect this distinction, we define a new coproduct, the *disconnecting coproduct* $\delta : S(V) \rightarrow S(V) \otimes S(V)$.

So we write a monomial of $S(V)$ as $u = \varphi^{n_1}(x_1) \dots \varphi^{n_k}(x_k)$, where all points x_i are distinct, and we define the coproduct of u as follows: $\delta\varphi^n(x) = 1 \otimes \varphi^n(x) + \varphi^n(x) \otimes 1$ if $k = 1$, and $\delta u = \delta(\varphi^{n_1}(x_1))\delta(\varphi^{n_2}(x_2)) \dots \delta(\varphi^{n_k}(x_k))$ if $k > 1$. Notice that this coproduct is coassociative and cocommutative but is not an algebra morphism, because $\delta(\varphi^2(x)) \neq (\delta(\varphi(x)))^2$. Since δ is coassociative and cocommutative, we may still define an associative, commutative and unital product $\hat{*}$, the *disconnecting convolution product*, on $\text{Lin}(S(V), \mathbb{C})$:

$$\forall (f, g) \in \text{Lin}(S(V), \mathbb{C}), f \hat{*} g := \pi \circ (f \otimes g) \circ \delta,$$

where π denotes the product: $\pi \circ (f \otimes g)(u \otimes v) = f(u)g(v)$. The unit of $\hat{*}$ is the same as the unit of $*$ (the projection map ε from $S(V)$ to $\mathbb{C} \subset S(V)$). To distinguish between the two products $*$ and $\hat{*}$, we write the operations involving $\hat{*}$ with a superscript $\hat{*}$: for example, we write $\log^{\hat{*}}$, and so on.

The relation between δ and Δ is investigated in [22]. The reduced coproduct $\underline{\delta}$ and the iterated coproduct δ^k are defined as in section 3. The enhanced Sweedler notation for the disconnecting coproduct is $\delta u = u_{\{1\}} \otimes u_{\{2\}}$ and $\underline{\delta}u = u_{\{1\}} \otimes u_{\{2\}}$.

The new coproduct δ enables us to give an algebraic definition of the connected form ρ_c corresponding to the unital form ρ :

$$\rho_c = \log^{\hat{*}}(\rho)$$

that is,

$$\forall u \in S(V), \rho_c(u) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (\rho - \varepsilon)(u_{\{1\}}) \dots (\rho - \varepsilon)(u_{\{n\}}).$$

This identity can be understood as a linked cluster theorem for forms. The same problem was addressed in [5, Sect.6] with another approach: in that article, the authors introduce two notions of truncated moment functions, indicated respectively by an exponent T and (T) . This corresponds roughly, at the Hopf algebraic level, to our distinction between the two coproducts Δ and δ . We thank the referee for pointing out to us this point and refer to [5] for further insights on truncated moment functions.

Pedantically, the set of connected forms is defined as the image of the group of unital forms under the map $\log^{\hat{*}}$. The two sets are in bijective correspondence and, reciprocally, we can express any unital form ρ in terms of ρ_c by

$$\rho = e^{\hat{*}\rho_c}$$

or:

$$\forall u \in S(V), \rho(u) = \varepsilon(u) + \sum_{n=1}^{\infty} \frac{1}{n!} \rho_c(u_{\{1\}}) \dots \rho_c(u_{\{n\}}).$$

For example, $\rho_c(\varphi^n(x)) = \rho(\varphi^n(x))$ for $n > 0$ and

$$\rho_c(\varphi^m(x)\varphi^n(y)) = \rho(\varphi^m(x)\varphi^n(y)) - \rho(\varphi^m(x))\rho(\varphi^n(y))$$

for $m > 0$ and $n > 0$. The connected form ρ_c is an infinitesimal form (that is, $\rho_c(1) = 0$). For u as above, $\rho_c(u)$ is defined as the sum of the connected terms of $\rho(u)$. This terminology is due to the fact that we can define Feynman diagrams to represent $\rho(u)$, and that $\rho_c(u)$ is obtained by summing the connected Feynman diagrams present in $\rho(u)$.

Note that the relation $\rho = e^{\hat{\ast}\rho_c}$ is the analogue of the relation $Z = e^W$ between the partition function and the free energy. As we prove now, the latter is a consequence of the former. The Hamiltonian density of a quantum field theory of local interactions is of the form $H(x)$. Therefore, $\delta H(x) = H(x) \otimes 1 + 1 \otimes H(x)$ is primitive and, by lemma 5.1, $Z = \rho(S) = e^W$, where $W = \rho_c(S)$ (see [22] for a detailed proof). This is an extension of the standard relation to the case of a general initial state. As a matter of fact, the coproduct δ was precisely defined for the Hamiltonian to be a primitive element. This ensures the standard relation between the partition function and the free energy (compare to [5]).

6.2. Example of the φ^3 theory. As we saw in section 2.3, the two-point Green function for a system described by the interaction Hamiltonian density $u = \varphi^3(x)$ is given by the expression:

$$G(x, y) = \frac{\langle 0 | T(\varphi(x)\varphi(y)e^{-iu}) | 0 \rangle}{\langle 0 | T(e^{-iu}) | 0 \rangle}.$$

We recall that the denominator cancels the divergence of the adiabatic switching of the interaction. In graphical terms, the denominator $\langle 0 | T(e^{-iu}) | 0 \rangle$ is the sum of all the vacuum Feynman diagrams (i.e. the diagrams that are linked neither to x nor to y). Another way to obtain a convergent expression is to use the connected Green function $G_c(x, y)$ which is the sum of all the connected diagrams in $G(x, y)$.

For a general form, the factorization of the adiabatic divergence is more complex [40, 49, 50] and it holds only for specific initial states [32]. For notational convenience, we do not write the denominator in the definition of the Green functions for a general form and we put

$$(5) \quad G(x, y) = \rho(\varphi(x)\varphi(y)e^{-i\int_{-\infty}^{\infty} H_{\text{int}}(t)}).$$

The connected Green function is defined as

$$G_c(x, y) = \rho_c(\varphi(x)\varphi(y)e^{-i\int_{-\infty}^{\infty} H_{\text{int}}(t)}).$$

The term $\int_{-\infty}^{\infty} H_{\text{int}}(t)$ can usually be written $\int dx P(x)$, where $P(x)$ is a polynomial in $\varphi(x)$. Therefore,

$$G(x, y) = \rho(\varphi(x)\varphi(y)) + \sum_{n=1}^{\infty} \frac{(-i)^n}{n!} \int dx_1 \dots dx_n \rho(\varphi(x)\varphi(y)P(x_1) \dots P(x_n)).$$

For example, in the φ^3 theory, we have $P(x) = \varphi^3(x)/3!$ and the first terms of the total Green function are

$$\begin{aligned} G(x, y) &= \rho(\varphi(x)\varphi(y)) - \frac{i}{6} \int dx_1 \rho(\varphi(x)\varphi(y)\varphi^3(x_1)) \\ &\quad - \frac{1}{72} \int dx_1 dx_2 \rho(\varphi(x)\varphi(y)\varphi^3(x_1)\varphi^3(x_2)) + \dots \end{aligned}$$

For notational convenience, we assume that $D_n(x_1, \dots, x_n) = r(\varphi(x_1) \dots \varphi(x_n)) = 0$ if n is odd. The connected Green function $G_c(x, y)$ is obtained by keeping the connected terms of the total Green function.

In standard quantum field theory, the expansion to the second order gives us

$$(6) \quad G_c(x, y) = D_2(x, y) - \frac{1}{2} \int dz dw D_2(x, z) D_2(y, w) D_2(z, w)^2 + \dots$$

For a general form, the expansion to the second order gives a more complex result:

(7)

$$\begin{aligned} G_c(x, y) = & D_2(x, y) - \int dz dw \left(\frac{1}{72} D_8(x, y, z, z, z, z, w, w, w) \right. \\ & + \frac{1}{12} D_2(x, z) D_6(y, z, z, w, w, w) \\ & + \frac{1}{12} D_6(x, z, z, w, w, w) D_2(y, z) + \frac{1}{12} D_6(x, y, z, w, w, w) D_2(z, z) \\ & + \frac{1}{8} D_6(x, y, z, z, w, w) D_2(z, w) + \frac{1}{12} D_4(x, y, z, z) D_4(z, w, w, w) \\ & + \frac{1}{8} D_4(x, y, z, w) D_4(z, z, w, w) + \frac{1}{4} D_4(x, z, z, w) D_4(y, z, w, w) \\ & + \frac{1}{6} D_2(x, z) D_2(y, z) D_4(z, w, w, w) + \frac{1}{4} D_2(x, z) D_2(y, w) D_4(z, z, w, w) \\ & + \frac{1}{2} D_2(x, z) D_4(y, z, w, w) D_2(z, w) + \frac{1}{4} D_2(x, z) D_4(y, z, z, w) D_2(w, w) \\ & + \frac{1}{2} D_4(x, z, w, w) D_2(y, z) D_2(z, w) + \frac{1}{4} D_4(x, z, z, w) D_2(y, z) D_2(w, w) \\ & + \frac{1}{4} D_4(x, y, w, w) D_2(z, z) D_2(z, w) + \frac{1}{8} D_4(x, y, z, w) D_2(z, z) D_2(w, w) \\ & + \frac{1}{4} D_4(x, y, z, w) D_2(z, w)^2 + \frac{1}{2} D_2(x, z) D_2(y, z) D_2(z, w) D_2(w, w) \\ & \left. + \frac{1}{2} D_2(x, z) D_2(y, w) D_2(z, w)^2 \right) + \dots \end{aligned}$$

These terms can be given the diagrammatic representation of figure 2.

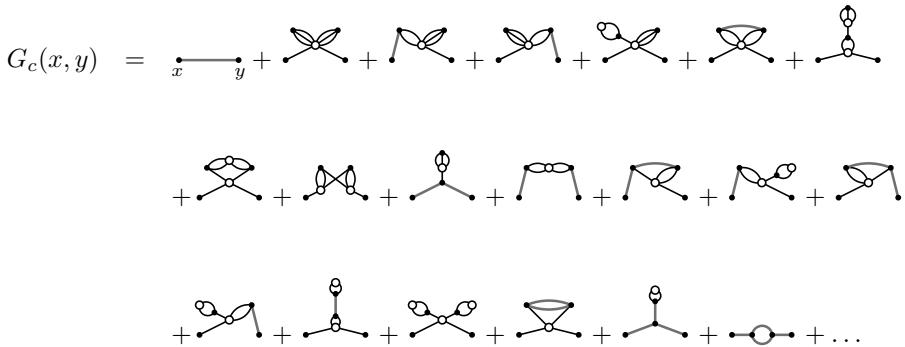


FIGURE 2. The first few terms of $G_c(x, y)$

It is clear from a comparison of equations (6) and (8) that the use of a form ρ which does not come from a quasi-free state increases significantly the combinatorial complexity. As a consequence, little was known about the structure of the connected

Green functions in the general case, especially as far as one-particle-irreducibility is concerned.

7. Symmetric functions and derivations

In [19, 20], Mestre and Oeckl have proposed a powerful Hopf algebraic tool to generate 1PI diagrams. The forthcoming developments of the present article aim at extending their work to correlated systems, and at describing in this general setting the decomposition of connected Green functions into one-particle irreducible Green functions.

In order to do so, the language of functional derivatives proves to be a very convenient framework. This leads naturally to the study of Hopf algebra derivations, which is the main topic of the present section. Namely, we describe some aspects of the algebra of symmetric functions in the Hopf algebraic setting and prove that this algebra has universal properties with respect to derivations acting on commutative Hopf algebras. We will show later that this formalism nicely encodes the properties of the generalized n -point propagators associated to arbitrary states, as described above. We do not seek the utmost generality in our constructions, but mention that they can be easily extended to more general (e.g. noncommutative) situations.

7.1. Definition of the derivations A_m . Let us consider, once again, the Hopf algebra $U = S(V)$. Two products arise therefore in $S(U) = S(S(V))$: the symmetric product in $S(V)$, denoted by juxtaposition, and the symmetric product in $S(U)$, denoted from now on by \vee to avoid any confusion.

To a unital form ρ on $U = S(V)$, we associate its convolution logarithm r and, for $\varphi(x_1) \dots \varphi(x_m) \in U$, we recall that $D_m(x_1, \dots, x_m) = r(\varphi(x_1) \dots \varphi(x_m))$, so that the generalized propagators D_m are symmetric functions of their arguments.

Recall that V is spanned by the symbols $\varphi(x)$, where x runs over points in \mathbb{R}^d . We choose an arbitrary total order on the points in \mathbb{R}^d , for example the lexicographical order on the d -tuples of coordinates. The operators $A_{m-1} : U \rightarrow S^m(U)$ are then defined by

$$(9) \quad \begin{aligned} A_{m-1}(u) &= \sum_{x_1 < \dots < x_m} D_m(x_1, \dots, x_m) \frac{\bar{\partial}u_{\{1\}}}{\partial\varphi(x_1)} \vee \dots \vee \frac{\bar{\partial}u_{\{m\}}}{\partial\varphi(x_m)} \\ &= \frac{1}{m!} \sum_{x_1, \dots, x_m} D_m(x_1, \dots, x_m) \frac{\bar{\partial}u_{\{1\}}}{\partial\varphi(x_1)} \vee \dots \vee \frac{\bar{\partial}u_{\{m\}}}{\partial\varphi(x_m)}, \end{aligned}$$

where $\frac{\bar{\partial}u}{\partial\varphi(x)}$ is defined by the following (topologically motivated) rule:

- $\frac{\bar{\partial}u}{\partial\varphi(x)} := \frac{\partial u}{\partial\varphi(x)}$ if u is not linear in $\varphi(x)$ or if $\frac{\partial u}{\partial\varphi(x)}$ is a scalar.
- $\frac{\bar{\partial}u}{\partial\varphi(x)} := 0$ otherwise.

These operators on U extend uniquely to derivations on $S(U)$, still written A_{m-1} , so that:

$$A_{m-1} : S^k(U) \longrightarrow S^{k+m-1}(U)$$

(recall, for completeness sake, that a derivation D satisfies the Leibniz rule, $D(xy) = D(x)y + xD(y)$, so that the action of a derivation on $S(U)$ follows by induction on the degrees from the knowledge of its action on U). Notice also that, by definition of δ , the terms with $x_i = x_j$ vanish in the sum of eq. (9).

In practice, forcing the null value in the linear case (when u is not a scalar multiple of $\varphi(x)$) will permit to avoid the creation of topologically disconnected graphs. We are particularly grateful to Achim Randelhoff who pointed out to us this property and allowed us to correct an error in a previous version of this article. The meaning of this observation should become clear later on, but the reader may immediately understand its implications by comparing the action of A_1^2 on $\varphi(x)\varphi(y)\varphi(z)^2$ with the action of the derivation \hat{A}_1^2 , where the action of \hat{A}_{m-1} is defined on U by:

$$\hat{A}_{m-1}(u) = \sum_{x_1 < \dots < x_m} D_m(x_1, \dots, x_m) \frac{\partial u_{\{1\}}}{\partial \varphi(x_1)} \vee \dots \vee \frac{\partial u_{\{m\}}}{\partial \varphi(x_m)}.$$

An important point for our forthcoming developments is that the A_i commute. This follows from the Schwarz commutation rules for derivatives and from the definition of the disconnecting coproduct δ . Graphically, these operations will allow us to construct inductively 1PI graphs (with a given set of vertices and with given multiplicities), and their algebraic properties will allow us to enumerate these graphs. This will be the purpose of the next section, whereas the following subsections introduce the abstract Hopf algebraic framework suited for this enumeration.

7.2. The Hopf algebra of symmetric functions. Let $X = \{x_1 \dots x_n \dots\}$ be a countable alphabet, and $\mathbb{C}[[X]]$ the algebra of formal power series over X . The group $\mathcal{S}_\infty = \lim_{\rightarrow} \mathcal{S}_n$ acts on $\mathbb{C}[[X]]$ by permutation of the letters of X ; the algebra of symmetric functions $\mathbb{S}ym$ is the subalgebra of \mathcal{S}_∞ -invariant series in $\mathbb{C}[[X]]$.

The algebra $\mathbb{S}ym$ is (up to completion with respect to the filtration induced by the grading of symmetric polynomials by their degree) a free commutative algebra over various families of generators. For our purposes, the most interesting ones are the families of power sums symmetric functions and complete symmetric functions associated respectively to the series

$$\mathbf{P}_\bullet := \sum_{k \in \mathbb{N}} P_k := 1 + \sum_{i \in \mathbb{N}} \frac{x_i}{1 - x_i}$$

and

$$\mathbf{C}_\bullet := \sum_{k \in \mathbb{N}} S_k := \prod_{i \in \mathbb{N}} \frac{1}{1 - x_i}.$$

In view of our forthcoming computations, it is actually convenient to work with an extension of $\mathbb{S}ym$, $q\mathbb{S}ym$: we write Q_k for $\frac{P_k}{k}$, $k \geq 1$, $Q_0 := q$ and $\mathbf{Q}_\bullet := \sum_{k \in \mathbb{N}} Q_k$, where q stands for an additional free variable. The series are related by the familiar Newton-type identity: $\mathbf{C}_\bullet = e^{\mathbf{Q}_\bullet - Q_0}$. We write \mathbf{S}_\bullet for the q -series $\mathbf{S}_\bullet := e^{Q_0} \mathbf{C}_\bullet = e^{\mathbf{Q}_\bullet}$.

The algebra $q\mathbb{S}ym$ carries a natural notion of grading (by the degrees of symmetric polynomials, with $\deg(Q_n) = n$), but it is convenient, for our purposes, and for reasons that will become clear later, to introduce an extra “auxiliary” grading by considering the family of the Q_k as a family of generators of $\mathbb{S}ym$ of auxiliary degree 1. This is best explained through an example: $Q_0^2 Q_3 Q_5 Q_9$ is of degree 17 and of auxiliary degree 5. The auxiliary degree is indicated with a superscript (whereas the degree is indicated by a subscript), so that, for example, the component of

degree n and auxiliary degree k in \mathbf{S}_\bullet is given by:

$$\mathbf{S}_n^k = \sum_{\alpha} \frac{Q_0^{\alpha_0}}{\alpha_0!} \frac{Q_1^{\alpha_1}}{\alpha_1!} \cdots \frac{Q_n^{\alpha_n}}{\alpha_n!}$$

where the sum runs over all $(n+1)$ -tuples of integers $\alpha = (\alpha_0, \dots, \alpha_n)$ with $\alpha_0 + \cdots + \alpha_n = k$ and $\alpha_1 + 2\alpha_2 + \cdots + n\alpha_n = n$. Notice that we distinguish carefully between \mathbf{S}_n^k and S_n^k , the latter standing for the k -th power of S_n . The following examples will be useful in the sequel: $\mathbf{S}_n^0 = \delta_{n,0} 1$, $\mathbf{S}_0^k = Q_0^k/k!$, $\mathbf{S}_n^1 = Q_n$, $\mathbf{S}_1^k = Q_0^{k-1} Q_1/(k-1)!$, $\mathbf{S}_2^2 = Q_0 Q_2 + Q_1^2/2$.

Generating series are a useful tool to handle computations with the \mathbf{Q}_n^k and the \mathbf{S}_n^k . Consider for example the series $\mathbf{S}_\bullet(a+b) := e^{(a+b)\mathbf{Q}_\bullet}$: its expansion as a series in the variables a and b yields:

PROPOSITION 7.1. *We have, for all $k, l \geq 0$:*

$$\binom{k+l}{k} \mathbf{S}_n^{k+l} = \sum_{m=0}^n \mathbf{S}_m^k \mathbf{S}_{n-m}^l.$$

In particular

$$\mathbf{S}_n^k = \frac{1}{k} \sum_{m=0}^n Q_m \mathbf{S}_{n-m}^{k-1}.$$

The Hopf algebraic properties of symmetric functions were recently exploited with great profit by Fauser and coll. [51, 52]. Similarly, we put a simple Hopf algebraic structure on $q\mathbb{S}\text{Sym}$ uniquely defined by requiring the power sums symmetric functions (i.e. Q_n for $n > 0$) to form, together with Q_0 , a series of primitive elements or, equivalently, by requiring the series \mathbf{S}_\bullet to be a group-like element. In other terms, the coproduct Δ on $q\mathbb{S}\text{Sym}$ is fully specified by requiring that $\Delta(Q_n) := Q_n \otimes 1 + 1 \otimes Q_n$, $n \geq 0$. In particular, the coproduct is compatible with the two graduations. When expliciting this property in \mathbf{S}_n^k , we get:

PROPOSITION 7.2. *The coproduct of \mathbf{S}_n^k is*

$$\Delta \mathbf{S}_n^k = \sum_{m=0}^n \sum_{i=0}^k \mathbf{S}_m^i \otimes \mathbf{S}_{n-m}^{k-i},$$

and its iterated coproduct is

$$\Delta^{p-1} \mathbf{S}_n^k = \sum_{\substack{n_1 + \cdots + n_p = n \\ k_1 + \cdots + k_p = k}} \mathbf{S}_{n_1}^{k_1} \otimes \cdots \otimes \mathbf{S}_{n_p}^{k_p},$$

Note that propositions 7.1 and 7.2 still hold if the variables Q_n do not commute.

7.3. On Hopf algebra derivations. Let $H = \bigoplus_{n \in \mathbb{N}} H_n$ be an arbitrary connected graded commutative Hopf algebra and $A_0, A_1, \dots, A_n, \dots$ an arbitrary sequence of derivations on H with degrees $0, 1, \dots, n, \dots$. That is, for any p , the restriction of A_n to H_p is a linear map from H_p to H_{p+n} , and A_n satisfies the Leibniz rule: for any h, l in H , $A_n(h \cdot l) = A_n(h) \cdot l + h \cdot A_n(l)$. We also assume that the A_n commute, so that the A_n generate a commutative subalgebra \mathcal{D} of $\text{End}(H)$ (for the composition of maps). Of course, we have in mind the particular derivations A_m acting on $S(U)$, but the following results hold in full generality.

There is therefore, since $q\mathbb{S}ym$ is free over the Q_n , a universal algebra map β from $q\mathbb{S}ym$ to $End(H)$ obtained by mapping Q_n to A_n . We write \mathbf{L}_\bullet for the image of \mathbf{S}_\bullet under this map, and L_n^k for the image of \mathbf{S}_n^k . Note that, for any p , L_n^k maps H_p to H_{p+n} . Of course, the identities that hold in $q\mathbb{S}ym$ for the variables Q_n and \mathbf{S}_n^k also hold in $End(H)$ for the variables A_n and L_n^k . More surprisingly however, the coalgebra structure of $q\mathbb{S}ym$ reflects the action of \mathcal{D} on H . We refer to [53, 16] for similar phenomena occurring in the study of Lie idempotents and renormalization in perturbative QFT.

PROPOSITION 7.3. *We have, for any $X \in q\mathbb{S}ym$ and any $h, h' \in H$:*

$$\beta(X)(hh') = \beta(X_{(1)})(h)\beta(X_{(2)})(h')$$

where $X_{(1)} \otimes X_{(2)}$ stands, as usual, for the coproduct of X in $q\mathbb{S}ym$.

The identity can be generalized by a straightforward recursion to compute $\beta(X)(h_1 \dots h_n)$. Notice first that the identity in the Proposition is obvious when X is a Q_n , since $\beta(Q_n) = A_n$ is, by hypothesis, a derivation. Now, assume that for X and Y in $q\mathbb{S}ym$ and arbitrary $h, h', l, l' \in H$ the above formula holds, that is:

$$\beta(X)(hh') = \beta(X_{(1)})(h)\beta(X_{(2)})(h'), \quad \beta(Y)(ll') = \beta(Y_{(1)})(l)\beta(Y_{(2)})(l').$$

It follows that:

$$\begin{aligned} \beta(Y) \circ \beta(X)(hh') &= \beta(Y)(\beta(X_{(1)})(h)\beta(X_{(2)})(h')) \\ &= \beta(Y_{(1)}) \circ \beta(X_{(1)})(h) \beta(Y_{(2)}) \circ \beta(X_{(2)})(h') \\ &= \beta((YX)_{(1)})(h) \beta((YX)_{(2)})(h'). \end{aligned}$$

In other terms, if two elements in $q\mathbb{S}ym$ satisfy the identity in the Proposition, their product also satisfies the identity. Since the Q_n satisfy the identity, and since their products span $q\mathbb{S}ym$, the Proposition follows.

Let us consider the particular example of $S(U)$ and of the derivations A_{m-1} . Let us write $A_{m-1}(u) = u_{m-1,1} \vee \dots \vee u_{m-1,m}$, with an enhanced Sweedler-type notation for the action of A_{m-1} from U to $S^m(U)$. We get:

$$\begin{aligned} L_n^k(u) &= \frac{1}{k} \sum_{m=1}^{n+1} L_{n-m+1}^{k-1} A_{m-1}(u) \\ &= \frac{1}{k} \sum_{m=1}^{n+1} \sum_{\substack{k_1 + \dots + k_m = k-1 \\ n_1 + \dots + n_m = n-m+1}} L_{n_1}^{k_1}(u_{m-1,1}) \vee \dots \vee L_{n_m}^{k_m}(u_{m-1,m}), \end{aligned}$$

where \vee is the symmetric product in $S(U)$. This is a generalization of lemma 13 in [19] and of proposition 15 in [20] where Mestre and Oeckl studied the case where $A_n = 0$ for $n \neq 0, 1$. Note that our notation is different from theirs.

8. One-particle irreducible decompositions

We consider now the derivation of an explicit decomposition of a connected Green function into 1PI Green functions for a general state.

As we have just noticed, the map $Q_m \mapsto A_m$ enables us to define L_n^k acting on $S(U) = S(S(V))$. For example $L_0^0(u) = u$, $L_m^1(u) = A_m(u)$ and

$$\begin{aligned}
L_0^2(u) &= \frac{1}{2} \sum_{x_1, x_2} D_1(x_1) D_1(x_2) \frac{\bar{\partial}^2 u}{\partial \varphi(x_1) \partial \varphi(x_2)}, \\
L_1^2(u) &= \sum_{x_1, x_2, x_3} D_2(x_1, x_2) D_1(x_3) \frac{\bar{\partial} u_{\{1\}}}{\partial \varphi(x_1)} \vee \frac{\bar{\partial}^2 u_{\{2\}}}{\partial \varphi(x_2) \partial \varphi(x_3)}, \\
L_2^2(u) &= \frac{1}{2} \sum_{x_1, x_2, x_3, x_4} D_2(x_1, x_2) D_2(x_3, x_4) \frac{\bar{\partial} u_{\{1\}}}{\partial \varphi(x_1)} \vee \frac{\bar{\partial}^2 u_{\{2\}}}{\partial \varphi(x_2) \partial \varphi(x_3)} \vee \frac{\bar{\partial} u_{\{3\}}}{\partial \varphi(x_4)} \\
&\quad + \frac{1}{2} \sum_{x_1, x_2, x_3, x_4} D_3(x_1, x_2, x_3) D_1(x_4) \frac{\bar{\partial} u_{\{1\}}}{\partial \varphi(x_1)} \vee \frac{\bar{\partial} u_{\{2\}}}{\partial \varphi(x_2)} \vee \frac{\bar{\partial}^2 u_{\{3\}}}{\partial \varphi(x_3) \partial \varphi(x_4)}, \\
L_1^3(u) &= \frac{1}{2} \sum_{x_1, x_2, x_3, x_4} D_1(x_1) D_1(x_2) D_2(x_3, x_4) \frac{\bar{\partial}^2 u_{\{1\}}}{\partial \varphi(x_1) \partial \varphi(x_3)} \vee \frac{\bar{\partial}^2 u_{\{2\}}}{\partial \varphi(x_2) \partial \varphi(x_4)} \\
&\quad + \frac{1}{2} \sum_{x_1, x_2, x_3, x_4} D_1(x_1) D_1(x_2) D_2(x_3, x_4) \frac{\bar{\partial}^3 u_{\{1\}}}{\partial \varphi(x_1) \partial \varphi(x_2) \partial \varphi(x_3)} \vee \frac{\bar{\partial} u_{\{2\}}}{\partial \varphi(x_4)}.
\end{aligned}$$

8.1. A tree interpretation. The operator L_n^k can be written as a sum over all the bipartite trees with k white vertices and $n+1$ black vertices. This description in terms of trees is important because, in standard QFT, a connected Green function can also be described as a tree of 1PI Green functions –a description we want to extend to the case of initial correlations. To give a more precise relation between L_n^k and bipartite trees, we consider the expression for L_n^k in terms of partitions α :

$$L_n^k = \sum_{\alpha} \frac{A_0^{\alpha_0}}{\alpha_0!} \cdots \frac{A_n^{\alpha_n}}{\alpha_n!},$$

where the sum runs over the sequences α of nonnegative integers with $\alpha_0 + \cdots + \alpha_n = k$ and $\alpha_1 + \cdots + n\alpha_n = n$. The monomial corresponding to a given α is represented by the sum of all bipartite trees with k white vertices and $n+1$ black vertices, such that α_i white vertices have valency $i+1$, for $i = 1, \dots, n$.

The terms of lowest degrees are

$$\begin{aligned}
L_0^0 &= \text{id} = \bullet, \\
L_0^1 &= A_0 = \bullet \circ, \\
L_0^2 &= \frac{1}{2!} A_0^2 = \circ \bullet \circ, \\
L_1^1 &= A_1 = \bullet \circ \bullet, \\
L_0^3 &= \frac{1}{3!} A_0^3 = \circ \bullet \circ, \\
L_1^2 &= A_0 A_1 = \bullet \circ \bullet \circ, \\
L_2^1 &= A_2 = \bullet \circ \bullet, \\
L_0^4 &= \frac{1}{4!} A_0^4 = \circ \bullet \circ \bullet, \\
L_2^2 &= A_0 A_2 + \frac{1}{2} A_1^2 = \bullet \circ \bullet \circ + \bullet \circ \bullet \circ, \\
L_1^3 &= \frac{1}{2} A_0^2 A_1 = \circ \bullet \circ \bullet \circ + \circ \bullet \circ \bullet, \\
L_3^1 &= A_3 = \bullet \circ \bullet, \\
L_0^5 &= \frac{1}{5!} A_0^5 = \circ \bullet \circ \bullet \circ, \\
L_3^2 &= A_0 A_3 + A_1 A_2 = \bullet \circ \bullet \circ + \bullet \circ \bullet \circ, \\
L_2^3 &= \frac{1}{2!} A_0^2 A_2 + \frac{1}{2!} A_0 A_1^2 = \circ \bullet \circ \bullet + \bullet \circ \bullet \circ + \bullet \circ \bullet \circ + \bullet \circ \bullet \circ, \\
L_1^4 &= \frac{1}{3!} A_0^3 A_1 = \bullet \circ \bullet \circ + \bullet \circ \bullet \circ, \\
L_4^1 &= A_4 = \bullet \circ \bullet.
\end{aligned}$$

To calculate the value of a tree of L_n^k : (i) Associate to each of the $k + n$ edges a variable x_i , with $i = 1, \dots, k + n$. (ii) To each white vertex v , associate the factor $D_m(x_{i_1}, \dots, x_{i_m})$, where m is the valency of v and x_{i_1}, \dots, x_{i_m} are the variables associated to the edges incident to v . (iii) There are $n + 1$ black vertices. Split u into $n + 1$ parts by $\delta^n u = u_{\{1\}} \otimes \dots \otimes u_{\{n+1\}}$. Number the black vertices from 1 to $n + 1$ and to vertex ℓ associate the factor

$$\frac{\bar{\partial}^m u_{\{\ell\}}}{\partial \varphi(x_{i_1}) \dots \partial \varphi(x_{i_\ell})},$$

where $x_{i_1}, \dots, x_{i_\ell}$ are the variables associated to the edges incident to the black vertex number ℓ . (iv) Multiply the factors corresponding to the black vertices with the product \vee in $S(U)$. (v) Divide the resulting value by the order of the symmetry group of the tree.

8.2. The 1PI components of forms. The last step before we can write a connected form in terms of 1PI forms is to give a reasonable definition of what is the 1PI component of a form, similarly to the definition of the connected components of forms. Several definitions are possible. The simplest one was proposed by Hall [4] and has recently provided detailed structural results [54]. Here we consider a definition which is strictly more general than Hall's and that leads to an interesting structure. In a graph, it is easy to describe what we mean by cutting a line or a set

of lines; this approach leads, in classical QFT (with 2-point Feynman propagators) to the definition of 1PI Feynman diagrams as connected diagrams that are still connected when an arbitrary propagator line is cut. We propose to generalize the notion by replacing the Feynman form (that is, the classical case where only 2-point Feynman propagators are considered) by an arbitrary unital form.

Our approach is rooted in the Hopf algebraic picture of QFT. Notice however that our constructions could be translated *mutatis mutandis* in the language of functional derivatives. For example, the derivatives $\frac{\partial}{\partial \varphi(x)}$ that we have used in the definition of the operators A_i were defined as usual derivatives (in the polynomial algebra over the symbols $\varphi(x)$) but could be understood alternatively as functional derivatives. The same observation holds for our forthcoming constructions.

In proposition 6.1, for any $u = \varphi^{n_1}(x_1) \dots \varphi^{n_k}(x_k)$, we have expanded $\rho(u)$ as a linear combination of terms such as $r(\varphi^{n_1^1}(x_1) \dots \varphi^{n_k^1}(x_k)) \dots r(\varphi^{n_1^l}(x_1) \dots \varphi^{n_k^l}(x_k))$. Let us consider a connected term

$$t := r(\varphi^{n_1^1}(x_1) \dots \varphi^{n_k^1}(x_k)) \dots r(\varphi^{n_1^l}(x_1) \dots \varphi^{n_k^l}(x_k))$$

in $e^{*r}(u)$.

DEFINITION 8.1. *The connected term t is said to be one-particle reducible if and only if, there exists $i \in \{1, \dots, l\}$ and $\{i_1, \dots, i_p\} \subset \{1, \dots, k\}$ such that*

- (1) $\varphi^{n_1^i}(x_1) \dots \varphi^{n_k^i}(x_k) = \varphi(x_{i_1}) \dots \varphi(x_{i_p})$
- (2) furthermore, in the remaining part of t ,

$$\begin{aligned} r(\varphi^{n_1^1}(x_1) \dots \varphi^{n_k^1}(x_k)) \dots r(\varphi^{n_1^{i-1}}(x_1) \dots \varphi^{n_k^{i-1}}(x_k)) \\ r(\varphi^{n_1^{i+1}}(x_1) \dots \varphi^{n_k^{i+1}}(x_k)) \dots r(\varphi^{n_1^l}(x_1) \dots \varphi^{n_k^l}(x_k)) \end{aligned}$$

the connected components of x_{i_1}, \dots, x_{i_p} are either empty or pairwise disjoint.

A connected term that is not one-particle reducible is said to be one-particle irreducible (1PI).

For example, $r(\varphi(x_1)\varphi(x_2))^2$ (a loop constructed out of two two-point propagators) is 1PI (in our situation, and also in the usual picture), and so is

$$r(\varphi(x_1)\varphi(x_2))r(\varphi(x_1)\varphi(x_3))r(\varphi(x_1)\varphi(x_2)\varphi(x_3)),$$

whereas

$$r(\varphi(x_1)\varphi(x_2))r(\varphi(x_1)\varphi(x_3))r(\varphi(x_1)\varphi(x_2)\varphi(x_4))$$

or

$$r(\varphi(x_1)\varphi(x_2))r(\varphi(x_1)\varphi^2(x_3))r(\varphi(x_2)\varphi^2(x_4))$$

are not.

DEFINITION 8.2. *The 1PI component ρ_I of a form is the sum of all the 1PI terms in the connected component of ρ .*

Recall that forms, their connected components, and also their 1PI components are linear maps from $U = S(V)$ to \mathbb{C} . However, any such linear map l can be uniquely extended to a multiplicative map, still written l from $S(U)$ to \mathbb{C} (that is, to a character of the algebra $S(U)$, in the algebraic terminology). Concretely, for $u_1, \dots, u_n \in U$, $l(u_1 \vee \dots \vee u_n) := l(u_1) \dots l(u_n)$. In particular, the connected and 1PI components of forms can be viewed as characters of the algebra $S(U)$, so that,

for example, an expression such as $\rho_c \circ A_m$ makes sense as the composition of a derivation of $S(U)$ and a map from $S(U)$ to \mathbb{C} .

Let us consider a simple example to illustrate these ideas, namely the structure equation linking connected and 1PI components of forms in the most common picture of pQFT: an interacting theory –say φ^3 – with Feynman diagrams built of 3-valent interaction vertices and 2-point propagators. A general Feynman diagram can be described as 1PI diagrams connected by $n \in \mathbb{N}$ Feynman propagators satisfying the property that cutting any of these propagators makes the original diagram disconnected. Taking into account the symmetry factor $n!$ arising from the fact that these Feynman propagators can be cut successively in an arbitrary order results into a structure equation relating the connected and 1PI Green functions. In Hopf algebraic terms:

$$\rho_c = \rho_I \circ e^F$$

where $F = A_1$ is the derivation of $S(U)$ associated to the Feynman propagator: for all $u_1, \dots, u_n \in U = S(V)$,

$$\begin{aligned} F(u_1 \vee \dots \vee u_n) &= \frac{1}{2} \sum_{x \neq y} \sum_{i \leq n} D_2(x, y) u_1 \vee \dots \vee u_{i-1} \vee \\ &\quad \left(\frac{\bar{\partial} u_{i,\{1\}}}{\partial \varphi(x)} \vee \frac{\bar{\partial} u_{i,\{2\}}}{\partial \varphi(y)} \right) \vee u_{i+1} \vee \dots \vee u_n, \end{aligned}$$

with D_2 the (quasi-)free 2-point Green function.

In the general case, replacing Feynman propagators D_2 by arbitrary propagators D_n , doesn't change the general principles of the proof. An arbitrary Feynman diagram for an interacting theory as the ones considered previously in the present section can still be cut into 1PI pieces connected by a family of n -point propagators, $n \in N$, in such a way that removing any of these n -point propagators splits the original diagram into n connected pieces. For a given n , the associated symmetry factor is $p_n!$, where p_n is the number of n -point propagators in the family. These observations result in a family of structure identities for 1PI diagrams at all orders, and an identity that should be understood as a structure theorem for the perturbative approach to interacting theories.

THEOREM 8.1. (*Structure of connected forms*) *For an arbitrary connected form ρ_c , we have*

$$\rho_c = \rho_I \circ \mathbf{L}_\bullet = \rho_I \circ e^{\left(\sum_{n \in \mathbb{N}} A_n \right)},$$

and

$$\rho_I = \rho_c \circ e^{\left(- \sum_{n \in \mathbb{N}} A_n \right)}.$$

We remark that the effect of A_0 is just a shift of the fields: for instance $e^{A_0}(\varphi^n(x)) = (\varphi(x) + D_1(x))^n$. For $u = \varphi^{n_1}(x_1) \dots \varphi^{n_p}(x_p)$, we have $A_m(u) = 0$ if $m \geq p$ because A_m splits u into $m+1$ pieces and the coordinates x_i of these pieces must all be different. More generally, $L_n^k(u)$ vanishes if $n \geq p$. Because of the trivial effect of A_0 we put $A_0 = 0$ and we get

$$\rho_c(u) = \rho_I(u) + \sum_{n=1}^{p-1} \sum_{k=1}^{n_1 + \dots + n_p} \rho_I(L_n^k(u)).$$

In particular, $\rho_c(\varphi^n(x)) = \rho_I(\varphi^n(x))$ and, for $x \neq y$,

$$\rho_c(\varphi^n(x)\varphi^m(y)) = \rho_I(\varphi^n(x)\varphi^m(y)) + mnD_2(x,y)\rho_I(\varphi^{n-1}(x))\rho_I(\varphi^{m-1}(y)).$$

The first equation of theorem 8.1 describes the connected Green functions in terms of 1PI Green functions. It is an extension to general states of the standard QFT result and of a theorem by Mestre and Oeckl [20]. The second equation is new even in the QFT context: it describes the 1PI Green functions as a linear combination of products of connected Green functions. In QFT, 1PI Green functions are expressed in terms of amputated connected Green functions. Here, we do not amputate the Green functions (this is not allowed for a general state because parts of the Green functions belong to the kernel of the differential operator used in the equation of motion of the free field).

The consequences of these identities for the QFT of interacting systems, and the fine study of connected and 1PI amplitudes are postponed to further work.

9. Conclusion

In this paper, we developed mathematical tools to extend the relation between connected Green functions and 1PI Green functions from the case of a quasi-free ground state to the case of a general state. Our main result is the structure of connected forms described by theorem 8.1. This work can be extended in two directions. On the physical side, the main structure identity can be used to derive resummation theorems that generalize Friedberg's formulas [55]. On the mathematical side, many of our results can be extended to the case of non commuting variables.

10. Acknowledgements

This work originated as a part of a long term joint project on the combinatorics of QFT with Alessandra Frabetti. We thank her very warmly: without the many intense and inspiring discussions that we had with her, this paper would not exist.

We thank Raymond Stora for numerous fruitful comments on a previous version of the manuscript. We also thank warmly Achim Randelhoff who pointed out to us that the definition of the derivations A_{m-1} in a previous version of this article was not suited to their later application to 1-particle irreducibility.

We are grateful to Muriel Livernet for an inspiring remark about the relation between Ref. [19] and one of her works [56] and a proof that some results of Ref. [19] are valid in a general algebraic setting. The help of Jean-Louis Loday is gratefully acknowledged by Ch. B., who is also grateful to John Challifour for sending a copy of his unpublished paper with Wightman.

References

- [1] D. W. Robinson. The ground state of the Bose gas. *Commun. Math. Phys.*, 1:159–74, 1965.
- [2] B. S. Kay and R. M. Wald. Theorems on the uniqueness and thermal properties of stationary, nonsingular, quasifree states on spacetimes with a bifurcate Killing horizon. *Phys. Repts.*, 207:49–136, 1991.
- [3] B. S. Kay. Sufficient conditions for quasifree states and an improved uniqueness theorem for quantum fields on space-times with horizons. *J. Math. Phys.*, 34:4519–39, 1993.
- [4] A. G. Hall. Non-equilibrium Green functions: Generalized Wick's theorem and diagrammatic perturbation theory with initial correlations. *J. Phys. A: Math. Gen.*, 8:214–25, 1975.
- [5] S.H. Djah, H. Gottschalk, and H. Ouerdiane. Feynman graph representation of the perturbation series for general functional measures. *J. Funct. Anal.*, 227:153–187, 2005.

- [6] D. Ruelle. *Statistical Mechanics, Rigorous Results*. W.A. Benjamin Inc., New York, 1969.
- [7] H. J. Borchers. Algebraic aspects of Wightman quantum field theory. In H. Araki, editor, *International Symposium on Mathematical Problems in Theoretical Physics*, volume 39 of *Lecture Notes in Physics*, pages 283–92, Berlin, 1975. Springer.
- [8] R. Stora. Lagrangian field theory. In C. DeWitt and C. Itzykson, editors, *Particle Physics – Les Houches 1971*, pages 1–79, New York, 1973. Gordon and Breach.
- [9] A. S. Wightman and J. L. Challifour. On the existence and properties of a two-dimensional quantum field. *Unpublished*, 69 pages, 1970.
- [10] W. Kutzelnigg and D. Mukherjee. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.*, 107:432–49, 1997.
- [11] Ch. Brouder, B. Fauser, A. Frabetti, and R. Oeckl. Quantum field theory and Hopf algebra cohomology. *J. Phys. A: Math. Gen.*, 37:5895–927, 2004.
- [12] D. Kreimer. On the Hopf algebra structure of perturbative quantum field theory. *Adv. Th. Math. Phys.*, 2:303–34, 1998.
- [13] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Commun. Math. Phys.*, 199:203–42, 1998.
- [14] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann–Hilbert problem I: the Hopf algebra structure of graphs and the main theorem. *Commun. Math. Phys.*, 210:249–73, 2000.
- [15] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann–Hilbert problem II: the β function, diffeomorphisms and the renormalization group. *Commun. Math. Phys.*, 216:215–41, 2001.
- [16] K. Ebrahimi-Fard, J. M. Gracia-Bondia, and F. Patras. A Lie theoretical approach to renormalization. *Commun. Math. Phys.*, 276:519–49, 2007.
- [17] B. Fauser. On the Hopf algebraic origin of Wick normal-ordering. *J. Phys. A: Math. Gen.*, 34:105–116, 2001.
- [18] A.C. Hirshfeld and P. Henselder. Star products and quantum groups in quantum mechanics and field theory. *Ann. Phys.*, 308:311–28, 2003.
- [19] A. Mestre and R. Oeckl. Combinatorics of n -point functions via Hopf algebra in quantum field theory. *J. Math. Phys.*, 47:052301, 2006.
- [20] A. Mestre and R. Oeckl. Generating loop graphs via Hopf algebra in quantum field theory. *J. Math. Phys.*, 47:122302, 2006.
- [21] W. D. van Suijlekom. Renormalization of gauge fields: A Hopf algebra approach. *Commun. Math. Phys.*, 276:773–98, 2007.
- [22] Ch. Brouder. Quantum field theory meets Hopf algebra. *Math. Nachr.*, 282:1664–90, 2009.
- [23] S. A. Joni and G.-C. Rota. Coalgebras and bialgebras in combinatorics. *Stud. Appl. Math.*, 61:93–139, 1979.
- [24] Ch. Brouder. Green function hierarchy for open shells. *Euro. Phys. Lett.*, 71:556–62, 2005.
- [25] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. II Fourier Analysis, Self-adjointness*. Academic Press, New York, 1975.
- [26] M. Reed and B. Simon. *Methods of Modern Mathematical Physics. I Functional Analysis*. Academic Press, New York, second edition, 1980.
- [27] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-Particle Systems*. McGraw-Hill, Boston, 1971.
- [28] M. Gell-Mann and F. Low. Bound states in quantum field theory. *Phys. Rev.*, 84:350–4, 1951.
- [29] G. Nenciu and G. Rasche. Adiabatic theorem and Gell-Mann-Low formula. *Helv. Phys. Acta*, 62:372–88, 1989.
- [30] Ch. Brouder, G. Stoltz, and G. Panati. Adiabatic approximation, Gell-Mann and Low theorem and degeneracies: A pedagogical example. *Phys. Rev. A*, 78:042102, 2008.
- [31] Ch. Brouder, G. Panati, and G. Stoltz. Gell-Mann and Low formula for degenerate unperturbed states. *Ann. Henri Poincaré*, 10:1285–309, 2010.
- [32] Ch. Brouder, G. Panati, and G. Stoltz. The many-body Green function of degenerate systems. *Phys. Rev. Lett.*, 103:230401, 2009.
- [33] Ch. Brouder and F. Patras. Hyperoctahedral Chen calculus for effective Hamiltonians. *J. Algebra* 322:4105–4120, 2009.
- [34] Ch. Brouder, Å. Mestre and F. Patras Tree expansion in time-dependent perturbation theory *J. Math. Phys.* 51, 072104, 2010.

- [35] G. Onida, L. Reining, and A. Rubio. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.*, 74:601–59, 2002.
- [36] E. K. U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory*. Adam Hilger, Bristol, 1991.
- [37] K.-C. Chou, Z.-B. Su, B.-L. Hao, and L. Yu. Equilibrium and nonequilibrium formalisms made unified. *Phys. Repts.*, 118:1–131, 1985.
- [38] R. Yaris and H. S. Taylor. Many-body Green's function method for electron scattering from open-shell systems. *Phys. Rev. A*, 12:1751–9, 1975.
- [39] C. Bloch and J. Horowitz. Sur la détermination des premiers états d'un système de fermions dans le cas dégénéré. *Nucl. Phys.*, 8:91–105, 1958.
- [40] T. Morita. Perturbation theory for degenerate problems of many-fermion systems. *Prog. Theor. Phys.*, 29:351–69, 1963.
- [41] S. Fujita. *Introduction to Non Equilibrium Quantum Statistical Mechanics*. Saunders, Philadelphia, 1966.
- [42] S. Majid. *Foundations of Quantum Group Theory*. Cambridge University Press, Cambridge, 1995.
- [43] M. E. Sweedler. *Hopf Algebras*. W. A. Benjamin, New York, 1969.
- [44] J.C. Collins. *Renormalization*. Cambridge University Press, Cambridge, 1984.
- [45] H. Scutaru. Transition probabilities between quasifree states. *J. Math. Phys.*, 39:6403–15, 1998.
- [46] V. A. Moskalenko, P. Entel, D. F. Digor, L. A. Dohotaru, and R. Citro. A diagram approach to the strong coupling in the single-impurity Anderson model. *Theor. Math. Phys.*, 155:914–35, 2008.
- [47] S. Hollands and R. M. Wald. Local Wick polynomials and time ordered products of quantum fields in curved spacetime. *Commun. Math. Phys.*, 223:289–326, 2001.
- [48] S. G. Thikhodeev. Relations between many-body Green's functions and correlation functions. *Sov. Phys. Doklady*, 27:492–3, 1982.
- [49] T. T. S. Kuo, S. Y. Lee, and K. F. Ratcliff. A folded-diagram expansion of the model-space effective Hamiltonian. *Nucl. Phys. A*, 176:65–88, 1971.
- [50] A. L. Kitaev. Nonstationary perturbation theory for a degenerate discrete level. *Theor. Math. Phys.*, 25:1224–7, 1975.
- [51] B. Fauser and P. D. Jarvis. A Hopf laboratory for symmetric functions. *J. Phys. A: Math. Gen.*, 37:1633–63, 2004.
- [52] B. Fauser, P. D. Jarvis, R. C. King, and B. G. Wybourne. New branching rules induced by plethysm. *J. Phys. A: Math. Gen.*, 39:2611–55, 2006.
- [53] F. Patras and C. Reutenauer. On Dynkin and Klyachko idempotents in graded bialgebras. *Adv. Appl. Math.*, 28:560–79, 2002.
- [54] Ch. Brouder. The structure of Green functions in quantum field theory with a general state. In B. Fauser, J. Tolksdorf, and E. Zeidler, editors, *Quantum Field Theory – Competitive Models*, pages 163–75, Basel, 2009. Birkhäuser.
- [55] R. Friedberg. Dual trees and resummation theorems. *J. Math. Phys.*, 16:20–30, 1975.
- [56] M. Livernet. A rigidity theorem for preLie algebras. *J. Pure Appl. Alg.*, 207:1–18, 2006.

INSTITUT DE MINÉRALOGIE ET DE PHYSIQUE DES MILIEUX CONDENSÉS, CNRS UMR7590, UNIVERSITÉS PARIS 6 ET 7, IPGP, 140 RUE DE LOURMEL, 75015 PARIS, FRANCE.

UNIVERSITÉ DE NICE, LABORATOIRE J.-A. DIEUDONNÉ, CNRS UMR 6621, PARC VALROSE, 06108 NICE CEDEX 02, FRANCE.

This page intentionally left blank

Multiple zeta values and periods: From moduli spaces to Feynman integrals

Francis Brown

ABSTRACT. The aim of these notes is to draw an analogy between period integrals on moduli spaces $\mathfrak{M}_{0,n}$ of genus 0 curves, and certain Feynman integrals when expressed in Schwinger coordinates. In the first half, we survey the main ideas in the proof [6] of a conjecture due to Goncharov and Manin, which states that the periods of $\mathfrak{M}_{0,n}$ are multiple zeta values. After reviewing the recent motivic approach to the study of Feynman integrals [3], we then show how similar ideas can be applied to the Feynman case to explain the prevalence of multiple zeta values and polylogarithms in perturbative quantum field theory. This leads to a new analytic approach to computing Feynman amplitudes and suggests some problems for future research.

1. Introduction

First, we define the moduli spaces we will consider. Let $n \geq 4$, and let $\mathfrak{M}_{0,n}$ denote the moduli space of genus 0 curves with n ordered marked points:

$$\mathfrak{M}_{0,n} = \{(z_1, \dots, z_n) \in \mathbb{P}^1 : z_i \text{ distinct}\} / \text{PSL}_2.$$

This is an affine variety whose dimension we will always denote by $\ell = n - 3$. We can place $z_1 = 1$, $z_2 = \infty$, $z_3 = 0$, and call the remaining points

$$t_1 = z_4, \quad t_2 = z_5, \quad \dots, \quad t_\ell = z_n.$$

We can therefore think of $\mathfrak{M}_{0,n}(\mathbb{C})$ as the complement of the following hyperplanes in affine space:

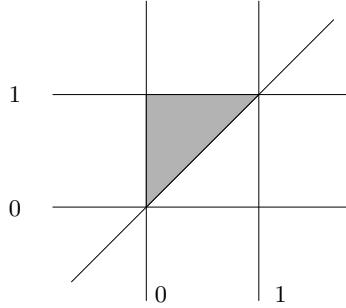
$$\mathfrak{M}_{0,n}(\mathbb{C}) = \{(t_1, \dots, t_\ell) \in \mathbb{C}^\ell : t_i \neq 0, 1, \quad t_i \neq t_j\}.$$

In particular, $\mathfrak{M}_{0,4}$ is just the projective line minus three points $\mathbb{P}^1 \setminus \{0, 1, \infty\}$, and $\mathfrak{M}_{0,5}(\mathbb{C})$ is the complement of five lines in \mathbb{C}^2 pictured below.

The simplest example of the kind of integrals we shall consider is

$$(1) \quad I = \int_X \frac{dt_1 dt_2}{(1 - t_1)t_2},$$

where the domain of integration X is the simplex $\{0 \leq t_1 \leq t_2 \leq 1\}$ and is shaded in the figure. It was shown by Leibniz in the 18th century that $I = \zeta(2)$. The integral I , considered as a Lebesgue integral, converges, but something undesirable occurs, namely, the integrand has singularities on the domain of integration in two places: $(0, 0)$ and $(1, 1)$. The boundary of the domain X consists of three lines:



$t_1 = 0$, $t_1 = t_2$ and $t_2 = 1$, and the singularities of the integrand consists of $t_1 = 1$, and $t_2 = 0$. These do not cross normally at the two points $(0,0)$ and $(1,1)$ (there are three divisors which meet at each point) so we must blow them up, which has the effect of moving the singularities of the integrand away from X . To see this, consider the following picture of $\overline{\mathcal{M}}_{0,5}$, where we keep track of what goes on at infinity.

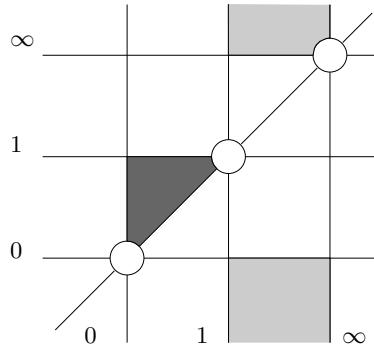


FIGURE 1. $\overline{\mathcal{M}}_{0,5}(\mathbb{R})$. The three bad points $(0,0)$, $(1,1)$ and (∞, ∞) have been replaced with real projective circles. The ten divisors split into two sets of five: a set B at finite distance which border the dark shaded cell, and a set A at infinity which border the light shaded cell.

After blowing up the three points $(0,0)$, $(1,1)$, (∞, ∞) in $\mathbb{P}^1 \times \mathbb{P}^1$, we obtain a smooth projective variety denoted $\overline{\mathcal{M}}_{0,5}$. The complement $\overline{\mathcal{M}}_{0,5} \setminus \mathcal{M}_{0,5}$ has 10 irreducible components which cross normally. There are now five divisors which bound the domain of integration X , and these form a pentagon. The remaining five divisors form another pentagon at infinity. Let $B = \bigcup_{i=1}^5 B_i$ denote the first set of divisors (said to be at finite distance), and let $A = \bigcup_{i=1}^5 A_i$ denote the remaining five at infinity. The boundary of the domain of integration is contained in B , and one can check that all the poles of the integrand ω are contained in A :

$$\partial X \subset B \quad \text{and} \quad \text{Sing } \omega \subset A .$$

Since A is disjoint from X , the integral I , considered as an integral on $\overline{\mathcal{M}}_{0,5}$,

$$I = \int_X \omega$$

now has the property that the integrand has no singularities on the domain of integration. This gives rise to the following cohomological interpretation. The domain of integration defines a relative homology class $[X] \in H_2(\overline{\mathfrak{M}}_{0,5}, B)$, and the integrand ω defines a class $[\omega] \in H^2(\overline{\mathfrak{M}}_{0,5} \setminus A)$. The underlying cohomology group (or motive) we must consider is therefore

$$(2) \quad H^2(\overline{\mathfrak{M}}_{0,5} \setminus A, B \setminus (B \cap A)) .$$

One can show that this is a \mathbb{Q} -vector space of dimension two. It has as a basis the forms

$$\omega_0 = 2 dt_1 dt_2 , \quad \omega_1 = \frac{dt_1 dt_2}{(1-t_1)t_2} .$$

Dually, one shows that a \mathbb{Q} -basis for the homology $H_2(\overline{\mathfrak{M}}_{0,5} \setminus A, B \setminus (B \cap A))$ is given by the class $[X]$ and by the class of the torus at infinity

$$T = \{|1-t_1| = \varepsilon, |t_2| = \varepsilon\} ,$$

where $0 < \varepsilon \ll 1$. A period matrix for (2) is given by integrating a basis of differential forms against a basis of homology classes. It follows from Leibniz's result, and the residue formula, that a period matrix is:

$$\begin{pmatrix} \int_X \omega_0 & \int_T \omega_0 \\ \int_X \omega_1 & \int_T \omega_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \zeta(2) & (2i\pi)^2 \end{pmatrix} .$$

In the general case, the periods of the moduli spaces $\mathfrak{M}_{0,n}$ are expressible in terms of multiple zeta values. Recall that if $n_1, \dots, n_r \in \mathbb{N}$, where $n_r \geq 2$, then the multiple zeta value $\zeta(n_1, \dots, n_r)$ is defined by the sum

$$\zeta(n_1, \dots, n_r) = \sum_{0 < k_1 < \dots < k_r} \frac{1}{k_1^{n_1} \dots k_r^{n_r}} .$$

Its *weight* is the quantity $n_1 + \dots + n_r$, and its *depth* is the quantity r . We will write $\mathcal{Z} = \mathbb{Q}[\zeta(2), \zeta(3), \dots]$ to denote the \mathbb{Q} -algebra generated by all multiple zeta values [24]. It is filtered by the weight.

There is a smooth compactification

$$\mathfrak{M}_{0,n} \subset \overline{\mathfrak{M}}_{0,n}$$

due to Deligne, Mumford, and Knudsen, such that $\overline{\mathfrak{M}}_{0,n} \setminus \mathfrak{M}_{0,n}$ is a smooth normal crossing divisor. Let $A, B \subset \overline{\mathfrak{M}}_{0,n} \setminus \mathfrak{M}_{0,n}$ denote two sets of divisors such that A and B share no irreducible components. We can consider the relative cohomology group:

$$H^\ell(\overline{\mathfrak{M}}_{0,n} \setminus A, B \setminus (B \cap A)) .$$

Goncharov and Manin [15] proved that this defines a mixed Tate motive unramified over \mathbb{Z} . A general conjecture due to Goncharov [11, 13] then predicts that the periods of any such motive are expressible in terms of multiple zeta values. Let X denote a smooth compact real submanifold of dimension ℓ whose boundary is contained in B , and let ω denote a rational ℓ -form on $\overline{\mathfrak{M}}_{0,n}$ whose singularities are contained in A . The integral

$$(3) \quad \int_X \omega ,$$

converges, and defines a period of this motive. In these notes we shall only consider the following restricted class of integrals which generalise (1):

$$(4) \quad I = \int_X \frac{\prod_{i=1}^{\ell} t_i^{a_i} (1-t_i)^{b_i}}{\prod_{1 \leq i < j \leq \ell} (t_i - t_j)^{c_{ij}}} dt_1 \dots dt_{\ell},$$

where X is the unit simplex $X = \{0 \leq t_1 \leq t_2 \leq \dots \leq t_{\ell} \leq 1\}$, and a_i, b_i, c_{ij} are integers which satisfy certain conditions to ensure convergence.

THEOREM 1. [6] *The period I is in \mathcal{Z} , of weight at most ℓ .*

A related result is given in [23]. In these notes, we will sketch some of the main ideas of the proof of the above theorem, and indicate how these techniques shed some light on the periods obtained from Feynman diagrams.

- (1) First we give an explicit construction of the spaces $\overline{\mathfrak{M}}_{0,n}$, and their underlying combinatorics, using what we call dihedral coordinates.
- (2) The main idea for computing the period integrals (3) is to do an induction by applying Stokes' formula. This involves taking primitives, which takes us outside the class of algebraic functions on $\mathfrak{M}_{0,n}$.
- (3) We define an algebra of generalised multiple polylogarithm functions on $\mathfrak{M}_{0,n}$ and explain how this algebra of functions is closed under the operation of taking primitives, and gives rise to an algorithm for computing the integrals (4).
- (4) We draw some analogies between the underlying geometry of the moduli spaces and that of graph polynomials. This suggests an analytic method for computing Feynman amplitudes for a certain class of graphs which have been verified to give multiple zeta values numerically. It also leads to the definition of a certain cobracket on the algebra of graphs, which has been considered by S. Bloch in [2].
- (5) Finally, we outline some further striking similarities between periods of moduli spaces and Feynman integrals which merit further study.

2. An explicit model for $\overline{\mathfrak{M}}_{0,n}$

From now on, let S denote a set with $n \geq 4$ elements which label the marked points on \mathbb{P}^1 . We shall write $\mathfrak{M}_{0,S}$ instead of $\mathfrak{M}_{0,n}$. First of all, we need to give an explicit description of the compactification $\overline{\mathfrak{M}}_{0,S}$ by gluing together a set of affine charts $\mathfrak{M}_{0,S}^{\delta}$, which depend on a certain parameter δ which is defined below. The $\mathfrak{M}_{0,S}^{\delta}$ are intermediary spaces

$$\mathfrak{M}_{0,S} \subset \mathfrak{M}_{0,S}^{\delta} \subset \overline{\mathfrak{M}}_{0,S},$$

and are defined explicitly below.

2.1. Dihedral Coordinates. The spaces $\mathfrak{M}_{0,S}^{\delta}$ are defined by explicit polynomials in certain variables which we call dihedral coordinates. To write these down, we first define a *dihedral structure* δ on S to be an identification of the elements of S with the sides of a regular n -gon, considered up to symmetries. In the case when $S = \{s_1, \dots, s_n\}$, we simply label the edges of a regular n -gon with the numbers $1, \dots, n$ for convenience.

Now let $\chi_{S,\delta}$ denote the set of chords in the labelled polygon (S, δ) . Each chord joins two vertices and meets four edges in this polygon. To simplify notations, we

can also label the vertices of (S, δ) . The convention is that the vertex labelled (i) in parentheses, meets edges i and $i+1$. Then an element $\{i, j\} \in \chi_{S, \delta}$ can be depicted as the chord joining (i) and (j) . Each such chord defines a *dihedral coordinate* which is a cross-ratio:

$$u_{ij} = \frac{(z_i - z_{j+1})(z_{i+1} - z_j)}{(z_i - z_j)(z_{i+1} - z_{j+1})}.$$

Since the cross-ratio is PSL_2 -invariant, it defines a function $u_{ij} : \mathfrak{M}_{0,S} \rightarrow \mathbb{P}^1 \setminus \{0, 1, \infty\}$. All these functions taken together give an embedding:

$$(u_{ij})_{\{i,j\} \in \chi_{S, \delta}} : \mathfrak{M}_{0,S} \longrightarrow \mathbb{A}^{n(n-3)/2},$$

where $n(n-3)/2$ is the number of chords in the n -gon (S, δ) . Finally, we define the *dihedral extension* $\mathfrak{M}_{0,S}^\delta$ of $\mathfrak{M}_{0,S}$ to be the Zariski closure of the image of this embedding.

EXAMPLES 2. Consider the case $n = 4$. The moduli space $\mathfrak{M}_{0,4}$ is just $\mathbb{P}^1 \setminus \{0, 1, \infty\}$, whose coordinate is denoted t_1 . There are two chords in a square, and the corresponding dihedral coordinates are $u_{13} = t_1$, $u_{24} = 1 - t_1$. The dihedral embedding is given by the map

$$\begin{aligned} (u_{13}, u_{24}) : \mathfrak{M}_{0,4} &\longrightarrow \mathbb{A}^2 \\ t_1 &\mapsto (1 - t_1, t_1) \end{aligned}$$

The Zariski closure of the image is the affine line \mathbb{A}^1 , and meets the coordinate axes in two places corresponding to the points $t_1 = 0, t_1 = 1$. Therefore $\mathfrak{M}_{0,4}^\delta \cong \mathbb{A}^1$ with two distinguished points $u_{13} = 0, u_{24} = 0$. Thus $\mathfrak{M}_{0,4}^\delta$ is obtained by adding back the points 0 and 1 to $\mathfrak{M}_{0,4}$, but not ∞ .

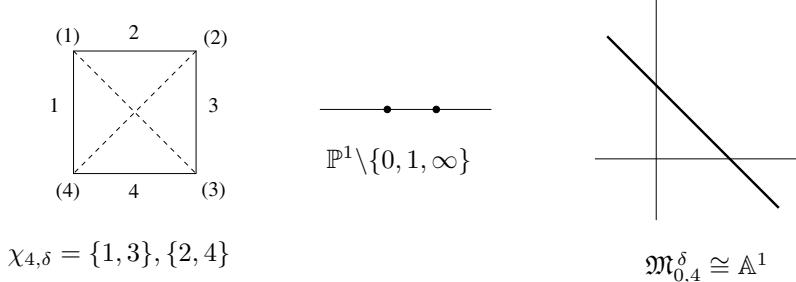


FIGURE 2

Now consider the case $n = 5$. A pentagon has five chords $\{13, 24, 35, 41, 52\}$, pictured below (left). The corresponding dihedral coordinates are:

$$\begin{aligned} u_{13} &= 1 - t_1, \quad u_{24} = \frac{t_1}{t_2}, \quad u_{35} = \frac{t_2 - t_1}{t_2(1 - t_1)} \\ u_{41} &= \frac{1 - t_2}{1 - t_1}, \quad u_{52} = t_2. \end{aligned}$$

This gives an embedding $\mathfrak{M}_{0,5} \hookrightarrow \mathbb{A}^5$. The dihedral extension $\mathfrak{M}_{0,5}^\delta$ is therefore an affine surface embedded in \mathbb{A}^5 , and has five canonical divisors given by $u_{ij} = 0$ for each chord $\{i, j\} \in \chi_{5,\delta}$. We will see later on that these form a pentagon. One can check that the effect of passing to $\mathfrak{M}_{0,5}^\delta$ is to blow up the two points $(0, 0)$ and $(1, 1)$

in the picture of $\mathfrak{M}_{0,5}$ below (middle). The extension $\mathfrak{M}_{0,5}^\delta$ (right) is obtained from $\mathfrak{M}_{0,5}$ by adding back the five divisors which bound the simplex $0 \leq t_1 \leq t_2 \leq 1$. In the introduction, these were referred to as the five divisors at finite distance.

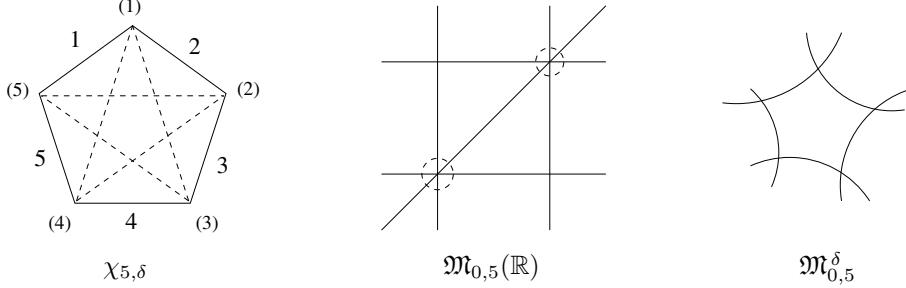


FIGURE 3

We can write down the five equations which define $\mathfrak{M}_{0,5}^\delta$. They are

$$u_{13} + u_{24}u_{25} = 1$$

along with the four other cyclic permutations of this equation in $\mathbb{Z}/5\mathbb{Z}$. There is a similar set of equations in the general case [6].

It is no accident that the five dihedral coordinates above are exactly the five terms which appear in the functional equation of the dilogarithm.

THEOREM 3. $\mathfrak{M}_{0,S}^\delta$ is a smooth affine scheme given by explicit equations defined over \mathbb{Z} . There is an embedding $\mathfrak{M}_{0,S} \subset \mathfrak{M}_{0,S}^\delta$. The complement $\mathfrak{M}_{0,S}^\delta \setminus \mathfrak{M}_{0,S}$ is a normal crossing divisor, whose irreducible components are

$$D_{ij} = \{u_{ij} = 0\} ,$$

indexed by each chord $\{i, j\} \in \chi_{S,\delta}$.

We do indeed retrieve the standard description of $\overline{\mathfrak{M}}_{0,S}$ in this fashion.

THEOREM 4. The Deligne-Mumford-Knudsen compactification is obtained by gluing the affine charts $\mathfrak{M}_{0,S}^\delta$ together:

$$\overline{\mathfrak{M}}_{0,S} = \bigcup_{\delta} \mathfrak{M}_{0,S}^\delta ,$$

where δ ranges over the set of $n!/2n$ dihedral structures on S .

An advantage of the description in terms of dihedral coordinates is that the underlying combinatorics of the moduli spaces is immediately transparent. We need to look a little more closely at this combinatorial structure.

2.2. Product structure of divisors. The key property is that each divisor D_{ij} on $\mathfrak{M}_{0,S}^\delta$ is itself a product of spaces of the same type, but of smaller dimension. It is this property which allows us to do inductions. The same feature will reappear when we consider the Feynman situation later.

To see this, cut the polygon (S, δ) along a chord $\{i, j\}$. It decomposes into two smaller polygons $(S_1 \cup \{e\}, \delta_1)$ and $(S_2 \cup \{e\}, \delta_2)$, where the new edge e corresponds

to the chord $\{i, j\}$. Then the divisor $D_{ij} = \{u_{ij} = 0\}$ is canonically isomorphic to a product:

$$D_{ij} \cong \mathfrak{M}_{0, S_1 \cup \{e\}}^{\delta_1} \times \mathfrak{M}_{0, S_2 \cup \{e\}}^{\delta_2}.$$

In the picture below, we have cut a hexagon along the chord $\{2, 5\}$. When we set the dihedral coordinate u_{25} to 0, the coordinates corresponding to the four chords which cross $\{2, 5\}$ (the dotted lines on the left), become 1. The set of dihedral coordinates then splits into two sets of two coordinates each (right). These correspond to the chords in each smaller quadrilateral. We deduce that in this case $D_{25} \cong \mathfrak{M}_{0,4}^{\delta} \times \mathfrak{M}_{0,4}^{\delta} \cong \mathbb{A}^1 \times \mathbb{A}^1$.

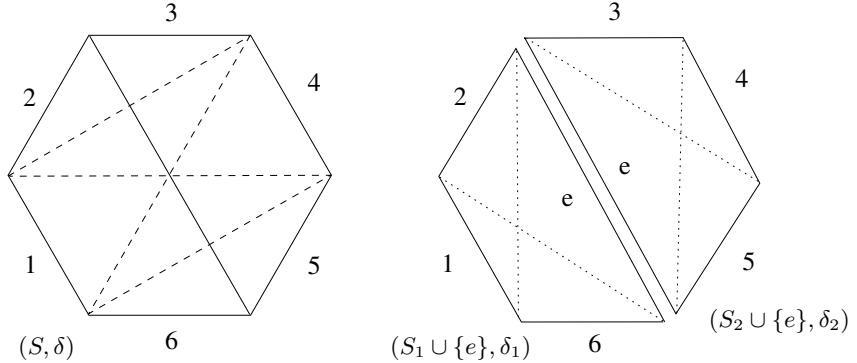


FIGURE 4

2.3. Fundamental cells and $\mathfrak{M}_{0,S}^{\delta}(\mathbb{R})$. In order to write down some period integrals, we need a domain of integration. To do this, consider the set of real points $\mathfrak{M}_{0,S}(\mathbb{R})$, which is not connected (fig. 1). Each connected component is a contractible cell, and can be written in dihedral coordinates:

$$X_{S,\delta} = \{0 < u_{ij} : \{i, j\} \in \chi_{S,\delta}\} \subset \mathfrak{M}_{0,S}(\mathbb{R})$$

In the examples considered earlier, $X_{4,\delta}$ corresponded to the unit interval $(0, 1) \subset \mathbb{R}$ in the case $\mathfrak{M}_{0,4}$, and $X_{5,\delta}$ corresponded to the unit simplex $\{0 < t_1 < t_2 < 1\}$ in the case $\mathfrak{M}_{0,5}$.

The variety $\mathfrak{M}_{0,S}(\mathbb{R})$ is covered by the set of all such cells

$$\mathfrak{M}_{0,S}(\mathbb{R}) = \bigsqcup_{\delta} X_{S,\delta},$$

where δ ranges over the set of dihedral structures on S . In other words, choosing a dihedral structure δ on S is just the same thing as choosing a connected component in the set of real points $\mathfrak{M}_{0,S}(\mathbb{R})$. The dihedral extension $\mathfrak{M}_{0,S}^{\delta}$ can be obtained by adding to $\mathfrak{M}_{0,S}$ all the divisors which bound this connected component, after desingularisation. Alternatively, $\mathfrak{M}_{0,S}^{\delta}$ is obtained by removing from $\overline{\mathfrak{M}}_{0,S}$ all divisors at infinity, except those which bound a given, fixed cell $X_{S,\delta}$. The dihedral extension $\mathfrak{M}_{0,S}^{\delta}(\mathbb{R})$ is precisely what we need to consider the closed cell

$$\overline{X}_{S,\delta} = \{0 \leq u_{ij} : \{i, j\} \in \chi_{S,\delta}\}.$$

It is bounded by the divisors D_{ij} and is called a *Stasheff polytope*.

2.4. Stasheff polytopes. The Stasheff polytope is a polytope defined combinatorially by parenthesized words or by the chords in a polygon [12]. Its three-dimensional version (fig. 6) is frequently called the associahedron. In this setting, the cell $\overline{X}_{S,\delta}$ is an explicit algebraic model of the Stasheff polytope, given by dihedral coordinates. Since these are indexed by the set of chords in an n -gon, we immediately deduce the following combinatorial description of $\overline{X}_{S,\delta}$:

- Its facets (codimension one faces) F_{ij} are indexed by the set of chords $\{i, j\}$. Each facet is a product

$$F_{ij} = \overline{X}_{S_1 \cup \{e\}, \delta_1} \times \overline{X}_{S_2 \cup \{e\}, \delta_2}$$

which corresponds to cutting an n -gon along the chord $\{i, j\}$.

- Two facets F_{ij} and F_{kl} meet if and only if the chords $\{i, j\}$ and $\{k, l\}$ do not cross.
- Faces of codimension k are given by sets of k non-crossing chords. In particular, the set of vertices of $\overline{X}_{S,\delta}$ are in a one-to-one correspondence with the set of triangulations of the n -gon (S, δ) .

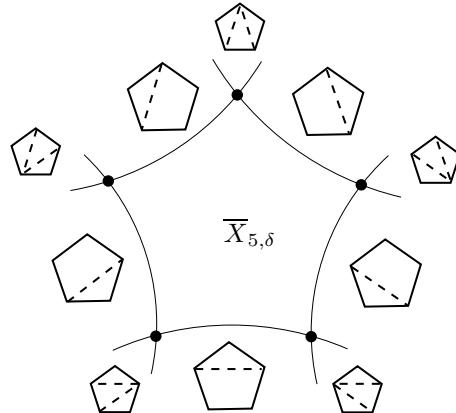
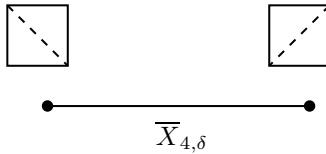


FIGURE 5

In order to see the product structure on the facets, we have to consider the associahedron $\overline{X}_{6,\delta}$ of dimension 3, pictured below. It has six facets which are pentagons, which correspond to the short chords in a hexagon:

$$\overline{X}_{5,\delta_1} \times \overline{X}_{3,\delta_2} \cong \overline{X}_{5,\delta_1} \times \{pt\} .$$

The remaining three facets are quadrilaterals, and are indexed by the long chords in a hexagon:

$$\overline{X}_{4,\delta_1} \times \overline{X}_{4,\delta_2} \cong [0, 1] \times [0, 1] .$$

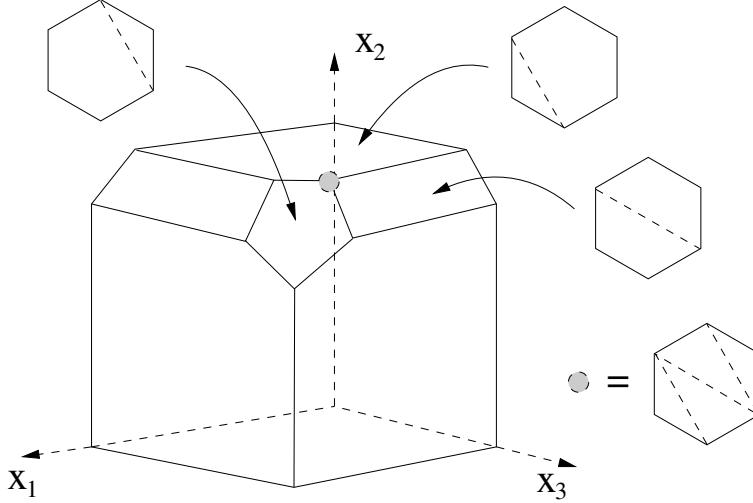


FIGURE 6

2.5. Period integrals over the fundamental cell. We can write down a set of canonical period integrals on $\mathfrak{M}_{0,S}$ as follows. First, one can prove that $\mathfrak{M}_{0,S}^\delta$ has a canonical regular ℓ -form $\omega_{S,\delta}$ which has neither zeros nor poles on $\mathfrak{M}_{0,S}^\delta$, *i.e.*, along $\mathfrak{M}_{0,S}$ and the divisors D_{ij} . It therefore defines a volume form on the set of real points $\mathfrak{M}_{0,S}^\delta(\mathbb{R})$.

For each chord $\{i, j\} \in \chi_{S,\delta}$, choose an integer $\alpha_{ij} \in \mathbb{Z}$. We define

$$(5) \quad I(\underline{\alpha}) = \int_{\overline{X}_{S,\delta}} \prod_{\{i,j\} \in \chi_{S,\delta}} u_{ij}^{\alpha_{ij}} \omega_{S,\delta} .$$

If all the parameters α_{ij} are non-negative, then the integrand has no singularities along the domain of integration. Since $\overline{X}_{S,\delta}$ is a compact subset of the open set $\mathfrak{M}_{0,S}^\delta(\mathbb{R}) \subset \overline{\mathfrak{M}}_{0,S}(\mathbb{R})$, and the integrand is continuous, it is clear that it converges. The converse is also true: (5) converges if and only if $\alpha_{ij} \geq 0$ for all $\{i, j\} \in \chi_{S,\delta}$. This family of integrals is indeed the desingularised version of (4). We now rewrite it using *cubical coordinates*: set

$$t_1 = x_1 \dots x_\ell, \quad \dots, \quad t_{\ell-1} = x_{\ell-1} x_\ell, \quad t_\ell = x_\ell$$

then $\mathfrak{M}_{0,n}(\mathbb{C})$ is the complement of hyperbolae:

$$\{(x_1, \dots, x_\ell) \in \mathbb{C}^\ell : x_i \neq 0, 1, \quad x_i \dots x_j \neq 1\} .$$

In cubical coordinates, the volume form can be written:

$$\omega_{S,\delta} = \frac{dx_1 \dots dx_\ell}{(1 - x_1 x_2)(1 - x_2 x_3) \dots (1 - x_{\ell-1} x_\ell)} .$$

EXAMPLES 5. In the case $\mathfrak{M}_{0,5}$, we associate 5 parameters h, i, j, k, l to each chord in a pentagon. The integral (5) can be written:

$$I(\underline{\alpha}) = \int_{\overline{X}_{5,\delta}} u_{52}^h u_{41}^i u_{35}^j u_{24}^k u_{13}^l \omega_{5,\delta} .$$

If we re-write it in cubical coordinates, we obtain

$$(6) \quad I(h, i, j, k, l) = \int_0^1 \int_0^1 \frac{x_1^h (1-x_1)^i x_2^k (1-x_2)^j}{(1-x_1 x_2)^{i+j-l}} \frac{dx_1 dx_2}{1-x_1 x_2} .$$

The main theorem implies that $I(h, i, j, k, l) \in \mathbb{Q} + \mathbb{Q} \zeta(2)$. This family of integrals was first considered by Dixon in 1905, and was used by Rhin and Viola to obtain the best irrationality measures for $\zeta(2)$ known to this date.

For arbitrary n , $I(\underline{\alpha})$ can be written as a generalised Selberg integral

$$\int_{[0,1]^\ell} \frac{\prod_i x_i^{a_i} (1-x_i)^{b_i}}{\prod_{i < j} (1-x_i \dots x_j)^{c_{ij}}} dx_1 \dots dx_\ell ,$$

where $a_i, b_i, c_{ij} \in \mathbb{Z}$. If these are instead considered as complex parameters, then this can be thought of as a generalised beta function.

3. Stokes' formula on the polytope $\overline{X}_{S,\delta}$

The basic idea for computing the integrals $I(\underline{\alpha})$ is to apply Stokes' formula. We wish to compute the period

$$I = \int_{\overline{X}_{S,\delta}} f ,$$

where f is a regular algebraic form on $\mathfrak{M}_{0,S}$ which has no singularities along $\overline{X}_{S,\delta}$. Suppose that f has a primitive P . Then

$$I = \int_{\partial \overline{X}_{S,\delta}} P = \sum_{\{i,j\} \in \chi_{S,\delta}} \int_{F_{ij}} P \Big|_{F_{ij}} .$$

Recall that each face F_{ij} is a product of Stasheff polytopes $\overline{X}_{T_1, \delta_1} \times \overline{X}_{T_2, \delta_2}$. Therefore I is a sum of products of integrals

$$\int_{\overline{X}_{T_1, \delta_1}} P_1 \times \int_{\overline{X}_{T_2, \delta_2}} P_2 .$$

We have therefore replaced the original integral I with a sum of products of integrals over polytopes of the same kind, but of smaller dimension. This gives an inductive procedure for computing I which exploits the combinatorial structure of $\overline{X}_{S,\delta}$.

The problem is that when we take the primitive of an algebraic form, we immediately have to introduce logarithms and other multivalued functions. Therefore, in order to make the argument work, we need an algebra of functions on $\mathfrak{M}_{0,S}$ which is sufficiently large that we can always take primitives inside it, but is sufficiently small that the numbers we obtain at the end are multiple zeta values and nothing more.

The point here is that the inductive process preserves the underlying geometry of the moduli spaces, in keeping with the general philosophy of periods [19], and this enables us to control the periods obtained at the end.

4. Generalised multiple polylogarithms

The class of functions we must consider in order to be able to take primitives are given by the multiple polylogarithms. We begin by recalling the definition of the classical polylogarithms, which were first defined by Leibniz. For any integer $n \geq 1$, define

$$\text{Li}_n(z) = \sum_{k=1}^{\infty} \frac{z^k}{k^n}, \quad \text{for } |z| < 1.$$

In particular, $\text{Li}_1(z) = -\log(1-z)$, and

$$\frac{d}{dz} \text{Li}_n(z) = \frac{\text{Li}_{n-1}(z)}{z},$$

for all $n \geq 2$. Using this, one shows that the functions $\text{Li}_n(z)$ have an analytic continuation to $\mathbb{C} \setminus \{0, 1\}$, which we identify with $\mathfrak{M}_{0,4}(\mathbb{C})$. These functions are multivalued in the sense that they define holomorphic functions on a universal covering space of $\mathfrak{M}_{0,4}(\mathbb{C})$.

There is a common generalisation of both the multiple zeta values and the classical polylogarithms which is due to Goncharov [13, 14]. For any $n_1, \dots, n_r \in \mathbb{N}$, the *multiple polylogarithm* is defined by:

$$(7) \quad \text{Li}_{n_1, \dots, n_r}(x_1, \dots, x_r) = \sum_{0 < k_1 < \dots < k_r} \frac{x_1^{k_1} \dots x_r^{k_r}}{k_1^{n_1} \dots k_r^{n_r}}.$$

The sum converges absolutely on the polydisk $|x_i| < 1$. By analytic continuation, these define multi-valued functions on the space

$$\{(x_1, \dots, x_r) \in \mathbb{C}^r : x_i \neq 0, 1, x_i x_{i+1} \dots x_j \neq 1\},$$

which, we recall, is just the moduli space $\mathfrak{M}_{0,r}(\mathbb{C})$ in cubical coordinates.

Since the functions (7) are multi-valued, we must be careful in fixing their branches, which means restricting to a single connected cell X in $\mathfrak{M}_{0,r}(\mathbb{R})$. Furthermore, the functions (7) have logarithmic singularities along each of the boundary hypersurfaces $x_i = 0, x_i = 1, x_i x_{i+1} \dots x_j = 1$, which do not cross normally. If we want to understand the limiting behaviour of these functions as we tend towards the boundary, we are once again forced to blow up the boundary components of the cell X . This leads us to consider the extension of these functions to the space $\mathfrak{M}_{0,r}^\delta(\mathbb{C})$ defined previously.

We can now define an algebra $L(\mathfrak{M}_{0,S}^\delta)$ of multivalued functions on $\mathfrak{M}_{0,S}(\mathbb{C})$ by taking all the multiple polylogarithms (7) and adding in some similar functions given explicitly in [6]. More precisely, we must consider all functions with unipotent monodromy, and these are given by Chen's theory of iterated integrals [9]. In particular, the algebra $L(\mathfrak{M}_{0,S}^\delta)$ contains all the dihedral coordinates u_{ij} , logarithms of dihedral coordinates $\log(u_{ij})$, polylogarithms $\text{Li}_n(u_{ij})$, and so on. For example, on $\mathfrak{M}_{0,5}$, the ring of regular functions is given by:

$$\mathbb{Q}[x_1, x_2, \frac{1}{x_1}, \frac{1}{x_2}, \frac{1}{1-x_1}, \frac{1}{1-x_2}, \frac{1}{1-x_1 x_2}]$$

and

$$\log(1-x_1 x_2), \log^2(x_1), \text{Li}_2(x_1 x_2), \text{Li}_2\left(\frac{1-x_1}{1-x_1 x_2}\right), \text{Li}_{1,1}(x_1, x_2), \dots$$

are all functions which lie in $L(\mathfrak{M}_{0,5}^\delta)$. Then $L(\mathfrak{M}_{0,S}^\delta)$ is a differential algebra defined over \mathbb{Q} , and has a natural grading. In cubical coordinates, the differentials are $\partial/\partial x_i$, and the weight of $\text{Li}_{n_1, \dots, n_r}$ is defined to be $n_1 + \dots + n_r$. In fact, $L(\mathfrak{M}_{0,S}^\delta)$ is a universal unipotent Picard-Vessiot extension of the ring of regular functions on $\mathfrak{M}_{0,S}$, and one can write down a basis for it explicitly.

The following theorem is the main technical tool for computing periods on the moduli spaces $\mathfrak{M}_{0,n}$ [6].

- THEOREM 6.**
- (1) *The weight 0 piece of $L(\mathfrak{M}_{0,S}^\delta)$ is the ring of regular functions on $\mathfrak{M}_{0,S}$.*
 - (2) *(Cohomological triviality). Every closed k -form on $\mathfrak{M}_{0,S}$ with coefficients in $L(\mathfrak{M}_{0,S}^\delta)$ has a primitive. This raises the weight by ≤ 1 .*
 - (3) *(Regularisation of multiple polylogarithms). Let $f \in L(\mathfrak{M}_{0,S}^\delta)$. We take the regularised value of f along D_{ij} , where $D_{ij} \in \mathfrak{M}_{0,S}^\delta \setminus \mathfrak{M}_{0,S}$. It is a product*

$$D_{ij} \cong \mathfrak{M}_{0,S_1 \cup \{e\}}^{\delta_1} \times \mathfrak{M}_{0,S_2 \cup \{e\}}^{\delta_2}$$

Then $\text{Reg}(f, D_{ij})$ is a sum of products of functions $f_1 f_2$, where

$$f_1 \in L(\mathfrak{M}_{0,S_1 \cup \{e\}}^{\delta_1}), \quad f_2 \in L(\mathfrak{M}_{0,S_2 \cup \{e\}}^{\delta_2}),$$

and elements of the algebra \mathcal{Z} of multiple zeta values.

Some explanation is required. Part (1) holds by definition, and says that the functions of weight 0 are just polynomials in the dihedral coordinates u_{ij} and their inverses. In order to explain part (2), we define a *polylogarithmic k -form* to be a sum of terms $f\omega$, where $f \in L(\mathfrak{M}_{0,S}^\delta)$ and where ω is a regular algebraic k -form on $\mathfrak{M}_{0,S}$. For example,

$$\phi = \text{Li}_2(x_1 x_2) dx_2 + \log(1 - x_1 x_2) dx_1$$

is a polylogarithmic 1-form on $\mathfrak{M}_{0,5}$ of weight 2. The theorem states that for any such form ϕ satisfying $d\phi = 0$, there is a polylogarithmic form P satisfying $dP = \phi$.

Part (3) tells us what happens when we take the limit of a generalised polylogarithm function f on $\mathfrak{M}_{0,n}$. Suppose that we have fixed a branch of f on a given cell $X_{S,\delta}$. When we take the limit of $f(x)$ when x tends towards a face F of the Stasheff polytope $\overline{X}_{S,\delta}$, the function f will diverge logarithmically. The regularised value of f is what remains of f after removing logarithmic divergences in the dihedral coordinates $\log^k(u_{ij})$. Recall that the face F is itself a product of moduli spaces $\mathfrak{M}_{0,r_1} \times \mathfrak{M}_{0,r_2}$. The theorem states that when we take the regularised limit of a generalised polylogarithm function f , we obtain multiple zeta values, and generalised polylogarithm functions on each smaller space \mathfrak{M}_{0,r_i} . In particular, the algebra $L(\mathfrak{M}_{0,S}^\delta) \otimes_{\mathbb{Q}} \mathcal{Z}$ is closed under taking regularised limits.

In conclusion, theorem 6 says that the analytic properties of multiple polylogarithms are determined by the combinatorics of Stasheff polytopes.

4.1. The iterated Stokes' formula once more. We can now return to the method of calculation of the integrals I . Recall that we wanted to integrate an algebraic function f :

$$I = \int_{\overline{X}_{S,\delta}} f .$$

By property (1) of the previous theorem, f is in $L(\mathfrak{M}_{0,S}^\delta)$ and is of weight 0. Next, by property (2), we can take primitives in the algebra $L(\mathfrak{M}_{0,S}^\delta)$, so

$$I = \int_{\overline{X}_{S,\delta}} f = \int_{\partial \overline{X}_{S,\delta}} P$$

for some primitive P of f . By property (3), when we restrict P to a face of the Stasheff polytope F_{ij} , we get multiple zeta values appearing, but we stay in the same class of polylogarithms. Therefore I can be written as a sum of terms of the form

$$\zeta \int_{\overline{X}_{T_1,\delta_1}} g_1 \int_{\overline{X}_{T_2,\delta_2}} g_2 ,$$

where $\zeta \in \mathcal{Z}$ is a multiple zeta value. Thus the induction goes through, and gives rise to a cascade of integrals. At each stage, the dimension of the domain of integration goes down by 1, and the weight of the integrand goes up by 1. At the final stage we end up with functions in $L(\mathfrak{M}_{0,3}^\delta) \otimes \mathcal{Z}$ which is just \mathcal{Z} , since $\mathfrak{M}_{0,3}$ reduces to a point. This shows that I is a linear combination of multiple zeta values, of weight at most ℓ .

This procedure gives an explicit algorithm for computing the integrals I . The process is in fact completely algebraic, and can be treated symbolically on a computer.

One has to be slightly careful with singularities of the integrands. The rule is that the functions we integrate *never* have poles along $\overline{X}_{S,\delta}$, but they *are* allowed to have logarithmic singularities along the boundary $\partial \overline{X}_{S,\delta}$.

5. Periods of Feynman motives

We now turn to a class of period integrals coming from massless Feynman diagrams in perturbative quantum field theory. The motivic point of view was initiated by Bloch, Esnault and Kreimer in [3]. After briefly reviewing their construction, we explain how our method of polylogarithmic integration for the moduli spaces $\mathfrak{M}_{0,n}$ can be expected to work in a similar manner for certain classes of Feynman diagrams [7].

5.1. Graph polynomials and periods. We begin by recalling the definition of graph polynomials. Let Γ be a Feynman graph, with N internal edges. In what follows, the external edges will play no role.

The graph polynomial ψ_Γ of Γ is a polynomial in variables $\alpha_1, \dots, \alpha_N$, with one variable corresponding to each edge of Γ . It is defined by the formula

$$\psi_\Gamma(\alpha_1, \dots, \alpha_N) = \sum_T \prod_{e \notin T} \alpha_e ,$$

where the sum is over all spanning trees of Γ (subgraphs of Γ which pass through every vertex, but do not contain a loop), and the product is over all edge variables which are not in T . The graph polynomial ψ_Γ is homogeneous of degree equal to the loop number of Γ .

EXAMPLE 7. Consider the three diagrams pictured below. The coffee bean diagram (left) has two independent loops and three edges, and has three spanning trees: $\{1\}$, $\{2\}$, and $\{3\}$. Therefore its graph polynomial is:

$$\psi_\Gamma = \alpha_1 \alpha_2 + \alpha_1 \alpha_3 + \alpha_2 \alpha_3 .$$

Similarly, the parachute in the middle has 4 edges and two loops, and its set of spanning trees is $\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 4\}, \{2, 4\}$, which implies that:

$$\psi_\Gamma = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) + \alpha_3\alpha_4 .$$

Finally, one can check that the wheel with three spokes diagram pictured on the right has exactly sixteen spanning trees, and so one has

$$\begin{aligned} \psi_\Gamma = & \alpha_1\alpha_2\alpha_6 + \alpha_1\alpha_4\alpha_6 + \alpha_2\alpha_5\alpha_6 + \alpha_4\alpha_5\alpha_6 + \alpha_1\alpha_3\alpha_6 + \alpha_2\alpha_3\alpha_6 \\ & + \alpha_3\alpha_4\alpha_6 + \alpha_3\alpha_5\alpha_6 + \alpha_1\alpha_3\alpha_4 + \alpha_1\alpha_3\alpha_5 + \alpha_2\alpha_3\alpha_4 \\ & + \alpha_2\alpha_3\alpha_5 + \alpha_2\alpha_4\alpha_5 + \alpha_1\alpha_4\alpha_5 + \alpha_1\alpha_2\alpha_5 + \alpha_1\alpha_2\alpha_4 . \end{aligned}$$

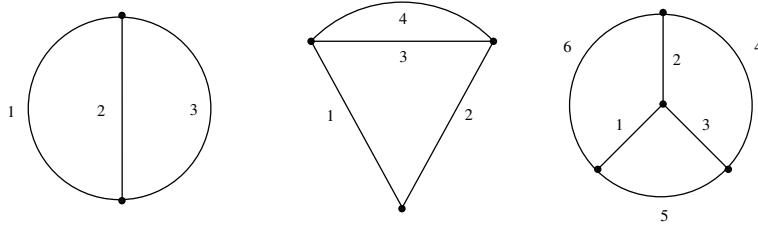


FIGURE 7. The coffee bean, or sunset diagram (left), the parachute (middle) and wheel with three spokes (right). The wheel with three spokes is primitive divergent, the other two are not, since they both contain divergent subgraphs.

Since the graph polynomial is homogeneous, it defines a certain hypersurface (the graph hypersurface) in projective space \mathbb{P}^{N-1} , which is denoted

$$X_\Gamma = \{\psi_\Gamma = 0\} .$$

In order to write down the residue of Γ , we make the assumption that Γ is primitive divergent, following the paper of Bloch, Esnault and Kreimer [3]. Let h_Γ denote the number of independent loops in Γ , and let N_Γ denote the number of edges. Then Γ is called primitive divergent if

$$N_\Gamma = 2h_\Gamma ,$$

and furthermore, Γ contains no divergent subgraphs: *i.e.*, for every strict subgraph $\gamma \subsetneq \Gamma$, we have $N_\gamma > 2h_\gamma$. In the examples above, the wheel with three spokes is primitive divergent (it is in fact the smallest non-trivial example of a primitive divergent graph), but the other two are not.

In the case when Γ is primitive divergent, we can cut a very long story short, and simply define the residue of Γ by the following projective integral:

$$(8) \quad \int_{\Delta_{2h-1}} \frac{\Omega_{2h-1}}{\psi_\Gamma^2} ,$$

where ψ_Γ is the graph polynomial of Γ , $\Delta_{2h-1} \subset \mathbb{P}^{2h-1}(\mathbb{R})$ is the standard simplex given by $\{(\alpha_1 : \dots : \alpha_N) : \alpha_i \geq 0\}$ in real projective space, and Ω_{2h-1} is given by

$$\Omega_{2h-1} = \sum_{i=1}^N (-1)^i \alpha_i d\alpha_1 \dots \widehat{d\alpha_i} \dots d\alpha_N .$$

The integrand of (8) is homogeneous, since the degree of ψ_Γ is h_Γ , and Γ is primitive divergent. One wants to realize this integral as a period of a mixed Hodge structure or motive. The integrand defines a cohomology class

$$\left[\frac{\Omega_{2h-1}}{\psi_\Gamma^2} \right] \in H^{2h-1}(\mathbb{P}^{2h-1} \setminus X_\Gamma),$$

and the domain of integration Δ defines a relative homology cycle in the group $H_{2h-1}(\mathbb{P}^{2h-1}, B)$, where $B = \cup_{i=1}^N \{\alpha_i = 0\}$. Just as in the case of moduli spaces, the singular locus of the integrand (the graph hypersurface) meets the boundary of the domain of integration, so one must blow up the intersections. This is done explicitly in [3]. Using the positivity of the coefficients in ψ_Γ , it turns out that the bad locus is a union of linear spaces, indexed by the set of subgraphs of Γ containing a loop:

$$X_\Gamma \cap B = \bigcup_{\gamma} \{\alpha_{e_1} = \dots = \alpha_{e_n} = 0 : \gamma = (e_1, \dots, e_n) \subsetneq \Gamma \text{ such that } h_\gamma > 0\}.$$

After blowing up the bad loci in order of decreasing codimension, it is proved in [3] that the strict transform \tilde{X}_Γ of X_Γ and the total transform \tilde{B} of B cross normally, and that the pull-back of the integrand of (8) is continuous on the strict transform $\tilde{\Delta}$ of Δ . In conclusion, one realizes the integral (8) as a period of the mixed Hodge structure:

$$(9) \quad \text{Mot}(\Gamma) = H^{2n-1}(\tilde{\mathbb{P}}^{2n-1} \setminus (\tilde{X}_\Gamma, \tilde{B} \setminus (\tilde{B} \cap \tilde{X}_\Gamma))).$$

We will be interested in the strict transform $\tilde{\Delta}$ of the domain of integration, which we call the Feynman polytope. It can be realized combinatorially by truncating the simplex Δ along faces indexed by certain subgraphs of Γ . The combinatorics of this polytope come into play in the relative cohomology group (9), and share many similarities with the Stasheff polytopes considered in §3.

EXAMPLE 8. Consider the coffee bean diagram, and its graph polynomial $\psi_\Gamma = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3$. Since all coefficients are positive, it meets the simplex $\Delta_3 = \{\alpha_1, \alpha_2, \alpha_3 \geq 0\}$ along the 3 faces:

$$\alpha_1 = \alpha_2 = 0, \quad \alpha_1 = \alpha_3 = 0, \quad \alpha_2 = \alpha_3 = 0,$$

which correspond to the subgraphs $\{1, 2\}$, $\{1, 3\}$ and $\{2, 3\}$, which all have one loop. In the case of the parachute, one must blow up the unit simplex $\Delta_4 = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4 \geq 0\}$ along the linear spaces:

$$\alpha_3 = \alpha_4 = 0, \quad \alpha_1 = \alpha_2 = \alpha_3 = 0, \quad \alpha_1 = \alpha_2 = \alpha_4 = 0,$$

which correspond to the subgraphs $\{3, 4\}$, $\{1, 2, 3\}$ and $\{1, 2, 4\}$. Note that the graph hypersurface also meets Δ_4 along the two faces $\alpha_1 = \alpha_3 = \alpha_4 = 0$ and $\alpha_2 = \alpha_3 = \alpha_4 = 0$, but these are contained in $\alpha_3 = \alpha_4 = 0$ and will already be taken care of after blowing up along $\alpha_3 = \alpha_4 = 0$.

In the previous example, we see that one only needs to blow up along linear subspaces corresponding to minimal subgraphs [3]. A subgraph $\gamma \subset \Gamma$ is said to be minimal if removing any edge of γ causes its number of loops to decrease, in other words,

$$h_{\gamma \setminus \{e\}} < h_\gamma \text{ for all edges } e \text{ of } \gamma.$$

Equivalently, γ is minimal if no edge of γ is contained in every spanning tree of γ .

5.2. Feynman polytopes. To any graph Γ we can associate a convex polytope $P(\Gamma)$, as follows. We begin with the unit simplex Δ_{N_γ} and label its facets with the edges of Γ . For each subgraph $\gamma \subset \Gamma$, let the face obtained by setting the edge variables of γ to 0 be denoted F_γ . Set

$$\tilde{F} = \{F_\gamma : \gamma \subsetneq \Gamma \text{ minimal, such that } h_\gamma = 1\}.$$

Then one must take the intersections of all faces in \tilde{F} , and truncate them, in increasing order of dimension. Since $F_{\gamma_1} \cap F_{\gamma_2} = F_{\gamma_1 \cup \gamma_2}$, and since the union of two minimal subgraphs is minimal, it suffices to truncate all minimal subgraphs γ with $h_\gamma > 0$. At the end we obtain a polytope $P(\Gamma)$ with the following properties (where, by abuse of notation, we denote the facets in the truncated polytope by F_γ also).

- (1) The set of facets $F(\Gamma)$ of $P(\Gamma)$ are indexed by the set

$$E(\Gamma) \cup \{\gamma \subsetneq \Gamma \text{ minimal such that } h_1(\gamma) > 0\}.$$

In other words, every facet is indexed either by an edge of Γ , or by a strict minimal subgraph of Γ which contains a loop.

- (2) (The forest formula). Two facets $F_{\gamma_1}, F_{\gamma_2} \in F(\Gamma)$ meet each other if

$$\gamma_1 \subset \gamma_2 \quad \text{or} \quad \gamma_2 \subset \gamma_1,$$

or if

$$\gamma_1 \cap \gamma_2 = \emptyset$$

and $F_{\gamma_1 \cup \gamma_2}$ is not already in $F(\Gamma)$.

- (3) (Product structure). Each facet F_γ is a product of Feynman polytopes (where $\Gamma/\!/\gamma$ denotes the graph obtained from Γ by contracting the edges in γ):

$$F_\gamma = P(\Gamma/\!/\gamma) \times P(\gamma).$$

If Γ has N edges, then its Feynman polytope $P(\Gamma)$ is of dimension $N - 1$. In particular, the Feynman polytope of a single edge or tadpole is a point. The product structure (3) was proved directly in [3] using properties of graph polynomials. It is an exercise to verify that it follows combinatorially from (2). The forest property (2) determines the full lattice of faces of the polytope $P(\Gamma)$ by induction.

EXAMPLES 9. Consider the example of the coffee bean graph. The domain Δ_3 is the two-dimensional simplex with faces labelled 1, 2, 3, and $P(\Gamma)$ is obtained by truncating all three faces 12, 13 and 23, because each corresponding subgraph contains a loop. We obtain a hexagon, pictured below. Note that the facet labelled $\gamma = 1$, for example, should be thought of as being isomorphic to $P(\gamma) \times P(\Gamma/\!/\gamma) \cong \{pt\} \times \{\text{interval}\}$.

In order to see the product structure more clearly on facets, consider the example of the parachute. Its polytope has seven facets and is obtained by truncating a tetrahedron with faces 1, 2, 3, 4 along the two corners 123, 124, and then along the edge 34, which correspond to the three minimal subgraphs which contain loops. The facet labelled 34 is a product $P(\gamma) \times P(\Gamma/\!/\gamma)$ of two intervals, where γ is the one-loop subgraph 34. The two facets labelled 1, 2 are hexagons, since the quotient graphs $\Gamma/\!/\{1\}$ and $\Gamma/\!/\{2\}$ are coffee bean graphs, whose corresponding polytopes are hexagons, by the above.

It is curious to note that no Feynman polytope can ever have pentagonal faces, as a simple calculation of the polytopes of all 3-edge graphs shows.

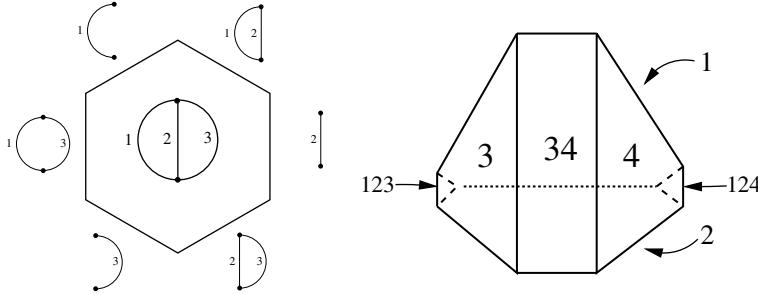


FIGURE 8. Left: the Feynman polytope of the coffee bean graph is a hexagon, obtained by truncating three corners of a triangle. The Feynman polytope of the parachute is obtained by truncating a tetrahedron at two corners and one edge, and has seven facets.

5.3. Calculation of Feynman amplitudes. We can apply the method of polylogarithmic integration on the polytopes $P(\Gamma)$ to compute all Feynman periods of a given primitive-divergent graph Γ . To give a rough idea of what this entails, consider a graph polynomial ψ , and let $\alpha_1, \dots, \alpha_N$ denote the coordinates corresponding to the edges. As a first approximation, we can naively integrate (8) with respect to the first two variables, and ignore questions of convergence. Since ψ is linear in each variable, we can write $\psi = \psi^1 \alpha_1 + \psi_1$, which gives

$$\int_0^\infty \int_0^\infty \frac{d\alpha_1 d\alpha_2}{\psi^2} = \int_0^\infty \int_0^\infty \frac{d\alpha_1 d\alpha_2}{(\psi^1 \alpha_1 + \psi_1)^2} = \int_0^\infty \frac{d\alpha_2}{\psi^1 \psi_1} .$$

Here we have replaced (8) with an affine integral. One can take the domain of integration to be the hyperplane $\alpha_N = 1$, for example. One can show that ψ^1 and ψ_1 can be interpreted as the graph polynomials obtained from Γ by removing or contracting the edge 1, and are therefore linear in each variable:

$$(10) \quad \psi^1 = \psi_{\Gamma \setminus 1} \quad \text{and} \quad \psi_1 = \psi_{\Gamma // 1} .$$

In a similar manner, write $\psi^1 = \psi^{12} \alpha_2 + \psi_2^1$, $\psi_1 = \psi_1^2 \alpha_2 + \psi_{12}$. By substituting and then decomposing into partial fractions, we obtain

$$\begin{aligned} \int_0^\infty \frac{d\alpha_2}{\psi^1 \psi_1} &= \int_0^\infty \frac{d\alpha_2}{(\psi^{12} \alpha_2 + \psi_2^1)(\psi_1^2 \alpha_2 + \psi_{12})} \\ &= \int_0^\infty \frac{d\alpha_2}{(\psi^{12} \psi_{12} - \psi_2^1 \psi_1^2)} \left[\frac{\psi^{12}}{\psi^{12} \alpha_2 + \psi_2^1} - \frac{\psi_1^2}{\psi_1^2 \alpha_2 + \psi_{12}} \right] \\ &= \frac{1}{(\psi^{12} \psi_{12} - \psi_2^1 \psi_1^2)} (\log \psi^{12} + \log \psi_{12} - \log \psi_2^1 - \log \psi_1^2) . \end{aligned}$$

The whole approach would be hopeless were it not for the miraculous fact that the determinant $\psi^{12} \psi_{12} - \psi_2^1 \psi_1^2$ is the square of certain polynomial D , which is itself linear in every variable $\alpha_3, \dots, \alpha_N$. This follows from a result due to Charles Dodgson on determinants of matrices. Therefore, at the next stage of the integration process, we can decompose the integrand into partial fractions with respect to α_3 , and integrate once more. A typical term is of the form:

$$\int_0^\infty \frac{\log(\psi_2^{13} \alpha_3 + \psi_{23}^1)}{(D^3 \alpha_3 + D_3)^2} d\alpha_3 .$$

Continuing in this way, we see that the singularities of the integrand leads us to consider a sequence of sets of polynomials

$$(11) \quad \{\psi\}, \{\psi^1, \psi_1\}, \{\psi^{12}, \psi_{12}, \psi_2^1, \psi_1^2, D\}, \dots$$

where $D^2 = (\psi^{12}\psi_{12} - \psi_2^1\psi_1^2)$. If, by eliminating variables in this manner, we obtain functions which are all linear with respect to some variable α_i at each stage, then the integral can be calculated using certain variants of polylogarithms which are known as hyperlogarithms [6, 13]. A typical function is of the form

$$\text{Li}_{n_1, \dots, n_r}(f_1, \dots, f_r)$$

where the f_i are rational functions in the Schwinger parameters α_i , whose numerators and denominators are products of elements in the sequence (11). Thus, whenever this algorithm terminates, it gives an analytic procedure for computing the corresponding integral in terms of such functions.

EXAMPLE 10. In the case of the wheel with 3 spokes, this integration can be carried out completely. Let ψ be its graph polynomial which has 16 terms, as given in example 7. Then one verifies that:

$$\begin{aligned} \psi^{12} &= \alpha_4 + \alpha_5 + \alpha_6 . \\ \psi_2^1 &= \alpha_3\alpha_6 + \alpha_3\alpha_4 + \alpha_3\alpha_5 + \alpha_4\alpha_5 + \alpha_4\alpha_6 . \\ \psi_1^2 &= \alpha_3\alpha_6 + \alpha_3\alpha_4 + \alpha_3\alpha_5 + \alpha_4\alpha_5 + \alpha_5\alpha_6 . \\ \psi_{12} &= \alpha_6(\alpha_3\alpha_4 + \alpha_3\alpha_5 + \alpha_4\alpha_5) . \\ D &= (\psi^{12}\psi_{12} - \psi_2^1\psi_1^2)^{1/2} = \alpha_3\alpha_4 + \alpha_3\alpha_5 + \alpha_3\alpha_6 + \alpha_4\alpha_5 . \end{aligned}$$

Since every polynomial is linear, we can do a further integration with respect to α_3 . Calculating all terms such as ψ_2^{13} , ψ_{123} , and all the corresponding Dodgson determinants such as $\psi_{12}^3 D_3 - \psi_{123} D^3$, we verify at the next stage of integration that we only obtain non-trivial terms of the form:

$$\alpha_5 + \alpha_6, \alpha_4 + \alpha_6, \alpha_4 + \alpha_5 + \alpha_6, \alpha_4 + \alpha_5, \alpha_4 - \alpha_5 .$$

These are linear in α_4 , and a further reduction with respect to α_4 gives:

$$\alpha_5 + \alpha_6, \alpha_5 - \alpha_6, 2\alpha_5 + \alpha_6 .$$

Since every polynomial is linear, we can always proceed to the next stage, and integrate our original Feynman period using polylogarithms. One can prove, by this very method, that

$$\int_{\Delta_5} \frac{\Omega_5}{\psi_{\Gamma}^2} = 6\zeta(3) .$$

The full details of this computation will be given in [7].

This process of reduction of polynomials only depends on the topology of the graph Γ , and enables us to check in advance whether a Feynman period will be computable in terms of polylogarithms or not. A similar reduction algorithm was considered by Stembridge [22] to study points of X_{Γ} over finite fields, and is taken up in a motivic setting in [3] to find a Tate filtration on the cohomology of X_{Γ} in some cases. Note that here we are interested in periods of the full relative cohomology group (9), and not just the cohomology of the complement of the graph hypersurface.

EXAMPLE 11. It is easy to check that the reduction always terminates in the case of the ring of functions on the moduli space $\mathfrak{M}_{0,n}$, because the map $\mathfrak{M}_{0,n+1} \rightarrow \mathfrak{M}_{0,n}$ which forgets a marked point is a fibration. This is the basic reason behind the cohomological triviality of theorem 6.

The algorithm described above will also work for certain convergent Feynman integrals where the propagators are raised to integer multiples of the dimensional regularization parameter (yielding the coefficients in its Taylor expansion), since this introduces logarithms into the integrand, and will also work in some massive cases too. In general, the method cannot be expected to work for all graphs. It is an interesting and important question to try to understand for which families of graphs it will.

6. Further analogies between the moduli and Feynman cases

In this final section, we explore some common aspects of moduli spaces and Feynman integrals which are of a more speculative nature, and would merit further study. First, we show that the Stokes' formula argument naturally leads to the definition of a cobracket on Feynman graphs. Next, we compare some large symmetry groups which have been discovered on both the number-theoretic and quantum field theory sides. Then, in §6.3, we introduce a simplified variant of our polylogarithmic integration process, which yields an algorithm for decomposing a period integral into master integrals. The master integrals for $\mathfrak{M}_{0,n}$ are called cellular zeta values and satisfy an enormous number of algebraic identities [8]. Finally, in §6.4, we discuss the possibility of using the Feynman motive of the complete graph as a universal motive for all Feynman graphs. This gives a framework to study relations between Feynman integrals, and their cancellation properties.

6.1. The boundary map and cobrackets for Feynman graphs. We show how the construction of the Feynman polytope naturally leads to a definition of a cobracket on the free \mathbb{Z} -module spanned by all Feynman graphs. This, and more general families of coproducts have been investigated in the recent work of S. Bloch and coworkers [2]. The main step in the iterated integration algorithm is the application of Stokes' theorem. If the domain of integration is a closed oriented polytope P , then it gets replaced with its boundary, which is the union of its facets:

$$P \mapsto \partial(P) = \sum_F F ,$$

each of which inherits an orientation. If each facet is a product of polytopes of the same kind, then we can write this

$$\partial : P \mapsto \sum P_i \times P_j .$$

It is clear that applying the boundary map twice gives zero:

$$\partial^2 = 0 ,$$

since facets meeting along a face of codimension 2 induce opposite orientations along that face. There is also a Leibniz-type rule on products:

$$\partial(P_i \times P_j) = \partial P_i \times P_j + P_i \times \partial P_j .$$

If we apply this idea to the Feynman polytopes defined above, we obtain a Lie cobracket on the space of Feynman graphs. Therefore, let

$$S = \mathbb{Z}[\Gamma : \Gamma \text{ a labelled Feynman graph with no external edges}] ,$$

where a labelling of Γ is an ordering of its internal edges. Define a map

$$(12) \quad \begin{aligned} \delta : S &\longrightarrow S \wedge S , \\ \delta(\Gamma) &= \sum_{\gamma} \gamma \wedge \Gamma // \gamma , \end{aligned}$$

where the sum is over all minimal subgraphs $\gamma \subset \Gamma$ (*i.e.* graphs such that removing any edge of γ strictly decreases the number of loops h_γ), along with the subgraphs which consist of a single edge.

EXAMPLE 12. The cobracket Γ for the parachute can be written:

$$\begin{aligned} \delta(1234) &= 1 \wedge 234 + 2 \wedge 134 + 3 \wedge 124 + 4 \wedge 123 \\ &\quad + 34 \wedge 12 + 123 \wedge 4 + 124 \wedge 3 , \end{aligned}$$

where each term in the right-hand side is a subgraph of Γ (when it occurs on the left of the \wedge) or quotient graph of Γ (on the right of \wedge). Rather than write out every such graph, we have denoted only the corresponding edges for simplicity (in the first term $1 \wedge 234$, 1 stands for a single edge labelled 1, and 234 is a coffee bean graph with edges labelled 234). There is exactly one term corresponding to each facet of the Feynman polytope for Γ (fig. 8).

The map δ is a Lie cobracket: in other words, $\delta^2 = 0$.

EXAMPLE 13. If one applies the same idea to the case of the Stasheff polytopes (which occur as the connected components in the moduli space $\overline{\mathcal{M}}_{0,n}(\mathbb{R})$), one retrieves the well-known cobracket on the graded vector space spanned by labelled n -gons. If P is a regular n -gon, then

$$\delta(P) = \sum_{\chi(P)} P_1 \wedge P_2 ,$$

where the sum is over all $n(n-3)/2$ chords of P , and P_1, P_2 are the smaller polygons obtained by cutting along each chord, as in figure 4.

6.2. Symmetry groups and products. A further similarity between integrals on moduli spaces as they occur in number theory, and certain Feynman integrals, is the existence of non-trivial symmetry groups.

To illustrate this, let us return to our example of a period integral on $\mathcal{M}_{0,5}$, given for a set of non-negative integers h, i, j, k, l by:

$$I(h, i, j, k, l) = \int_0^1 \int_0^1 \frac{x^h(1-x)^i y^k(1-y)^j}{(1-xy)^{i+j-l}} \frac{dxdy}{1-xy} .$$

The form of the integral is preserved by the two birational automorphisms:

$$\sigma : (x, y) \mapsto (y, x) \quad \text{and} \quad \tau : (x, y) \mapsto \left(1-xy, \frac{1-y}{1-xy}\right) .$$

One can check by direct computation that the element τ is of order 5, preserves the open unit square $(0, 1) \times (0, 1)$, and that the elements σ, τ generate a dihedral group of order 10 (in the general case, the integral (5) on $\mathcal{M}_{0,n}$ has a natural action by a dihedral group of order $2n$). The truly surprising fact is that there is

a further symmetry, related to the transformations of the hypergeometric function, discovered by Dixon in 1905:

$$(13) \quad \frac{1}{j!k!} I(h, i, j, k, l) = \frac{1}{(k+l-i)!(i+j-l)!} I(h, i, k+l-i, i+j-l, l).$$

This transformation, together with σ and τ , generate a group which is isomorphic to the full symmetric group on five letters \mathfrak{S}_5 . Similarly, consider

$$I(h, i, j, k, l; r, s, t) = \int_{[0,1]^3} \frac{x^h(1-x)^i y^t(1-y)^j z^l(1-z)^k}{(1-xy)^{i+j-r}(1-yz)^{j+k-s}} \frac{dxdydz}{(1-xy)(1-yz)},$$

which is a special case of the more general period integral (5) for $\mathfrak{M}_{0,6}$. If we assume that $t = h+i+1 = l+k+1$, then this integral, suitably multiplied by gamma-factors, carries a highly non-trivial group action of order 1920. This family of integrals, after a change of variables, is the same family as that considered by Rhin and Viola [20], and one can prove that it evaluates to a rational linear combination of 1 and $\zeta(3)$. They exploited the symmetries of this family of integrals to obtain the best known rational approximations to $\zeta(3)$. The existence of such symmetry groups is of number-theoretic interest, because, roughly speaking, the richer the group of transformations, the smaller the estimate for the irrationality measure one obtains.

Now consider an example from massless ϕ^4 theory. Let G be the wheel with three spokes diagram, and Ψ_G its graph polynomial (example 7). The master two-loop two point function can be written in terms of the integral:

$$(14) \quad I(n_1, \dots, n_6, \varepsilon) = \int_{\Delta_5} \prod_{i=1}^6 \alpha_i^{n_i \varepsilon} \frac{\Omega_5}{\Psi_G^{2-\varepsilon}},$$

where n_i are integers satisfying $\sum_{i=1}^6 n_i = -3$. Using our polylogarithmic integration method [7], one can prove that the coefficients in the Taylor expansion of this integral in ε are multiple zeta values (this result was previously obtained by Bierenbaum and Weinzierl using Mellin-Barnes techniques [1]). The graph G is the complete graph on 4 vertices, and has a tetrahedral symmetry group \mathfrak{S}_4 which acts on this family of integrals by permuting the variables α_i . There is another obvious transformation which maps $\alpha_i \mapsto \alpha_i^{-1}$ for $1 \leq i \leq 6$ and which preserves Ψ_G . This follows from the fact that G is self-dual. The surprising fact is that there is a further symmetry, which comes from the star-triangle relation, such that all symmetries together generate a group of order 1440. As in the previous examples, this multiplies the bare integral $I(n_1, \dots, n_6, \varepsilon)$ by certain gamma-factors, and introduces auxiliary parameters which are certain linear combinations of the n_i . This is studied in detail in [20].

The existence of these remarkable symmetry groups suggests that integrals with many automorphisms play a special role in both number theory and physics, and should be related to an underlying theory of higher-dimensional hypergeometric functions.

6.2.1. Products. Another common theme is the existence of relations between products of integrals. On the one hand, multiple zeta values are known to satisfy a rich set of quadratic relations known as the double shuffle relations. These express any product of multiple zeta values as a linear combination of multiple zeta values. Similarly, on the physics side, it is well-known that the leading term of a 2-vertex reducible primitive Feynman diagram splits as the product of periods of smaller

Feynman diagrams. A geometric approach to understanding such relations was initiated in [6] and studied in detail in [8]. The idea is to consider a product of forgetful maps

$$(15) \quad f = f_1 \times f_2 : \mathfrak{M}_{0,r+s+3} \rightarrow \mathfrak{M}_{0,r+3} \times \mathfrak{M}_{0,s+3} .$$

The dimensions of the spaces on the left and right-hand sides are equal. Given a differential r and s form ω_1 and ω_2 on $\mathfrak{M}_{0,r+3}$ and $\mathfrak{M}_{0,s+3}$ respectively, the pull-back $f^*(\omega_1 \otimes \omega_2)$ is a differential $r+s$ form on $\mathfrak{M}_{0,r+s+3}$. This gives:

$$(16) \quad \int_{X_{r+3}} \omega_1 \int_{X_{s+3}} \omega_2 = \int_{f^{-1}(X_{r+s+3})} f^*(\omega_1 \otimes \omega_2) .$$

The domain $f^{-1}(X_{r+s+3})$ decomposes as a union of Stasheff polytopes, and the integrand $f^*(\omega_1 \otimes \omega_2)$ can be decomposed as a linear combination of elements in a suitable basis of differential forms with logarithmic singularities. The right-hand side is therefore a sum of period integrals of the form (5). In short, we obtain a quadratic relation which expresses a product of period integrals on moduli spaces as a sum of integrals in a higher-dimensional moduli space. This can be written explicitly [8].

A similar idea can be used to prove the hypergeometric transformation (13) geometrically. If one interprets the factorials as a period integral for $\mathfrak{M}_{0,4}$, then the left-hand side of (13) is essentially a product of period integrals over $X_4 \times X_5$ in $\mathfrak{M}_{0,4} \times \mathfrak{M}_{0,5}$. This is in turn isomorphic to a rectangular facet F of the Stasheff polytope X_6 for $\mathfrak{M}_{0,6}$ (see figure 6). Applying a dihedral rotation to X_6 sends this face F to a different rectangular facet $\sigma F \cong X'_4 \times X'_5$, and the integral breaks into a different product of integrals on $\mathfrak{M}_{0,4} \times \mathfrak{M}_{0,5}$, which gives the right-hand side of (13). The calculation is given in [6], §7.7.

It is natural to ask whether these geometric ideas for generating relations between periods can be extended to the case of Feynman integrals, and whether one can exploit the product structure on the faces of the Feynman polytopes to obtain identities between Feynman integrals.

6.3. Contiguity and IBP, master integrals and cohomology. A common practice in the evaluation of Feynman integrals with non-trivial numerators is to apply integration by parts identities to reduce to a finite set of master integrals [21]. However, this process can be rather complicated, and in some simple cases, a substantial shortcut may be afforded by using logarithmic de Rham forms.

6.3.1. The integration by parts method. To illustrate this on a simple example, consider again the period integral $I(h, i, j, k, l)$ on $\mathfrak{M}_{0,5}$ (6). Suppose we wish to calculate this integral for large positive integers h, i, j, k, l . One common approach used in physics would be to reduce this integral to a linear combination of integrals with smaller values of the parameters using contiguity relations.

For example, the identity $1 - (x)(y) = (1 - xy)$ implies that

$$I(h, i, j, k, l) - I(h+1, i, j, k+1, l) = I(h, i, j, k, l+1) .$$

Likewise, an integration by parts with respect to x gives

$$(h+1)I(h, i, j, k, l+1) = i I(h+1, i-1, j, k, l) - (i+j-l) I(h+1, i, j, k+1, l) ,$$

whenever $i > 0$. By writing down many such equations, and solving them, it is possible to express $I(h, i, j, k, l)$ as a linear combination of integrals with smaller

values of h, i, j, k, l . By induction, one can prove in this way that $I(h, i, j, k, l)$ is a rational combination of the two master integrals

$$I(0, 0, 0, 0, 0) = \int_{[0,1]^2} \frac{dxdy}{1-xy} \quad \text{and} \quad I(0, 0, 0, 0, 1) = \int_{[0,1]^2} dxdy .$$

This is carried out in [20]. In practice, this is cumbersome because it is not clear a priori whether we have a sufficient number of contiguity relations in order to be able to solve the problem (see the discussion in [21]).

6.3.2. *A cohomological method.* A different approach is to determine a basis for the relative cohomology group. The differential form

$$\omega = \frac{x^h(1-x)^i y^k(1-y)^l}{(1-xy)^{i+j-l}} \frac{dxdy}{1-xy} \in H^2(\mathfrak{M}_{0,5}^\delta, B) .$$

As mentioned in the introduction, a basis for the relative de Rham cohomology $H^2(\mathfrak{M}_{0,5}^\delta, B)$ is given by

$$\omega_1 = \frac{dxdy}{1-xy} \quad \text{and} \quad \omega_0 = dxdy .$$

This basis is computed using the de Rham complex of the simplicial complex

$$\mathfrak{M}_{0,5}^\delta \leftarrow \coprod_{i=1}^5 B_i \leftarrow \coprod_{1 \leq i < j \leq 5} B_i \cap B_j ,$$

where the $B_i \cong \mathfrak{M}_{0,4}^\delta = \mathbb{A}^1$ are faces of the shaded pentagon at finite distance in figure 1. The key point is that $\mathfrak{M}_{0,5}^\delta$ is affine, and therefore, by a theorem due to Grothendieck, the cohomology is spanned by the global regular forms. It immediately follows that:

$$(17) \quad [\omega] = a [\omega_1] + b [\omega_0] \in H^2(\mathfrak{M}_{0,5}^\delta, B) , \quad a, b \in \mathbb{Q} ,$$

and hence $I(h, i, j, k, l) = a I(0, 0, 0, 0, 0) + b I(0, 0, 0, 0, 1)$. The constants a, b can be computed as follows. By taking residues on both sides of (17), we deduce that

$$a = \text{Res}_{y=0}(\text{Res}_{x=1/y} \omega) \in \mathbb{Z} .$$

To compute b , observe that $\omega - a\omega_1$ is an exact form on $\mathfrak{M}_{0,5}^\delta$ (since $[\omega_1]$ generates $H^2(\mathfrak{M}_{0,5}^\delta) \cong \mathbb{Q}$), and hence has an algebraic primitive $\eta \in \Omega^1(\mathfrak{M}_{0,5}^\delta)$ such that $d\eta = \omega - a\omega_1$. We can now apply our method using Stokes' formula and iterated integration. The point is that the integration process uses only algebraic functions in this case, and no polylogarithms. At the end we obtain

$$b = \int_{[0,1]^2} d\eta \in \mathbb{Q} .$$

6.3.3. *A variant of the polylogarithmic integration algorithm.* The above can be reformulated as a general algorithm for reducing a period integral on $\mathfrak{M}_{0,n}^\delta$ in terms of ‘master’ integrals. Let ω be a differential form of degree ℓ on $\mathfrak{M}_{0,\ell+3}$. First, write down a basis $\omega_1, \dots, \omega_N$ for $H^\ell(\mathfrak{M}_{0,\ell+3})$. For instance, one can take the set of ℓ -forms in simplicial coordinates:

$$\frac{dt_1 \dots dt_\ell}{(t_1 - \varepsilon_1)(t_2 - \varepsilon_2) \dots (t_\ell - \varepsilon_\ell)} ,$$

where $\varepsilon_1 \in \{0, 1\}$, $\varepsilon_2 \in \{0, 1, t_1\}$, \dots , $\varepsilon_\ell \in \{0, 1, t_1, \dots, t_{\ell-1}\}$. We know that there exist rational numbers $a_i \in \mathbb{Q}$, such that

$$[\omega] = \sum_{i=1}^N a_i [\omega_i] .$$

The a_i can be found explicitly by computing residues on both sides. The difference $\omega - \sum_{i=1}^N a_i \omega_i$ is exact, and so we can find η , a regular $\ell - 1$ form on $\mathfrak{M}_{0,\ell+2}$, such that $d\eta = \omega - \sum_{i=1}^N a_i \omega_i$. It follows that

$$\int_{X_{\ell+3}} \omega = a_i \sum_{i=1}^N \int_{X_{\ell+3}} \omega_i + \int_{X_{\ell+3}} d\eta .$$

Applying Stokes' formula to the last term, and using the fact that the boundary $\partial X_{\ell+3}$ of the Stasheff polytope is a union of products of Stasheff polytopes of smaller dimension, we deduce that:

$$\int_{X_{\ell+3}} \omega = a_i \sum_{i=1}^N \int_{X_{\ell+3}} \omega_i + J .$$

where J is a sum of products of periods of moduli spaces of smaller dimension. We can proceed by induction on J . This reduces our initial integral to an explicit linear combination of products of master integrals of the form

$$\int_{X_{r+3}} \omega_i , \quad 1 \leq r \leq \ell ,$$

where the ω_i are elements of a basis for $H^r(\mathfrak{M}_{0,r+3})$. The master integrals can be computed once and for all in advance using the method of polylogarithmic integration detailed earlier.

Note that we have not worried about convergence. It is in fact possible to construct an explicit basis of convergent regular forms on $\mathfrak{M}_{0,n}$, or, equivalently, a basis for $H^\ell(\mathfrak{M}_{0,n}^\delta)$. This is somewhat more involved, and is carried out in [8]. The corresponding master integrals are called cellular zeta values, and satisfy many interesting algebraic relations using the trick (16).

In conclusion, a similar idea will work for certain classes of Feynman diagrams for which one can determine a basis for the relative cohomology group (9). In particular, those which have linear cohomology in the sense of [2]. The corresponding ‘master integrals’ can be interpreted as entries of the period matrix, and the algorithm described above gives an alternative reduction method to IBP methods, but which is now guaranteed to terminate.

6.4. The complete graph and the universal Feynman motive. Rather than have to deal with each Feynman diagram individually, it is natural to ask if every Feynman integral can be written as the period of a single geometric object. This is indeed possible, and suggests an approach to studying the apparent cancellation of periods when summing over a large number of Feynman graphs.

First observe that for $n \geq 1$,

$$(18) \quad \int_0^\infty \frac{a}{(ax+b)^{n+1}} dx = \frac{1}{n b^n} \quad \text{and} \quad \int_0^\infty \frac{b x^{n-1}}{(ax+b)^{n+1}} dx = \frac{1}{n a^n} .$$

Let Ψ_G be the graph polynomial of G , and α_e an edge parameter. Recall from (10) that $\Psi_G = \Psi_{G//e} + \alpha_e \Psi_{G\setminus e}$. Since Ψ_G is an affine function of α_e , the previous equation implies that:

$$(19) \quad \int_0^\infty \frac{\Psi_{G\setminus e}}{\Psi_G^{n+1}} d\alpha_e = \frac{1}{n \Psi_{G//e}^n} \quad \text{and} \quad \int_0^\infty \frac{\Psi_{G//e} \alpha_e^{n-1}}{\Psi_G^{n+1}} d\alpha_e = \frac{1}{n \Psi_{G\setminus e}^n}.$$

It follows by induction that the Feynman period of any subquotient graph of G is a Feynman period of G . We only consider subgraphs here. Therefore, fix a graph Γ , and suppose that γ is obtained from Γ by deleting the (ordered) set of edges e_1, \dots, e_k :

$$\gamma = \Gamma \setminus \{e_1, \dots, e_k\}.$$

Let σ denote the ordering on $\{e_1, \dots, e_k\}$. To γ, σ we associate the polynomial

$$N_{\gamma, \sigma} = \alpha_{e_1}^k \Psi_{\Gamma//e_1} \alpha_{e_2}^{k-1} \Psi_{\Gamma\setminus e_1//e_2} \alpha_{e_3}^{k-2} \Psi_{\Gamma\setminus\{e_1, e_2\}//e_3} \dots \alpha_{e_k} \Psi_{\Gamma\setminus\{e_1, e_2, \dots, e_{k-1}\}//e_k}.$$

It follows by induction using (18) that

$$\frac{1}{\Psi_\gamma^2} = k! \int_{[0, \infty]^k} \frac{N_{\gamma, \sigma}}{\Psi_\Gamma^{k+2}} d\alpha_{e_1} \dots d\alpha_{e_k}.$$

If we sum over all possible orderings of $\{e_1, \dots, e_k\}$:

$$N_\gamma = \sum_{\sigma \in S_n} N_{\gamma, \sigma},$$

then the same formula holds without the factor $k!$. Since every simple graph with at most n vertices is a subgraph of the complete graph K_n , we can express every Feynman period as a period of the complete graph.

$$(20) \quad \begin{aligned} \mathbb{Z}[\gamma \text{ with at most } n \text{ vertices}] &\longrightarrow \mathbb{Q}\left[\alpha_1, \dots, \alpha_r, \frac{1}{\Psi_{K_n}}\right], \\ \gamma &\mapsto \frac{N_\gamma}{\Psi_{K_n}^{r-e(\gamma)}}. \end{aligned}$$

where $r = \binom{n}{2}$ is the number of edges of K_n , and $e(\gamma)$ is the number of edges of γ . Note that each polynomial N_γ is not canonical, since it depends on the choice of embedding of γ into K_n . We can define the universal Feynman motive to be the ‘motive’ $\text{mot}(K_n)$ (9) of the complete graph K_n .

Now suppose we need to sum the Feynman periods over a large set S of graphs, with at most n vertices. Then it only suffices to calculate a single integral whose denominator is given by the complete graph polynomial, and whose numerator is some polynomial:

$$\sum_{\gamma \in S} \int \frac{1}{\psi_\gamma^2} = \int \frac{F}{\Psi_{K_n}^{\binom{n}{2}}}, \quad \text{where} \quad F = \sum_{\gamma \in S} N_\gamma \Psi_{K_n}^{e(\gamma)}.$$

The complete graph is therefore a natural place to study master integrals, the cancellation problem, and algebraic relations between Feynman graphs.

References

- [1] I. Bierenbaum, S. Weinzierl, *The massless two-loop two-point function*, Eur. Phys. J. C, Part. Fields 32, No.1, 67-78 (2003).
- [2] S. Bloch, *Motives associated to graphs*, Japan J. Math. 2 (2007), 165–196
- [3] S. Bloch, H. Esnault, D. Kreimer, *On motives associated to graph polynomials*, Comm. Math. Phys. 267 (2006), no. 1, 181-225.
- [4] D. Broadhurst, *Exploiting the 1,440-fold Symmetry of the Master Two-Loop Diagram*, Z. Phys. C - Particles and Fields 32, 249-253 (1986).
- [5] D. Broadhurst, D. Kreimer, *Knots and numbers in ϕ^4 theory to 7 loops and beyond*, Int. J. Mod. Phys. C 6, 519 (1995).
- [6] F. C. S. Brown, *Multiple zeta values and periods of moduli spaces $\mathfrak{M}_{0,n}$* , Ann. Scient. Éc. Norm. Sup., 4^e série, t. 42, p. 373-491.
- [7] F. C. S. Brown, *The massless higher-loop two point function*, Comm. in Math. Physics 287, Number 3, 925-958 (2009)
- [8] F. C. S. Brown, S. Carr, L. Schneps, *Cell zeta values*, arXiv:0910.0122v1, to appear in Compos. Math.
- [9] K. T. Chen, *Iterated path integrals*, Bull. Amer. Math. Soc. **83**, (1977), 831-879.
- [10] P. Deligne, D. Mumford, *The irreducibility of the space of curves of a given genus*, Publ. Math. IHES **36**, (1969), 75-109.
- [11] P. Deligne, A. B. Goncharov, *Groupes fondamentaux motiviques de Tate mixte*, Ann. Sci. Ecole Norm. Sup., Sér. IV **38**, No. 1, (2005), 1-56.
- [12] S. Devadoss, *Tesselations of moduli spaces and the mosaic operad*, in Contemp. Math. **239** (1999), 91-114.
- [13] A. B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, preprint (2001), arXiv:math.AG/0103059v4.
- [14] A. B. Goncharov, *Periods and mixed motives*, arXiv:math.AG/0202154 (2001).
- [15] A. B. Goncharov, Y. I. Manin, *Multiple ζ -motives and moduli spaces $\mathcal{M}_{0,n}$* , Compositio Math. **140** (2004), 1-14.
- [16] R. M. Hain, *Classical Polylogarithms*, Proc. Sympos. Pure Math. **55**, (1994), Part 2.
- [17] C. Itzykson, J.B. Zuber, *Quantum field theory*, McGraw-Hill (1980), xxii, 705 p.
- [18] F. F. Knudsen, *The projectivity of the moduli space of stable curves II. The stacks $\overline{\mathcal{M}}_{0,n}$* , Math. Scand. **52** (1983), 163-199.
- [19] M. Kontsevich, D. Zagier, *Periods*, dans Mathematics unlimited - 2001 and beyond, Ed. Engquist and Schmidt, pp. 771-808, Springer (2001).
- [20] G. Rhin, C. Viola, *The group structure for $\zeta(3)$* , Acta. Arith. **97**, no. 3 (2001), 269-293.
- [21] V. A. Smirnov, *Evaluating Feynman integrals*, Springer Tracts in Modern Physics 211. Berlin: Springer. ix, 247 p. (2004). [ISBN 3-540-23933-2/hbk]
- [22] J. Stembridge, *Counting points on varieties over finite fields related to a conjecture of Kontsevich*, Ann. Combin. 2 (1998) 365–385.
- [23] T. Terasoma, *Selberg integrals and multiple zeta values*, Compos. Math **133**, (2002), no. 1, 1-24.
- [24] M. Waldschmidt, *Valeurs zêta multiples : une introduction*, J. Théorie des Nombres de Bordeaux **12** (2002), 581-595.

From quantum electrodynamics to posets of planar binary trees

Frédéric Chapoton and Alessandra Frabetti

ABSTRACT. This paper is a brief mathematical excursion which starts from quantum electrodynamics and leads to the Möbius function of the Tamari lattice of planar binary trees, within the framework of groups of tree-expanded series.

First we recall Brouder's expansion of the photon and the electron Green's functions on planar binary trees, before and after the renormalization. Then we recall the structure of Connes and Kreimer's Hopf algebra of renormalization in the context of planar binary trees, and of their dual group of tree-expanded series. Finally we show that the Möbius function of the Tamari posets of planar binary trees gives rise to a particular series in this group.

Introduction

Planar binary trees are among the most classical objects in combinatorics. Being counted by the Catalan numbers, they are in bijection with more than one hundred other combinatorial objects, cf. [Sta99], such as noncrossing partitions and Dyck paths. In the last fifteen years, they have begun to play a key role in some algebraic structures, ranging from groups to Hopf algebras and to operads, cf. [Lod01, HNT05, AS06]. In this article, we would like to show how a problem in quantum electrodynamics can lead to such algebraic structures.

As shown by Christian Brouder in [Bro00], in quantum electrodynamics the Green functions can be described perturbatively with an expansion over planar binary trees. In the quantum field theory describing a scalar self-interacting field, the Green functions are naturally expanded over rooted trees (non-planar and non-binary). In contrast to this, in quantum electrodynamics Brouder adopted binary trees to describe the coupling of two interacting fields, the left path for the electron field and the right path for the photon field, and adopted planar trees to respect the non-commutativity of the product among the Green functions' coefficients, which are 4×4 matrices.

The expansion of the Green functions over planar binary trees requires some simple grafting operations among trees, [Bro00]. Instead, the renormalization of these Green functions gives rise to some Hopf algebras on planar binary trees, cf. [BF01], analogue to the famous Connes-Kreimer Hopf algebra encoding the renormalization of the scalar Φ^3 theory, cf. [Kre98, CK00a]. These Hopf algebras

can naturally be seen as coordinate rings of some pro-algebraic groups of formal series expanded over trees. One of them turns out to be a related to an operad on planar binary trees, cf. [Fra08].

There are in fact two operads on planar binary trees in the literature. In the present context we are concerned with the dupcial operad, cf. [Lod06]. The other one is the dendriform operad [Lod01], which has been the subject of a lot of attention recently. In particular, it has been shown that the dendriform operad is deeply related to the family of posets called Tamari lattices, cf. [HT72]. In this article, we show that Tamari lattices are also directly related to the dupcial operad. To do it, we compute the inverse of some series in the group of tree-expanded series associated with the dupcial operad.

The paper is organized as follows. In the first section we recall Brouder's expansion of the photon and the electron Green's functions on planar binary trees, before and after the renormalization. In the second section we review the Hopf algebras of renormalization for quantum electrodynamics in the context of planar binary trees, and their dual groups of tree-expanded series. In the last section we show that the Möbius function of the Tamari posets of planar binary trees gives rise to a particular series in the group of tree-expanded series which generalises the group of formal diffeomorphisms on a line.

Acknowledgements. This work was partially financed by the ANR-06-BLAN-0380 "HopfCombOp" and by the ANR-05-JCJC-0044-01 "AHBE".

1. Quantum electrodynamics and tree-expanded series

Quantum electrodynamics (QED) is the quantisation of classical electromagnetism, that is the field theory describing the attraction/repulsion between electrons, by means of the photons as mediators. In this section we review the perturbative expansion of the 2-point Green's functions of the electron and the photon on planar binary trees, as proposed by C. Brouder in [Bro00].

1.1. Dyson–Schwinger equations for the electron and the photon propagators. Denote by $\mathbb{R}^{1,3}$ the Minkowski space, with flat diagonal metric $g_{\mu\nu} = (1, -1, -1, -1)$, relativistic space-time coordinates $x = x^\mu \in \mathbb{R}^{1,3}$ and momenta $p = p_\mu$.

Let $\psi : \mathbb{R}^{1,3} \rightarrow \mathbb{C}^4$ be the fermionic field describing the electron, with mass m and electric charge e , and let $A = (A^\mu) : \mathbb{R}^{1,3} \rightarrow \mathbb{C}^4$ be the massless bosonic field describing the electromagnetic potential (the photon). The dynamic of the system of interacting electrons and photons is described by the Lagrangian density

$$\mathcal{L}_{QED}(\psi, A; e, m) = \mathcal{L}_{Dirac}(\psi; m) + \mathcal{L}_{Maxwell}^\xi(A^\mu) + \mathcal{L}_{int}(\psi, A^\mu; e)$$

where

$$\mathcal{L}_{Dirac}(\psi; m) = \bar{\psi}(x^\mu)(i\gamma^\mu \partial_\mu - m)\psi(x^\mu),$$

is the Lagrangian describing the free electrons (γ^μ are the Dirac 4x4 matrices and $\bar{\psi} = \psi^* \gamma^0$ is the anti-fermionic field, e.g. the positron),

$$\mathcal{L}_{Maxwell}^\xi(A) = -\frac{1}{4}(\partial^\mu A^\nu - \partial^\nu A^\mu)^2 - \frac{\xi}{2}(\partial_\nu A^\mu)^2,$$

is the Lagrangian describing the free photons (ξ is a gauge parameter introduced as a fictive mass to avoid the problems due to the absence of the photon mass), and

$$\mathcal{L}_{int}(\psi, A^\mu; e) = -e \bar{\psi} \gamma^\mu \psi A^\mu$$

is the term describing the interaction (of order 3), with coupling constant given by the electric charge e .

In this paper we are concerned with the connected 2-point Green functions

$$\begin{cases} S(x, y) = \langle \psi(x) \psi(y) \rangle, & \text{for the electron} \\ D_{\mu\nu}(x, y) = \langle A_\mu(x) A_\nu(y) \rangle, & \text{for the photon} \end{cases}$$

whose square module represent the probability that the quantum field moves from the point y to the point x in space-time. (We suppose that the 1-point Green's functions are zero.) If the fields are free, that is, the Lagrangian of the system is the sum $\mathcal{L}_{Dirac}(\psi; m) + \mathcal{L}_{Maxwell}^\xi(A^\mu)$ without the interaction part, the 2-point Green functions coincide with the free propagators. These are 4×4 matrix-valued distributions on $\mathbb{R}^{1,3}$, whose Fourier transforms on the momenta space are

$$\begin{cases} S_0(p) = \frac{1}{\gamma^\mu p_\mu - m + i\epsilon}, \\ D_{0,\mu\nu}(p) = -\frac{g_{\mu\nu}}{p^2 + i\epsilon} + \left(1 - \frac{1}{\xi}\right) \frac{p_\mu p_\nu}{(p^2 + i\epsilon)^2}. \end{cases}$$

If the fields interact according to the Lagrangian \mathcal{L}_{QED} , the 2-point Green functions (that we shall keep calling propagators) satisfy a system of two functional equations called *Dyson-Schwinger equations*, analogue to the equation of motion in the case of classical fields. For an isolated system, C. Brouder presented these two equations in the following form, cf. [Bro00]:

$$(1.1) \quad \begin{cases} S(p) = S_0(p) + i e^2 S_0(p) \int \frac{d^4 q}{(2\pi)^4} \gamma^\lambda D_{\lambda\lambda'}(q) \frac{\delta S(p-q)}{e \delta A_{\lambda'}(q)}, \\ D_{\mu\nu}(p) = D_{0,\mu\nu}(p) - i e^2 D_{0,\mu\lambda}(p) \int \frac{d^4 q}{(2\pi)^4} \text{tr} \left[\gamma^\lambda \frac{\delta S(q)}{e \delta A_{\lambda'}(-p)} \right] D_{\lambda'\nu}(p). \end{cases}$$

In this equations, the electron propagator $S(p)$ depends *a priori* on an external electromagnetic field $A_\lambda(q)$, and the functional derivative $\frac{\delta S(p)}{\delta A_\lambda(q)}$ detects this dependence. Then, since the equations are valid for an isolated system, we imply that $S(p)$ and its functional derivatives are evaluated at $A_\lambda(q) = 0$.

The solution of the equations (1.1) is a perturbative series in the powers of the *fine structure constant* $\alpha = \frac{e^2}{4\pi}$,

$$(1.2) \quad \begin{cases} S(p) = S_0(p) + \sum_{n \geq 1} \alpha^n S_n(p), \\ D_{\mu\nu}(p) = D_{0,\mu\nu}(p) + \sum_{n \geq 1} \alpha^n D_{n,\mu\nu}(p), \end{cases}$$

obtained recursively from the initial datas, which are the free propagators $S_0(p)$, $D_{0,\mu\nu}(p)$ and all the functional derivatives $\frac{\delta S_0(p)}{\delta A_\lambda(q)}$, $\frac{\delta^2 S_0(p)}{\delta A_\lambda(q_1) \delta A_\lambda(q_2)}$, etc. At each

new order of perturbation, the coefficient of the series is an analytical expression involving the previous coefficients. In fact, for the series (1.2), the system (1.1) is equivalent to

$$(1.3) \quad \left\{ \begin{array}{l} S_n(p) = iS_0(p) \sum_{k+l=n-1} \int \frac{d^4q}{(2\pi)^4} \gamma^\lambda D_{k,\lambda\lambda'}(q) \frac{\delta S_l(p-q)}{e \delta A_{\lambda'}(q)}, \\ D_{n,\mu\nu}(p) = -iD_{0,\mu\lambda}(p) \sum_{k+l=n-1} \int \frac{d^4q}{(2\pi)^4} \text{tr} \left[\gamma^\lambda \frac{\delta S_k(q)}{e \delta A_{\lambda'}(-p)} \right] D_{l,\lambda'\nu}(p). \end{array} \right.$$

1.2. Renormalization factors and Feynman graphs. Equations (1.3) can be formally solved, but they present a major problem of quantum field theory: at any order of perturbation $n > 0$, even for $n = 1$, the perturbative coefficients $S_n(p)$ and $D_{n,\mu\nu}(p)$ contain divergent integrals. The physical explanation is that the quantum fields ψ and A , and the measurable parameters e and m , have an energy level which is far too high with respect to the classical ones, and the classical Lagrangian $\mathcal{L}_{QED}(\psi, A; e, m)$ is not sufficient to describe them. As a consequence, the Lagrangian must be modified into the so-called the *renormalized Lagrangian*

$$\begin{aligned} \mathcal{L}_{QED}^{ren}(\psi, A; e, m) = & Z_2(e) \bar{\psi}(x^\mu) (i\gamma^\mu \partial_\mu - (m + \delta m(e))) \psi(x^\mu) \\ & - Z_3(e) \frac{1}{4} (\partial^\mu A^\nu - \partial^\nu A^\mu)^2 - \frac{\xi}{2} (\partial_\nu A^\mu)^2 \\ & - Z_1(e) e \bar{\psi} \gamma^\mu \psi A^\mu, \end{aligned}$$

where the so-called *renormalization factors* $Z_1(e)$, $Z_2(e)$, $Z_3(e)$ and $\delta m(e)$ are series in the powers of α which compensate the divergences and allow to produce finite quantities. If we call *bare parameters* the rescaled parameters

$$(1.4) \quad \left\{ \begin{array}{l} m_0 = m + \delta m(e), \\ e_0 = \frac{e Z_1(e)}{Z_2(e) \sqrt{Z_3(e)}}, \end{array} \right.$$

and *bare fields* the rescaled fields

$$\left\{ \begin{array}{l} \psi_0 = \sqrt{Z_2(e)} \psi, \\ A_0^\mu = \sqrt{Z_3(e)} A^\mu, \end{array} \right.$$

it turns out that

$$\mathcal{L}_{QED}^{ren}(\psi, A; e, m) = \mathcal{L}_{QED}(\psi_0, A_0; e_0, m_0).$$

Then, using the Dyson-Schwinger equations (1.1) one can compute the *bare propagators* $S(p; e_0, m_0)$ and $D_{\mu\nu}(p; e_0, m_0)$ relative to the bare fields, and F. Dyson showed in [Dys49] that the *renormalized propagators* can be found simply as

$$(1.5) \quad \left\{ \begin{array}{l} S^{ren}(p; e, m) = Z_2(e)^{-1} S(p; e_0, m_0), \\ (D_{\mu\nu}^T)^{ren}(p; e, m) = Z_3(e)^{-1} D_{\mu\nu}^T(p; e_0, m_0), \end{array} \right.$$

where $D_{\mu\nu}^T(p) = D_{\mu\nu}(p) - \frac{1}{\xi} \frac{p_\lambda p_\mu}{(p^2 + ie)^2}$ is the transversal part of the propagator, while the longitudinal part $\frac{1}{\xi} \frac{p_\lambda p_\mu}{(p^2 + ie)^2}$ is not affected by the renormalization.

One way to compute the renormalization factors is to expand the propagators over all possible connected *Feynman diagrams* with 2 external legs. The amplitude of the graphs are the simplest integrals appearing in the perturbative coefficients of the integral expansion of the propagators. Those which diverge must be renormalized. The divergence extracted from a graph forms the *counterterm* of the graph, and the renormalization factors are constructed by reassembling the counterterms of all concerned graphs.

Feynman graphs describe all possible virtual interactions between electrons and photons in the indeterminate quantum fluctuation. They are constructed on some fixed types of edges, which represent the free propagators, and with some fixed types of vertices, which represent the allowed interactions, and therefore are completely determined by the Lagrangian. If we denote by Γ the connected Feynman graphs with 2 external legs, and we distinguish the electron graphs (e) from the photon graphs (γ), the perturbative expansion of the QED propagators is

$$(1.6) \quad \left\{ \begin{array}{l} S(p) = S_0(p) + \sum_{\Gamma^e} \alpha^{L(\Gamma)} S_\Gamma(p), \\ D_{\mu\nu}(p) = D_{0,\mu\nu}(p) + \sum_{\Gamma^\gamma} \alpha^{L(\Gamma)} D_{\Gamma,\mu\nu}(p), \end{array} \right.$$

where $L(\Gamma)$ denotes the number of loops of the graph. The coefficients $S_\Gamma(p)$ and $D_{\Gamma,\mu\nu}(p)$ are some analytical expressions called the *amplitude* of the graph. The integral expansion (1.2) and the diagrammatic expansion (1.6) of the solution of Eqs. (1.1) are of course related: at each order of perturbation we have

$$\left\{ \begin{array}{l} S_n(p) = \sum_{L(\Gamma)=n} S_\Gamma(p), \\ D_{n,\mu\nu}(p) = \sum_{L(\Gamma)=n} D_{\Gamma,\mu\nu}(p). \end{array} \right.$$

The system (1.1), applied to the series expanded as (1.6), gives an algorithm to construct the amplitudes, called the *Feynman's rules*, cf. for instance Appendix A in [IZ80]. Roughly speaking, Feynman's rules are a dictionary between graphical signs (edges, vertices, loops) and analytical expressions (free propagators, factors of the coupling constant, integrals, overall symmetry factors). Feynman's rules have many advantages: they are intuitive and easy to memorize, they allow to compute the coefficients of the propagators each independently of any other, and finally they allow to compute the counterterms of the graphs and therefore the coefficients of the renormalization factors. Based on this method, one can prove for instance the *Ward-Takahashi identity* $Z_2(e) = Z_1(e)$, which imply that the charge is renormalized as

$$(1.7) \quad e_0 = e Z_3(e)^{-\frac{1}{2}}.$$

The computation of the counterterms requires a specific algorithm, the so-called *BPHZ formula*, which takes into account the subdivergences of the complicated graphs. The intricate combinatorics of this algorithm is nowadays completely clear, since A. Connes and D. Kreimer proved in [CK00a, CK00b] that it is equivalent to a Hopf algebra structure on the set of Feynman graphs.

However the method based on Feynman graphs has also two main disadvantages. On one side, the number of Feynman diagrams increases exponentially with the number of vertices, making difficult to keep control of the summability of the perturbative expansions on graphs. On the other side, the amplitude of a graph with many vertices turns out to be a very complicated analytical expression. The computation of the propagators requires hundreds of integrals even at low orders of perturbation, and the computation of the counterterms, performed through the BPHZ formula, makes the computation of each renormalized integral even longer.

1.3. Tree-expansion of QED propagators. In [Bro00], C. Brouder proposed an alternative perturbative solution of the Dyson-Schwinger equations (1.1), which has the main advantage of reducing drastically the number of computations required at each order of perturbation. To do it, he based the expansion of the perturbative series on the set of *planar binary rooted trees*, instead of Feynman graphs. These trees are planar graphs without loops, and a preferred external edge called the root. For a tree t , we denote by $|t|$ the number of its internal vertices. The tree $|$, with no internal vertices, is called the *root tree*. The tree γ , with one internal vertex, is called the *vertex tree*. Other examples of trees, those with number of internal vertices equal to two and three, are

$$\gamma, \quad \gamma, \quad \gamma, \quad \gamma, \quad \gamma, \quad \gamma, \quad \gamma.$$

Let us denote by Y_n the set of planar binary trees with n internal vertices, by $Y = \bigcup_{n=0}^{\infty} Y_n$ the set of all trees, and by $\bar{Y} = \bigcup_{n=1}^{\infty} Y_n$ the set of all trees with at least one internal vertex. Finally, let us denote by $\vee : Y_n \times Y_m \rightarrow Y_{n+m+1}$ the operation which grafts two trees on a new root, that is

$$s \vee t = \begin{array}{c} s \quad t \\ \gamma \end{array}.$$

If we suppose that the bare 2-point Green's functions of QED are formal series of the form

$$(1.8) \quad \left\{ \begin{array}{l} S(p; e_0, m_0) = S_0(p) + \sum_{t \in \bar{Y}} e_0^{2|t|} S_t(p), \\ D_{\mu\nu}(p; e_0, m_0) = D_{0,\mu\nu}(p) + \sum_{t \in \bar{Y}} e_0^{2|t|} D_{t,\mu\nu}(p), \end{array} \right.$$

then the system (1.1) determines the coefficients of these series as

$$(1.9) \quad \left\{ \begin{array}{l} S_t(p) = iS_0(p) \int \frac{d^4 q}{(2\pi)^4} \gamma^\lambda D_{t^l, \lambda\lambda'}(q) \frac{\delta S_{tr}(p-q)}{e_0 \delta(A_0)_{\lambda'}(q)}, \\ D_{t,\mu\nu}(p) = -iD_{0,\mu\nu}(p) \int \frac{d^4 q}{(2\pi)^4} \text{tr} \left[\gamma^\lambda \frac{\delta S_{t^l}(q)}{e_0 \delta(A_0)_{\lambda'}(-p)} \right] D_{tr, \lambda'\nu}(p), \end{array} \right.$$

where, for any tree $t \neq |$, the trees t^l and t^r are such that $t = t^l \vee t^r$. The expansion on trees (1.8) is related to the integral expansion (1.2) because the coefficients of the integral series at the perturbative order n are the sum of the corresponding coefficients over the trees t such that $2|t| = n$. It is also related to the expansion (1.6) on Feynman graphs because the coefficients depending on a tree t are the sum of the coefficients depending on some Feynman graphs which can be found with an

algorithm given in [BF01]. Roughly speaking, a tree corresponds to a finite sum of Feynman diagrams.

A priori, a BPHZ formula for the amplitude of the trees is not known, nor its equivalent algebraic version based on a Hopf algebra of planar binary trees. Therefore, to find the tree-expansion for the renormalized propagators, it is necessary to solve directly the Dyson-Schwinger equations to which they are constrained. These are best given on the inverse of the photon propagator, that is, on the *vacuum polarization*

$$\Pi_{\lambda\mu}(p) = (p_\lambda p_\mu - p^2 g_{\lambda\mu}) - \xi p_\lambda p_\mu - [D^{-1}]_{\lambda\mu}(p).$$

Then, the *renormalized Dyson-Schwinger equations* are given by

$$(1.10) \quad \begin{cases} S^{ren}(p; e, m) = Z_2^{-1} S_0(p) - \delta m S_0(p) S^{ren}(p) \\ \quad + i e^2 S_0(q) \int \frac{d^4 q}{(2\pi)^4} \gamma^\lambda D^{ren}{}_{\lambda\lambda'}(q) \frac{\delta S^{ren}(p-q)}{e \delta A_{\lambda'}(q)}, \\ \Pi_{\lambda\mu}^{ren}(p; e, m) = (1 - Z_3)(p_\lambda p_\mu - p^2 g_{\lambda\mu}) \\ \quad - i e^2 Z_2 \int \frac{d^4 q}{(2\pi)^4} \text{tr} \left[\gamma_\lambda \frac{\delta S^{ren}(q)}{e \delta A^\mu(-p)} \right]. \end{cases}$$

Assuming that the propagators are expanded on planar binary trees according to

$$(1.11) \quad \begin{cases} S^{ren}(p; e, m) = S_0(p) + \sum_{t \in \bar{Y}} e^{2|t|} S_t^{ren}(p), \\ D_{\mu\nu}^{ren}(p; e, m) = D_{0,\mu\nu}(p) + \sum_{t \in \bar{Y}} e^{2|t|} D_{t,\mu\nu}^{ren}(p), \end{cases}$$

and that the renormalization factors are also expanded on planar binary trees, that is,

$$(1.12) \quad \begin{aligned} Z_2(e) &= 1 + \sum_{|t|>0} e^{2|t|} C_2(t), \quad \text{with } C_2(\mid) = 1, \\ Z_3(e) &= 1 - \sum_{|t|>0} e^{2|t|} C_3(t), \quad \text{with } C_3(\mid) = 1, \\ \delta m(e) &= \sum_{|t|>0} e^{2|t|} C_m(t), \quad \text{with } C_m(\mid) = 0, \end{aligned}$$

then the system (1.10) determines the coefficients of the renormalized propagators as

$$(1.13) \quad \begin{cases} S_t^{ren}(p) = C_2(I(t)) - S_0(p) (C_m * S^{ren})(t) \\ \quad + i S_0(p) \int \frac{d^4 q}{(2\pi)^4} \gamma^\lambda D_{t^l, \lambda\lambda'}^{ren}(q) \frac{\delta S_{t^r}^{ren}(p-q)}{e \delta A_{\lambda'}(q)}, \\ D_{t,\mu\nu}^{ren}(p) = D_{\mid \vee t_r, \mu\nu}^{ren}(p) (p_\mu p_\nu - p^2 g_{\mu\nu}) D_{t_l, \mu\nu}^{ren}(p), \text{ if } t = t_l \vee t_r \text{ with } t_l \neq \mid, \\ D_{\mid \vee t, \mu\nu}^{ren}(p) = C_3(\mid \vee t) D_{0,\mu\lambda}^T(p) \\ \quad - i D_{0,\mu\lambda}^T(p) \int \frac{d^4 q}{(2\pi)^4} \text{tr} \left[\gamma^\lambda \frac{\delta(S^{ren}(q) * C_2)(t)}{e \delta A_{\lambda'}(-p)} \right] D_{0,\lambda'\nu}^T(p). \end{cases}$$

In these equations we make use of a convolution $*$ dual to the pruning coproduct on trees defined recursively by $\Delta_e^{\text{inv}}(|) = | \otimes |$ and

$$(1.14) \quad \Delta_e^{\text{inv}}(t) = | \otimes t + \sum_{\Delta_e^{\text{inv}}(t_r)} (t_l \vee (t_r)_{(1)}) \otimes (t_r)_{(2)}, \quad \text{for } t = t_l \vee t_r.$$

The map I appearing in the first equation is the antipode associated to this coproduct.

The coproduct Δ_e^{inv} turns out to be related to the operation *under* among trees, introduced by J.-L. Loday in [Lod02]: given two planar binary trees $u, v \neq |$, the tree u under v , denoted by $u \setminus v$, is the tree obtained by grafting the right-most leaf of u under the root of v , that is,

$$u \setminus v = u' \overset{v}{\backslash} .$$

Then, Eq. (1.14) can be simply rewritten as

$$\Delta_e^{\text{inv}}(t) = \sum_{t=u \setminus v} u \otimes v,$$

where we also suppose that Δ_e^{inv} is a multiplicative map which respects the decomposition of a tree $t = t_l \vee t_r$ into its two multiplicative factors $t = (t_l \vee |) \setminus t_r$.

In fact, observing carefully the equations for the photon propagator, one can see that another operation among trees is implicitly used: the decomposition of a tree t into two multiplicative factors t^l and $| \vee t^r$. The product among trees which corresponds to this decomposition is the operation *over*, cf. [Lod02]: given two planar binary trees $u, v \neq |$, the tree u over v , denoted by u/v , is the tree obtained by grafting the root of u over the left-most leaf of v , that is,

$$u/v = u \overset{v}{\backslash} .$$

Associated to this product there is of course a coproduct

$$\Delta_\gamma^{\text{inv}}(t) = \sum_{t=u/v} u \otimes v.$$

It turns out that the relationship between the bare propagators $S(p; e_0, m_0)$, $D_{\mu\nu}(p; e_0, m_0)$ and the renormalized propagators $S^{ren}(p; e, m)$, $D_{\mu\nu}^{ren}(p; e, m)$, all expanded on planar binary trees, is governed by two Hopf algebra structures on the set of trees based on the two coproducts Δ_e^{inv} , $\Delta_\gamma^{\text{inv}}$ and also on another coproduct Δ^α analogue to the Connes-Kreimer coproduct on Feynman diagrams.

2. Renormalization Hopf algebras and groups of tree-expanded series

2.1. QED renormalization Hopf algebras on trees. In the papers [BF01] and [BF03] it is shown that the relationship between the QED propagators before and after the renormalization, excluding the mass renormalization and considering the expansion over planar binary trees, is described by a semidirect coproduct of Hopf algebras on trees, involving a Hopf algebra \mathcal{H}^e for the electron propagator, one for the photon propagator \mathcal{H}^γ , and one for the electric charge \mathcal{H}^α .

As algebras, \mathcal{H}^e and \mathcal{H}^γ are both isomorphic to the free non-commutative algebra generated by all trees, where we identify the root tree $|$ to the formal unit. Therefore $\mathcal{H}^e = \mathcal{H}^\gamma = \mathbb{Q}\langle\overline{Y}\rangle \cong \mathbb{Q}\langle Y \rangle / (| - 1)$. The coalgebra structures are given

by the pruning coproducts $\Delta_\gamma^{\text{inv}} : \mathcal{H}^\gamma \longrightarrow \mathcal{H}^\gamma \otimes \mathcal{H}^\gamma$ and $\Delta_e^{\text{inv}} : \mathcal{H}^e \longrightarrow \mathcal{H}^e \otimes \mathcal{H}^e$ defined as the dual operations of the two grafting products *over* and *under* on trees:

$$\begin{aligned} \text{over: } t/s &= \overset{t}{\backslash} s, \quad \Delta_\gamma^{\text{inv}}(u) = \sum_{t/s=u} t \otimes s, \\ \text{under: } t \backslash s &= t' \overset{s}{/}, \quad \Delta_e^{\text{inv}}(u) = \sum_{t \backslash s=u} t \otimes s. \end{aligned}$$

The counits $\varepsilon : \mathcal{H}^\gamma \longrightarrow \mathbb{Q}$ and $\varepsilon : \mathcal{H}^e \longrightarrow \mathbb{Q}$ are dual to the unit $|$, that is $\varepsilon(t) = \delta_{t,|}$. Note that the Hopf algebras \mathcal{H}^e and \mathcal{H}^γ are neither commutative nor cocommutative.

On the other side, \mathcal{H}^α is the abelian quotient of the algebra $\mathbb{Q}Y$ of all trees endowed with the over product (which is not commutative). Thus the root tree $|$ is the unit, and if we set $V(t) = \gamma \backslash t$, the algebra \mathcal{H}^α is in fact isomorphic to the polynomial algebra $\mathbb{Q}[V(t), t \in Y]$. The coproduct $\Delta^\alpha : \mathcal{H}^\alpha \longrightarrow \mathcal{H}^\alpha \otimes \mathcal{H}^\alpha$ is defined on the generators by the assignment

$$(2.1) \quad \Delta^\alpha(V(t)) = 1 \otimes V(t) + \delta^\alpha(V(t)),$$

where $\delta^\alpha : \mathcal{H}^\alpha \longrightarrow \mathcal{H}^\alpha \otimes \mathcal{H}^\alpha$ is a right coaction of \mathcal{H}^α on itself (w.r.t. the coproduct Δ^α), defined recursively as

$$(2.2) \quad \delta^\alpha(V(t)) = (V \otimes \text{Id}) [\Delta^\alpha(t^l) / \delta^\alpha(V(t^r))],$$

where $t = t^l / V(t^r)$. The counit $\varepsilon : \mathcal{H}^\alpha \longrightarrow \mathbb{Q}$ is of course the unital algebra morphism with value $\varepsilon(V(t)) = 0$ on the generators.

According to [BF01] and [BF03], the massless renormalization of the electron propagator expanded over planar binary trees is described by the semidirect (or smash) coproduct Hopf algebra $\mathcal{H}^{\text{qed}} = \mathcal{H}^\alpha \ltimes \mathcal{H}^e$, while the massless renormalization of the photon propagator is described by the Hopf algebra \mathcal{H}^α itself, seen as a Hopf subalgebra of the semidirect coproduct $\mathcal{H}^\alpha \ltimes \mathcal{H}^\gamma$. These semidirect coproduct Hopf algebras can be defined, according to [Mol77], because there is a coaction of \mathcal{H}^α on \mathcal{H}^e and another one on \mathcal{H}^γ , namely two variations of δ^α , which preserve the algebra structures.

The results can be summarized as follows:

- the bare and renormalized amplitudes $S(p; e_0, m_0)$, $S^{\text{ren}}(p; e_0, m_0)$ and the renormalization factor $Z_2(e)$ can be seen as characters of the Hopf algebra \mathcal{H}^e with values in an algebra of regularized amplitudes \mathcal{A}_{reg} ;
- the bare and renormalized amplitudes $D_{\mu\nu}(p; e, m)$, $D_{\mu\nu}^{\text{ren}}(p; e, m)$ and the renormalization factor $Z_3(e)$ can be seen as characters of the Hopf algebra \mathcal{H}^γ , with values in the same algebra \mathcal{A}_{reg} ;
- the bare charge $e_0(e)$ can be seen as a character of the Hopf algebra \mathcal{H}^α , with values in the same algebra \mathcal{A}_{reg} ;
- the relationship between the tree-expanded coefficients of the propagators before and after the renormalization for massless QED is given by the convolution associated to the coproducts of the Hopf algebras $\mathcal{H}^{\text{qed}} = \mathcal{H}^\alpha \ltimes \mathcal{H}^e$, for the electron, and $\mathcal{H}^\alpha \subset \mathcal{H}^\alpha \ltimes \mathcal{H}^\gamma$ for the photon.

It is well known that the characters of a commutative Hopf algebra form a group. If the algebra is the inductive limit of finite-dimensional graded algebras, as is the case here, the group is pro-algebraic (and even pro-unipotent). In the next

section we describe the pro-algebraic groups formed by the characters of the Hopf algebras related to the QED renormalization, abelianized when necessary.

2.2. Groups of tree-expanded series. The abelian quotients \mathcal{H}_{ab}^γ and \mathcal{H}_{ab}^e of the two Hopf algebras \mathcal{H}^γ and \mathcal{H}^e are the coordinate rings of two pro-algebraic groups, denoted by G^γ and G^e respectively, cf. [Fra08]. For a given unital, associative and commutative algebra \mathcal{A} , consider the set

$$G_Y^{\text{inv}} = \left\{ \mathbf{A} = \sum_{t \in Y} \mathbf{A}_t \ t, \ \mathbf{A}_{|} = 1 \right\},$$

of tree-expanded series with coefficients \mathbf{A}_t in \mathcal{A} and invertible constant term. The groups G^γ and G^e are given on G^{inv} respectively by the over and the under products on series, induced by the analogous operations on trees:

$$\begin{aligned} \text{over} \quad \mathbf{A}/\mathbf{B} &= \sum_{s,t \in Y} \mathbf{A}_s \ \mathbf{B}_t \ s/t, \quad \text{for } G^e, \\ \text{under} \quad \mathbf{A}\backslash\mathbf{B} &= \sum_{s,t \in Y} \mathbf{A}_s \ \mathbf{B}_t \ s\backslash t, \quad \text{for } G^\gamma. \end{aligned}$$

In both cases, the unit is given by the series $|$.

We can identify a tree-expanded series $\mathbf{A} = \sum_{t \in Y} \mathbf{A}_t \ t$ with a “generalized” series on a variable x , by setting $\mathbf{A}(x) = \sum_{t \in Y} \mathbf{A}_t \ x^t$, where the monomial x^t is a formal symbol. Then, it is easy to see that the groups G^e and G^γ are non-abelian generalizations of the abelian group

$$G^{\text{inv}} = \left\{ \mathbf{A}(x) = \sum_{n \geq 0} \mathbf{A}_n \ x^n, \ \mathbf{A}_0 = 1 \right\},$$

of usual invertible series in one variable, with coefficients \mathbf{A}_n in \mathcal{A} , and considered with the multiplication.

Similarly, the Hopf algebra \mathcal{H}^α can be realized as a group of tree-expanded series, endowed, this time, with an operation which generalizes the composition (substitution) of formal series. To do this, we first consider the set

$$G_Y^{\text{dif}} = \left\{ \mathbf{A} = \sum_{t \in \overline{Y}} \mathbf{A}_t \ t, \ \mathbf{A}_{\gamma} = 1 \right\}$$

of tree-expanded formal diffeomorphisms with coefficients \mathbf{A}_t in \mathcal{A} . The composition of two tree-expanded series $\mathbf{A} = \sum_t \mathbf{A}_t \ t$ and $\mathbf{B} = \sum_s \mathbf{B}_s \ s$ is given by

$$\mathbf{A} \circ \mathbf{B} = \sum_{\substack{t \in \overline{Y} \\ s_1, s_2, \dots, s_{|t|} \in \overline{Y}}} \mathbf{A}_t \ \mathbf{B}_{s_1} \mathbf{B}_{s_2} \dots \mathbf{B}_{s_{|t|}} \ \mu_t(s_1, s_2, \dots, s_{|t|}),$$

where $\mu_t(s_1, s_2, \dots, s_{|t|})$ is the tree obtained by inserting the trees $s_1, \dots, s_{|t|}$ in the vertices of t , ordered from left to right and from the leaves to the root of t . In fact, the operation μ is the operadic composition of the Duplication operad, also called OverUnder operad, cf. [Lod06, AL07, van03], and the group G_Y^{dif} is an example of a general construction which works for any connected operad, cf. [Cha, van04, Fra08]. The unit of the group G_Y^{dif} is the series γ .

As before, we can identify a tree-expanded series $A = \sum_{t \in \bar{Y}} A_t t$ with a “generalized” series in one variable, $A(x) = \sum_{t \in \bar{Y}} A_t x^t$. Then, if we define the power of the series $B(x)$ by a tree $t \neq \perp$ as $B(x)^t = \mu_t(B(x), B(x), \dots, B(x))$, the composition of tree-expanded series can also be seen as a substitution, that is $(A \circ B)(x) = A(B(x)) = \sum_{t \in \bar{Y}} A_t B(x)^t$. In [Fra08] it is then shown that G_Y^{dif} is a group which projects onto the group

$$G^{\text{dif}} = \left\{ A(x) = \sum_{n \geq 1} A_n x^n, A_1 = 1 \right\}$$

of usual formal diffeomorphisms on a line (tangent to the identity).

The group G_Y^{dif} contains two proper subgroups

$$G_Y^\rho = G^\gamma / \gamma = \left\{ \rho_A = \sum_{t \in Y} A_t \gamma^t, A_\perp = 1 \right\},$$

$$G_Y^\lambda = \gamma \setminus G^e = \left\{ \lambda_A = \sum_{t \in Y} A_t \gamma^t, A_\perp = 1 \right\},$$

whose projection onto usual series is the product $x G^{\text{inv}}$, which is isomorphic to G^{dif} itself.

In [Fra08] it was proved that the group G^α is a proper subgroup of G_Y^ρ of the form

$$(2.3) \quad G^\alpha = \left\{ \alpha_A = (\perp - \gamma \setminus A)^{-1} / \gamma, A \in \mathcal{A}[[Y]] \right\},$$

where $\mathcal{A}[[Y]]$ denotes the set of tree-expanded formal series $A = \sum_{t \in Y} A_t t$ with coefficients A_t in the algebra \mathcal{A} . In fact, for any tree-expanded formal series $A = \sum_{t \in Y} A_t t \in \mathcal{A}[[Y]]$, the series $\perp - \gamma \setminus A$ belongs to the set G_Y^{inv} , and therefore to the group G^γ of tree-expanded invertible series with respect to the product $/$. The projection of G^α onto usual series is the group $x(1 - x\mathcal{A}[[x]])^{-1}$ which coincides, again, with the whole group G^{dif} .

3. Tree-expanded series in relation with Tamari posets

In this section, we will consider some examples of tree-expanded series in G_Y^{dif} . We will compute the inverse of some of them with respect to composition and show that some others are related to the Tamari lattices.

We will need the following involution on tree-expanded series. Let $S = \sum_{n \geq 1} S_n$ be a tree-expanded series in G_Y^{dif} , decomposed with respect to the order of the trees. The *suspension* of S is the series $\tilde{S} = \sum_{n \geq 1} (-1)^{n-1} S_n$. The suspension of a composition $\tilde{A} \circ \tilde{B}$ is the composition of the suspensions $\tilde{A} \circ \tilde{B}$.

We will also use the following properties of the composition:

$$(3.1) \quad \forall A, C, D \quad (A/D) \circ C = (A \circ C)/(D \circ C) \quad \text{and} \quad (A \setminus D) \circ C = (A \circ C) \setminus (D \circ C).$$

These properties follow from the relation of G_Y^{dif} with the Duplication operad.

Let us consider the tree-expanded series A defined inductively by

$$(3.2) \quad A = \gamma + w A / \gamma + \gamma \setminus A + w A / \gamma \setminus A,$$

where w is a formal parameter. This is in fact the sum over all trees where each tree t appears with coefficient w to the power the number of right-oriented leaves of t minus 1:

$$(3.3) \quad A = \text{Y} + w \text{Y} + \text{Y} + w^2 \text{Y} + w \text{Y} + w \text{Y} + w \text{Y} + \text{Y} + \dots$$

Let us consider also the tree-expanded series B defined by

$$(3.4) \quad B = \text{Y} - w \text{Y}/B - B \setminus \text{Y} - w \text{Y}/B \setminus \text{Y},$$

where w is a formal parameter. This is in fact the sum over $p \geq 0$ and $q \geq 0$ of $(-1)^{p+q} w^p c_p / \text{Y} \setminus d_q$ where c_p is the left comb with p vertices and d_q is the right comb with q vertices:

$$(3.5) \quad B = \text{Y} - w \text{Y} - \text{Y} + w^2 \text{Y} + w \text{Y} + \text{Y} + \dots$$

PROPOSITION 3.1. *In the group G_Y^{dif} , one has $A \circ B = \text{Y}$, that is A is the inverse of B . This identity can be specialized to any complex value of w .*

PROOF. Let us compute the composition $A \circ B$ starting from the definition of A . By properties of the composition, one gets, by multiplication on the right by B , that

$$A \circ B = B + w(A \circ B)/B + B \setminus (A \circ B) + w(A \circ B)/B \setminus (A \circ B).$$

Comparing this to the following form of the definition of B :

$$(3.6) \quad \text{Y} = B + w \text{Y}/B + B \setminus \text{Y} + w \text{Y}/B \setminus \text{Y},$$

one can see that Y and $A \circ B$ both satisfy the same induction and have the same initial term, hence they are equal. \square

Let us now consider the tree-expanded series C and D defined by

$$(3.7) \quad C = \text{Y} + C/\text{Y} = \text{Y} + \text{Y}/C = \text{Y} + \text{Y} + \text{Y} + \dots$$

and

$$(3.8) \quad D = \text{Y} + D \setminus \text{Y} = \text{Y} + \text{Y} \setminus D = \text{Y} + \text{Y} + \text{Y} + \dots$$

So C is in fact the sum over all left combs and D is the sum over all right combs. The series C (resp. D) belongs to the subgroup of G_Y^{dif} formed by series indexed by left combs only (resp. right combs only). It is quite clear (by inversion in these subgroups) that $C^{-1} = \tilde{C}$ and $D^{-1} = \tilde{D}$.

PROPOSITION 3.2. *One has*

$$(3.9) \quad C + C \setminus D = D + C/D.$$

The series $E = C \circ D^{-1}$ is characterized by the equation

$$(3.10) \quad E = \text{Y} + E/\text{Y} - E \setminus \text{Y}.$$

PROOF. It is enough to see that the left hand side of (3.9) is exactly the sum of all trees of the shape $c^p/\text{Y} \setminus d^q$ where c_p is a left comb (possibly empty) and d_q is a right comb (idem). By symmetry, this is also true for the right-hand side, hence they are equal. The equation for $C \circ D^{-1}$ is obtained from this by composition on the right by D^{-1} . \square

The series E appears surprisingly to be related to the Möbius function of the Tamari lattice. Recall that the Tamari lattice is a classical partial order [HT72] on the set Y_n of trees of order n , where the left comb c_n is the unique minimal element and the right comb d_n is the unique maximal element. The partial order is defined by transitive closure of the following relation: $t \leq t'$ if t' is obtained from t by replacing a local configuration \swarrow by a local configuration \searrow .

PROPOSITION 3.3. *One has the following description of the series E and D^{-1} :*

$$(3.11) \quad E = \sum_{n \geq 1} \sum_{t \in Y_n} \mu(c_n, t) t$$

and

$$(3.12) \quad D^{-1} = \swarrow + \sum_{n \geq 1} \sum_{t \in Y_n} \mu(c_n, \swarrow \setminus t) \swarrow \setminus t,$$

where μ is the Möbius function of the Tamari lattice.

PROOF. The first statement is equivalent to the second one, by using the following property of Tamari lattices, see for instance [BW97, Lemma 2.1]: any interval $[c_{p+q}, t'/t'']$ in a Tamari lattice is isomorphic to the product of the intervals $[c_p, t']$ and $[c_q, t'']$ in smaller Tamari lattices. This shows that the generating series of Möbius number for all trees is the composition of C by the generating series of Möbius number for trees of the shape $\swarrow \setminus t$.

The second statement can be deduced from the computation of the Möbius function of the Tamari lattice by Björner and Wachs: see [BW97, Corollary 9.5]. In their notations, c_n corresponds to the word $(0, \dots, 0)$ and a tree of shape $\swarrow \setminus t$ to a word w_t beginning with the letter $n - 1$. One gets that the Möbius number $\mu((0, \dots, 0), w_t)$ is non zero if and only if the word w_t is $(n - 1, n - 2, \dots, 2, 1)$, which corresponds to the tree $t = d_n$ in our notations. \square

Remark: By using the projection morphism from G_Y^{dif} to G^{dif} , one can see that the sum of coefficients in E of all trees of a fixed order $n > 1$ is zero, which also follows from the proposition.

Let us then introduce the tree-expanded series $R = \swarrow + \swarrow \setminus A$ and $L = \swarrow + A/\swarrow$. In fact, R is the sum of all trees of the shape $\swarrow \setminus t$ and L is the sum of all trees of the shape t/\swarrow . The series R (resp. L) belongs to the subgroup of G_Y^{dif} formed by series indexed by trees of the shape $\swarrow \setminus t$ only (resp. by trees of the shape t/\swarrow only).

PROPOSITION 3.4. *One has*

$$(3.13) \quad R = \swarrow + R \setminus L,$$

$$(3.14) \quad L = \swarrow + R/L.$$

The composition $R \circ L^{-1}$ is equal to the suspension of E .

PROOF. The first two formulas follows directly from a standard combinatorial argument using a decomposition of trees. For instance, the first formula can be deduced from the existence of an unique maximal decomposition of a tree as an iterated \setminus product.

By multiplication on the right by L^{-1} , one has

$$(3.15) \quad R \circ L^{-1} = L^{-1} + (R \circ L^{-1}) \setminus \swarrow,$$

$$(3.16) \quad \swarrow = L^{-1} + (R \circ L^{-1}) / \swarrow.$$

Hence by elimination of L^{-1} one has

$$(3.17) \quad R \circ L^{-1} = \gamma + (R \circ L^{-1}) \setminus \gamma - (R \circ L^{-1}) / \gamma.$$

By definition of the suspension, the suspension of E and $R \circ L^{-1}$ satisfy the same induction, hence they are equal. \square

Remark: one can also deduce from this a similar description of L^{-1} .

References

- [AL07] M. Aguiar and M. Livernet. The associative operad and the weak order on the symmetric groups. *J. Homotopy Relat. Struct.*, 2(1):57–84 (electronic), 2007.
- [AS06] M. Aguiar and F. Sottile. Structure of the Loday-Ronco Hopf algebra of trees. *J. Algebra*, 295(2):473–511, 2006.
- [BF01] C. Brouder and A. Frabetti. Renormalization of QED with planar binary trees. *Eur. Phys. J.*, C19:715–741, 2001.
- [BF03] C. Brouder and A. Frabetti. QED Hopf algebras on planar binary trees. *J. Algebra*, 267(1):298–322, 2003.
- [Bro00] C. Brouder. On the trees of quantum fields. *Eur. Phys. J.*, C12:535–549, 2000.
- [BW97] A. Björner and M. L. Wachs. Shellable nonpure complexes and posets. II. *Trans. Amer. Math. Soc.*, 349(10):3945–3975, 1997.
- [Cha] F. Chapoton. Rooted trees and an exponential-like series. arXiv:math.QA/0209104.
- [CK00a] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann-Hilbert problem I: the Hopf algebra structure of graphs and the main theorem. *Commun. Math. Phys.*, 210:249–273, 2000.
- [CK00b] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann-Hilbert problem II: the β function, diffeomorphisms and the renormalization group. *Commun. Math. Phys.*, 2000. hep-th/0003188.
- [Dys49] F. J. Dyson. The S matrix in quantum electrodynamics. *Phys. Rev.*, 75:1736–55, 1949.
- [Fra08] A. Frabetti. Groups of tree-expanded series. *J. Algebra*, 319(1):377–413, 2008.
- [HNT05] F. Hivert, J.-C. Novelli, and J.-Y. Thibon. The algebra of binary search trees. *Theoret. Comput. Sci.*, 339(1):129–165, 2005.
- [HT72] S. Huang and D. Tamari. Problems of associativity: A simple proof for the lattice property of systems ordered by a semi-associative law. *J. Combinatorial Theory Ser. A*, 13:7–13, 1972.
- [IZ80] C. Itzykson and J.-B. Zuber. *Quantum Field Theory*. McGraw-Hill, New York, 1980.
- [Kre98] Dirk Kreimer. On the Hopf algebra structure of perturbative quantum field theories. *Adv. Theor. Math. Phys.*, 2(2):303–334, 1998.
- [Lod01] Jean-Louis Loday. Dialgebras. In *Dialgebras and related operads*, volume 1763 of *Lecture Notes in Math.*, pages 7–66. Springer, Berlin, 2001.
- [Lod02] J.-L. Loday. Arithmetree. *J. Algebra*, 258(1):275–309, 2002. Special issue in celebration of Claudio Procesi’s 60th birthday.
- [Lod06] J.-L. Loday. Generalized bialgebras and triples of operads, 2006.
- [Mol77] R. K. Molnar. Semi-direct products of Hopf algebras. *J. Algebra*, 47(1):29–51, 1977.
- [Sta99] Richard P. Stanley. *Enumerative combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.
- [van03] P. van der Laan. Operads and the Hopf algebras of renormalisation, 2003.
- [van04] P. van der Laan. *Operads. Hopf algebras and coloured Koszul duality*. PhD thesis, Universiteit Utrecht, 2004.

UNIVERSITÉ DE LYON ;, UNIVERSITÉ LYON 1 ;, CNRS, UMR5208, INSTITUT CAMILLE JORDAN,, 43 BLVD DU 11 NOVEMBRE 1918,, F-69622 VILLEURBANNE-CEDEX, FRANCE

E-mail address: chapoton@math.univ-lyon1.fr, frabetti@math.univ-lyon1.fr

Sweedler's duals and Schützenberger's calculus

Gérard H. E. Duchamp and Christophe Tollu.

ABSTRACT. We describe the problem of Sweedler's duals for bialgebras as essentially characterizing the domain of the transpose of the multiplication. This domain is the set of what could be called “representative linear forms” which are the elements of the algebraic dual which are also representative on the multiplicative semigroup of the algebra.

When the algebra is free, this notion is indeed equivalent to that of rational functions of automata theory. For the sake of applications, the range of coefficients has been considerably broadened, i.e. extended to semirings, so that the results could be specialized to the boolean and multiplicity cases. This requires some caution (use of “positive formulas”, iteration replacing inversion, stable submodules replacing finite-rank families for instance). For the theory and its applications has been created a rational calculus which can, in return, be applied to harness Sweedler's duals. A new theorem of rational closure and application to Hopf algebras of use in Physics and Combinatorics is provided. The concrete use of this “calculus” is eventually illustrated on an example.

1. Introduction

This paper is entirely devoted to questions of rationality which arose, seemingly independently, in automata theory (Schützenberger's calculus) and in the dualization of multiplication (Sweedler's duals). As in the classical (univariate) case, rationality shows itself twofold : on functions and on expressions.

The rationality framework exposed here can be considered as the noncommutative analogue of linear recurrences [16]. It is known, (see [9, 23]), that it is equivalent to state that the coefficients of the Taylor expansion of a function satisfy a linear recurrence or that the generating function itself is rational (*i.e.* the quotient of two polynomials). This equivalence has a counterpart in the theory of rational expressions and this (nowadays classical) theory can be considered as “localized at zero” (*i.e.* analogous to the theory of rational functions without a pole at zero).

The paper is organized as follows.

In section 2, one sets out the theory of representative functions which were introduced in [29] and are now standard in the theory of algebraic groups [20]. Most of the material of this section (preparatory to the subsequent ones) is not new and can be found in several domains (but we believe that, beyond the needs of exposition, the description of the link itself will be of some use to these domains). The scope of proposition (2.1) can be harmlessly extended to semigroups [1].

As representative functions on the free monoid are the core of automata theory [2, 14, 27, 28], the domain of their scalars can (and, in fact, had to) be considerably enlarged to include structures allowing matrix (with unit) computations. This is the domain of semirings [17, 18], the scalars of automata theory.

Section 4 prepares the link between representative functions and rational expressions by means of the notion of star (which is the “positive analogue” of the inverse [6, 19, 24, 25, 26]) and star closure (which is the analogue of rational closure in classical algebra¹).

The framework is then ready for a correct exposition of rational expressions which is the main concern of section 4.

In section 5, we apply what has been constructed to the dualization of bialgebras, then recovering known results.

In section 6, we show how to apply this “rational calculus” to solve the carrier problem in combinatorial physics.

Acknowledgements

The authors wish to acknowledge support from Agence Nationale de la Recherche (Paris, France) under Program No. ANR-08-BLAN-0243-2 as well as support from “Projet interne au LIPN 2009” “Polyzêta functions”.

2. Representative functions on a semigroup

The aim of this section is to discuss the dualization of bi-algebras and Hopf algebras. This problem, solved by Sweedler’s duals, is the following.

Let k be a field and $(B, ., \Delta, 1_B, \varepsilon)$ be a k -bialgebra ; we know that, if B is finite-dimensional (resp. graded in finite dimensions), the dual (resp. graded dual) endowed with the transpose structure is a bialgebra and that, in case B is a Hopf algebra this statement carries over. Now the question can be asked.

What is the good notion of restricted dual for the general (i.e. ungraded finite or infinite dimensional) case ?

Analysing the dualization of the structural operations $(., \Delta, 1_B, \varepsilon)$ of B , one sees at once that only the dualization of the multiplication is problematic as, in the general case, the codomain of the transpose of $.$ is larger than $B^* \otimes B^*$.

The first result follows (and somehow extends) [1]. To state it, we need the notion of (left and right) shifts of functions on a semigroup. Let k be a field, $(S, .)$ a semigroup and $f \in k^S$. For each $s \in S$ define $f_s : x \mapsto f(sx)$ (right shift of f) and ${}_s f : x \mapsto f(xs)$ (left shift of f), ${}_s f_t : x \mapsto f(txs)$ (bi-shift of f). Then, we have.

PROPOSITION 2.1. (see also [1] section 2.2) *Let k be a field, $(S, .)$ a semigroup and $f \in k^S$. The following are equivalent :*

- i) *The family $(f_s)_{s \in S}$ is of finite rank in k^S*
- ii) *The family $({}_s f)_{s \in S}$ is of finite rank in k^S*
- iii) *The family $({}_s f_t)_{s, t \in S}$ is of finite rank in k^S*
- iv) *There exists a double family $(g_i, h_i)_{1 \leq i \leq n}$ of functions such that*

$$(1) \quad \left(\forall x, y \in S \right) \left(f(xy) = \sum_{i=1}^n g_i(x)h_i(y) \right)$$

¹and “sous-algèbres pleines” in spectral theories [4].

v) There exists $\lambda \in k^{1 \times n}$, $\gamma \in k^{n \times 1}$ and $\mu : (S, \cdot) \rightarrow (k^{n \times n}, \times)$ a morphism of semigroups such that $(\forall s \in S)(f(s) = \lambda\mu(s)\gamma)$.

Moreover, if S admits a neutral (i.e. is a monoid), its image by μ of (v) above can be chosen to be the unity matrix.

Proof — Omitted

The elements of k^S which fulfill the above conditions will be called *representative functions* on S and denoted $R(k, S, \cdot)$ [1, 8, 13].

REMARK 2.2. i) When k is only a PID, proposition (2.1) above still holds with the five equivalent conditions, orbits and ranks being computed in \bar{k}^S (\bar{k} is the fraction field of k).

ii) If S is finite, $R(k, S, \cdot) = k^S$ and if S is a group, one has

$$R(k, S, \cdot) = k^S \iff S \text{ is finite}$$

iii) If S is a semigroup, the equivalences above are false in general as shown by the following counterexample. Let G be a finite group and endow $S = \mathbb{N} \times G$ with the law $(n, g) * (m, h) = (0, gh)$. It can be easily checked that $(S, *)$ is a semigroup and that $R(k, S, *) = k^S$.

iv) When W is a shift-invariant subspace of k^S and $f \in W \cap R(k, S, \cdot)$, the families $(f_s)_{s \in S}, (sf)_{s \in S}, (sft)_{s, t \in S}$ are of course in $W \cap R(k, S, \cdot)$ (and are of finite rank). Two useful examples of such “relative representative functional spaces” are with $W = C(S)$ (continuous functions, S being a topological semigroup) and $W = S^*$ (linear forms, S being an algebra with its product as semigroup law).

v) If $T \subset S$ is a subsemigroup with finite set-theoretical complement ($S \setminus T$ is finite) then $f \in k^S$ is representative iff $f|_T$ is so. In particular, if $S = T^{(1)}$, the monoid obtained by adjunction of a unity to T , one has that $f \in k^S$ is representative iff $f|_T$ is.

vi) The proof of proposition (2.1) can also be found in [20] where it does not use the structure of group.

3. Semirings

Throughout the text “monoid” stands for “semigroup with unit”.

Semirings are the structures adapted to matrix (with unity) computation. A semiring $(k, +, \times)$ consists of the following data

- a set k
- two binary laws $+$, \times on k

such that

- $(k, +)$ and (k, \times) are monoids, the first being commutative, their neutrals will be denoted respectively 0_k and 1_k
- \times is left and right distributive over $+$
- 0_k is an annihilator i. e. $(\forall x \in k)(0_k \cdot x = x \cdot 0_k = 0_k)$

EXAMPLE 3.1. i) Any ring.

ii) The boolean semiring $\mathbb{B} = \{0, 1\}$ endowed with the laws $x \oplus y = x + y - xy$ and $x \otimes y = xy$.

iii) The semiring $([-\infty, +\infty[, \max, +)$, called “(max, plus)-semiring” in the literature.

iv) In the semiring $([0, +\infty[, +, \times)$, the laws are continuous at infinity and then can

be completed. We obtain a semiring $([0, +\infty], +, \times)$ which is suited for multiplicities arising in repeated additions of positive values during iterations.

EXAMPLE 3.2. The following example is fundamental and will be used in the definition of CM-modules. If $(M, +)$ is a commutative monoid, then $(\text{End}(M), +, \circ)$ (defined as though M were a group) is a semiring. The units are respectively, the constant mapping $M \ni m \mapsto 0_M$ for $+$ and $\text{Id}_M : m \mapsto m$ for \circ .

The structure of semiring defines a category larger than that of rings, the morphisms being defined similarly. Let $(k_i, +_i, \times_i)$, $i = 1, 2$ be two semirings, a mapping $\phi : k_1 \rightarrow k_2$ is called a morphism of semirings iff it is a morphism for the two structures of monoids (additive and multiplicative), then compatible with the laws and units of k_1 and k_2 .

The definition of modules (here called CM-modules as they are constructed on Commutative Monoids as vector structure) follows also the classical pattern.

The structure of a (left) k -CM-module is given by the following data

- a commutative monoid $(M, +)$
- a morphism (the scaling morphism) of semirings $s : k \rightarrow \text{End}(M)$.

The structure of (right) k -CM-module is defined by replacing $\text{End}(M)$ by $\text{End}^{\text{op}}(M)$ the opposite semiring (constructed with the opposite multiplicative law). Bi- and multimodules are defined as in [3] and follow the general philosophy of “structures with operators”.

EXAMPLE 3.3. Let X be a set, then k^X , the set of all functions $X \rightarrow k$ is naturally endowed with a structure of $k - k$ bimodule defined as in the case when k is a ring. So is $k^{(X)}$, the set of finitely supported functions of k^S (actually a sub- k - k bimodule of k^S).

The free monoid generated by a set X (finite or infinite) is the set of words (i.e. finite sequences of elements of X comprising the empty one denoted by 1_{X^*}) endowed with the concatenation law.

4. Shift operators and rational closure

Let $(M, .)$ be a (commutative or not) monoid. For a function $f : M \rightarrow k$ and $a \in M$, we define the following shift operators [1, 2, 14, 21]

- $f_a : x \mapsto f(ax)$ (right shift)
- ${}_a f : x \mapsto f(xa)$ (left shift)

We also have to describe the analog, for CM-algebras, of *full subalgebras* and *full subalgebraic closures* (see [4] Ch. 1.1.4) and this requires the notion of *summability* [2].

DEFINITION 4.1. A family $(f_i)_{i \in I}$ of functions $M \rightarrow k$ is called *summable* iff, for each $m \in M$, $(f_i(m))_{i \in I}$ is finitely supported. Then, the mapping $m \mapsto \sum_{i \in I} f_i(m)$ is denoted $\sum_{i \in I} f_i$ and called the *sum* of $(f_i)_{i \in I}$.

As a consequence, it is easily checked that, if M is *locally finite* ([14] Vol. A VII.4) and $f : M \rightarrow k$ is without constant term (i.e. $f(1_M) = 0_k$), then the family $(f^n)_{n \in \mathbb{N}}$ (convolutional powers) is summable and its sum

$$(2) \quad \sum_{n \in \mathbb{N}} f^n$$

will be denoted f^* and called the *star* of f .

NOTE 4.2. *There is a lot of literature about the star problem (see [17, 18]). For a general discussion of star-type solutions in a semiring, see [11].*

Now, we are in the position of stating the Kleene-Schützenberger theorem.

THEOREM 4.3. *Let $M = X^*$ be a free monoid, k a semiring and $f \in k^M$. The following are equivalent*

- i) *the family $(f_w)_{w \in M}$ belongs to a finitely generated shift-invariant left-submodule*
 - ii) *the family $(_w f)_{w \in M}$ belongs to a finitely generated shift-invariant right-submodule*
 - iii) *there exist a row $\lambda \in k^{1 \times n}$, a column $\gamma \in k^{n \times 1}$, and a representation (of monoids) $\mu : M \rightarrow (k^{n \times n}, \times)$ such that $(\forall w \in M)(f(w) = \lambda \mu(w) \gamma)$.*
- If X is finite, then (i-iii) above are also equivalent to:*
- iv) *f lies in the rational closure of X (i.e. the smallest subalgebra of $k\langle\langle X \rangle\rangle$ closed under the star operation and containing X).*

REMARK 4.4. i) *Rational elements in the sense of (iii) infinitely many (linearly) independent shifts. That is why finitely generated shift-invariant submodules are needed in the general case. As an example one may consider*

$$S = (a^*)^2 \in \mathbb{N}\langle\langle a \rangle\rangle.$$

One can check at once that $a^{-k}S = k.(a^) + S$ and then for all k , $a^{-(k+1)}S \notin \text{span}(a^{-s}S)_{s=0}^k$.*

- ii) *One can remove the hypothesis of freeness of M if k is a field. Indeed, in this case, the submodule can be taken as generated by the shifts (right or left) of f and the representation is automatically compatible with the relations of M .*
- iii) *Here the star is used as the localization at one (i.e. with positive formulas) of the inverse function. Indeed, with coefficients in a ring, if we are at the neighbourhood of 1, the condition $(1-x)(1+y) = 1$ (resp. $(1+y)(1-x) = 1$) is equivalent to $y = x + xy$ (resp. $y = x + yx$). These self-reproducing positive conditions are taken as the definition of "y is a star of x" in a semiring (see [11]).*
- iv) *The condition (iv) in theorem (4.3) is known as Kleene-Schützenberger theorem as, when k is specialized to \mathbb{B} , this is actually Kleene's theorem. In this sense, this theorem lies at the frontier of harmonic analysis (the set of representative functions is dense in the Fourier space of compact groups), spectral theory (the notion of full subalgebra closure comes from this theory [4]) and theoretical computer science (the notion of a semiring was developed as a computational model of iteration and the notion of a semiring was developed to cope with general scalars as diverse as the ones arising in stochastic automata theory and shortest path problems).*

In the general case (X not necessarily finite), Kleene-Schützenberger's theorem has to be modified as follows.

THEOREM 4.5. *Let $M = X^*$ be a free monoid, k a semiring and $f \in k^M$. The following are equivalent*

- i) *the family $(f_w)_{w \in M}$ belongs to a finitely generated shift-invariant left-submodule*
- ii) *the family $(_w f)_{w \in M}$ belongs to a finitely generated shift-invariant right-submodule*
- iii) *there exist a row $\lambda \in k^{1 \times n}$, a column $\gamma \in k^{n \times 1}$, and a representation (of monoids) $\mu : M \rightarrow (k^{n \times n}, \times)$ such that $(\forall w \in M)(f(w) = \lambda \mu(w) \gamma)$.*

iv) the function f lies in the rational closure of $\overline{kX} = \{\sum_{x \in X} \alpha(x)x\}_{\alpha \in k^X}$ (i.e. the smallest subalgebra of $k\langle\langle X \rangle\rangle$ closed under the star operation which contains \overline{kX}).

REMARK 4.6. The rational closure of X is, in fact, the intersection of the set of elements characterized by (i-iii) (i.e. Sweedler's dual of $k\langle X \rangle$), and the algebra $\bigcup_{\substack{F \subset X \\ F \text{ finite}}} k\langle\langle F \rangle\rangle$ of the series whose support involves a finite alphabet.

5. Rational expressions

The construction of [7] was localized at zero, we extend it here to any localization i.e. for any mapping $\Lambda : X \rightarrow k$.

As the rational closure involves a unary law (the star) partially defined, the definition of universal formulas for this closure needs some caution. Indeed, we need to build in parallel a “character” (the constant term) $const$ so that all proper expressions should have a star.

One first defines, as in [7] the completely free expressions (or formulas) as the terms of the universal algebra defined on $X \cup \{0_E\}$ (0_E , which does not belong to X will serve as a null or void expression and will be mapped to the zero series). This algebra will be denoted $\mathcal{E}^{cf}(X, k)$. More precisely

- If $x \in X \cup \{0_E\}$ then $x \in \mathcal{E}^{cf}(X, k)$.
- If $E, E_1, E_2 \in \mathcal{E}^{cf}(X, k)$, and $\lambda \in k$ then

$$E_1 + E_2 \in \mathcal{E}^{cf}(X, k), E_1 \cdot E_2 \in \mathcal{E}^{cf}(X, k) \\ \lambda E \in \mathcal{E}^{cf}(X, k), E\lambda \in \mathcal{E}^{cf}(X, k).$$

- If $E \in \mathcal{E}^{cf}(X, k)$ then $E^* \in \mathcal{E}^{cf}(X, k)$.

The partial function $const : \mathcal{E}^{cf}(X, k) \rightarrow k$ (constant term) is constructed as follows:

- (1) If $x \in X$ $const(x) = \Lambda(x)$ and $const(0_E) = 0_k$.
- (2) If $E, E_i \in \mathcal{E}^{cf}(X, k)$, $i = 1, 2$ and $\lambda \in k$ then

$$const(E_1 + E_2) = const(E_1) + const(E_2)$$

$$const(E_1 \cdot E_2) = const(E_1) \cdot const(E_2)$$

$$const(\lambda E) = \lambda const(E), const(E\lambda) = const(E)\lambda.$$

- (3) If $const(E) = 0_k$ then $const(E^*) = 1_k$.

The domain of $const$ will be called *rational expressions* and denoted $\mathcal{E}_\Lambda(X, k)$. For example $0_E^* \in \mathcal{E}_\Lambda(X, k)$.

Let now $\Theta : X \rightarrow k\langle\langle A \rangle\rangle$ be a mapping such that, for every $x \in X$, the constant term of $\Theta(x)$, i.e. the coefficient of 1_{A^*} in $\Theta(x)$, is equal to $\Lambda(x)$ (in symbols, $\forall x \in X$, $[1_{A^*}]\Theta(x) = \Lambda(x)$). Following recursively (1-2-3) above, we can construct a polymorphism $\phi_\Theta : \mathcal{E}_\Lambda(X, k) \rightarrow k\langle\langle A \rangle\rangle$ which is a morphism for the laws (2 internal and 2 external) and the star. Moreover $\delta_{1_{A^*}} \circ \phi_\Theta = const$ (i.e. $const$ can be considered as a “constant term function” for the expressions). The image of ϕ_Θ is exactly the rational closure of the set $\{\Theta(x)\}_{x \in X}$.

6. Dual laws and bialgebras

Let $\Delta : \mathcal{A} \rightarrow \mathcal{A} \otimes \mathcal{A}$ be any comultiplication (i.e. \mathcal{A} is a k -coalgebra). It is known that its dual $(\mathcal{A}^*, {}^t\Delta)$ is an algebra and if \mathcal{A} is coassociative (resp. cocomutative, counital), \mathcal{A} is associative (resp. commutative, unital) [1].

We would like here to enlarge the framework of [12].

If \mathcal{A} is an algebra, let us call *dual law* on \mathcal{A}^* a law of the form ${}^t\Delta$ for some (not necessarily coassociative) comultiplication on \mathcal{A} .

In [12] were considered the *dual laws* on $k\langle\langle X \rangle\rangle \simeq k\langle X \rangle^*$ in order to prove that the Hadamard and Infiltration products, which were known to preserve rationality, were essentially the only (along with an interpolation between the two) alphabetic (associative and unital) dual laws between series. The notion of *dual law* provides an implementation scheme for the automata so that the rationality preservation is naturally effective.

THEOREM 6.1. *Let \mathcal{A} be a k -algebra and $\Delta : \mathcal{A} \rightarrow \mathcal{A} \otimes \mathcal{A}$ be a comultiplication which is a morphism of algebras. Then*

- i) *If k is a field, Sweedler's dual \mathcal{A}° of \mathcal{A} is closed under the dual law ${}^t\Delta$.*
- ii) *If k is a semiring and $\mathcal{A} = k\langle X \rangle$, \mathcal{A}° is closed under the dual law ${}^t\Delta$.*

NOTE 6.2. i) *The theorem is no longer true if $\Delta : \mathcal{A} \rightarrow \mathcal{A} \otimes \mathcal{A}$ is arbitrary (i.e. not necessarily a morphism) as shows the following counterexample. With $\Delta : \mathbb{Q}[x] \rightarrow \mathbb{Q}[x] \otimes \mathbb{Q}[x]$ such that $\Delta(x) = \frac{1}{n!}(x^n \otimes x^n)$, one has*

$${}^t\Delta\left(\frac{1}{1-x}, \frac{1}{1-x}\right) = \exp(x).$$

ii) *In (i) above, the restriction on scalars (to be a field) can be extended to inductive limits of PIDs.*

iii) *Comultiplications different from morphisms can preserve rationality. For example, let $\Delta : k\langle X \rangle \rightarrow k\langle X \rangle \otimes k\langle X \rangle$ be a morphism and $\Delta_1 : k\langle X \rangle \rightarrow k\langle X \rangle \otimes k\langle X \rangle$ be a linear mapping which coincides with Δ except for a finite number of words of X^* . It can be checked that Δ_1 , although not a morphism, preserves rationality.*

Let us now return to the case of a bialgebra $(B, ., \Delta, 1_B, \varepsilon)$. The following proposition says that, if a linear form on B is transformed by ${}^t m$ ($m : B \otimes B \rightarrow B$ is just the multiplication mapping) into an element of $B^* \otimes B^*$, it must be of the form exhibited in (iv) of Proposition 2.1. Let us state that precisely.

PROPOSITION 6.3. *Let $(B, m, \Delta, 1_B, \varepsilon)$ be a bialgebra and $f \in B^*$.*

- 1) *The following are equivalent*
- i) ${}^t m(f) \in B^* \otimes B^*$ (for the canonical embedding $B^* \otimes B^* \hookrightarrow (B \otimes B)^*$)
- ii) $f \in R(k, B, .)$
- iii) $\ker(f)$ contains a finite-codimension one-sided ideal
- iv) $\ker(f)$ contains a finite-codimension two-sided ideal
- v) $\text{There exist } \lambda \in k^{1 \times n}, \gamma \in k^{n \times 1} \text{ and } \mu : (B, ., \Delta) \rightarrow (k^{n \times n}, +, \times) \text{ a morphism of } k\text{-algebras (associative with units) such that } (\forall x \in B)(f(x) = \lambda\mu(x)\gamma).$
- 2) *Moreover, let B^0 be the set of linear forms which are decomposable as in 1 – (i) above. Then $(B^0, {}^t\Delta, {}^t m, {}^t \varepsilon, {}^t 1_B)$ is a bialgebra, and if B admits an antipode σ (i.e. is a Hopf algebra), one has ${}^t \sigma(B^0) \subset B^0$ and $(B^0, {}^t\Delta, {}^t m, {}^t \varepsilon, {}^t 1_B, {}^t \sigma)$ is a Hopf algebra.*

7. An application of rational expressions to Combinatorial Physics

In a joint work with J. Katriel, one of us gave the solution of the problem of matrix coefficients in the Fock space of carriers between two levels. We give here a brief review of these continued fractions-type formulas and provide a sketch of their proof.

In [22] was considered, as a Fock space, a general vector space V over a field k with basis $|e_n\rangle$ $n = 0, 1, \dots$, equipped with its natural grading

$$(3) \quad V = \bigoplus_{n \in \mathbb{Z}} V_n \text{ with } V_n := ke_n; \quad V_{-n-1} := \{0\} \text{ for } n \geq 0$$

and scalar product defined by $\langle e_n | e_m \rangle = \delta_{n,m}$. Let f, g be two linear operators on V of degrees $-1, +1$, respectively. Generically, they read

$$(4) \quad \text{for } n \geq 0; \quad f|e_0\rangle := 0; \quad f|e_{k+1}\rangle = \alpha_{k+1}|e_k\rangle; \quad g|e_k\rangle := \beta_{k+1}|e_{k+1}\rangle$$

We consider the words in f, g :

$$(5) \quad w(f, g) := f^{p_1}g^{q_1}f^{p_2}g^{q_2} \cdots f^{p_n}g^{q_n}$$

the degree (excess) $\pi_e(w)$ of which is $\sum_{k=1}^n (q_k - p_k)$ (this is, for algebraists, the degree of the graded operator $f^{p_1}g^{q_1}f^{p_2}g^{q_2} \cdots f^{p_n}g^{q_n}$). This provides a representation μ of a two-letter free monoid $\{b_-, b_+\}^*$ on V by $\mu(b_-) = f$; $\mu(b_+) = g$, which is graded for the weight on $\{b_-, b_+\}^*$. In order to keep the reading of a word from left to right, one performs the action on the right. Thus V becomes a $\{b_-, b_+\}^*$ right module by

$$(6) \quad e_0.b_- = 0; \quad e_{n+1}.b_- = \alpha_{n+1}e_n; \quad e_n.b_+ = \beta_{n+1}e_{n+1}, \quad n \geq 0$$

one is interested by the matrix elements².

$$(7) \quad \langle e_n. \{b_-, b_+\}^i | e_m \rangle = \omega_{n \rightarrow m}^{(i)}.$$

Define $W_{n \rightarrow m}^{(i)}$ as follows

$$(8) \quad W_{n \rightarrow m}^{(i)} = \{w \in \{b_-, b_+\}^i \mid (\pi_e(w) = m - n) \text{ and } (w = uv \implies \pi_e(u) \geq -n)\}.$$

The following proposition characterizes $W_{n \rightarrow m}^{(i)}$ as an universal transporter between level n and level m .

PROPOSITION 7.1. *i) If all the weights $\alpha_n; n \geq 1, \beta_n, n \geq 0$ are nonzero, $W_{n \rightarrow m}^{(i)}$ is exactly the set of words of length i such that $\langle e_n.w | e_m \rangle \neq 0$.
ii) In all cases the latter is a subset of $W_{n \rightarrow m}^{(i)}$ i. e.*

$$(9) \quad \{w \in \{b_-, b_+\}^i \mid \langle e_n.w | e_m \rangle \neq 0\} \subset W_{n \rightarrow m}^{(i)}$$

Indeed, $W_{n \rightarrow m}^{(i)}$ admits factorizations as “noncommutative continued fractions”. In order to do this, we recall for the reader the definition of Dyck, positive Dyck and negative restricted (by depth) Dyck codes, respectively:

$$(10) \quad \begin{aligned} D &= \{w \in \pi_e^{-1}(0) \mid w = uv \text{ and } \varepsilon \notin \{u, v\} \implies \pi_e(u) \neq 0\} \\ D_+ &= \{w \in D \mid w = uv \text{ and } \varepsilon \notin \{u, v\} \implies \pi_e(u) > 0\} \\ D_-^{(n)} &= \{w \in D_- \mid \min_{uv=w}(\pi_e(u)) = -n\}. \end{aligned}$$

Now, identifying the subsets defined previously with their characteristic series in $\mathbb{N}\langle\langle b_-, b_+ \rangle\rangle$, one has the following factorizations (here, equalities between rational expressions of $b\pm, D_+$ and $D_-^{(n)}$).

²An interpretation of a similar coefficient in terms of paths, namely for the computation of

$$\int_{-\infty}^{+\infty} x^i p_n p_m w(x) dx$$

where $(p_n)_{n \in \mathbb{N}}$ is a family of orthogonal polynomials for the weight $w(x)$ can be found in [30].

PROPOSITION 7.2. Set $W_{n \rightarrow m} = \sum_{i \geq 0} W_{n \rightarrow m}^{(i)}$.

1) One has the following factorizations for $W_{n \rightarrow n+k}$

$$(D_+ + D_-^{(n)})^* (b_+ D_+^*)^k = \left[\prod_{i=0}^{\rightarrow (k-1)} ((D_-^{(n+i)})^* b_+) \right] (D_+ + D_-^{(n+k)})^* \text{ if } k \geq 0$$

$$((\mathbb{1}D)_+ b_-)^k (D_+ + D_-^{(n+k)})^* = (D_+ + D_-^{(n)})^* \left[\prod_{i=1}^{\rightarrow (-k)} ((b_- (D_-^{(n-i)})^*)) \right] \text{ otherwise}$$

2) One has the following self-reproducing equations for the Dyck codes

$$(12) \quad D_+ = b_+ (D_+)^* b_- ; \quad D_-^{(n)} = b_- (D_-^{(n-1)})^* b_+ ; \quad D_-^{(0)} = \emptyset$$

The statements of this proposition can be considered as a “noncommutative continued fraction” expansion of $W_{n \rightarrow m}$.

As a result [22], with

$$(13) \quad \mu(W_{m \rightarrow n}^{(i)})|n\rangle = \omega_{n \rightarrow m}^{(i)}|m\rangle \text{ and } T_{n \rightarrow n+k} := \sum_{i \geq 0} t^i \omega_{n \rightarrow n+k}^{(i)},$$

using proposition (7.2) and the representations $\mu_t(b\pm) = tb\pm$ (observe that $\mu_1 = \mu$), one can expand $T_{n \rightarrow n+k}$ as a product of continued fractions. Let

$$F_n^+ = \cfrac{1}{1 - \cfrac{t^2 \alpha_{n+1} \beta_{n+1}}{1 - \cfrac{t^2 \alpha_{n+2} \beta_{n+2}}{1 - \cfrac{t^2 \alpha_{n+3} \beta_{n+3}}{1 - \dots}}}} = \cfrac{1}{1 - E_n^+}$$

$$(14) \quad F_n^- = \cfrac{1}{1 - \cfrac{t^2 \alpha_n \beta_n}{1 - \cfrac{t^2 \alpha_{n-1} \beta_{n-1}}{1 - \cfrac{t^2 \alpha_{n-2} \beta_{n-2}}{1 - \dots}}}} = \cfrac{1}{1 - E_n^-}$$

and

$$(15) \quad F_n = \cfrac{1}{1 - E_n^+ - E_n^-}.$$

Then, if $k \geq 0$, we obtain

$$(16) \quad T_{n \rightarrow n+k} = t^k F_{n+k} \prod_{i=0}^{k-1} F_{n+i}^- = t^k F_n \prod_{i=1}^k F_{n+i}^+$$

and, if $k \leq 0$

$$(17) \quad T_{n \rightarrow n+k} = t^{-k} F_{n+k} \prod_{i=0}^{-k-1} F_{n-i}^+ = t^{-k} F_n \prod_{i=1}^{-k} F_{n-i}^-$$

8. Conclusion

The dualization problem solved by Sweedler's duals has striking relations with language theory (for this last point, see [2]). A true “engineer-like” calculus was developed in order to handle the rational closure mentioned above. This set of formulas is mainly based on a recursion to compute the “star of a matrix” (the

formulas and a complete discussion can be found in [11] and are reminiscent of general formulas giving the inverse of a matrix decomposed in blocks). This calculus is powerful enough to be the main ingredient in investigating rationality properties within various domains (see [9, 10] for noncommutative geometry, [5] for functions on the free group and [2, 14] for automata theory) and sufficiently expressive to give exact developments of some transfer coefficients in Combinatorial Physics [22]. These rational expressions are generic in the sense that any Sweedler's dual can be described by them. In this way, as for Dirac's notation, we can hope to inherit the computational skill developed through fourty years of practice.

References

- [1] E. Abe, *Hopf algebras*. Cambridge Univ. Press, 1980.
- [2] J. Berstel, C. Reutenauer, *Rational series and their languages*. EATCS Monographs on Theoretical Computer Science, Springer, 1988.
- [3] N. Bourbaki, *Algebra, chapter III*, Springer (1970)
- [4] N. Bourbaki, *Théories spectrales*, Hermann (1967)
- [5] G. Cauchon, *Séries de Malcev-Neumann sur le groupe libre et questions de rationalité*, Theoret. Comp. Sci. **98** (1992) 79-97.
- [6] A. Cayley, *On certain results related to quaternions*, Phil. Mag. **26** (1845) 141-145.
- [7] J.-M. Champarnaud, G. Duchamp, *Derivatives of rational expressions and related theorems*, Theoret. Comp. Sci. **313** (2004) 31.
- [8] V. Chari, A. Pressley, *A guide to quantum groups*. Cambridge Univ. Press, 1994.
- [9] A. Connes, *Noncommutative geometry*. Acad. Press, 1994.
- [10] G. Duchamp, C. Reutenauer, *Un critère de rationalité provenant de la géométrie non-commutative*, Inventiones Mathematicae, **128** (1997) 613-622.
- [11] G. Duchamp, H. H. Kacem, É. Laugerotte, *Algebraic elimination of ϵ -transitions*, DMTCS, **7** (2005) 51-70.
- [12] G. Duchamp, M. Flouret, É. Laugerotte, J.-G. Luque, *Direct and dual laws for automata with multiplicities*, Theoret. Comp. Sci. **267** (2001) 105-120.
- [13] G. H. E. Duchamp, P. Blasiak, A. Horzela, K. A. Penson, A. I. Solomon, *Hopf Algebras in General and in Combinatorial Physics: a practical introduction*, arXiv : 0802.0249
- [14] S. Eilenberg, *Automata, languages and machines*. Acad. Press, New-York, 1974.
- [15] M. Fliess, Sur le plongement de l'algèbre des séries rationnelles non commutatives dans un corps gauche. *CRAS Ser. A* **271** (1970) 926-927.
- [16] M. Fliess, Matrices de Hankel. *Jour. of Pure and Appl. Math.* **53** (1994) 197-222.
- [17] Golan J. S., *Power Algebras over Semirings with Applications in Mathematics and Computer science*. Kluwer Academic Publishers, 1999.
- [18] Golan J. S., *Semirings and Affine Equations over Them: Theory and Applications*. Kluwer Academic Publishers, 2003.
- [19] A. Heyting, Die Theorie der linearen Gleichungen in einer Zahlenspezies mit nichtkommutativer Multiplikation. *Math. Ann.* **98** (1927) 465-490.
- [20] G. P. Hochschild, Basic theory of algebraic groups and Lie algebras, Springer 1981,
- [21] G. Jacob, *Représentations et substitutions matricielles dans la théorie matricielle des semigroupes*, Thèse, Univ. de Paris (1975).
- [22] J. Katriel, G. Duchamp, *Ordering relations for q -boson operators, continued fractions techniques, and the q -CBH enigma*, J. Phys. A: Math. Gen. **28** (1995) 7209-7225.
- [23] S. K. Lando, *Lectures on generating functions*, A. M. S. (2003).
- [24] J. Lewin, *Fields of fractions for group algebras of free groups*, Trans. Amer. Math. Soc. **192** (1974) 339-346.
- [25] D.S. Passman, *The algebraic structure of group rings*, John Wiley - Interscience, (1977).
- [26] A.R. Richardson, *Simultaneous linear equations over a division ring*, Proc. Lond. Math. Soc. **28** (1928) 395-420.
- [27] M.P. Schützenberger, *On the definition of a family of automata*, Information and Control, **4** (1961) 275-270.

- [28] M.P. Schützenberger, *On a theorem of R. Jungen*, Proc. Amer. Math. Soc. **13** (1962) 885-889.
- [29] M.E. Sweedler, *Hopf algebras*, W.A. Benjamin, New York, 1969.
- [30] X.G. Viennot, *Une théorie combinatoire des polynômes orthogonaux*, Lect. Notes LACIM UQAM, Montreal (1984).
http://web.mac.com/xgviennot/iWeb/Xavier_Viennot

LIPN - UMR 7030 CNRS - UNIVERSITÉ PARIS 13 F-93430 VILLETTANEUSE, FRANCE

This page intentionally left blank

Primitive elements of the Hopf algebra of free quasi-symmetric functions

L. Foissy

ABSTRACT. Using the dendriform and bidendriform Cartier-Quillen-Milnor-Moore theorem, we construct a basis of the space of primitive elements of the Hopf algebra of free quasi-symmetric functions, indexed by a certain set of trees, and inductively computable.

Introduction

The Hopf algebra of free quasi-symmetric functions **FQSym**, also known as the Malvenuto-Reutenauer Hopf algebra, is introduced in [7]. It is a graded self-dual Hopf algebra, with the set of all permutations as a basis. Certain interesting properties are shown in [2]: in particular, it is shown that it is both free and cofree, and a basis of the space of its primitive elements is given, using the self-duality and a monomial basis. Note that computing the primitive elements of degree n by this method implies to inverse a certain $n! \times n!$ matrix.

The aim of this paper is to describe another basis of $Prim_{coAss}(\mathbf{FQSym})$, which can be inductively computed. We use for this the dendriform structure of **FQSym**. Recall that a dendriform algebra is an associative algebra such that its product can be split into two non-associative products \prec and \succ , with good compatibilities [5, 6, 8]. It is known that **FQSym**, or more precisely its augmentation ideal, is dendriform. More precisely, it is a dendriform Hopf algebra, in the sense of [8]. This implies, by the dendriform Cartier-Quillen-Milnor-Moore theorem, that $Prim_{coAss}(\mathbf{FQSym})$ is a brace algebra.

We introduced in [4] the notion of bidendriform bialgebra and showed that **FQSym** is bidendriform. The bidendriform Cartier-Quillen-Milnor-Moore theorem implies that **FQSym** is freely generated, as a dendriform algebra, by the space $Prim_{coDend}(\mathbf{FQSym})$ of primitive elements in the codendriform sense. Combining this result with the dendriform Cartier-Quillen-Milnor-Moore theorem, we show that $Prim_{coAss}(\mathbf{FQSym})$ is freely generated by $Prim_{coDend}(\mathbf{FQSym})$ (as a brace algebra). We recall in section 2 a description of free brace algebras. If $(v_i)_{i \in I}$ is a basis of the vector space V , then the free brace algebra generated by V has a basis indexed by planar rooted trees decorated by I , and the brace structure is described in this basis by the help of graftings. Hence, for any basis of $Prim_{coDend}(\mathbf{FQSym})$,

it is possible to recover a basis of $Prim_{coAss}(\mathbf{FQSym})$, indexed by a certain set of planar rooted trees.

Let, for all $n \in \mathbb{N}$:

$$\begin{cases} p_n &= \dim(Prim_{coAss}(\mathbf{FQSym})_n), \\ q_n &= \dim(Prim_{coDend}(\mathbf{FQSym})_n). \end{cases}$$

We then prove in section 3 that for $n \geq 2$, $q_n = (n-2)p_{n-1}$. We then give $n-2$ applications from $Prim_{coAss}(\mathbf{FQSym})_{n-1}$ to $Prim_{coDend}(\mathbf{FQSym})_n$, which give all elements of $Prim_{coDend}(\mathbf{FQSym})_n$. These applications are given by the insertion of $n+1$ at a given place in elements of the symmetric group S_n , seen as words in letters $1, \dots, n$.

Combining the results of the second and third sections, we define inductively in the fourth section a new basis of $Prim_{coAss}(\mathbf{FQSym})_n$, indexed by certain planar decorated rooted trees. The trees which are only a root give a basis of $Prim_{coDend}(\mathbf{FQSym})_n$.

Notations.

- (1) K is a commutative field of any characteristic.
- (2) If V is a K -vector field which is \mathbb{N} -graded, we shall denote by V_k the space of homogeneous elements of V of degree k .

1. Bidendriform bialgebras and \mathbf{FQSym}

1.1. Bidendriform bialgebras. We introduced in [4] the following definition:

DEFINITION 1. A bidendriform bialgebra is a family $(A, \prec, \succ, \Delta_\prec, \Delta_\succ)$ such that:

- (1) A is a K -vector space and:

$$\begin{array}{c|c} \prec : \begin{cases} A \otimes A \longrightarrow A \\ a \otimes b \longrightarrow a \prec b, \end{cases} & \Delta_\prec : \begin{cases} A \longrightarrow A \otimes A \\ a \longrightarrow \Delta_\prec(a) = a'_\prec \otimes a''_\prec, \end{cases} \\ \succ : \begin{cases} A \otimes A \longrightarrow A \\ a \otimes b \longrightarrow a \succ b, \end{cases} & \Delta_\succ : \begin{cases} A \longrightarrow A \otimes A \\ a \longrightarrow \Delta_\succ(a) = a'_\succ \otimes a''_\succ. \end{cases} \end{array}$$

- (2) (Dendriform axioms). (A, \prec, \succ) is a dendriform algebra: for all $a, b, c \in A$,

$$(1) \quad (a \prec b) \prec c = a \prec (b \prec c + b \succ c),$$

$$(2) \quad (a \succ b) \prec c = a \succ (b \prec c),$$

$$(3) \quad 5a \prec b + a \succ b \succ c = a \succ (b \succ c).$$

- (3) (Codendriform axioms). $(A, \Delta_\prec, \Delta_\succ)$ is a codendriform coalgebra: for all $a \in A$,

$$(4) \quad (\Delta_\prec \otimes Id) \circ \Delta_\prec(a) = (Id \otimes \Delta_\prec + Id \otimes \Delta_\succ) \circ \Delta_\prec(a),$$

$$(5) \quad (\Delta_\prec \otimes Id) \circ \Delta_\prec(a) = (Id \otimes \Delta_\prec) \circ \Delta_\succ(a),$$

$$(6) \quad (\Delta_\prec \otimes Id + \Delta_\succ \otimes Id) \circ \Delta_\succ(a) = (Id \otimes \Delta_\succ) \circ \Delta_\succ(a).$$

(4) (Bidendriform axioms). For all $a, b \in A$,

$$\begin{aligned} (\Delta_\succ)(a \succ b) &= a'b'_\succ \otimes a'' \succ b''_\succ + a' \otimes a'' \succ b + b'_\succ \otimes a \succ b''_\succ + ab'_\succ \otimes b''_\succ + a \otimes b, \\ (\Delta_\prec)(a \prec b) &= a'b'_\prec \otimes a'' \prec b''_\prec + a' \otimes a'' \prec b + b'_\prec \otimes a \prec b''_\prec, \\ (\Delta_\prec)(a \succ b) &= a'b'_\prec \otimes a'' \succ b''_\prec + ab'_\prec \otimes b''_\prec + b'_\prec \otimes a \succ b''_\prec, \\ (\Delta\Omega)(a \prec b) &= a'b'_\prec \otimes a'' \prec b''_\prec + a'b \otimes a'' + b'_\prec \otimes a \prec b''_\prec + b \otimes a. \end{aligned}$$

Remarks.

- (1) If A is a bidendriform bialgebra, then $K \oplus A$ is naturally a Hopf algebra, extending $\prec + \succ$ and $\Delta_\prec + \Delta_\succ$ on $K \oplus A$.
- (2) If A is a bidendriform bialgebra, it is also a dendriform hopf algebra in the sense of [8, 9], with coassociative coproduct given by $\tilde{\Delta} = \Delta_\prec + \Delta_\succ$. The compatibilities of dendriform Hopf algebras are given by (7) + (9) and (8) + (10).

If A is a bidendriform algebra, we define:

$$\text{Prim}_{\text{codend}}(A) = \text{Ker}(\Delta_\prec) \cap \text{Ker}(\Delta_\succ).$$

The following result is proved in [4] (theorem 35 and corollary 17):

THEOREM 2 (Bidendriform Cartier-Quillen-Milnor-Moore theorem). *Let A be a \mathbb{N} -graded bidendriform bialgebra, such that $A_0 = (0)$. Then A is freely generated as a dendriform algebra by $\text{Prim}_{\text{coDend}}(A)$. Moreover, consider the following formal series:*

$$R(X) = \sum_{n=1}^{+\infty} \dim(A), \quad Q(X) = \sum_{n=1}^{+\infty} \dim(\text{Prim}_{\text{coDend}}(A_n)) X^n.$$

$$\text{Then } Q(X) = \frac{R(X)}{(R(X) + 1)^2}.$$

1.2. An example: the Hopf algebra **FQSym.** See [1, 2, 7]. The algebra **FQSym** is the vector space generated by the elements $(\mathbf{F}_u)_{u \in \mathbb{S}}$, where \mathbb{S} is the disjoint union of the symmetric groups S_n ($n \in \mathbb{N}$). Its product and its coproduct are given in the following way: for all $u \in S_n$, $v \in S_m$, putting $u = (u_1 \dots u_n)$,

$$\begin{aligned} \Delta(\mathbf{F}_u) &= \sum_{i=0}^n \mathbf{F}_{st(u_1 \dots u_i)} \otimes \mathbf{F}_{st(u_{i+1} \dots u_n)}, \\ \mathbf{F}_u \cdot \mathbf{F}_v &= \sum_{\zeta \in sh(n, m)} \mathbf{F}_{(u \times v) \cdot \zeta^{-1}}, \end{aligned}$$

where $sh(n, m)$ is the set of (n, m) -shuffles, and st is the standardisation. Its unit is $1 = \mathbf{F}_\emptyset$, where \emptyset is the unique element of S_0 . Moreover, **FQSym** is a \mathbb{N} -graded Hopf algebra, by putting $|\mathbf{F}_u| = n$ if $u \in S_n$.

Examples.

$$\begin{aligned} \mathbf{F}_{(1 2)} \mathbf{F}_{(1 2 3)} &= \mathbf{F}_{(1 2 3 4 5)} + \mathbf{F}_{(1 3 2 4 5)} + \mathbf{F}_{(1 3 4 2 5)} + \mathbf{F}_{(1 3 4 5 2)} + \mathbf{F}_{(3 1 2 4 5)} \\ &\quad + \mathbf{F}_{(3 1 4 2 5)} + \mathbf{F}_{(3 1 4 5 2)} + \mathbf{F}_{(3 4 1 2 5)} + \mathbf{F}_{(3 4 1 5 2)} + \mathbf{F}_{(3 4 5 1 2)}; \end{aligned}$$

$$\begin{aligned} \Delta(\mathbf{F}_{(1 2 5 4 3)}) &= 1 \otimes \mathbf{F}_{(1 2 5 4 3)} + \mathbf{F}_{(1)} \otimes \mathbf{F}_{(1 4 3 2)} + \mathbf{F}_{(1 2)} \otimes \mathbf{F}_{(3 2 1)} \\ &\quad + \mathbf{F}_{(1 2 3)} \otimes \mathbf{F}_{(2 1)} + \mathbf{F}_{(1 2 4 3)} \otimes \mathbf{F}_{(1)} + \mathbf{F}_{(1 2 5 4 3)} \otimes 1. \end{aligned}$$

Let $(\mathbf{FQSym})_+ = Vect(\mathbf{F}_u / u \in S_n, n \geq 1)$ be the augmentation ideal of \mathbf{FQSym} . We define $\prec, \succ, \Delta_\prec$ and Δ_\succ on $(\mathbf{FQSym})_+$ in the following way: for all $u \in S_n, v \in S_m$, by putting $u = (u_1 \dots u_n)$,

$$\begin{aligned} \mathbf{F}_u \prec \mathbf{F}_v &= \sum_{\substack{\zeta \in sh(n,m) \\ \zeta^{-1}(n+m)=n}} \mathbf{F}_{(u \times v) \cdot \zeta^{-1}}, \\ \mathbf{F}_u \succ \mathbf{F}_v &= \sum_{\substack{\zeta \in sh(n,m) \\ \zeta^{-1}(n+m)=n+m}} \mathbf{F}_{(u \times v) \cdot \zeta^{-1}}, \\ \Delta_\prec(\mathbf{F}_u) &= \sum_{i=u^{-1}(n)}^{n-1} \mathbf{F}_{st(u_1 \dots u_i)} \otimes \mathbf{F}_{st(u_{i+1} \dots u_n)}, \\ \Delta_\succ(\mathbf{F}_u) &= \sum_{i=1}^{u^{-1}(n)-1} \mathbf{F}_{st(u_1 \dots u_i)} \otimes \mathbf{F}_{st(u_{i+1} \dots u_n)}. \end{aligned}$$

Examples.

$$\begin{aligned} \mathbf{F}_{(1 \ 2)} \prec \mathbf{F}_{(1 \ 2 \ 3)} &= \mathbf{F}_{(1 \ 3 \ 4 \ 5 \ 2)} + \mathbf{F}_{(3 \ 1 \ 4 \ 5 \ 2)} + \mathbf{F}_{(3 \ 4 \ 1 \ 5 \ 2)} + \mathbf{F}_{(3 \ 4 \ 5 \ 1 \ 2)}, \\ \mathbf{F}_{(1 \ 2)} \succ \mathbf{F}_{(1 \ 2 \ 3)} &= \mathbf{F}_{(1 \ 2 \ 3 \ 4 \ 5)} + \mathbf{F}_{(1 \ 3 \ 2 \ 4 \ 5)} + \mathbf{F}_{(1 \ 3 \ 4 \ 2 \ 5)} \\ &\quad + \mathbf{F}_{(3 \ 1 \ 2 \ 4 \ 5)} + \mathbf{F}_{(3 \ 1 \ 4 \ 2 \ 5)} + \mathbf{F}_{(3 \ 4 \ 1 \ 2 \ 5)}, \end{aligned}$$

$$\begin{aligned} \Delta_\prec(\mathbf{F}_{(1 \ 2 \ 5 \ 4 \ 3)}) &= \mathbf{F}_{(1 \ 2 \ 3)} \otimes \mathbf{F}_{(2 \ 1)} + \mathbf{F}_{(1 \ 2 \ 4 \ 3)} \otimes \mathbf{F}_{(1)}, \\ \Delta_\succ(\mathbf{F}_{(1 \ 2 \ 5 \ 4 \ 3)}) &= \mathbf{F}_{(1)} \otimes \mathbf{F}_{(1 \ 4 \ 3 \ 2)} + \mathbf{F}_{(1 \ 2)} \otimes \mathbf{F}_{(3 \ 2 \ 1)}. \end{aligned}$$

The following result is proved in [4], theorem 38:

THEOREM 3. $((\mathbf{FQSym})_+, \prec, \succ, \Delta_\prec, \Delta_\succ)$ is a connected bidendriform bialgebra.

Moreover, $(\mathbf{FQSym})_+$ is \mathbb{N} -graded, by putting the elements of S_n homogeneous of degree n . By theorem 2, with $q_n = \dim(Prim_{coDend}(\mathbf{FQSym})_n)$, we obtain:

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|---|---|---|---|----|-----|-------|--------|---------|-----------|------------|
| q_n | 1 | 0 | 1 | 6 | 39 | 284 | 2 305 | 20 682 | 203 651 | 2 186 744 | 25 463 925 |

2. Recovering $Prim_{coAss}(\mathbf{FQSym})$ from $Prim_{coDend}(\mathbf{FQSym})$

2.1. Dendriform Cartier-Quillen-Milnor-Moore theorem and variations. Recall that a brace algebra is a K -vector space A together with a n -multilinear operation for all $n \geq 2$:

$$\langle \dots \rangle : \begin{cases} A^{\otimes n} &\longrightarrow A \\ a_1 \otimes \dots \otimes a_n &\longrightarrow \langle a_1, \dots, a_n \rangle, \end{cases}$$

satisfying certain relations; see [8, 9] for more details. For example:

$$\langle a_1, \langle a_2, a_3 \rangle \rangle = \langle a_1, a_2, a_3 \rangle + \langle \langle a_1, a_2 \rangle, a_3 \rangle + \langle a_2, a_1, a_3 \rangle.$$

The following theorem is proved in [8, 9]; more precisely, the first point of this theorem is proposition 2-8 and theorem 3-4 of [8] and the second point is theorem 4-6 of [9]:

THEOREM 4 (Dendriform Cartier-Quillen-Milnor-Moore theorem). *Let A be a dendriform Hopf algebra. We denote $Prim_{coAss}(A) = Ker(\tilde{\Delta})$.*

(1) *$Prim_{coAss}(A)$ is a brace algebra, with brackets given by:*

$$\langle p_1, \dots, p_n \rangle =$$

$$\sum_{i=0}^{n-1} (-1)^{n-1-i} (p_1 \prec (p_2 \prec (\dots \prec p_i) \dots) \succ p_n \prec (\dots (p_{i+1} \succ p_{i+2}) \succ \dots) \succ p_{n-1}).$$

(2) *If A is freely generated as a dendriform algebra by a subvector space $V \subseteq Prim_{coAss}(A)$, then $Prim_{coAss}(A)$ is freely generated as a brace algebra by V .*

Let us precise the relation between $Prim_{coAss}(A)$ and $Prim_{coDend}(A)$ if A is a bidendriform bialgebra. Combining the dendriform and the bidendriform Cartier-Quillen-Milnor-Moore theorems:

THEOREM 5. *Let A be a \mathbb{N} -graded bidendriform bialgebra, with $A_0 = (0)$. Then $Prim_{coAss}(A)$ is, as a brace algebra, freely generated by $Prim_{coDend}(A)$.*

PROOF. By the bidendriform Cartier-Quillen-Milnor-Moore theorem, as a dendriform algebra A is freely generated by $Prim_{coDend}(A)$. By the second point of the dendriform Cartier-Quillen-Milnor-Moore theorem, $Prim_{coAss}(A)$ is freely generated as a brace algebra by the space $Prim_{coDend}(A)$. \square

PROPOSITION 6. *If A is \mathbb{N} -graded dendriform Hopf algebra, such that $A_0 = (0)$, then A is generated as a dendriform algebra by $Prim_{coAss}(A)$. Moreover, consider the following formal series:*

$$R(X) = \sum_{n=1}^{\infty} \dim(A_n) X^n, \quad P(X) = \sum_{n=1}^{+\infty} \dim(Prim_{coDend}((A)_n)) X^n.$$

$$\text{Then } P(X) = \frac{R(X)}{1 + R(X)}.$$

PROOF. *First step.* Let $p_1, \dots, p_n \in Prim_{coAss}(A)$. We define by induction on n :

$$\omega(p_1, \dots, p_n) = \begin{cases} p_1 & \text{if } n = 1, \\ p_n \prec \omega(p_1, \dots, p_{n-1}) & \text{if } n \geq 2. \end{cases}$$

An easy induction on n allows to show the following result, using (8)+(10):

$$\tilde{\Delta}(\omega(p_1, \dots, p_n)) = \sum_{i=1}^{n-1} \omega(p_1, \dots, p_i) \otimes \omega(p_{i+1}, \dots, p_n).$$

We denote by $\tilde{\Delta}^n : A \longrightarrow A^{n+1}$ the iterated coproducts of A . It comes by induction:

$$\tilde{\Delta}^m(\omega(p_1, \dots, p_n)) = \begin{cases} 0 & \text{if } m \geq n, \\ p_1 \otimes \dots \otimes p_n & \text{if } m = n-1. \end{cases}$$

Second step. We consider the tensor (non counitary) coalgebra:

$$C = \bigoplus_{n=1}^{\infty} Prim_{coAss}(A)^{\otimes n}.$$

It is a coalgebra for the deconcatenation coproduct. As $Prim_{coAss}(A)$ is \mathbb{N} -graded, C is a graded coalgebra with formal series:

$$S(X) = \frac{1}{1 - P(X)} - 1 = \frac{P(X)}{1 - P(X)}.$$

By the first step, the following application is a morphism of graded coalgebras:

$$\Psi : \begin{cases} C & \longrightarrow A \\ p_1 \otimes \dots \otimes p_n & \longrightarrow \omega(p_1, \dots, p_n). \end{cases}$$

Third step. Suppose that $Ker(\Psi)$ is non zero. As it is a coideal of C , it contains primitive elements of C , that is to say elements of $Prim_{coAss}(A)$. As Ψ is obviously monic on $Prim_{coAss}(A)$, this is impossible. So $Ker(\Psi) = (0)$ and Ψ is monic.

Let $a \in A$. As $A_0 = (0)$, for a certain $N(a) \in \mathbb{N}^*$, $\tilde{\Delta}^{N(a)}(a) = 0$. We prove that $a \in Im(\Psi)$ by induction on $N(a)$. If $N(a) = 1$, then $a \in Prim_{coAss}(A)$ and the result is obvious. Suppose that the result is true for all $b \in A$ such that $N(b) < N(a)$. As $\tilde{\Delta}^{N(a)}(a) = 0$, necessarily $\tilde{\Delta}^{N(a)-1}(a) \in Prim_{coAss}(A)^{\otimes N(a)}$. We put:

$$\tilde{\Delta}^{N(a)-1}(a) = a_1 \otimes \dots \otimes a_n, \quad b = a - \omega(a_1, \dots, a_n).$$

By the first step, $\tilde{\Delta}^{N(a)-1}(b) = 0$, so $N(b) < N(a)$. By the induction hypothesis, $b \in Im(\Psi)$. As $\omega(a_1, \dots, a_n) \in Im(\Psi)$, $a \in Im(\Psi)$.

Last step. As Ψ is an isomorphism of graded coalgebras, $S(X) = R(X)$. Hence:

$$R(X) = \frac{P(X)}{1 - P(X)},$$

$$\text{so } R(X) - R(X)P(X) = P(X) \text{ and } P(X) = \frac{R(X)}{1 + R(X)}. \quad \square$$

2.2. Free brace algebras. Using a description of the free dendriform algebra generated by a set \mathcal{D} with planar decorated forests, we gave a description of the free brace algebra $Brace(\mathcal{D})$ in [3]. A basis of this brace algebra is given by the set $T^{\mathcal{D}}$ of planar rooted trees decorated by \mathcal{D} . For example:

$$\begin{aligned} Brace(\mathcal{D})_1 &= Vect(\bullet_a, a \in \mathcal{D}), \\ Brace(\mathcal{D})_2 &= Vect(\mathfrak{l}_a^b, a, b \in \mathcal{D}), \\ Brace(\mathcal{D})_3 &= Vect(\overset{c}{\bullet} \overset{b}{V}_a^b, \mathfrak{l}_a^b, a, b, c \in \mathcal{D}), \\ Brace(\mathcal{D})_4 &= Vect(\overset{d}{\bullet} \overset{c}{\bullet} \overset{b}{V}_a^b, \overset{d}{\bullet} \overset{c}{V}_a^b, \overset{d}{\bullet} \overset{c}{V}_a^b, \overset{d}{\bullet} \overset{c}{V}_a^b, \mathfrak{l}_a^b, a, b, c, d \in \mathcal{D}), \dots \end{aligned}$$

The brace bracket satisfies, for all $t_1, \dots, t_{n-1} \in T^{\mathcal{D}}$, $d \in \mathcal{D}$:

$$\langle t_1, \dots, t_{n-1}, \bullet_d \rangle = B_d(t_{n-1} \dots t_1),$$

where $B_d(t_{n-1} \dots t_1)$ is the tree obtained by grafting the trees t_{n-1}, \dots, t_1 (in this order) on a common root decorated by d . For example, if $a, b, c, d \in \mathcal{D}$,

$$\langle \bullet_a, \mathfrak{l}_b^c, \bullet_d \rangle = \overset{c}{\bullet} \overset{b}{V}_d^a.$$

As a consequence, if A is a connected bidendriform bialgebra and if $(q_d)_{d \in \mathcal{D}}$ is a basis of $Prim_{coDend}(A)$, then a basis of $Prim_{coAss}(A)$ is given by $(p_t)_{t \in T^{\mathcal{D}}}$ defined inductively by:

$$\begin{cases} p_{\bullet d} &= q_d, \\ p_{B_d^+(t_1 \dots t_n)} &= \langle p_{t_n}, \dots, p_{t_1}, q_d \rangle. \end{cases}$$

3. Recovering $Prim_{coDend}(\mathbf{FQSym})$ from $Prim_{coAss}(\mathbf{FQSym})$

For all $n \in \mathbb{N}^*$, we put:

$$\begin{cases} p_n &= \dim(Prim_{coAss}(\mathbf{FQSym})_n), \\ q_n &= \dim(Prim_{coDend}(\mathbf{FQSym})_n). \end{cases}$$

PROPOSITION 7. For all $n \geq 2$, $q_n = (n-2)p_{n-1}$.

PROOF. We put:

$$R(X) = \sum_{n=1}^{\infty} n!X^n, \quad P(X) = \sum_{n=1}^{\infty} p_n X^n, \quad Q(X) = \sum_{n=1}^{\infty} q_n X^n.$$

By theorem 2 and proposition 6:

$$P(X) = \frac{R(X)}{1+R(X)}, \quad Q(X) = \frac{R(X)}{(1+R(X))^2}.$$

Hence:

$$P'(X) = \frac{R'(X)}{(1+R(X))^2}.$$

Moreover:

$$\begin{aligned} R'(X) &= \sum_{n=1}^{\infty} nn!X^{n-1} \\ &= \sum_{n=1}^{\infty} (n+1)!X^{n-1} - \sum_{n=1}^{\infty} n!X^{n-1} \\ &= \frac{R(X) - X}{X^2} - \frac{R(X)}{X} \\ &= \frac{R(X) - X(1+R(X))}{X^2}. \end{aligned}$$

We deduce:

$$X^2 P'(X) = \frac{R(X) - X(1+R(X))}{(1+R(X))^2} = Q(X) - \frac{X}{1+R(X)} = Q(X) - X + XP(X).$$

So:

$$X^2 P'(X) + XP(X) = \sum_{n=1}^{\infty} (n-1)p_n X^{n+1} = Q(X) - X = \sum_{n=2}^{\infty} q_n X^n.$$

As a conclusion, for all $n \geq 2$, $q_n = (n-2)p_{n-2}$. \square

DEFINITION 8. Let $i \in \mathbb{N}^*$. We define $\Phi_i : \mathbf{FQSym} \rightarrow \mathbf{FQSym}$ in the following way: for all $n \in \mathbb{N}$, for all $\sigma = (\sigma_1, \dots, \sigma_n) \in S_n$,

$$\Phi_i(\mathbf{F}_{\sigma}) = \begin{cases} 0 & \text{if } i \geq n, \\ \mathbf{F}_{(\sigma_1, \dots, \sigma_i, n+1, \sigma_{i+1}, \dots, \sigma_n)} & \text{if } i < n. \end{cases}$$

THEOREM 9. *Let $n \geq 2$. The following application is bijective:*

$$\Phi : \begin{cases} (Prim_{coAss}(\mathbf{FQSym})_{n-1})^{n-2} & \longrightarrow \ Prim_{coDend}(\mathbf{FQSym})_n \\ (p_1, \dots, p_{n-2}) & \longrightarrow \ \Phi_1(p_1) + \dots + \Phi_{n-2}(p_{n-2}). \end{cases}$$

PROOF. *First step.* We prove that Φ takes values in $Prim_{coDend}(\mathbf{FQSym})$. Let $p \in Prim_{coAss}(\mathbf{FQSym})$ and $1 \leq i \leq n-2$. For all $k \in \mathbb{N}$, let π_k be the projection on \mathbf{FQSym}_k . By definition of Δ_{\prec} and Δ_{\succ} , for all $\sigma \in S_{n-1}$:

$$\begin{aligned} \Delta_{\prec}(\Phi_i(\mathbf{F}_{\sigma})) &= \left(\sum_{j=i+1}^{n-2} \pi_j \otimes \pi_{n-1-j} \right) \circ \tilde{\Delta}(\mathbf{F}_{\sigma}), \\ \Delta_{\succ}(\Phi_i(\mathbf{F}_{\sigma})) &= \left(\sum_{j=1}^i \pi_j \otimes \pi_{n-1-j} \right) \circ \tilde{\Delta}(\mathbf{F}_{\sigma}). \end{aligned}$$

By linearity, we obtain:

$$\begin{aligned} \Delta_{\prec}(p) &= \left(\sum_{j=i+1}^{n-2} \pi_j \otimes \pi_{n-1-j} \right) \circ \tilde{\Delta}(p) = 0, \\ \Delta_{\succ}(p) &= \left(\sum_{j=1}^i \pi_j \otimes \pi_{n-1-j} \right) \circ \tilde{\Delta}(p) = 0. \end{aligned}$$

This proves the first step.

Second step. We now prove that Φ is monic. Let $(p_1, \dots, p_{n-2}) \in Ker(\Phi)$. Let be $1 \leq i \leq n-2$. We define:

$$\varpi_i : \begin{cases} \mathbf{FQSym}_n & \longrightarrow \ \mathbf{FQSym}_n \\ \mathbf{F}_{\sigma} & \longrightarrow \ \begin{cases} 0 & \text{if } \sigma^{-1}(n) \neq i+1, \\ \mathbf{F}_{\sigma} & \text{if } \sigma^{-1}(n) = i+1. \end{cases} \end{cases}$$

Then, in an obvious way, $\varpi_i(\Phi(p_1, \dots, p_{n-2})) = \Phi_i(p_i) = 0$. As Φ_i is obviously monic on \mathbf{FQSym}_{n-1} (because $i \leq n-2$), $p_i = 0$. So Φ is monic.

Last step. Since $dim \left((Prim_{coAss}(\mathbf{FQSym})_{n-1})^{n-2} \right)$ coincides by proposition 7 with $dim(Prim_{coDend}(\mathbf{FQSym})_n)$, Φ is bijective. \square

4. An inductive basis of $Prim_{coAss}(\mathbf{FQSym})$

We now combine results of the second and third sections to obtain an basis of the space $Prim_{coAss}(\mathbf{FQSym})$. We first define inductively some set of partially planar decorated trees $\mathbb{T}(n)$ in the following way:

- (1) $\mathbb{T}(0)$ is the set of non decorated planar trees. The weight of an element of $\mathbb{T}(0)$ is the number of its vertices.
- (2) Suppose that $\mathbb{T}(n)$ is defined. Then $\mathbb{T}(n+1)$ is the set of planar trees defined by :
 - (a) The elements of $\mathbb{T}(n+1)$ are partially decorated planar trees.
 - (b) The vertices of the elements of $\mathbb{T}(n+1)$ can eventually be decorated by a pair (t, k) , with $t \in \mathbb{T}(n)$ and k an integer in $\{1, \dots, weight(t) - 1\}$.

- (c) The weight of an element of $\mathbb{T}(n)$ is the sum of the number of its vertices and of the weights of the trees of $\mathbb{T}(n)$ that appear in its decorations.

Inductively, for all $n \in \mathbb{N}$, $\mathbb{T}(n) \subseteq \mathbb{T}(n+1)$. We put $\mathbb{T} = \bigcup_{n \in \mathbb{N}} \mathbb{T}(n)$.

Examples.

- (1) Elements of \mathbb{T} of weight 1: \dots .
- (2) Elements of \mathbb{T} of weight 2: \mathbb{I} .
- (3) Elements of \mathbb{T} of weight 3: $\mathbb{V}, \mathbb{I}, \bullet_{(\mathbb{I},1)}$.
- (4) Elements of \mathbb{T} of weight 4:
 - (a) $\mathbb{V}, \mathbb{V}, \mathbb{V}, \mathbb{V}, \mathbb{I}$,
 - (b) $\bullet_{(\mathbb{V},1)}, \bullet_{(\mathbb{V},2)}, \bullet_{(\mathbb{I},1)}, \bullet_{(\mathbb{I},2)}, \bullet_{(\bullet_{(\mathbb{I},1)},1)}, \bullet_{(\bullet_{(\mathbb{I},1)},2)}$,
 - (c) $\mathbb{I}^{(\mathbb{I},1)}, \mathbb{I}_{(\mathbb{I},1)}$.

We can then define a basis $(p_t)_{t \in \mathbb{T}}$ of $\text{Prim}_{\text{coAss}}(\mathbf{FQSym})$ inductively in the following way:

- (1) $p_{\bullet} = \mathbf{F}_{(1)}$.
- (2) If $t = \bullet_{(t',i)}$, then $p_t = \Phi_i(p_{t'})$.
- (3) If t is not a single root, let t_1, \dots, t_{n-1} be the children of its roots, from left to right, and t_n its root. Then $p_t = \langle p_{t_{n-1}}, \dots, p_{t_1}, p_{t_n} \rangle$.

By the preceding results:

THEOREM 10. $(p_t)_{t \in \mathbb{T}}$ is a basis of $\text{Prim}_{\text{coAss}}(\mathbf{FQSym})$. Moreover, a basis of $\text{Prim}_{\text{coDend}}(\mathbf{FQSym})$ is given by the p_t 's, where t is a single root.

Examples.

- (1) $p_{\bullet} = \mathbf{F}_{(1)}$.
- (2) $p_{\mathbb{I}} = -\mathbf{F}_{(21)} + \mathbf{F}_{(12)}$.
- (3) (a) $p_{\bullet_{(\mathbb{I},1)}} = -\mathbf{F}_{(231)} + \mathbf{F}_{(132)}$.
 (b) $p_{\mathbb{V}} = \mathbf{F}_{(231)} - \mathbf{F}_{(132)} - \mathbf{F}_{(312)} + \mathbf{F}_{(213)}$.
 (c) $p_{\mathbb{I}} = \mathbf{F}_{(321)} - \mathbf{F}_{(231)} - \mathbf{F}_{(213)} + \mathbf{F}_{(123)}$.
- (4) (a) $p_{\bullet_{(\bullet_{(\mathbb{I},1)},1)}} = -\mathbf{F}_{(2431)} + \mathbf{F}_{(1432)}$.
 (b) $p_{\bullet_{(\bullet_{(\mathbb{I},1)},2)}} = -\mathbf{F}_{(2341)} + \mathbf{F}_{(1342)}$.
 (c) $p_{\bullet_{(\mathbb{V},1)}} = \mathbf{F}_{(2431)} - \mathbf{F}_{(1432)} - \mathbf{F}_{(3412)} + \mathbf{F}_{(2413)}$.
 (d) $p_{\bullet_{(\mathbb{V},2)}} = \mathbf{F}_{(2341)} - \mathbf{F}_{(1342)} - \mathbf{F}_{(3142)} + \mathbf{F}_{(2143)}$.
 (e) $p_{\bullet_{(\mathbb{I},1)}} = \mathbf{F}_{(3421)} - \mathbf{F}_{(2431)} - \mathbf{F}_{(2413)} + \mathbf{F}_{(1423)}$.
 (f) $p_{\bullet_{(\mathbb{I},2)}} = \mathbf{F}_{(3241)} - \mathbf{F}_{(2341)} - \mathbf{F}_{(2143)} + \mathbf{F}_{(1243)}$.
 (g) $p_{\mathbb{V}} = -\mathbf{F}_{(2341)} + \mathbf{F}_{(1342)} + \mathbf{F}_{(3142)} + \mathbf{F}_{(3412)} - \mathbf{F}_{(2143)} - \mathbf{F}_{(2413)} - \mathbf{F}_{(4213)} + \mathbf{F}_{(3214)}$.

- (h) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2 \end{smallmatrix} = -\mathbf{F}_{(2431)} - \mathbf{F}_{(4231)} + \mathbf{F}_{(2341)} + \mathbf{F}_{(3241)} + \mathbf{F}_{(1432)} + \mathbf{F}_{(4132)} + \mathbf{F}_{(4312)} - \mathbf{F}_{(1342)} - \mathbf{F}_{(3142)} - \mathbf{F}_{(3412)} - \mathbf{F}_{(3214)} + \mathbf{F}_{(2314)}.$
- (i) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2 \end{smallmatrix} = -\mathbf{F}_{(3241)} + \mathbf{F}_{(2341)} + \mathbf{F}_{(2143)} + \mathbf{F}_{(2413)} + \mathbf{F}_{(4213)} - \mathbf{F}_{(1243)} - \mathbf{F}_{(1423)} - \mathbf{F}_{(4123)} - \mathbf{F}_{(2314)} - \mathbf{F}_{(3214)} + \mathbf{F}_{(1324)} + \mathbf{F}_{(3124)}.$
- (j) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2 \end{smallmatrix} = -\mathbf{F}_{(3421)} + \mathbf{F}_{(2431)} + \mathbf{F}_{(4231)} - \mathbf{F}_{(3241)} + \mathbf{F}_{(2314)} - \mathbf{F}_{(1324)} - \mathbf{F}_{(3124)} + \mathbf{F}_{(2134)}.$
- (k) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2 \end{smallmatrix} = -\mathbf{F}_{(4321)} + \mathbf{F}_{(3421)} + \mathbf{F}_{(3241)} - \mathbf{F}_{(2341)} + \mathbf{F}_{(3214)} - \mathbf{F}_{(2314)} - \mathbf{F}_{(2134)} + \mathbf{F}_{(1234)}.$
- (l) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2,1 \end{smallmatrix} = \mathbf{F}_{(2341)} + \mathbf{F}_{(2431)} + \mathbf{F}_{(4231)} - 2\mathbf{F}_{(1342)} - \mathbf{F}_{(1432)} - \mathbf{F}_{(4132)} - \mathbf{F}_{(3142)} - \mathbf{F}_{(3412)} + \mathbf{F}_{(1243)} + \mathbf{F}_{(2143)} + \mathbf{F}_{(2413)}.$
- (m) $p \begin{smallmatrix} 1 \\ \downarrow \\ 2,1 \end{smallmatrix} = \mathbf{F}_{(3421)} - \mathbf{F}_{(2431)} - \mathbf{F}_{(2314)} + \mathbf{F}_{(1324)}.$

References

1. Marcelo Aguiar and Frank Sottile, *Structure of the Malvenuto-Reutenauer Hopf algebra of permutations*, Adv. Math. **191** (2005), no. 2, 225–275, math.CO/0203282.
2. Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon, *Some generalizations of quasi-symmetric functions and noncommutative symmetric functions*, Springer, Berlin, 2000, math.CO/0105065.
3. Loïc Foissy, *Les algèbres de Hopf des arbres enracinés, II*, Bull. Sci. Math. **126** (2002), 249–288.
4. ———, *Bidendriform bialgebras, trees, and quasi-symmetric functions*, J. Pure Appl. Algebra **209** (2007), 429–459.
5. Jean-Louis Loday, *Dialgebras*, Lecture Notes in Math., no. 1763, Springer, Berlin, 2001, math.QA/0102053.
6. Jean-Louis Loday and Maria O. Ronco, *Hopf algebra of the planar binary trees*, Adv. Math. **139** (1998), no. 2, 293–309.
7. Claudia Malvenuto and Christophe Reutenauer, *Duality between quasi-symmetric functions and the Solomon descent algebra*, J. Algebra **177** (1995), no. 3, 967–982.
8. Maria O. Ronco, *Primitive elements of a free dendriform algebra*, Contemp. Math. **267** (2000), 245–263.
9. ———, *Eulerian idempotents and Milnor-Moore theorem for certain non-cocommutative Hopf algebras*, J. Algebra **254** (2002), no. 1, 152–172.

LABORATOIRE DE MATHÉMATIQUES - FRE3111, UNIVERSITÉ DE REIMS *Moulin de la Housse*
- BP 1039 - 51687 REIMS Cedex 2, France

E-mail address: e-mail : loic.foissy@univ-reims.fr

A Renormalisation Group approach to Stochastic Löewner Evolutions

Roland Friedrich

ABSTRACT. In this notes we shall describe the relation of a certain class of simple random curves arising in 2D statistical mechanics models in the scaling limit, which can be described dynamically by Stochastic Löewner Evolutions (SLE), and the equivalent Renormalisation Group (RG) theoretic interpretation in Conformal Field Theory, as a fixed point of the RG flow. Further, we shall recall the relation of this random curves with String Theory, and how one can derive a general measure on such random paths, by using weighted regularised determinants, which come from sections of twisted line bundles.

1. Les contes aléatoires de la Vallée de Chevreuse

This text is a somewhat extended version of a talk I gave at the workshop “RENORMALIZATION” at the Max-Planck-Institute for Mathematics in Bonn, in December 2006.

Its aim is to explain, how Stochastic Löewner Evolutions (SLE) are connected with the Renormalisation Group (RG), but also what the underlying global geometric picture is. In doing so, I shall restrict myself to the case of the unit disc, the original model, and leave out the general theory for arbitrary Riemann surfaces, as developed in [11, 12, 18, 19]. During the process of writing, I felt free to expand on results which I obtained several years ago, or facts I was aware of for quite some time. A detailed companion text to the material presented here, is [10].

There has been two major directions to construct measures on random loops or intervals, namely the one pursued by P. Malliavin [22] and the other one by O. Schramm, who initiated SLE [27], but also G. Lawler, W. Werner [21], (LSW), and M. Aizenman [3]. Both was taking place roughly around 1999.

The (technical) foundations of SLE were laid down by Schramm and S. Rohde in [26], and subsequently applied by LSW to open probabilistic problems.

However it was the article [21], building on previous work of LW, which came up with a wealth of amazing statements and objects, like the “restriction property”, the appearance of the “Schwarzian derivative”, “SLE(κ, ρ)”, the “loop soup” etc. That was a beautiful and rare challenge, for Conformal Field Theory.

Geographically, things were happening on a very short scale. At Orsay, W. Werner, and students actively developed the field.

Not far away, M. Bauer and D. Bernard [6] took up the challenge. They had definitively a major impact with their version of the CFT-SLE correspondence, i.e., the Saclay version of SLE.

One should note, that in the pre-SLE era, B. Duplantier had also derived stunning statements about random processes, by using methods from statistical mechanics and CFT. Some of his results certainly, challenged the mathematical community, to re-derive it “rigorously”, and to merge it with SLE.

The IHES version of SLE is the one we are presenting here. It builds, from the very beginning on determinant line bundles, i.e. domain-wise regularised determinants, and the so-called “Virasoro Uniformisation” (VU), which had been co-invented by Kontsevich, quite some time ago, and which generalises parts of the work of Kirillov, Yurev and Neretin. Also, (VU) is at the heart of the fundamental approach to CFT on Riemann surfaces, by Tsuchiya, Ueno, Yamada, Kawamoto and Namikawa [29].

The aim of the IHES version was to give a unified treatment which would incorporate both approaches mentioned at the beginning, but also explain all the general results physically, as proposed by LSW.

What came out, is indeed not only able to link parts of String Theory and CFT with SLE, but also to generalise it to arbitrary Riemann surfaces and other supports of the conformally invariant measures [11, 12, 18, 19]. This approach has by now become a standard, from which all types of SLE can naturally be deduced.

After the appearance of O. Schramm’s original paper [27], the topic was actively picked up in the Vallée de Chevreuse, where fundamental, and by now standard, contributions and extension to the subject of SLE were made. In particular, its genuine connection with conformal field theory were successfully established, as exemplified by the formula, that gives the connection of SLE_κ with the “central charge” c of the corresponding CFT [6, 13, 18]:

$$c_\kappa = \frac{(6 - \kappa)(3\kappa - 8)}{2\kappa}.$$

2. Different point of views

Let us start with the following example, as depicted in Figure 1. There we have a hexagonal lattice approximating a closed disc. The yellow coloured hexagons compose the interior in the discretisation of the disc and the white ones, the boundary, indicated by the circle. Further we have drawn a simple polygonal path connecting

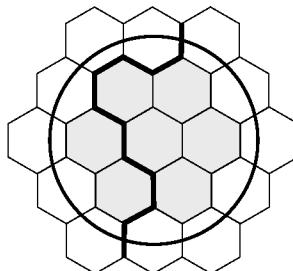


FIGURE 1. A simple curve on a hexagonal lattice connecting two boundary points. Topological set-up.

two boundary points and strictly laying in the interior of the disc.

From a discrete, combinatorial and enumerative point of view, we can ask several natural questions. E.g. how many discrete simple path are connecting the same two boundary points, thereby always running inside the circle or what is the (Euclidean) length of the longest path. It is important to note, that here boundary and interior are a priori just topological notions, merely giving the roughest binary framework to start with.

If we would keep the circle fixed but replace our lattice by another of a finer grid size, we still can ask the same questions but some of them become increasingly unnatural. Also, the combinatorial complexity of the problem would start to explode, so that instead of keeping track of all possibilities, we would be forced to resort to statistical methods. But what does that exactly mean and what are the relevant quantities we should look for.

Well, by now many people and groups invented their questions and methods, as the above problem could arise e.g. in Polymer Physics.

A particular challenge would be to construct probability measures on such discrete paths and to understand how all this scales as the mesh size becomes smaller and smaller. Also, what should the continuum object be and what are the natural measures to start with, as the same set can support different ones.

Again, depending on the context, e.g. dynamical generation of paths or “static sets”, several classes of measures have been considered and proposed.

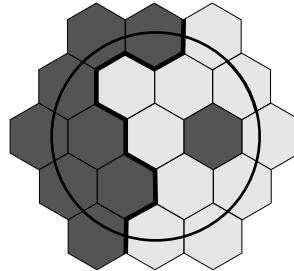


FIGURE 2. A simple curve on a hexagonal lattice connecting two boundary points. Boundary condition set-up.

As we shall discuss in the next section, Statistical Mechanics and in particular Lattice Models, provide a rich source of such curves in form of interfaces. But, there are some fine points one has to be aware of, as the simple topological set-up will be altered by the notion of boundary conditions, as depicted in Figure 2, and further explained latter. However, boundary conditions intrinsically reveal that we are taking up a macroscopic stance on the problem, which shifts our focus on what to consider as being natural and interesting.

3. Motivation from Statistical Mechanics

Besides percolation, as originally, the two-dimensional Ising model has by now become a widely adapted way in the literature to motivate SLE and to connect it with statistical mechanics respectively CFT, [12, 11, 18].

So let us start with a simply connected and bounded domain D in the plane $\mathbb{R} \cong \mathbb{C}$ and a triangular lattice (TL) of mesh size $\delta > 0$, “just” covering the closed

set \overline{D} . We shall assume that the boundary of the domain is sufficiently smooth and the mesh size enough fine to avoid cumbersome complications.

A configuration of spins is a function σ on the set of vertices $V(TL)$ with values ± 1 . We may now pass to the dual hexagonal lattice, and colour the hexagons enclosing a vertex with value -1 black and those with value $+1$ white. In the percolation setting this would correspond to site (or vertex) percolation on a triangular grid.

A sure way to get an interface (domain wall), as illustrated in Figure 3, is to take two marked points A and B on the boundary ∂D and then to fix the value of the spin sitting in the centre of exactly one hexagon black, if that one intersects the boundary segment from B to A and similarly white, for that boundary part which belongs to the segment from A to B , assuming a counter-clockwise orientation.

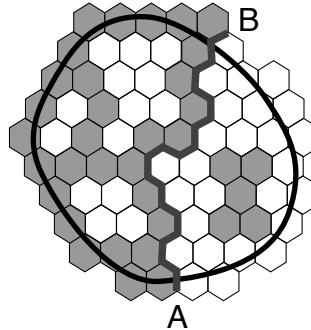


FIGURE 3. The interface associated with the two-dimensional Ising model.

The curve which arises is a Jordan curve (no branching) connecting vertices such that on the left we always have spin -1 and to the right $+1$.

In the case of our model with this particular choice of boundary conditions, we have an object that persists on all finite scales and should therefore be a good candidate to proceed to an object which can be regarded as a “natural” scaling limit, i.e. a macroscopic observable.

However as a look at Figure 3 reveals, once the particular (red) interface is fixed, we could change some of the values of the spins corresponding to boundary parts, without altering the shape of the curve itself, which implies that the set of all configurations having that particular interface, is much bigger than just those compatible with the fixed “domain wall” boundary conditions.

The situation at the discrete level is summarised schematically in Figure 4. The (yellow-blue) square represents the space of all possible configurations, which is fibered over the space of all possible boundary values, i.e. the state space, and corresponds to the case of “free boundary values”.

The state space can be subdivided into two disjoint non-empty sets, which are labelled as “curves” respectively “no curves” and which contain those boundary conditions which may produce a curve γ connecting the two marked points A and B , respectively not. It is important to note that the fibre over a state from “curves” may contain configurations which do not include a domain wall running from A to B . This is indicated by the (blue) triangle $\{\sigma\}_\gamma$ over the set “curves”.

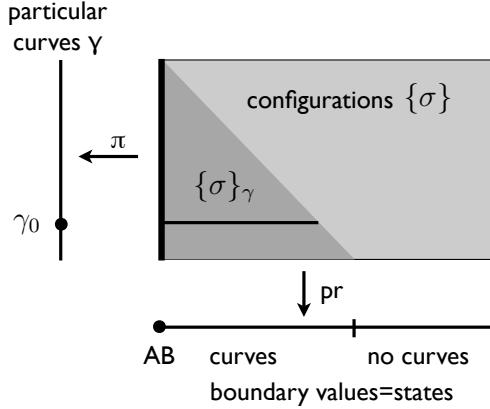


FIGURE 4. The configuration space as a fibered space over the state space, principal fibration, or the non-local fibration over curves.

There is one special state, labelled AB , which corresponds to the boundary conditions described as “domain wall boundary conditions”, (cf. Fig. 3). Now, the fibre over the state AB is composed of configurations that do always contain a Jordan arc running from A to B and therefore we may call it the “special fibre”.

We have yet another (partial) fibration. Namely the space of configurations with a domain wall (blue triangle) is fibered over the set of particular curves. In our Figure 4, this is depicted by the vertical line labelled “particular curves” with the projection map π . It is important to note, that say for a curve γ_0 , the intersection of the fibres $\pi^{-1}(\gamma_0)$ and $pr^{-1}(AB)$ is not just one point (configuration), but indeed does contain several.

3.1. Gibbs measures with “proper” boundary conditions. The Gibbs measure on configurations naturally induces a probability measure on the random curves we just considered. However, the boundary conditions we choose, play an essential role.

So, in the thermodynamical equilibrium at the absolute temperature $T > 0$ the Gibbs measure is described by

$$d\mu_G(\sigma) := \frac{1}{Z} e^{-\beta E[\sigma]}, \quad \beta := \frac{1}{T} \quad (\text{inverse temperature}) ,$$

where the functional $E[\sigma]$, denotes the energy of a configuration and the numerical pre-factor $Z > 0$ serves as normalisation, such that the sum of all elementary events adds up to 1. It is defined as

$$Z := \sum_{\{\sigma\}} e^{-\beta E[\sigma]} ,$$

with the sum extending over all configurations $\{\sigma\}$ and it is called the partition function. The measure for a fixed domain D depends on two parameters, namely the temperature but also the underlying mesh size, i.e. $\mu_G = \mu_{G(T,\delta)}$.

If we choose fixed AB -boundary conditions for the calculation of the partition function, then we are restricting ourselves to the special fibre over AB , and take its Boltzmann-weighted volume $Z_{\sigma_{AB}}$ as normalisation factor. For free boundary

conditions we would restrict the summation to the set of all configurations which contain a domain wall, i.e. to the set $\{\sigma\}_\gamma$ (blue triangle) with Z_{σ_γ} as the corresponding partition function.

The natural equivalence relation on the set of configurations $\{\sigma\}_\gamma$ is given by declaring two realisations σ_1 and σ_2 as being equivalent if they include the same chordal domain wall, connecting the two marked boundary points, cf. Fig. 3. Again, we can restrict the equivalence relation to the fibre $\text{pr}^{-1}(AB)$ only. In any case, the situation yields the fibration with base space the particular curves and the equivalent configuration corresponding to the fibres, respectively to their intersection with $\text{pr}^{-1}(AB)$.

The probability measure on the quotient space, with simple events the particular curves γ , is the image measure, either with respect to $Z_{\sigma_{AB}}$ or Z_{σ_γ} , i.e. $\pi_*\mu_{G_{AB}}$ resp. $\pi_*\mu_{G_{\sigma_\gamma}}$.

Although the geometric set of simple random curves is the same (topological set-up), by taking different boundary conditions we arrive at having two different measures on the same set.

Now, and this is important to note, the proofs (analytic or numeric for the Ising model) [28, 15] concerning the scaling limit of the above measures, i.e. the approach of the thermodynamic limit along the critical temperature T_c , deal with the family of measures on the special fibre $\text{pr}^{-1}(AB)$.

Further, as we shall discuss in the next section, on general grounds it was / is expected that the scaling limit is a measure, supported on simple paths, which is conformally invariant.

Let us close with the following remarks. The procedure just described, is not particular to the Ising model. We can choose other models, e.g. Q -states Potts model (i.e. different Boltzmann-weights) as well as other lattices (not necessarily hexagonal) with appropriate boundary conditions (e.g. wired and free). For lattice models there are other “good” probabilistic events. In the Ising model slightly below the critical temperature T_c we observe the sea of nested Jordan curves of domain boundaries. Again, by passing to T_c and simultaneously rescaling we obtain a dense

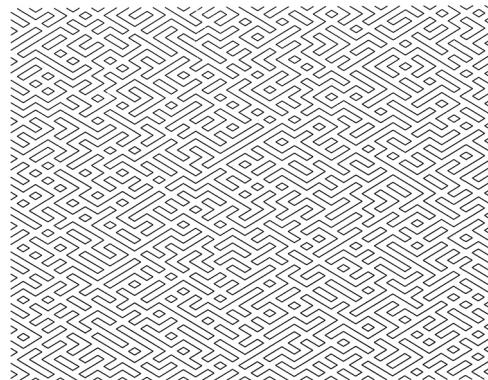


FIGURE 5. Central charge zero random field of loops on a torus, which has a dual description as a percolation model. (Result from work on Random Fields and SLE, by Friedrich and Kontsevich from 2000/2001.)

collection of closed, non-intersecting loops, carrying a scale invariant probability distribution which would be also derived from the series of Gibbs measures as we explained in the case of chordal lines. (cf. [19])

4. Measures as Domain Functionals, the ALPS-correspondence and SLE

So-far our discussion centred around a fixed domain and the measures associated with it, supported on simple curves and connecting two previously chosen distinct boundary points. Now, we could ask, how the measure would change if we keep the domain, but vary the points, or even more generally, also vary the underlying domains. So, what we have in fact is a domain functional

$$(1) \quad D_{A,B} \mapsto \mu_{D_{A,B}}$$

where D denotes the domain, and $A, B \in \partial D$, two distinguished (and ordered) boundary points.

In two dimensions, there exists a very strong statement, the Riemann mapping theorem, which tells that any two simply connected domains are conformally related, and if the domains are bounded by Jordan curves, then this mapping can be extended as a homeomorphism to the whole boundary, according to Carathéodory's theorem.

Now it was conjectured [2, 20], that these measures should be conformally invariant in the scaling limit, which can be expressed as the following commutative diagram,

$$\begin{array}{ccc} D_{A,B} & \xrightarrow{f} & D'_{A',B'} \\ F_{\text{ALPS}} \downarrow & & \downarrow F_{\text{ALPS}} \\ \mu_{D_{A,B}} & \xrightarrow{f_*} & \mu_{D'_{A',B'}} \end{array}$$

with the morphism f being a conformal equivalence which respects the ordered marked points, i.e. $f(A) = A'$ resp. $f(B) = B'$, and as objects Jordan domains with two marked and ordered points, $\text{JDom}_{\bullet,\bullet}$. As we shall see latter, this category can be enlarged to include also slit domains. f_* denotes the induced measurable mapping, i.e. $F_{\text{ALPS}}(f) =: f_*$, where F_{ALPS} stands for the “Aizenman–Langlands–Pouillet–Saint-Aubin” correspondence (functor).

Let us note that because of conformal invariance and because the set of morphisms between any two objects in $\text{JDom}_{\bullet,\bullet}$ has more than one object, in fact a real continuum, it is enough to study just one reference model, e.g. the upper half-plane \mathbb{H} with 0 and ∞ as the marked points.

The category $\text{JDom}_{\bullet,\bullet}$ has several equivalent parametrisations. Namely, if we consider first the set J^∞ of smooth Jordan curves, then one has the double quotient [1]

$$\text{SU}(1,1) \backslash \text{Diff}_+(S^1) / \text{SU}(1,1)$$

as the base space, and as fibre model of the total space, the torus (minus the diagonal), i.e., $S^1 \times S^1 \setminus \{\text{diagonal}\}$. Sections then correspond to domains with two distinct marked boundary points.

As these measures are translational invariant, it is enough to look at J_0^∞ , i.e., all smooth Jordan curves surrounding the origin. The uniformising application with domain the unit disc \mathbb{D} , is unique if we require it to preserve the origin 0, and to

have either a strictly positive derivative at 0 or to map 1 onto one of the marked points.

For symmetry reasons we take the first normalisation. Then the manifold \mathcal{M} of all such univalent maps corresponds to a contractible subset in (cf. [11])

$$\text{Aut}(\mathcal{O}) = \{a_1 z + a_2 z^2 + a_3 z^3 \dots, 0 < a_1\} \hookrightarrow \mathbb{C}^{\mathbb{N}^*},$$

which itself is a contractible space. By the Bieberbach-DeBranges theorem the coefficients of elements in \mathcal{M} satisfy $a_n \leq n \cdot a_1$. Similarly, we have

$$\left\{ \begin{array}{l} \text{conformally} \\ \text{invariant} \\ \text{measures} \end{array} \right\}_{\kappa > 0} \longrightarrow \mathcal{M} \times (S^1 \times S^1 \setminus \{\text{diagonal}\}) \xrightarrow{\pi} \mathcal{M} \hookrightarrow \text{Aut}(\mathcal{O}).$$

Then the trivial bundle, with fibre (almost) a torus, parametrise, up to one real positive constant κ , conformally invariant probability measures on simple paths which connect two distinct boundary points, as we shall see.

It was O. Schramm's original insight [27], to use Löwner's slit mapping to describe these random traces dynamically and to classify then all possible conformally invariant measures on them, i.e. to show the existence of κ .

Physically, the existence of such a parameter can be understood as labelling models from Statistical Mechanics that contribute a measure, as previously explained. However, it is a deep and fundamental fact of the two-dimensional conformally invariant realm, that one parameter is enough, as we shall see in Section 5.

4.1. Schramm-Loewner Evolutions. One of the basic observations in deriving the “driving function” in the dynamical approach to random Jordan arcs is, besides a symmetry argument, continuity and the previously introduced notion of conformal invariance, the following Markovian type property (stated for the category $\text{JDom}_{\bullet, \bullet}$):

For a domain D with non-degenerate boundary, let $\mathcal{W}(D_{A,B})$ be the set of Jordan arcs in D with endpoints A and B . Denote by $\{\mu_{D_{A,B}}\}$ a family of probability measures on Jordan arcs in the complex plane such that

$$\mu_{D_{A,B}}(\mathcal{W}(D_{A,B})) = 1.$$

Then the Markovian-type property says for γ a random Jordan arc, that if γ' is a sub-arc of γ which has A as one endpoint and whose other endpoint we denote by A' , then the conditional distribution of γ given γ' is

$$(2) \quad \mu_{D_{A,B}|\gamma'} = \mu_{(D \setminus \gamma')_{A',B}}.$$

So, the only compatible driving function which satisfies the above requirements has to be proportional to standard one-dimensional Brownian motion, which leads to the following facts [27, 26].

The chordal SLE $_{\kappa}$ curve γ in the upper half-plane \mathbb{H} describes the growth of simple random curves emerging from the origin and aiming at infinity, as follows:

DEFINITION 4.1 (Schramm-Loewner Equation). *For $z \in \mathbb{H}$, $t \geq 0$ define $g_t(z)$ by $g_0(z) = z$ and*

$$(3) \quad \frac{\partial g_t(z)}{\partial t} = \frac{2}{g_t(z) - W_t}.$$

The maps g_t are normalised such that $g_t(z) = z + o(1)$ when $z \rightarrow \infty$ and $W_t := \sqrt{\kappa} B_t$ where $B_t(\omega)$ is the standard one-dimensional Brownian motion, starting at 0 and with variance $\kappa > 0$. Given the initial point $g_0(z) = z$, the ordinary differential equation (3) is well defined until a random time τ_z when the right-hand side in (3) has a pole. There are two sets of points that are of interest, namely the preimage of infinity $\tau^{-1}(\infty)$ and its complement. For those in the complement we define:

$$(4) \quad K_t := \overline{\{z \in \mathbb{H} : \tau(z) < t\}} .$$

The family $(K_t)_{t \geq 0}$, called hulls, is an increasing family of compact sets in $\overline{\mathbb{H}}$ where g_t is the uniformising map from $\mathbb{H} \setminus K_t$ onto \mathbb{H} . Further there exists a continuous process $(\gamma_t)_{t \geq 0}$ with values in $\overline{\mathbb{H}}$ such that $\mathbb{H} \setminus K_t$ is the unbounded connected component of $\mathbb{H} \setminus \gamma[0, t]$ with probability one. This process is the trace of the SLE_κ and it can be recovered from g_t , and therefore from W_t , by

$$(5) \quad \gamma_t = \lim_{z \rightarrow W_t, z \in \mathbb{H}} g_t^{-1}(z) .$$

The constant κ characterises the nature of the resulting curves. For $0 < \kappa \leq 4$, SLE_κ traces over simple curves, for $4 < \kappa < 8$ self-touching curves (curves with double points, but without crossing its past) and, finally, if $8 \leq \kappa$ the trace becomes space filling.

Now, for another simply connected domain D with two boundary points $A, B \in \partial D$ the chordal SLE_κ in D from A to B is defined as

$$K_t(D_{A,B}) := h^{-1}(K_t(\mathbb{H}, 0, \infty))$$

where $K_t(\mathbb{H}, 0, \infty)$ is the hull as in (4) and h is the conformal map from D onto \mathbb{H} with $h(A) = 0$ and $h(B) = \infty$.

Before we end this section, let us rewrite (3) in Itô form, by setting $f_t(z) := g_t(z) - W_t$, which now satisfies the stochastic differential equation

$$(6) \quad df_t(z) = \frac{2}{f_t(z)} dt - dW_t .$$

For a non-singular boundary point $x \in \mathbb{R}$, we can read off the generator A for the Itô-diffusion $X_t := f_t(x)$ as

$$A = 2 \frac{1}{x} \frac{d}{dx} - \frac{\kappa}{2} \frac{d^2}{dx^2} .$$

Defining the first order differential operators

$$(7) \quad \ell_n := -x^{n+1} \frac{d}{dx} \quad n \in \mathbb{Z} ,$$

we obtain

$$A = \frac{\kappa}{2} \ell_{-1}^2 - 2\ell_{-2} .$$

Let us note, that the differential operators (7) form a representation of the Witt algebra [13, 14]. We shall come back to this matters latter, where we shall see, how the A -harmonic functions correspond to null-vectors in a Verma module of the Virasoro algebra.

5. Renormalisation Group flow and conformally invariant measures

Before we proceed, we recall Zamolodchikov's c -Theorem for two-dimensional field theories [30].

Let us be given a (Euclidean) Field Theory with action functional $S[\mathbf{g}, a]$ depending on an (infinite) set of dimensionless parameters, $\mathbf{g} = (g_1, g_2, \dots)$, the “coupling constants”, and an (ultraviolet) “cut-off” a such that the action is obtained as an integral of local densities, i.e. $S = \int \sigma(\mathbf{g}, a, x) dx$.

The fundamental assumption is the existence of a one-parameter group of motions R_t in the space Q of coupling constants \mathbf{g} , $R_t : Q \rightarrow Q$, with the property that a field theory described by an action $S[R_t \mathbf{g}, e^t a]$ is equivalent to the original theory with the action $S[\mathbf{g}, a]$ modulo correlations. This means that all correlation functions calculated in the two theories are the same at scales $x \gg e^t a$ and $t > 0$. The components of the vector fields which generate the renormalisation group (RG) flow are called “ β -functions”, i.e.

$$(8) \quad \frac{dg_i}{dt} = \beta_i(\mathbf{g}) .$$

Then the following properties hold true for the RG:

- (1) There exists a positive function $c(g) \geq 0$ which decreases monotonically, i.e.,

$$(9) \quad \frac{d}{dt} c = \beta_i(g) \frac{\partial}{\partial g_i} c(g) \leq 0$$

with equality only obtained at the fixed points of the RG-flow, i.e., at $g = g_*$, $(\beta_i(g_*) = 0)$.

- (2) The “critical” fixed points are stationary for $c(g)$, i.e. $\beta_i(g) = 0 \Rightarrow \partial c / \partial g_i = 0$. Further, at the critical fixed points the corresponding 2D field theory is a Conformal Field Theory, with generators L_n , $n \in \mathbb{Z}$, of the infinite symmetry algebra, the Virasoro algebra, satisfying the commutation relations

$$(10) \quad [L_n, L_m] = (n - m)L_{n+m} + \frac{\tilde{c}}{12}(n^3 - n)\delta_{n+m,0} ,$$

with \tilde{c} the “central charge” and which is a function of the fixed points, i.e., $\tilde{c} = \tilde{c}(g_*)$.

- (3) The value of $c(g)$ at the fixed point g_* coincides with the corresponding central charge in (10), i.e., $c(g_*) = \tilde{c}(g_*)$.

Now, in Figure 6, we have summarised the process of approaching the scaling limit of the two-parameter family of Gibbs measures $d\mu_G(N, T)$, $N \sim 1/\delta^2$, for a given Boltzmann weight, i.e., for a particular lattice model. Then, as just discussed, the corresponding fixed point in the RG flow is described by a CFT, characterised by its central charge c . Additionally the limiting measure should be conformally invariant and supported on random simple chordal paths, i.e., by SLE for some κ .

Now, the “Main Identity” which functionally relates Schramm's diffusion constant κ of the SLE process with CFT, has been derived by various people very early [6, 14, 18]. The identity reads

$$(11) \quad c = \frac{(\kappa - 6)(3\kappa - 8)}{2\kappa} .$$

Therefore, we know from the c -Theorem how to relate the RG flow and SLE.

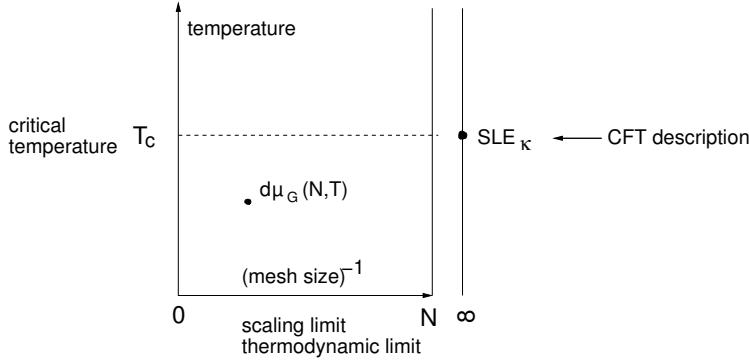


FIGURE 6. The SLE_κ claim: SLE_κ is the unstable fixed point of the renormalisation group flow.

The link we have just derived here is of considerable mathematical interest, as it connects dynamical systems, representation theory and stochastic analysis very deeply. It is certainly worth, to be pursued further.

6. The String model for the Wilson loop and SLE

The Functional Integral approach to Schramm-Loewner Evolutions was introduced in [11, 12, 18] and further extended in [7] to Liouville Theory.

The physical problem for the Wilson loop in (Classical) String Theory is to sum over all two-dimensional real surfaces having a closed Jordan curve C in \mathbb{R}^n as boundary.

Historically, this sum Z , the partition function, has been defined and calculated (approximately) according to the classical “Dual String” methods, by applying the Nambu-Goto action, by Eguchi, Durhuus, Olesen, Nielsen, Petersen, Brink, Di Vecchia, Howe, Deser, Zumino, Lüscher, Symanzik and Weisz for “dual strings” however it was Polyakov’s approach and his specific action, centred around the Dirichlet integral, which made the problem much more approachable. His action was then subsequently extended to the same problem, i.e. to non-closed strings, by others like Friedan, but decisively by O. Alvarez [5]. In any case, the underlying physical object is a Functional Integral, with the action being the specific part, and the mathematical problem left over, to interpret it “rigorously”. For mathematical details of String Theory, see, e.g. [4].

The “partition function” is obtained from the well-known functional integral

$$(12) \quad Z := \int_{\{g\}} \int_{\{X\}} e^{-S[g, X]} [Dg][DX],$$

which is properly computed over all embeddings $\{X\}$ and all Riemannian metrics $\{g\}$.

The action $S = S[g, X]$, for surfaces Σ with boundary $\partial\Sigma$, contains besides the Dirichlet Energy $D[g, X]$ of the embedding, also other terms, to make the theory

renormalisable, and in its simplest version reads in local co-ordinates,

$$(13) \quad S[g, X] := D[g, X] + \text{metric terms}$$

$$(14) \quad = \frac{1}{2} \int_{\Sigma} g^{ij} \frac{\partial X^{\mu}}{\partial x^i} \frac{\partial X^{\mu}}{\partial y} \sqrt{\det g} dx dy$$

$$(15) \quad + \frac{1}{2\pi} \int_{\Sigma} K_g \sqrt{\det g} dx dy + \frac{1}{2\pi} \int_{\partial\Sigma} k_g |dx|$$

where K_g is the Gauss curvature of the metric g and k_g the geodesic curvature of ∂M according to g .

Subsequently we shall not discuss the metric contributions further as they basically enter via the Gauss-Bonnet theorem. Also, we shall only treat regular metrics, although singularities have a profound effect [7, 19].

Now, in Conformal Field Theory, the marginals, i.e., the integral over all embeddings for a fixed metric is the important quantity, as we recall. This will also reveal the analogy with our earlier discussion in the statistical mechanics approach.

6.1. Spaces of mappings and the $H^{1/2}$ space on the circle. The Sobolev space $H^{1/2}(S^1, \mathbb{R})/\mathbb{R}$ of $L^2(S^1)$ real functions with mean-value zero on the circle can be identified with the sequence space

$$\ell_2^{1/2} = \{ u \equiv (u_0, u_1, u_2, \dots) \mid u_i \in \mathbb{C} \text{ and } \{\sqrt{n} u_n\} \text{ is square summable} \}.$$

Part of the importance of this Hilbert space comes from the facts that one can interpret its vectors as boundary values of real harmonic functions on the unit disc, \mathbb{D} , with finite Dirichlet energy but also as it characterises the subset of quasi-symmetric (qs.) homeomorphisms of the set of all homeomorphisms of S^1 , [23]. The Poisson integral representation gives then a harmonic extension of the space $H^{1/2}$, which is also an isometric isomorphism of Hilbert spaces. For the disc the extension can explicitly be written down in terms of the formula of Douglas.

Let us consider a compact real two-dimensional smooth surface Σ with non-empty and non-degenerate boundary $\partial\Sigma$, homeomorphic to the unit circle. Let us also fix a metric g , the background metric, on Σ and let us consider a map

$$h : \Sigma \rightarrow \mathbb{R}^n,$$

mapping $\partial\Sigma$ diffeomorphically and orientation preserving onto the contour C .

For h harmonic with respect to the metric g we would have for the Laplacian Δ_g ,

$$(16) \quad \Delta_g h \equiv 0.$$

The fact which permits to progress further in the endeavour of defining the path integral is first that any embedding $X : \Sigma \rightarrow \mathbb{R}^n$ compatible with the boundary conditions, i.e., $X(\partial\Sigma) = C$ can be decomposed as

$$(17) \quad X = h_X + X_0$$

where h_X is the unique harmonic map with $h_X|_{\partial\Sigma} \equiv X|_{\partial\Sigma}$ and $X_0|_{\partial\Sigma} = 0$. This yields the affine space

$$h_X + H_0^{1/2}(\Sigma, \mathbb{R}^n)$$

with $H_0^{1/2}(\Sigma, \mathbb{R}^n)$ denoting the Sobolev space of all maps from Σ to \mathbb{R}^n with vanishing boundary values.

An appropriate Hilbert space of states would be $\mathcal{H} := H^{1/2}(\partial\Sigma, C)$. Let $f \in \mathcal{H}$

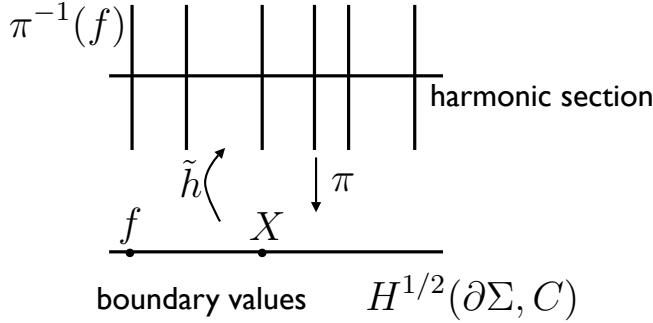


FIGURE 7. The harmonic section \tilde{h} (harmonic extension) of the space of all embeddings of the boundary $\partial\Sigma$ of the surface. The boundary values compose the space of states.

and let us denote by \tilde{f} its harmonic extension, i.e. the unique harmonic function with boundary value f . The situation is schematised in Figure 7.

In the next few lines we shall deal only with one component of the field. The integral over the fibre $\pi^{-1}(f)$, (cf. Fig. 7),

$$(18) \quad \Psi[f] := \int_{\pi^{-1}(f)} e^{-S[g, X]} [DX]$$

is the continuum version of the partition sum with a specific choice of boundary conditions. Then the following calculation, with the previous notational conventions, yields for the Dirichlet energy (* Hodge star),

$$(19) \quad 2 \cdot S[X] = \int_{\Sigma} dX \wedge *dX = \int_{\Sigma} d(\tilde{f} + X_0) \wedge *d(\tilde{f} + X_0)$$

$$(20) \quad = \int_{\Sigma} d\tilde{f} \wedge *d\tilde{f} + \int_{\Sigma} dX_0 \wedge *dX_0 + 2 \int_{\Sigma} dX_0 \wedge d\tilde{f}$$

Since $\Delta_g \tilde{f} = d * d\tilde{f} = 0$ by definition for the harmonic function \tilde{f} , we obtain

$$(21) \quad S[\tilde{f} + X_0] = S[X_0] + S[\tilde{f}] = S[X_0] + \frac{1}{2} \int_{\partial\Sigma} f * d\tilde{f} .$$

Therefore the partition function factorises into

$$\int_{\pi^{-1}(f)} e^{-S[g, X]} [DX] = e^{-\frac{1}{2} \int_{\partial\Sigma} f * d\tilde{f}} \cdot \int_{\{X_0\}} e^{-S[X_0]} [DX_0] ,$$

which gives (for one component)

$$(22) \quad \Psi[f] = \left[\frac{\det(\Delta_g)}{\text{Area}(\Sigma, g)} \right]^{-1/2} \cdot e^{-\frac{1}{2} \int_{\partial\Sigma} f * d\tilde{f}} ,$$

and where it is understood that the determinants are regularised. Note that we are dealing with bordered surfaces, and therefore we do not have zero modes.

To obtain the full marginal $Z[g]$, we have to integrate over all possible boundary values. The appropriate measure μ , has been considered from a Gaussian point of view by G. Segal and I. Frenkel, and for “Unitarising Measures for the Virasoro

algebra”, by P. Malliavin, H. Airault and A. Thalmaier [22, 1]; (cf. the support of measures in [1]). Technically

$$(23) \quad \int_{\tilde{h}(\mathcal{H})} e^{-D[g, \tilde{h}]} [D\tilde{h}] := \int_{\varphi \in \text{Hom}_{qs}(S^1)} e^{-E[\tilde{\varphi}, g_0]} d\mu_{\text{SFM}}(\varphi) =: j[\mathbf{t}, C]$$

where we have denoted the measure on $\text{Hom}_{qs}(S^1)$ by $d\mu_{\text{SFM}}$, $E[\tilde{\varphi}, g_0]$ stands for the Dirichlet integral of the harmonic extension induced by φ , which also requires now Σ to have a complex structure, and finally \mathbf{t} denotes the conformal class of the metric g_0 compatible with the structure on Σ , i.e. a point in Teichmüller space (cf. [4]).

The expression (23) is independent of the choices made, with the crucial exception that it depends on the conformal class of the metric.

So, for every fixed metric g , the marginal integral in (12) over embeddings equals

$$(24) \quad Z[g] = \int_{\text{Emb}(\Sigma)} e^{-S[X, g]} [DX] := j[\mathbf{t}, C] \cdot \left[\frac{\det(\Delta_g)}{\text{Area}(\Sigma, g)} \right]^{-n/2}$$

If we restrict to “planar” surfaces with the Wilson loop C as boundary, i.e. to embeddings of the unit disc into \mathbb{R}^n , then the dependence on the conformal class is simple, as all complex structures on the disc are equivalent. It gives also (part) of the relevant partition function for SLE.

However, the regularisation procedure introduces the so-called conformal anomaly, which renders the expressions covariant. Also, for singular metrics interesting effects show up which correspond to exotic versions of the SLE-process [7], e.g. SLE(κ, ρ) [21], or measure the degree of non-commutativity of SLE, cf. [12, 21].

Now, in CFT the important objects are the marginals $Z[g]$, and they are naturally grouped together as a determinant line bundle, as we shall discuss next.

6.2. Determinant line bundles and flat connections. In String Theory but also in Conformal Field Theory, the partition function is considered to be a section of a determinant line bundle. Here we shall briefly recall how one can derive a measure on random paths, by using regularised determinants [12, 11, 18, 19].

Since to every Jordan domain we can associate the determinant of the Laplacian (with respect to the Euclidean metric and Dirichlet boundary conditions), i.e., $\det(\Delta_D) := \det(\Delta_{g_{\text{Eucl.}}})$, we get a trivial bundle over \mathcal{M} , where $f \in \mathcal{M}$ denotes the uniformising map from the unit disc \mathbb{D} onto the domain D , containing the origin.

$$\begin{array}{c} \det(\Delta_{f(\mathbb{D})}) \\ \pi \downarrow \\ \mathcal{M} \end{array}$$

The uniformising map provides us also with a natural connection which allows us to compare the regularised determinants at different points. It has its origin in Polyakov’s string theory [25] and was then subsequently extended in the works [5, 24, 16]. Let us consider the space \mathcal{F} of all flat metrics on \mathbb{D} which are conformal to the Euclidean metric, obtained by pull-back. Namely, for $D \in \text{JDom}$ let $f : \mathbb{D} \rightarrow D$ be a conformal equivalence, and define

$$\phi := \log |f'| .$$

This gives a correspondence of harmonic functions on \mathbb{D} with the category JDom and by Weyl rescaling with \mathcal{F} via

$$ds = |f'| |dz| = e^\phi |dz| .$$

To fix the $\text{SU}(1, 1)$ -freedom, which gives classes of isometric metrics, we divide the state space $H^{1/2}(S^1)$ by the Möbius group of the disc. Henceforth we shall work with the equivalence classes, so, e.g. 0 corresponds to the orbit of the Euclidean metric under $\text{SU}(1, 1)$.

The connection reads [24]:

$$(25) \quad \det(\Delta_D) = e^{-\frac{1}{6\pi} \oint_{S^1} (\frac{1}{2} \phi * d\phi + \phi |dz|)} \cdot \det(\Delta_{\mathbb{D}})$$

Next we would like to show a group property which is essential, as it translates latter into the Markov property, on which conformally invariant measures on paths hinge. Let us consider the sequence of conformal maps between domains \mathbb{D}, D, G :

$$\mathbb{D} \xrightarrow{f} D \xrightarrow{g} G .$$

Then the relation of $\det(\Delta_G)$ and $\det(\Delta_{\mathbb{D}})$ is obtained via $\frac{d}{dz} g(f(z)) = g'(f(z)) \cdot f'(z)$, and

$$\log |g'(f(z)) \cdot f'(z)| = \underbrace{\log |g'(f(z))|}_{=: \psi(z)} + \underbrace{\log |f'(z)|}_{=: \phi(z)} .$$

Further by using the property of harmonic functions, i.e.,

$$\oint_{S^1} \frac{1}{2} (\phi \partial_n \psi + \psi \partial_n \phi) = 0 ,$$

gives

$$\det(\Delta_G) = \underbrace{e^{-\frac{1}{6\pi} \oint_{S^1} (\frac{1}{2} \psi(f(z)) * d\psi(f(z)) + \psi(f(z)) |dz|)}}_{I.} \cdot \underbrace{e^{\frac{1}{6\pi} \oint_{S^1} (\frac{1}{2} \phi * d\phi + \phi |dz|)} \cdot \det(\Delta_{\mathbb{D}})}_{II.}$$

where

$$\begin{aligned} I. &= \oint_{\partial D} \left(\frac{1}{2} \tilde{\psi} * d\tilde{\psi} + \tilde{\psi} |dw| \right) \quad \text{with } \tilde{\psi}(w) := \log |g'(w)| , \\ II. &= \det(\Delta_D) \end{aligned}$$

This also shows, that we can consider the determinant for the unit disc as the origin in an infinite affine space.

Let us note, that one should be careful with some of the signs in the literature (cf. [5, 24, 16]). Then according to expression (24), for the partition function with two-dimensional target space, we have to take the inverse on both sides in (25), i.e.,

$$(26) \quad Z_D = e^{\frac{1}{6\pi} \oint_{S^1} (\frac{1}{2} \phi * d\phi + \phi |dz|)} \cdot Z_{\mathbb{D}} .$$

The mapping f which maps the unit disc \mathbb{D} onto the slit unit disc $\mathbb{D} \setminus [1 - \sqrt{t}, 1]$, can be given by:

$$f(z) = \frac{\sqrt{1+t} - \sqrt{t + \left(\frac{z-1}{z+1}\right)^2}}{\sqrt{1+t} + \sqrt{t + \left(\frac{z-1}{z+1}\right)^2}} .$$

But for such singular disturbances of the boundary the variation formula (25) “breaks down”, as the Weyl rescaled metric acquires singularities on the boundary. In order to compute nevertheless the variation of the partition function, one has

to resort to a regularisation procedure, which involves the Schwarzian derivative (cf. [12]).

However, as the expression (22) showed, we have also to specify boundary conditions, i.e. a state in which our system should be.

So, if we translate the “domain-wall” boundary conditions, we get another contribution, which contains the information of the points where the boundary conditions change. In the continuum this corresponds to jump discontinuities along the boundary for the Dirichlet problem. We shall treat here the unit disc with jumps of width 2λ at 1 and -1 . The general case is then obtained by composition with a conformal map.

The harmonic function compatible with these boundary conditions is

$$(27) \quad u(re^{it}) = \Re \left[\frac{2\lambda}{\pi i} (\log(re^{it} - 1) - \log(re^{it} + 1)) \right]$$

where we have assumed counter-clockwise orientation of the circle. However, the resulting contribution according (22) has to be regularised in order to have a finite quantity, Z_λ , depending on the points of discontinuity and the height of the jump(s).

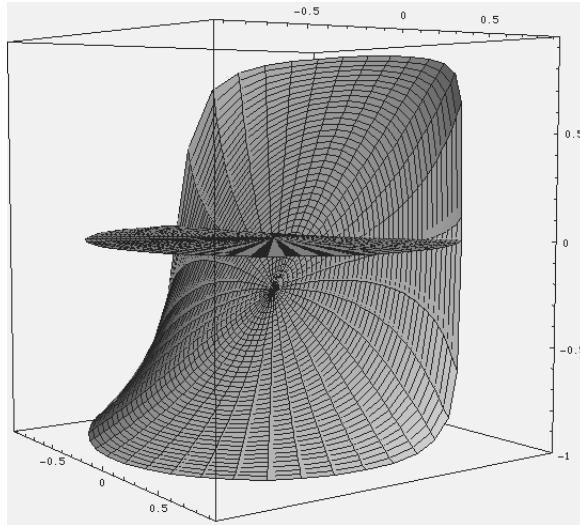


FIGURE 8. Harmonic function which satisfies the jump boundary conditions with $\lambda = 1$ on the arc from $\pi/4$ to π and -1 on the rest, for the Dirichlet problem. In the centre the unit disc is displayed.

Nevertheless, this reveals the global geometry of our partition functions in case of the Bosonic free field, i.e. for $c = 1$:

$$\begin{array}{ccc} \pi^*(\det(\Delta_D)^{-1}) \otimes Z_\lambda & \longrightarrow & \mathcal{M} \times (S^1 \times S^1 \setminus \{\text{diagonal}\}) \\ & & \pi \downarrow \\ & & \det(\Delta_D)^{-1} \longrightarrow \mathcal{M} \hookrightarrow \text{Aut}(\mathcal{O}) \end{array}$$

However, what we see is that in the continuum compared with the lattice, there are infinitely many possibilities from which we have to choose the proper

ones. The possible choices compose the spectrum and are dictated by the scaling dimensions h of the “boundary fields”, which can be obtained from highest-weight representations of the Virasoro algebra [9, 8]. They relate to λ as $h = \text{const.} \cdot \lambda^2$. In the next section we shall derive this purely from probability theory.

To summarise, there are a number of variants of SLE, consisting of a Riemannian bordered surface X (oriented, otherwise general topology) with marked points x_1, \dots, x_n on the boundary and y_1, \dots, y_m in the bulk. Analytic coordinates (or merely 1-jets, i.e. dependence only on the first derivative) at the marked points are given. Then the partition function Z is a positive function of such configurations. It has a tensor dependence on analytic co-ordinates (i.e. it transforms as $\prod_i (dz_i)^{h_i} \prod_j |dw_j|^{2h_j}$, z_i local co-ordinate at x_i , w_j local coordinate at y_j), and depends on the metric as $\det(\Delta_D)^{-c}$, where c is the central charge, a real constant. Further, it should be positive and equal to the renormalised partition function in the lattice approximation.

These partition functions are null vectors of canonical Virasoro representations, and they correspond to correlators in Conformal Field Theory [13, 14, 11, 12, 18, 19, 6]. The (simplest) version for the disc with two marked points is now stated.

THEOREM 6.1 (2003 [12, 18, 11, 19]). *Let a configuration (D, A, B) , consisting of a simply connected domain D , with metric g smooth up to the boundary and two marked boundary points A, B , with analytic local co-ordinates, be given. The partition function Z_{SLE_κ} of chordal SLE_κ is:*

$$Z_{\text{SLE}_\kappa} = |\det_D|^{\otimes c} \otimes |T_A^* \partial D|^{\otimes h} \otimes |T_B^* \partial D|^{\otimes h} = \det(\Delta_D)^{-c} \cdot Z_h \equiv \langle \psi(A) \psi(B) \rangle ,$$

where

$$(28) \quad \begin{aligned} c &= 1 - \frac{3}{2} \cdot \frac{(\kappa - 4)^2}{\kappa} = -\frac{(\kappa - 6)(3\kappa - 8)}{2\kappa} \text{ is the central charge, and} \\ h &= h(\kappa) = \frac{6 - \kappa}{2\kappa} \text{ the highest-weight.} \end{aligned}$$

$\langle \rangle$ denotes the unnormalised correlator of the boundary condition changing operators, corresponding to the boundary field ψ , of weight h .

Let us comment on the above Theorem. The fundamental and central relation is (28), which ensures that the partition function really behaves as a section of a bundle. Note also the different meanings of \det .

Let us discuss the case $\kappa = 3$, briefly. According to formula (28) we get $c = 1/2$ and $h = h_{1;2} = 1/2$ which would correspond to the Ising model. The weight of the tensor is a half-order differential and the two-point function can be expressed by means of the Szegő kernel for spin $\frac{1}{2}$, i.e., for fermions. This little, but important, example should be seen as a general paradigm which tells us that a correlator with multiple insertions can be derived and expressed in terms of the Green’s function, as a Vandermonde determinant, as follows from Wick’s theorem.

7. Doob’s h -transform, Virasoro null-vectors and SLE

This is a key section as it gives a unified treatment of the directions initiated by Schramm, Lawler and Werner, and it further connects with the one started by Malliavin. Our theory relies on the “Virasoro Uniformisation” (VU), which in case of the disc, corresponds to the theory developed by A.A. Kirillov, D. Yur’ev [17]

and Y. Neretin. In addition they took also the underlying infinite Kähler geometry into account. We shall now build on [11].

Let us begin by briefly explaining the main idea. In Figure 9, we have illustrated the path space which is identical for all models from statistical mechanics, producing simple curves, e.g. domain walls. However, every such model contributes its own measure on the path space.

If we parametrise it, we get from the measure in the statistical mechanics model different families of probability densities, whose evolution we have again indicated. As our earlier discussion showed, it is the central charge, resp. κ which labels different measures.

Also, to every point corresponds a partition function, which is composed of the regularised determinant and the information on the boundary conditions. The measures on the path space and the transformation properties of this partition functions under the stochastic evolution, i.e., the lifted process, have to form a conformally invariant pair. In particular, the conformal image of any trajectory has to be again a martingale, up to reparametrisation. This requirement gives us the necessary link between the parameters involved, i.e., κ, c, h , but also explains the appearance of degenerate highest-weight representations of the Virasoro algebra.

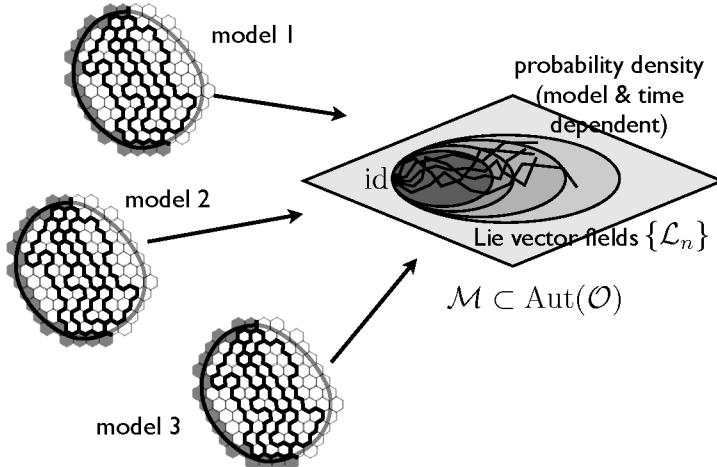


FIGURE 9. Different models with different Gibbs measures produce different probability densities on the same path space in $\mathcal{M} \subset \text{Aut}(\mathcal{O})$. The time evolution of the κ -densities is governed by a second-order differential operator in Hörmander form, given in terms of the explicit Lie fields \mathcal{L}_n .

7.0.1. Path space and Weil-Petersson metric. Recall the space \mathcal{M} . An element of it can be written as $c_1z + c_2z^2 + c_3z^3 + \dots$, with $c_1 > 0$. Now, if we fix a parametrisation t of the traces in our domain, we can cut them accordingly, and obtain so a family of univalent functions of the form

$$f(t, z) = e^{\gamma(t)}z + c_2(t)z^2 + \dots,$$

with $\gamma(t)$ a strictly monotone real function reflecting the path along which we are cutting. Further, it follows from the monotonicity that every simple path gives again

a simple path in \mathcal{M} . Also, any two different traces in a domain give two different traces in \mathcal{M} , i.e., the map is injective, and unique up to reparametrisation. This can be seen as follows.

The traces are compact subsets of the domain, which is a Hausdorff space. If two traces, with the exception of the endpoints are different, there are at least two interior points which are different. Hence, we can choose open disjoint neighbourhoods. The family of mappings for one of the curves at some instance maps the point onto the boundary, for the other never. Then, it follows from the Open Mapping theorem and the Identity Theorem for analytic functions that the corresponding traces in $\text{Aut}_+(\mathcal{O})$ must have points which are not the same, which in turns shows injectivity.

So, we have that \mathcal{M} is a subset of the semi-direct product $\mathbb{R}_+ \ltimes \text{Aut}_+(\mathcal{O})$, where

$$\text{Aut}_+(\mathcal{O}) := z \left(1 + \sum_{k=1}^{\infty} c_k z^k \right),$$

and it is enough to study the traces in $\text{Aut}_+(\mathcal{O})$. We shall think of the spaces as completed with respect to the natural filtration. The respective Lie groups and algebras are given below:

$$\begin{array}{ll} \text{Aut}_+(\mathcal{O}) & \text{Der}_+(\mathcal{O}) = z^2 \mathbb{C}[[z]] \partial_z \\ \cap & \cap \\ \text{Aut}(\mathcal{O}) & \text{Der}_0(\mathcal{O}) = z \mathbb{C}[[z]] \partial_z \\ \cap & \cap \\ & \text{Der}(\mathcal{O}) = \mathbb{C}[[z]] \partial_z \end{array}$$

and the exponential map $\exp : \text{Der}_+(\mathcal{O}) \rightarrow \text{Aut}_+(\mathcal{O})$, being an isomorphism. Further, the group $\text{Aut}(\mathcal{O})$ acts on itself by composition.

Now, this space has a natural affine structure with co-ordinates $\{c_k\}$, and the identity map corresponding to the origin 0.

Let us introduce the following spaces, which are in a sense dual, as they represent the power series developments around infinity, namely

$$\begin{aligned} \text{Aut}(\mathcal{O}_\infty) &:= \{ bz + b_0 + \frac{b_1}{z} + \dots, \quad b \neq 0 \} . \\ (29) \quad \text{Aut}_+(\mathcal{O}_\infty) &:= \{ z + b_0 + \frac{b_1}{z} + \dots, \} . \end{aligned}$$

Now, the complex Virasoro algebra $\text{Vir}_{\mathbb{C}}$ we are considering is spanned by the polynomial vector fields $e_n = -ie^{in\theta} \frac{d}{d\theta}$, $n \in \mathbb{Z}$, and \mathfrak{c} , with commutation relations $[\mathfrak{c}, e_n] = 0$ and

$$[e_m, e_n] = [e_m, e_n] + \omega_{c,h}(e_m, e_n) \cdot \mathfrak{c} ,$$

with the extended Gelfand-Fuks cocycle

$$\omega_{c,h}(v_1, v_2) := \frac{1}{2\pi} \int_0^{2\pi} \left((2h - \frac{c}{12}) v_1'(\theta) - \frac{c}{12} v_1'''(\theta) \right) v_2(\theta) d\theta ,$$

and v_1, v_2 being complex valued vector fields on S^1 . It has been shown [17], that there exists a two-parameter family of Kähler metrics on this space, with the form

being at the origin

$$(30) \quad w_{c,h} := \sum_{k=1}^{\infty} \left(2hk + \frac{c}{12}(k^3 - k) \right) dc_k \wedge d\bar{c}_k ,$$

building on the Gelfand-Fuks cocycle. This metrics also generate the Weil-Petersson (WP) metric. So, we can measure “the distance of two SLE curves”, i.e., their shapes, and more generally, do Differential Geometry on the path space.

7.0.2. *Analytic line bundles $E_{c,h}$ over \mathcal{M} .* The complex vector fields e_n on the circle, which form the Witt algebra, have a representation in terms of the Lie fields \mathcal{L}_{e_n} which act transitively on $\text{Aut}_+(\mathcal{O})$.

However, to have an action of $\text{Vir}_{\mathbb{C}}$, one has to introduce a determinant line bundle. So, on the infinite complex manifold $\text{Aut}_+(\mathcal{O})$ there exists an analytic line bundle $E_{c,h}$ which carries a transitive action of the Virasoro algebra. The line bundle $E_{c,h}$ is in fact trivial, with total space $E_{c,h} = \text{Aut}_+(\mathcal{O}) \times \mathbb{C}$. We can parametrise it by pairs (f, λ) , where f is a univalent function and $\lambda \in \mathbb{C}$.

It carries the following action

$$(31) \quad L_{v+\tau\mathbf{c}}(f, \lambda) = (\mathcal{L}_v f, \lambda \cdot \Psi(f, v + \tau\mathbf{c})) ,$$

where

$$(32) \quad \Psi_{c,h}(f, v + \tau\mathbf{c}) := h \oint \left[\frac{wf'(w)}{f(w)} \right]^2 v(w) \frac{dw}{w} + \frac{c}{12} \oint w^2 S(f, w) \frac{dw}{w} + i\tau c ,$$

and where

$$S(f, w) := \{f; w\} := \frac{f'''(w)}{f'(w)} - \frac{3}{2} \left(\frac{f''(w)}{f'(w)} \right)^2 ,$$

denotes the Schwarzian derivative of f with respect to w . The central element \mathbf{c} acts fibre-wise linearly by multiplication with $i\mathbf{c}$.

We have summarised the geometry in the following commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathbb{C} & \longrightarrow & \text{Vir}_{\mathbb{C}} & \longrightarrow & \text{Witt} & \longrightarrow 0 \\ & & \parallel & & \downarrow & & \downarrow & \\ 0 & \longrightarrow & \mathbb{C} & \longrightarrow & \Theta_{E_{c,h}} & \longrightarrow & \Theta_{\mathcal{M}} & \longrightarrow 0 \\ & & \parallel & & \downarrow & & \downarrow & \\ 0 & \longrightarrow & \mathbb{C} & \longrightarrow & E_{c,h} & \longrightarrow & \mathcal{M} & \longrightarrow 0 \end{array}$$

where Θ_X denotes the respective tangent sheaf. Generally, to have an action means to have a morphism from a Lie algebra \mathfrak{g} to the tangent sheaf, and transitive means that the map $\mathfrak{g} \otimes \mathcal{O}_X \rightarrow \Theta_X$ is surjective. Algebraically, the situation corresponds to so-called Harish-Chandra pairs (\mathfrak{g}, K) .

7.0.3. *Vir-modules and Verma modules $V(c, h)$.* One can endow the space of holomorphic sections $|\sigma\rangle \in \mathcal{O}(E_{c,h}) \equiv \Gamma(\text{Aut}_+(\mathcal{O}), E_{c,h})$ of the line bundle $E_{c,h}$ with a $\text{Vir}_{\mathbb{C}}$ -module structure.

Namely, let \mathcal{P} be the set of (co-ordinate dependent) polynomials on \mathcal{M} , defined by

$$P(c_1, \dots, c_N) : \text{Aut}(\mathcal{O})/\mathfrak{m}^{N+1} \rightarrow \mathbb{C} ,$$

with \mathfrak{m} the unique maximal ideal. In the classical literature on complex variables the above quotient space is called the “coefficient body”. \mathcal{P} corresponds to the

sections $\mathcal{O}(\mathcal{M})$ of the structure sheaf $\mathcal{O}_{\mathcal{M}}$ of \mathcal{M} and it carries an action of the representation of the Witt algebra in terms of the Lie fields $\mathcal{L}_n \equiv \mathcal{L}_{e_n}$.

In affine co-ordinates $\{c_n\}$, these fields can be written as [17], e.g.

$$\mathcal{L}_n = \frac{\partial}{\partial c_k} + \sum_{k=1}^{\infty} (k+1)c_k \frac{\partial}{\partial c_{n+k}} \quad n \geq 1 .$$

Now, $\mathcal{O}(E_{c,h})$ can be identified with the polynomial sections of the structure sheaf, i.e. with $\mathcal{O}(\mathcal{M})$, by choice of a free generator, and the previous trivialisation.

The action of $\text{Vir}_{\mathbb{C}}$ on sections of $E_{c,h}$ can then be written in co-ordinates, with the notation $\partial_n \equiv \frac{\partial}{\partial c_n}$, according to formulae (31, 32) as (cf. [17]),

$$\begin{aligned} L_n &= \partial_n + \sum_{k=1}^{\infty} c_k \partial_{k+n}, \quad n > 0 \\ L_0 &= h + \sum_{k=1}^{\infty} k c_k \partial_k , \\ L_{-1} &= \sum_{k=1}^{\infty} ((k+2)c_{k+1} - 2c_1 c_k) \partial_k + 2h c_1 , \\ (33) \quad L_{-2} &= \sum_{k=1}^{\infty} ((k+3)c_{k+2} - (4c_2 - c_1^2)c_k - a_k) \partial_k + h(4c_2 - c_1^2) + \frac{c}{2}(c_2 - c_1^2) , \end{aligned}$$

where the a_k are the Laurent coefficients of $1/f$, and c the central charge. The conformal invariance will naturally lead to highest-weight modules.

Now a polynomial $P(c_1, \dots, c_N) \in \mathcal{O}(E_{c,h})$ is a singular vector for $\{L_n\}$, $n \geq 1$, if

$$\left(\partial_k + \sum_{k \geq 1} (k+1)c_k \partial_{k+n} \right) P(c_1, \dots, c_N) = 0 .$$

Then, the highest-weight vector is the constant polynomial 1, and satisfies $L_0 \cdot 1 = h \cdot 1$, $\mathfrak{c} \cdot 1 = c \cdot 1$.

The dual $\mathcal{O}^*(E_{c,h})$, i.e. the space of linear functionals $\langle \sigma |$ on $\mathcal{O}(E_{c,h})$, can again be identified with $\mathcal{O}(\mathcal{M})$ via the pairing,

$$\langle P, Q \rangle := P(\partial_1, \dots, \partial_N) Q(c_1, \dots, c_N) |_{c_1 = \dots = c_N = 0} ,$$

and the action of Vir can then similarly be given explicitly in co-ordinates as in equations (33), with the roles of c_k and ∂_k interchanged.

Analogously, the dual space $\mathcal{O}^*(E_{c,h})$ can be endowed with a $\text{Vir}_{\mathbb{C}}$ action. The singular vector for the corresponding (irreducible) Verma module $V_{c,h}$ with respect to $\{L_{-n}\}$, $n \geq 1$, is again the constant polynomial 1.

7.0.4. Hypo-Ellipticity, sub-Riemannian Geometry and Conformal martingales. After this preparatory explanations, we are in position to make the link with SLE.

The Loewner equation (6) can be seen as a family of ordinary complex stochastic differential equations, labelled by the points of the upper half-plane \mathbb{H} . But as the individual processes are not independent, in the sense that at every instant t , they constitute a random conformal map for the points, where the individual processes are still defined, the interdependence can be translated into a power series expansion around the point which is the least affected in time, namely the point at infinity, or generally, the point at which the trace is aiming. So, we naturally get a random

dynamical system on the infinite coefficient body, i.e. a stochastic process $f_\infty(t)$ on univalent maps.

Now, by taking the projective limit we obtain the generator \hat{A}_∞ of the flow on $\text{Aut}_+(\mathcal{O}_\infty)$, corresponding to the Löwner equation (6) for some fixed κ , [6]:

$$(34) \quad \hat{A}_\infty = \varprojlim \left(\frac{\kappa}{2} \frac{\partial^2}{\partial b_1^2} + 2 \sum_{k=2}^N P_k(b_1, \dots, b_N) \frac{\partial}{\partial b_k} \right) ,$$

which is driven by one-dimensional standard Brownian motion, and where the polynomials P_k in the drift vector are defined on the coefficient body, with the $N \times N$ diffusion matrix (modulo κ)

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \end{pmatrix}$$

The generator of the diffusion process can be written in Hörmander form in terms of the tangent vector fields in the affine co-ordinates $\{b_k\}$, applying the notational conventions for operators acting on polynomial sections of the structure sheaf:

$$L_1^\infty := \frac{\partial}{\partial b_1}, \quad \text{and} \quad L_2^\infty := - \sum_{k=2}^{\infty} P_k(b) \frac{\partial}{\partial b_k} .$$

But as they satisfy the commutation relations of the Witt algebra, they span the whole tangent space and since we know that this Lie algebra acts transitively, the strong Hörmander condition is satisfied. Therefore, the resulting flow is hypoelliptic [18, 11], and the corresponding geometry sub-Riemannian.

Let us lift the process $f_\infty(t)$ on \mathcal{M} induced by SLE, to the complex manifold $E_{c,h}$ by using sections σ ; (cf. Figure 10). However, there are some constraints this map has to satisfy, which are given by the transformation properties of the partition function under conformal maps. This means, that the sections have to be flat with respect to the Hermitian connection $\nabla_{c,h}$ (cf. eq. (30)), and the physical connection (26), (energy-momentum tensor), which determines c and h , for a given model. Again, we have indicated this in Figure 10, where the group action which relates different points on the trajectory, corresponds to the transition functions of the Löwner chain, and in the total space to the conformal anomaly, which relates the points in the fibre, i.e., the values of the determinants over the respective domains. This is equivalent to the Markov property, which was foundational in the derivation of the driving function for SLE.

Still, there is one more thing. The random trajectories $\sigma_t := \sigma(f_\infty(t))$ should be (local) martingales. This is necessary, if we intend to couple the dynamical picture with the static one, i.e. “ensemble averages should be equal to time averages”.

To do so, we have to couple the parameters κ, c, h , which we shall obtain from the Doob-Getoor h -transform. Namely, if we find harmonic sections, $\sigma_{\text{hr.}}$, then $\sigma(f_\infty(t))$ is a local martingale, for the lifted, now Virasoro generators \hat{L}_n .

But, if we restrict to polynomial sections, generated by the \hat{L}_n , $n < 1$, by acting on the constant polynomial 1, which then successively extends over the coefficient body $\text{Aut}(\mathcal{O})/\mathfrak{m}^{N+1}$, then as a direct calculation shows, using e.g. (33), the module

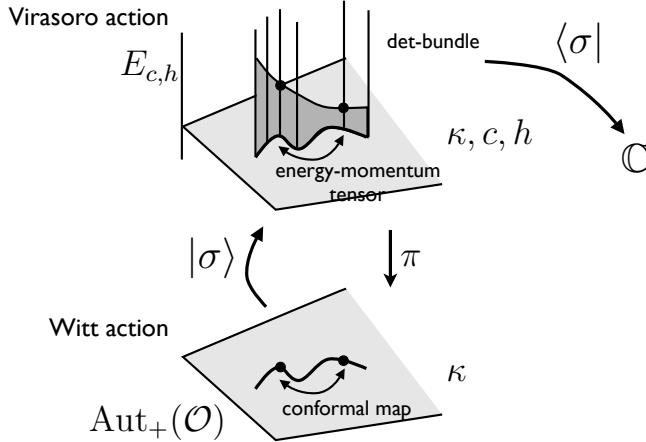


FIGURE 10. The stochastic process in the base $\text{Aut}_+(\mathcal{O})$ is lifted to the total space $E_{c,h}$ via the section $|\sigma\rangle$. The operators acting on the total space, depend on the parameters κ, c, h , coming from the partition function $Z_{c,h}$. The compatibility conditions, i.e., to be a martingale, require c, h to depend functionally on κ . $\langle\sigma|$ is the dual vector.

contains a null vector, exactly if

$$c_\kappa = \frac{(6 - \kappa)(3\kappa - 8)}{2\kappa} \quad \text{and} \quad h_\kappa = \frac{6 - \kappa}{2\kappa}.$$

But, this is nothing else than the relation in equation (28), derived from CFT. Therefore, all this polynomials are in the kernel of the lifted generator,

$$\frac{\kappa}{2} \hat{L}_1^2 - 2\hat{L}_2,$$

which acts as a differential operator.

Therefore, the representation theoretic notion of “being degenerate at level two”, translates in probabilistic language into a generalised Doob h -transform.

This is the way, how conformally invariant measures on simple paths and models from Statistical Mechanics couple, as demonstrated for the special case, the disc. But this is also valid in the general situation.

Conclusions. We hope that our leisurely promenade has convinced the reader of the beauty and richness of the interplay between the various fields from mathematics and physics.

The connections which emerged with probability theory are certainly one of the more remarkable facts of this decade. Also, the insight that determinant line bundles are a very rewarding starting point to look at SLE, has become, as a glance at the now emerging literature reveals, a ‘sine qua non’ in the field. However, it is amusing, that although fundamental parts of the material given here have been public for quite some time, all the sudden people start, e.g. to regularise “their determinants”.

Finally, it should be immediate that the theory which we presented, builds on an enormous quantity of other theories and mathematical tools, which have been established long before SLE, for other purposes.

Therefore, SLE is, according to our opinion, yet another good example for the restless “globalisation of mathematics” and its evolution into a highly integrated organism.

References

- [1] H. Airault, P. Malliavin and A. Thalmaier, *Canonical Brownian motion on the space of univalent functions and resolution of Beltrami equations by a continuity method along stochastic flows*, J. de Mathématiques Pures et Appliquées, Vol. **83**, Issue 8, (2004).
- [2] M. Aizenman, *The geometry of critical percolation and conformal invariance*, StatPhys 19 (Xiamen 1995), (1996).
- [3] M. Aizenman A. Burchard, *Hölder Regularity and Dimension Bounds for Random Curves*, Duke Math. J. **99**, 419 (1999).
- [4] S. Albeverio, J. Jost, S. Paycha and S. Scarlatti, *A mathematical introduction to string theory*, London Math. Soc. LNS **225**, Cambridge Univ. Press, (1997).
- [5] O. Alvarez, *Theory of Strings with Boundaries: Fluctuation, topology and quantum geometry*, Nucl. Phys. B **216**, (1983).
- [6] M. Bauer and D. Bernard, *2D growth processes: SLE and Loewner chains*, Phys. Rep. **432**, 115 (2006).
- [7] R. Bauer and R. Friedrich, *The Correlator Toolbox, Metrics and Moduli*, Nucl. Phys. B **733**, (2006).
- [8] A. A. Belavin, A. M. Polyakov and A. B. Zamolodchikov, *Infinite conformal symmetry in two-dimensional quantum field theory*, Vol. **241**, Issue 2, (1984).
- [9] J. Cardy, *Boundary Conditions, Fusion Rules and The Verlinde Formula*, Nucl. Phys. B **324**, (1989).
- [10] R. Friedrich, *The Global Geometry of Stochastic Læwner Evolutions*, Conference Proceedings: The 1st MSJ-SI, Probabilistic Approach to Geometry, (2008).
- [11] R. Friedrich, *On Connections of Conformal Field Theory and Stochastic Læwner Evolution*, arXiv, (2004).
- [12] R. Friedrich and J. Kalkkinen, *On Conformal Field Theory and Stochastic Loewner Evolution*, arXiv 2003, Nucl. Phys. B, **687**, 279-302 (2004).
- [13] R. Friedrich and W. Werner, *Conformal fields, restriction properties, degenerate representations and SLE*, C.R. Acad. Sci. Paris, Ser. I **335** (2002).
- [14] R. Friedrich and W. Werner, *Conformal restriction, highest-weight representations and SLE*, Comm. Math. Phys., **243**, (2003).
- [15] A. Gamsa and J. Cardy, *Schramm-Loewner evolution in the three-state Potts model—a numerical study*, JSTAT, (2007).
- [16] A. Hassell and S. Zelditch, *Determinants of Laplacians in Exterior Domains*, IMRN, Int. Math. Res. Notices, No. **18**, (1999).
- [17] A.A. Kirillov and D. Yur'ev, *Representation of the Virasoro algebra by the orbit method*, JGP, Vol. 5, n. **3**, (1988).
- [18] M. Kontsevich, *Arbeitstagung 2003-CFT, SLE and phase boundaries*, MPIM2003-60a, (2003).
- [19] M. Kontsevich and Y. Suhov, *On Malliavin measures, SLE, and CFT*, Proceedings of the Steklov Institute of Mathematics, Vol. 258(**1**), (2007).
- [20] R. Langlands, P. Pouliot, and Y. Saint-Aubin, *Conformal invariance in two-dimensional percolation*. Bul l. Amer. Math. Soc., **30**(1), (1994).
- [21] G. Lawler, O. Schramm and W. Werner, *Conformal restriction. The chordal case*, J. Amer. Math. Soc. **16**, (2003).
- [22] P. Malliavin, *La diffusion canonique au-dessus du groupe des difféomorphismes du cercle*, C.R. Acad. Sci. Paris, Vol. **329**, Issue 4, (1999).
- [23] S. Nag and D. Sullivan, *Teichmüller theory and the universal period mapping via quantum calculus and the $H^{1/2}$ space on the circle*, Osaka J. Math. **32** (1995).

- [24] B. Osgood, R. Philips and P. Sarnak, *Extremals of Determinants of Laplacians*, and, *Compact Isospectral Sets of Surfaces*, J. Funct. Anal., **80**, (1988).
- [25] A. Polyakov, *Quantum geometry of bosonic strings*, Phys. Lett. B, **103** (1981).
- [26] S. Rohde and O. Schramm, *Basic properties of SLE*, Ann. Math. **161**, (2005).
- [27] O. Schramm, Scaling limits of loop-erased random walks and uniform spanning trees, Israel J. Math. **118**, (2000).
- [28] S. Smirnov, *Towards conformal invariance of 2D lattice models*, Proc. Int. Congr. Mathematicians, vol. 2, Eur. Math. Soc., pp. 1421-51, (2006).
- [29] N. Kawamoto, Y. Namikawa, A. Tsuchiya, and Y. Yamada, *Geometric Realization of Conformal Field Theory on Riemann Surfaces*, Commun. Math. Phys. **116**, 247-308, (1988).
- [30] A.B. Zamolodchikov, “Irreversibility” of the flux of the renormalization group in a 2D field theory, JETP Lett. **43**, (1986).

MPI, BONN

E-mail address: `friedrich@cbs.mpg.de`

This page intentionally left blank

On the causal gauge principle

José M. Gracia-Bondía

ABSTRACT. Work by the Zürich school of causal (Epstein–Glaser) renormalization has shown that renormalizability in the presence of massless or massive gauge fields (as primary entities) explains gauge invariance and, in some instances, the presence of a Higgs-like particle, without the Brout–Englert–Higgs–Kibble (BEHK) mechanism. We review it, in a pedagogical vein, with a pointer to go beyond.

CONTENTS

| | |
|--|-----|
| 1. Introduction | 115 |
| 2. Overview of the CGI method | 116 |
| 3. The abelian model | 119 |
| 4. Three MVBS | 128 |
| 5. The Weinberg–Salam model within CGI | 129 |
| 6. Discussion | 131 |
| References | 132 |

1. Introduction

By now spontaneous symmetry breaking (SSB) of local symmetry is a well-established paradigm of high-energy physics. At the end of the 60s and beginning of the 70s, it allowed the incorporation of (electro)weak interactions into the framework of renormalizable field theory. In connection with the contemporaneous rise of the Standard Model (SM), it enjoys immense historical success.

However, allusion to unsatisfactory or mysterious aspects of the Higgs sector of the SM does pop up in the literature —see for instance [1, Sect. 22.10]. The Higgs self-coupling terms are completely ad-hoc, unrelated to other aspects of the theory, and do not seem to constitute a gauge interaction. Moreover they raise the hierarchy problem [2, Ch. 11]. The most frequent interpretation of the BEHK mechanism clashes with cosmology [3].

Debate on the proper interpretation of the mechanism (whether the symmetry is “broken” or just “hidden”, whether the Higgs field truly has a non-zero vacuum expectation value (VEV) or not [4], and so on) seems endless. This breeds some

skepticism, even among earlier and doughty practitioners. At the end of his Nobel lecture [5], Veltman chose to declare: “*While theoretically the use of spontaneous symmetry breakdown leads to renormalizable Lagrangians, the question of whether this is really what happens in Nature is entirely open*”.

Indeed, since the *deus ex machina* fields involved in broken or hidden symmetry are unobservable, the status question for the BEHK contraption cannot be resolved by the likely sighting of the Higgs particle in the LHC.¹

The subject has also been obscured all along by theoretical prejudice. In the SM the Higgs field carries the load of giving masses to *all* matter and force fields. For instance, it is said that mass terms for the vector bosons are incompatible with gauge invariance. It ain’t so: such mass terms fit in gauge theory by use of Stückelberg fields [8, 9].

Skepticism would be idle, nevertheless, in the absence of alternative theoretical frameworks. Assuming an agnostic stance, we pose the question: is it possible to formulate the main results of flavourdynamics, and to frame suggestions of new physics, without recourse to unobservable processes? In tune with the phenomenological SM Lagrangian [10], this amounts to regard massive vector bosons (MVB) as fundamental entities.

So let us stop pretending we know the origins of mass. Higgs-like scalar fields will still come in handy for either renormalizability or unitarity; however, their gauge variations need not be the conventional ones. Fermions can be assigned Dirac masses, and couplings with the scalar field proportional to those; this contradicts in no way the chiral nature of their interactions in the SM.

An approach with the mentioned traits is already found in the literature in the work by Scharf, Dütsch and others, under the label of the “quantum gauge invariance” principle. A few references to it are [11–14] and mainly the book [15]. The “quantum Noether principle” of [16, 17] coincides essentially with it. Both are based on the rigorous causal scheme for renormalization [18] by Epstein and Glaser (EG).

Henceforth we refer to the approach as causal gauge invariance (CGI). The usual plan of the article is found at the end of the next section, when the stakes hopefully have been made clearer.

2. Overview of the CGI method

The spirit of CGI is very much that of the [19]. Let s denote the nilpotent BRS operation. To realize gauge symmetry, one should incorporate BRS symmetry ab initio in a “quantum” Lagrangian \mathcal{L} , such that (very roughly speaking) $s\mathcal{L} \sim 0$, and proceed to build from there. We do this for MVBs.

The starting point for the analysis is the Bogoliubov–Epstein–Glaser functional scattering matrix on Fock space, in the form of a power series:

$$(1) \quad \mathbb{S}(g) = 1 + \sum_{n=1}^{\infty} \frac{i^n}{n!} \int dx_1 \dots dx_n T_n(x_1, \dots, x_n) g(x_1) \dots g(x_n).$$

The coupling constants of the model are replaced by test functions —we wrote just one of them for simplicity. The theory is then constructed basically by using

¹This situation has recently called the attention of knowledgeable philosophers of science [6, 7]: in epistemological terms, they argue that the mechanism had heuristic value in the context of discovery; but much less so in the context of justification.

causality and Poincaré invariance to recursively determine the form of the time-ordered products T_n from the T_m with $m < n$; in this sense the procedure is inverse to the “cutting rules”. Only those fields should appear in T_n that already are present in T_1 . The procedure yields a finite perturbation theory without regularization; ultraviolet divergences are avoided by proper definition of the n -point functions as distributions.

(Ultimately one would be interested in the adiabatic limit $g(x) \uparrow g$. This is delicate, however, due to infrared problems. We look at the theory before that limit is taken.)

With the proviso that two forms of T_n are equivalent if they differ by s -coboundaries, CGI is formulated by the fact that sT_n must be a divergence. Roughly speaking, we must have

$$(2) \quad \begin{aligned} sT_n(x_1, \dots, x_n) &= i \sum_{l=1}^{l=n} T[T_1(x_1), \dots, \partial_l \cdot Q(x_l), \dots, T_1(x_n)], \\ &=: i \sum_{l=1}^{l=n} \partial_l \cdot Q_n(x_1, \dots, x_n). \end{aligned}$$

for vectors Q_n , called Q -vertices, with ∂_l denoting the partial divergences with respect to the x_l coordinates and T a time-ordering operator. In this way renormalization and gauge invariance are linked in the EG scheme. (We said “roughly” because (2) suggests that T and spacetime derivatives commute, which is not generally the case for on-shell fields.)

Note that T_1 only contains the first-order part of the Lagrangian. Nevertheless, already the first order condition

$$sT_1 = i\partial \cdot Q_1$$

constrains significantly the form of the Lagrangian. Later on, we show leisurely how the CGI method works for tree graphs belonging to T_2 . This is almost all what is required for the purposes of this paper: for ordinary gauge theories, the treatment of T_3 is pretty simple, and higher orders not needed at all.

Keep in mind that one works here with *free* fields. Interacting fields can be arrived at in the Epstein–Glaser procedure, somewhat *a posteriori*, using their definition by Bogoliubov as logarithmic functional derivatives of $\mathbb{S}(g)$ with respect to appropriate sources. Their gauge variations resemble more those of standard treatments; but we do not use them. Thus s “sees” only the (massive or massless) gauge fields, and the attending (anti-)ghost and Stückelberg fields. This is why everything flows from the quantum gauge structure of the boson sector. Coupling to fermions, which ought not be organized in multiplets *a priori*, comes almost like an afterthought.

As it turns out, the procedure is quite restrictive, and in particular only a few models for MVB theory pass muster. These exhibit very definite mass and interaction patterns, in particular quartic self-interaction for the scalar particles.

We next compile the results.

Consider then a model with t intermediate vector bosons A_a in all, of which any may be in principle massive or massless. Let us say there are r massive ones with masses m_a , $1 \leq a \leq r$ and s massless ones, and $t = r + s$. We assume there is *one* (at most) physical scalar particle H of mass m_H : *entes non sunt multiplicanda*

praeter necessitatem. The BRS extension of the Wigner representation theory for MVBs requires Stückelberg fields B_a [13], beyond the fermionic ghosts u_a, \tilde{u}_a ; in case A_a is massless, we of course let B_a drop out. Adopting the Feynman gauge, the gauge variations are as follows:

$$\begin{aligned}
 sA_a^\mu(x) &= i\partial^\mu u_a(x); \\
 sB_a(x) &= im_a u_a(x); \\
 su_a(x) &= 0; \\
 s\tilde{u}_a(x) &= -i(\partial \cdot A_a(x) + m_a B_a(x)). \\
 (3) \quad sH(x) &= 0.
 \end{aligned}$$

This operator is nilpotent on-shell.

The total bosonic interaction Lagrangian, in a notation close to that of [15], is of the form

$$(4) \quad \mathcal{L}_{\text{int}} = gT_1 + \frac{g^2 T_2}{2},$$

where g is an overall dimensionless coupling constant;

$$T_1 = f_{abc}(T_{1abc}^1 + T_{1abc}^2 + T_{1abc}^3 + T_{1abc}^4) + C(T_1^5 + T_1^6 + T_1^7 + T_1^8 + T_1^9)$$

includes the cubic couplings, and

$$T_2 = T_2^1 + T_2^2 + T_2^3 + T_2^4 + T_2^5 + T_2^6 + T_2^7$$

includes the quartic ones. The list of cubic couplings not involving H is given by:

$$\begin{aligned}
 T_{1abc}^1 &= [A_a \cdot (A_b \cdot \partial) A_c - u_b (A_a \cdot \partial \tilde{u}_c)]; \\
 T_{1abc}^2 &= \frac{m_b^2 + m_c^2 - m_a^2}{4m_b m_c} [B_b (A_a \cdot \partial B_c) - B_c (A_a \cdot \partial B_b)]; \\
 T_{1abc}^3 &= \frac{m_b^2 - m_a^2}{2m_c} (A_a \cdot A_b) B_c; \\
 (5) \quad T_{1abc}^4 &= \frac{m_a^2 + m_c^2 - m_b^2}{2m_c} \tilde{u}_a u_b B_c;
 \end{aligned}$$

The list of cubic couplings of the Higgs-like particle is:

$$\begin{aligned}
 T_1^5 &= m_a [B_a (A_a \cdot \partial H) - H (A_a \cdot \partial B_a)]; \\
 T_1^6 &= m_a^2 (A_a \cdot A_a) H; \\
 T_1^7 &= -m_a^2 \tilde{u}_a u_a H; \\
 T_1^8 &= -\frac{1}{2} m_H^2 B_a^2 H; \\
 (6) \quad T_1^9 &= -\frac{1}{2} m_H^2 H^3.
 \end{aligned}$$

Remarks: in (5) and (6) we sum over repeated indices; the f_{abc} are completely skewsymmetric in their three indices, and fulfil the Jacobi identity; T_1^1 yields the cubic part in the classical Yang–Mills Lagrangian; C is a constant independent of a . The dimension of the Lagrangian must be M^4 in natural units, and the boson field dimension in our formulation is 1 for *both* spins: the dimension of C is M^{-1} . Note the diagonality of the couplings of the Higgs-like particle. Crossed terms like $(A_a \cdot A_b)H$ for $a \neq b$, and others like $B_a B_b B_c, B_a H^2 \dots$, that could be envisaged, are held to vanish by CGI.

The list of quartic couplings:

$$\begin{aligned}
 T_2^1 &= -\frac{1}{2}f_{abc}f_{ade}(A_b \cdot A_d)(A_c \cdot A_e); \\
 T_2^2 &= \left[\frac{(m_d^2 + m_e^2 - m_a^2)(m_c^2 + m_e^2 - m_b^2)}{4m_d m_c m_e^2} f_{ade} f_{bce} + C^2 m_a m_b \delta_{ad} \delta_{bc} \right] \\
 &\quad \times (A_a \cdot A_b) B_c B_d; \\
 T_2^3 &= -\frac{1}{4}C^2 m_H^2 B_a^2 B_b^2 \quad \text{irrespective of } a, b \leq r; \\
 T_2^4 &= C f_{abc} \frac{m_b^2 - m_a^2}{m_c} (A_a \cdot A_b) B_c H; \\
 T_2^5 &= C^2 m_a^2 (A_a \cdot A_a) H^2; \\
 T_2^6 &= -\frac{1}{2}C^2 m_H^2 B_a^2 H^2 \quad \text{irrespective of } a \leq r; \\
 (7) \quad T_2^7 &= -\frac{1}{4}C^2 m_H^2 H^4.
 \end{aligned}$$

The foregoing lists can be gleaned from [15]; every coefficient of the interaction Lagrangian is in principle determined in terms of the f_{abc} and the pattern of masses. We are not through, because CGI implies *constraints*, in general non-linear and extremely restrictive, on *allowed patterns* of masses for the gauge fields. But we may anticipate a few more comments. The first term T_2^1 in (7) just yields the quartic part in the classical Yang–Mills Lagrangian, as expected. In case all the A_a are massless, there is no need to add physical or unphysical scalar fields for renormalizability, and only T_1^1 and T_2^1 survive in the theory; they of course coincide respectively with the first and second order part of the usual Yang–Mills Lagrangian. In particular, CGI gives rise to gluodynamics. (It must be said, though, that the physical equivalence of couplings differing in a divergence is less compelling in this case, since there is no asymptotic limit for the Bogoliubov–Epstein–Glaser $\mathbb{S}(g)$ -matrix; CGI offers no tools to deal with this infrared problem.) Remarkably, with independence of the masses, CGI unambiguously leads to generalized Yang–Mills theories on reductive Lie algebras; apparently this was realized first by Stora [20].

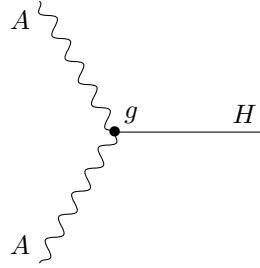
The plan of the rest of the article is as follows. Notice that the case $r = 1, s = 0$ leads to an abelian model in which all the terms with the Higgs-like field H survive. We use this example in Section 3 to illustrate in some detail —missing in [15]— how the second-order condition determines the couplings. Section 4 deals with *three* gauge fields —there are no models with two gauge fields to speak of, since $\mathfrak{u}(1) \oplus \mathfrak{u}(1)$ is the only two-dimensional reductive Lie algebra. For that we need to invoke the mentioned *mass relations* (reference [15] unfortunately contains misprints in this respect). Section 5 elaborates on the reconstruction of the SM in CGI, looking at the fermion sector as well. The paper ends with a discussion.

3. The abelian model

Consider a theory with a neutral gauge field A of mass m and a *physical* neutral scalar field H of mass m_H , and basic coupling AAH .

Since massive quantum electrodynamics is known to be renormalizable without an extra scalar field, this is perhaps not very interesting; but our aim here is merely showing the workings of the causal gauge principle.

3.1. The first-order analysis. For T_1 , take the most general Ansatz containing cubic terms in the fields and leading to a renormalizable theory. With the



benefit of hindsight, we write down on the first line the terms destined to survive:

$$(8) \quad \begin{aligned} T_1/m = & (A \cdot A)H + b\tilde{u}uH + c(H(A \cdot \partial B) - B(A \cdot \partial H)) + dB^2H + eH^3 \\ & + a(A \cdot A)B + b_2\tilde{u}uB + b_3u(A \cdot \partial \tilde{u}) + d_1B^3 + d_3BH^2. \end{aligned}$$

The factor m is natural according to our previous discussion on dimensions. The symmetric combination $HA \cdot \partial B + BA \cdot \partial H$ has been excluded for the following reason:

$$A \cdot (B\partial H + H\partial B) = \partial \cdot (BHA) - (\partial \cdot A)BH,$$

and in view of (3), the $(\partial \cdot A)BH$ term is s -exact apart from terms of already present in (8). Concretely,

$$s(\tilde{u}BH) = -(\partial \cdot A)BH - mB^2H - m\tilde{u}uH.$$

We calculate next sT_1/m in (8) and obtain for the first group of terms:

$$(9) \quad \begin{aligned} & 2\partial \cdot (uHA) - 2u(\partial \cdot A)H - 2uA \cdot \partial H - bu(\partial \cdot A)H \\ & - bmuBH + c\partial \cdot u(H\partial B - B\partial H) \\ & + cm[HA \cdot (\partial u) - uA \cdot \partial H] + 2dmuBH. \end{aligned}$$

We have used $s(uC) = -usC$ for any C . In detail:

$$-is(A \cdot AH) = (\partial u \cdot A)H = 2[\partial \cdot (uHA) - u(\partial \cdot A)H - uA \cdot \partial H].$$

Next

$$-is(\tilde{u}uH) = -u(\partial \cdot A)H - muBH.$$

Next

$$-is(A \cdot (H\partial B - B\partial H)) = \partial u \cdot (H\partial B - B\partial H) + m[HA \cdot (\partial u) - uA \cdot \partial H].$$

Finally $-is(B^2H) = 2muBH$.

Similarly, for the second group of terms we obtain:

$$\begin{aligned} & 2a\partial \cdot (uBA) - 2au(\partial \cdot A)B - 2auA \cdot \partial B + amuA \cdot A \\ & - b_2u(\partial \cdot A)B - b_2muB^2 + b_3(\partial u \cdot u \partial \tilde{u} + uA \cdot \partial(\partial \cdot A + mB)) \\ & + 3d_1muB^2 + d_3muH^2. \end{aligned}$$

All terms of that group are excluded because their contributions to sT_1 are not pure divergences. For instance, the first one corresponds to the term in $uA \cdot A$, that can be canceled only by setting $a = 0$.

On the other hand, the second term in the second line in (9) can be recast as

$$\partial \cdot (u(H\partial B - B\partial H)) + (m^2 - m_H^2)uBH.$$

For the following terms we have

$$A \cdot (\partial u)H - uA \cdot \partial H = \partial \cdot (uHA) - u(\partial \cdot A)H - 2uA \cdot \partial H.$$

In all,

$$\begin{aligned} -isT_1/m &= \partial \cdot (C + D) - (2 + cm + b)u(\partial \cdot A)H \\ &\quad - (2 + 2cm)u(A \cdot \partial H) + (2dm - bm + c(m^2 - m_H^2))uBH; \end{aligned}$$

with the vectors C, D given by $C := (2 + cm)uHA$; $D := cu(H\partial B - B\partial H)$. The terms that are not a divergence must cancel. This at once leads to:

$$c = -\frac{1}{m}; \quad b = -1; \quad d = -\frac{m_H^2}{2m^2}; \quad \text{thus} \quad C = uHA; \quad D = \frac{-u}{m}(H\partial B - B\partial H).$$

In summary, we have obtained the cubic couplings in the Lagrangian:

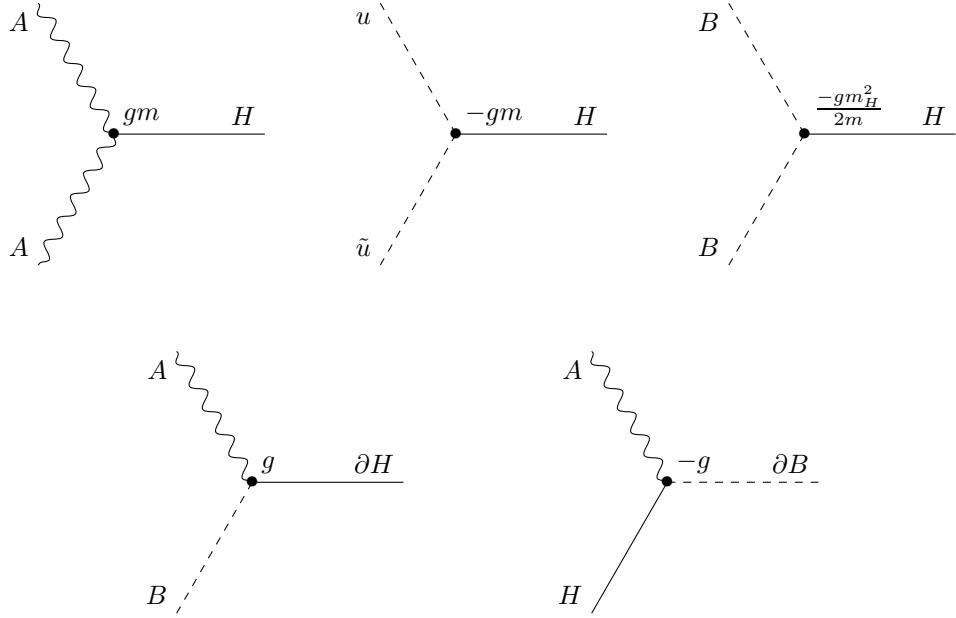


FIGURE 1. cubic vertices

To this we should add the H^3 coupling, whose coefficient is still indeterminate. Moreover:

$$(10) \quad sT_1 = i\partial \cdot Q_1 \quad \text{with} \quad Q_1 = muHA - u(H\partial B - B\partial H).$$

3.2. The second-order analysis. The next step is less trivial. Equation (2) certainly makes sense outside the diagonals, for then the T product is calculated like an ordinary product. But the extension to the diagonals, which is simply $x_1 = x_2$ for $n = 2$, can produce local correction terms. At this order, the advanced and retarded products are given by:

$$(11) \quad \begin{aligned} A_2(x_1, x_2) &= T_2(x_1, x_2) - T_1(x_1)T_1(x_2); \\ R_2(x_1, x_2) &= T_2(x_1, x_2) - T_1(x_2)T_1(x_1); \end{aligned}$$

Here $T_2(x_1, x_2)$ is still unknown, but it is clear that A_2 will have support on the past light cone of x_2 , and R_2 on its future light cone; hence the nomenclature. Consider then $D_2(x, y) := (R_2 - A_2)(x, y) = [T_1(x), T_1(y)]$, whose support is within the light cone (we say D_2 is causal). We have thus

$$(12) \quad \begin{aligned} sD_2(x, y) &= [sT_1(x), T_1(y)] + [T_1(x), sT_1(y)] \\ &= i\partial_x [Q_1(x), T_1(y)] + i\partial_y [T_1(x), Q_1(y)]; \end{aligned}$$

so that D_2 moreover is gauge-invariant. The crucial step in EG renormalization is the *splitting* of D_2 into the retarded part R_2 and the advanced part A_2 ; once this is done, T_2 is found at once from (11). The issue is how to preserve gauge invariance in this distribution splitting. For this, we split D_2 and the commutators—with the derivatives—in the previous equation; then gauge invariance:

$$sR_2(x, y) = i\partial_x R_{2/1}(x, y) + i\partial_y R_{2/2}(x, y)$$

can only be (and is) violated for $x = y$, that is, by local terms in $\delta(x - y)$. However, if in turn local renormalization terms $N_2, N_{2/1}, N_{2/2}$ can be found in such a way that

$$s(R_2(x, y) + N_2(x, y)) = i\partial_x(R_{2/1} + N_{2/1}) + i\partial_y(R_{2/2} + N_{2/2}),$$

with an obvious notation, then CGI to second order holds.

To the purpose we consider only tree diagrams. In view of (12), we systematically proceed to study the divergences coming from cross-terms between (10) and

$$(13) \quad T_1 = m \left[(A \cdot A)H + u\bar{u}H - \frac{1}{m} A \cdot (H\partial B - B\partial H) - \frac{m_H^2}{2m^2} B^2 H + eH^3 \right].$$

Factors containing derivatives give rise to normalization contributions after distribution splitting.

The most difficult part of the coming calculation asks for divergences of terms with commutators $[\partial^\mu B(x), \partial^\nu B(y)]$ and $[\partial^\mu H(x), \partial^\nu H(y)]$. Following [21], we look at Section 4 in [11] in order to prepare the computation. There, for general functions F, E we find the formulas:

$$(14) \quad \begin{aligned} \partial_\mu^x [F(x)E(y)\delta(x - y)] + \partial_\mu^y [F(y)E(x)\delta(x - y)] \\ = \partial_\mu F(x)E(x)\delta(x - y) + F(x)\partial_\mu E(x)\delta(x - y) \end{aligned}$$

$$(15) \quad \begin{aligned} \text{and } F(x)E(y)\partial_\mu^x\delta(x - y) + F(y)E(x)\partial_\mu^y\delta(x - y) \\ = F(x)\partial_\mu E(x)\delta(x - y) - \partial_\mu F(x)E(x)\delta(x - y). \end{aligned}$$

We may prove both from the following observation: since

$$F(x)E(y)\delta(x - y) = F(x)E(x)\delta(x - y),$$

it must be that

$$\partial_\mu^x (F(x)E(y)\delta(x - y)) = \partial_\mu^x (F(x)E(x)\delta(x - y));$$

which forces

$$(16) \quad E(y)\partial_\mu^x\delta(x - y) = E(x)\partial_\mu^x\delta(x - y) + \partial_\mu E(x)\delta(x - y).$$

Now,

$$\begin{aligned} \partial_\mu^x [F(x)E(y)\delta(x - y)] + \partial_\mu^y [F(y)E(x)\delta(x - y)] \\ = \partial_\mu F(x)E(x)\delta(x - y) + F(x)E(y)\partial_\mu^x\delta(x - y) \\ + \partial_\mu F(x)E(x)\delta(x - y) - F(y)E(x)\partial_\mu^x\delta(x - y) \\ = \partial_\mu F(x)E(x)\delta(x - y) + F(x)E(y)\partial_\mu^x\delta(x - y) \\ - F(x)E(x)\partial_\mu^x\delta(x - y) = \partial_\mu F(x)E(x)\delta(x - y) + F(x)\partial_\mu E(x)\delta(x - y); \end{aligned}$$

where we have used (16) twice. Analogously,

$$\begin{aligned} F(x)E(y)\partial_\mu^x\delta(x - y) + F(y)E(x)\partial_\mu^y\delta(x - y) &= F(x)E(x)\partial_\mu^x\delta(x - y) \\ &+ F(x)\partial_\mu E(x)\delta(x - y) - F(y)E(x)\partial_\mu^x\delta(x - y) \\ &= F(x)\partial_\mu E(x)\delta(x - y) - \partial_\mu F(x)E(x)\delta(x - y), \end{aligned}$$

using (16) twice again.

We finally start the advertised computation. Coming from respectively the second term of $Q_1(x)$ in (10) and third of $T_1(y)$ in (13), now we find for $i[Q_1(x), T_1(y)]$:

$$\begin{aligned} &iu(x)H(x)[\partial^\mu B(x), \partial^\nu B(y)]A_\nu(y)H(y) \\ &= u(x)H(x)A_\nu(y)H(y)\partial_x^\mu\partial_y^\nu D(x-y). \end{aligned}$$

The identity $[B(x), B(y)] = -iD(x-y)$ for scalar fields has been employed. Next we need to tackle the divergence of the splitting of $\partial_x^\mu\partial_y^\alpha D$. Splitting of the Jordan–Pauli propagator D gives rise to the retarded propagator D^{ret} . Now, each derivation increases by one the singular order of a distribution. Thus, although $\partial_x^\mu\partial_y^\nu D^{\text{ret}}$ is a well-defined distribution, its singular order is $-2 + 2 = 0$, therefore allowing a normalization term in the split distribution:

$$\partial_x^\mu\partial_y^\nu D^{\text{ret}}(x-y) \rightarrow \partial_x^\mu\partial_y^\nu D^{\text{ret}}(x-y) + C_B g^{\mu\nu}\delta(x-y).$$

After applying ∂_μ , simply from

$$\partial_\mu\partial_x^\mu D^{\text{ret}}(x-y) = -m^2 D^{\text{ret}}(x-y) + \delta(x-y),$$

the total singular part is of the form

$$C_B\partial_x^\nu[F(x)E(y)\delta(x-y)] + F(x)E(y)\partial_y^\nu\delta(x-y), \quad \text{with } F = uH; E = HA_\nu.$$

Adding the term with x and y interchanged, and using the identities (14) and (15), it comes finally the short rule for this kind of singular term:

$$F(x)E(y)\partial_x^\mu\partial_y^\nu D(x-y) \rightarrow [(C_B + 1)(\partial^\nu F)E + (C_B - 1)F\partial^\nu E]\delta(x-y).$$

Therefore we obtain in the end

$$\begin{aligned} (17) \quad &(C_B + 1)[H^2(A \cdot \partial u) + uH(A \cdot \partial H)]\delta(x-y) \\ &+ (C_B - 1)[uH^2(\partial \cdot A) + uH(A \cdot \partial H)]\delta(x-y). \end{aligned}$$

By the same token, coming now from respectively the third and fourth terms in $Q_1(x)$ and $T_1(y)$, and performing entirely similar operations, we obtain

$$\begin{aligned} (18) \quad &(C_H + 1)[B^2(A \cdot \partial u) + uB(A \cdot \partial B)]\delta(x-y) \\ &+ (C_H - 1)[uB^2(\partial \cdot A) + uB(A \cdot \partial B)]\delta(x-y). \end{aligned}$$

There is no good reason for $C_H \neq C_B$; see further on.

There are no singular contributions from the first term in $Q_1(x)$. The second term there will contribute for the commutators with the fourth and fifth terms in $T_1(y)$. Concretely, there is the term

$$\begin{aligned} &-iu(x)H(x)[\partial^\mu B(x), B(y)]A_\nu(y)\partial^\nu H(y) \\ &= -u(x)H(x)A_\nu(y)\partial^\nu H(y)\partial_x^\mu D(x-y), \end{aligned}$$

plus the analogous one in $[T_1(x), Q_1(y)]$. We are led to the singular part

$$(19) \quad -2uH(A \cdot \partial H)\delta(x-y).$$

The short rule here is $\partial^\mu D \rightarrow 2\delta$.

Next, we obtain

$$\begin{aligned} &\frac{im_H^2}{m}u(x)H(x)[\partial^\mu B(x), B(y)]B(y)H(y) \\ &= \frac{m_H^2}{m}u(x)H(x)B(y)H(y)\partial_x^\mu D(x-y), \end{aligned}$$

leading to the singular part

$$(20) \quad \frac{2m_H^2}{m} u B H^2 \delta(x - y).$$

From the last term in Q_1 , combining with the first term in $T_1(y)$, we obtain in all the singular part

$$(21) \quad 2muB(A \cdot A)\delta(x - y).$$

Combining both third terms, we consider

$$\begin{aligned} & -iu(x)B(x)[\partial^\mu H(x), H(y)]A_\nu(y)\partial^\nu B(y) \\ & = -u(x)B(x)A_\nu(y)\partial^\nu B(y)\partial_x^\mu D_{m_H}(x - y). \end{aligned}$$

We have in all the singular part:

$$(22) \quad -2uB(A \cdot \partial B)\delta(x - y).$$

Coming from respectively the third term in Q and the fifth term in T_1 , there is the commutator

$$\begin{aligned} & \frac{-im_H^2}{2m} u(x)B(x)[\partial^\mu H(x), H(y)]B^2(y) \\ & = -\frac{m_H^2}{m} u(x)B(x)B^2(y)\partial_x^\mu D_{m_H}(x - y). \end{aligned}$$

After collecting the similar term and taking the divergences, this leads to

$$(23) \quad -\frac{m_H^2}{m^3} u B^3 \delta(x - y).$$

Coming respectively from the third and sixth term, there is the commutator

$$\begin{aligned} & 3imeu(x)B(x)[\partial^\mu H(x), H(y)]H^2(y) \\ & = 3emu(x)B(x)H^2(y)\partial_x^\mu D_{m_H}(x - y). \end{aligned}$$

After taking the divergences, this leads to a total singular part

$$(24) \quad 6emuB H^2 \delta(x - y).$$

Next we list all possible normalization terms. Among them, the two first ones are coming from second-order tree graphs with two derivatives on the inner line. In other words, they come from $s[T_1(x), T_1(y)]$. Indeed, in this causal distribution, combining the third terms in the expression of T_1 , there appears the term

$$\begin{aligned} & iA_\mu(x)H(x)[\partial^\mu B(x), \partial^\nu B(y)]A_\nu(y)H(y) \\ & = A_\mu(x)H(x)A_\nu(y)H(y)\partial_x^\mu\partial_x^\nu D(x - y). \end{aligned}$$

This leads us to a normalization term $C_B(A \cdot A)H^2\delta(x - y)$. By the same token, the reader may verify that combining the fourth terms in the expression of T_1 there appears the normalization term $C_H(A \cdot A)B^2\delta(x - y)$.

However, any term of the same form, compatible with Poincaré covariance, discrete symmetries, ghost number and power counting represents in principle a

legitimate normalization. Thus we introduce the list of (re)normalization terms we need:

$$\begin{aligned} N_2^1 &= C_B(A \cdot A)H^2\delta(x - y); \\ N_2^2 &= C_H(A \cdot A)B^2\delta(x - y); \\ N_2^3 &= -\frac{m_H^2}{4m^2}B^4\delta(x - y); \\ N_2^4 &= \left(\frac{m_H^2}{m^2} + 3e\right)B^2H^2\delta(x - y). \end{aligned}$$

In view of (1) they generate new couplings. There is also a N_2^5 term in H^4 , that we omit for now. For convenience, we have anticipated the coefficients in N_2^3, N_2^4 , which are of the second class. The normalization terms amount to new vertices with four external legs. We compute the coboundaries:

$$\begin{aligned} sN_2^1 &= 2C_BH^2(A \cdot \partial u)\delta(x - y); \\ sN_2^2 &= 2C_H[B^2(A \cdot \partial u) + muB(A \cdot A)]\delta(x - y); \\ sN_2^3 &= -\frac{m_H^2}{m}uB^3\delta(x - y); \\ sN_2^4 &= \left(\frac{2m_H^2}{m} + 6em\right)uBH^2\delta(x - y). \end{aligned}$$

The cancellation now is easy to obtain: let $C_B = C_H = 1$. This means that we have only to worry about the first two terms in (17) and similarly in (18). Now, respectively the term (19) cancels the second one in (17) and the term (22) cancels the second one in (18). The two remaining terms in (17) and (18), together with (20), (21), (23) and (24) are exactly accounted for thanks to the normalization summands.

Therefore we have determined T_1 and T_2 , except that e still remains indeterminate. But please read on.

3.3. Higher-order analysis. For the higher-order analysis, it is convenient to have the expansion of the inverse \mathbb{S} -matrix:

$$\mathbb{S}^{-1}(g) =: 1 + \sum_1^\infty \frac{i^n}{n!} \int d^4x_1 \dots \int d^4x_n \overline{T}_n(x_1, \dots, x_n) g(x_1) \dots g(x_n).$$

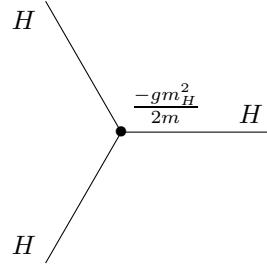
For instance, the second order term $\overline{T}_2(x_1, x_2)$ in the expansion of $\mathbb{S}^{-1}(g)$ is given by

$$\overline{T}_2(x_1, x_2) = -T_2(x_1, x_2) + T_1(x_1)T_1(x_2) + T_1(x_2)T_1(x_1).$$

Then, say,

$$\begin{aligned} A_3(x_1, x_2, x_3) &= \overline{T}_1(x_1)T_2(x_2, x_3) + \overline{T}_1(x_2)T_2(x_1, x_3) + \overline{T}_2(x_1, x_2)T_1(x_3) \\ &\quad + T_3(x_1, x_2, x_3); \\ R_3(x_1, x_2, x_3) &= T_1(x_3)\overline{T}_2(x_1, x_2) + T_2(x_1, x_3)\overline{T}_1(x_2) + T_2(x_2, x_3)\overline{T}_2(x_1) \\ &\quad + T_3(x_1, x_2, x_3). \end{aligned}$$

Just as before, $D_3 := R_3 - A_3$ depends only on known quantities, is causal in x_3 and is gauge invariant. Splitting it, we can calculate T_3 . We refer to [15] for the outcome of the analysis in our case, which turns out to be quite simple. The the missing cubic term is given by:



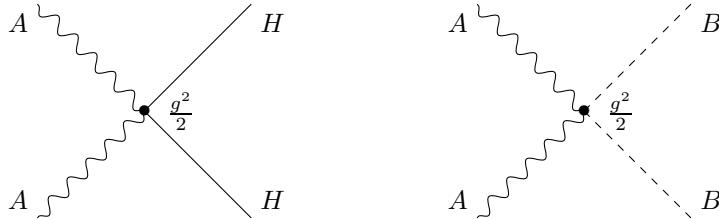
Also, it is seen that we need the new normalization term $N_2^5 = fH^4\delta$. One finds $f = -m_H^2/4m^2$, and we are home.

3.4. Summary of the abelian model. Thus we write down the final (interaction) Lagrangian associated to the abelian theory of the previous section. There are two physical fields A^μ, H , of respective masses m, m_H , and an assortment of ghosts u, \tilde{u}, B , which in our Feynman gauge all possess mass m . We obtained six cubic couplings (proportional to g) and five quartic ones (proportional to g^2). It is remarkable that CGI generates the latter from the former. Only four terms out of the eleven involve couplings exclusively among the physical fields.

$$\begin{aligned} \mathcal{L}_{\text{int}}(x) = & gm(A \cdot A)H - gm\tilde{u}uH + gB(A \cdot \partial H) \\ & - gH(A \cdot \partial B) - \frac{gm_H^2}{2m}H^3 - \frac{gm_H^2}{2m}B^2H \\ & + \frac{g^2}{2}(A \cdot A)H^2 + \frac{g^2}{2}(A \cdot A)B^2 \\ & - \frac{g^2m_H^2}{8m^2}H^4 - \frac{g^2m_H^2}{4m^2}H^2B^2 - \frac{g^2m_H^2}{8m^2}B^4. \end{aligned}$$

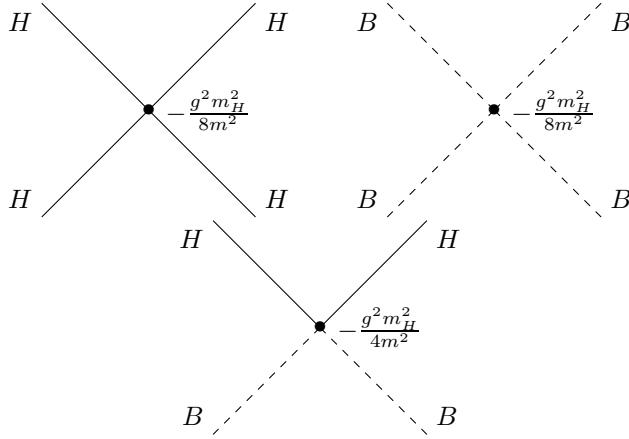
With $C = 1/m$ this tails down perfectly with (4) together with (5), (6) and (7). By construction the total Lagrangian is BRS invariant in the sense defined here. (It has been proved recently in a rigorous way [22] in the EG framework for interacting fields that “classical” BRS invariance implies gauge invariance for all tree graphs at all orders.)

We exhibit the quartic interaction vertices graphically.



Notice that the purely scalar couplings are

$$\begin{aligned} & -g\frac{m_H^2}{2m}H(B^2 + H^2) - g^2\frac{m_H^2}{8m^2}(B^2 + H^2)^2 \\ & = -\frac{g^2m_H^2}{8m^2}(B^2 + H^2)(B^2 + H^2 + \frac{4m}{g}H). \end{aligned}$$



Performing now an asymptotic analysis (that is, taking the Stückelberg field $B = 0$) it becomes

$$-\frac{g^2 m_H^2}{8m^2} (H^4 + \frac{4m}{g} H^3).$$

4. Three MVBS

Let us now seek all gauge theories with *three* gauge fields. The only interesting Lie algebra entering the game is

$$\mathfrak{g} = \mathfrak{su}(2);$$

in this case obviously total antisymmetry implies the Jacobi identity.

The case $m_1 = m_2 = m_3 = 0$ is certainly possible, and then neither scalar Higgs nor Stückelberg fields are necessary.

The simplest of the *mass relations* we referred to in Section 2 is the following: if $f_{abc} \neq 0$ and $m_a = 0$, then necessarily $m_b = m_c$. We see at once from this that if $m_1 = 0$ must be $m_2 = m_3$: the case $m_1 = m_2 = 0, m_3 \neq 0$ is downright impossible.

The only other mass relation one needs to check to verify that models with two or three MVBS and one Higgs-like field are correct in our sense is

$$(25) \quad \begin{aligned} 4C^2 m_b^2 m_a^2 &= 2(m_a^2 + m_b^2) \sum_{d:m_d=0} (f_{abd})^2 \\ &+ \sum_{k:m_k \neq 0} \frac{(f_{abk})^2}{m_k^2} [(m_a^2 + m_b^2 + m_k^2)^2 - 4(m_a^2 m_b^2 + m_k^4)]. \end{aligned}$$

With $C^{-1} = \pm m_2$, the model with the mass pattern $m_2 = m_3 \neq 0, m_1 = 0$ passes muster.

If we assume that all masses are different from zero, then necessarily $m_1 = m_2 = m_3$. Indeed, equation (25) implies

$$4m_a^2 m_b^2 m_c^2 C^2 = [(m_a^2 + m_b^2 + m_c^2)^2 - 4(m_a^2 m_b^2 + m_c^4)]$$

where (a, b, c) is any permutation of $(1, 2, 3)$. Therefore,

$$m_1^2 m_2^2 + m_3^4 = m_2^2 m_3^2 + m_1^4 = m_3^2 m_1^2 + m_2^4.$$

This yields

$$(m_1^2 m_2^2 + m_3^4) - (m_2^2 m_3^2 + m_1^4) = (m_3^2 - m_1^2)(m_1^2 - m_2^2 + m_3^2) = 0,$$

$$(m_2^2 m_3^2 + m_1^4) - (m_3^2 m_1^2 + m_2^4) = (m_1^2 - m_2^2)(m_2^2 - m_3^2 + m_1^2) = 0,$$

whose only all-positive solution is $m_1 = m_2 = m_3 =: m$; and then $4m^6 C^2 = m^4$ yields $C^{-1} = \pm 2m$.

Physically, the two cases just examined correspond respectively to the Georgi–Glashow “electroweak” theory without neutral currents; and to the $\mathfrak{su}(2)$ Higgs–Kibble model. Reference [13] claims that more than one Higgs-like particle for the $\mathfrak{su}(2)$ Higgs–Kibble model is not allowed. It is well known that the first mass pattern obtained here is arrived at by SSB when the Higgs sector is chosen to be a $SU(2)$ isovector; and the second one when it is a complex doublet. But in our derivation SSB played no role.

5. The Weinberg–Salam model within CGI

Scharf and coworkers (see references in the introduction) followed a “deductive” approach to the SM, with the only assumption that m_1, m_2, m_3 are all positive, plus existence of the photon, that is, $m_4 = 0$. There is no point in repeating that. Suffice to say that a structure constant like f_{124} is found to be non-zero, thus $m_1 = m_2$; and also the mass constraints imply $m_3 > m_1$. Defining

$$\cos \theta_W := m_1/m_3,$$

it is possible now to take for the non-zero structure constants

$$|f_{123}| = \cos \theta_W \quad \text{and} \quad |f_{124}| = \sin \theta_W.$$

With this, simply bringing (4) together with equations with (5), (6) and (7), one retrieves the boson part of the SM Lagrangian, as given for example in [23].

Thus it appears that the ordinary version of the Higgs sector for the gauge group $SU(2) \times U(1) \simeq U(2)$ is “chosen” by CGI. Of course, one can argue for it from other considerations within the SSB framework, or refer to experiment. The fact remains that CGI is more restrictive than the BHEK mechanism. The comment in the final discussion on the open problem of determining which patterns of broken symmetry are allowed in CGI for general gauge groups.

5.1. Coupling to matter. Things remain interesting when considering the fermion sector. The basic interaction between carriers and matter in a gauge theory is of the form

$$g(b^a A_{a_\mu} \bar{\psi} \gamma^\mu \psi + b'^a A_{a_\mu} \bar{\psi} \gamma^\mu \gamma^5 \psi),$$

with $\bar{\psi}$ the Dirac adjoint spinor and b, b' appropriate coefficients. In dealing with the SM our fermions are the known ones, fulfilling as free fields the Dirac equation: we do not assume chiral fermions *ab initio*. Their gauge variation is taken to be

zero. Thus for the SM one makes the Ansatz

$$\begin{aligned}
T_1^F = & b_1 W_\mu^+ \bar{e} \gamma^\mu \nu + b'_1 W_\mu^+ \bar{e} \gamma^\mu \gamma^5 \nu + b_2 W_\mu^- \bar{\nu} \gamma^\mu e + b'_2 W_\mu^- \bar{\nu} \gamma^\mu \gamma^5 e \\
& + b_3 Z_\mu \bar{e} \gamma^\mu e + b'_3 Z_\mu \bar{e} \gamma^\mu \gamma^5 e + b_4 Z_\mu \bar{\nu} \gamma^\mu \nu + b'_4 Z_\mu \bar{\nu} \gamma^\mu \gamma^5 \nu \\
& + b_5 A_\mu \bar{e} \gamma^\mu e + b'_5 A_\mu \bar{e} \gamma^\mu \gamma^5 e + b_6 A_\mu \bar{\nu} \gamma^\mu \nu + b'_6 A_\mu \bar{\nu} \gamma^\mu \gamma^5 \nu \\
& + c_1 B^+ \bar{e} \nu + c'_1 B^+ \bar{e} \gamma^5 \nu + c_2 B^- \bar{\nu} e + c'_2 B^- \bar{\nu} \gamma^5 e \\
& + c_3 B_Z \bar{e} e + c'_3 B_Z \bar{e} \gamma^5 e + c_4 B_Z \bar{\nu} \nu + c'_4 B_Z \bar{\nu} \gamma^5 \nu \\
(26) \quad & + c_5 H \bar{\nu} \nu + c'_5 H \bar{\nu} \gamma^5 \nu + c_6 H \bar{e} e + c'_6 H \bar{e} \gamma^5 e.
\end{aligned}$$

Here e stands for an electron, muon or neutrino or a (suitable combination of) quarks d, s, b ; and ν for the neutrinos or the quarks u, c, t ; the charge difference is always minus one. For instance in the “vertex” $W_\mu^+ \bar{e} \gamma^\mu \nu$ a “positron” exchanges a W^+ boson and becomes a “neutrino”. Charge is conserved in each term.

The method to determine the coefficients in (26) *remains the same*; only, it is simpler in practice. We limit ourselves to a few remarks. The direct equation

$$sT_1^F = i\partial \cdot Q_1^F$$

already allows to determine $c_1, c'_1, c_2, c'_2, c_3, c'_3, c_4, c'_4$, as well as the vanishing of b'_5 and b'_6 , assuming nonvanishing fermion masses. (For ν representing a true neutrino, we expect the term with coefficient b'_6 to vanish anyway, since the photon should not couple to uncharged particles. The same is true for b_6 .) Thus the photon has no axial-vector couplings, “because” there is no Stückelberg field for it, that is, because it is massless. The reader will have no trouble in finding the explicit form of Q_1^F , that can be checked with [15, Eq. 4.7.4]. At second order, one needs to take into account the interplay of contractions between Q_1 and T_1^F , as well as the “purely fermionic” ones between Q_1^F and T_1^F . There are no contractions between Q_1^F and T_1 , since the former does not contain derivatives. Also, *no new normalization terms* with fermionic fields may be forthcoming in sN_2 or $\partial_x \cdot N_{2/1}, \partial_y \cdot N_{2/2}$, since a term $\sim \varphi_1 \varphi_2 \bar{\psi} \psi \delta$ would be nonrenormalizable by power counting: the only way to cancel local terms is that the coefficient of every generated Wick monomial add up to zero.

At the end of the day, the physical Higgs couplings are proportional to the mass, and *chirality* of the interactions is a *consequence* of CGI [11, 12]. For leptons it yields:

$$\begin{aligned}
T_1^F = & \frac{1}{2\sqrt{2}} W_\mu^+ \bar{e} \gamma^\mu (1 \pm \gamma_5) \nu + \frac{1}{2\sqrt{2}} W_\mu^- \bar{\nu} \gamma^\mu (1 \pm \gamma_5) e + \frac{1}{4 \cos \theta_W} Z_\mu \bar{e} \gamma^\mu (1 \pm \gamma_5) e \\
& - \sin \theta_W \tan \theta_W Z_\mu \bar{e} \gamma^\mu e - \frac{1}{4 \cos \theta_W} Z_\mu \bar{\nu} \gamma^\mu (1 \pm \gamma_5) \nu + \sin \theta_W A_\mu \bar{e} \gamma^\mu e \\
& + i \frac{m_e - m_\nu}{2\sqrt{2}m_W} B^+ \bar{e} \nu \pm i \frac{m_e + m_\nu}{2\sqrt{2}m_W} B^+ \bar{e} \gamma^5 \nu - i \frac{m_e - m_\nu}{2\sqrt{2}m_W} B^- \bar{\nu} e \\
& \pm i \frac{m_e + m_\nu}{2\sqrt{2}m_W} B^- \bar{\nu} \gamma^5 e \pm i \frac{m_e}{2m_W} B_Z \bar{e} e \pm i \frac{m_\nu}{2m_W} B_Z \bar{e} \gamma^5 \\
& + \frac{m_\nu}{2m_W} H \bar{\nu} \nu + \frac{m_e}{2m_W} H \bar{e} e,
\end{aligned}$$

as it should. In summary we have recovered the SM, with its rationale upside-down.

6. Discussion

People define $e = g \sin \theta_W$; $g' = g \tan \theta_W$. Therefore,

$$\sec \theta_W = \frac{\sqrt{g^2 + g'^2}}{g},$$

and selecting the chiral projector and apart from the standard factors, the effective coupling of the term in $W_\mu^+ \bar{e} \gamma^\mu \nu$ and conjugate is g ; that of the term $A_\mu \bar{e} \gamma^\mu e$ is e ; that of the term $Z_\mu \bar{\nu} \gamma^\mu \nu$ is $-\sqrt{g^2 + g'^2}/2g$; and so on. Thus one can artfully write things as if g, g' are two different coupling constant associated to the emerging representation of the gauge group. But we have seen that the coefficients come from the pattern of masses, which in our viewpoint is fixed by nature. In order to bring home the point, let us make the *Gedankenexperiment* of building the SM from the Georgi–Glashow model, by adding a vector boson, sitting on an invariant abelian subgroup. Implicitly we allow for two different coupling constants (plus mixing of the old photon and the new MVB). But in that case there is no reason for $m_Z > m_W$. It is more natural to assume that the SM stems from the Higgs–Kibble model, keeping one coupling constant, whereby two of the three masses are moderately “pulled down” by mixing with the new photon. This goes to the heart of the experimental situation; other weak isospin values do not enter the game. In other words, no support comes from our quarter to the idea that the SM as it stands is “imperfectly unified”. The argument is bolstered by the fact that the true group of the electroweak interaction is $U(2)$, not $SU(2) \times U(1)$.²

In usual presentations of the SM the $U(2)$ symmetry is said to be “broken”, among other reasons, because there is only one conserved quantity, electric charge, instead of four. In CGI the interaction appears to respect the $U(2)$ symmetry. But of course symmetry is broken already at the level of the free Lagrangian, due to different masses (electric charge conservation reflects the residual symmetry $m_1 = m_2$). This is to say that not all bases of the Lie algebra are equivalent, since there is a natural basis dictated by the pattern of masses. The role of the mass constraints is precisely to pick out this basis.³

Let us recapitulate. CGI is a tool for the actual construction of Lagrangians. We limited ourselves to polynomial couplings. At first order in the coupling constant, CGI fixes some of the couplings of the vector bosons and ghost (fermionic and bosonic) fields. At second order, it requires additional quartic couplings, as well as some extra ingredient, which here is made out of physical scalars or Higgs-like fields. Third-order invariance goes on to fix the remaining couplings of the Higgs-like fields. One obtains in that way potentials of the symmetry-breaking kind, although SSB does not enter the picture. The class of allowed models is however smaller than within the prevalent paradigm. Ockham’s razor, already invoked in Section 2 in relation with the number of Higgs fields, seems even more pertinent here.

²To our knowledge, this was noticed first in [24].

³In the early seventies, speculations on fermionic patterns of masses from SSB were rife — see for instance [25]. They have been since all but abandoned. They might perhaps become a respectable subject of study again. By now we may only say that in CGI differences between fermion masses are related to differences between boson masses, in the sense that (disregarding family mixing) models in which all bosons share the same mass would entail identity of all fermion masses as well.

On the historical side, it is difficult to imagine the development of electroweak unification during the sixties without the SSB crutch. Massive vector bosons were beyond the pale then. The only contemporaneous article (still instructive today) I know of, willing and eager to start from them as fundamental entities is [26]; it did not have enough impact. Around ten years later, after the invention of SSB, cogent arguments based on tree unitarity —see [27] and references therein— weighed in favour of gauge theories with broken symmetry (plus abelian mass terms for invariant abelian subgroups). These arguments do not appear to distinguish between potentials constructed through CGI and those postulated in the standard procedure; the question deserves further investigation.

Since they lead to the same phenomenological Lagrangian, there seems to be no way as yet —within ordinary particle physics, at least— to distinguish between the SM as presented in textbooks and its causal version. This is good, because it shows that CGI is solidly anchored in physics. It is also bad: “a difference, to be a difference, has to make a difference”. Still, a constructive CGI program was in principle attractive because the apparent severity of the constraints on the masses of the gauge fields. Ambauen and Scharf [28] argued that the $SU(5)$ grand unification model by Georgi and Glashow with its standard pattern of Higgs fields [29, Chap. 18], is not causally gauge invariant; and the situation in this respect for a while was murky. A systematic comparison between CGI and the general theory of broken local symmetries [30] has been performed only recently [31].

Acknowledgments. I am most grateful for discussions to Luis J. Boya, Florian Scheck and Joseph C. Várilly. Special thanks are due to Michael Dütsch, who patiently explained to me several aspects of the quantum gauge principle, according to the Zürich school. I acknowledge support from CICYT, Spain, through grant FIS2005-02309.

References

- [1] I. J. R. Aitchison and A. J. G. Hey, *Gauge theories in particle physics: QCD and the electroweak theory*, IOP Publishing, Bristol, 2004.
- [2] C. Burgess and G. Moore, *The Standard Model: a primer*, Cambridge University Press, Cambridge, 2007.
- [3] M. J. G. Veltman, Phys. Rev. Lett. **34** (1975) 777.
- [4] H. Cheng and E-C. Tsai, Phys. Rev. D **40** (1989) 1246.
- [5] M. J. G. Veltman, Rev. Mod. Phys. **72** (2000) 341.
- [6] J. Earman, Philos. Sci. **71** (2004) 1227.
- [7] H. Lyre, Intl. Studies Philos. Sci. **22** (2008) 119.
- [8] H. Ruegg and M. Ruiz-Altaba, Int. J. Mod. Phys. A **19** (2004) 3265.
- [9] J. M. Gracia-Bondía, “BRS invariance for massive boson fields”, to appear in the Proceedings of the Summer School “Geometrical and topological methods for quantum field theory”, Cambridge University Press, Cambridge, 2009; hep-th/0808.2853.
- [10] W. Kilian, *Electroweak symmetry breaking: the bottom-up approach*, Springer, New York, 2003.
- [11] M. Dütsch and G. Scharf, Ann. Phys. (Leipzig) **8** (1999) 359.
- [12] A. Aste, G. Scharf and M. Dütsch, Ann. Phys. (Leipzig) **8** (1999) 389.
- [13] M. Dütsch and B. Schroer, J. Phys. A **33** (2000) 4317.
- [14] D. R. Grigore, J. Phys. A **33** (2000) 8443.
- [15] G. Scharf, *Quantum gauge theories. A true ghost story*, Wiley, New York, 2001.
- [16] T. Hurth and K. Skenderis, Nucl. Phys. B **541** (1999) 566.
- [17] T. Hurth and K. Skenderis, in *New developments in quantum field theory*, P. Breitenlohner, D. Maison and J. Wess, eds., Springer, Berlin, 2000; pp. 86–105.
- [18] H. Epstein and V. Glaser, Ann. Inst. Henri Poincaré **XIXA** (1973) 211.

- [19] L. Alvarez-Gaumé and L. Baulieu, Nucl. Phys. B **212** (1983) 255.
- [20] R. Stora, “Local gauge groups in quantum field theory: perturbative gauge theories”, talk given at the workshop “Local quantum physics”, Erwin Schrödinger Institute, Vienna, 1997.
- [21] M. Dütsch, private communication.
- [22] M. Dütsch, Ann. Phys. (Leipzig) **14** (2005) 438.
- [23] M. Veltman, *Diagrammatica*, Cambridge University Press, Cambridge, 1994.
- [24] F. Scheck, *Leptons, hadrons and nuclei*, North-Holland, Amsterdam, 1983.
- [25] H. Georgi and S. L. Glashow, Phys. Rev. D **6** (1972) 2977.
- [26] V. I. Ogievetskij and I. V. Polubarinov, Ann. Phys. (New York) **25** (1963) 358.
- [27] J. M. Cornwall, D. N. Levin and G. Tiktopoulos, Phys. Rev. D **10** (1974) 1145.
- [28] M. Ambauen and G. Scharf, “Violation of quantum gauge invariance in Georgi–Glashow $SU(5)$ ”, hep-th/0409062.
- [29] H. Georgi, *Lie algebras in particle physics*, Westview Press, Boulder, 1999.
- [30] Ling-Fong Li, Phys. Rev. D **9** (1974) 1723.
- [31] M. Dütsch, J. M. Gracia-Bondía, F. Scheck and J. C. Várilly, “Higgs-mechanism-free model building”, to appear.

DEPARTAMENTO DE FÍSICA TEÓRICA,, UNIVERSIDAD DE ZARAGOZA, ZARAGOZA 50009, SPAIN,
AND, DEPARTAMENTO DE FÍSICA,, UNIVERSIDAD DE COSTA RICA, SAN PEDRO 2060, COSTA RICA

This page intentionally left blank

Abstract integration, Combinatorics of Trees and Differential Equations

Massimiliano Gubinelli

ABSTRACT. This is a review paper on recent work about the connections between rough path theory, the Connes-Kreimer Hopf algebra on rooted trees and the analysis of finite and infinite dimensional differential equations. We try to explain and motivate the theory of rough paths introduced by T. Lyons in the context of differential equations in presence of irregular noises. We show how it is used in an abstract algebraic approach to the definition of integrals over paths which involves a cochain complex of finite increments. In the context of such abstract integration theories we outline a connection with the combinatorics of rooted trees. As interesting examples where these ideas apply we present two infinite dimensional dynamical systems: the Navier-Stokes equation and the Korteweg-de-Vries equation.

Keywords: Rough Path theory, Connes-Kreimer Hopf algebra, driven differential equations, Navier-Stokes equation, Korteweg-de-Vries equation.

1. Introduction

Rough path theory has been developed by T. Lyons for the analysis of the map $\Phi : x \mapsto y$ which sends a vector-valued driving signal x to the solution of the differential equation

$$(1) \quad dy_t = \sum_a f_a(y_t) dx_t^a, \quad y_0 = \bar{y}$$

where \bar{y} is the initial condition. This equation has a well defined meaning as a non-autonomous ODE when x_t is differentiable in t . In applications, however, it is interesting to consider more general driving signals, e.g. when taking into account random perturbations of dynamical systems where x_t is the Brownian motion which is almost surely non-differentiable. Indeed the application to stochastic differential equations (SDEs) has been the main motivation in the developments of the theory. The standard approach is then to understand eq. (1) as the integral equation

$$(2) \quad y_t = \bar{y} + \int_0^t \sum_a f_a(y_s) dx_s^a,$$

provide a well-defined meaning for the *stochastic integral* in the r.h.s. and the proceed to solve the equation by standard fixed-point methods. Usually the integral above can be an Itô integral or a Stratonovich one, but other more exotic choices are possible in other stochastic contexts (e.g. the Skorohod integral, the “normal ordered” integral, etc...). Whenever we speak of a solution y to the SDE (1) we mean suitable random function y for which $f(y)$ can be integrated against dx in an appropriate sense.

Lyons’ basic observation was that, in the case of a smooth control x , the solution of the differential equation is a well-behaved function of the *iterated integrals* X of x :

$$(3) \quad X_{ts}^{\bar{a}} = \int_s^t \int_s^{u_n} \cdots \int_s^{u_2} dx_{u_1}^{a_1} \cdots dx_{u_n}^{a_n}$$

where we denote with \bar{a} the multi-index (a_1, \dots, a_n) with $|\bar{a}| = n$ its length. In some sense the iterated integrals encode the local behavior of the path x well enough to faithfully recover its effect on the solution y . In a system perspective, the iterated integrals provide a canonical set of coordinates for the analysis of non-linear systems much like the standard Fourier coefficients are natural coordinates for linear ones.

Like Fourier coefficients, iterated integrals enjoy nice relations upon concatenation of paths: given three times $s < u < t$ we have the celebrated Chen’s relations [1] between iterated integrals:

$$(4) \quad X_{ts}^{a_1 \cdots a_n} = X_{tu}^{a_1 \cdots a_n} + X_{us}^{a_1 \cdots a_n} + \sum_{k=1}^{n-1} X_{tu}^{a_1 \cdots a_k} X_{us}^{a_{k+1} \cdots a_n}.$$

These non-linear equations plays a fundamental role in the development of the theory.

The plan of the paper is the following: in Sect. 2 we motivate the simplest instance of a rough path in the context of integration against smooth approximations of an irregular path. Sect. 3 we describe an algebraic approach to integration and its use to define integrals against a rough path. This approach is not standard from the point of view of rough paths. More conventional expositions of the theory can be found in [2, 3] and [4]. In Sect. 4 we describe the combinatorics of iterated integrals of very general type via the Connes-Kreimer Hopf algebra on rooted trees following [5]. This will ultimately allow to define and solve differential equation associated to such integrals. Finally in Sect. 5 we apply the objects and the related combinatorics of algebraic integrals to the study of infinite dimensional differential equations via two prototypical examples: the 3d Navier-Stokes equation as studied in [6] and the 1d periodic Korteweg-de-Vries equation [7].

2. Renormalizable theories of integration

The analysis of the integral equation (2) in the case that x is a non-differentiable function can be split in two parts: the definition of the integral on the r.h.s. and the fix-point argument. Let us concentrate on the first part. Assume given a continuous function $x : [0, T] \rightarrow \mathbb{R}^n$ which we assume only γ -Hölder continuous, i.e. for which the following estimate holds

$$|x_t - x_s| \leq C|t - s|^\gamma, \quad t, s \in [0, T]$$

(where the least constant C define the Hölder norm $\|x\|_\gamma$) and for simplicity let us restrict in this section to the case $1/3 < \gamma < 1$. Let us pose the problem of giving a “reasonable” definition of the integral

$$I[\varphi]_{ts} = \int_s^t \varphi(x_u) dx_u$$

for some smooth one-form $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Proceeding by approximation we consider a family $\{x(\varepsilon)\}_{\varepsilon > 0}$ of smooth approximations to the path x and let

$$I[\varphi](\varepsilon)_{ts} = \int_s^t \varphi(x(\varepsilon)_u) dx(\varepsilon)_u.$$

It is easy to convince ourselves that in general we do not have any control of these integrals if $x(\varepsilon) \rightarrow x$ when $\varepsilon \rightarrow 0$ in the γ -Hölder norm. It is then a remarkable fact that all the possible integrals obtained varying the function φ all converge at the same time (provided φ is sufficiently smooth) when the approximated iterated integral of order two:

$$X(\varepsilon)_{ts}^{a_1 a_2} = \int_s^t \int_s^u dx(\varepsilon)_v^{a_1} dx(\varepsilon)_u^{a_2}$$

converges to a function $X : [0, T]^2 \rightarrow \mathbb{R}^n \otimes \mathbb{R}^n$ in the sense that

$$\sup_{0 \leq s < t \leq T} \frac{|X(\varepsilon)_{ts}^{a_1 a_2} - X_{ts}^{a_1 a_2}|}{|t - s|^{2\gamma}} \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

So in some sense we can claim that the integration theory against the path x is well defined as long as we are able to control the convergence of the iterated integral of order two $X^{a_1 a_2}$: all the other integrals (and more fundamentally also all the higher order iterated integrals) will turn out to be nice functions of the data given by the path x and $X^{a_1 a_2}$.

It is suggestive to understand this phenomenon as a very simple example of “renormalizable theory” where all the quantities of interests have well defined meanings as functionals of a finite number of fundamental objects whose intrinsic determination remains outside the scope of the theory itself.

In this limited context we face the appearance of the simplest non-trivial example of a *rough path*: the couple $(X^a, X^{a_1 a_2})$ where $X_{ts}^a = x_t^a - x_s^a$ is a γ -rough path, i.e. a path and some additional information in the form of “iterated integrals”, for which a complete theory of integration and differential equations can be constructed (as we will see shortly). The fact that it is enough to consider only the second order integral is due to our hypothesis that $\gamma > 1/3$.

Note that once the limit has been taken the object $X^{a_1 a_2}$ is no more an iterated integral (in a classical sense) and can be characterized more abstractly by the following two properties

(1) the Chen’s relation:

$$(5) \quad X_{ts}^{a_1 a_2} = X_{tu}^{a_1 a_2} + X_{us}^{a_1 a_2} + X_{tu}^{a_1} X_{us}^{a_2}$$

(2) a regularity condition

$$(6) \quad \sup_{0 \leq s < t \leq T} \frac{|X_{ts}^{a_1 a_2}|}{|t - s|^{2\gamma}} < \infty.$$

and more interestingly there could be more than one possible choice for this object compatible with these two conditions leading to different integration theories.

3. Algebraic integration

To understand how the iterated integrals X come into play in the definition of $I[\varphi]$ we need some tool which allows us to analyze the “local” (with respect to the parameter) behavior of the integral. Specifically we want to expand the integral in a short interval as

$$(7) \quad I[\varphi]_{ts} = \varphi_a(x_s) X_{ts}^a + X_{ts}^{ba} \partial_b \varphi_a(x_s) + r_{ts}$$

where r stands for some remainder term which we hope will be of higher order in $|t - s|$ than the other terms. The main property of the integral $I[\varphi]$ is its trivial behavior under splitting of the integration interval: $I[\varphi]_{ts} = I[\varphi]_{tu} + I[\varphi]_{us}$ for $s < u < t$, this of course means also that the integral can be written as the increment of a function $I[\varphi]_{ts} = f_t - f_s$ (for example taking $f_t = I[\varphi]_{0t}$). In [8] we introduced a cochain complex (\mathcal{C}_*, δ) which encodes this basic property of integrals. For $n \geq 1$ let \mathcal{C}_n be set of functions $a \in C([0, T]^n; \mathbb{R})$ such that $a_{t_1 \dots t_n} = 0$ if $t_i = t_{i+1}$ for some $1 \leq i \leq n-1$. Elements in \mathcal{C}_n will be called n -increments. Define a coboundary $\delta : \mathcal{C}_n \rightarrow \mathcal{C}_{n+1}$ as

$$\delta a_{t_1 \dots t_{n+1}} = \sum_{k=1}^n (-1)^{k+1} a_{t_1 \dots \hat{t}_k \dots t_n}$$

where \hat{t}_k means that this particular argument is omitted. Explicitly we have $\delta a_{ts} = a_t - a_s$ for $a \in \mathcal{C}_1$ and $\delta a_{tus} = a_{ts} - a_{tu} - a_{us}$ for $a \in \mathcal{C}_2$. Let $\mathcal{ZC}_n = \text{Ker} \delta \cap \mathcal{C}_n$ and $\mathcal{BC}_n = \text{Im} \delta \cap \mathcal{C}_n$. It is easy then to verify that $\delta \delta = 0$ and that the complex (\mathcal{C}_*, δ) is exact, i.e. that $\text{Im} \delta = \text{Ker} \delta$ at any \mathcal{C}_n , $n \geq 1$. In particular a 2-increment $a \in \mathcal{C}_2$ is the increment of a function $f \in \mathcal{C}_1$ if and only if $\delta a = 0$. So at \mathcal{C}_2 the coboundary measures the degree of “exactness” of 2-increments. Moreover a key fact is the following: if δa is suitably small then there exists *only one* function f (modulo constants) such that

$$(8) \quad \delta f = a + r$$

where the remainder is small. To be more precise we need to introduce the relevant notion of “smallness”. We say that $a \in \mathcal{C}_2^\gamma$ if

$$\|a\|_\gamma := \sup_{s < t} \frac{|a_{ts}|}{|t - s|^\gamma} < \infty.$$

And similarly $b \in \mathcal{C}_3^\gamma$ if

$$\|b\|_\gamma := \sup_{s < u < t} \frac{|b_{tus}|}{|t - s|^\gamma} < \infty.$$

Both \mathcal{C}_2^γ and \mathcal{C}_3^γ are Banach spaces when endowed with the norms $\|\cdot\|_\gamma$. Define $\mathcal{C}_n^{1+} = \cup_{z > 1} \mathcal{C}_n^z$. Moreover if $g \in \mathcal{C}_n$ and $h \in \mathcal{C}_m$ then we write gh for the element of \mathcal{C}_{n+m-1} such that $(gh)_{t_1 \dots t_{n+m-1}} = g_{t_1 \dots t_n} h_{t_n \dots t_{n+m-1}}$.

The key result in this theory of increments is the existence of a *sewing map* which provide a natural inverse operation to δ :

THEOREM 3.1 (The sewing map). *There exists a unique bounded and linear map $\Lambda : \mathcal{ZC}_3^{1+} \rightarrow \mathcal{C}_2^{1+}$ such that $\delta \Lambda h = h$ for any $h \in \mathcal{ZC}_3^{1+}$.*

In particular the map $(1 - \Lambda \delta) : \mathcal{C}_2 \rightarrow \mathcal{C}_2$ projects (in essentially a unique way) suitable elements of \mathcal{C}_2 to $\mathcal{ZC}_2 = \mathcal{BC}_1$: the space of increments of functions. Then going back to the decomposition (8) and assuming that $\delta a \in \mathcal{C}_3^{1+}$ we can form

the exact 2-increment $a - \Lambda\delta a = \delta f$ and obtain that the remainder r is given by $\Lambda\delta a \in \mathcal{C}_2^{1+}$. It is also easy to see that, the decomposition (8) is unique if we require $r \in \mathcal{C}_2^{1+}$, indeed if two such decomposition exists, (f, r) and (f', r') , their difference satisfy $\delta(f - f') = r - r' \in \mathcal{C}_2^z$ for some $z > 1$ and $f - f'$ would be a function of z -Hölder class and since $z > 1$ that this function would take the constant value zero.

With this notions at hand we realize that eq. (7) is nothing more than an instance of a decomposition similar to (8). As we have already seen we can determine both $I[\varphi]$ and r at once using only the well-defined remaining terms in the r.h.s., so letting $a_{ts} = \varphi_a(x_s)X_{ts}^a + \partial_b\varphi_a(x_s)X_{ts}^{ba}$, in order to be able to apply Λ , we need to require that $\delta a \in \mathcal{C}_3^{1+}$. By some easy computation using the fact that δ satisfy some kind of Leibniz rule, we get

$$\delta a_{tus} = -\delta\varphi_a(x)_{tu}X_{us}^a - \delta\partial_b\varphi_a(x)_{tu}X_{ts}^{ba} + \partial_b\varphi_a(x_s)\delta X_{tus}^{ba}$$

If we now exploit the Chen's relation for X^{ba} we can simplify this expression further

$$(9) \quad \delta a_{tus} = -[\delta\varphi_a(x)_{tu} - \partial_b\varphi_a(x_s)\delta X_{tu}^b]X_{us}^a - \delta\partial_b\varphi_a(x)_{tu}X_{ts}^{ba}$$

Assuming that φ is sufficiently smooth (C^2 is enough) we obtain easily that

$$\delta\varphi_a(x) - \partial_b\varphi_a(x)\delta X^b \in \mathcal{C}_2^{2\gamma}, \quad \delta\partial_b\varphi_a(x) \in \mathcal{C}_2^\gamma$$

so that taking into account all the regularities we end up with $\delta a \in \mathcal{C}_3^{3\gamma}$. Then our assumption on γ ensure that $3\gamma > 1$ and we can prove that there exists a unique couple (f, r) with $f \in \mathcal{C}_1$ and $r \in \mathcal{C}_2^{1+}$ such that

$$(10) \quad \delta f = \varphi_a(x_s)X_{ts}^a + \partial_b\varphi_a(x_s)X_{ts}^{ba} + r_{ts}.$$

By construction f depends only on the γ -rough path X and on φ . We can then define the integral by

$$I[\varphi] = \delta f = (1 - \Lambda\delta)[X^a\varphi_a(x) + X^{ba}\partial_b\varphi_a(x)].$$

To motivate the fact that the decomposition (10) is a sort of renormalized integral we can make the following observation. Take a partition $\{t_i\}$ of $[s, t]$ of size Δ and consider the Riemann sums $S_\Delta = \sum_i \varphi(x_{t_i})X_{t_{i+1}t_i}$. In general we have no mean to say that these sums are convergent as $\Delta \rightarrow 0$. However given a γ -rough path X we can perform a subtraction to these sums and define

$$S'_\Delta = \sum_i \left[\varphi_a(x_{t_i})X_{t_{i+1}t_i}^a + \partial_b\varphi_a(x_{t_i})X_{t_{i+1}t_i}^{ba} \right]$$

then using the decomposition (10) as $\Delta \rightarrow 0$ we have the limit

$$S'_\Delta = \sum_i (\delta f)_{t_{i+1}t_i} - \sum_i r_{t_{i+1}t_i} = (\delta f)_{ts} - \sum_i o(|t_{i+1}t_i|) \rightarrow (\delta f)_{ts}$$

since the first sum telescopes and the second is easily show to converge to zero.

Moreover the regular dependence on the data ensure that smooth approximations $I[\varphi](\varepsilon)$ converge to $I[\varphi]$ as here defined as long as we can prove the convergence of approximating path (and its second order iterated integrals) to the rough path X .

Exploiting the sewing map, the integral equation (2) in presence of a γ -rough path with $\gamma > 1/3$ can be understood as a fixed-point equation for an unknown $y \in \mathcal{C}_1$:

$$\delta y = (1 - \Lambda\delta)[X^a\varphi_a(y) + X^{ac}\partial_b\varphi_a(y)\varphi_c^b(y)]$$

these equation can be solved by a standard iteration method in a suitable subspace of \mathcal{C}_1 ([8] for details).

It is also possible to construct higher order iterated integrals starting from low order ones where for iterated integrals we means simply object which obey Chen's relations (4). Take for example the third order object $X^{a_1 a_2 a_3} \in \mathcal{C}_2$, eq. (4) can be written as a statement about the coboundary of $X^{a_1 a_2 a_3}$:

$$\delta X_{tus}^{a_1 a_2 a_3} = X_{tu}^{a_1 a_2} X_{us}^{a_3} + X_{tu}^{a_1} X_{us}^{a_2 a_3}$$

and in the above hypothesis on X it is easy to check that the r.h.s. belongs to $\mathcal{Z}\mathcal{C}_2^{3\gamma} \subset \mathcal{Z}\mathcal{C}_2^{1+}$ so that it is in the domain of Λ and we can *define*

$$X^{a_1 a_2 a_3} := \Lambda[X^{a_1 a_2} X^{a_3} + X^{a_1} X^{a_2 a_3}]$$

as the unique solution in \mathcal{C}_2^{1+} of the Chen's relation. Iteratively this allows to construct all the higher order iterated integrals. Some more work allows to prove a uniform estimate on the growth of the norms involved in the procedure [2, 8]:

$$(11) \quad \|X^{\bar{a}}\|_{|\bar{a}|^\gamma} \leq C_1 \frac{C_2^{|\bar{a}|}}{(|\bar{a}|!)^\gamma}$$

for any multiindex a .

4. Trees

What happens if $\gamma \leq 1/3$? The obstruction to the exactness of the increment a in eq. (9) will no more belong to the domain of the sewing map. Indeed in eq. (7) we cannot anymore expect that the remainder belongs to \mathcal{C}_2^{1+} and at the very least this would affect our argument for uniqueness. We are then forced to proceed and expand further the integral, or from another point of view, to add some counterterms to remove large contributions to r .

To understand the general structures that we need if we proceed further in the expansions it is better to take as working bench the more difficult case of the differential equation (2). The series solution will be indexed by rooted trees: a phenomenon which is present already for solutions to the ODE $dy/dt = f(y)$ as noted in the classical work of Cayley [9] and later beautifully exploited by Butcher [10] and subsequent works of Hairer and Wanner [11, 12].

4.1. Labelled rooted trees. Given a finite set \mathcal{L} , the \mathcal{L} -labeled rooted trees are the finite graphs where labels of \mathcal{L} are attached to each vertex and where there is a special vertex called *root* such that there is a unique path from the root to any other vertex.

Some examples of rooted trees labeled by $\mathcal{L} = \{1, 2, 3\}$ are



Trees do not distinguish the order of branches to each vertex. Given k \mathcal{L} -decorated rooted trees τ_1, \dots, τ_k and a label $a \in \mathcal{L}$ we define $\tau = [\tau_1, \dots, \tau_k]_a$ as the tree obtained by attaching the k roots of τ_1, \dots, τ_k to a new vertex with label a which will be the root of τ . Any decorated rooted tree can be constructed using the simple decorated tree \bullet_a ($a \in \mathcal{L}$) and the operation $[\dots]$, e.g.

$$[\bullet] = \bullet \quad [\bullet, [\bullet]] = \begin{array}{c} \bullet \\ \backslash \quad / \\ \bullet \quad \bullet \end{array}, \quad \text{etc. . .}$$

Denote $\mathcal{T}_{\mathcal{L}}$ the set of all \mathcal{L} decorated rooted trees and let \mathcal{T} the set of rooted trees without decoration (i.e. for which the set of labels \mathcal{L} is made of a single element). There is a canonical map $\mathcal{T}_{\mathcal{L}} \rightarrow \mathcal{T}$ which simply forget all the labels and every function on \mathcal{T} can be extended, using this map to a function on $\mathcal{T}_{\mathcal{L}}$ for any set of labels \mathcal{L} . Let $|\cdot| : \mathcal{T} \rightarrow \mathbb{R}$ the weight which counts the number of vertices of the (undecorated) tree defined recursively as

$$|\bullet| = 1, \quad |[\tau_1, \dots, \tau_k]| = 1 + |\tau_1| + \dots + |\tau_k|$$

moreover the *factorial* $\gamma(\tau)$ of a tree τ is

$$\gamma(\bullet) = 1, \quad \gamma([\tau_1, \dots, \tau_k]) = |[\tau_1, \dots, \tau_k]| \gamma(\tau_1) \cdots \gamma(\tau_k)$$

finally the *symmetry factor* σ is given by the recursive formula $\sigma(\tau) = 1$ for $|\tau| = 1$ and

$$(12) \quad \sigma([\tau^1 \cdots \tau^k]_a) = \frac{k!}{\delta(\tau^1, \dots, \tau^k)} \sigma(\tau^1) \cdots \sigma(\tau^k)$$

where $\delta(\tau^1, \dots, \tau^k)$ counts the number of different ordered k -uples (τ^1, \dots, τ^k) which corresponds to the same (unordered) collection $\{\tau^1, \dots, \tau^k\}$ of subtrees. The factor $k!/\delta(\tau^1, \dots, \tau^k)$ counts the order of the subgroup of permutations of k elements which does not change the ordered k -uple (τ^1, \dots, τ^k) . Then $\sigma(\tau)$ is the order of the subgroup of permutations on the vertex of the tree τ which do not change the tree (taking into account also the labels). Another equivalent recursive definition for σ is

$$\sigma([\tau^1]^{n_1} \cdots [\tau^k]^{n_k}]_a) = n_1! \cdots n_k! \sigma(\tau^1)^{n_1} \cdots \sigma(\tau_k)^{n_k}$$

where τ^1, \dots, τ^k are distinct subtrees and n_1, \dots, n_k the respective multiplicities.

The algebra $\mathcal{AT}_{\mathcal{L}}$ is the commutative free polynomial algebra generated by $\{1\} \cup \mathcal{T}_{\mathcal{L}}$ over \mathbb{R} , i.e. elements of $\mathcal{AT}_{\mathcal{L}}$ are finite linear combination with coefficients in \mathbb{R} of formal monomials in the form $\tau_1 \tau_2 \cdots \tau_n$ with $\tau_1, \dots, \tau_n \in \mathcal{T}_{\mathcal{L}}$ or of the unit $1 \in \mathcal{AT}_{\mathcal{L}}$. The tree monomials are called *forests* and are collectively denoted $\mathcal{F}_{\mathcal{L}}$; we include the empty forest $1 \in \mathcal{F}_{\mathcal{L}}$. The algebra $\mathcal{AT}_{\mathcal{L}}$ is endowed with a graduation g given by $g(\tau_1 \cdots \tau_n) = |\tau_1| + \dots + |\tau_n|$ and $g(1) = 0$. This graduation induces a corresponding filtration of $\mathcal{AT}_{\mathcal{L}}$ in finite dimensional linear subspaces $\mathcal{A}_n \mathcal{T}_{\mathcal{L}}$ generated by the set $\mathcal{F}_{\mathcal{L}}^n$ of forests of degree $\leq n$.

Any map $f : \mathcal{T}_{\mathcal{L}} \rightarrow A$ where A is some commutative algebra, can be extended in a unique way to a homomorphism $f : \mathcal{AT}_{\mathcal{L}} \rightarrow A$ by setting: $f(\tau_1 \cdots \tau_n) = f(\tau_1) f(\tau_2) \cdots f(\tau_n)$.

In the following we will use letters $\tau, \rho, \sigma, \dots$ to denote trees in $\mathcal{T}_{\mathcal{L}}$ or forests in $\mathcal{F}_{\mathcal{L}}$, the degree $g(\tau)$ of a forest $\tau \in \mathcal{F}_{\mathcal{L}}$ will also be written as $|\tau|$. Roman letters $a, b, c, \dots \in \mathcal{L}$ will denote vector indexes (i.e. labels) while \bar{a}, \bar{b}, \dots will denote multi-indexes with values in \mathcal{L} : $\bar{a} = (a_1, \dots, a_n) \in \mathcal{L}^n$ with $|\bar{a}| = n$ the size of this multi-index.

4.2. Iterated integrals and the Connes-Kreimer Hopf algebra. Given a smooth path $x \in C^1([0, T], \mathbb{R}^n)$ we can canonically associate to it a family of 2-increments X^{τ} indexed by trees labelled by $\mathcal{L} = \{1, \dots, n\}$ by the iterated integrals

$$(13) \quad X_{ts}^{\bullet a} = x_t - x_s, \quad X_{ts}^{[\tau_1 \cdots \tau_n]_a} = \int_s^t X_{us}^{\tau_1} \cdots X_{us}^{\tau_n} dx_u^a$$

Iterated integrals like (3) correspond to “linear” trees $\tau = [[\cdots [\bullet_{a_n}]_{a_{n-1}} \cdots]_{a_2}]_{a_1}$ and by abuse of notation we will continue to write $X^{a_1 \cdots a_n}$ for such X^τ . The generalization of the Chen’s relations (4) involves the Hopf algebra structure on on labelled trees essentially introduced by Connes and Kreimer which we now briefly describe.

On the algebra $\mathcal{AT}_\mathcal{L}$ we can define a counit $\varepsilon : \mathcal{AT}_\mathcal{L} \rightarrow \mathbb{R}$ as an algebra homomorphism such that $\varepsilon(1) = 1$ and $\varepsilon(\tau) = 0$ otherwise and a coproduct $\Delta : \mathcal{AT}_\mathcal{L} \rightarrow \mathcal{AT}_\mathcal{L} \otimes \mathcal{AT}_\mathcal{L}$ as the algebra homomorphism such that

$$(14) \quad \Delta(\tau) = 1 \otimes \tau + \sum_{a \in \mathcal{L}} (B_+^a \otimes \text{id})[\Delta(B_-^a(\tau))]$$

on trees $\tau \in \mathcal{T}_\mathcal{L}$, where $B_+^a(1) = \bullet_a$ and $B_+^a(\tau_1 \cdots \tau_n) = [\tau_1 \cdots \tau_n]_a$ and B_-^a is the inverse of B_+^a or is equal to zero if the tree root does not have label a , i.e.

$$B_-^a(B_+^b(\tau_1 \cdots \tau_n)) = \begin{cases} \tau_1 \cdots \tau_n & \text{if } a = b \\ 0 & \text{otherwise.} \end{cases}$$

Endowed with ε and Δ the algebra $\mathcal{AT}_\mathcal{L}$ become a bialgebra, there exists also an antipode S which complete the definition of the Hopf algebra structure on $\mathcal{T}_\mathcal{L}$ as described by Connes-Kreimer [13] (in the unlabeled case).

We will often use Sweedler’s notation for the coproduct $\Delta\tau = \sum \tau_{(1)} \otimes \tau_{(2)}$ but we also introduce a counting function $c : \mathcal{T}_\mathcal{L} \times \mathcal{T}_\mathcal{L} \times \mathcal{F}_\mathcal{L} \rightarrow \mathbb{N}$ such that $\Delta\tau = \sum_{\rho \in \mathcal{T}_\mathcal{L}, \sigma \in \mathcal{F}_\mathcal{L}} c(\tau, \rho, \sigma) \rho \otimes \sigma$ moreover it will be useful to consider also the reduced coproduct $\Delta'\tau = \Delta\tau - 1 \otimes \tau - \tau \otimes 1$ with counting function c' .

The generalization of eq. (4) reads as follows:

THEOREM 4.1 (Tree multiplicative property). *The map X satisfy the algebraic relation*

$$(15) \quad \delta X^\sigma = X^{\Delta'(\sigma)}, \quad \sigma \in \mathcal{AT}_\mathcal{L}$$

Let us give an example in one dimension ($d = 1$) where trees are not decorated. The forests with $|\tau| \leq 3$ are

$$\bullet, \bullet\bullet, \bullet\bullet\bullet, \bullet\bullet\bullet\bullet, \bullet\bullet\bullet\bullet\bullet, \bullet\bullet\bullet\bullet\bullet\bullet$$

The action of the reduced coproduct on these forests and the corresponding action of the coboundary on the iterated integrals are given by

$$\begin{aligned} \Delta' \bullet &= \bullet \otimes \bullet \\ \Delta'(\bullet\bullet) &= 2\bullet \otimes \bullet \\ \Delta' \bullet\bullet &= \bullet \otimes \bullet + \bullet \otimes \bullet \\ \Delta'(\bullet\bullet\bullet) &= \bullet \otimes \bullet\bullet + \bullet\bullet \otimes \bullet + \bullet \otimes \bullet\bullet\bullet \\ \Delta'(\bullet^3) &= 3\bullet^2 \otimes \bullet + 3\bullet \otimes \bullet^2 \\ \Delta' \bullet\bullet\bullet &= \bullet \otimes \bullet\bullet + 2\bullet\bullet \otimes \bullet \\ \delta X_{tus}^\bullet &= X_{tu}^\bullet X_{us}^\bullet \\ \delta X_{tus}^{\bullet\bullet} &= 2X_{tu}^\bullet X_{us}^\bullet \\ \delta X_{tus}^{\bullet\bullet\bullet} &= X_{tu}^\bullet X_{us}^{\bullet\bullet} + X_{tu}^{\bullet\bullet} X_{us}^\bullet \\ \delta X_{tus}^{\bullet\bullet\bullet\bullet} &= X_{tu}^\bullet X_{us}^{\bullet\bullet\bullet} + X_{tu}^{\bullet\bullet\bullet} X_{us}^\bullet \\ \delta X_{tus}^{\bullet\bullet\bullet\bullet\bullet} &= X_{tu}^\bullet X_{us}^{\bullet\bullet\bullet\bullet} + X_{tu}^{\bullet\bullet\bullet\bullet} X_{us}^\bullet \\ \delta X_{tus}^{\bullet\bullet\bullet\bullet\bullet\bullet} &= 3X_{tu}^{\bullet\bullet} X_{us}^\bullet + 3X_{tu}^\bullet X_{us}^{\bullet\bullet} \\ \delta X_{tus}^{\bullet\bullet\bullet\bullet\bullet\bullet\bullet} &= X_{tu}^\bullet X_{us}^{\bullet\bullet\bullet\bullet} + 2X_{tu}^{\bullet\bullet\bullet\bullet} X_{us}^\bullet \end{aligned}$$

As a further example consider the iterated integrals T^τ associated to the identity path

$$T_{ts}^\bullet = t - s, \quad T_{ts}^{[\tau_1 \cdots \tau_n]} = \int_s^t T_{us}^{\tau_1} \cdots T_{us}^{\tau_n} du$$

By induction it is not difficult to prove that $T_{ts}^\tau = (t-s)^{|\tau|}(\tau!)^{-1}$, so applying Thm. 4.1 to T^τ we get a remarkable binomial-like formula for the Connes-Kreimer coproduct

$$(16) \quad (a+b)^{|\tau|} = \sum \frac{\tau!}{\tau^{(1)}!\tau^{(2)}!} a^{|\tau^{(1)}|} b^{|\tau^{(2)}|}.$$

Iterated integrals of this sort appears naturally when trying to expand in series the solution of driven differential equations: for any analytic vectorfield $f : \mathbb{R}^d \rightarrow \mathbb{R}^n$ and any smooth path $x \in C^1([0, T], \mathbb{R}^n)$, the solution of the differential equation $dy_t = \sum_{a \in \mathcal{L}} f_a(y_t) dx_t^a$, with $y_0 = \eta$ admit locally the series representation

$$(17) \quad \delta y_{ts} = \sum_{\tau \in \mathcal{T}_{\mathcal{L}}} \frac{1}{\sigma(\tau)} \phi^f(\tau)(y_s) X_{ts}^\tau, \quad y_0 = \eta$$

where we recursively define functions $\phi^f(\tau)$ as

$$\phi^f(\bullet_a)(\xi) = f_a(\xi), \quad \phi^f([\tau^1 \cdots \tau^k]_a)(\xi) = \sum_{\bar{b} \in \mathcal{IL}_1 : |\bar{b}|=k} f_{a;b_1 \dots b_k}(\xi) \prod_{i=1}^k [\phi^f(\tau^i)(\xi)]^{b_i}.$$

where $f_{a;b_1 \dots b_k} = \partial_{b_1 \dots b_k}^k f_a$ are the k -th order partial derivatives of the function f_a .

Note that using the regularity of the path x the iterated integrals X^τ can be simplified: indeed Chen [1] proved that products of iterated integrals can be always expressed as linear combination of iterated integrals via the *shuffle product*:

$$(18) \quad X_{ts}^{a_1 \dots a_n} X_{ts}^{b_1 \dots b_m} = \sum_{\bar{c} \in \text{Sh}(\bar{a}, \bar{b})} X_{ts}^{c_1 \dots c_{n+m}}$$

where given two multi-indexes $\bar{a} = (a_1, \dots, a_n)$ and $\bar{b} = (b_1, \dots, b_n)$ their *shuffles* $\text{Sh}(\bar{a}, \bar{b})$ consists of the set of all the possible permutations of the $(n+m)$ -uple $(a_1, \dots, a_n, b_1, \dots, b_m)$ which does not change the ordering of the two subsets \bar{a}, \bar{b} . Using relation (18) every X^τ can be reduced to a linear combination of standard iterated integrals $\{X^{\bar{a}}\}_{\bar{a}}$.

It is, however, interesting that many constructions related to integrals and solution of driven differential equations do not depend on eq. (18) being valid. For example in the theory of the Itô integral eq. (18) does not hold and more complex algebraic relations have to be considered: if x is a multidimensional Brownian motion and X defined via Itô integration we have

$$X_{ts}^a X_{ts}^b = X_{ts}^{ab} + X_{ts}^{ba} + \delta_{ab}(t-s).$$

Let us clarify the algebraic framework in which we consider possible integration theories. It turns out that the only data we really need to build a family $\{X^\tau\}_{\tau \in \mathcal{T}_{\mathcal{L}}}$ satisfying (15) is given by a set of linear maps $\{I^a\}_{a \in \mathcal{L}}$ defined on a domain $\mathcal{D}_I \subset \mathcal{C}_2^+$ (the unital algebra obtained extending \mathcal{C}_2 with a unit e and considering the pointwise product) to \mathcal{C}_2 satisfying two properties:

- (1) $I(hf)_{ts} = I(h)_{ts} f_s$ for all $h \in \mathcal{D}_I, f \in \mathcal{C}_1$ and where $(hf)_{ts} = h_{ts} f_s$;
- (2) $\delta I(h)_{tus} = I(e)_{tu} h_{us} + \sum_i I(h^{1,i})_{tu} h_{us}^{2,i}$ when $h \in \mathcal{D}_I$, with $\delta h_{tus} = \sum_i h_{tu}^{1,i} h_{us}^{2,i}$ and $h^{1,i} \in \mathcal{D}_I$.

Then given a family $\{I^a\}_{a \in \mathcal{L}}$ of such integral maps on a common algebra $\mathcal{D} \subseteq \mathcal{C}_2$ we can associate to them a family $\{X^\tau\}_{\tau \in \mathcal{F}_\mathcal{L}}$ recursively as

$$X_{ts}^{\bullet a} = I^a(e)_{ts}, \quad X_{ts}^{[\tau^1 \cdots \tau^k]_a} = I^a(X^{\tau^1 \cdots \tau^k})_{ts}, \quad X_{ts}^{\tau^1 \cdots \tau^k} = X_{ts}^{\tau^1} \cdots X_{ts}^{\tau^k}.$$

In this way we establish an algebra homomorphism from $\mathcal{AT}_\mathcal{L}$ to a sub-algebra of \mathcal{C}_2 generated by the X^τ -s. This homomorphism send the operation B_+^a on $\mathcal{AT}_\mathcal{L}$ to the integral map I^a on \mathcal{C}_2 in such a way that Theorem 4.1 holds.

Given X^τ for any $\tau \in \mathcal{T}_\mathcal{L}$ with $|\tau| \leq n$ and such that $\|X^\tau\|_{\gamma|\tau|} \leq C$ for some constant C and $\gamma \in (1/(n+1), 1)$ we are able to extend to any τ with $|\tau| > n$ the “integrals” $X^\tau \in \mathcal{C}_2^{\gamma|\tau|}$ in a unique manner solving the equation $\delta X^\tau = X^{\Delta'\tau}$ with the aid of the sewing map. This extension will satisfy the bound

$$(19) \quad \|X^\tau\|_{\gamma|\tau|} \leq C_1 C_2^{|\tau|} q_\gamma(\tau)$$

for any τ with C_1, C_2 two finite constants and $q_\gamma(\tau)$ a function satisfying

$$q_\gamma(\tau) = \frac{1}{2\gamma|\tau| - 2} \sum' q_\gamma(\tau^{(1)}) q_\gamma(\tau^{(2)}).$$

The actual asymptotic behavior of q_γ for $|\tau| \rightarrow \infty$ is not yet known, but we conjecture that it should hold the equivalence

$$(20) \quad q_\gamma(\tau) \simeq C^{|\tau|} (\tau!)^{-\gamma}$$

for some constant C (see [5]). This asymptotic behavior is satisfied on the subset of linear trees and is consistent with our results on tree-indexed iterated integrals in the context of 3d Navier-Stokes equation [6] (see also Sect. 5).

We call the object $\{X^\tau\}_\tau$ satisfying (15) and (19) for some $\gamma \in (0, 1)$ a $(\gamma\text{-})$ branched rough path. By the above consideration it is clear that it is determined by the finite subset $\{X^\tau\}_{\gamma|\tau| \leq 1}$.

Assuming that the integrals $\{I^a\}_{a \in \mathcal{L}}$ generate a γ -branched rough path in the above sense, then we can extend the integral to a larger class of integrands. Take a set of constants $\{h_0^\tau\}_\tau$ and consider the path $h \in \mathcal{C}_1$ defined as $h_t = \sum_\tau h_0^\tau X_{t0}^\tau$, then

$$(21) \quad \delta h_{ts} = \sum_\tau h_0^\tau \sum' X_{ts}^{\tau^{(1)}} X_{s0}^{\tau^{(2)}} = \sum_{\tau, \rho, \sigma} c'(\tau, \rho, \sigma) h_0^\tau X_{ts}^\rho X_{s0}^\sigma = \sum_\tau h_s^\tau X_{ts}^\tau$$

where we introduced paths $h^\tau \in \mathcal{C}_1$ by $h_s^\tau = \sum_{\sigma, \rho, \tau} c'(\rho, \tau, \sigma) h_0^\rho X_{s0}^\sigma$. We must have

$$0 = \delta \sum_\tau X^\tau h^\tau = \sum_\tau X^{\tau^{(1)}} X^{\tau^{(2)}} h^\tau - X^\tau \delta h^\tau$$

so we can check that

$$(22) \quad \delta h^\tau = \sum_{\rho, \tau, \sigma} c'(\rho, \tau, \sigma) X^\sigma h^\rho$$

also holds. At this level these relations are formal since they require an infinite number of terms to hold, however exploiting our analysis of the (\mathcal{C}_*, δ) complex and the sewing map we can work modulo \mathcal{C}_2^{1+} and neglect the terms in the expansions which involve X^τ -s with $\gamma|\tau| > 1$. A more accurate analysis (which can be found in [5]) reveals that for the purpose of integrating and solving differential equations we can actually work modulo larger terms but for the sake of clarity we refrain to do this here.

We call a *controlled path* (by X) any path h which satisfy modulo \mathcal{C}_2^{1+} the eq. (21) together with eq. (22) for all trees τ with $\gamma|\tau| < 1$. The integral I can then be extended to integrate any controlled path. Indeed since $he = eh + \delta h$ it is a consistent definition to set

$$\begin{aligned} I^a(h) &= I^a(he) = I^a(e)h + I^a(\delta h) = X^{\bullet a}h + \sum_{\gamma|\tau|<1} I^a(X^\tau)h^\tau + I^a(\mathcal{R}) \\ &= X^{\bullet a}h + \sum_{\gamma|\tau|<1} X^{[\tau]_a}h^\tau + \mathcal{R} \end{aligned}$$

where we exploited the properties of the integral I^a and the definition of X and where $\mathcal{R} \in \mathcal{C}_2^{1+}$ is a generic remainder term (possibly different from line to line). Controlled paths are also stable under mapping by sufficiently smooth functions f (the degree of differentiability depends on γ): i.e. if y is a controlled path, then $z = f(y)$ is again a controlled path with an explicit formula for the coefficients z^τ in term of the derivatives of f and of the coefficients y^τ . Using these properties we can consider the differential equation for the integrals I and vectorfield f :

$$\delta y = I^a(f_a(y)), \quad y_0 = \eta$$

in the space of controlled paths y and solve it by a fixed-point argument.

5. Infinite dimensional dynamical systems

In this section we would like to show how the ideas and the tools described before could be applied in the context of infinite dimensional dynamical system by introducing operator-valued iterated integrals (and rough paths) which by their non-commutative nature are intrinsically indexed by trees. In particular we will discuss two different examples: the 1d periodic Korteweg-de-Vries (KdV) equation and the 3d Navier Stokes (NS) equation. In the first case we will exploit the increment complex to analyze perturbatively the solution for irregular initial data, in the second case we will use series expansion over trees of the solution to analyze the dynamics for large times .

As before we will restrict ourself to an overview of the results and to a sketch of the arguments. The interested readers can find a rigorous discussion elsewhere [7, 6] and he can refer to [14] for extension of these consideration covering the analysis of stochastic partial differential equations driven by irregular noises.

5.1. The KdV equation. The 1d periodic KdV equation is the partial differential equation

$$(23) \quad \partial_t u(t, \xi) + \partial_\xi^3 u(t, \xi) + \frac{1}{2} \partial_\xi u(t, \xi)^2 = 0, \quad u(0, \xi) = u_0(\xi), \quad (t, \xi) \in \mathbb{R} \times \mathbb{T}$$

on the torus $\mathbb{T} = [-\pi, \pi]$. When the initial data is not smooth this equation must be interpreted in the integral form

$$u(t) = U(t)u_0 + \frac{1}{2} \int_0^t U(t-s) \partial[u(s)]^2 ds$$

where U is the “free propagator” given by the linear equation $\partial_t U(t) = \partial^3 U(t)$, ∂ denoting the spatial derivative, and where we consider the solution u as a path in

a space of functions. By going to the interaction picture $\tilde{v}_t = U(-t)u(t)$ we get

$$\tilde{v}_t = u_0 + \frac{1}{2} \int_0^t U(-s) \partial [U(s) \tilde{v}_s]^2 ds$$

Then the Fourier coefficients $\{v_t(k)\}_{k \in \mathcal{Z}}$ of \tilde{v}_t satisfy the equation
(24)

$$v_t(k) = v_0(k) + \frac{ik}{2} \sum_{k=k_1+k_2, k_n \neq 0} \int_0^t e^{-i(k^3 - k_1^3 - k_2^3)s} v_s(k_1) v_s(k_2) ds, \quad t \in [0, T], k \neq 0.$$

We restrict our attention to initial conditions such that $v_0(0) = 0$. By calling \dot{X} the bilinear operator in the r.h.s. this equation takes the abstract form

$$v_t = v_s + \int_s^t \dot{X}_\sigma(v_\sigma, v_\sigma) d\sigma, \quad t, s \in [0, T].$$

where the paths take values in the Hilbert space H_α of complex Fourier coefficients $\varphi(k)$ with $\varphi(0) = 0$ and $\varphi(-k) = \overline{\varphi(k)}$ endowed with the scalar product $\langle \varphi_1, \varphi_2 \rangle_\alpha = \sum_{k \neq 0} k^{2\alpha} \varphi_1(-k) \varphi_2(k)$.

By iteratively substituting the unknown in this integral equation we obtain an expansion whose first terms looks like

$$(25) \quad \begin{aligned} v_t &= v_s + \int_s^t d\sigma \dot{X}_\sigma(v_s, v_s) + 2 \int_s^t d\sigma \dot{X}_\sigma(v_s, \int_s^\sigma d\sigma_1 \dot{X}_{\sigma_1}(v_s, v_s)) \\ &\quad + \int_s^t d\sigma \dot{X}_\sigma(\int_s^\sigma d\sigma_1 \dot{X}_{\sigma_1}(v_s, v_s), \int_s^\sigma d\sigma_2 \dot{X}_{\sigma_2}(v_s, v_s)) \\ &\quad + 4 \int_s^t d\sigma \dot{X}_\sigma(v_s, \int_s^\sigma d\sigma_1 \dot{X}_{\sigma_1}(v_s, \int_s^{\sigma_1} d\sigma_2 \dot{X}_{\sigma_2}(v_s, v_s))) + r_{ts} \end{aligned}$$

where r_{ts} stands for the remaining terms in the expansion. Denote with \mathcal{BT} the set of (unlabeled) planar rooted trees with at most two branches at each node. A planar tree is a rooted tree endowed with an ordering of the branches at each node. Then each of the terms in this expansion can be associated to a tree in \mathcal{BT} and we can define recursively multi-linear operators X^τ as

$$\begin{aligned} X_{ts}^\bullet(\varphi_1, \varphi_2) &= \int_s^t \dot{X}_\sigma(\varphi_1, \varphi_2) d\sigma; \\ X_{ts}^{[\tau^1]}(\varphi_1, \dots, \varphi_{m+1}) &= \int_s^t \dot{X}_\sigma(X_{\sigma s}^{\tau^1}(\varphi_1, \dots, \varphi_m), \varphi_{m+1}) d\sigma \end{aligned}$$

and

$$X_{ts}^{[\tau^1 \tau^2]}(\varphi_1, \dots, \varphi_{m+n}) = \int_s^t \dot{X}_\sigma(X_{\sigma s}^{\tau^1}(\varphi_1, \dots, \varphi_m), X_{\sigma s}^{\tau^2}(\varphi_{m+1}, \dots, \varphi_{m+n})) d\sigma.$$

Eq. (25) has the form

$$(26) \quad \delta v_{ts} = X^\bullet(v^{\times 2})_{ts} + X^\bullet(v^{\times 3})_{ts} + X^\bullet(v^{\times 4})_{ts} + X^\bullet(v^{\times 4})_{ts} + r_{ts}$$

as an equation in \mathcal{C}_2 where increments take values in H_α and where we let $v_s^{\times n} = (v_s, \dots, v_s)$ (n times). The operators X^τ satisfy multiplicative relations

$$\delta X^\bullet(\varphi_1, \varphi_2, \varphi_3) = X^\bullet(X^\bullet(\varphi_1, \varphi_2), \varphi_3),$$

$$\delta X^\bullet(\varphi_1, \varphi_2, \varphi_3, \varphi_4) = X^\bullet(X^\bullet(\varphi_1, \varphi_2, \varphi_3), \varphi_4) + X^\bullet(X^\bullet(\varphi_1, \varphi_2), \varphi_3, \varphi_4),$$

and

$$\begin{aligned}\delta X^{\bullet}(\varphi_1, \varphi_2, \varphi_3, \varphi_4) &= X^{\bullet}(X^{\bullet}(\varphi_1, \varphi_2), X^{\bullet}(\varphi_3, \varphi_4)) \\ &\quad + X^{\bullet}(\varphi_1, \varphi_2, X^{\bullet}(\varphi_3, \varphi_4)) + X^{\bullet}(\varphi_3, \varphi_4, X^{\bullet}(\varphi_1, \varphi_2))\end{aligned}$$

where we used the symmetry of the operator \dot{X} to obtain this last equation. These relations have much in common with the analogous relations for branched rough paths, however here the additional information of the position of the various arguments must be taken into account in the combinatorics of the reduced coproduct. It would be interesting to determine a Hopf algebra structure on \mathcal{BT} which could account this additional information in a general way.

Note that \dot{X} is unbounded on H_{α} while it is possible to prove that X^{\bullet} and X^{\bullet} are bounded, in particular for the simplest of them, X^{\bullet} , we have the bound

$$|X_{ts}^{\bullet}(\varphi_1, \varphi_2)|_{\alpha} \leq C|t-s|^{\gamma}|\varphi_1|_{\alpha}|\varphi_2|_{\alpha}$$

with the parameters γ and α satisfying

$$(27) \quad \gamma < 1/2, \alpha \in \mathbb{R} \text{ with } \gamma - \alpha < 1 \text{ and } 3\gamma - \alpha < 3/2.$$

Moreover we are able to decompose the second member of the hierarchy X^{\bullet} as

$$(28) \quad X^{\bullet}(\varphi_1, \varphi_2, \varphi_3) = \hat{X}^{\bullet}(\varphi_1, \varphi_2, \varphi_3) + \Phi(\varphi_1, \varphi_2)\varphi_3 + \Phi(\varphi_1, \varphi_3)\varphi_2$$

where for the same range of γ, α as in (27) we have

$$|\hat{X}_{ts}^{\bullet}(\varphi_1, \varphi_2, \varphi_3)|_{\alpha} \leq C|t-s|^{2\gamma}|\varphi_1|_{\alpha}|\varphi_2|_{\alpha}|\varphi_3|_{\alpha}$$

while the operator $\Phi_{ts} : H_{\alpha} \times H_{\alpha} \rightarrow \mathcal{C}$ is bounded only for $\alpha \geq -1/2$ and

$$|\Phi_{ts}(\varphi_1, \varphi_2)| \leq C|t-s| |\varphi_1|_{\alpha} |\varphi_2|_{\alpha}.$$

The analysis of the higher order operators has not yet been performed. However already at this stage something can be said if we take $3\gamma > 1$ since we are naturally led to consider eq. (26) as a increment equation and rewrite it using the sewing map and the operators X (only up to second order) obtaining the equation

$$(29) \quad \delta v = (1 - \Lambda\delta)[X^{\bullet}(v^{\times 2}) + X^{\bullet}(v^{\times 3})]$$

which can be solved by fixed point methods in H_{α} for any $\alpha > -1/2$ (cfr. (27) and the condition on Φ).

The condition $\alpha \geq -1/2$ is essentially imposed by the operator Φ appearing in the decomposition (28) of X^{\bullet} . This constraint it is linked with a resonance phenomenon which appears in the scattering by the non-linear term involving four waves and which hint to the fact that the KdV equation is not uniformly well-posed in H_{α} for $\alpha < -1/2$ [15, 16].

We have replaced the differential and integral approach to the study of this equation by an approach based on an operator valued rough path and the increment complex. It is interesting then to look how the properties of the dynamical system reflect in this unusual approach. As an example let us consider conservation laws.

The KdV equation formally conserves the H_0 norm. This conservation law imposes additional algebraic relations to the operators X : it is not difficult to prove that we have

$$(30) \quad \langle \varphi_1, X_{ts}^{\bullet}(\varphi_2, \varphi_3) \rangle_0 + \langle \varphi_2, X_{ts}^{\bullet}(\varphi_1, \varphi_3) \rangle_0 + \langle \varphi_3, X_{ts}^{\bullet}(\varphi_1, \varphi_2) \rangle_0 = 0$$

and that

$$(31) \quad 2\langle \varphi, X_{ts}^2(\varphi, \varphi, \varphi) \rangle_0 + \langle X_{ts}(\varphi, \varphi), X_{ts}(\varphi, \varphi) \rangle_0 = 0,$$

where all the test functions belong to H_0 . To see that these two relations imply the H_0 conservation law for solutions we will prove that $\delta\langle v, v \rangle_0 = 0$ when v satisfy

$$(32) \quad \delta v = X^\bullet(v^{\times 2}) + X^\bullet(v^{\times 3}) + \mathcal{R}$$

cfr. eq. (29). Let us compute explicitly: $[\delta\langle v, v \rangle_0]_{ts} = \langle v_t, v_t \rangle_0 - \langle v_s, v_s \rangle_0 = 2\langle \delta v_{ts}, v_s \rangle_0 + \langle \delta v_{ts}, \delta v_{ts} \rangle_0$. Substituting in this expression the equation (32) we get

$$\begin{aligned} [\delta\langle v, v \rangle_0]_{ts} &= 2\langle X_{ts}^\bullet(v_s, v_s) + X_{ts}^\bullet(v_s, v_s, v_s), v_s \rangle_0 \\ &\quad + \langle X_{ts}^\bullet(v_s, v_s), X_{ts}^\bullet(v_s, v_s) \rangle_0 + \mathcal{R}. \end{aligned}$$

The relation (30) implies that $\langle v_s, X_{ts}^\bullet(v_s, v_s) \rangle_0 = 0$ while eq. (31) allows to cancel the X^\bullet term with the quadratic X^\bullet term. After the cancellations the increment of the H_0 norm squared is then $[\delta\langle v, v \rangle_0] \in \mathcal{C}_2^{1+}$ but this means that it must be zero and that $\langle v_t, v_t \rangle_0 = \langle v_0, v_0 \rangle_0$ for any $t \geq 0$.

5.2. Navier-Stokes-like equations. We consider the NS equation in \mathbb{R}^3 which going in Fourier space can be written

$$(33) \quad v_t(k) = e^{-|k|^2 t} v_0(k) + i \int_0^t e^{-|k|^2(t-s)} \int_{\mathbb{R}^3} dk' \langle k, v_s(k-k') \rangle P_k v_s(k') ds$$

where v_t is the Fourier transform of the velocity field, $\langle \cdot, \cdot \rangle$ is the scalar product in \mathbb{C}^3 and $P_k : \mathcal{C}^3 \rightarrow \mathcal{C}^3$ is the projection on the directions orthogonal to the vector $k \in \mathbb{R}^3$, i.e. $P_k a = a - \langle k, a \rangle k |k|^{-1}$. Eq. (33) will be studied in the spaces $\Phi(\alpha)$, $\alpha \in [2, 3]$ where $v \in \Phi(\alpha)$ if $v \in C(\mathbb{R}^3; \mathcal{C}^3)$ with $k \cdot v(k) = 0$ and $\|v\|_\alpha = \sup_{k \in \mathbb{R}^3} |k|^\alpha |v(k)| < \infty$. We will write $\alpha = 2 + \varepsilon$ with $\varepsilon \in [0, 1)$. The spaces $\Phi(\alpha)$ can contain solutions with infinite energy and enstrophy so classical results about existence and uniqueness do not apply. Sinai [17, 18, 19], studied eq. (33) in $\Phi(\alpha)$ with $\alpha > 2$, showing that there is existence of unique local solutions and that these solutions survive for arbitrary large time if the initial condition is small enough. Related works on NS are those of Le Jan and Sznitman [20], Cannone and Planchon [21] and the reviews of Bhattacharya et al. in [22] and Waymire [23].

Following [6] we will describe the representation for these solutions as series indexed by planar binary trees. The use of trees to rigorously analyze the NS equation has been somewhat pioneered by Gallavotti [24].

The NS equation can be cast in the abstract form

$$(34) \quad u_t = S_t u_0 + \int_0^t S_{t-s} B(u_s, u_s) ds.$$

where S is a bounded semi-group on $\Phi(\alpha)$ and B is a symmetric bilinear operator which is usually defined only on a subspace of $\Phi(\alpha)$. Here we cannot proceed as in the KdV case by going to the interaction picture since S is only a semi-group, so we must cope with the convolution directly. In [6] we showed that the solutions of this equation in the case of the 3d NS equation have the norm convergent series representation

$$(35) \quad u_t = S_t u_0 + \sum_{\tau \in \mathcal{BT}} X_{t0}^\tau(u_0^{\times \theta(\tau)})$$

where $\theta(\tau)$ is a degree function defined by $\theta(\bullet) = 2$, $\theta([\tau]) = 1 + \theta(\tau)$, $\theta([\tau_1 \tau_2]) = \theta(\tau^1) + \theta(\tau^2)$ and the $\theta(\tau)$ -multilinear operators X^τ have recursive definition

$$X_{ts}^\bullet(\varphi^{\times 2}) = \int_s^t S_{t-u} B(S_{u-s} \varphi, S_{u-s} \varphi) du$$

$$X_{ts}^{[\tau^1]}(\varphi^{\times(\theta(\tau^1)+1)}) = \int_s^t S_{t-u} B(X_{us}^{\tau^1}(\varphi^{\times\theta(\tau^1)}), \varphi) du$$

and

$$X_{ts}^{[\tau^1 \tau^2]}(\varphi^{\times(\theta(\tau^1)+\theta(\tau^2))}) = \int_s^t S_{t-u} B(X_{us}^{\tau^1}(\varphi^{\times\theta(\tau^1)}), X_{us}^{\tau^2}(\varphi^{\times\theta(\tau^2)})) du$$

and by induction we can prove that, for any $\varepsilon \in [0, 1)$ these operators are bounded by

$$(36) \quad |X_{ts}^\tau(h^{\theta(\tau)})(k)| \leq C_\tau \frac{e^{-|k|^2(t-s)/(|\tau|+1)}}{|k|^\alpha} (t-s)^{|\tau|\varepsilon/2} \|h\|_\alpha^{\theta(\tau)}$$

where the constants C_τ can be chosen as

$$(37) \quad C_\tau = A^{|\tau|} (\tau!)^{-\varepsilon/2}.$$

for some other constant $A > 0$ depending only on ε .

Due to the presence of the convolution integral these X operators does not behaves nicely with respect to the coboundary δ . In [14] we introduced cochain complex $(\tilde{C}_*, \tilde{\delta})$ adapted to the study of such convolution integrals where the coboundary $\tilde{\delta}$ is obtained from δ by a "twisting" involving the semigroup. There exists also a corresponding convolutional sewing map $\tilde{\Lambda}$ which provide an appropriate inverse to $\tilde{\delta}$. Algebraic relations for these iterated integrals have then by-now familiar expressions, e.g.:

$$\tilde{\delta} X^\bullet(\varphi^{\times 3}) = X^\bullet(X^\bullet(\varphi^{\times 2}), S\varphi)$$

and so on. Indeed X can be considered as a branched rough path and the asymptotic behavior (37) supports somewhat the conjectured asymptotics (20).

The series representation (35) together with the bounds (36) imply that

$$(38) \quad |u_t(k)| \leq |S_t u_0(k)| + \sum_{\tau \in \mathcal{BT}} \frac{B^{|\tau|}}{\sigma(\tau) \tau!} \frac{e^{-(k^2 t)/(|\tau|+1)}}{|k|^\alpha} t^{\varepsilon|\tau|/2} \|u_0\|_\alpha^{\theta(\tau)}$$

By induction we can prove that $\gamma(\tau) \geq 2^{|\tau|-1}$ where equality holds for the binary trees for which every path from the root to the leaves has the same length. This estimate together with

$$(39) \quad (|\tau|+1)/2 \leq \theta(\tau) \leq |\tau|+1$$

easily proven by induction on $|\tau|$ give

$$(40) \quad |u_t(k)| \leq |S_t u_0(k)| + \sum_{n \geq 1} Z_n B^n \frac{e^{-(k^2 t)/(n+1)}}{|k|^\alpha} t^{\varepsilon n/2} \|u_0\|_\alpha^{(n+1)/2} (1 + \|u_0\|_\alpha)^{(n+1)/2}$$

for some different constant B and where Z_n is the number of trees in \mathcal{BT} with n vertices for which we have the estimate $Z_n \leq D^n (n+1)^{-3/2}$ for a constant $D > 0$. So the series (35) is controlled by the geometric series (40) and converges in norm t is small or $\|h\|_\alpha$ is small. In the case $\varepsilon = 0$ the dependence in time of the r.h.s.

is bounded and so the series converges for all time if the initial condition is small enough. The series gives also additional informations on the global solution when $\varepsilon = 0$:

- a) *dissipation*: for fixed $k \in \mathbb{R}^3 \setminus \{0\}$, $\lim_{t \rightarrow \infty} |u_t(k)| = 0$;
- b) *smoothness*: for fixed $t > 0$, there exists two constants C_3, C_4 such that $|v_t(k)| \leq C_3 e^{-C_4 |k| \sqrt{t}}$ as $|k| \rightarrow \infty$.

In particular this second property can be proved by the Laplace method applied to the majorizing series (40).

Another interesting way to analyze the NS series (35) is to note that different classes of trees give different contributions. We define *simple* trees the trees with at most one branch at each vertex, i.e. of the form $[\cdots [\bullet] \cdots]$. *Short* trees are instead trees for which at each vertex we have two branches, each of which carries (asymptotically) a fixed proportion (α or $1 - \alpha$) of the vertice and without loosing generality we consider $\alpha \in (0, 1/2)$. We will denote \mathcal{BT}_0 the set of simple trees and \mathcal{BT}_α the set of short trees corresponding to the proportion α . The distinction between these classes of trees is relevant when discussing the asymptotic behavior of the tree factorial. Indeed for $\tau \in \mathcal{BT}_0$ we have $\tau! = |\tau|!$ while for any $\alpha \in (0, 1/2)$ there exists constants D_1, D_2, D_3, D_4 such that, for any $\tau \in \mathcal{BT}_\alpha$ we have

$$D_3 |\tau|^{-1} D_4^{|\tau|} \leq \gamma(\tau) \leq D_1 |\tau|^{-1} D_2^{|\tau|}.$$

This different behavior is responsible for different convergence properties of the sum (35) when restricted to simple or short trees. Indeed it is possible to prove that the series restricted to simple trees is convergent for all times whatever the size of the initial condition while the estimate for short trees do not ensure this important property. In some sense the difficulty of finding global solution of NS is due to the presence of arbitrarily large short trees in the expansion. A similar phenomenon is observed in [19] and exploited in [25] to prove blow-up for the complex Navier-Stokes equation by a renormalization group argument.

References

- [1] Kuo Tsai Chen. Iterated path integrals. *Bull. Amer. Math. Soc.*, 83(5):831–879, 1977.
- [2] Terry J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.
- [3] Terry Lyons and Zhongmin Qian. *System control and rough paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002. Oxford Science Publications.
- [4] Antoine Lejay. An introduction to rough paths. In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, pages 1–59. Springer, Berlin, 2003.
- [5] M. Gubinelli. Ramification of rough paths. *J. Diff. Eq.*, 2008. to appear.
- [6] Massimiliano Gubinelli. Rooted trees for 3D Navier-Stokes equation. *Dyn. Partial Differ. Equ.*, 3(2):161–172, 2006.
- [7] M. Gubinelli. Rough solutions of the periodic Korteweg-de Vries equation. 2006.
- [8] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.
- [9] Professor Cayley. On the analytical forms called trees. *American Journal of Mathematics*, 4(1/4):266–268, 1881.
- [10] J. C. Butcher. An algebraic theory of integration methods. *Math. Comp.*, 26:79–106, 1972.
- [11] E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing (Arch. Elektron. Rechnen)*, 13(1):1–15, 1974.
- [12] Ch. Brouder. Trees, renormalization and differential equations. *BIT*, 44(3):425–438, 2004.
- [13] Alain Connes and Dirk Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Comm. Math. Phys.*, 199(1):203–242, 1998.
- [14] M. Gubinelli and S. Tindel. Rough evolution equations. 2006.

- [15] Michael Christ, James Colliander, and Terrence Tao. Asymptotics, frequency modulation, and low regularity ill-posedness for canonical defocusing equations. *Amer. J. Math.*, 125(6):1235–1293, 2003.
- [16] J. Colliander, M. Keel, G. Staffilani, H. Takaoka, and T. Tao. Sharp global well-posedness for KdV and modified KdV on \mathbb{R} and T . *J. Amer. Math. Soc.*, 16(3):705–749 (electronic), 2003.
- [17] Ya. G. Sinai. On local and global existence and uniqueness of solutions of the 3D Navier-Stokes system on \mathbb{R}^3 . In *Perspectives in analysis*, volume 27 of *Math. Phys. Stud.*, pages 269–281. Springer, Berlin, 2005.
- [18] Yakov Sinai. Power series for solutions of the 3D-Navier-Stokes system on \mathbb{R}^3 . *J. Stat. Phys.*, 121(5-6):779–803, 2005.
- [19] Ya. G. Sinaĭ. A diagrammatic approach to the 3D Navier-Stokes system. *Uspekhi Mat. Nauk*, 60(5(365)):47–70, 2005.
- [20] Y. Le Jan and A. S. Sznitman. Stochastic cascades and 3-dimensional Navier-Stokes equations. *Probab. Theory Related Fields*, 109(3):343–366, 1997.
- [21] Marco Cannone and Fabrice Planchon. On the regularity of the bilinear term for solutions to the incompressible Navier-Stokes equations. *Rev. Mat. Iberoamericana*, 16(1):1–16, 2000.
- [22] Rabi N. Bhattacharya, Larry Chen, Scott Dobson, Ronald B. Guenther, Chris Orum, Mina Ossiander, Enrique Thomann, and Edward C. Waymire. Majorizing kernels and stochastic cascades with applications to incompressible Navier-Stokes equations. *Trans. Amer. Math. Soc.*, 355(12):5003–5040 (electronic), 2003.
- [23] Edward C. Waymire. Probability & incompressible Navier-Stokes equations: an overview of some recent developments. *Probab. Surv.*, 2:1–32 (electronic), 2005.
- [24] Giovanni Gallavotti. *Foundations of fluid dynamics*. Texts and Monographs in Physics. Springer-Verlag, Berlin, 2002. Translated from the Italian.
- [25] Dong Li and Ya. G. Sinai. Blow ups of complex solutions of the 3D Navier-Stokes system and renormalization group method. *J. Eur. Math. Soc. (JEMS)*, 10(2):267–313, 2008.

CEREMADE (CNRS UMR 7534) UNIVERSITÉ PARIS DAUPHINE, FRANCE

E-mail address: gubinelli@ceremade.dauphine.fr

This page intentionally left blank

Rooted trees appearing in products and co-products

Ralf Holtkamp

ABSTRACT. We review basic concepts related to rooted trees and their combinatorics as they are needed for the introduction of several bialgebra or Hopf algebra structures on vector space bases of rooted trees. In view of the applications, we have to distinguish between several notions of such trees, such as abstract/planar, partially labeled/non-labeled etc. We present a unifying approach, inspired by the theory of operads. Then we focus on important operations on trees, such as grafting and cutting trees, which appear in various examples.

1. Introduction

Rooted trees are examples of combinatorial objects which appear in a lot of different contexts in mathematics. Their appearance as algebra bases or vector space bases for various interesting Hopf algebra structures is one aspect, which still leads to a lot of new research. We mention here the Hopf algebras of A. Connes and D. Kreimer (cf. [2]), the Hopf algebras of R. Grossman and R. G. Larson (cf. [5]), the dendriform Hopf algebras of J.-L. Loday and M. Ronco (cf. [14]), and Hopf algebra structures introduced by C. Brouder and A. Frabetti (cf. [1]).

The aim of this note is not to delve into Hopf algebras, but to describe the foundations, i.e. to present an approach to rooted trees useful for the study of the mentioned Hopf algebras.

Our goal in Section 2 is to put several different types of rooted trees – planar (ordered), or abstract (unordered), usually partially labeled – under a common roof, which is inspired by operad theory. In Section 3 we look at products or operations on trees, which are related to the concatenation of words. In particular, we look at parenthesized words, which correspond directly to binary trees. The relevant operations like the grafting product on trees extend to operations $V \otimes V \rightarrow V$ or $(V \otimes \dots \otimes V) \rightarrow V$, where V is some vector space spanned by trees or forests of trees. Dually defined are co-operations on V (into tensor products of V). In Section 4, we look at the substitution operation for trees. There are various related operations. We use a right comb presentation of planar binary trees, and we give an explicit bijection between planar forests with $n - 1$ vertices and corresponding planar binary trees with n leaves. Then we concentrate on the concept of cuts, as it is used for the coproduct of the Connes-Kreimer Hopf algebra. We finish with some remarks on nonsymmetric operads and the Stasheff polytopes.

The author thanks the referee for helpful suggestions.

2. Types of rooted trees - an overview

We distinguish between several types of trees. First we recall the notions of rooted trees and planar rooted trees (compare [21, 22]). We skip the definition of graphs (see [6]). A naive notion of a graph will suffice. For a more sophisticated notion, involving half-edges, see [16], §5.3.

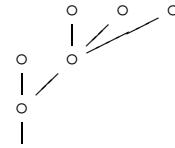
2.1. Rooted Trees. A finite connected graph $\emptyset \neq T = (\text{Ve}(T), \text{Ed}(T))$ with a distinguished vertex ρ_T is called a rooted tree, if the following condition holds: For every vertex $\lambda \in \text{Ve}(T)$ there is exactly one path connecting λ and ρ_T (equivalently, T has no cycles, compare the definition given in the appendix of [21]).

The vertex ρ_T is called the root of T . Thinking of the edges as oriented towards the root, at each vertex there are incoming edges and one outgoing edge. The standard convention is that the root has no outgoing edges, but here we add to the root an outgoing edge that is not connected to any further vertex. If we want to exclude this edge, we speak of the other edges as inner edges.

If λ and λ' in $\text{Ve}(T)$ are connected by an edge oriented from λ' to λ , then λ is called the father of λ' , and λ' is called a child of λ .

The vertices with no child are called leaves, and the set of leaves of T is denoted by $\text{Le}(T)$.

2.2. Example. The following graph is a rooted tree:



Here we draw the root at the bottom. If the number $\#\text{Ve}(T)$ of vertices of T is at least 2, T is uniquely described by the unordered list of non-empty rooted trees T_1, \dots, T_r , $r \geq 1$, of the full subtrees whose roots are the children of the root ρ_λ of T (cf. Definition 2.16).

2.3. Planar rooted trees. Let T be a rooted tree. For any given vertex λ of T , we say that λ is an r -ary vertex, if the number of incoming edges is r .

Binary trees are rooted trees where all vertices are 2-ary, except for the leaves, which are 0-ary. Analogously, m -ary trees are defined.

For T a rooted tree of any type, we write the set $\text{Ve}(T)$ of vertices as a disjoint union $\bigcup_{r \in \mathbb{N}} \text{Ve}^r(T)$, where $\text{Ve}^r(T)$ consists of all r -ary vertices of T . The elements of $\text{Ve}^*(T) = \bigcup_{r \geq 1} \text{Ve}^r(T) = \text{Ve}(T) - \text{Le}(T)$ are called internal vertices of T .

The height of a vertex $\lambda \in \text{Ve}(T)$ is the number of edges separating it from ρ_T . The height of a rooted tree T is the maximum height of its vertices.

A rooted tree T together with a chosen order of incoming edges at each vertex is called a planar rooted tree (or ordered rooted tree). In our drawings, this is an ordering from left to right, see Example (2.4).

Planar binary trees and planar m -ary trees are defined as well. One also considers incomplete planar m -ary trees, defined recursively as follows: Every such tree is either the empty tree or uniquely determined by the ordered list of m possibly empty incomplete planar m -ary trees, see Example (2.5).

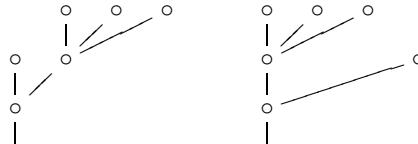
Ordinary rooted trees are also called abstract (or unordered) rooted trees to stress that they are non-planar.

Consequently, we denote by PTree the set of planar rooted trees, and by ATree the set of abstract rooted trees. Some authors also use the notations $OT :=$ PTree and $UT :=$ ATree; OT for ordered trees, UT for unordered trees.

By A_m Tree and P_m Tree we denote the corresponding sets of m -ary (complete) rooted trees, and in the binary case $m = 2$ we also use the notation ABTree and PBTree.

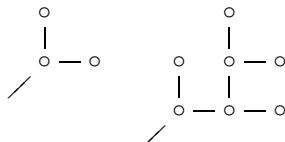
2.4. Example. For every vertex of a planar rooted tree, the chosen order of incoming edges corresponds to an ordering of edges from left to right. Every drawing of a rooted tree provides us with a planar structure, which we have to forget when dealing with abstract rooted trees.

The drawings

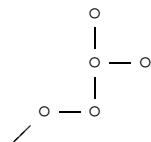


represent the same abstract rooted tree T (of height 2), but different planar rooted trees T^1, T^2 .

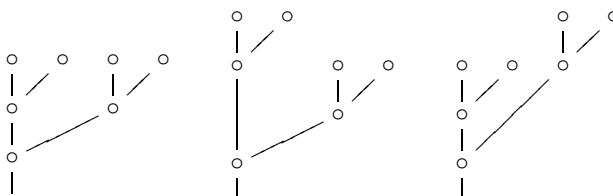
2.5. Example. Examples for binary trees can be drawn as follows:



An incomplete binary tree is represented by the following picture:



2.6. Leveled trees. We mention that one can also associate levels to all vertices of a given rooted tree, to distinguish for example between the rooted trees



Such leveled trees and also incomplete trees are not going to be considered in the rest of this article.

2.7. Integer sequences.

For $n, p \in \mathbb{N}$, let

$$\text{PTree}_n := \{T \in \text{PTree} : \#\text{Ve}(T) = n\}$$

be the set of planar rooted trees with n vertices, and let

$$\text{PTree}^p := \{T \in \text{PTree} : \#\text{Le}(T) = p\}$$

be the set of planar rooted trees with p leaves.

Furthermore, let

$$\text{PTree}_n^p := \text{PTree}_n \cap \text{PTree}^p.$$

We make the analogous definitions for abstract rooted trees, and also for the abstract binary and for the planar binary rooted trees.

The number $\#\text{PTree}_n$ of planar rooted trees with n vertices is the n -th Catalan number

$$c_n = \frac{(2(n-1))!}{n!(n-1)!}.$$

The numbers c_n , for $n = 1, \dots, 11$, are

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796.$$

with generating series $f(t) = \sum_{n=1}^{\infty} c_n t^n$ given by

$$\frac{1 - \sqrt{1 - 4t}}{2}.$$

The numbers c_n also count planar binary rooted trees with n leaves.

The numbers $a_n = \#\text{ATree}_n$ of abstract rooted trees with n vertices, for $n = 1, \dots, 11$, are

$$1, 1, 2, 4, 9, 20, 48, 115, 286, 719, 1842.$$

The generating series $f(t) = \sum_{n=1}^{\infty} a_n t^n$ fulfills the equation

$$f(t) = \frac{t}{\prod_{n \geq 1} (1 - t^n)^{a_n}}$$

or equivalently the equation

$$f(t) = t \exp \left(\sum_{k \geq 1} \frac{f(t^k)}{k} \right)$$

(cf. [6]).

The numbers $b_n = \#\text{ABTree}^n$ of abstract binary rooted trees with n leaves, for $n = 1, \dots, 11$, are

$$1, 1, 1, 2, 3, 6, 11, 23, 46, 98, 207, 451$$

with generating series $f(t)$ given by (cf. [18]) the equation

$$f(t) = t + \frac{1}{2} f(t)^2 + \frac{1}{2} f(t^2).$$

The series can thus be written in the form

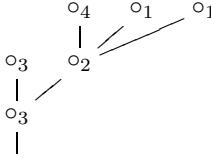
$$\begin{aligned} f(t) &= 1 - \sqrt{1 - f(t^2) - 2t} \\ &= 1 - \sqrt{\sqrt{1 - f(t^4) - 2t^2} - 2t} = \dots \end{aligned}$$

Information on these integer sequences can be found in the On-Line Encyclopedia of Integer Sequences [20] (see Sequence A000108 for c_n , Sequence A000081 for a_n , Sequence A001190 for b_n).

2.8. Labeled trees. Let M be a set and T a rooted tree (of a given type). The set M is considered to be a set of labels (or colors).

Then a labeling of T is a map $\nu : \text{Ve}(T) \rightarrow M$. The rooted tree T together with such a labeling is called a labeled rooted tree (of the given type).

2.9. Example. In our drawing, we can put labels at the vertices:



Here we have used the set $M = \{1, 2, 3, 4\}$.

2.10. Admissible labelings. Let M be a labeling $\nu : \text{Ve}(T) \rightarrow M$ as above. Suppose that $M = \bigcup_{r \in \mathbb{N}} M_r$, where M_0, M_1, M_2, \dots is a given sequence of sets.

Then the labeling ν is called (M) -admissible, if $\nu(\lambda) \in M_r$ for every r -ary vertex λ (for every r).

The set of planar rooted trees $T \in \text{PTree}$ with admissible labeling from $(M_r)_{r \in \mathbb{N}}$ is denoted by $\text{PTree}((M_r))$ or simply by $\text{PTree}(M)$. Similarly, $\text{ATree}(M)$ is used for abstract rooted trees with (M) -admissible labelings.

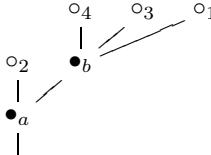
2.11. n -trees. Suppose furthermore that $M_0 = \mathbb{N}$ and that only such labelings are allowed that map the leaves of any rooted tree with n leaves bijectively on the set $\{1, 2, \dots, n\}$. Planar or abstract rooted trees with such an (M) -admissible labeling are called (planar or abstract) n -trees.

A (planar or abstract) rooted tree T is called reduced, if $\text{Ve}^1(T) = \emptyset$, i.e. if there are no 1-ary vertices in T .

The set of planar reduced rooted trees is denoted by PRTree .

2.12. Example. Let $M_1 = \emptyset$. Then rooted trees (of any type) with an (M) -admissible labeling are necessarily reduced.

The following drawing represents a planar n -tree with $a \in M_2, b \in M_3$:



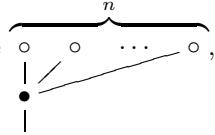
Note that, in our drawing, we used different colors/labels $\bullet_a, \bullet_b, \dots$ for internal vertices and \circ_1, \circ_2, \dots for the leaves.

We can identify the set PTree with the set $\text{PTree}((M_r = \{\circ\}, r \in \mathbb{N}))$, i.e. we consider non-labeled rooted trees as trivially labeled trees.

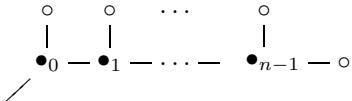
2.13. Example. The following abstract (or planar) rooted tree is not reduced:



one of its vertices is 1-ary.

The rooted tree , called n -corolla, is reduced for every $n \geq 2$.

A planar binary tree of the form

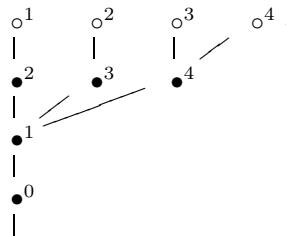


is called a right comb of height n . The right comb without labels of height n is denoted by R_n . One considers left combs as well. As abstract trees, the right comb and the left comb (of a given height) are equal.

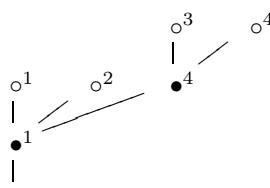
2.14. Reduction map. For any (planar or abstract) rooted tree T , there exists a unique rooted tree $\text{red}(T)$, called the reduction of T , defined as follows: Its set of vertices is $\text{Ve}(T) - \text{Ve}^1(T)$, and for any pair λ, λ' of vertices of $\text{red}(T)$ there is an oriented path (or, equivalently, a path not passing the root) from λ to λ' in $\text{red}(T)$ if and only if there is such a path in T .

Induced is a map $\text{red} : \text{PTree} \rightarrow \text{PRTree}$ and a similar map for abstract rooted trees. Admissible labelings (of all vertices of arity $\neq 1$) are preserved.

2.15. Example. Consider the rooted tree



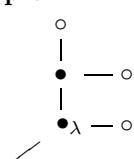
Its reduction is



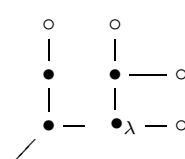
2.16. Full subtrees. Let T be a planar (or abstract) rooted tree, and let $\lambda \in \text{Ve}(T)$ be a vertex. The vertex λ determines a subgraph T_λ of T , called the full subtree of T with root λ , such that: $\lambda \in \text{Ve}(T_\lambda)$, and for every vertex $\lambda' \in \text{Ve}(T_\lambda)$ all incoming edges (vertices included) of λ' in T belong also to T_λ .

2.17. Example.

The rooted tree



is a full subtree of



2.18. The empty tree. It is sometimes useful to call the empty graph \emptyset a rooted tree (and to define its height to be -1). It is then formally adjoined to the sets of trees described above, e.g., we consider the vector space with basis $\text{PTree} \cup \{\emptyset\}$ over the field K .

3. Operations of concatenation type

We recall the free objects of the categories of abelian semigroups, (not necessarily abelian) semigroups, and magmas. Words occur as elements of these free objects (cf. [19] as a general reference for operations and co-operations on words).

Elements of the free abelian semi-group $W_{\text{Com}}(X)$ over X are commutative words $x_{i_1}^{\nu_1} x_{i_2}^{\nu_2} \cdots x_{i_r}^{\nu_r}$, $i_1 < i_2 < \dots < i_r$, $r \geq 1$, $\nu_i \in \mathbb{N}$. Adjoining a unit 1 (empty word) we get the free abelian semi-group $W_{\text{Com}}^1(X)$ with unit.

The free semigroup $W_{\text{As}}(X)$ over X or the free monoid $W_{\text{As}}^1(X)$ over X is equipped with the concatenation $W_{\text{As}}^1(X) \times W_{\text{As}}^1(X) \rightarrow W_{\text{As}}^1(X)$, $(v, w) \mapsto v.w$ denoted by a lower dot. The elements of $W_{\text{As}}^1(X)$ are words $w = w_1.w_2 \dots w_r$, $w_i \in X$ for all i . Here $r \in \mathbb{N}$ is the length of w .

A magma is just a set M equipped with a binary operation (usually denoted by $\cdot : M \times M \rightarrow M$). The elements of the free magma $W_{\text{Mag}}(X)$ over X are parenthesized words. We are going to identify these words with planar binary rooted trees.

3.1. Forests. A commutative word (not necessarily non-empty) of abstract rooted trees, written as a disjoint union $T^1 \cup T^2 \cup \dots \cup T^r$ is called a (rooted) forest.

A word $T^1.T^2 \dots T^r$ of planar rooted trees is called a planar (rooted) forest.

We denote by PForest and AForest the corresponding sets of forests. The concepts and notations of labeled trees carry over to forests.

One may consider other, less canonical, combinations like abstract forests of planar rooted trees. If the context is clear, the words abstract and planar are often omitted.

3.2. Parenthesized strings. There exists a correspondence between planar rooted trees and (irreducible) parenthesized strings. We sketch this correspondence in the general setting of planar rooted trees with (M) -admissible labeling, where M_0 is given by $\{x_1, x_2, \dots\}$, and all other M_r , $r \geq 1$ are given by $\{y_1, y_2, \dots\}$. D. Kreimer's definition of irreducible parenthesized words in [11] is completely analogous. Here we make a difference between labels for the leaves ($x_i \in X$) and labels for the internal vertices ($y_i \in Y$), though.

Given a planar rooted tree T with (M) -admissible labeling, we recursively construct the corresponding parenthesized string.

If T consists of its root ρ_T , then ρ_T is a leaf labeled by some x_i . The corresponding parenthesized string is (x_i) , i.e. an opening bracket followed by the letter x_i followed by a closing bracket.

Else, let ρ_T be labeled by y_i , and let $T^1 \dots T^n$ be the forest of labeled rooted trees which remains after removing the root with its incoming edges. Assume that for each T_j , $1 \leq j \leq n$, we have already constructed the corresponding parenthesized string w_j . Then $(y_i w_1 \dots w_n)$ is the parenthesized string associated to T .

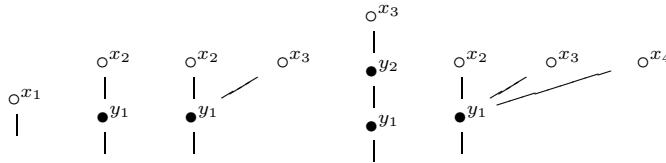
We get a string of letters and balanced brackets such that the leftmost opening bracket is matched by the rightmost closing bracket (irreducibility) and such that each letter has exactly one opening bracket on its lefthandside.

It is easy to see that the rooted tree can be reconstructed from its string.

Reducible words are defined by concatenation of irreducible ones, thus they correspond to forests.

For abstract rooted trees, there is a completely similar construction. The only difference is that some words have to be identified due to the missing order of incoming edges.

3.3. Example. The parenthesized strings (x_1) , $(y_1(x_2))$, $(y_1(x_2)(x_3))$, $(y_1(y_2(x_3)))$, and $(y_1(x_2)(x_3)(x_4))$ represent the rooted trees



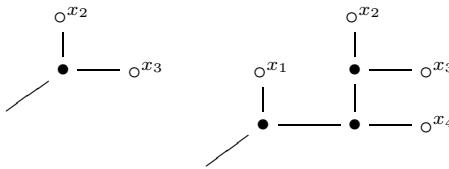
An empty pair of brackets without label is also allowed. It represents the empty tree.

3.4. Example. In the following we consider binary trees.

Let $M_0 = \{x_1, x_2, \dots\}$, and let M_2 be a one element set.

For every pair of brackets, the position of the closing bracket is forced once the position of the opening bracket is given. Thus one can omit the brackets and just use a letter c to mark and replace any combination of an opening bracket followed by the common label of an internal vertex.

For example the binary trees



can be represented by the strings cx_2x_3 , $cx_1c^2x_2x_3x_4$.

3.5. Malcev representation. Since the free magma generated by a set of variables X consists of parenthesized strings given by planar binary trees, we can call the set of planar binary trees with leaves labeled by X the free magma generated by X . The product \cdot in the free magma is thus given by a map $(T, T') \mapsto \vee(T, T')$ of planar binary trees (the grafting-operation defined in the following paragraph, here in the case $n = 2$).

If a field K is given, we can pass from the free magma generated by X to the free magma algebra (similarly to passing from semi-groups or groups to semi-group algebras or group-algebras).

The representation given in Example (3.4) is the Malcev representation of the free magma algebra over $X = \{x_1, x_2, \dots\}$ in the free associative algebra generated by $\{c, x_1, x_2, \dots\}$. The free magma multiplication \cdot corresponds to the operation $(v, w) \mapsto cvw$ in the free associative algebra.

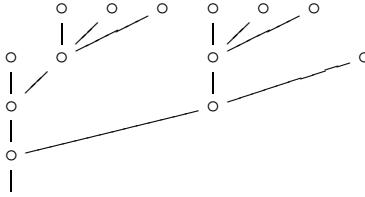
3.6. Grafting product. Given a forest $T^1 \dots T^n$ of $n \geq 0$ planar rooted trees with an (M) -admissible labeling, together with a label $\rho \in M$, there is a rooted tree $T = \vee_\rho(T^1 \cdot T^2 \dots T^n)$ defined by introducing a new n -ary root and grafting the trees T^1, \dots, T^n onto this new root. The new root gets the label ρ . The specified order in the forest determines the order of incoming edges at ρ_T . The rooted tree T is called the grafting product of $T^1 \dots T^n$ over ρ .

If there is no choice for a label ρ (i.e. there is only one label available) we simply write $\vee(T^1 \dots T^n)$.

Analogously defined is the grafting product for (forests of) abstract rooted trees. In the literature it is often denoted by $B_+(T^1 \dots T^n)$, e.g. in [12].

It is common to identify a unit for these operations with the empty tree \emptyset . To avoid difficulties with respect to uniqueness, we still do not allow the members T^i of a forest $T^1 \dots T^n$ to be empty. (We allow empty forests, though.)

3.7. Example. The grafting product $\vee(T^1, T^2)$ of T^1, T^2 from (2.4) is just



3.8. Degrafting map. Every planar rooted tree T is the grafting product of a forest $\neg T$ uniquely determined by T .

In the case of an n -ary root ρ , the forest $\neg T$ consists of the n full subtrees whose roots are the childs of the root ρ of T . Especially, all rooted trees in the forest $\neg T$ have heights less than the height of T .

There is a canonical de-grafting map \neg from labeled non-empty planar rooted trees to forests of labeled planar rooted trees (given by deleting the root together with its label). The analogously defined operator on forests of abstract rooted trees is often denoted by B_- .

3.9. Reversed words and mirrored trees. For words $w = w_1 \cdot w_2 \dots w_n$ the reversed word \bar{w} is defined by $w_n \cdot w_{n-1} \dots w_1$. Similarly, given a planar rooted tree T , there is a unique rooted tree \bar{T} recursively defined by

$$\overline{\vee(T^1 \dots T^n)} = \vee(\bar{T}^n \dots \bar{T}^1),$$

where $\overline{\vee(\emptyset)} = \vee(\emptyset)$ and $\bar{\emptyset} = \emptyset$. In other words, \bar{T} is obtained by mirroring T along the root axis.

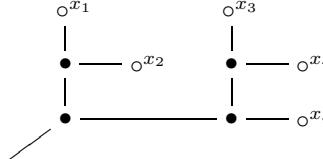
It holds that $\overline{(\bar{T})} = T$. The rooted trees T^1 and T^2 from (2.4) are in correspondence via $T \mapsto \bar{T}$.

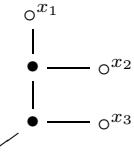
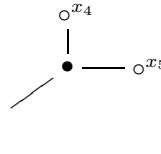
3.10. Splits and deconcatenation. Dually defined to the concatenation of words is the deconcatenation. Recall that the result of the deconcatenation, applied to a word w of length r , is a sum where each summand is given by a split of w into a left part $w_{(1)}$ and a right part $w_{(2)}$.

Dually defined to the grafting products on planar rooted trees are splitting co-operations. We give one example, which occurs in the case of planar binary trees as the dual co-operation of the binary grafting operation. It is studied in [10].

Let T be a planar binary tree with r leaves numbered from left to right by $1, 2, \dots, r$. Let i , $1 \leq i < r$, be an integer. We split the tree T into two trees $T_{(1)}^i$ and $T_{(2)}^i$ by cutting in between the leaves i and $i + 1$. More precisely, the tree $T_{(1)}^i$ is the reduction of the part of T which is on the left side of the path from leaf i to the root (including the path). The tree $T_{(2)}^i$ is the reduction of the analogous part on the right side of the path from leaf $i + 1$ to the root.

Look at $i = 3$ and $T =$



Then $T_{(1)}^3 =$  , $T_{(2)}^3 =$ 

The result of the co-operation applied to T is the sum, over all i , of the summands $T_{(1)}^i \otimes T_{(2)}^i$.

4. Operations related to the substitution

If one considers a letter which appears in a (commutative or non-commutative) word as placeholder for a word itself, and substitutes this letter by the corresponding word, one naturally obtains a new word. Similarly defined is the following substitution procedure for trees:

4.1. Substitution. Let T^1, T^2 be planar (or abstract) rooted trees and let b be a leaf of T^1 . Then the substitution of T^2 in T^1 at b , denoted by $T^1 \circ_b T^2$, is obtained by replacing the leaf b of T^1 by the root of T^2 .

4.2. Over and under. Given a planar binary tree T in PBTree, or a planar binary tree T in PBTree(M_0, M_2), let $\alpha = \alpha(T)$ denote the first leaf of T (i.e. the leftmost leaf in a drawing which puts all leaves on one line).

Let $\omega = \omega(T)$ denote the last (i.e. rightmost) leaf of T .

Given a second planar binary tree S , we define

$$T \setminus S := T \circ_{\omega(T)} S.$$

The operation \setminus is called under-operation and was introduced in [14]. Clearly \setminus is associative and $S \setminus T \setminus Z$ is well-defined.

The analogous operation \circ_α given by $T \circ_\alpha S = T \circ_{\alpha(T)} S$ plays the role of an associative multiplication in a Hopf algebra defined by C. Brouder and A. Frabetti (cf. [1]).

The opposite multiplication \circ_α^{op} , defined by $\circ_\alpha^{op}(S, T) := \circ_\alpha(T, S)$, is the over-operation $S/T := T \circ_{\alpha(T)} S$ of [14].

Using the mirror-operation $T \mapsto \bar{T}$, one can express S/T as $\overline{(T) \setminus (S)}$.

The tree $\begin{array}{c} \circ \\ \backslash \quad / \\ S \end{array}$ consisting of the root serves as a unit for all these operations, e.g.

$$S \backslash \begin{array}{c} \circ \\ \backslash \quad / \\ S \end{array} = S = \begin{array}{c} \circ \\ \backslash \quad / \\ S \end{array} \backslash S.$$

4.3. Right comb presentation. Given planar binary trees T^1, \dots, T^n , $n \geq 1$, and a sequence w of n labels from M_2 , we define

$$\vee \rightarrow_w (T^1 \dots T^n)$$

to be the planar binary tree which can be obtained from the right comb R_n of height n as follows: We replace the first leaf $\alpha(R_n)$ by T^1 , the second by T^2 and so on, leaving the $(n+1)$ -th leaf (i.e. $\omega(R_n)$) unaltered. We just write $\vee \rightarrow(T^1 \dots T^n)$ if there is no choice of labels.

We define $\vee \rightarrow(\emptyset) = \begin{array}{c} \circ \\ \backslash \quad / \\ \emptyset \end{array}$. In particular, R_n can then be written as

$$\vee \rightarrow \underbrace{\begin{array}{c} \circ \\ \backslash \quad / \\ \dots \end{array}}_n \circ).$$

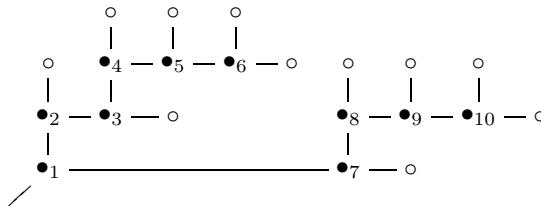
for every n .

It is easy to see that the smallest set which contains $\begin{array}{c} \circ \\ \backslash \quad / \\ \emptyset \end{array}$ and is closed under $\vee \rightarrow_w$ operations contains all planar binary trees.

For every planar binary tree with (M) -admissible labeling, this right comb presentation is unique. The left comb presentation is similarly defined. The right (or left) comb presentation induces a map φ_r (φ_l respectively) from planar binary trees to planar forests of planar rooted trees (not necessarily binary) such that

$$\varphi_r \left(\begin{array}{c} \circ \\ \backslash \quad / \\ \emptyset \end{array} \right) = \emptyset, \quad \varphi_r \left(\vee \rightarrow_{w_1 \dots w_n} (T^1 \dots T^n) \right) = \vee_{w_1} (\varphi_r(T^1)) \dots \vee_{w_n} (\varphi_r(T^n))$$

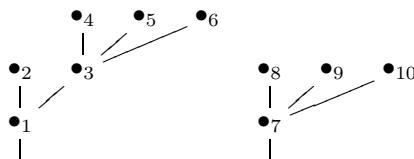
4.4. Example. The following planar binary tree
in $\text{PBTree}^{11}(M_2 = \{1, \dots, 10\})$



can be written in right comb presentation as

$$\vee \rightarrow_{1,7} \left(\vee \rightarrow_{2,3} \left(\begin{array}{c} \circ \\ \backslash \quad / \\ \emptyset \end{array}, \vee \rightarrow_{4,5,6} \left(\begin{array}{c} \circ \\ \backslash \quad / \\ \circ \end{array}, \begin{array}{c} \circ \\ \backslash \quad / \\ \circ \end{array} \right) \right) \right), \quad \vee \rightarrow_{8,9,10} \left(\begin{array}{c} \circ \\ \backslash \quad / \\ \circ \end{array}, \begin{array}{c} \circ \\ \backslash \quad / \\ \circ \end{array}, \begin{array}{c} \circ \\ \backslash \quad / \\ \circ \end{array} \right)$$

Its image under φ_r is this forest:



4.5. Corollary. Let $M = (M_0, \emptyset, M_2, \emptyset, \emptyset, \dots)$, where M_0 is a one element set and M_2 is arbitrary, and let \tilde{M} be the constant sequence (M_2, M_2, \dots) .

Then the map φ_r (or φ_l) provides a bijection from the set $\text{PBTree}^n(M)$ of planar binary trees with (M) -admissible labelings and n leaves onto the set of planar forests with (overall) $n - 1$ vertices, with (\tilde{M}) -admissible labelings.

The number of planar forests with $n - 1$ vertices and (\tilde{M}) -admissible labelings as well as the number of the corresponding planar binary trees is given by

$$c_n \cdot (\#M_2)^{n-1}$$

where c_n is the n -th Catalan number.

The bijection between non-labeled planar binary trees with n leaves and non-labeled planar trees with n vertices occurs as a special case (when we graft the corresponding forest onto a new root).

4.6. Cutting trees. Let T be a (planar or abstract) rooted tree with root ρ , and $C \subseteq \text{Ve}(T)$. We call C an admissible cut of T , if for every vertex $\lambda \in C$ all vertices of the full subtree with root λ are also in C . The case $C = \emptyset$ is called the empty cut. The case $C = \text{Ve}(T)$ is called the full cut.

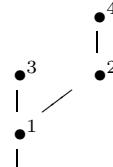
Given such an admissible cut, let $R^C(T)$ be the not necessarily non-empty tree (with root ρ , if $R^C(T) \neq \emptyset$), obtained by removing all vertices of C (together with their outgoing edges).

From T we can remove (the subgraph) $R^C(T)$ to get a (planar or abstract) forest $C(T)$ with set of vertices C .

The pair $(C(T), R^C(T))$ is called result of the cut C .

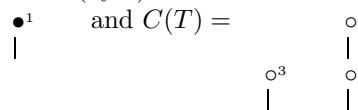
An admissible cut of T can also be defined as a non-empty subset of the set of (inner) edges of T such that for every vertex $v \in \text{Ve}(T)$ on the path to the root there is at most one edge selected, cf. [2]. This definition leads to the same pair $(C(T), R^C(T))$ and is in fact equivalent, once we add the full and empty cut.

4.7. Example. Let T be the following planar rooted tree:

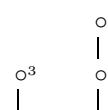


We are going to indicate (by \circ) which vertices are selected.

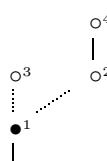
We get $R^C(T) =$



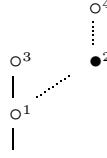
and $C(T) =$



as the result of the admissible cut



Not an admissible cut is:



4.8. Hopf algebra of Connes and Kreimer. In [11], D. Kreimer discovered a commutative Hopf algebra for the use of renormalization of quantum field theories. It was further studied by A. Connes and D. Kreimer, cf. [2], and [3].

The vector space, over a field K , of all abstract forests can be considered as a commutative polynomial algebra $K[\text{ATree}]$, graded with respect to the canonical degree function induced by

$$\deg T = n = \#\text{Ve}(T), \text{ for } T \in \text{ATree}_n,$$

and the unit 1 is identified with the empty tree \emptyset . The Connes-Kreimer Hopf algebra is obtained when $K[\text{ATree}]$ is provided with the coproduct Δ_{CK} , defined on the basis by

$$B_+(T^1 \dots T^n) \mapsto B_+(T^1 \dots T^n) \otimes \emptyset + (\cdot, B_+) (\Delta_{\text{CK}}(T^1) \otimes \dots \otimes \Delta_{\text{CK}}(T^n)).$$

The graded dual of this commutative Hopf algebra is isomorphic (via a graded isomorphism) to a noncommutative cocommutative Hopf algebra on trees introduced by R. Grossman and R. G. Larson (cf. [5]), see [17], [7].

There is an alternative description of the coproduct Δ_{CK} using the concept of admissible cuts (see 4.6). Since any non-full admissible cut of $T = \vee(T^1 \dots T^n)$ corresponds to n admissible cuts (of T^1, \dots, T^n), it is not hard to prove by induction that the image of T under Δ_{CK} is given by

$$\sum_{\text{C admissible cut}} C(T) \otimes R^C(T),$$

see [2].

The analogous construction for planar forests of planar trees (instead of abstract forests of abstract trees) yields a graded Hopf algebra structure which is isomorphic to the graded Hopf algebra of planar binary trees introduced by J.-L. Loday and M. Ronco in [14], see [8], [4].

4.9. Example. The coproduct Δ_{CK} maps

$$\overset{\circ}{\text{I}} = \vee(\emptyset) \mapsto \overset{\circ}{\text{I}} \otimes \emptyset + \emptyset \otimes \overset{\circ}{\text{I}}$$

For $f = 2$  we compute that $\Delta_{\text{CK}}(f)$ is given by

$$\begin{aligned} 2 \left(\overset{\circ}{\text{I}} \otimes \emptyset + \emptyset \otimes \overset{\circ}{\text{I}} + \overset{\circ}{\text{I}} \otimes \overset{\circ}{\text{I}} \right) - \overset{\circ}{\text{I}} \overset{\circ}{\text{I}} \otimes \emptyset - \emptyset \otimes \overset{\circ}{\text{I}} \overset{\circ}{\text{I}} - 2 \overset{\circ}{\text{I}} \otimes \overset{\circ}{\text{I}} \\ = f \otimes \emptyset + \emptyset \otimes f \end{aligned}$$

For $h = 2$  -  -  we compute that $\Delta_{\text{CK}}(h)$ is given by

$$\begin{aligned}
& 2 \left(\text{Diagram 1} \otimes \emptyset + \emptyset \otimes \text{Diagram 1} + \text{Diagram 2} \otimes \text{Diagram 2} \right) \\
& - \text{Diagram 3} \otimes \emptyset - \emptyset \otimes \text{Diagram 3} - \text{Diagram 4} \otimes \text{Diagram 4} \\
& - \text{Diagram 5} \otimes \emptyset - \emptyset \otimes \text{Diagram 5} - \left(\text{Diagram 6} + \text{Diagram 7} \right) \otimes \text{Diagram 8} - \text{Diagram 9} \otimes \left(\text{Diagram 10} + \text{Diagram 11} \right) \\
& = h \otimes \emptyset + \emptyset \otimes h + \text{Diagram 12} \otimes f - f \otimes \text{Diagram 13}
\end{aligned}$$

4.10. Free nonsymmetric operads. Let a collection $(M_r)_{r \geq 2}$ of sets be given, and set $M_0 := \{\circ\}$, $M_1 := \emptyset$.

Define, for $n \geq 2$, $\Gamma(M)(n)$ to be the vector space of all linear combinations of reduced planar trees with n leaves equipped with an (M) -admissible labeling. Let

$\Gamma(M)(0) = 0$, $\Gamma(M)(1) = K \cdot |$, where $| := \begin{array}{c} \bullet \\ \uparrow \end{array}$ is the tree consisting of the root.

The sequence $\overline{\Gamma}(M)$ of vector spaces, together with the unit $|$ and composition maps determined by the \circ_i -operations $T^1 \circ_i T^2$ (given by the substitution of T^2 in T^1 at the i -th leaf, see 4.1) is the free non- Σ operad generated by the collection $M = (M_r)_{r \geq 2}$.

Families of generalized bialgebra structures on algebras over families of free nonsymmetric operads are considered in [9].

4.11. Example. Let M_2 consist of one generator α , and let $M_r = \emptyset$ ($r \geq 3$).

Then all elements of $\underline{\Gamma}(M)$ are linear combinations of planar binary trees. For $n \geq 1$, we can identify a basis of $\underline{\Gamma}(M)(n)$ with the set of (non-labeled) planar binary trees with n leaves. Especially, $\dim \underline{\Gamma}(M)(n) = c_n$.

The tree  corresponds to the binary operation α , and we get ternary operations $\alpha \circ_1 \alpha$ and $\alpha \circ_2 \alpha$ as compositions:

This is the non- Σ operad Mag of (non-unitary) magma algebras.

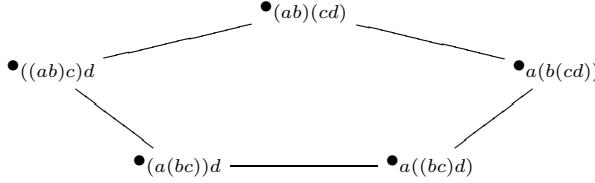
4.12. The associahedron. Let the collection $M = (M_r)_{r \geq 0}$ be given by $M_0 := \{\circ\}$, $M_1 := \emptyset$, and M_r , a one element set for each $r \geq 2$. Then the generated free nonsymmetric operad $\Gamma(M)$ is known as the operad of Stasheff polytopes.

The Stasheff polytope or associahedron K_{n+1} is a convex polytope of dimension $n-1$, $n \geq 1$, with one vertex for each planar binary tree with $n+1$ leaves. More

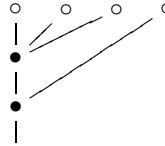
exactly, K_{n+1} is a cell complex in dimension $n-1$ with the elements of $\text{PBTree}^{n+1} = \text{PRTree}_{2n+1}^{n+1}$ as 0-cells.

The associahedra were created by J. Stasheff [23] to study higher homotopies for associativity. If we consider the parenthesized strings (of 3 letters) given by the two planar binary trees with 3 leaves, we can get from one to the other by shifting a bracket, in other words (cf. [16], I.1), applying an associating homotopy $h(x, y, z)$ from $x(yz)$ to $(xy)z$.

In K_4 for example, the five planar binary trees with 4 leaves have to be arranged in a pentagon



such that each side corresponds to an application of $h(x, y, z)$. These 5 sides can be labeled by the 5 reduced planar trees (with 4 leaves and 2 internal vertices) indicating the associating homotopy. For example, the edge between $(a(bc))d$ and $((ab)c)d$ corresponds to the tree



The $(n+1)$ -corolla represents the top dimensional cell of the polytope K_{n+1} .

By definition of the cell complex K_{n+1} the cells of dimension k are in bijection with the elements of $\text{PRTree}_{2n-k+1}^{n+1}$, for $k = 0, \dots, n-1$.

The polytope K_2 is a point which corresponds to the unique element  of $\text{PBTree}^2 = \text{PRTree}_3^2$, and K_3 is an interval.

The facets (i.e. codimension one cells) of K_{n+1} are of the form $K_{r+1} \times K_{s+1}$, $r, s \geq 1$, $r+s=n$, with label obtained by grafting the s -corolla to the i -th leaf of the r -corolla, $1 \leq i \leq r$. One gets inclusion maps $\circ_i : K_{r+1} \times K_{s+1} \rightarrow K_{r+s+1}$ (cf. [16], I.1.6).

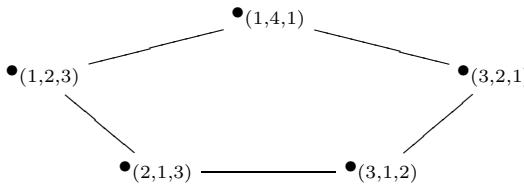
Pentagons and squares are the facets of K_5 .

In fact, the realization of K_{n+1} as a convex polytope was an open problem at first. Several solutions were given (cf. [24]). A simple realization, given in [13], associates to each planar binary tree $T \in \text{PBTree}^{n+1}$ a coordinate tuple $x(T)$ in (a hyperplane of) \mathbb{R}^n as follows:

For $1 \leq i < n+1 = \#\text{Le}(T)$, the i -th internal vertex of T is the highest internal vertex which belongs to both the paths from the i -th and the $(i+1)$ -th leaf to the root.

Consider the subtree with root given by the i -th internal vertex of T , and let a_i be the number of leaves on the left side, b_i the number of leaves on the right side (of the subtree's root). Then the i -th entry of $x(T)$ is given by $a_i b_i$.

For example, the coordinate tuples we get for K_4 are:



It is shown in [13], Theorem 1.1, that the convex hull of the points $x(T)$, $T \in \text{PBTree}^{n+1}$, is a realization of the Stasheff polytope of dimension $n - 1$.

It is possible to give an orientation to all the edges of the Stasheff polytope, see [15]. When 0-cells are represented as parenthesized words, arrows are directed such that they correspond to shifting a bracket from left $((xx)x)$ to right $(x(xx))$. The induced partial ordering on the set PBTree is called the Tamari order, cf. [22].

References

- [1] C. Brouder and A. Frabetti, QED Hopf algebras on planar binary trees, *J. Algebra* 267 (2003), no. 1, 298-322.
- [2] A. Connes and D. Kreimer, Hopf Algebras, Renormalization and Noncommutative Geometry, *Comm. Math. Phys.* 199 (1998), 203-242.
- [3] A. Connes and D. Kreimer, Renormalization in QFT and the Riemann-Hilbert Problem I, *Comm. Math. Phys.* 210 (2000), 249-273.
- [4] L. Foissy, Les algèbres de Hopf des arbres enracinés décorés II, *Bull. Sci. Math.* 126(2002), 249-288.
- [5] R. Grossman, R. G. Larson, Hopf algebraic structures of families of trees, *J. Algebra* 126 (1989), 184-210.
- [6] F. Harary, "Graph theory", Addison-Wesley, Reading, Mass., 1969.
- [7] M. Hoffman, Combinatorics of rooted trees and Hopf algebras, *Trans. AMS* 355 (2003), 3795-3811.
- [8] R. Holtkamp, Comparison of Hopf algebras on Trees, *Arch. Math.* 80 (2003), 368-383.
- [9] R. Holtkamp, On Hopf algebra structures over free operads, *Adv. Math.* 207 (2006), 544-565.
- [10] R. Holtkamp, J.-L. Loday, M. Ronco, Coassociative magmatic bialgebras and the Fine numbers, *J. Algebraic Combin.* 28 (2008), 97-114.
- [11] D. Kreimer, On the Hopf Algebra structure of perturbative quantum field theories, *Adv. Theor. Math. Phys.* 2 (1998), 303-334.
- [12] D. Kreimer, On overlapping divergences, *Comm. Math. Phys.* 204 (1999), 669-689.
- [13] J.-L. Loday, Realization of the Stasheff polytope, *Arch. Math.* 83 (2004), 267-278.
- [14] J.-L. Loday and M. Ronco, Hopf Algebra of the Planar Binary Trees, *Adv. Math.* 139 (1998), 299-309.
- [15] J.-L. Loday and M. Ronco, Order structure on the algebra of permutations and of planar binary trees, *J. Alg. Comb.* 15(3), (2002), 253-270.
- [16] M. Markl, S. Shnider, and J. Stasheff, "Operads in Algebra, Topology and Physics", vol. 96 of Math. Surveys and Monographs. AMS, Providence, Rhode Island, 2002.
- [17] F. Panaite, Relating the Connes-Kreimer and Grossman-Larson Hopf algebras built on rooted trees, *Lett. Math. Phys.* 51 (2000), 211-219.
- [18] V.M. Petrogradsky, Enumeration of algebras close to absolutely free algebras and binary trees, *J. Algebra* 290 (2005), 337-371.
- [19] C. Reutenauer, "Free Lie Algebras", London Math. Soc. Monographs, Oxford University Press, New York, 1993.
- [20] N. Sloane (Edt.), The On-Line Encyclopedia of Integer Sequences, 2008, <http://www.research.att.com/~njas/sequences/index.html>
- [21] R.P. Stanley, "Enumerative Combinatorics", Vol. 1, Cambridge University Press, Cambridge, 1997.
- [22] R.P. Stanley, "Enumerative Combinatorics", Vol. 2, Cambridge University Press, Cambridge, 1999.

- [23] J.D. Stasheff, Homotopy associativity of H -spaces I,II, *Trans. Amer. Math. Soc.* 108 (1963), 275-292, *ibid.* 293-312.
- [24] J.D. Stasheff, From operads to physically inspired theories, *in* Loday et.al.(ed): Operads, proceeding of renaissance conferences, AMS, 1997.

UNIVERSITY OF BOCHUM, 44780 BOCHUM, GERMANY
E-mail address: ralf.holtkamp@ruhr-uni-bochum.de

This page intentionally left blank

Magnus expansions and beyond

Arieh Iserles

ABSTRACT. In this brief review we describe the coming of age of Magnus expansions as an asymptotic and numerical tool in the investigation of linear differential equations in a Lie-group and homogeneous-space setting. Special attention is afforded to the many connections between modern theory of *geometric numerical integration* and other parts of mathematics: from abstract algebra to differential geometry and combinatorics, all the way to classical numerical analysis.

1. Lie-group equations

Numerical solution of evolutionary differential equations is as old as the theory of differential equations itself: although proper numerical analysis of differential equations commenced with Leonhard Euler, earlier *ad hoc* numerical ideas abound in the works of Sir Isaac Newton and of Gottfried von Leibnitz. (A brief, yet outstanding historical synopsis can be found in (Hairer, Nørsett & Wanner 1986).) In the last fifty years numerical analysis of differential equations has developed in leaps and bounds, in parallel with the evolution in computing power and speed.

On the face of it, all is well in the numerical kingdom. However, a closer look reveals a worrying gap between the efforts of numerical analysts and of pure mathematicians. Thus, pure mathematicians expand a very great deal of effort to analyse qualitative properties of differential equations but they usually fall short of fleshing out numbers. At the same time, numerical analysts are extraordinarily successful in producing numbers and figures with appropriately small errors but these numbers and figures typically fail to respect qualitative properties of differential equations. This disparity between analysis and computation motivated in the last decade the emergence of a new paradigm of *geometric numerical integration (GNI)*: to seek computational methods that render exactly important qualitative features of differential equations. Examples of qualitative features whose preservation under discretization is important include the symplectic structure of Hamiltonian and Lie–Poisson systems (Hairer, Lubich & Wanner 2003, Hairer, Lubich & Wanner 2006, Leimkuhler & Reich 2004, Marsden & West 2001), volume conservation of divergence-free differential systems (McLachlan & Quispel 2002), Lie symmetries (Hubert 2000) and, closer to the subject matter of this review, evolution on a Lie group or a homogeneous space (Iserles, Munthe-Kaas, Nørsett & Zanna 2000).

It is vital to emphasize that this respect for qualitative and geometric features under discretization has a threefold purpose. Firstly, qualitative attributes are central to the mathematical description of differential equations. Secondly, they often have profound physical significance – after all, conservation of volume, angular momentum or particle labelling is critical once we model scientific phenomena with differential equations. Finally, both numerical experience and a wealth of mathematical results indicate that conservation of geometric features leads to numerical algorithms that produce significantly smaller error in long-term integration (Hairer et al. 2006).

The concern of this review is in differential equations evolving in homogeneous spaces. We recall that \mathcal{M} is a homogeneous space, acted upon by the Lie group \mathcal{G} , if there exists a map $\lambda : \mathcal{G} \times \mathcal{M} \rightarrow \mathcal{M}$ such that

$$\begin{aligned} \lambda(g_1, \lambda(g_2, x)) &= \lambda(g_1 \cdot g_2, x) \quad \forall g_1, g_2 \in \mathcal{G}, x \in \mathcal{M}, \\ \forall x_1, x_2 \in \mathcal{M} \quad \exists g \in \mathcal{G} : \quad \lambda(g, x_1) &= x_2. \end{aligned}$$

A differential equation evolving in \mathcal{M} can be always written in the form

$$(1.1) \quad y' = \lambda_*(A(t, y))y, \quad t \geq t_0, \quad y(t_0) \in \mathcal{M},$$

where $\lambda_* : \mathfrak{g} \rightarrow \mathcal{X}(\mathcal{M})$ and $A : [t_0, \infty) \times \mathcal{M} \rightarrow \mathfrak{g}$. Here \mathfrak{g} is the Lie algebra corresponding to the Lie group \mathcal{G} , while $\mathcal{X}(\mathcal{M})$ is the linear space (actually, a Lie algebra equipped with a suitable bracket) of vector fields on \mathcal{M} (Iserles et al. 2000). For (1.1) to make sense as a differential equation, we require that A is Lipschitz in a suitable norm, but in practice it is usual to impose greater smoothness. Not to be let astray by considerations marginal to our main narrative, we assume in the sequel that A is a C^∞ function.

Important examples of (1.1) include

- *Lie-group equations*

$$(1.2) \quad y' = A(t, y)y, \quad t \geq t_0, \quad A : [t_0, \infty) \times \mathcal{G} \rightarrow \mathfrak{g},$$

when \mathcal{G} acts on itself, $\mathcal{M} = \mathcal{G}$.

- *Isospectral flows*

$$y' = [B(t, y), y], \quad t \geq t_0, \quad y(0) \in \text{Sym}(n), \quad B : [t_0, \infty) \times \text{Sym}(n) \rightarrow \mathfrak{so}(n),$$

where $\text{Sym}(n)$ and $\mathfrak{so}(n)$ are the $n \times n$ real symmetric and skew-symmetric matrices, respectively. In that case it is known that the invariants of the system are the eigenvalues of $y(0)$ and that the underlying group action is of $\text{SO}(n)$, acting by similarity.

- Ordinary differential equations evolving on a sphere,

$$y' = A(y(t)) \times y(t), \quad t \geq t_0, \quad A : \mathbb{R}^3 \rightarrow \mathbb{R}^3.$$

- Differential flows on *Stiefel* and *Grassmann manifolds*, *Frenet–Serret equations* and many other differential systems of interest.

Typically, the manifold \mathcal{M} , once embedded in a Euclidean space, is a nonlinear structure. Classical discretization methods are notoriously bad in respecting nonlinear invariants. The main paradigm of *Lie-group methods* is to rephrase the underlying equation (1.1) as an *algebra action* by \mathfrak{g} . Since \mathfrak{g} is a linear space, we can expect all reasonable discretization methods to respect its structure.

Except for minor implementational details, there is no loss of generality in considering the Lie-group equation (1.2) in place of the homogeneous-space equation (1.1): once we can

discretize in \mathcal{G} , we can extend the procedure to all manifolds acted upon by the group through algebra actions (Iserles et al. 2000).

We say that the smooth map $\phi : \mathcal{G} \rightarrow \mathfrak{g}$ is a *trivialisation* if it is one-to-one in a neighbourhood of the identity and $\phi(I) = 0$. (Here I and 0 are the identity of \mathcal{G} and the zero of \mathfrak{g} , respectively.)

Similarly to classical discretization methods for ordinary differential equations, Lie-group methods advance in a step-by-step fashion. Thus, we aim to compute $y_N \approx y(t_N)$, where $t_N = t_{N-1} + h_{N-1}$, $h_{N-1} > 0$, $N \in \mathbb{N}$. To lift the equation from the group to its algebra, we let

$$\tilde{y}(t) = \phi(\Omega(t))y_N, \quad t \in [t_N, t_{N+1}],$$

where Ω is our new unknown, which evolves in the Lie algebra \mathfrak{g} . Substituting into (1.2), we obtain the *trivialised equation*

$$(1.3) \quad \Omega' = d\phi_\Omega^{-1} A(t, \phi(\Omega)y_N), \quad t \geq t_N, \quad \Omega(t_N) = 0.$$

The mapping $d\phi_\Omega^{-1}$ need be computed individually for each trivialisation. The most natural important and ubiquitous is the exponential trivialisation $\phi(\Omega) = e^\Omega$, where the exponential map from the algebra to the group is defined in the usual manner, as the flow of the corresponding vector field. Note that once \mathfrak{g} is represented by matrices and $[\cdot, \cdot]$ is the standard matrix commutator, the exponential map becomes the standard matrix exponential $e^\Omega = \sum_{m=0}^{\infty} \frac{1}{m!} \Omega^m$. In that case we obtain the *dexpinv equation*

$$(1.4) \quad \Omega' = \text{dexp}_\Omega^{-1} A = \sum_{m=0}^{\infty} \frac{B_m}{m!} \text{ad}_\Omega^m A,$$

where $\{B_m\}_{m \geq 0}$ are Bernoulli numbers and ad_Ω is the adjoint operator (Hausdorff 1906).

Another useful trivialisation applies only in the important case when \mathcal{G} is a *quadratic* matrix Lie group. Assuming that \mathcal{G} is represented by matrices, this means that there exists $p \in \text{GL}(n)$ such that

$$\mathcal{G} = \{x \in \text{GL}(n) : xpx^\top = p\}.$$

Examples include the orthogonal group $O(n)$, the symplectic group $Sp(n)$ and the Lorentz group $O_{3,1}$. The corresponding Lie algebra is

$$\mathfrak{g} = \{b \in \mathfrak{gl}(n) : bp + pb^\top = 0\}.$$

In the instance of quadratic Lie groups we might use the *Cayley trivialisation*

$$\phi(x) = (I - \frac{1}{2}x)^{-1}(I + \frac{1}{2}x), \quad x \in \mathfrak{g}$$

and the outcome is the *dcayinv equation*

$$(1.5) \quad \Omega' = \text{dcay}_\Omega^{-1} A = A - \frac{1}{2}[\Omega, A] - \frac{1}{4}\Omega A \Omega$$

– it is easy to prove that in a quadratic Lie algebra $b, c \in \mathfrak{g}$ implies $bcb \in \mathfrak{g}$ (Lewis & Simo 1994).

2. Lie-algebraic expansions

2.1. The Magnus expansion. We consider the Lie-group equation (1.2), except that, for the time being, we stipulate that $A = A(t)$. The outcome is the linear equation

$$(2.1) \quad y' = A(t)y, \quad t \geq t_0, \quad y(t_0) = y_0 \in \mathcal{G}.$$

Our point of departure is a paper by Magnus (1954), which presented an intriguing expansion of the solution of (2.1) subject to exponential trivialisation. Thus, the function Ω in (1.4) can be expanded in the form

(2.2)

$$\begin{aligned}
 \Omega(t) = & \int_{t_N}^t A(\xi) d\xi - \frac{1}{2} \int_{t_N}^t \int_{t_N}^{\xi_1} [A(\xi_2), A(\xi_1)] d\xi_2 d\xi_1 \\
 & + \frac{1}{12} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_1} [A(\xi_3), [A(\xi_2), A(\xi_1)]] d\xi_3 d\xi_2 d\xi_1 \\
 & + \frac{1}{4} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} [[A(\xi_3), A(\xi_2)], A(\xi_1)] d\xi_3 d\xi_2 d\xi_1 \\
 & - \frac{1}{8} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_3} [[[A(\xi_4), A(\xi_3)], A(\xi_2)], A(\xi_1)] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\
 & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} [A(\xi_4), [[A(\xi_3), A(\xi_2)], A(\xi_1)]] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\
 & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_2} [[A(\xi_4), [A(\xi_3), A(\xi_2)]], A(\xi_1)] d\xi_4 d\xi_3 d\xi_2 d\xi_1 \\
 & - \frac{1}{24} \int_{t_N}^t \int_{t_N}^{\xi_1} \int_{t_N}^{\xi_2} \int_{t_N}^{\xi_1} [[A(\xi_4), A(\xi_2)], [A(\xi_3), A(\xi_1)]] d\xi_4 d\xi_3 d\xi_2 d\xi_1 + \dots
 \end{aligned}$$

The above expansion has a number of interesting features. Firstly, it is guaranteed to remain within a Lie algebra, since it is a linear combination of terms which, as one can immediately verify, live in \mathfrak{g} . Secondly, if \mathfrak{g} is Abelian then all the terms except for the leading integral are zero. Thirdly, the complexity of terms grows rapidly (indeed, exponentially) and there is little point in continuing this expansion (or indeed proving its convergence) unless we can identify a simple recursive or combinatorial rule to generate all terms in a transparent manner. Fourthly, the use of a truncated *Magnus expansion* (2.2) as a numerical tool is fraught with potential difficulties because multivariate quadrature is notoriously expensive (Davis & Rabinowitz 1984).

Wilhelm Magnus himself neither derived a general formula for (2.2) nor proved conditions for convergence. This did not prevent the Magnus expansion from being used in literally hundreds of papers, mostly in theoretical physics, quantum chemistry and stochastic analysis, as a perturbative tool in the solution of linear systems of the form (2.1).

Combinatorial formulæ for general terms in the Magnus expansion have been presented by several authors (Bialynicki-Birula, Mielenik & Plebański 1969, Fomenko & Chakon 1990, Mielenik & Plebański 1970, Strichartz 1987) but they are probably too complicated and unwieldy for practical use at high orders. A practical recursive algorithm, allowing for easy and transparent generation of Magnus expansion terms, has been presented by Iserles & Nørsett (1999) and it is based on a homomorphism between a subset of planar rooted trees and Magnus terms. It rests upon the following two composition rules:

- (1) The term $A(t)$ is associated with the trivial tree \bullet ;

(2) If $C_{\tau_1}(t)$ and $C_{\tau_2}(t)$ are associated with the rooted trees τ_1 and τ_2 respectively then

$$\left[\int_{t_N}^t C_{\tau_1}(\xi) d\xi, C_{\tau_2}(t) \right] \quad \text{is associated with} \quad \begin{array}{c} \tau_1 \\ \backslash \\ \bullet \\ \backslash \quad \backslash \\ \tau_2 \end{array}.$$

Let \mathcal{T} denote the set of all trees that can be generated by this procedure. Moreover, we denote by $\mathcal{T}_m \subset \mathcal{T}$, $m \in \mathbb{Z}_+$, the set of all such trees of height m (i.e., with m ‘vertical’ edges), hence $\mathcal{T} = \bigcup_{m \geq 0} \mathcal{T}_m$.

It is easy to confirm that each $\tau \in \mathcal{T}_m$, $m \in \mathbb{N}$, can be expressed in the form

$$(2.3) \quad \tau = \begin{array}{c} \tau_s \\ \backslash \\ \tau_3 \\ \backslash \\ \tau_2 \\ \backslash \\ \tau_1 \\ \backslash \\ \bullet \end{array},$$

where $\tau_k \in \mathcal{T}_{\ell_k}$ and $\sum_{k=1}^s \ell_k + s = m$. We define a function $\alpha : \mathcal{T} \rightarrow \mathbb{Q}$ inductively as follows. For the sole element in \mathcal{T}_0 we let $\alpha(\bullet) = 1$, otherwise we use (2.3) and set

$$\alpha(\tau) = \frac{B_s}{s!} \prod_{j=1}^s \alpha(\tau_j),$$

where B_s is the s th Bernoulli number.

THEOREM 1. *The expansion*

$$(2.4) \quad \Omega(t) = \sum_{m=0}^{\infty} \sum_{\tau \in \mathcal{T}_m} \alpha(\tau) \int_{t_N}^t C_{\tau}(\xi) d\xi$$

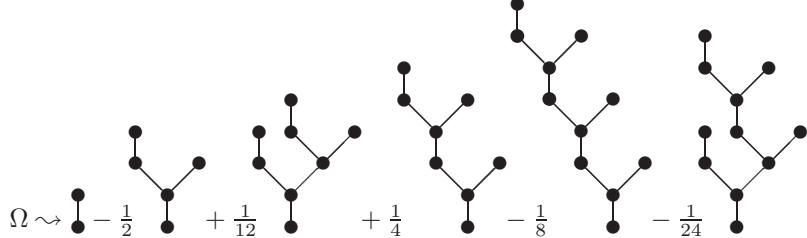
converges to the solution of the dexpinv equation (1.4) (Iserles & Nørsett 1999). Given a norm $\|\cdot\|$, the optimal convergence condition is

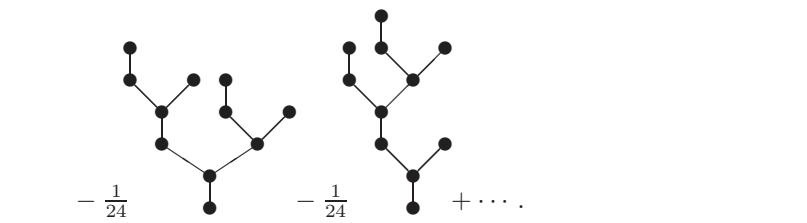
$$(2.5) \quad \int_{t_N}^t \|A(\xi)\| d\xi \leq \pi$$

(cf. (Moan & Niesen 2008) for the real case, (Casas 2007) for the complex one and for infinite-dimensional Hilbert space setting).

Appending a root to trees in \mathcal{T} , a procedure which, consistently with our composition rules, stands for integration, we can render the Magnus expansion (2.2) concisely and transparently in terms of rooted trees,

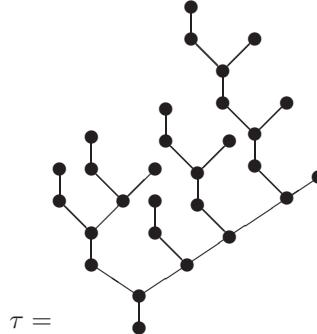
$$\Omega \sim \Omega_0 - \frac{1}{2} \Omega_1 + \frac{1}{12} \Omega_2 + \frac{1}{4} \Omega_3 - \frac{1}{8} \Omega_4 - \frac{1}{24} \Omega_5 - \frac{1}{24} \Omega_6 + \dots$$





With minimal training, it becomes surprisingly easy to ‘read’ the structure of an expansion term, expressed in the terminology of commutators and integrals, directly from the underlying tree.

As an example of synthesising a complex tree from simpler constituents, consider the following member of \mathcal{T}_9 ,

$$\tau =$$


(There are ten ‘vertical’ edges but we do not count the one emanating from the root.) Moreover, $s = 4$ and we identify the four “sub-trees” as

$$\tau_1 =$$


$$, \quad \tau_2 =$$


$$, \quad \tau_3 =$$


$$, \quad \tau_4 =$$


with

$$\alpha(\tau_1) = \frac{1}{12}, \quad \alpha(\tau_2) = 1, \quad \alpha(\tau_3) = -\frac{1}{2}, \quad \alpha(\tau_4) = \frac{1}{4}.$$

Since $B_4/4! = -\frac{1}{720}$, we deduce that

$$\alpha(\tau) = -\frac{1}{720} \times \frac{1}{12} \times 1 \times \left(-\frac{1}{2}\right) \times \frac{1}{4} = \frac{1}{69120}.$$

2.2. Truncating Magnus expansions. Once we wish to use (2.4) as either numerical or perturbative tool, we need to truncate the expansion. The most obvious truncation is

$$\Omega_r(t) = \sum_{m=0}^{r-1} \sum_{\tau \in \mathcal{T}_m} \alpha(\tau) \int_{t_N}^t C_\tau(\xi) d\xi, \quad t \in [t_N, t_{N+1}].$$

Since $\tau \in \mathcal{T}_m$ implies that $C_\tau(t) = \mathcal{O}(h_N^m)$ (recall that $t_{N+1} = t_N + h_N$), we deduce that $\Omega_r(t) = \Omega(t) + \mathcal{O}(h_N^{r+1})$. This, however, is suboptimal (Iserles, Nørsett & Rasmussen 2001).

The underlying idea in the above truncation was to allocate to each C_τ , $\tau \in \mathcal{T}_m$, the grade $\beta(\tau) = m$ and allow the grades to propagate according to the rule

$$\tau = \begin{array}{c} \tau_1 \\ | \\ \bullet \\ \swarrow \quad \searrow \\ \tau_2 \end{array} \Rightarrow \beta(\tau) = \beta(\tau_1) + \beta(\tau_2) + 1.$$

In that case $C_\tau(t) = \mathcal{O}(h_N^p)$ where $p \geq \beta(\tau)$. However, if last inequality is sharp then we might be including superfluous terms in the truncation.

The remedy is to allocate grades differently. Thus, we say that $\tau \in \mathcal{T}$ is of grade $\gamma(\tau)$ if $C_\tau(t) = \mathcal{O}(h_N^p)$ for $p \geq \gamma(\tau)$ and, for at least one $A : C^\infty([t_N, t_{N+1}] \rightarrow \mathfrak{g})$, it is true that $C_\tau(t) = \mathcal{O}(h_N^{\gamma(\tau)})$.¹ We let \mathcal{F}_m stand for the set of all $\tau \in \mathcal{T}$ such that $\gamma(\tau) = m$ and truncate

$$\Omega_{[r]}(t) = \sum_{m=0}^{r-1} \sum_{\tau \in \mathcal{F}_m} \alpha(\tau) \int_{t_N}^t C_\tau(\xi) d\xi, \quad t \in [t_N, t_{N+1}].$$

The tail is again $\mathcal{O}(h_N^{r+1})$ but, by construction, $\Omega_{[r]}$ contains the least possible number of terms! Specifically, the γ function propagates according to the recursion

$$\tau = \begin{array}{c} \tau_1 \\ | \\ \bullet \\ \swarrow \quad \searrow \\ \tau_2 \end{array} \Rightarrow \gamma(\tau) = \begin{cases} \gamma(\tau_1) + \gamma(\tau_2) + 1, & \tau_1 \neq \tau_2, \\ \gamma(\tau_1) + \gamma(\tau_2) + 2, & \tau_1 = \tau_2. \end{cases}$$

This minor difference with $\beta(\tau)$ leads to significant saving in the number of terms for large r . Asymptotically,

$$\limsup_{m \rightarrow \infty} (\#\mathcal{T}_m)^{1/m} = 4, \quad \limsup_{m \rightarrow \infty} (\#\mathcal{F}_m)^{1/m} \approx 3.1167417747$$

(Iserles et al. 2001).

2.3. Alternative Lie-algebraic expansions. Magnus expansions represent just one, although arguably the most important, instance of a Lie-algebraic expansion. One alternative are the *Fer expansions*

$$y(t_{N+1}) = \exp\left(\int_{t_N}^{t_{N+1}} B_0(\xi) d\xi\right) \exp\left(\int_{t_N}^{t_{N+1}} B_1(\xi) d\xi\right) \exp\left(\int_{t_N}^{t_{N+1}} B_2(\xi) d\xi\right) \cdots y(t_N)$$

where $B_0 = A$, while each B_m for $m \in \mathbb{N}$ can be computed from B_{m-1} (Fer 1958). Although it can also be rendered conveniently in the terminology of rooted trees (Iserles et al. 2000),

¹It is useful to regard A as a matrix – because of the Ado theorem every finite-dimensional Lie algebra admits a faithful representation, hence this represents no loss of generality (Varadarajan 1984).

it is possible to prove that its numerical implementation is consistently more expensive than that of Magnus expansions (Celledoni, Iserles, Nørsett & Orel 2002).

A version of Magnus expansions, using rooted trees with coloured leaves, has been developed for equations of the form $y' = A(t)y - yB(t)$, where both A and B evolve in \mathfrak{g} (Iserles 2001a). Equations of this kind feature in the computation of Floquet exponents. Another Magnus expansion was presented for the computation of the (nonlinear) $\mathfrak{so}(n)$ action in the computation of the isospectral *double-bracket equations* $y' = [y, [M, y]]$, $y(0) \in \text{Sym}(n)$, where $M \in \mathfrak{so}(n)$ (Iserles 2002).

It is possible to extend Magnus expansions to a general nonlinear setting $y' = A(t, y)y$, $y(t_0) \in \mathcal{G}$, where $A : [t_0, 0) \times \mathcal{G} \rightarrow \mathfrak{g}$ (Zanna 1999, Casas & Iserles 2006), but this is arguably less efficient and more complicated than the use of *Runge–Kutta–Munthe–Kaas methods* (Munthe-Kaas 1998).

An important alternative to the expansions of the kind reviewed in this paper is provided by expansions based on *canonical coordinates of the second kind*

$$y(t) = e^{\sigma_1(t)C_1}e^{\sigma_2(t)C_2}\dots e^{\sigma_q(t)C_q}y(t_N),$$

where $q \geq \dim \mathfrak{g}$, $\{C_1, \dots, C_q\}$ is a frame of \mathfrak{g} and $\sigma_1, \dots, \sigma_q$ are scalar functions (Crouch & Grossman 1993, Owren & Marthinsen 2001).

All the above are based on the exponential trivialisation. Once, however, we expand Lie-group and homogeneous-space equations acted by quadratic Lie-groups, there are clear advantages in cost and complexity in using the Cayley trivialisation and the *dcayinv* equation (1.5). The entire theory of this section can be extended to this setting once we associate expansion terms with *bicolour rooted trees*. Specifically, we have the expansion

$$(2.6) \quad \Omega(t) = \sum_{m=0}^{\infty} \frac{(-1)^m}{2^m} \sum_{\tau \in \mathcal{S}_m} (-1)^{\delta(\tau)} D_{\tau}(t), \quad t \geq t_N,$$

where expansion terms are assembled according to the following composition rules,

- (1) We commence from $\int_{t_N}^t A(\xi) d\xi$ and associate to it the tree 
- (2) Having already associated $D_{\tau}(t)$ with the tree τ , we associate

$$\int_{t_N}^t [D_{\tau}(\xi), A(\xi)] d\xi \quad \text{with} \quad \begin{array}{c} \tau \\ \diagdown \quad \diagup \\ \bullet \quad \bullet \end{array}$$

- (3) Having already associated $D_{\tau_1}(t)$ and $D_{\tau_2}(t)$ with the trees τ_1 and τ_2 , respectively, we associate

$$\int_{t_N}^t D_{\tau_1}(\xi) A(\xi) D_{\tau_2}(\xi) d\xi \quad \text{with} \quad \begin{array}{c} \tau_1 \quad \tau_2 \\ \diagdown \quad \diagup \\ \circ \quad \bullet \end{array}.$$

\mathcal{S}_m stands for the set of all such trees with $m + 1$ ‘vertical’ edges (corresponding to integration), while $\delta(\tau)$ stands for the number of white nodes in the rooted tree τ (Iserles 2001b).

The *Cayley expansion* is

$$\begin{aligned}
 \Omega \sim & \left(\text{tree}_1 - \frac{1}{2} \text{tree}_2 + \frac{1}{4} \text{tree}_3 - \frac{1}{4} \text{tree}_4 - \frac{1}{8} \text{tree}_5 + \frac{1}{8} \text{tree}_6 \right. \\
 & \left. + \frac{1}{8} \text{tree}_7 + \frac{1}{8} \text{tree}_8 + \dots \right)
 \end{aligned}$$

It is easy to verify that the term corresponding to the sixth tree belongs to the quadratic Lie algebra, but this is not the case with regard to either the seventh or the eighth tree. Fortunately, the sum of these two terms lies in the algebra. We will return to this issue in Section 3.

2.4. Rooted trees and Hopf algebras. Rooted trees and forests have played an important role in numerical mathematics as a convenient organising principle to elucidate combinatorial structure of algorithms. An incomplete list includes

- Runge–Kutta methods and their expansion in *B-series* (Butcher 1963, Butcher 1972, Hairer & Wanner 1974);
- Symplectic and partitioned Runge–Kutta methods (Abia & Sanz-Serna 1993, Hairer et al. 2006);
- Splitting and composition of vector fields (Murua & Sanz-Serna 1999, McLachlan & Quispel 2002);
- Expansions of iterated integrals (Chen 1977);
- Nonlinear functional equations with proportional delay (Iserles 1994);
- Volume preservation under discretization (Chartier & Murua 2007, Iserles, Quispel & Tse 2007).

Some time ago Grossman & Larson (1989) identified the *Butcher group*, the algebraic structure underlying B-series, with a *Hopf algebra*. The fundamental importance of algebras of rooted trees and their connections with the representation of the renormalisation group of quantum field theory has been recognised in (Connes & Kreimer 1998, Connes & Kreimer 2000), and this has led to a great deal of interest and further research, not least on geometric integrators: cf. for example (Munthe-Kaas & Wright 2008, Murua 2006).

3. Quadrature and graded algebras

3.1. Magnus quadrature. Consider again the expansion (2.2), this time from a practical standpoint. Given a matrix function A and letting $t = t_{N+1}$, we wish to compute integrals therein by quadrature. Each term is an integral of a nested commutator over a multivariate polytope and it is well known that multivariate quadrature is excessively expensive in terms both of function evaluations and algebraic operations (Davis & Rabinowitz 1984). Indeed,

viewed from the standpoint of conventional quadrature, the computation of truncated expansion $\Omega_{[r]}(t_{N+1})$ for $r \geq 2$ is impractical. However, the special structure of Magnus expansion terms lends itself to quadrature formulae that are surprisingly economical in terms of function evaluations.

For simplicity, we assume in this section that $t_N = 0$ and let $h = h_N$.

THEOREM 2. (Iserles & Nørsett 1999) *Let c_1, c_2, \dots, c_ν be distinct points in $[0, 1]$ and set $A_k = hA(c_k h)$, $k = 1, \dots, \nu$. Given a term*

$$(3.1) \quad I[\mathcal{P}] = \int_{h\mathcal{P}} \mathbf{L}(A(\xi_1), A(\xi_2), \dots, A(\xi_p)) d\xi_p \cdots d\xi_2 d\xi_1$$

that features in the Magnus expansion (2.4) (here $h\mathcal{P}$ is a polytope, while \mathbf{L} is a multilinear form, consisting of nested commutators), we approximate it by the quadrature formula

$$(3.2) \quad Q[\mathcal{P}] = \sum_{\ell \in C_p^\nu} b_\ell \mathbf{L}(A_{l_1}, A_{l_2}, \dots, A_{l_p}),$$

where C_p^ν is the set of all words of length p from the alphabet $\{1, 2, \dots, \nu\}$, while the weights b_ℓ can be computed by an appropriate explicit formula. Then

$$Q[\mathcal{P}] = I[\mathcal{P}] + \mathcal{O}(h^{\sigma+1}),$$

where σ is the order of classical interpolatory quadrature in $[0, 1]$ with constant weight function and the nodes c_1, c_2, \dots, c_ν .

In particular, if c_1, \dots, c_ν are the nodes of Gaussian quadrature (that is, zeros of the Legendre polynomial P_ν , shifted to $[0, 1]$) then $\sigma = 2\nu$.

Note that the calculation of $Q[\mathcal{P}]$ requires just ν function values. Moreover, *the same function values are recycled for all the polytopes!* In other words, to discretize $\Omega_{[r]}(h)$ consistently with error $\mathcal{O}(h^{r+1})$, we require just $\lfloor (r+1)/2 \rfloor$ function evaluations altogether.

The *quid pro quo* is the huge number of algebraic operations needed to compute $Q[\mathcal{P}]$ for all terms in the truncated Magnus expansion, both because the number of expansion terms grows exponentially fast and since C_p^ν is a large set of cardinality that, again, grows exponentially with p . The cost of linear algebra can be reduced dramatically by employing an idea originally due to Munthe-Kaas & Owren (1999). To this end we need to assume that the nodes c_1, c_2, \dots, c_ν are symmetric with respect to the point $\frac{1}{2}$ – not an undue hardship since our natural choice, Gaussian points, obey this.

We replace $\{A_1, A_2, \dots, A_\nu\}$ by their linear combinations $\{B_1, B_2, \dots, B_\nu\}$ such that

$$\sum_{l=1}^{\nu} (c_k - \frac{1}{2})^{l-1} B_l = A_k, \quad k = 1, 2, \dots, \nu.$$

While $A_k = \mathcal{O}(h)$, it is possible to prove that $B_k = \mathcal{O}(h^k)$, $k = 1, \dots, \nu$. Consequently,

$$(3.3) \quad \mathbf{L}(A_{k_1}, A_{k_2}, \dots, A_{k_p}) = \mathcal{O}(h^p), \quad \mathbf{L}(B_{k_1}, B_{k_2}, \dots, B_{k_p}) = \mathcal{O}(h^{k_1+k_2+\dots+k_p}).$$

We say that $\mathbf{L}(A_{k_1}, A_{k_2}, \dots, A_{k_p})$ is of *grade* p , while $\mathbf{L}(B_{k_1}, B_{k_2}, \dots, B_{k_p})$ is of grade $\sum_{i=1}^p k_i$.

Since each A_k is a linear combination of B_l s while \mathbf{L} is a multilinear function, we can replace (3.2) with

$$(3.4) \quad \mathcal{Q}[\mathcal{P}] = \sum_{\ell \in \tilde{\mathbf{C}}_p^\nu} \tilde{b}_\ell \mathbf{L}(B_{l_1}, B_{l_2}, \dots, B_{l_p}),$$

where the weights \tilde{b}_ℓ can be obtained from the b_ℓ s by labourous, yet transparent calculation.

Rewriting (3.2) in the form (3.4) is a prerequisite for three steps that, in their totality, discard the great majority of terms, thereby rendering the cost of linear algebra considerably more affordable.

Firstly, recalling that σ is the order of the quadrature, hence that we already incur an error of $\mathcal{O}(h^{\sigma+1})$ in the formation of $\mathcal{Q}[\mathcal{P}]$, we can discard in (3.4) all terms which are $\mathcal{O}(h^{\kappa+1})$ for $\kappa \geq \sigma$. This means that we replace \mathbf{C}_p^ν by $\tilde{\mathbf{C}}_p^\nu \subset \mathbf{C}_p^\nu$ such that $\ell \in \tilde{\mathbf{C}}_p^\nu$ if $\sum_{i=1}^p l_i \leq \sigma$.

Secondly, it has been proved by Iserles et al. (2001) that both the Magnus expansion (2.4) and the truncated expansion $\Omega_{[r]}$ are *time symmetric*: they can be expanded in odd powers of h . Because of the symmetry of the nodes c_1, \dots, c_ν with respect to the midpoint, this feature is inherited by the quadrature formula. This means that linear combinations of terms of even grade vanish and we may discard them from the linear combination. We thus sum over $\tilde{\mathbf{C}}_p^\nu \subset \tilde{\mathbf{C}}_p^\nu$ where $\ell \in \tilde{\mathbf{C}}_p^\nu$ if $\sum_{i=1}^p l_i$ is even: this gets rid of roughly half the terms!

Finally, we exploit in an organised manner the range of redundancies inherent in a Lie algebra by virtue of the skew symmetry of the Lie bracket and the Jacobi identity. The suitable formalism is that of *graded Lie algebra* and we have already alluded to it in our definition of a grade of an expansion term.

Let $\mathbf{G} = \{G_i\}_{i \in I}$ be a set of *generators*, where I is either finite or countable. A Lie algebra \mathfrak{f} is *free* over I if

- i. It is true that $G_i \in \mathfrak{f}$, $i \in I$;
- ii. For any Lie algebra \mathfrak{h} and any function $I \rightarrow \mathfrak{h}$ there exists a unique Lie-algebra homomorphism $\pi : \mathfrak{f} \rightarrow \mathfrak{h}$ such that $\pi(G_i) = H_i \in \mathfrak{h}$ for all $i \in I$.

Let \mathfrak{f} be a free Lie algebra. A function $\omega : \mathbf{G} \rightarrow \mathbb{N}$ is said to be a *grading* of \mathfrak{f} . It is extended to $\omega : \mathfrak{f} \rightarrow \mathbb{N}$ by recursion,

$$H = [H_1, H_2] \quad \Rightarrow \quad \omega(H) = \omega(H_1) + \omega(H_2).$$

Grading imposes an order upon \mathfrak{f} . Given $G, H \in \mathfrak{f}$, we say that $G \prec H$ if either $\omega(G) < \omega(H)$ or $\omega(G) = \omega(H)$ and G precedes H in the lexicographic order imposed by the recursive construction of \mathfrak{f} from the generators \mathbf{G} .

The *Hall basis* \mathcal{H} of \mathfrak{f} is constructed recursively so as to contain terms of least grade. Specifically, it consists of

- (1) The set \mathbf{G} of generators;
- (2) $[G_i, G_j]$ if $G_i \prec G_j$;
- (3) Elements of the form $[H_i, [H_j, H_k]]$ where $H_i, H_j, H_k, [H_j, H_k] \in \mathcal{H}$ and $H_j \preceq H_i \prec [H_j, H_k]$.

Let $I = \{1, 2, \dots, \nu\}$. We denote by \mathcal{K}_ν^m the linear space of all terms of grade $m \in \mathbb{N}$ and observe that $\mathfrak{f} = \bigoplus_{m=1}^{\infty} \mathcal{K}_\nu^m$. The algorithm that has led us to the Hall basis can be extended transparently to provide a basis of \mathcal{K}_ν^m .

We consider free Lie algebras generated by $\{A_1, A_2, \dots, A_\nu\}$ and $\{B_1, B_2, \dots, B_\nu\}$ respectively, with the gradings $\omega(A_k) \equiv 1$ and $\omega(B_k) = k$. Note the significance of grades

to our narrative: $\omega(G) = k$ means that $G = \mathcal{O}(h^k)$. Note further that we need to retain in $\mathcal{Q}[\mathcal{P}]$ only terms H such that $\omega(H) \leq \sigma$. Thus, and in addition to the steps that have taken us from \mathbf{C}_p^ν to $\tilde{\mathbf{C}}_p^\nu$ to $\widehat{\mathbf{C}}_p^\nu$, we express all terms of grades $\leq \sigma$ as linear combinations of elements in the Hall bases of \mathcal{K}_ν^m for $m \leq \sigma$. The savings implicit in this procedure are underlined by a generalisation of the Witt formula,

$$\dim \mathcal{K}_\nu^m = \frac{1}{m} \sum_{d|m} \mu(d) \left(\sum_{i=1}^r \lambda_i^{m/d} \right)$$

(Munthe-Kaas & Owren 1999), where μ is the *Möbius function*

$$\mu(d) = \begin{cases} 1, & d = 1, \\ (-1)^q, & d = p_1 p_2 \cdots p_q, p_1 < p_2 < \cdots < p_q \text{ prime,} \\ 0, & \text{otherwise,} \end{cases}$$

$r = \max_{i=1, \dots, \nu} \omega(G_i)$ and $\lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}$ are the zeros of $1 - \sum_{i=1}^\nu z^{\omega(G_i)}$. Our conclusion is that $\dim \mathcal{K}_\nu^m$ is surprisingly small, and this can be exploited to reduce the volume of linear algebra.

TABLE 1. The number of terms in different quadrature formulæ for Gaussian nodes.

| ν | 1 | 2 | 3 | 4 | 5 |
|--|---|---|----|------|---------|
| $\sigma = 2\nu$ | 2 | 4 | 6 | 8 | 10 |
| $\sum_{m=1}^{2\nu-1} \rho_m$ | 1 | 5 | 80 | 3304 | 1256567 |
| $\sum_{m=1}^{2\nu-1} \bar{\rho}_m$ | 1 | 3 | 10 | 33 | 111 |
| $\sum_{\substack{m=1 \\ m \text{ odd}}}^{2\nu-1} \bar{\rho}_m$ | 1 | 2 | 7 | 22 | 73 |

We let $\rho_m = \dim \mathcal{K}_\nu^m$ for the generators $\mathbf{G} = \{A_1, A_2, \dots, A_\nu\}$ and $\bar{\rho}_m = \dim \mathcal{K}_\nu^m$ for $\mathbf{G} = \{B_1, B_2, \dots, B_\nu\}$. Choosing Gaussian nodes c_1, c_2, \dots, c_ν (hence $\sigma = 2\nu$) and expressing all terms in $\mathcal{Q}[\mathcal{P}]$ in a Hall basis, we compare in Table 1 the number of terms using the two sets of generators and, for the latter, discarding all terms of even grade. The savings, in particular for large ν , are remarkable. Insofar as numerical analysis is concerned, they represent a difference between a mere curiosity and an effective computational algorithm.

Even this, though, is not the final word in applying algorithmic ingenuity, underpinned by algebraic theory, to the design of effective numerical Magnus integrators. Blanes, Casas & Ros (2000) have developed a technique which, exploiting further symmetries and structure of the underlying problem, allows to aggregate terms and reduce further the cost of computation per step. For example, for $\nu = 3$ we commence from Gaussian nodes, whereby

$$A_1 = hA\left(\left(\frac{1}{2} - \frac{\sqrt{15}}{10}\right)h\right), \quad A_2 = hA\left(\frac{1}{2}h\right), \quad A_3 = hA\left(\left(\frac{1}{2} + \frac{\sqrt{15}}{10}\right)h\right)$$

and

$$B_1 = A_2, \quad B_2 = \frac{\sqrt{15}}{3}(A_3 - A_1), \quad B_3 = \frac{10}{3}(A_3 - 2A_2 + A_1).$$

Once we exploit all the three mechanisms that have led us to the bottom row of Table 1, we obtain the sixth-order method

$$\begin{aligned} \tilde{\Omega} = B_1 + \frac{1}{12}B_3 - \frac{1}{12}[B_1, B_2] + \frac{1}{240}[B_2, B_3] + \frac{1}{360}[B_1, [B_1, B_3]] - \frac{1}{240}[B_2, [B_1, B_2]] \\ + \frac{1}{720}[B_1, [B_1, [B_1, B_2]]]. \end{aligned}$$

Note that only odd-grade terms are present and that we need to compute seven commutators: this is consistent with Table 1.

Alternatively, introducing $\mathcal{O}(h^7)$ changes that have no bearing on the order of the method, we compute

$$\begin{aligned} C_1 &= [B_1, B_2], \\ C_2 &= [B_1, 2B_3 + C_1], \\ C_3 &= [-20B_1 - B_3 + C_1, B_2 - \frac{1}{60}C_2], \\ \bar{\Omega} &= B_1 + \frac{1}{12}B_3 + \frac{1}{240}C_3, \end{aligned}$$

another sixth-order approximation but requiring just three commutators per step.

3.2. Cayley expansions and Hierarchical algebras. Wishing to do unto the truncated Cayley expansion (2.6) what we have done for Magnus expansions, we are stumped by the presence of terms of the form $BCB \in \mathfrak{g}$, where $B, C \in \mathfrak{g}$ and, more generally, by the fact that a quadratic Lie algebra is closed under the symmetric product $BCD + DCB$ for all $B, C, D \in \mathfrak{g}$. Graded free Lie algebras are inadequate for the task in hand, and this has motivated the introduction of *hierarchical algebras* (Iserles & Zanna 2000).

Let $(\mathfrak{g}, +)$ be an Abelian group over a field of zero characteristic and assume the existence of a countable family of m -nary operations

$$[\cdot, \dots, \cdot]_m : \overbrace{\mathfrak{g} \times \mathfrak{g} \times \dots \times \mathfrak{g}}^{m \text{ times}} \rightarrow \mathfrak{g}, \quad m \in \mathbb{N},$$

subject to the following three axioms,

Alternate symmetry:: For every $m \in \mathbb{N}$ and $F_1, \dots, F_m \in \mathfrak{g}$

$$[F_1, F_2, \dots, F_m]_m + (-1)^m [F_m, F_{m-1}, \dots, F_1]_m = 0,$$

where 0 is the zero of \mathfrak{g} ;

Multilinearity:: Each $[\cdot, \dots, \cdot]_m$ is linear in all its components;

Hierarchy condition:: For every $m, n \in \mathbb{N}$, $k \in \{1, 2, \dots, m\}$,

$$F_1, \dots, F_{k-1}, F_{k+1}, \dots, F_m \in \mathfrak{g} \quad \text{and} \quad E_1, \dots, E_n \in \mathfrak{g}$$

it is true that

$$\begin{aligned} &[F_1, \dots, F_{k-1}, [E_1, \dots, E_n]_n, F_{k+1}, \dots, F_m]_m \\ &= [F_1, \dots, F_{k-1}, E_1, \dots, E_n, F_{k+1}, \dots, F_m]_{m+n-1} \\ &\quad - (-1)^n [F_1, \dots, F_{k-1}, E_n, \dots, E_1, F_{k+1}, \dots, F_m]_{m+n-1}. \end{aligned}$$

Such a \mathfrak{g} , equipped with the above structure, is called a *hierarchical algebra*.

THEOREM 3. (Iserles & Zanna 2000)

- (1) Every hierarchical algebra is a Lie algebra with commutator $[F_1, F_2] = [[F_1, F_2]]_2$;
 (2) Every quadratic Lie algebra is a hierarchical algebra with

$$[[F_1, F_2, \dots, F_m]]_m = F_1 F_2 \cdots F_m - (-1)^m F_m F_{m-1} \cdots F_1, \quad m \in \mathbb{N}.$$

We define a *free hierarchical algebra (FHA)* \mathfrak{f} in the usual manner. An element $F \in \mathfrak{f}$ is said to be *primitive* if it is of the form $F = [[G_{i_1}, G_{i_2}, \dots, G_{i_r}]]_r$ for some $r \in \mathbb{N}$, where $\mathbf{G} = \{G_1, \dots, G_\nu\}$ are the generators of a FHA.

LEMMA 4. (Iserles & Zanna 2000) *Every element of a free hierarchical algebra can be represented as a linear combination of primitive elements.*

Grading can be imposed on a FHA exactly like before. Let

$$u(t) = 1 - \sum_{k=1}^{\nu} z^{\omega(G_k)},$$

where $\mathbf{G} = \{G_1, \dots, G_\nu\}$. It is possible to prove with much effort that, formally,

$$\sum_{m=1}^{\infty} t^{m-1} \dim \mathcal{K}_\nu^m = \frac{1}{2} \left[\frac{1}{u(t)} + \frac{u(t)}{u(t^2)} \right],$$

where, as before, \mathcal{K}_ν^m stands for the linear space spanned by grade- m terms.

Assuming further that the zeros $\lambda_1, \lambda_2, \dots, \lambda_r$ of u (which we have already encountered in the previous subsection, recall that $r = \max \omega(G_i)$) are distinct, it follows that

$$\begin{aligned} \dim \mathcal{K}_\nu^{2m} &= \frac{1}{2} \sum_{k=1}^r \frac{\lambda_k^{-m-1}}{u'(\lambda_k)} \left\{ 2 - \lambda_k^{-m} - \frac{1}{2}[u(\lambda_k^{1/2}) + u(-\lambda_k^{1/2})] \right\}, \\ \dim \mathcal{K}_\mu^{2m+1} &= \frac{1}{2} \sum_{k=1}^r \frac{\lambda_k^{-m-3/2}}{u'(\lambda_k)} \left\{ -\lambda_k^{-m-\frac{1}{2}} + \frac{1}{2}[u(\lambda_k^{1/2}) - u(-\lambda_k^{1/2})] \right\}. \end{aligned}$$

Since Hall-type bases can be obtained by recursion within a FHA setting, we have all the building blocks to construct efficient quadratures associated with Cayley expansions. For example, a sixth-order method (with the same B_1, B_2, B_3 as for the sixth-order Magnus method of the last subsection) reads (in terms of the m -nary products – this somewhat complicates the presentation but renders it consistent with our framework)

$$\begin{aligned} \tilde{\Omega} &= [[B_1]]_1 + \frac{1}{12} [[B_3]]_1 - \frac{1}{12} [[B_1, B_2]]_2 - \frac{1}{24} [[B_1, B_1, B_1]]_3 + \frac{1}{240} [[B_2, B_3]]_2 \\ &\quad + \frac{1}{240} [[B_1, B_2, B_2]]_3 - \frac{1}{240} [[B_1, B_1, B_3]]_3 - \frac{1}{240} [[B_2, B_1, B_2]]_3 - \frac{1}{160} [[B_1, B_3, B_1]]_3 \\ &\quad + \frac{1}{120} [[B_1, B_1, B_1, B_2]]_4 - \frac{1}{240} [[B_1, B_1, B_2, B_1]]_4 + \frac{1}{240} [[B_1, B_1, B_1, B_1]]_5. \end{aligned}$$

Note that, again, only terms of odd grade are required. Although more complicated than $\tilde{\Omega}$, to say nothing of $\bar{\Omega}$, this approach has an important advantage in the case of quadratic Lie algebras, in particular when the dimension is large, since solving linear systems is considerably cheaper than computing a matrix exponential.

References

- Abia, L. & Sanz-Serna, J. M. (1993), ‘Partitioned Runge–Kutta methods for Hamiltonian problems’, *Maths Comp.* **60**, 617–634.
 Bialynicki-Birula, I., Mielnik, B. & Plebański, J. (1969), ‘Explicit solution of the continuous Baker–Campbell–Hausdorff problem and a new expression for the phase operator’, *Ann. Phys.* **51**, 187–200.

- Blanes, S., Casas, F. & Ros, J. (2000), 'Improved high order integrators based on the Magnus expansion', *BIT* **40**, 434–450.
- Butcher, J. C. (1963), 'Coefficients for the study of Runge–Kutta integration processes', *J. Austr. Math. Soc.* **3**, 185–201.
- Butcher, J. C. (1972), 'An algebraic theory of integration methods', *Maths Comp.* **26**, 79–106.
- Casas, F. (2007), 'Sufficient conditions for the convergence of the magnus expansion', *J. Phys. A: Math. Theor.* **40**, 15001–15007.
- Casas, F. & Iserles, A. (2006), 'Explicit Magnus expansions for nonlinear equations', *J. Phys. A: Math. Gen.* **39**, 5445–5461.
- Celledoni, E., Iserles, A., Nørsett, S. P. & Orel, B. (2002), 'Complexity theory for Lie-group solvers', *J. Complexity* **18**, 242–286.
- Chartier, P. & Murua, A. (2007), 'Preserving first integrals and volume forms of additively split systems', *IMA J. Num. Anal.* **27**, 381–405.
- Chen, T.-K. (1977), 'Iterated path integrals', *Bull. Amer. Math. Soc.* **83**, 831–879.
- Connes, A. & Kreimer, D. (1998), 'Hopf algebras, renormalization and noncommutative geometry', *Comm. Math. Phys.* **199**, 203–242.
- Connes, A. & Kreimer, D. (2000), 'Renormalization in quantum field theory and the Riemann–Hilbert problem I. The Hopf algebra structure of graphs and the main theorem', *Comm. Math. Phys.* **210**, 249–273.
- Crouch, P. E. & Grossman, R. (1993), 'Numerical integration of ordinary differential equations on manifolds', *J. Nonlinear Sci.* **3**, 1–33.
- Davis, P. J. & Rabinowitz, P. (1984), *Methods of Numerical Integration*, 2nd edn, Academic Press, Orlando, FL.
- Fer, F. (1958), 'Résolution de l'équation matricielle $\dot{U} = pU$ par produit infini d'exponentielles matricielles', *Bull. Classe des Sci. Acad. Royal Belg.* **44**, 818–829.
- Fomenko, A. T. & Chakon, R. V. (1990), 'Recurrence formulas for homogeneous terms of a convergent series that represents a logarithm of a multiplicative integral on Lie groups', *Funct. Anal. Applcs* **24**, 41–49.
- Grossman, R. & Larson, R. G. (1989), 'Hopf-algebraic structure of families of trees', *J. Algebra* **126**, 184–210.
- Hairer, E. & Wanner, G. (1974), 'On the Butcher group and general multi-valued methods', *Computing* **13**, 1–15.
- Hairer, E., Lubich, C. & Wanner, G. (2003), 'Geometric numerical integration illustrated by the Störmer–Verlet method', *Acta Numerica* **12**, 399–450.
- Hairer, E., Lubich, C. & Wanner, G. (2006), *Geometric Numerical Integration*, 2nd edn, Springer-Verlag, Berlin.
- Hairer, E., Nørsett, S. P. & Wanner, G. (1986), *Solving Ordinary Differential Equations I: Nonstiff Problems*, Springer-Verlag, Berlin.
- Hausdorff, F. (1906), 'Die symbolische Exponentialformel in der Gruppentheorie', *Berichte der Sächsischen Akademie der Wissenschaften (Math. Phys. Klasse)* **58**, 19–48.
- Hubert, E. (2000), 'Factorization-free decomposition algorithms in differential algebra', *J. Symbolic Comput.* **29**, 641–662.
- Iserles, A. (1994), 'On nonlinear delay differential equations', *Trans. Amer. Math. Soc.* **15**, 441–477.
- Iserles, A. (2001a), 'A Magnus expansion for the equation $Y' = AY - YB$ ', *J. Comp. Maths* **19**, 15–26.
- Iserles, A. (2001b), 'On Cayley-transform methods for the discretization of Lie-group equations', *Found. Comp. Maths* **1**, 129–160.
- Iserles, A. (2002), 'On the discretization of double-bracket flows', *Found. Comp. Maths* **2**, 305–329.
- Iserles, A. & Nørsett, S. P. (1999), 'On the solution of linear differential equations in Lie groups', *Royal Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 983–1020.
- Iserles, A. & Zanna, A. (2000), 'On the dimension of certain graded Lie algebras arising in geometric integration of differential equations', *LMS J. Comput. & Maths* **3**, 44–75.
- Iserles, A., Munthe-Kaas, H. Z., Nørsett, S. P. & Zanna, A. (2000), 'Lie-group methods', *Acta Numerica* **9**, 215–365.
- Iserles, A., Nørsett, S. P. & Rasmussen, A. F. (2001), 'Time symmetry and high-order Magnus methods', *Appld Num. Maths* **39**, 379–401.
- Iserles, A., Quispel, G. R. W. & Tse, P. S. P. (2007), 'B-series methods cannot be volume-preserving', *BIT* **47**, 351–378.
- Leimkuhler, B. & Reich, S. (2004), *Simulating Hamiltonian Dynamics*, Cambridge University Press, Cambridge.
- Lewis, D. & Simo, J. C. (1994), 'Conserving algorithms for the dynamics of Hamiltonian systems on Lie groups', *J. Nonlinear Sci.* **4**, 253–299.
- Magnus, W. (1954), 'On the exponential solution of differential equations for a linear operator', *Comm. Pure Appl. Math.* **7**, 649–673.

- Marsden, J. E. & West, M. (2001), ‘Discrete mechanics and variational integrators’, *Acta Numerica* **10**, 357–514.
- McLachlan, R. I. & Quispel, G. R. W. (2002), ‘Splitting methods’, *Acta Numerica* **11**, 341–434.
- Mielnik, B. & Plebański, J. (1970), ‘Combinatorial approach to Baker–Campbell–Haussdorff exponents’, *Ann. Inst. Henri Poincaré A* **12**, 215–254.
- Moan, P. C. & Niesen, J. (2008), ‘Convergence of the Magnus series’, *Found. Comput. Maths* **8**, 291–301.
- Munthe-Kaas, H. (1998), ‘Runge–Kutta methods on Lie groups’, *BIT* **38**, 92–111.
- Munthe-Kaas, H. & Owren, B. (1999), ‘Computations in a free Lie algebra’, *Royal Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 957–982.
- Munthe-Kaas, H. Z. & Wright, W. M. (2008), ‘On the Hopf algebraic structure of Lie group integrators’, *Found. Comput. Maths* **8**, 227–257.
- Murua, A. (2006), ‘The Hopf algebra of rooted trees, free Lie algebras, and Lie series’, *Found. Comput. Maths* **6**, 387–426.
- Murua, A. & Sanz-Serna, J. M. (1999), ‘Order conditions for numerical integrators obtained by composing simpler integrators’, *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **357**, 1079–1100.
- Owren, B. & Marthinsen, A. (2001), ‘Integration methods based on canonical coordinates of the second kind’, *Numer. Math.* **87**, 763–790.
- Strichartz, R. S. (1987), ‘The Campbell–Baker–Hausdorff–Dynkin formula and solutions of differential equations’, *J. Func. Anal.* **72**, 320–345.
- Varadarajan, V. S. (1984), *Lie Groups, Lie Algebras, and their Representations*, GTM 102, Springer-Verlag, New York.
- Zanna, A. (1999), ‘Collocation and relaxed collocation for the Fer and the Magnus expansions’, *SIAM J. Num. Anal.* **36**, 1145–1182.

DEPARTMENT OF APPLIED MATHEMATICS AND THEORETICAL PHYSICS, CENTRE FOR MATHEMATICAL SCIENCES, UNIVERSITY OF CAMBRIDGE, WILBERFORCE RD, CAMBRIDGE CB3 0WA, UK

Wilsonian renormalization, differential equations and Hopf algebras

T. Krajewski and P. Martinetti

ABSTRACT. In this paper, we present an algebraic formalism inspired by Butcher's B-series in numerical analysis and the Connes-Kreimer approach to perturbative renormalization. We first define power series of non linear operators and propose several applications, among which the perturbative solution of a fixed point equation using the non linear geometric series. Then, following Polchinski, we show how perturbative renormalization works for a non linear perturbation of a linear differential equation that governs the flow of effective actions. Furthermore, we define a general Hopf algebra of Feynman diagrams adapted to iterations of background field effective action computations. As a simple combinatorial illustration, we show how these techniques can be used to recover the universality of the Tutte polynomial and its relation to the q -state Potts model. As a more sophisticated example, we use ordered diagrams with decorations and external structures to solve the Polchinski's exact renormalization group equation. Finally, we work out an analogous construction for the Schwinger-Dyson equations, which yields a bijection between planar ϕ^3 diagrams and a certain class of decorated rooted trees.

CONTENTS

| | |
|--|-----|
| 1. Introduction | 187 |
| 2. Basics of wilsonian renormalization | 189 |
| 3. Rooted trees and power series of non linear operators | 202 |
| 4. Renormalization, effective actions and Feynman diagrams | 213 |
| 5. Conclusion and outlook | 234 |
| Acknowledgements | 235 |
| References | 235 |

1. Introduction

The last century has witnessed the development of Quantum Field Theory (QFT) as the adequate framework for the formulation of many physical theories, ranging from elementary particles to condensed matter physics. Historically, it was

first devised to describe the interaction of charged particles with the electromagnetic field, leading to Quantum Electrodynamics (QED). However, QED was shown to be plagued by a surprising disease: When expanded in powers of the electric charge e , corrections to the leading terms turn out to be divergent. The occurrence of these infinities can be traced to virtual processes involving very high energies. Then, if we limit the energy these processes involve to a scale Λ and choose the electric charge $e(\Lambda)$ as a suitable function of Λ , all the divergent quantities can be made finite, provided they are expressed in terms of a few low energy parameters to be determined by experiment, like the masses of the particles and the strength of their interactions. This is the basic idea of renormalization: Processes occurring at a given scale must be formulated using parameters adapted to these scale.

Later on, QFT progressively invaded almost all fields of theoretical physics. Renormalizability was a guideline for the construction of the standard model of elementary particles while critical phenomena in condensed matter were also understood using renormalization. More in general, QFT describes fluctuating systems with continuous degrees of freedom and is best understood through paths integrals of the type

$$(1) \quad \int [D\phi] e^{-S[\phi]},$$

where the integration is over a space of functions. When considered as a perturbation of a Gaussian integral, this path integral is expanded over Feynman diagrams. Because of the continuous nature of the path integral, some of these diagrams yield divergent quantities, as dictated by dimensional analysis, and can only acquire a meaning through renormalization. In a certain sense, it can be said that the founding fathers of QFT (Dirac, Feynman, Schwinger, Tomonaga, Dyson, Weinberg, t’Hooft, Veltman, Wilson and many others) have achieved an infinite dimensional analogue of Newton’s work: Driven by fundamental physics, they devised a new form of calculus based on path integrals, Feynman diagrams and renormalization.

However, it is fair to say that this new calculus is not yet fully understood. In particular, it hides an algebraic structure analogous to diffeomorphisms, as uncovered by Connes and Kreimer a decade ago (see [1], [2] and [3] as well as the first chapter of the book [4] and [5] for recent reviews). To be specific, divergent Feynman diagrams generate a commutative Hopf algebra whose associated group is a refinement of the group of diffeomorphisms of the coupling constant. This formalism is versatile enough to encompass the infinite dimensional analogue of the change of parameters provided by the computation of the Wilsonian effective action, as we show in this paper. Our contribution is organized as follows.

In the first section, we give a terse account of renormalization. We first recall the elementary particle physics point of view and introduce Feynman diagrams. Then, we introduce Wilson’s viewpoint, first for spin systems and then for general QFT. This is textbook physics and can be found in modern texts on QFT (to quote only a few authors who strongly influenced our presentation, see the textbook [6] by Zee for an eclectic overview of QFT applied to particle and condensed matter physics, the treatise by Zinn-Justin [7] for a general treatment of QFT in the path integral formalism and the review by Rivasseau [8] for an introduction to renormalization, accessible to a large audience). We close this section by recalling a renormalization group proof of the central limit theorem of probability theory that may help mathematicians grasp Wilson’s ideas.

The next section is devoted to rooted trees and their Hopf algebra. These were introduced as fundamental tools in the numerical analysis of a differential equation by Butcher [9]. We revisit this construction and interpret trees as indices for power series of non linear operators, with composition law given by the convolution product. We illustrate this construction with the non linear geometric series that provides a perturbative solution of a general fixed point equation. This can be used to resum tree like structures as shown on a elementary example based on Gaußian random matrices.

In the last section, we deal with renormalization. We first formulate perturbative renormalization for a general renormalization group differential equation. Then, we introduce a general Hopf algebra of Feynman diagrams that encodes successive effective action computations. We illustrate the use of this Hopf algebra on two examples. First, we give a new proof of the universality of Tutte polynomial and of its relation to the Potts model. Then, we use the Hopf algebra to solve Polchinski's equation that describes the flow of effective actions in QFT. Finally, in the last section we investigate an analogous structure based on the Schwinger-Dyson equations. As a byproduct, we obtain a new bijection between planar ϕ^3 diagrams and a certain class of decorated rooted trees.

2. Basics of wilsonian renormalization

2.1. What is renormalization? In its most general acceptance, renormalization can be defined as the general methods that implement the change of parameters in the description of a physical system necessary in order to take into account the indirect effects of unobserved degrees of freedom, usually living on a much shorter length scale. The simplest example we can think of is the propagation of light in a medium, as described by geometrical optics. Whereas in the vacuum the speed of light is c , in some medium it is reduced to c/n , with n the refractive index, as a consequence of the interaction of light with the atoms of the medium. While the latter are not described by geometrical optics, their effect on the propagation of light is simply taken into account by the substitution $c \rightarrow c/n$.

As a second example, let us consider the motion of a ping-pong ball of inertial mass m_{inert} , gravitational mass m_{grav} and volume V in a perfect fluid of density ρ , as popularized by Connes [4]. A straightforward application of Newton's law of dynamics yields its acceleration through

$$(2) \quad m_{\text{inert}}\mathbf{a} = m_{\text{grav}}\mathbf{g} - \rho V\mathbf{g},$$

where the first term is the ball's weight and the second one Archimedes' buoyant force. If the ball is ten time lighter than the equivalent volume of water, then it experiences an acceleration $\mathbf{a} = -9\mathbf{g}$, which is obviously nonsensical. The hole in the previous reasoning is that we did not take into account the inertia of the surrounding fluid. Using a suitable ansatz for the fluid's velocity field, it can be shown [4] that the latter implies a shift of the inertial mass of the ball

$$(3) \quad m_{\text{inert}} \rightarrow m_{\text{inert}} + \frac{\rho V}{2},$$

which yields the more acceptable acceleration of $\mathbf{a} = -3/2\mathbf{g}$. Both these elementary examples illustrate the main feature of renormalization: In a simple theory (geometrical optics, Newton's law of dynamics), the effect of unobserved degrees of

freedom, which should be described by a more sophisticated theory (electrodynamics, fluid mechanics), are taken into account by a mere change of parameters in the simpler theory.

The very same type of ideas show off in QFT, which grew up from the marriage between special relativity and quantum mechanics. First, recall that one of the main lessons from special relativity is the equivalence between matter and energy, as neatly summarized by Einstein's celebrated relation $E = mc^2$. This is currently observed in high energy physics experiments: Incoming particles (say leptons, like a pair e^+e^-) with a total given energy collide and result in some outgoing particles whose nature may be totally different (for instance, a pair $q\bar{q}$). Of course, total energy is always conserved so that heavy particles may appear only if the incoming energy is high enough.

On the other hand, quantum mechanics provides us with a description of the microscopic world. Its predictions, carefully verified by experiments, yield the probability of finding the system in a final state $|f\rangle$ knowing that it was initially in the state $|i\rangle$. The transition amplitude from an initial state to the final one is often expressed as a sum over all (unobserved) intermediate states $|n\rangle$,

$$(4) \quad \langle f|i\rangle = \sum_n \langle f|n\rangle \langle n|i\rangle.$$

In QFT, these initial and final states represent a fixed number of elementary particles of given energy but the summation over all the intermediate states involves an arbitrary number of virtual particles of arbitrarily high energies. It is fundamental to note that none of the intermediate states are observed since the experiment only answers the question "*What is the probability of finding the system in the state $|f\rangle$ at the end of the experiment if we know that it was initially in the state $|i\rangle$?*".

In most cases of interest, the sum over intermediate states diverges as the upper limit of the energy Λ of the virtual particles go to infinity. However, for the so-called renormalizable theories, it turns out that these infinities cancel when the parameters of the theory (masses $m(\Lambda)$ and coupling constants $g(\Lambda)$) are suitably chosen as functions of Λ , which may then be taken to infinity leaving physically relevant quantities finite. This is the process of renormalization, as it occurs in high energy physics: The effect of unobserved virtual states of high energy is taken into account in some of the parameters of the theory. In a certain sense, the cut-off Λ sets a limit for the validity of our field theory and the sole effect of the very high energy modes beyond Λ is to adapt the parameters of the theory to energy of the process we consider. Alternatively, this can also be considered as an ignorance of the very small scale structure of space-time since high energy correspond to small distances by Fourier transform.

2.2. Feynman diagrams and perturbative renormalization. As far as renormalization is concerned, QFT is best studied in the framework of the Euclidian path integral. In the simplest case, the fields are defined as functions ϕ from space-time \mathbb{R}^D to \mathbb{R} and the dynamics is governed by the action functional,

$$(5) \quad S[\phi] = \int d^D x \left(\frac{1}{2} \partial_\mu \phi \partial^\mu \phi + \frac{1}{2} m^2 \phi^2 + \frac{g}{N!} \phi^N \right),$$

where m is a mass and g a coupling constant for a N -particle interaction. At the quantum level, all the information is encoded in the Green's functions,

$$(6) \quad G(x_1, \dots, x_n, m, g) = \mathcal{N} \int [D\phi] e^{-\frac{S[\phi]}{\hbar}} \phi(x_1) \cdots \phi(x_n),$$

where the integration is over the space of all fields and \mathcal{N} is a normalization constant to be defined later. $\hbar = 6.02 \cdot 10^{-34}$ J·s is Planck's constant and measures the deviation of the quantum theory from the classical one. In most of the paper, we choose unit such that $\hbar = 1$. Note that this is only the simplest model of a scalar field theory. More complicated action functionals are required to account for the real world physics: Spinors ψ for fermionic particles and gauge connections A for their interactions, as is the case for QED and the Standard Model of elementary particles. Nevertheless, we restrict our attention in the sequel to the simplest example of a scalar field theory.

To give a precise meaning to (6), it is convenient to expand $G(x_1, \dots, x_n, m, g)$ as a power series in g using Feynman diagrams. The latter are best introduced on simpler analogue, with the space of fields replaced by a finite dimensional vector space. Thus, the equivalent of (6) is

$$(7) \quad G_{i_1, \dots, i_n}(A, V) = \mathcal{N} \int d\phi e^{-S(\phi)} \phi_{i_1} \cdots \phi_{i_n},$$

with the action

$$(8) \quad S(\phi) = \frac{1}{2} \phi \cdot A^{-1} \cdot \phi + V(\phi).$$

The quadratic term is defined by a positive definite symmetric matrix A ,

$$(9) \quad \phi \cdot A^{-1} \cdot \phi = \sum_{i,j} (A^{-1})_{ij} \phi_j \phi_i,$$

and the interaction potential V is a polynomial in all the fields

$$(10) \quad V(\phi) = \sum_N \sum_{i_1, \dots, i_N} \frac{g_{i_1, \dots, i_N}}{N!} \phi_{i_1} \cdots \phi_{i_N}.$$

In this case we choose $\mathcal{N}^{-1} = \det A / 2\pi$ and expand $e^{-V(\phi)}$ as a power series in the couplings g_{i_1, \dots, i_N} . Thus, (7) amounts to the computation of the average of a monomial using a Gaussian weight and is given by Wick's theorem: It is a sum over all possible pairings of the variables in the monomial, each pairing of ϕ_i with ϕ_j being weighted by A_{ij} . For example,

$$(11) \quad \phi_i \phi_j \phi_k \phi_l \rightarrow A_{ij} A_{kl} + A_{ik} A_{jl} + A_{il} A_{jk}.$$

Then, each term of the expansion of (7) is associated with a diagram with n external legs and vertices of valence N ,

$$(12) \quad G_{i_1, \dots, i_n}(A, V) = \sum_{\gamma} \frac{G_{i_1, \dots, i_n}^{\gamma}(A, V)}{S_{\gamma}}.$$

The contribution of each diagram is computed using the Feynman rules:

- associate the indices i_1, \dots, i_n to the external legs and indices j_k to the internal half edges;
- associate a matrix element $A_{j_k j_l}$ to any edge connecting the indices j_k and j_l ;

- associate a coupling $-g_{j_1, \dots, j_N}$ to a N -valent vertex whose half edges have indices j_1, \dots, j_N ;
- sum over all the indices j_k .

Besides, one has to divide by the symmetry factor S_γ which is the cardinal of the automorphism group of the diagram, leaving the external legs fixed. For example,

$$(13) \quad i_1 - \text{Diagram} - i_2 \quad \rightarrow \quad \frac{1}{2} \sum_{\substack{j_1, j_2, j_3 \\ j_4, j_5, j_6}} A_{i_1, j_1} g_{j_1, j_2, j_3} A_{j_2, j_4} A_{j_3, j_5} g_{j_4, j_5, j_6} A_{j_6, i_2}.$$

This simple finite dimensional model already captures some important algebraic aspects of perturbation theory as will be discussed in the last part devoted to the Hopf algebras based on Feynman diagrams.

At a formal level, the Green's functions (6) can be computed as a power series in g by replacing ϕ_i by a function $x \mapsto \phi(x)$, the matrix element A_{ij} by the propagator

$$(14) \quad K(x, y) = \int_{\mathbb{R}^D} \frac{d^D p}{(2\pi)^D} \frac{e^{ip \cdot (x-y)}}{p^2 + m^2}.$$

and $V(\phi)$ by the interaction term

$$(15) \quad \frac{g}{N!} \int_{\mathbb{R}^D} d^D x \phi^N(x) = \frac{g}{N!} \int_{\mathbb{R}^D} \frac{d^D p_1}{(2\pi)^D} \dots \frac{d^D p_N}{(2\pi)^D} (2\pi)^D \delta(p_1 + \dots + p_N) \tilde{\phi}(p_1) \dots \tilde{\phi}(p_N),$$

with

$$(16) \quad \tilde{\phi}(p) = \int_{\mathbb{R}^D} d^D x e^{-ip \cdot x} \phi(x)$$

the Fourier transform of $\phi(x)$. Heuristically, the Feynman diagrams can be thought of as quantum mechanical processes with particles on their external legs and virtual particles of momenta p propagating on the internal lines. It is important to notice that although momentum is conserved at each vertex and along each line, the particles that propagate along the loops¹ may have arbitrary momenta.

The integral along loop momenta extends on the whole of \mathbb{R}^D and may lead to divergencies since the Fourier transform of propagator does not decrease fast enough at large p . A connected diagram with I internal lines, V vertices and $L = I - V + 1$ loops behaves in dimension D like p^ω with $\omega = LD - 2I$ the superficial degree of divergence. Using $NV = 2I + E$, the latter can be written as

$$(17) \quad \omega = V \left(\frac{ND}{2} - D - N \right) + E \left(1 - \frac{D}{2} \right) + D$$

When the coefficient of V is strictly positive, divergencies occur for any Green's function and such a theory cannot be renormalized. If this coefficient is strictly negative, then there are only a finite number of divergent diagrams. These are the superrenormalizable theories which are nevertheless of limited interest in physics. Finally, the critical situation defines the renormalizable theories, like ϕ^4 in $D = 4$ or ϕ^3 in $D = 6$. In this case, the divergencies occur only for the diagrams with $E \leq N$ and may be absorbed into the coefficients of a polynomial interaction of degree N . At the notable exception of gravity, all the interactions of elementary particles can be formulated using renormalizable interactions.

¹What is called loop in the physics terminology adopted here is called a cycle by graph theorists.

To prevent the propagation of Fourier modes of momenta $\geq \Lambda$, let us alter the propagator by introducing a cut-off Λ ,

$$(18) \quad K(x, y) \rightarrow K_\Lambda(x, y) = \int_{\frac{1}{\Lambda^2}}^{\infty} d\alpha \int_{\mathbb{R}^D} \frac{d^D p}{(2\pi)^D} e^{ip \cdot (x-y)} e^{-\alpha(q^2 + m^2)}.$$

This procedure is known as the regularization and can be performed in various ways: Besides the method used here, one could also discretize the theory on a lattice or evaluate the diagrams in complex dimension z and recover the divergences as poles when $z \rightarrow D$. In principle all these methods are equivalent but we restrict ourselves to the momentum space cut-off presented here since it is suited to the Wilsonian point of view we adopt in this paper.

For a renormalizable theory like the ϕ^4 theory, one can trade the parameters g and m for some cut-offs dependent ones $g_0(\Lambda)$ and $m_0(\Lambda)$ and further introduce an additional wave function renormalization $Z(\Lambda)$ in such a way that $Z^{\frac{n}{2}}(\Lambda)G_\Lambda(x_1, \dots, x_n, m_0(\Lambda), g_0(\Lambda))$ admits a finite limit when $\Lambda \rightarrow \infty$. To obtain definite physical predictions, the bare parameters $g_0(\Lambda)$ and $m_0(\Lambda)$ and the wave function renormalization $Z(\Lambda)$ must be determined in terms of normalization conditions involving renormalized parameters m_r and g_r measured at a low energy scale μ . Thus, we define the renormalized Green's functions as

$$(19) \quad G_r(x_1, \dots, x_n, m_r, \mu, g_r) = \lim_{\Lambda \rightarrow \infty} Z^{\frac{n}{2}}(\Lambda, m_r, g_r, \mu) G_\Lambda(x_1, \dots, x_n, m_0(\Lambda, m_r, g_r, \mu), g_0(\Lambda, m_r, g_r, \mu))$$

Note that we are dealing here with perturbative renormalization only, so that the previous equality must be understood as an equality between formal power series in g_r . In fact, $g_0(\Lambda)$, $m_0(\Lambda)$ and $Z(\Lambda)$ are themselves formal power series in g_r that can be computed in terms of Feynman diagrams using the Bogoliubov-Parasiuk-Hepp-Zimmermann (BPHZ) formula. Roughly speaking, the contribution of a divergent diagram with 2 or 4 external legs to the renormalisation of the parameters is encoded in its counterterm $C(\gamma)$ which is determined recursively by the relation

$$(20) \quad C(\gamma) = T \left(\sum_{\substack{\{\gamma_i, \dots\} \\ \gamma_i \cap \gamma_j = \emptyset}} \prod_i C(\gamma_i) \frac{\gamma}{\prod_i \gamma_i} \right),$$

with T taking the divergent part of the diagram. This sum runs over all sets, including the empty one, of disjoint, divergent, one particle irreducible subdiagrams of γ (i.e. diagrams that cannot be disconnected by cutting an arbitrary internal line). The reduced diagram on the RHS is obtained by shrinking each γ_i to a single vertex and finally take the divergent part of the whole sum, with a finite part determined by the normalization conditions at the low energy scale μ . In the framework of dimensional regularization, this operation is elegantly written as a Birkhoff decomposition for a loop in the space complex dimension with values in a group associated to a commutative Hopf algebra (see the work of Connes and Kreimer [2] and [3]).

In the construction of the renormalized theory, the renormalized coupling constant g_r and mass m_r are the parameters of the theory which are directly accessible. They are determined by experiments performed at a low energy scale μ . However, if we consider the same theory but with renormalized parameters determined at a

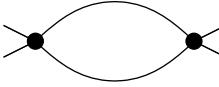


FIGURE 1. One loop diagram for the four point function

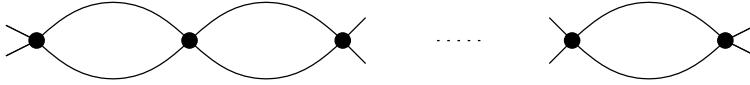


FIGURE 2. Chain of one loop diagrams for the four point function

slightly different scale μ' , then we are performing a mere change in the parametrization of the same theory. Consequently, the renormalized parameters measured at different energy scale are not independent. For instance, the renormalized coupling constant depends on the scale through a differential equation of the type

$$(21) \quad \mu \frac{dg}{d\mu} = \beta(g).$$

The β function can be determined by requiring that the bare parameters, which are the "true" parameters of the theory entering in the path integral, are independent of μ . Accordingly, the low energy renormalized theory has a subtle form of scale invariance inherited from the behavior of the high energy virtual processes. The zeroes of β define the fixed points of the renormalization group and play a fundamental role. In particular, the origin $g = 0$ is a fixed point and it is (UV) attractive if $\beta(g) < 0$. This last case corresponds to the asymptotically free theories for which the coupling constant goes to 0 at high energies. These theories are believed to be consistent non perturbatively, that is, the expansion of the Green's functions can be given a meaning beyond that of an asymptotic series at $g = 0$. In the other case, there are strong arguments against a non perturbative definition.

For instance, at one loop level for the ϕ^4 model in dimension 4, the relation between the bare and the renormalized theory follows from the computation of the diagram given in figure 1,

$$(22) \quad -g_r = -g_0 + \beta g_0^2 \log \frac{\Lambda}{\mu} + O(g_0^2),$$

with β a positive constant. Here, g_r is identified with the four point function evaluated on a specific configuration of momenta of order μ , using the bare theory with a cut-off Λ .

From this relation, one readily extracts the β function as $\beta(g) = \beta g^2$, so that the theory is not asymptotically free. If we now attempt to get a non perturbative information from (22) by summing the geometric series associated to the chain of one loop diagrams depicted on figure 2 we get, as $g_0 > 0$ for the consistency of the path integral,

$$(23) \quad g_r = \frac{g_0}{1 + \beta g_0 \log \frac{\Lambda}{\mu}} \leq \frac{1}{\beta \log \frac{\Lambda}{\mu}},$$

so that the $g_r \rightarrow 0$ as $\Lambda \rightarrow \infty$. This is the triviality problem of ϕ^4 theory. Note that the argument would not apply if β were negative.

To close this preliminary section on renormalization, let us give a brief historical sketch to show how these ideas gradually appeared in physics.

- **Early computations in QED**

It was already noticed by the founding fathers of QED in the thirties that the first non trivial terms of the perturbative expansion are plagued by infinities. However, certain combination of divergent quantities were shown to be finite, thus leading to successfully tested predictions. For instance, the anomalous magnetic moment of the electron, which is nowadays one of the most precisely tested prediction of physics. This set the stage for the general proof of the renormalizability of QED in the fifties by Dyson, Feynman, Schwinger and Tomonaga.

- **Perturbative QED and its β function**

For QED with massless fermions an altered form of scale invariance was discovered: If we measure the low energy electric charge e at various energies μ , we describe the same physics provided we allow for a μ dependent charge e_μ . QED has a positive β function so that the electric charge increases at high energy. This also indicates a breakdown of (perturbative) QED at very high energy, which is no surprise since QED is not a complete theory as it does not take into account nuclear forces.

- **BPHZ (Bogoliubov-Parasiuk-Hepp-Zimmermann) formula**

Systematic combinatorial rules for renormalization were found in the sixties allowing to express the counterterm of a diagrams in terms of counterterms of its subdiagrams and divergent integrals. This is the content of the BPHZ formula and solved the problem of overlapping divergences.

- **Electroweak interactions**

In the late sixties, Glashow, Salam and Weinberg proposed a model based on spontaneously broken non-abelian gauge theories combining electromagnetic and weak interactions. Note that the latter was previously described by the non renormalizable Fermi Lagrangian. This electroweak model was later shown to be renormalizable by 't Hooft and Veltman and the predicted vector bosons were observed at CERN in 1983. Spontaneous symmetry breaking yielding massive vector bosons is achieved through the introduction of the Higgs field, still to be experimentally discovered at the time of writing (mid 2008).

- **Strong interactions**

In the seventies, non-abelian gauge theories with a limited number of fermions were shown to be asymptotically free, meaning that the scale dependent coupling constant is driven to zero at very high energy. This is in agreement with experiments performed at SLAC where quarks inside the hadrons appear to behave as free particles at high energies. Accordingly, the strong force is also described by a renormalizable field theory, Quantum ChromoDynamics (QCD). However, at low energy the QCD coupling constant increases thus impeding a perturbative treatment.

- **Critical phenomena**

Critical phenomena occur in second order phase transitions where the correlation length tends to diverge, so that many degrees of freedom interact. In the seventies, the seminal work of Wilson led to an understanding of critical phenomena, making use of techniques of renormalizable QFT.

- **Effective field theories**

Later on, the Wilsonian viewpoint on renormalization shed a new light on high energy phenomenology. Indeed, most of the models currently used are thought of as effective field theories, yielding accurate physical predictions within a certain range of energy. This is also the point of view adopted in string theory: Whereas the latter is valid at an energy of the order 10^{19} GeV (Planck's energy) or beyond, the standard model coupled to gravity is understood as a low energy effective action valid at a few hundred GeV. This is just as Fermi's theory that can be understood as a low energy effective theory derived from the electroweak model.

- **Constructive QFT**

In fact, the renormalization program in QFT we previously outlined only makes sense at the perturbative level. This means that physical quantities are computed as asymptotic series in a small coupling constant but nothing is said about the convergence of this expansion. Constructive QFT is a branch of mathematical physics aiming at defining the theory using general principles and the information encoded in its perturbative expansion (see [11]). Here too, the Wilsonian point of view led to substantial progress.

- **Stochastic dynamics**

QFT appear to be a powerful tool each time a system is modeled by a partial differential equation involving a random driving force. For example, the random growth of a surface is described by the Kardar-Parisi-Zhang equation whose solution is expressed as a path integral (see [6] for a nice pedagogical introduction). Then, renormalization theory provides some efficient tools to investigate the universal behavior of such a system.

The ubiquity of renormalization theory in the previous situations can be understood as follows. In each case, the physical description is supposed to be valid on a wide range of length scales, including extremely small ones that are not directly probed by experiments. Then, renormalization is the set of methods that prescribes how the parameters of the theory have to be adapted to the degrees of freedom that live on a given scale to take into account the effect of the unobserved degrees of freedom that live on smaller scale (or higher energy).

2.3. Coarse graining for spin systems. The idea that renormalization is an adaptation of the theory to its natural scale mostly follows from the work of Wilson on critical phenomena. To illustrate how the Wilsonian point of view appeared in the study of critical phenomena in statistical mechanics, let us consider a system of spins σ_i located on the sites of a lattice of spacing a . To simplify the discussion much as possible, let us assume that the spins can only take q different values on each site with a Hamiltonian H involving short range interactions (say a nearest neighbor one) that favors spin alignment. For instance, one can take

$$(24) \quad H(\sigma) = -J \sum_{\langle i,j \rangle} \delta_{\sigma_i, \sigma_j},$$

where the sum runs over all pairs of nearest neighbors and J is a positive constant. This is the celebrated Ising model for $q = 2$ and the q -state Potts model in the general case.

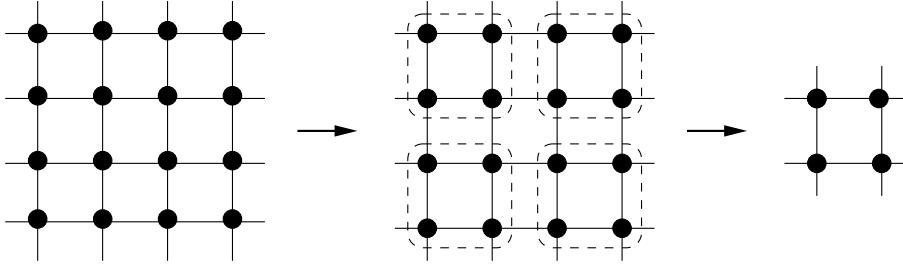


FIGURE 3. Coarse graining

The basic object of interest is the partition function

$$(25) \quad \mathcal{Z} = \sum_{\sigma} e^{-\beta H(\sigma)},$$

where the sum runs over all spin configurations, $\beta = \frac{1}{kT}$, with $k = 1.381 \cdot 10^{-23}$ J·K⁻¹ Boltzmann's constant and T the temperature. For these models, there is a competition between the tendency of the spin to align in a way to minimize their energy and thermal fluctuations which tend to favor random orientations. Some specific systems like the two dimensional Ising model exhibit a phase transition with a high temperature $T > T_c$ disordered phase and a low temperature $T < T_c$ ordered one, with T_c the critical temperature. As the temperature approaches the critical one, the correlation length tend to diverge, thus leading to long range collective phenomena. Moreover, it has been observed that critical systems fall into universality classes made of systems that have a common asymptotic behavior. For instance, in a given universality class, the correlation length may diverge as $T \rightarrow T_c$ with the same exponent,

$$(26) \quad \xi \sim \frac{1}{|T - T_c|^\nu}.$$

Wilson's renormalization group idea amounts to perform the summation in (25) step by step in order to end up with a much simpler description. Specifically, one successively averages over spin configurations on lattices of increasing spacing (see figure 3)

$$(27) \quad \sigma \xrightarrow{\mathcal{T}} \sigma'$$

together with a renormalization group transformation of the Hamiltonian

$$(28) \quad H \xrightarrow{\mathcal{R}} H'.$$

The latter is defined by

$$(29) \quad e^{-\beta H'(\sigma')} = \sum_{T(\sigma)=\sigma'} e^{-\beta H(\sigma)},$$

in such a way that the partition function remains invariant

$$(30) \quad \mathcal{Z} = \sum_{\sigma} e^{-\beta H(\sigma)} = \sum_{\sigma'} e^{-\beta H'(\sigma')}.$$

As we perform the coarse graining transformation from σ to σ' , the lattice size is doubled and it is convenient to also resize the unit of length by a factor of $s = 2$,

so as to stay on the same lattice. Accordingly, the correlation length is divided by s and after sufficiently many iterations of the renormalization group transformation, one can hope to end up with a tractable problem with a correlation length of the order of the lattice spacing. Moreover, the sequence of Hamiltonians defined by

$$(31) \quad H_{n+1} = \mathcal{R}(H_n)$$

undergoes some drastic simplifications if it converges towards a fixed point H^* of \mathcal{R} . Let us further assume that the linearization of \mathcal{R} in the vicinity of H^* can be diagonalized with eigenvalues that are written in the form $\lambda_i = s^{y_i}$ and eigenvectors \mathcal{O}_i . Starting with a Hamiltonian close to the fixed point,

$$(32) \quad H = H^* + \sum t_i \mathcal{O}_i,$$

the iterations of \mathcal{R} lead us to distinguish to three cases.

- Relevant terms associated with $y_i > 0$ tend to grow and one has to fine tune the initial Hamiltonian very close to the critical surface (i.e. the basin of attraction of the fixed point) in order to avoid divergences. Relevant terms govern the deviation from criticality and there is usually one of them proportional to $T - T_c$.
- Irrelevant terms corresponding to $y_i < 0$ die off very quickly and do not play any significant role. Information pertaining to them is erased so that universality is recovered.
- Marginal terms with zero eigenvalue evolve very slowly but after a deeper study they finally end up in one of the two previous classes, apart from some exceptional cases.

As an elementary illustration, let us determine the critical behavior of the correlation length as an inverse power of $|T - T_c|$. We start with a system close to criticality, so that its relevant parameter $t = T - T_c$ with eigenvalue y is very small and its correlation length ξ large. We now iterate n times the renormalization group transformation in such a way that the relevant parameter t' becomes of order unity and the correlation length ξ' close to the lattice spacing,

$$(33) \quad a \sim \xi' \sim \frac{\xi}{s^n} \quad \text{and} \quad 1 \sim t' \sim s^{ny} t.$$

Eliminating s^n we obtain the critical behavior of the correlation,

$$(34) \quad \xi \sim \frac{1}{|T - T_c|^{\frac{1}{y}}},$$

where we have discarded a slowly varying multiplicative constant .

To end this very brief introduction to Wilson's ideas, let us sketch a renormalization group proof of the central limit theorem in probability theory, which has the merit of explaining the universal nature of the Gaussian probability law (see [12] and [13]). Consider $N = 2^n$ independent identically distributed (IID) random variables ξ_i with probability density ρ_0 , assuming that its variance is 1 and a vanishing expectation value. Recall that the central limit theorem states that

$$(35) \quad \zeta_N = \frac{\xi_1 + \dots + \xi_N}{\sqrt{N}}$$

converges in law towards the normal law with density $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ as $N \rightarrow \infty$.

In the spirit of coarse graining, let us compute ζ_N by iterating n times the transformation that associates to two IID random variables ξ' and ξ'' the new one defined by

$$(36) \quad \xi = \frac{\xi' + \xi''}{\sqrt{2}}.$$

The probability density $\mathcal{R}(\rho)$ of ξ is expressed in terms of that of ξ' and ξ'' as

$$(37) \quad [\mathcal{R}(\rho)](x) = \sqrt{2} \int dy \rho(y) \rho(\sqrt{2}x - y),$$

so that the probability density of ζ_N can be written as

$$(38) \quad \rho_n = \mathcal{R}^n(\rho_0).$$

The renormalization map \mathcal{R} admits the Gaußian

$$(39) \quad \rho_*(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

as a fixed point. To investigate the stability of this fixed point, let us linearize \mathcal{R} in the neighborhood of the Gaußian. Writing the perturbation as

$$(40) \quad \rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left(1 + \epsilon(x) \right),$$

we have, up to order ϵ^2 terms,

$$(41) \quad \mathcal{R}[\rho](x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left[1 + \frac{2}{\sqrt{\pi}} \int dy e^{-y^2} \epsilon \left(y + x/\sqrt{2} \right) \right] + O(\epsilon^2).$$

The eigenvectors of the linearization of \mathcal{R} are obtained by taking ϵ proportional to a Hermite polynomial H_n . This is easy to check using their generating function

$$(42) \quad e^{-\frac{s^2}{2} + xs} = \sum_n H_n(x) \frac{s^n}{n!}.$$

The associated eigenvalues are $\lambda_n = 2^{1-\frac{n}{2}}$, so that all the terms involving Hermite polynomials of order $n > 2$ are irrelevant. For $n = 0, 1, 2$, the coefficients of the expansion of ϵ have to vanish in order to make sure that the iteration of \mathcal{R} takes the probability density towards the fixed point. They simply state that the probability distribution has to be normalized ($n = 0$), with vanishing expectation value ($n = 1$) and with variance 1 ($n = 2$). In fact, only $n = 0$ and $n = 1$ are relevant, $n = 2$ is marginal. Imposing these conditions on the probability density we start with is analogous to the fixing the otherwise divergent quantities in QFT by low energy measurements.

2.4. Wilsonian renormalization in QFT. The very same ideas can be applied to QFT, with the partition function replaced by the path integral

$$(43) \quad \mathcal{Z} = \int [D\phi]_\Lambda e^{-S[\phi]}.$$

The integration measure involves an UV cut-off Λ , restricting the integration over fields whose Fourier modes vanish for momenta above Λ . It is analogous to a system defined on a lattice and reflects the fact that rapid variations of the fields on spatial scales below $a \sim \frac{1}{\Lambda}$ do not make sense in this context.

In analogy with spin blocking, one integrates over fast modes in order to obtain a equivalent field theory with a lower cut-off $\Lambda' = \frac{\Lambda}{s}$. To proceed, let us separate the field ϕ into its fast component ϕ'' with momenta between Λ' and Λ and its slow

one ϕ' with momenta below Λ' . By performing the integration over ϕ'' , we obtain the low energy effective action

$$(44) \quad e^{-S'[\phi']} = \int [D\phi'']_{\Lambda', \Lambda} e^{-S[\phi' + \phi'']}. \quad \text{[Equation 44]}$$

The partition function may be computed either with S or with S' ,

$$(45) \quad \mathcal{Z} = \int [D\phi]_{\Lambda} e^{-S[\phi]} = \int [D\phi']_{\Lambda'} e^{-S'[\phi']}. \quad \text{[Equation 45]}$$

The equality between the two terms follows from a substitution of S' in terms of S followed by the change of variables $\phi = \phi' + \phi''$. The remaining integration is cancelled by a normalization factor. These are self-understood constants in all formulae that do not depend on the actions and are defined by imposing that effective actions and partition functions are trivial for trivial actions.

Besides, this remains true if we insert in the path integral any functional of ϕ' , so that one can conclude that low energy physics remains unchanged if we simultaneously lower the cut-off and replace S by S' . All the low energy effects of the physics at the intermediate scale Λ' have been encoded in S' , which may therefore have a rather complicated form even if S is simple.

As for spin systems, one is led to iterate the renormalization group transform to obtain a sequence of effective actions with decreasing cut-offs. In the context of QFT, it more transparent to adopt a continuous formulation, leading to a family of cut-off dependent effective actions S_{Λ} , governed by a differential equation

$$(46) \quad \Lambda \frac{dS_{\Lambda}}{d\Lambda} = \beta(\Lambda, S_{\Lambda}). \quad \text{[Equation 46]}$$

The action with started with is encoded in the initial condition, $S_{\Lambda_0} = S_0$, valid at a very high energy scale Λ_0 . To derive the renormalization group equation (46), let us note that S_{Λ} must be obtained from S_{Λ_0} by integrating over all the modes between Λ and Λ_0 . This is achieved in the path integral approach by

$$(47) \quad e^{-S_{\Lambda}[\phi]} = \int [D\chi] e^{-\frac{1}{2}\chi \cdot A_{\Lambda, \Lambda_0}^{-1} \cdot \chi - S_{\Lambda_0}[\chi + \phi]}, \quad \text{[Equation 47]}$$

where A_{Λ, Λ_0} is a smooth cut-off dependent propagator that implements a smooth analogue of the integration over the fields with Fourier modes between Λ and Λ_0 . For example, we shall often use a cut-off inspired by Schwinger's representation

$$(48) \quad A_{\Lambda, \Lambda_0}(p, q) = \delta(p, q) \int_{\frac{1}{(\Lambda_0)^2}}^{\frac{1}{\Lambda^2}} d\alpha e^{-\alpha p^2}, \quad \text{[Equation 48]}$$

with the mass term relegated in the effective action. Note that Λ plays the role of an IR cut-off whereas Λ_0 is an UV one. By differentiating with respect to Λ , we get Polchinski's equation (see [10]),

$$(49) \quad \Lambda \frac{\partial S_{\Lambda}}{\partial \Lambda} = \frac{1}{2} \int dp dq \Lambda \frac{\partial A_{\Lambda, \Lambda_0}}{\partial \Lambda}(p, q) \left(\frac{\delta^2 S}{\delta \tilde{\phi}(p) \delta \tilde{\phi}(q)} - \frac{\delta S}{\delta \tilde{\phi}(p)} \frac{\delta S}{\delta \tilde{\phi}(q)} \right), \quad \text{[Equation 49]}$$

which is conveniently captured by Feynman diagrams as illustrated on figure 4. Although A_{Λ, Λ_0} depends on both the UV and IR cut-offs, its derivative only depends on Λ .

$$\Lambda \frac{d}{d\Lambda} \text{ (circle)} = \frac{1}{2} \text{ (circle)} \times \text{ (circle)} + \frac{1}{2} \times \text{ (circle with a loop)}$$

FIGURE 4. Diagrammatic form of Polchinski's equation

Here we have use the Fourier transform $\tilde{\phi}(p)$ of the field $\phi(x)$ but an analogous equation can be written in position space,

$$(50) \quad \Lambda \frac{\partial S_\Lambda}{\partial \Lambda} = \frac{1}{2} \int dx dy \Lambda \frac{\partial A_{\Lambda, \Lambda_0}}{\partial \Lambda}(x, y) \left(\frac{\delta^2 S}{\delta \phi(x) \delta \phi(y)} - \frac{\delta S}{\phi(x)} \frac{\delta S}{\delta \phi(y)} \right),$$

with

$$(51) \quad A_{\Lambda, \Lambda_0}(x, y) = \int_{\mathbb{R}^D} \frac{d^D p}{(2\pi)^D} e^{ip \cdot (x-y)} A_{\Lambda, \Lambda_0}(p, -p).$$

To close this terse section on renormalization, let us draw a few high energy physics consequences from the Wilsonian approach. In this context, Λ_0 is the UV cut-off introduced in order to render the Feynman diagrams convergent and S_0 identifies with the bare action. S_Λ is a low energy effective action, which has to be further used in Feynman diagrams having all their internal momenta below a fixed low energy scale Λ . As we try to get rid of Λ_0 by taking it to infinity, some of the parameters entering in S_Λ tend to diverge. The latter have to be fixed by low energy measurement and renormalization simply amounts to choosing a bare action S_0 in such a way that the flow will meet these low energy conditions. Since the limit $\frac{\Lambda_0}{\Lambda} \rightarrow \infty$ corresponds to an iteration of a large number of renormalization group transforms, this is very close to what has been done for critical phenomena. From this point of view, one can understand the ubiquity of renormalizable theories depending on a few parameters as an analogue of universality in statistical mechanics. Whatever S_0 we start with, provided it is natural (i.e. its parameters are numbers $\simeq 1$ times powers of Λ_0 dictated by dimensional analysis), we always end up with renormalizable theories at energies $\Lambda \ll \Lambda_0$. By the same token, universality implies that the renormalized theory is independent of the way we implement the cut-off.

Besides, this analogy can be pushed further: All the renormalizable interactions that appear in high energy physics can be thought of as low energy effective theories derived using a renormalization group flow from a yet unknown theory, valid at a very high energy scale Λ_0 . The latter theory therefore plays in particle physics a role analogue to lattice models in condensed matter physics. As such, it may not even be a field theory and rely on some utterly new ideas, like string theory, spinfoam models or noncommutative geometry. The important point is that it comes equipped with a high energy cutoff Λ_0 , of the order of Planck's length, which is physical and should not be taken to infinity. At energies slightly below Λ_0 , the physics is well captured by a QFT whose field content, action and symmetries are postulated, but should be determined by the unknown theory. Starting with this action, we move towards lower energies Λ using the renormalization group equation and end up with an effective action determined by a few relevant parameters to be experimentally determined. Irrelevant terms corresponding to non renormalizable Lagrangians are also present, though very small, and are innocuous because the cut-off remains finite. Therefore, there is no special problem with gravity, the latter being simply

the leading irrelevant correction to the Standard Model, and the Einstein-Hilbert Lagrangian may be itself corrected by higher order terms. Of course this does by no means solve the problem of quantum gravity, the later presumably amounts to finding the unknown theory we should start with.

One of the main lessons from the Wilsonian approach to renormalization in high energy physics is that the cut-off acquires a physical meaning: It sets the upper limit of validity of the theory. This is in sharp contrast with the point of view adopted a few decades ago where the cut-off was a mere scaffolding needed to build the theory and that has to be removed as soon as possible.

This approach also raises another question: How do we make sure that the relevant parameters, like the masses, are small and not of the order of Λ_0 ? For lattice systems, an operator can tune the bare Hamiltonian sufficiently close to the critical surface so that we end up with finite results even after a large number of iterations. In particle physics, finding such a process remains an open question and is particularly important for a scalar boson like the Higgs field. It seems that such a problem can be solved by introducing supersymmetric particles but at the time of writing (mid 2008) no such particles have been found.

3. Rooted trees and power series of non linear operators

3.1. Rooted trees in numerical analysis. Soon after Connes and Kreimer presented an algebraic framework base on rooted trees for renormalization [1], it was realized by Brouder [14] that perturbative renormalization has some intriguing connection with the algebraic techniques used in the numerical analysis of differential equations.

To begin with, let us expand in powers of $s - s_0$ the solution of the differential equation in \mathbb{R}^n ,

$$(52) \quad \frac{dx}{ds} = X(x), \quad x(s_0) = x_0,$$

as

$$(53) \quad x(s) = \sum_n \frac{(s - s_0)^n}{n!} \frac{d^n x}{ds^n} \Big|_{s=s_0}.$$

To proceed, we compute the successive derivatives of x with respect to s ,

$$(54) \quad \begin{aligned} \frac{dx^i}{ds} &= X^i \\ \frac{d^2x^i}{ds^2} &= \sum_j \frac{\partial X^i}{\partial x^j} X^j \\ \frac{d^3x^i}{ds^3} &= \sum_{j,k} \frac{\partial X^i}{\partial x^j} \frac{\partial X^j}{\partial x^k} X^k + \frac{\partial^2 X^i}{\partial x^j \partial x^k} X^j X^k \\ \frac{d^4x^i}{ds^4} &= \sum_{j,k,l} \frac{\partial X^i}{\partial x^j} \frac{\partial X^j}{\partial x^k} \frac{\partial X^k}{\partial x^l} X^l + 3 \frac{\partial^2 X^i}{\partial x^j \partial x^k} \frac{\partial X^k}{\partial x^l} X^j X^l \\ &\quad + \frac{\partial^3 X^i}{\partial x^j \partial x^k \partial x^l} X^j X^k X^l + \frac{\partial X^i}{\partial x^j} \frac{\partial^2 X^j}{\partial x^k \partial x^l} X^k X^l \end{aligned}$$

All the terms in this expansion turn out to be in one-to-one correspondence with rooted trees with at most four vertices. For example, the second term in the derivative $\frac{d^2x}{ds^2}$ corresponds to the tree depicted in figure 5.

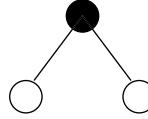


FIGURE 5. A rooted tree with one branching

This is familiar from Runge-Kutta methods in numerical analysis. Indeed, a Runge-Kutta method is defined by a square matrix $(a_{ij})_{1 \leq i, j \leq n}$ and a vector $(b_i)_{1 \leq i \leq n}$ of real numbers, arranged into a table

$$(55) \quad \begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1n} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ c_n & a_{n1} & a_{n2} & \cdots & a_{nn} \\ \hline & b_1 & b_2 & \cdots & b_n \end{array}$$

where $c_i = \sum_j a_{ij}$. Then, the solution of the differential equation

$$(56) \quad \frac{dx}{ds} = X(x), \quad \text{with} \quad x(s_0) = x_0,$$

is given at $s_1 = s_0 + h$ by $x_1 = x_0 + h \sum_{i=1}^n b_i X(y_i)$, where y_i is determined by $y_i = x_0 + h \sum_{j=1}^n a_{ij} X(y_j)$. When $a_{ij} = 0$ for $j > i$, the computation of y_i in terms of trees is straightforward. Then, trees turn out to be useful to compare the result of Runge-Kutta method with the expansion (54).

It is an amazing fact discovered by Butcher that Runge-Kutta methods can be composed [9]. One can perform a first approximate computation using a first method to obtain x as a function of the initial condition x_0 . Then, we can consider x as a new initial condition and proceed to a different approximate computation resulting in x' . This amounts to a single computation of x' with initial condition x_0 , using a third method which is the product of the previous ones.

3.2. Hopf algebra structure. The occurrence of rooted trees in numerical analysis through Runge-Kutta methods and their composition law is conveniently described using the Hopf algebra of rooted trees. To begin with, recall that a tree is a connected graph without cycle and a rooted tree is a tree with a distinguished vertex called the root. A rooted tree is oriented from the root to its terminal vertices, the leaves. Except in the very last paragraph devoted to a bijection between planar diagrams and trees, we always consider isomorphism classes of trees.

Let \mathcal{H}_T be the commutative algebra generated by all (isomorphism classes of) rooted trees. It admits a Hopf algebra structure defined on the generators as follows.

- The coproduct $\Delta : \mathcal{H}_T \rightarrow \mathcal{H}_T \otimes \mathcal{H}_T$ is

$$(57) \quad \Delta(t) = t \otimes 1 + 1 \otimes t + \sum_{c \text{ admissible cut}} P_c(t) \otimes R_c(t)$$

A cut of an edge is admissible cut if any path from any leaf to the root is cut at most once. $R_c(t)$ is the connected component that contain the root after the cut, it is made of vertices located above the cut. $P_c(t)$ is the product of remaining trees corresponding to vertices below the cut with

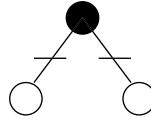


FIGURE 6. An admissible cut on the rooted tree of figure 5

as new root those vertices immediately below the cut. For example, for the simplest tree with one branching, we have

$$(58) \quad \Delta \left(\begin{array}{c} \bullet \\ \circ \circ \end{array} \right) = 1 \otimes \begin{array}{c} \bullet \\ \circ \circ \end{array} + \begin{array}{c} \bullet \\ \circ \circ \end{array} \otimes 1 + 2 \bullet \otimes \begin{array}{c} \bullet \\ \circ \end{array} + \bullet \bullet \otimes \bullet.$$

- The counit $\epsilon : \mathcal{H}_T \rightarrow \mathbb{C}$ is trivial, $\epsilon(t) = 0$ unless $t = 1$.
- The antipode $S : \mathcal{H}_T \rightarrow \mathcal{H}_T$ is defined as a sum over all cuts,

$$(59) \quad S(t) = - \sum_{c \text{ cut}} (-1)^{n_c(t)} \Pi_c(t),$$

with $n_c(t)$ the number of edges of the cut.

Since the Hopf algebra \mathcal{H}_T is commutative, its characters form a group G_T for the convolution product

$$(60) \quad \alpha * \beta = (\alpha \otimes \beta) \circ \Delta$$

with unit ϵ and inverse $\alpha^{-1} = \alpha \circ S$. Recall that a character is an algebra morphism from \mathcal{H}_T into a fixed commutative ring,

$$(61) \quad \begin{cases} \alpha(\lambda a + \mu b) &= \lambda \alpha(a) + \mu \alpha(b) \\ \alpha(ab) &= \alpha(a)\alpha(b), \end{cases}$$

for any $a, b \in \mathcal{H}_T$ and $\lambda, \mu \in \mathbb{C}$. Let us emphasize that the characters must take their values in a commutative ring, in order to check that the product of characters is still a character. Most of the time we work with complex numbers but it may be useful to consider rings of functions.

A character of \mathcal{H}_T is entirely specified by its values on the trees, which may be arbitrary, so that it encodes an infinite sequence of complex numbers. In fact, G_T is a Lie group whose Lie algebra \mathcal{G}_T consists in all infinitesimal characters. The latter are linear maps $\delta : \mathcal{H}_T \rightarrow \mathbb{C}$, satisfying the infinitesimal form of multiplicativity,

$$(62) \quad \delta(ab) = \epsilon(a)\delta(b) + \delta(a)\epsilon(b)$$

for any $a, b \in \mathcal{H}_T$. Alternatively, infinitesimal characters can be defined as derivations with values in the trivial bimodule provided by ϵ .

The convolution exponential $\exp_* : \mathcal{G}_T \rightarrow G_T$, defined by

$$(63) \quad \exp_*(\delta) = \epsilon + \delta + \frac{\delta * \delta}{2} + \cdots + \frac{\delta^{*n}}{n!} + \cdots,$$

and its inverse $\log_* : G_T \rightarrow \mathcal{G}_T$, given by

$$(64) \quad \log_*(\alpha) = \alpha - \epsilon - \frac{(\alpha - \epsilon) * (\alpha - \epsilon)}{2} + \cdots + (-1)^{n-1} \frac{(\alpha - \epsilon)^{*n}}{n} + \cdots,$$

allow to identify \mathcal{G}_T as the infinitesimal form of G_T . Convergence of these series is not an issue, because both δ and $\alpha - \epsilon$ vanish on 1, so that only a finite number of terms survive when evaluated on a tree t . Because infinitesimal characters vanish

on products of trees, they often only involve very simple combinatorial problems. Then, less trivial identities can be obtained by exponentiation. We shall illustrate this fact on several examples in the sequel.

To investigate the the structure of the group of characters G_T , it is convenient to first define a decreasing sequence K_T^n of subgroups of G_T by requiring that elements of K_T^n are characters that vanish on all trees of order $\leq n$. This is a sequence of normal subgroups of G_T , so that one can define an increasing sequence of quotient groups $G_T^n = G_T / K_T^{n+1}$. G_T^n is a finite dimensional nilpotent Lie group with Lie algebra \mathcal{G}_T^n and one recovers G_T and its Lie algebra as the projective limits of the corresponding sequences. Besides, one can construct recursively these groups since G_T^{n+1} is an abelian extension of G_T^n . Recall that given a group G and an abelian group K , a abelian extension of G by K is a new group \tilde{G} defined as follows. As a set $\tilde{G} = G \times K$, but its product law differs from the direct product law by a 2-cocycle θ ,

$$(65) \quad (\alpha, \lambda) \cdot (\beta, \mu) = (\alpha\beta, \lambda + \mu + \theta(\alpha, \beta)),$$

for any $\alpha, \beta \in G$ and $\lambda, \mu \in K$. The associativity of the multiplication on \tilde{G} , is ensured by the cocycle condition

$$(66) \quad \theta(\beta, \gamma) - \theta(\alpha\beta, \gamma) + \theta(\alpha, \beta\gamma) - \theta(\alpha, \beta) = 0$$

for all $\alpha, \beta\gamma \in G$.

For any rooted tree t of order $n + 1$, let us define

$$(67) \quad \theta_t(\alpha, \beta) = \sum_{\substack{\text{c admissible cuts of } t \\ \text{c} \neq \text{empty} \ c \neq \text{full}}} \alpha(t_1) \cdots \alpha(t_n) \beta \left(\frac{t}{t_1 \cdots t_n} \right),$$

where we exclude the empty and the full cuts. Using the reduced coproduct $\Delta' = \Delta - 1 \otimes \text{Id} - \text{Id} \otimes 1$, the cocycle reads,

$$(68) \quad \theta_t(\alpha, \beta) = (\alpha \otimes \beta) \circ \Delta'(t),$$

and thus identifies with Δ' considered as a multiplicative function from the algebra generated by trees of order $n + 1$ to the tensor product of to copies of the algebra generated by trees of order less or equal to n . This allows for a quick check of the cocycle condition since it follows directly from the associativity of Δ .

\mathcal{H}_T is graded by the number of vertices $|t|$ and fits into the general framework of graded and commutative Hopf algebras. Recall that a graded and commutative Hopf algebra is a direct sum $\mathcal{H} = \bigoplus_{n \in \mathbb{N}} \mathcal{H}_n$ with $\mathcal{H}_m \cdot \mathcal{H}_n \subset \mathcal{H}_{mn}$ and $\Delta \mathcal{H}_n \subset \bigoplus_{m \in \mathbb{N}} \mathcal{H}_m \otimes \mathcal{H}_{n-m}$. This framework allows us to implement inductive reasoning. If we further assume that \mathcal{H}_0 is one dimensional, a general structure theorem [15] states that \mathcal{H} is a free commutative algebra so that all the previous results still hold, since most of the proofs rely on inductive reasoning.

3.3. Power series of non linear operators. When solving the differential equation (52) using a sum indexed by rooted trees, we associate each tree with a combination of the derivative of X . In order to generalize this construction, it is fruitful to view trces as indices for formal power series of non linear operators.

To proceed, let us consider any smooth map X from a Banach space \mathcal{E} to itself and let us define X raised to the power of the tree t as

$$(69) \quad X^t = \overrightarrow{\prod}_{v \in t} X^{(n_v)}$$

where n_v is number of edges leaving the vertex v and $X^{(n)}$ is the n^{th} order differential of X . By convention, the tree is oriented from the root to the leaves and we compose the operators following this order. When evaluated at a point $x \in \mathcal{E}$, $X^{(n)}(x)$ is n -linear map and the product refers to the composition of those linear maps from the root to the leaves. As a result, X^t is a smooth map from \mathcal{E} to itself. For instance, we have

$$(70) \quad X^{\bullet} = X' [X' [X]] \quad \text{and} \quad X^{\circlearrowleft} = X'' [X, X].$$

Note that this corresponds exactly to the combinations of partial derivatives used for solving the differential equation (52). Let us note that there is an interesting similarity with the parenthesized words introduced by Kreimer in the very first paper on the Hopf algebras of renormalization [17].

Given an arbitrary character $\alpha \in G_T$, let us define the formal power series

$$(71) \quad \Psi_\alpha(X) = \sum_t \alpha(t) \frac{X^t}{S_t},$$

with S_t the symmetry factor of the tree, which is the cardinal of its automorphism group. This has to be understood as a formal power series of smooth maps, i.e. when evaluated on any $x \in \mathcal{E}$ it yields a formal power series of elements of \mathcal{E} , which is ordered in increasing number of vertices of the trees. Let us also note that the series always starts with the empty tree which is nothing but the unit of the algebra \mathcal{H}_T . By convention, any operator raised to the power of the empty tree is the identity, $X^\emptyset = \text{id}$.

If we now consider another $\beta \in G_T$, then we can compose $\Psi_\alpha(X)$ with $\psi_\beta(X)$, Taylor expand and reorganize the series in terms of rooted trees. It turns out that the result is naturally expressed using the product law in G_T ,

$$(72) \quad \underbrace{\Psi_\alpha(X) \circ \Psi_\beta(X)}_{\text{composition}} = \underbrace{\Psi_{\beta * \alpha}(X)}_{\text{convolution}}.$$

This composition law is identical to the composition law of B-series [16], underlying Butcher's work on the algebraic structure of Runge-Kutta methods. To prove this relation, we write $\delta = \log_*(\alpha)$ and consider both sides of as a function of $s \in \mathbb{R}$ for $\alpha_s = \exp_*(s\delta)$. Then, we derive both sides with respect to s and check that the derivatives agree using simple combinatorial identities. Let us note that if X is linear, all derivatives $X^{(n)}$ vanish for $n > 1$ so that only trees without branchings appear. In this case eq. (72) amounts to the ordinary multiplication of power series of a linear operator.

It is also convenient to allow the trees to carry some decorations on their vertices. Then, all the structure remains similar, with $R_c(t)$ and $P_c(t)$ carrying decorations inherited from t in the definition of the coproduct. The main interest of decorated rooted trees lies in the fact that they allow us to define power series of several non linear operators, each operator being labeled by its decoration.

The most useful series is the geometric one, defined by summing over all trees with a weight one,

$$(73) \quad (\text{id} - X)^{-1} = \sum_t \frac{X^t}{S_t}.$$

As suggested by the notation, it yields an inverse of $\text{id} - X$ for the composition law. Accordingly, it provides us with a perturbative solution of the fixed point equation

$$(74) \quad x = x_0 + X(x)$$

as a sum over trees

$$(75) \quad x = (\text{id} - X)^{-1}(x_0) = \sum_t \frac{X^t}{S_t}(x_0).$$

Moreover, let us note that the tree like structure of the geometric expansion proves to be extremely helpful in dealing with convergence issues. To proceed, let us assume that the n^{th} order differential is a bounded n -linear map satisfying the estimate $\|X^{(n)}(x_0)\| \leq a_n$, with a_n a sequence of real numbers such that

$$(76) \quad f(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n$$

converges. Then, the series

$$(77) \quad x = (\text{id} - hX)^{-1}(x_0) = \sum_t h^{|t|} \frac{X^t}{S_t}(x_0),$$

is bounded in norm by the series associated to the equation $z = hf(z)$, which is known to be an analytic function of h at the origin by the implicit function theorem.

Beyond the geometric series, it may also be useful to consider the binomial series. For any character α and complex number a , let us define

$$(78) \quad (\alpha)^{*a} = \epsilon + a(\alpha - \epsilon) + \frac{a(a-1)}{2}(\alpha - \epsilon)*(\alpha - \epsilon) + \dots + \frac{a(a-1)\dots(a-n+1)}{n!}(\alpha - \epsilon)^{*n} + \dots$$

Obviously, the binomial series fulfills all the identities we expect $(\alpha)^{*a} * (\alpha)^{*b} = (\alpha)^{*a+b}$ and $((\alpha)^{*a})^{*b} = (\alpha)^{*ab}$ at the notable exception of $(\alpha)^{*a} * (\beta)^{*a} = (\alpha * \beta)^{*a}$, which does not hold since the convolution product is not commutative.

If we take α to be the character that takes the value 1 on the tree with one vertex and vanishes on any other non trivial tree, then the iterations of the coproduct are easy to compute,

$$(79) \quad (\alpha)^{*a}(t) = \sum_{n=d_t}^{|t|} N(n, t) \frac{a(a-1)\dots(a-n+1)}{n!},$$

where $N(n, t)$ is the number of surjective maps from the vertices of t to $\{1, \dots, n\}$, strictly increasing from the root to the leaves. d_t is the depth of the tree which is the length of the longest path from the root to the leaves.

Accordingly, one can apply the binomial series to a non linear operator,

$$(80) \quad (\text{id} + X)^a = \sum_t \sum_{n=d_t}^{|t|} N(n, t) \frac{a(a-1)\dots(a-n+1)}{n!} \frac{X^t}{S_t}$$

the coefficients in from of $\frac{X^t}{S_t}$ can be considered as tree-lie generalizations of the binomial coefficients, to which they reduce on trees without branchings. For instance, up to order 4 we have

$$(81) \quad (\text{id} + X)^a = 1 + aX + \frac{a(a-1)}{2} \left(X'[X] + \frac{1}{2} X''[X, X] \right) + \frac{a(a-1)(a-2)}{6} \left(X'[X'[X]] + X''[X, X] \right) + \dots$$

This allows, for instance, to compute the square root of a diffeomorphism

$$(82) \quad \begin{aligned} \sqrt{\text{id} + X} &= \frac{1}{2}X - \frac{1}{8}X'[X] + \frac{1}{16}X'[X'[X]] - \frac{5}{128}X'[X'[X'[X]]] \\ &+ \frac{1}{128}X''[X, X'[X]] - \frac{1}{2 \cdot 64}X'[X''[X, X]] + \frac{1}{6 \cdot 64}X'''[X, X, X] + \dots \end{aligned}$$

which fulfills $\sqrt{\text{id} + X} \circ \sqrt{\text{id} + X} = \text{id} + X$ up to terms of fifth order in X .

Let us also introduce the number $\tilde{N}(n, t)$ of surjective maps from the vertices of t to $\{1, \dots, n\}$, increasing (but not necessarily in a strict manner) from the root to the leaves. Then, by applying the binomial series to the geometric one, we get

$$(83) \quad (\text{id} - X)^{-a} = ((\text{id} - X)^{-1})^a = \sum_t \sum_{n=1}^{|t|} \tilde{N}(n, t) \frac{a(a-1) \cdots (a-n+1)}{n!} \frac{X^t}{S_t}.$$

Then, by comparing with the expression of $(\text{id} - X)^{-a}$ derived from (80), we get

$$(84) \quad \begin{aligned} \sum_{n=1}^{|t|} \tilde{N}(n, t) \frac{a(a-1) \cdots (a-n+1)}{n!} \\ = (-1)^{|t|} \sum_{n=d_t}^{|t|} N(n, t) (-1)^n \frac{a(a+1) \cdots (a+n+1)}{n!} \end{aligned}$$

By evaluating for various values of a , we get various relations between $N(n, t)$ and $\tilde{N}(n, t)$. For example, if $a = -1$ we obtain the sum rule

$$(85) \quad 1 = (-1)^{|t|} \sum_{n=d_t}^{|t|} (-1)^n N(n, t),$$

which is a convenient way to check the computation of $N(n, t)$. For $a = -2$, we get

$$(86) \quad 2 + \tilde{N}(2, t) = \sum_{n=d_t}^{|t|} (-1)^{n+|t|} (n+1) N(t, n),$$

which counts the number of terms in the coproduct.

The geometric series and its cousins can be used in two complementary ways. In many cases, the equations we encounter in physics can be written as fixed point equations of the type (74), so that rooted trees allow for an efficient determination of their perturbative solutions. For instance, this is the case for the self-consistency condition derived in mean field theory or the Schwinger-Dyson equations in QFT. In mathematics, this applies to the solution of an ordinary differential equation

written in integral form, as we shall work out in section 3.5, or to the Lagrange inversion formula.

Alternatively, one can be given a perturbative expansion exhibiting a natural tree like structure. If one can find an operator X such that this expansion can be written as a geometric series (73), then this expansion is known to be a solution of the fixed point equation (74). Accordingly, the solution of this equation provides a good candidate for the sum of the tree like expansion. We illustrate this procedure in the next section.

3.4. Wigner's semi-circle law from rooted trees. Let us give a very simple illustration of way one can use rooted trees to recover Wigner's semi-circle law of an ensemble of Gaussian Hermitian $N \times N$ matrices. This technique is nothing but a rooted tree formalisation of the treatment that can be found in [6].

We are interested in the large N behavior of the density of eigenvalues, defined as

$$(87) \quad \rho(\lambda) = \lim_{N \rightarrow \infty} \int [DM] e^{-N \text{Tr} V(M)} \sum_i \frac{\delta(\lambda - \lambda_i)}{N}$$

with λ_i denoting the eigenvalues of M and $V(M) = \frac{1}{2}M^2$. Also,

$$[DM] = \prod_i dM_{ii} \prod_{i < j} d\text{Re}M_{ij} d\text{Im}M_{ij}$$

is the standard Lebesgue measure on the space of $N \times N$ Hermitian matrices and δ the Dirac distribution.

Using $\delta(x) = \frac{1}{\pi} \text{Im} \left(\frac{1}{x - i\epsilon} \right)$ with $\epsilon \rightarrow 0^+$, we can rewrite ρ as

$$(88) \quad \rho(\lambda) = \lim_{N \rightarrow \infty} \frac{1}{N\pi} \text{Im} \left\{ \int [DM] \text{Tr} \left(\frac{1}{\lambda - i\epsilon - M} \right) e^{-N \text{Tr} V(M)} \right\}.$$

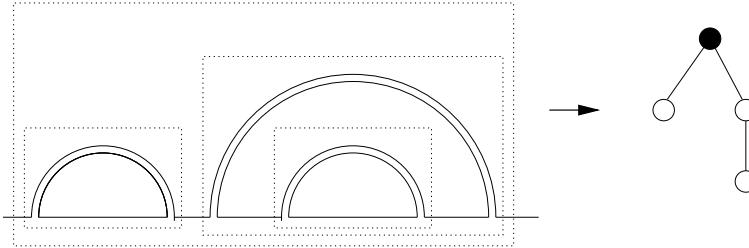
Introducing $z = \lambda - i\epsilon$ to simplify the notation, we expand the resolvent as a geometrical series

$$(89) \quad \frac{1}{z - M} = \frac{1}{z} + \frac{1}{z} M \frac{1}{z} + \frac{1}{z} M \frac{1}{z} M \frac{1}{z} + \dots$$

The expectation value $G(z)$ of $\frac{1}{z - M}$ is easily computed using Wick's theorem with the Gaussian matrix integral. It leads to Feynman rules with a single line propagator for $\frac{1}{z}$ a double line propagator $\frac{1}{N}$ for the propagation of the matrix field and a vertex set to 1 for the interaction of 2 single lines with a double line. Then, each face of the resulting diagram yields a factor N because of the summation over matrix indices. Thus, the leading contribution in the large N limit is given by those diagrams that have as many faces as possible for a fixed number of propagator. These diagrams are nothing but the celebrated planar diagrams.

The planar Feynman diagrams that contribute to $G(z)$ are in one to one correspondence with planar rooted trees (i.e. rooted trees counted as different if they differ by a permutation of their vertices). The bijection is obtained by drawing boxes that are either disjoint or nested around the matrix propagators, with the diagram always in the outermost box, even if there is no matrix propagator. This outermost box defines the root and the tree is drawn from the root to the leaves by following the natural hierachal structure of the inclusion of the boxes.

Isomorphic planar trees correspond to Feynman diagram having the same contribution to $G(z)$. Moreover, recall that a rooted tree is an isomorphism class of

FIGURE 7. A planar diagram contributing to $G(z)$

planar rooted tree and the sum over all representatives of the class of t equals $\frac{X^t(0)}{S(t)}$, with

$$(90) \quad X(x) = \frac{1}{z - x}.$$

Note that the n^{th} order derivative of X generates a factor $n!$ at each vertex with n outgoing edges. When multiplied together along the tree and divided by the symmetry factor, the resulting number counts the cardinality of the isomorphism class of planar trees corresponding to t .

Therefore, $G(z)$ is the sum of the non linear geometric series

$$(91) \quad G(z) = \sum_t \frac{X^t(0)}{S(t)} I_N,$$

with I_N the $N \times N$ identity matrix. The sum of the geometric series is the solution x of the quadratic fixed point equation $X(x) = x$ with leading term $\frac{1}{z}$, so that

$$(92) \quad G(z) = \frac{z - \sqrt{z^2 - 4}}{2} I_N.$$

and we recover Wigner's semicircle law as

$$(93) \quad \rho(\lambda) = \frac{1}{2\pi} \sqrt{4 - \lambda^2},$$

with support in $[-2, +2]$.

If we expand $G(z)$ in powers of $\frac{1}{z}$, we recover the ubiquitous Catalan numbers $C_n = \frac{(2n)!}{(n!)^2(n+1)}$,

$$(94) \quad G(z) = \sum_{n=0}^{\infty} C_n \frac{1}{z^{2n+1}} I_N.$$

This is not surprising since the Catalan number C_n count, amongst many other things [18], the number of non crossing partitions of a set of $2n$ elements into blocks of 2. More sophisticated examples of summation over rooted trees applied to random matrices are described in the lecture by P. Di Francesco [19]. In particular, it is shown how some similar techniques can be used for a quartic potential $V(M) = \frac{1}{2}M^2 + \frac{g}{4}M^4$.

3.5. Perturbative solution of differential equation. To solve perturbatively a time dependent non linear differential equation

$$(95) \quad \frac{dx}{ds} = X_s(x),$$

with boundary condition $x(s_0) = x_0$, it is convenient to write it in integral form,

$$(96) \quad x(s) = x_0 + \int_{s_0}^s ds' X_{s'}(x(s')).$$

This is a fixed point equation for the integral non-linear operator given by $x(s) \mapsto \int_{s_0}^s ds' X(x(s'))$ with the function $s \mapsto x(s)$ as the unknown. Using the geometric series, it solution is expanded over trees as

$$(97) \quad x(s) = \sum_t \frac{1}{S_t} \int_{I_{s,s_0}^t} d^{|t|} s \overrightarrow{\prod}_{v \in t} X_{s_v}^{(n_v)}(x_0).$$

To explain the meaning of these terms, recall that any tree is orientated from the root to the leaves. Then, associate to each vertex v a real variable s_v in such a way that $s_0 \leq s_v \leq s_{v+}$ if v_+ is a vertex immediately above v and $s_{v+} = s$ for the root. This defines what we call a "treeplex" $I_{s,s_0}^t \subset \mathbb{R}^{|t|}$, that reduces to a ordinary simplex if t is a tree without branching. Then compose all the multilinear operators $X_{s_v}^{(n_v)}(x_0)$ from the root to the leaves and integrate over the variables s_v . Obviously, for linear X we recover the ordinary time ordered products, with an integral over a simplex instead of a treeplex.

If we write the solution of (95), as $\Omega_{s,s_0}(x_0) = \sum_t \Omega_{s,s_0}^t(x_0)$, the semigroup property holds

$$(98) \quad \Omega_{s_2,s_1} \circ \Omega_{s_2,s_1} = \sum_{\text{admissible cuts } c} \Omega_{s_2,s_1}^{t_0} \circ_c [\Omega_{s_1,s_0}^{t_1}, \dots, \Omega_{s_1,s_0}^{t_n}],$$

where t_0 is the part of t that contains its root after the cut and t_1, \dots, t_n are the branches we cut. The symbol \circ means composition followed by Taylor expansion and \circ_c is the same operation, to be performed at each vertex immediately above the cut. This relation is similar to the definition of the coproduct and an analogous equation exists for the antipode, related to the inversion of Ω_{s,s_0} .

Comparing both sides of (98) yields a geometric interpretation of the coproduct: Cut I_{s_2,s_0}^t by a plane $s_{\text{root}} = s_1$, so that it is written as a disjoint union

$$(99) \quad I_{s_2,s_0}^t = \bigcup_{c \text{ admissible cut}} \mathfrak{S}_c (I_{s_1,s_0}^{t_1} \times \dots \times I_{s_1,s_0}^{t_n} \times I_{s_2,s_1}^{t_0})$$

with $P_c(t) = t_1 \dots t_n$ and $R_c(t) = t_0$. \mathfrak{S}_c a suitable permutation of the labels preserving the ordering of the tree. For example, for the simplest tree with one branching, this is illustrated on figure 8 and the polytopes in this decomposition are in direct correspondence with the terms in (58).

For a time independent equation, the integral over I_{s,s_0}^t factorizes and defines the tree factorial

$$(100) \quad \int_{I_{s,s_0}^t} d^{|t|} s = \frac{1}{t!}.$$

The latter can be computed recursively: If t is a tree obtained by grafting the trees t_1, \dots, t_n to the root, then

$$(101) \quad t! = (|t_1| + \dots + |t_n| + 1) t_1! \dots t_n!.$$

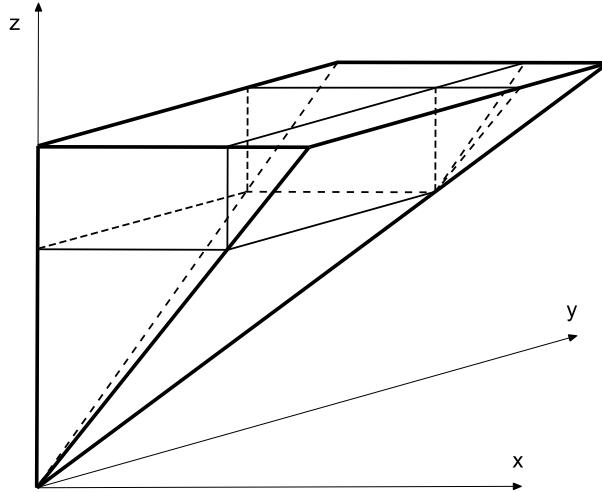


FIGURE 8. Geometric interpretation of the coproduct

For a tree without branching we recover the ordinary factorial $t! = |t|!$.

Accordingly, the map Ω_{s,s_0} is a power series of the non linear operator X

$$(102) \quad \Omega_{s,s_0} = \sum_t \frac{1}{S_t} \frac{(s-s_0)^{|t|}}{t!} X^t,$$

which is a non linear analogue of the exponential. The exponential character can be written as

$$(103) \quad \exp_* \{(s-s_0)\delta\} (t) = \frac{(s-s_0)^{|t|}}{t!}$$

with $\delta \in \mathcal{G}_T$ the infinitesimal character that takes the value 1 on the simplest non empty tree and vanishes otherwise. It satisfies the linear equation

$$(104) \quad \frac{d}{ds} \exp_* \{(s-s_0)\delta\} = \delta * \exp_* \{(s-s_0)\delta\} = \exp_* \{(s-s_0)\delta\} * \delta.$$

Thus, we have traded a non linear equation for a linear one, at the price of working with power series of non linear operators.

The relation (103) provides a combinatorial definition of the tree factorial, in analogy with what we did for the binomial series (78). Because δ is an infinitesimal character, there are tremendous simplifications in the convolution powers of δ evaluated on the iterated coproduct and only one term survives

$$(105) \quad \delta^{*n}(t) = N(n, t)$$

if $n = |t|$ and vanishes otherwise. Therefore, $\frac{1}{t!} = \frac{N(|t|, t)}{|t|!}$ so that

$$(106) \quad t! = \frac{\# \{\text{labelings of the set of vertices in } \{1, \dots, n\}\}}{\# \{\text{labelings of the set of vertices decreasing from the root to the leaves}\}}.$$

Turning back to numerical analysis, the Runge-Kutta method (55) one naturally associates a character α constructed as follows. Decorate the vertices of the tree t by indices in $\{1, 2, \dots, n\}$ and associate each edge from i to j with the matrix

element a_{ij} . Multiply all these matrix elements together with a factor b_i associated to the root and sum over all indices. Then, the approximate solution after one step reads

$$(107) \quad x_1 = \sum_t \frac{h^{|t|}}{S_t} \alpha(t) X^t(x_0).$$

It is a method of order n if α coincides with the exponential on all trees of order at most n . This defines another formal power series of non linear operators and the composition law of Runge-Kutta methods is represented by the convolution of their characters.

4. Renormalization, effective actions and Feynman diagrams

4.1. Renormalization for a differential equation. Before dealing with path integrals and Feynman diagrams, it is fruitful to first investigate the perturbative renormalization for an ordinary differential equation [20]. Recall that the differential renormalization group equation (46) reads

$$(108) \quad \Lambda \frac{dS_\Lambda}{d\Lambda} = \beta(\Lambda, S_\Lambda),$$

with boundary condition $S_{\Lambda_0} = S_0$ at a very high energy scale Λ_0 . In physics, S_Λ is an effective action that encodes the description of the physical system under scrutiny, valid up to an energy scale Λ , while S_0 is the bare action which agrees with the effective action at the cut-off scale Λ_0 . Ignoring temporarily the path integral interpretation of the renormalization group equation, let us simply consider a general equation of the form (108) for a variable S in a Banach space \mathcal{E} . Following the general interpretation of renormalization, we are interested in the low energy behavior of S_Λ for $\Lambda \ll \Lambda_0$.

Because all the physical quantities can be measured in unit of mass², we further assume that the space \mathcal{E} comes equipped with an action of the dilatation group \mathbb{R}^{*+} . This action is written in exponential form as $e^{s\mathcal{D}}$, with \mathcal{D} a linear operator. It is often convenient to assume that \mathcal{D} is diagonal, with eigenvalues corresponding to the canonical dimensions. Dimensional analysis is assumed to be compatible with the renormalisation group equation,

$$(109) \quad \beta(e^s \Lambda, e^{s\mathcal{D}} S) = e^{s\mathcal{D}} \beta(\Lambda, S).$$

To take dimensional analysis into account, it is convenient to introduce the dimensionless variables $s = \log(\Lambda/\Lambda_{\text{ref}})$ with Λ_{ref} a fixed reference scale, $u(s) = \Lambda^{-\mathcal{D}} \cdot S_\Lambda$ and $X(u) = \Lambda_{\text{ref}}^{-\mathcal{D}} \cdot \beta(\Lambda_{\text{ref}}, S_\Lambda)$. In terms of these new variables, the renormalization group differential equation takes the simpler form

$$(110) \quad \frac{du}{ds} = -\mathcal{D}u + X(u),$$

with boundary condition $u(s_0) = u_0$.

The perturbative solution of this equation is easily obtained by writing it in integral form

$$(111) \quad u(s) = e^{-(s-s_0)\mathcal{D}} u_0 + \int_{s_0}^s ds' e^{(s-s')\mathcal{D}} X(u(s')).$$

²In the system of units $\hbar = c = 1$ which we adopt here.

Making use of the geometric series for the integral operator given by $\mathcal{F} \cdot u(s) = \int_{s_0}^s ds' e^{(s-s')\mathcal{D}} X(u(s'))$,

$$(112) \quad u = (\text{id} - \mathcal{F})^{-1}(u_0)$$

we expand $u(s)$ as a sum over trees, using the following rules to compute the contribution of each tree t . First, orient all the edges from the root to the leaves and then associate

- variables s_v to the vertices that leave in the the treeplex $I_{s_0, s}^t$.
- linear operators $e^{-(s_{v+} - s_{v-})\mathcal{D}}$ to edges from v_+ to v_- (recall that edges are oriented from the root to the leaves);
- multilinear operators $X^{(n_v)}(e^{-(s_v - s_0)\mathcal{D}} u_0)$ to vertices with n_v outgoing edges.

Finally, compose from the root to the leaves all operators associated to edges and vertices, multiply on the left by $e^{-(s - s_{\text{root}})\delta}$ and integrate over the variables associated to the vertices. This is very similar to the expansion we encountered in the previous section for a time dependent equation, at the notable exception of the insertion of the of the linear operators $e^{-(s_{v+} - s_{v-})\mathcal{D}}$ along the edges.

To simplify the notations, let us abbreviate as $X_{s_v}^{(n)}$ the n -linear map obtained by differentiating X n times at $e^{-(s_v - s_0)\mathcal{D}} u_0$. For example, the linear tree of order two contributes as

$$(113) \quad \int_{s_0}^s ds_2 \int_{s_0}^{s_2} ds_1 \quad e^{-(s - s_2)\mathcal{D}} X'_{s_2} \left[e^{-(s_2 - s_1)\mathcal{D}} X_{s_1} \right],$$

whereas the order 3 tree with one branching yields

$$(114) \quad \frac{1}{2} \int_{s_0}^s ds_3 \int_{s_0}^{s_3} ds_2 \int_{s_0}^{s_3} ds_1 \quad e^{-(s - s_3)\mathcal{D}} X''_{s_3} \left[e^{-(s_3 - s_2)\mathcal{D}} X_{s_2}, e^{-(s_3 - s_1)\mathcal{D}} X_{s_1} \right]$$

In the simplest case of a linear X , the contributions of all branched trees vanish and we recover Duhamel's formula (or in the physics terminology, the Schwinger-de Witt expansion) of the exponential of $-\mathcal{D} + X$ in powers of X ,

In analogy with perturbative renormalization in QFT, let us try to get rid of the cut-off by taking the limit $s_0 \rightarrow \infty$. If we discard all the non linear terms, the analysis us easy to perform:

- variables u^{irr} that belong to the negative eigenspaces of \mathcal{D} (irrelevant variables) tend to 0;
- variables u^{mar} that belong to the zero eigenspace of \mathcal{D} (marginal variables) remain constant;
- variables u^{rel} that belong to the positive eigenspaces of \mathcal{D} (relevant variables) grow to infinity.

In perturbation theory, this picture is slightly modified by the addition of the non linear terms. Marginal variables have a polynomial dependence on s , analogous to the $\log \Lambda$ terms in QFT. Moreover, irrelevant terms at a given order depend on the marginal and relevant ones at a lower order, so that they also experience divergences. This phenomena is analogous to the subdivergences of QFT.

The solution of the problem of UV divergencies is natural from the Wilsonian viewpoint: Instead of imposing the boundary condition for the divergent variables at the very high energy s_0 tending to infinity, impose the latter at a fixed low energy s_r . This corresponds to a determination of the relevant (and marginal) parameters

by a measurement at low energy. To implement this change of boundary condition, let us denote by \mathcal{P} the projector onto the relevant and marginal variables and by \mathcal{F}' the non linear operator

$$(115) \quad \mathcal{F}'[u](s) = \mathcal{P} \int_{s_r}^{s_0} ds' e^{(s-s')\mathcal{D}} X(u(s')).$$

Therefore, the renormalized theory, expressed as a function of the renormalized parameters (irrelevant parameters at s_0 , relevant and marginal ones at s_r), is obtained as a fixed point for the renormalized operator $\mathcal{F} + \mathcal{F}'$. To compare the renormalized and unrenormalized expansions, let us write

$$(116) \quad [\text{id} - (\mathcal{F} + \mathcal{F}')]^{-1} = [\text{id} - \mathcal{F}]^{-1} \circ [\text{id} - \mathcal{F}' \circ (\text{id} - \mathcal{F})^{-1}]^{-1}.$$

In the QFT language, the LHS is the renormalized theory while the RHS is the unrenormalized one composed with the counterterms. The latter are the boundary conditions on the relevant and marginal parameters to be imposed at the high energy s_0 in such a way that these parameters match their fixed values at low energy s_r . At the level of the tree expansions, the expansion of the renormalized theory can be pictured as a sum over trees with both white vertices for \mathcal{F} and black ones for \mathcal{F}' , while the unrenormalized theory only involves white vertices. Counterterms only involve trees with a black root, followed by black and white vertices that correspond to the renormalization of the subdivergences. This is the analogue of the BPHZ formula for a differential equation.

In the renormalized expansion, all the s_0 dependence has dropped from the divergent exponentials and their integrals, so that the contribution of each tree admits a finite limit as $s_0 \rightarrow \infty$. Moreover, the initial conditions u_0^{irr} for the irrelevant parameters all enter through the combination $e^{-\mathcal{D}(s-s_0)} u_0^{\text{irr}}$, which goes to 0 as $s_0 \rightarrow \infty$. In this way we recover universality: The renormalized theory does not depend on the choice of the irrelevant parameters at high energy.

When formulated in terms of differential equations, the renormalization procedure is general enough to encompass Polchinski's equation (49), as far as perturbation is concerned. Solving the latter using rooted trees yields a smart proof of the renormalizability of ϕ^4 theory in $D = 4$ that does not refer to Feynman diagrams [21], in addition to the original proof by Polchinski [10].

Though rooted trees are a convenient device that often leads to convergent expansions, it is fair to say that this is seldom the case for differential equation with boundary condition at ∞ . To illustrate the issue raised by non perturbative computations within the differential renormalization group approach, let us consider a very simple example with one marginal and one irrelevant coupling,

$$(117) \quad \begin{cases} \frac{dg}{ds} = \beta g^2 & (\mathcal{D} = 0), \\ \frac{du}{ds} = u + \gamma g & (\mathcal{D} = -1). \end{cases}$$

This system can be solved using the perturbative techniques we have presented. However, because of the extreme simplicity of these equations, it is much more convenient to determine an exact solution. Let us first investigate the case of the marginal coupling constant g . In the unrenormalized case, the boundary condition is imposed at very high energy, so that the unrenormalized solution and its

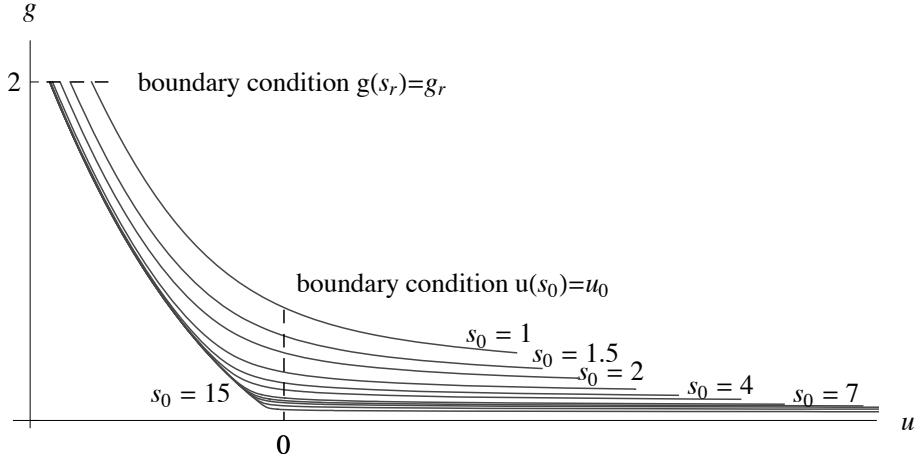


FIGURE 9. Convergence towards the renormalized theory for the differential equation (117)

expansion read

$$(118) \quad g(s) = \frac{g_0}{1 - \beta g_0(s - s_0)} = g_0 \sum_{n=0}^{\infty} (\beta g_0)^n (s - s_0)^n.$$

In terms of rooted trees, the LHS is the sum over rooted trees with vertices having only at most two outgoing edges (since the operator $g \mapsto \beta g^2$ is quadratic) and a weight corresponding to the tree factorial.

As expected, the coefficients of the expansion, which play the role of the Feynman diagrams, diverge as $s_0 \rightarrow \infty$ at fixed g_0 . Perturbative renormalization is achieved by imposing boundary conditions at a fixed low energy scale $g(s_r) = g_r$, so that the renormalization coupling constant reads

$$(119) \quad g(s) = \frac{g_r}{1 - \beta g_r(s - s_r)} = g_r \sum_{n=0}^{\infty} (\beta g_r)^n (s - s_r)^n.$$

The transition from the unrenormalized coupling to renormalized one is achieved by choosing g_0 in such a way that g obeys the new boundary condition $g(s_r) = g_r$, which is easily seen to be

$$(120) \quad g_0 = \frac{g_r}{1 - \beta g_r(s_0 - s_r)} = g_r \sum_{n=0}^{\infty} (\beta g_r)^n (s_r - s_0)^n.$$

Then, as we substitute this expression of g_0 in (118) we recover (119). At the perturbative level, this substitution is best seen using rooted trees and illustrates the way BPHZ renormalization must be performed with the trees.

Once the renormalized coupling constant (119) is known, the solution of the second equation follows from the method of variation of constants,

$$(121) \quad u(s) = u_0 e^{s - s_0} + \int_{s_0}^s ds' \frac{\gamma g_r e^{-(s' - s)}}{1 - \beta g_r(s' - s_r)}.$$

First, let us notice that the u is independent of u_0 as $s_0 \rightarrow \infty$. This is universality in this simple context: The renormalized theory does not depend on the irrelevant terms at high energy. At $s_r = 0$ and $s_0 \rightarrow \infty$, the exact solution can be expanded as

$$(122) \quad u(g_r) = -\gamma \int_0^\infty ds e^{-s} \frac{g_r}{1 - g_r \beta s} \simeq - \sum_{n=0}^{\infty} \beta^n n! (g_r)^{n+1}.$$

Though it provides accurate approximations of u for small g_r , it is a divergent power series, Borel summable only if $\beta < 0$. Nevertheless, if $\beta < 0$ the coupling $g(s)$ remains bounded and u is well defined as a functional of $g : s \mapsto g(s)$,

$$(123) \quad u[g] = \int_0^\infty ds e^{-s} g(s)$$

This is a particularly simple example of an effective expansion [11], whose counterpart in QFT is the fact that an asymptotically free theory, characterized by $\beta < 0$, may be well defined when expanded in a scale dependent coupling instead $g(s)$ of a single renormalized coupling g_r . In case $\beta > 0$, (121) can still be given a mathematical meaning by deforming the contour of integration in the s plane. However, in this case there are extra ambiguities related to the monodromy around the pole so that the construction of the theory requires extra information not contained in the perturbation theory. Nevertheless, the theory may lose its physical meaning because it is no longer real.

The occurrence of the previous type of $n!$ behavior that prevents a naive summation of the perturbative solution with boundary conditions at infinity is due to the phenomenon of resonances. Consider a general renormalization group equation

$$(124) \quad \frac{du}{dt} = -\mathcal{D}u + X(u)$$

and assume that \mathcal{D} has been diagonalized in a basis (e_i) with eigenvalues λ_i . We further assume that X can be written as a polynomial (or an analytic function)

$$(125) \quad X(u) = \sum_{i,j_1 \dots j_n} e_i X_{j_1 \dots j_n}^i (u_{j_1})^{m_1} \dots (u_{j_n})^{m_n}.$$

A monomial in (125) is resonant if there is a non trivial relation of the type

$$(126) \quad \lambda_i = m_1 \lambda_{j_1} + \dots + m_n \lambda_{j_n},$$

with $m_k \in \mathbb{N}$.

As we try to solve the renormalization group equation perturbatively in u , the Poincaré-Dulac theorem states that, by a series of successive changes of variables of the type $v = u + F(u)$ with F having at least a second order term, the renormalization group equation takes the simpler form

$$(127) \quad \frac{dv}{dt} = -\mathcal{D}v + Y(v)$$

with only resonant monomials in Y . Because the resonance condition (126) can be formulated as

$$(128) \quad e^{-s\mathcal{D}} \circ Y = Y \circ e^{-s\mathcal{D}},$$

all the exponentials but the the first one cancel in the perturbative expansion of the solution given in (112). Therefore, the iterated integral along the tree generates

a power law s^n and leads to a factorial behavior $n!$ when integrated together with the exponential along the root.

Let us conclude this section by two remarks that may be the subject of some future work.

- It could be interesting to derive the normal form of the QFT renormalization group equation (49), with a view towards a non perturbative analysis. This could be done along the lines initiated in [22] for a toy model.
- Reduction to normal forms is a non linear generalization of the Lie algebra decomposition into commuting semi-simple and nilpotent elements [23]. As such, it could be conveniently described in the algebraic framework presented in the sections devoted to rooted trees and Feynman diagrams.

4.2. Hopf algebra of Feynman diagrams and effective actions. In the view of the algebraic structure we have presented in the previous sections, let us investigate Wilsonian renormalization in QFT. The basic renormalization group transform is the integration over fast modes that induces the transformation of effective actions $S \rightarrow S'$,

$$(129) \quad e^{S'[\phi]} = \mathcal{N} \int [D\chi] e^{-\frac{1}{2} \chi \cdot A^{-1} \cdot \chi} e^{S[\phi + \chi]}.$$

A is a suitable propagator that implements the integration over momenta lying between Λ and Λ_0 . In what follows, we consider A to be arbitrary propagator, with suitable UV and IR cut-offs. \mathcal{N} is a normalization factor, independent of S and ϕ , that ensures that $S' = 0$ for $S = 0$. In the sequel, we do not display such normalization factors anymore. Note that we have used S instead of $-S$ in the exponential to avoid some irrelevant signs.

The renormalization group transformation of effective actions is best formulated using the background field method. The rationale of the background field method is to split the field into a fixed classical field ϕ and a quantum field χ over which we integrate, so that we end up with an effective action function for the classical field ϕ . In our setting, the fast modes play the role of the quantum field whereas the low energy ones remain as an argument of the low energy effective action. Using this technique, S' is expanded over connected vacuum Feynman diagrams (i.e. diagrams without external leg) as

$$(130) \quad S'[\phi] = \sum_{\gamma} \frac{A^{\gamma}(S)}{S_{\gamma}} [\phi].$$

The sum starts with the trivial empty diagram whose contribution is $S[\phi]$. For other diagrams, $A^{\gamma}(S)$ is computed using the background field technique Feynman rules: Propagators are determined by A and n -valent vertices are n^{th} order functional differentials of S at ϕ . As is always the case for diagrammatic expansions, one has to further divide by the symmetry factor S_{γ} , which is nothing but the cardinal of the automorphism group of the diagram. All the dependence on ϕ is encoded in the vertices so that there is no need for external sources and we restrict ourselves to vacuum diagrams. Note that there is no restriction on S since the convergence of each Feynman diagram is ensured by the cut-off. Strictly speaking, this is true only for a sharp cut-off, when considering a cut-off with an exponential decrease we have to assume that the momentum dependence of the functional derivatives of S does not grow faster than a polynomial.

To illustrate the background field technique, let us give a simple example with a finite dimensional integral and a positive definite symmetric matrix A_{ij} as a propagator.

$$\text{Diagram: } \begin{array}{c} \text{A horizontal line with two vertices, each connected to a vertical line.} \\ \rightarrow \frac{1}{12} \sum_{\substack{i_1, i_2, i_3 \\ j_1, j_2, j_3}} \frac{\partial^3 S}{\partial \chi^{i_1} \partial \chi^{i_2} \partial \chi^{i_3}}(\phi) A_{i_1, j_1} A_{i_2, j_2} A_{i_3, j_3} \frac{\partial^3 S}{\partial \chi^{j_1} \partial \chi^{j_2} \partial \chi^{j_3}}(\phi) \end{array}$$

In the spirit of Wilsonian renormalization, one has to pursue this procedure and consider S' to be a starting point for a new integration over yet slower modes. This is the reason why we are forced to deal with actions of a very general form. Indeed, even if S is a simple ϕ^4 interaction, this is no longer the case for S' that contains arbitrarily complicated vertices. To handle this type of computation, it is necessary to modify the Hopf algebra defined by Connes and Kreimer as follows. Let \mathcal{H}_F be the free commutative algebra generated by all connected vacuum Feynman diagrams with vertices of arbitrary valence, including univalent and bivalent ones. It is equipped with a Hopf algebra structure with trivial counit and coproduct

$$(131) \quad \Delta(\gamma) = \gamma \otimes 1 + 1 \otimes \gamma + \sum_{\substack{\gamma_1, \dots, \gamma_n \\ \text{disjoint subdiagrams}}} \gamma_1 \cdots \gamma_n \otimes \gamma / (\gamma_1 \cdots \gamma_n),$$

where the sum runs over all possible mutually disjoint connected subdiagrams. In our setting, a subdiagram is a subset lines and vertices related to them. The reduced diagram $\Gamma / (\gamma_1 \cdots \gamma_n)$ is obtained by shrinking each γ_i to a single vertex in Γ . Obviously, \mathcal{H}_F is graded by the number of internal lines I but not by the number of vertices. However, the number of loops L also provides a grading because of the identity $L = I - V + 1$ for a connected diagram, but this grading does not yield a connected algebra since all the trees have grading 0. To these gradings is associated the two parameter group of automorphisms

$$(132) \quad \varphi_{a,b}(\gamma) = a^{I_\gamma} b^{L_\gamma} \gamma,$$

with I_γ (resp. L_γ) being the number of internal lines (resp. loops) of γ . Because of the grading, one can apply the general theory of graded and commutative Hopf algebras: The antipode may be defined recursively and characters of \mathcal{H}_F form a group G_F , with a Lie algebra \mathcal{G}_F of infinitesimal characters. Note that elements of both G_F and \mathcal{G}_F are uniquely defined by their values on the diagrams, which may assume any values in a given commutative algebra. As for trees, the exponential defines a map $\exp_* : \mathcal{G}_F \rightarrow G_F$ by

$$(133) \quad \exp_*(\delta) = \epsilon + \delta + \frac{\delta * \delta}{2} + \cdots + \frac{\delta^{*n}}{n!} + \cdots,$$

which is a solution of the differential equation

$$(134) \quad \frac{d}{ds} \exp_*(s\delta) = \delta * \exp_*(s\delta) = \exp_*(s\delta) * \delta$$

reducing to the unit of G_T at $s = 0$.

The elements of G_F can be considered as weights of the for the Feynman diagrams in the expansion of the transformation of effective actions $\Psi_\alpha^A : S \rightarrow S'$

$$(135) \quad S' = \Psi_\alpha^A(S) = \sum_\gamma \alpha(\gamma) \frac{A^\gamma(S)}{S_\gamma}$$

Given any two characters $\alpha, \beta \in G_F$, there is a composition law analogous to the one we encounter for power series of non linear operators in (72),

$$(136) \quad \underbrace{\Psi_\beta^A \circ \Psi_\alpha^A}_{\text{composition}} = \underbrace{\Psi_{\alpha * \beta}^A}_{\text{convolution}} .$$

This composition law describes how the weight of the Feynman diagrams entering in an iteration of two effective action computations can be obtained by inserting diagrams with weight α into diagrams with weight β . To prove this relation, we introduce $\delta = \log_*(\alpha)$ and $\alpha_s = \exp_*(s\delta)$. At first order in s , this relation does not involve multiple insertions of diagrams and is easily verified by simple combinatorics. Then, we integrate the differential equation to recover (136). In a certain sense, it is a Wilsonian analogue of the morphism from the group of graphs into the diffeomorphisms of the coupling constant presented in [3].

In physics, there are many characters that appear naturally. First, one may consider α to be the character that selects 1PI diagrams and β the one that selects the trees. Then, $\alpha * \beta$ takes the value one on all diagrams and the composition law (136) simply states that the one recovers the complete effective action by first computing 1PI diagrams and then inserting into trees. As another simple example, let us consider the character $\alpha_{s,h}(\gamma) = s^{I_\gamma} h^{L_\gamma}$, so that the corresponding effective action transformation $S \rightarrow S'$ can be expressed as

$$(137) \quad S'[\phi] = h \log \left\{ \int [D\chi] e^{-\frac{1}{2}\chi \cdot (hsA)^{-1} \cdot \chi} e^{\frac{S[\phi+\chi]}{h}} \right\} .$$

As we expand S' in term of Feynman diagrams, we find a weight hs for every internal line and h^{-1} for every vertex together with a global factor of h . Thus a connected diagram with I_γ internal lines, V_γ vertices and $L_\gamma = I_\gamma - V_\gamma + 1$ loops is weighted by $s^{I_\gamma} h^{I_\gamma - V_\gamma + 1} = s^{I_\gamma} h^{L_\gamma}$. Other characters of interest are the polynomials in N that arise in the $O(N)$ invariant models.

To handle more complicated situations, it may necessary to decorate the Feynman diagrams. For instance, if we want to perform the second effective action with a different propagator B , then we have to decorate the internal lines with an index that differentiates A and B . With decorations on the internal lines, the Hopf algebra structure remains similar with the subgraph that we extract in (131) carrying the decorations. Then, we can perform a first effective action computation with propagator A and character α , followed by a second one with B and β . The results amounts to a single computation with decorated internal lines that selects the type of propagators we use. Then, the analogue of (136) reads

$$(138) \quad \Psi_\beta^B \circ \Psi_\alpha^A = \Psi_{\alpha * \beta}^{A,B},$$

Both α and β are characters of the Hopf algebra with two decorations on the internal lines, but α and (resp. β) vanish on diagrams that do not contain A (resp. B). The previous relation is a generalization of the convolution of Gaußian integrals

$$(139) \quad \begin{aligned} \int [D\chi] \int [D\xi] \exp \left\{ -\frac{1}{2}\chi A^{-1} \chi - \frac{1}{2}\xi B^{-1} \xi + S[\phi + \chi + \xi] \right\} &= \\ \int [D\zeta] \exp \left\{ -\frac{1}{2}\zeta (A + B)^{-1} \zeta + S[\phi + \zeta] \right\}, \end{aligned}$$

which we obtain from (138) when the characters α and β only take the values 0 or 1, depending on whether the diagrams only involve the corresponding propagator. Obviously, this construction can be generalized to an arbitrary number of propagators so that we can iterate the effective action construction. Unfortunately, this construction is not sufficient to handle Polchinski's equation since the latter involve an infinite number of iterations corresponding to the differential equation. Moreover, the various types interactions (relevant, marginal and irrelevant) play fundamentally different roles so that it is necessary to distinguish them in the Feynman expansion. We postpone these questions to the section 4.4, devoted to the iterative solution of Polchinski's equation. Before we deal with the latter, let us investigate some algebraic properties of the \mathcal{H}_F and illustrate them in the next section on some properties of the Tutte polynomial.

First of all, let us notice that there are only two diagrams of the degree one (i.e. with one internal line): The bridge with two vertices and the self-loop with a single vertex. Let us define two infinitesimal characters: δ_{tree} takes the value one on the bridge and vanishes otherwise, and δ_{loop} takes the value one on the self-loop and vanishes otherwise,

$$(140) \quad \begin{cases} \delta_{\text{tree}}(\text{---}) = 1 \\ \delta_{\text{tree}}(\gamma) = 0 \text{ if } \gamma \neq \text{---} \end{cases} \quad \text{and} \quad \begin{cases} \delta_{\text{loop}}(\text{○}) = 1 \\ \delta_{\text{loop}}(\gamma) = 0 \text{ if } \gamma \neq \text{○} \end{cases}$$

By iterating the coproduct, it is easy to see that

$$(141) \quad \exp_* \{a \delta_{\text{tree}} + b \delta_{\text{loop}}\}(\gamma) = a^{I_\gamma} b^{L_\gamma},$$

where I_γ is the number of internal lines and L_γ the number of loops. By coassociativity, the iteration can be performed either on the left or on the right. If we perform it on the left, then we contract the internal lines and we get a factor of a if the line is a self-loop or b if it is not. When performed on the right we delete internal lines and we also get factors of a and b depending on whether the line is a bridge (i.e. a line whose removal disconnects the diagram) or not. Note that all contraction/deletion schemes yield the same result and can be mixed together: At one step we can delete and then contract at another.

Any infinitesimal character $\delta \in \mathcal{G}_F$ defines two derivations of \mathcal{H}_F that correspond to the infinitesimal form of right and left multiplication by $\exp_*(s\delta)$ on functions over G_T at first order in s ,

$$(142) \quad \begin{cases} f \triangleleft \delta = (\delta \otimes \text{Id}) \circ \Delta(f), \\ \delta \triangleright f = (\text{Id} \otimes \delta) \circ \Delta(f), \end{cases}$$

for any $f \in \mathcal{H}_T$. Obviously, left and right action for different characters commute. In case of the infinitesimal characters δ_{tree} and δ_{loop} , these operations have natural interpretations in terms of deletions and contraction of edges. Recall that a self-loop is an edge with both ends linked to the same vertex and a bridge is an edge whose removal disconnects the diagram. Note that an edge connecting a single vertex to the rest of the diagram is considered as a bridge.

- $\delta_{\text{loop}} \triangleright \gamma$ is a sum over all the diagrams obtained from γ by cutting a line which is not a bridge.
- $\delta_{\text{tree}} \triangleright \gamma$ is a sum over all the diagrams obtained from γ by cutting a bridge.
- $\gamma \triangleleft \delta_{\text{loop}}$ is a sum over all the diagrams obtained from γ by contracting a self-loop with one edge.

- $\gamma \triangleleft \delta_{\text{tree}}$ is a sum over all the diagrams obtained from γ by contracting a line which is not a self-loop.

As will be illustrated in the next section, these rules allow to translate simple diagrammatic manipulations into differential equations.

4.3. An interlude on the Tutte polynomial. The Tutte polynomial is a two variable polynomial associated to graphs that plays a fundamental role in graph theory and statistical mechanics (see [24] for a recent review). In what follows, we shall illustrate its universality property in graph theory and its relation to the Potts model partition function using the techniques we introduced in the previous section.

The Tutte polynomial is a polynomial in x and y which is multiplicative on disjoint unions and that can be defined using the rank-nullity formula on connected graphs,

$$(143) \quad P_{\text{Tutte}}(x, y, \gamma) = \sum_{A \subset E} (y-1)^{n(A)} (x-1)^{r(E)-r(A)},$$

where the sum runs over all subsets of the set of edges E of γ . In the QFT language, the nullity and the rank of a connected diagram can be expressed in terms of the number of internal lines and loops

$$(144) \quad \begin{cases} n(\gamma) &= L_\gamma \\ r(\gamma) &= I_\gamma - L_\gamma \end{cases}$$

Moreover, subsets of the sets of edges are in one-to-one correspondence with possibly disconnected subdiagrams, so that the Tutte polynomial can be written as the evaluation at $s = 1$ of the character

$$(145) \quad \alpha = \exp_* s \{ \delta_{\text{tree}} + (y-1) \delta_{\text{loop}} \} * \exp_* s \{ (x-1) \delta_{\text{tree}} + \delta_{\text{loop}} \},$$

where we have used (144) and (141) to write the rank and nullity using the exponentials. From this expression, it is obvious that the Tutte polynomial character can be expressed as

$$(146) \quad \alpha = \overbrace{\exp_* s \{ \delta_{\text{tree}} + (y-1) \delta_{\text{loop}} \} * \exp_* s \{ -\delta_{\text{tree}} + \delta_{\text{loop}} \}}^{\text{Tutte polynomial character at } x=0 \text{ on the subgraphs}} * \overbrace{\exp_* s \{ \delta_{\text{tree}} - \delta_{\text{loop}} \} * \exp_* s \{ (x-1) \delta_{\text{tree}} + \delta_{\text{loop}} \}}_{\text{Tutte polynomial character at } y=0 \text{ on the reduced graphs}},$$

which is nothing but the convolution identity of Kook, Reiner and Stanton (see [25]).

Then, using (134), α can be characterized as the solution of the differential equation

$$(147) \quad \frac{d\alpha}{ds} = x \alpha * \delta_{\text{tree}} + y \delta_{\text{loop}} * \alpha + [\delta_{\text{tree}}, \alpha]_* - [\delta_{\text{loop}}, \alpha]_*$$

that reduces to the trivial character at $s = 0$.

This characterisation of the Tutte polynomial as the solution of a differential equation can be used to derive its universality property. The latter is formulated as follows. Consider a four variables graph polynomial $Q(x, y, a, b, \gamma)$ that is multiplicative on disjoint unions and one vertex unions (i.e. diagrams joined at a single vertex), and has the following behavior under contraction and deletion of an arbitrary internal line:

- if γ' is obtained by deleting in γ a bridge, then

$$(148) \quad Q(x, y, a, b, \gamma) = x Q(x, y, a, b, \gamma'),$$

- if γ' is obtained by contracting in γ a self-loop, then

$$(149) \quad Q(x, y, a, b, \gamma) = y Q(x, y, a, b, \gamma'),$$

- if γ' is obtained by deleting in γ a regular line (i.e. a line that is neither a bridge nor a self-loop) and γ'' obtained by contracting the same line, then

$$(150) \quad Q(x, y, a, b, \gamma) = a Q(x, y, a, b, \gamma') + b Q(x, y, a, b, \gamma'').$$

Then, the character defined by

$$(151) \quad \beta(\gamma) = s^{I_\gamma} Q(x, y, a, b, \gamma)$$

obeys the differential equation

$$(152) \quad \frac{d\beta}{ds} = x \beta * \delta_{\text{tree}} + y \delta_{\text{loop}} * \beta + a [\delta_{\text{tree}}, \beta]_* - b [\delta_{\text{loop}}, \beta]_*.$$

This follows from the discussion at the end of last section, where we have formulated algebraically the contraction/deletion of internal lines. Note that the contraction of a self-loop cancels in the fourth term while the deletion of a bridge cancel in the third term because of the multiplicativity on one vertex unions. Acting with the automorphism $\varphi_{a^{-1}, b^{-1}}$ defined in (132), we obtain the Tutte polynomial differential equation (147) with modified parameters $\frac{x}{a}$ and $\frac{y}{b}$

$$(153) \quad \begin{aligned} \frac{d}{ds} \varphi_{a^{-1}, b^{-1}}(\beta) &= \frac{x}{a} \varphi_{a^{-1}, b^{-1}}(\beta) * \delta_{\text{tree}} + \frac{y}{b} \delta_{\text{loop}} * \varphi_{a^{-1}, b^{-1}}(\beta) \\ &+ [\delta_{\text{tree}}, \varphi_{a^{-1}, b^{-1}}(\beta)]_* - [\delta_{\text{loop}}, \varphi_{a^{-1}, b^{-1}}(\beta)]_*. \end{aligned}$$

Accordingly, $\beta = \varphi_{a, b}(\alpha)$, with α the corresponding Tutte polynomial character, so that

$$(154) \quad Q(x, y, a, b, \gamma) = a^{r(\gamma)} b^{n(\gamma)} P_{\text{Tutte}}\left(\frac{x}{a}, \frac{y}{b}, \gamma\right).$$

This is the universality property of the Tutte polynomial: Any graph polynomial multiplicative on disjoint and one vertex unions satisfying some simple deletion/contraction identities can be expressed using the Tutte polynomial.

As a second illustration, let us derive the relation of the Tutte polynomial with the q -state Potts model partition function using Feynman diagrams and effective actions. To this aim, we write the Tutte polynomial as a product of two characters (145) that we interpret as a composition of two background field effective action computation using (136). To simplify the notations, we set $u = x - 1$, $v = y - 1$ and $q = uv$, so that the factorization of the Tutte polynomial character reads, at $s = 1$,

$$(155) \quad \alpha = \exp_* \{ \delta_{\text{tree}} + v \delta_{\text{loop}} \} * \exp_* \{ u \delta_{\text{tree}} + \delta_{\text{loop}} \}.$$

Besides, for our purpose it is only necessary to consider effective actions for a single 0 dimensional field which is nothing but a real variable. The first character is $\exp_* \{ \delta_{\text{tree}} + v \delta_{\text{loop}} \}$ and its induces a transformation of effective actions $S \mapsto S'$ that weights the Feynman diagrams with the character $\alpha(\gamma) = v^{L_\gamma}$. Then, using (137) we get

$$(156) \quad S'[\phi] = v \log \left\{ \int [D\chi] e^{-\frac{1}{2v} \chi^2} e^{\frac{S[\phi+\chi]}{v}} \right\}.$$

Next, we consider S' as a starting point for a new effective action computation $S' \rightarrow S''$ with a weight $\exp_* \{u \delta_{\text{tree}} + \delta_{\text{loop}}\}$, which amounts to weight internal lines by u and loops by $\frac{1}{u}$, so that

$$(157) \quad S''[\psi] = \frac{1}{u} \log \left\{ \int [D\phi] e^{-\frac{1}{2}\phi^2} e^{uS[\psi+\phi]} \right\}.$$

As we replace S' by its expression in terms of S in the last equation, we get

$$(158) \quad S''[\psi] = \frac{1}{u} \log \left\{ \int [D\phi] e^{-\frac{1}{2}\phi^2} \left(\int [D\chi] e^{-\frac{1}{2v}\chi^2} e^{\frac{S[\psi+\phi+\chi]}{v}} \right)^q \right\}.$$

According to the general rule of effective action composition (136), this amounts to a single effective action computation with Feynman diagrams weighted by their Tutte polynomial. In particular, if we take a universal action like $S[\phi] = \lambda e^\phi$ and evaluate S'' at $\psi = 0$, we get the Tutte polynomial multiplied by a power of λ

$$(159) \quad \frac{1}{u} \log \left\{ \int [D\phi] e^{-\frac{1}{2}\phi^2} \left(\int [D\chi] e^{-\frac{1}{2v}\chi^2} e^{\frac{\lambda}{v} e^{\phi+\chi}} \right)^q \right\} = \sum_{\gamma \text{ conn.}} \frac{\lambda^{V_\gamma}}{S_\gamma} P_{\text{Tutte}}(u+1, v+1, \gamma),$$

with V_γ the number of vertices of γ . The choice of the universal action $S[\phi] = \lambda e^\phi$ is motivated by the fact that it generates all diagrams with a factor 1. If restrict ourselves to a monomial like $\lambda^{\frac{\phi^n}{n!}}$, then we generate only the diagrams with n -valent vertices.

Unfortunately, it is not yet possible to interpret (159) as a generating function for the Tutte polynomial, since this would require us to be able to identify the contribution of each diagram in the perturbative expansion. To proceed, we have to write the integrals over ϕ and χ as a single integral which is a perturbation of a Gaussian integral. This can be easily done when q is an integer by introducing q independent fields χ

$$(160) \quad \left(\int [D\chi] e^{-\frac{1}{2v}\chi^2} e^{\frac{1}{v}S[\phi+\chi]} \right)^q = \int \prod_{1 \leq i \leq q} [D\chi_i] e^{-\frac{1}{2v} \sum_i (\chi_i)^2} e^{\frac{1}{v} \sum_i S[\chi_i + \phi]}$$

Then, the integral over ϕ and χ in (159) reads

$$(161) \quad \int [D\phi] \int \prod_{1 \leq i \leq q} [D\chi_i] e^{-\frac{1}{2}\phi^2} e^{-\frac{1}{2v} \sum_i (\chi_i)^2} e^{\frac{1}{v} \sum_i S[\chi_i + \phi]}$$

It is convenient to trade χ_i for $\xi_i = \chi_i + \phi$ so that the integral over ϕ is Gaussian

$$(162) \quad \int [D\phi] \int \prod_{1 \leq i \leq q} [D\xi_i] e^{-\frac{1}{2}(1+u)\phi^2 + \frac{\phi}{v} \sum_i \xi_i} \times e^{-\frac{1}{2v} \sum_i \xi_i^2} \times e^{\frac{1}{v} \sum_i S[\xi_i]}.$$

Performing the Gaussian integral over ϕ and discarding an overall constant which is absorbed in the normalization, we get

$$(163) \quad \int \prod_{1 \leq i \leq q} [D\xi_i] e^{-\frac{1}{2v} \{ (\sum_i (\xi_i)^2 - \frac{1}{v(1+u)} (\sum_i \xi_i)^2) \}} \times e^{\frac{1}{v} \sum_i S[\xi_i]}.$$

This can be written as a perturbed Gaussian integral over a multiplet of fields $\xi = (\xi_i)$, so that (159) reads

$$(164) \quad \frac{1}{u} \log \left\{ \int [D\xi] e^{-\frac{1}{2}\xi \cdot A^{-1}\xi} e^{V(\xi)} \right\} = \sum_{\gamma \text{ connected}} \frac{\lambda^{V_\gamma}}{S_\gamma} P_{\text{Tutte}}(u+1, v+1, \gamma),$$

with a $q \times q$ propagator

$$(165) \quad A = v + M$$

where M is the $q \times q$ matrix whose entries are all equal to 1 and an interaction

$$(166) \quad V(\xi) = \frac{\lambda}{v} \sum_i \exp(\xi_i).$$

The expansion of (164) over Feynman diagrams can be considered as a generating functions for the Tutte polynomials, since each diagram is weighted by $P_{\text{Tutte}}(u + 1, v + 1, \gamma)$. Moreover, the Feynman rules for this expansion are as follows,

- each vertex contributes to a factor $\frac{\lambda}{v}$ and is equipped with an index i taking q values, because of the q independent fields in the interaction;
- each propagator contributes a factor of $1 + v$ if it connects vertices with the same index and a factor of 1 otherwise,
- sum over all the indices.

Up to a factor, this is nothing but the partition function $Z(\beta, J, \gamma)$ for the q -state Potts model defined on the graph γ . Recall that the latter is a lattice model that is defined on an arbitrary graph γ by assigning spins σ_v taking q values to the vertices of the diagram. The energy of a configuration of spins $\sigma = (\sigma_v)$ is defined as a sum over all the edges of the diagram

$$(167) \quad H(\sigma) = -J \sum_{\text{edges } e} \delta_{\sigma_v, \sigma'_v},$$

with v and v' the vertices linked by e and $J > 0$ a constant that favors spin alignment. Its partition function is defined as

$$(168) \quad Z(\beta, J, \gamma) = \sum_{\sigma} e^{-\beta H(\sigma)}$$

and it follows from the previous discussion that it is proportional to the Tutte polynomial

$$(169) \quad Z(\beta, J, \gamma) = uv^{V_{\gamma}} P_{\text{Tutte}}(u + 1, v + 1, \gamma)$$

with $v = e^{-\beta J} - 1$.

4.4. Iterative solution of Polchinski's equation. Let us now come back to Polchinski's equation (49) that describes the scale evolution of the effective action S_{Λ} . Recall that S_{Λ} is obtained from a bare action S_{Λ_0} valid at a very high energy scale by integrating Fourier modes between Λ and Λ_0 . A large class of propagators that implement this integration and preserve both translational and rotational symmetry can be constructed in momentum space as

$$(170) \quad A_{\Lambda, \Lambda_0}(p, q) = \left\{ \chi \left(\frac{p^2}{\Lambda_0^2} \right) - \chi \left(\frac{p^2}{\Lambda^2} \right) \right\} \times \frac{\delta(p + q)}{p^2}.$$

The cut-off function χ a smooth positive decreasing function with values close to 1 on $[0, 1]$ and close to 0 on $[1, +\infty[$. A typical example is $\chi(s) = e^{-s}$ that yields a propagator

$$(171) \quad A_{\Lambda, \Lambda_0}(p, q) = \delta(p + q) \int_{\frac{1}{(\Lambda_0)^2}}^{\frac{1}{\Lambda^2}} d\alpha e^{-\alpha p^2},$$

that leads to the convenient expression of Feynman diagrams based on Symanzik's polynomials for ϕ^4 theory (see [8] and the remark at the end of this section). Note that we have not included the mass term in the propagator to simplify the forthcoming dimensional analysis. Here we treat the mass term as an interaction, which does not modify the UV behavior of the theory. The derivative is independent of Λ_0

$$(172) \quad \Lambda \frac{\partial A_{\Lambda, \Lambda_0}}{\partial \Lambda}(p, q) = \frac{2}{\Lambda^2} \chi' \left(\frac{p^2}{\Lambda^2} \right) \times \delta(p + q),$$

and implements the integration over an infinitesimal shell. With such a propagator, Polchinski's equation can be cast in the general form (108) with the non linear operator

$$(173) \quad \beta(\Lambda, S) = \int dp dq \Lambda \frac{\partial A_{\Lambda, \Lambda_0}}{\partial \Lambda}(p, q) \left(\frac{\delta^2 S}{\delta \tilde{\phi}(p) \delta \tilde{\phi}(q)} - \frac{\delta S}{\delta \tilde{\phi}(p)} \frac{\delta S}{\delta \tilde{\phi}(q)} \right)$$

In high energy physics with $c = \hbar = 1$, all quantities are measured in terms of masses so that if we scale all masses by a factor e^s , the action transforms as $S \rightarrow S' = e^{sD} S$ with $S' [e^{s(D/2+1)} \tilde{\phi}'] = S[\tilde{\phi}]$ and $\tilde{\phi}'(p) = \tilde{\phi}(e^s p)$ in a space-time of dimension D . Then, it is easy to check that

$$(174) \quad \beta(e^s \Lambda, e^{sD} S) = e^{sD} \beta(\Lambda, S),$$

so that all the results of section 4.1 apply. In particular, a translation invariant term of the form

$$(175) \quad \int dp_1 \cdots dp_N \delta(p_1 + \cdots + p_N) \Gamma(p_1, \dots, p_N) \tilde{\phi}(p_1) \cdots \tilde{\phi}(p_N)$$

where Γ is homogenous of degree n , scales with a factor e^{sd} with

$$(176) \quad d = D - N \left(\frac{D}{2} - 1 \right) - n.$$

Therefore, we recover the fact that the mass term is always relevant and that for ϕ^4 theory in dimension 4 the coupling constant is marginal and all terms with a high number of fields and/or high order derivative interactions are irrelevant.

Accordingly, Polchinski's equation can be written in integral form,

$$(177) \quad S_\Lambda = S_{\Lambda_0} + \int_{\Lambda_0}^{\Lambda} d\Lambda' \beta(\Lambda', S_{\Lambda'})$$

with $\dot{A} = \frac{\partial A}{\partial \Lambda}$ and

$$(178) \quad \beta(\Lambda, S) = \frac{1}{2} \int dp dq \dot{A}_\Lambda(p, q) \left(\frac{\delta^2 S}{\delta \tilde{\phi}(p) \delta \tilde{\phi}(q)} - \frac{\delta S}{\delta \tilde{\phi}(p)} \frac{\delta S}{\delta \tilde{\phi}(q)} \right)$$

Then, its iterative solution is expanded over ordered Feynman diagrams. The latter are Feynman diagrams with a hierarchy of internal lines defined by drawing boxes on the diagram such that the boxes are either disjoint or nested and such that a box differ from the next one it is included in by an internal line. Alternatively, the order on the internal lines can be pictured by a rooted tree t drawn on the set of internal lines of the diagram. The contribution $A_{\Lambda, \Lambda_0}^{\gamma_{\text{ord}}}(S_0)$ to the effective action S_Λ is computed as follows,

- associate an intermediate cut-off Λ_i to any internal line;

- compute the value of the Feynman diagram using the background field Feynman rules with vertices derived from S_0 and propagators \dot{A}_{Λ_i} ;
- integrate the variables Λ_i over the treeplex I_{Λ, Λ_0}^t , defined in section 3.5.

Accordingly, the low energy effective action reads

$$(179) \quad S_{\Lambda} = \sum_{\gamma_{\text{ord}}} \frac{A_{\Lambda, \Lambda_0}^{\gamma_{\text{ord}}}(S_0)}{S_{\gamma_{\text{ord}}}}$$

with $S_{\gamma_{\text{ord}}}$ the cardinal of the automorphisms group of the diagram that preserve the ordering.

Ordered diagrams generate a commutative Hopf algebra $\mathcal{H}_F^{\text{ord}}$ with coproduct similar to that of \mathcal{H}_F given in (131), except that it has to preserve the box structure. This coproduct is very similar to the rooted tree coproduct, as it reproduces on the diagram the admissible cuts of the tree as an extraction of disjoint boxes. For example, with Δ' the non trivial part of Δ :

$$\begin{aligned} \Delta' \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) &= \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \otimes \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) + \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \otimes \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right), \\ \Delta' \left(\begin{array}{c} \bullet \\ \text{---} \\ \text{---} \end{array} \right) &= \bullet \otimes \text{---} + \text{---} \otimes \bullet, \\ \Delta' \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) &= 2 \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) \otimes \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right) + \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right)^2 \otimes \left(\begin{array}{c} \text{---} \\ \text{---} \end{array} \right), \\ \Delta' \left(\begin{array}{c} \bullet \\ \text{---} \\ \text{---} \end{array} \right) &= 2 \bullet \otimes \text{---} + \bullet^2 \otimes \bullet. \end{aligned}$$

Because the effective action could as well be computed in a single step using the propagator A_{Λ, Λ_0} ,

$$(180) \quad \sum_{\gamma} \frac{A_{\Lambda, \Lambda_0}^{\gamma}(S_0)}{S_{\gamma}} = \sum_{\gamma_{\text{ord}}} \frac{A_{\Lambda, \Lambda_0}^{\gamma_{\text{ord}}}(S_0)}{S_{\gamma_{\text{ord}}}},$$

there is a Hopf algebra morphism from \mathcal{H}_F to $\mathcal{H}_F^{\text{ord}}$ defined by summing over all the orders on the diagram,

$$(181) \quad \frac{\gamma}{S_{\gamma}} \rightarrow \sum_{\text{orders on } \gamma} \frac{\gamma^{\text{ord}}}{S_{\gamma}^{\text{ord}}}.$$

Moreover, there is also a morphism from the algebra $\mathcal{H}_F^{\text{ord}}$ to the algebra of rooted trees which associates n ordered diagram with its tree

$$(182) \quad \gamma^{\text{ord}} \rightarrow t$$

If we pullback the tree factorial by the composition of these two morphisms we get

$$(183) \quad \frac{1}{S_{\gamma}} \rightarrow \sum_{\text{orders on } \gamma} \frac{1}{t! S_{\gamma}^{\text{ord}}},$$

in accordance with the combinatorial interpretation of the tree factorial in (106).

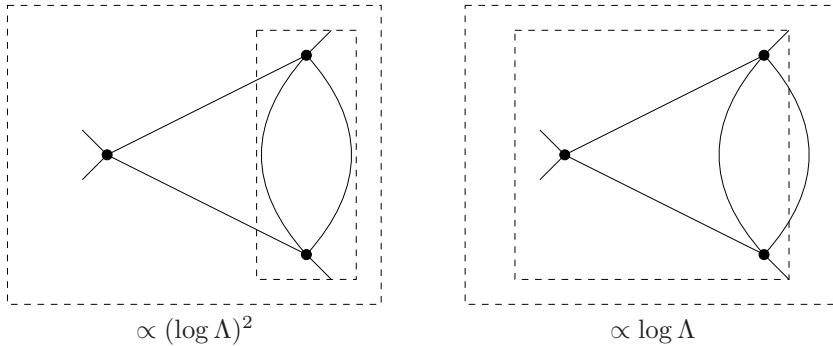


FIGURE 10. Two decompositions of the one loop four point function

From a renormalization group viewpoint, the use of ordered diagrams is necessary since a given diagram can be decomposed as different ordered ones that behave slightly differently. For instance, the 2 loop contribution to effective low energy coupling in ϕ^4 theory in 4 dimensions depicted in figure 10 can be ordered in two kind of ways: Either we consider it a 4 point function inserted in a 4 point function or as a 6 point function inserted in a 2 point function. In the first case we get a leading contribution proportional to $(\log \Lambda)^2$ whereas in the second case it is proportional $\log \Lambda$.

To implement dimensional analysis at the level of the Hopf algebras \mathcal{H}_F and $\mathcal{H}_F^{\text{ord}}$ we introduce decorated diagrams with generalized external structures. Let us decompose the space of action functionals as a direct sum $\mathcal{E} = \mathcal{E}' \oplus \mathcal{E}''$, with \mathcal{E}' containing the relevant and marginal terms, which span a finite dimensional space. The decorations are labels $i \in I$ for a basis of eigenstates of the scaling operator \mathcal{D} in \mathcal{E}' , which we write as \mathcal{O}_i and an extra index for \mathcal{E}'' . For instance, for ϕ^4 theory in dimension 4, a basis of \mathcal{E}' is given by the operators

$$(184) \quad \left\{ \int d^4x \frac{\phi^4}{4!}, \int d^4x \frac{\phi^2}{2}, \int d^4x \frac{(\partial\phi)^2}{2} \right\}$$

and we can identify the coupling constant, square mass and wave function parameter as the components of the action in this basis. These labels are generalization of the external structures introduced by Connes and Kreimer in [2], the main difference being that we have to treat all the interactions and not only the renormalizable ones.

A decorated Feynman diagram with a generalized external structure (with an order on the internal lines or not) is a couple $(\tilde{\gamma}, i)$ where the vertices of $\tilde{\gamma}$ are decorated with indices in I . These diagrams generate a commutative Hopf algebra $\mathcal{H}_F^{\text{dec}}$ with coproduct

$$(185) \quad \Delta(\tilde{\gamma}, i) = (\tilde{\gamma}, i) \otimes 1 + 1 \otimes (\tilde{\gamma}, i) + \sum_{\substack{\tilde{\gamma}_1, \dots, \tilde{\gamma}_n \text{ disjoint} \subset \tilde{\gamma} \\ i_1, \dots, i_n \text{ decorations} \in I}} (\tilde{\gamma}_1, i_1) \cdots (\tilde{\gamma}_n, i_n) \otimes \left(\frac{\tilde{\gamma}}{(\tilde{\gamma}_1, i_1) \cdots (\tilde{\gamma}_n, i_n)}, i \right),$$

with the vertices of the reduced diagram decorated by the generalized external structures of the subdiagrams.

Using the background field technique, such a diagram is evaluated with vertices obtained by projecting onto the eigenspace of the decoration carried by the vertex for the fluctuating field χ and the final result projected onto the eigenspace associated to i for the background field ϕ . Then, all the composition identities such as (136) hold since they simply amount to inserting the decomposition $\mathcal{E} = \mathcal{E}' \oplus \mathcal{E}$ for each action functional.

The unrenormalized low energy effective action S_Λ^{un} is computed by integrating modes with a propagator A_{Λ, Λ_0} starting with a bare action S_{Λ_0} , so that

$$(186) \quad S_\Lambda^{\text{un}} = \Psi_\alpha^{A_{\Lambda, \Lambda_0}}(S_{\Lambda_0}),$$

where α is the character that takes the value 1 on all diagrams. The renormalization procedure amounts to choose S_{Λ_0} in such a way that at the low energy scale Λ the relevant and marginal parameters are fixed to S_R . If we denote by β the character that takes the value 1 on the diagram with relevant and marginal external structures, then the bare action is determined by

$$(187) \quad S_R = \Psi_\beta^{A_{\Lambda, \Lambda_0}}(S_{\Lambda_0}),$$

Then, inverting this equation and substituting in (186), we get the renormalized low energy effective theory as

$$(188) \quad S_\Lambda^{\text{ren}} = \Psi_{\beta^{-1} * \alpha}^{A_{\Lambda, \Lambda_0}}(S_R).$$

In this discussion, we have used the characters only as a means of selecting some of the diagrams. In fact, the evaluation of the diagrams themselves leads to characters thanks to the decorations and external structures. This is the point of view adopted by Connes and Kreimer and is definitely more powerful since it allows to use the algebraic structure based on the Hopf algebra. Unfortunately, in our case it leads to a rather complicated formalism when we compose two effective action computations since we have to keep track of the location of the insertion of diagrams.

Finally, let us conclude by a simple example of an application of the previously outlined Lie algebraic techniques to the computation of the first Symanzik polynomial. In quantum field theory, a Feynman diagram γ with n edges can be evaluated, in dimension D , as

$$(189) \quad \int \frac{d^n \alpha}{[U_\gamma(\alpha)]^{\frac{D}{2}}} e^{-\frac{V_\gamma(\alpha, p)}{U_\gamma(\alpha)}}$$

where $U_\gamma(\alpha)$ and $V_\gamma(\alpha, p)$ are (Symanzik) polynomials in the Schwinger parameters $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ and the momenta p . Explicitly, the first Symanzik polynomial is expressed as

$$(190) \quad U_\gamma(\alpha) = \sum_{\substack{t \\ \text{spanning trees}}} \prod_{i \notin t} \alpha_i,$$

where a spanning tree is a tree drawn on γ that touches all the vertices.

For example, for the diagram depicted in figure 11, the first Symanzik polynomial is

$$(191) \quad U_{\text{diag}}(\alpha) = \alpha_1 \alpha_3 + \alpha_1 \alpha_4 + \alpha_2 \alpha_3 + \alpha_2 \alpha_4 + \alpha_3 \alpha_4.$$

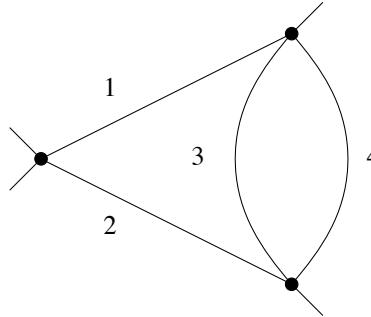


FIGURE 11. A two loop diagram with labelled edges

In the Hopf algebra of Feynman diagrams with labelled edges consider the infinitesimal characters such that

$$(192) \quad \delta_{\text{tree}}(\bullet\bullet) = 1 \quad \text{and} \quad \delta_{\text{loop}}(\bullet\circlearrowleft) = \alpha_i$$

and that vanish otherwise. Then, the first Symanzik polynomial can be written as

$$(193) \quad U_\gamma(\alpha) = e^{\delta_{\text{tree}}} * e^{\delta_{\text{loop}}}$$

Indeed, the first factor extracts all the trees and contracts them. Then, the second factor ensures that the reduced diagram has a unique vertex, which means that we extract only spanning trees, and assigns to it the product of the α over the non contracted edges. Then,

$$(194) \quad U_\gamma(\alpha) = e^{\delta_{\text{tree}}} * e^{\delta_{\text{loop}}} * e^{-\delta_{\text{tree}}} * e^{\delta_{\text{tree}}} = e^{\sum_n \delta_{n \text{ loop}}} * e^{\delta_{\text{tree}}}$$

where $\delta_{n \text{ loop}}$ is the infinitesimal character that takes the value $\sum_i \alpha_i$ on the one loop diagram with n edges and vanishes otherwise. This last relation follows from the Lie algebraic identity,

$$(195) \quad \frac{1}{n!} \underbrace{\left[\delta_{\text{tree}}, [\dots [\delta_{\text{tree}}, \delta_{\text{loop}}] \dots] \right]}_{n \text{ iterations}} = \delta_{n \text{ loop}}$$

interpreted in terms of the contraction/deletion relations (4.2).

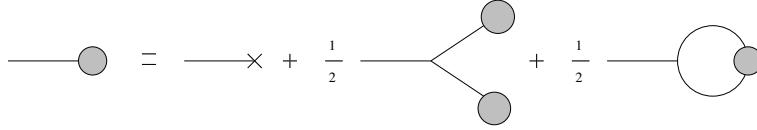
Thus, $U_\gamma(\alpha)$ can be evaluated by summing over all contraction schemes of the loops, generated by the iteration of the coproduct in the computation of $e^{\sum_n \delta_{n \text{ loop}}}$. For example, for the diagram depicted in figure 11,

$$(196) \quad U_{\text{diag}}(\alpha) = \frac{1}{2} [(\alpha_1 + \alpha_2 + \alpha_3)\alpha_4 + (\alpha_1 + \alpha_2 + \alpha_4)\alpha_3 + (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)]$$

These terms correspond to the following contraction schemes

$$(197) \quad \{(123, 4), (124, 3), (34, 12)\}.$$

In general, such an expression for the first Symanzik polynomial may prove useful in the study of the UV divergences, since the latter are governed by the integrals over loops.

FIGURE 12. Schwinger-Dyson equation in ϕ^3 theory

4.5. An algebraic structure related to Schwinger-Dyson equations.

Schwinger-Dyson equations are a set of functional differential equations playing a role analogous to the equation of motion in QFT (see [26] for a general discussion). When solved order by order, they allow to recover the perturbative expansion in terms of coupling constants. More precisely, they are functional differential equations for the derivative of the generating function $\mathcal{W}[J]$ of the connected Feynman diagrams. $\mathcal{W}[J]$ is expressed as a functional of the source J as

$$(198) \quad \mathcal{W}[J] = h \log \mathcal{Z}[J]$$

with $\mathcal{Z}[J]$ defined as

$$(199) \quad \mathcal{Z}[J] = \mathcal{N} \int [D\phi] e^{\frac{-S[\phi] + J \cdot \phi}{h}}$$

and the normalization constant defined such that $\mathcal{Z}[0] = 0$. Note that we have introduced h so that we can measure deviation from the classical theory obtained at $h = 0$.

One of the structural properties of the path integral is its invariance under change of variables in the space of fields over which we integrate. In the simplest case of an infinitesimal translation $\phi(x) \rightarrow \phi(x) + \epsilon(x)$, the invariance of \mathcal{Z} implies

$$(200) \quad \left(-S' \left[\frac{\delta}{\delta J} \right] + J \right) \mathcal{Z}[J] = 0$$

Replacing $\mathcal{Z}[J]$ by $e^{\frac{\mathcal{W}}{h}}$, it yields an equation for $\frac{\delta \mathcal{W}}{\delta J}$ whose complexity increases with the powers of the monomials entering into S . In the simplest case of the ϕ^3 interaction,

$$(201) \quad S[\phi] = \int d^D x \left(\frac{1}{2} \phi (-\Delta + m^2) \phi + \frac{g}{3!} \phi^3 \right),$$

the Schwinger-Dyson equation reads

$$(202) \quad \frac{\delta \mathcal{W}}{\delta J(x)} = \int d^D y \left[K(x, y) J(y) + \frac{g}{2} K(x, y) \left(\frac{\delta \mathcal{W}}{\delta J(y)} \right)^2 + \frac{gh}{2} K(x, y) \frac{\delta^2 \mathcal{W}}{\delta J(y) \delta J(y)} \right]$$

with K the kernel of the inverse of $-\Delta + m^2$. Let us represent the propagator K by a line, J by a cross and \mathcal{W} by a blob out of which we draw a line for each differentiation with respect to J . Then, eq. (202) is neatly summarized in figure 12. These equations generalize to more complicated interactions: if we consider a ϕ^n interaction we have a term for each partition of $n-1 = n_1 + \dots + n_k$ corresponding to the product of the n_i^{th} functional derivatives of \mathcal{W} . Of course, the functional differentials are distributions and the equations may become singular when the latter are evaluated at coinciding points, leading again to the UV divergences. In this

section, we shall ignore this problem, for instance by working with the regularized theory, and focus on the formal aspects of the Schwinger-Dyson equations.

The equation (202) is a fixed point equation for $\varphi = \frac{\delta \mathcal{W}}{\delta J}$ and may be solved using a geometric series over rooted trees. Indeed, rewrite (202) as

$$(203) \quad \varphi = \varphi_0 + \mathcal{F}_{\text{tree}}[\varphi] + \mathcal{F}_{\text{loop}}[\varphi],$$

and expand φ as a sum over rooted trees with black vertices for $\mathcal{F}_{\text{tree}}$ and white ones for $\mathcal{F}_{\text{loop}}$. Note that $\mathcal{F}_{\text{tree}}$ is bilinear, so that black vertices have one or two outgoing edges whereas white ones have only one because $\mathcal{F}_{\text{loop}}$ is linear. As we compose these operators along the trees, we obtain sums of Feynman diagrams with a distinguished external line and terminal lines equipped by crosses. If we iterate $\mathcal{F}_{\text{tree}}$ alone we only obtain tree-level diagrams while $\mathcal{F}_{\text{loop}}$ generates loops.

The correspondence between these Feynman diagrams and rooted trees expansions can be made more precise if we introduce a hierarchy on the vertices of the diagram, in complete analogy with the ordering of the lines in the iterative solution of Polchinski's equation. The ordering of the vertices is encoded in a spanning rooted tree drawn on the diagram constructed as follows. The first vertex we meet starting from the distinguished external leg is defined to be the highest one and is the root of the tree. As we remove this vertex from the diagram, the latter may fall into several connected components. If the diagram is disconnected after the removal of the vertex, we color the root in black, otherwise we leave it blank. Next, in each component, we choose as distinguished external line any line that was previously connected to the vertex we removed. Then, the lines point towards new vertices that are just below the vertex we removed in the hierarchy. No order is assumed between the vertices belonging to different connected components. Then, we iterate the operation, considering any component as a new diagram with a distinguished vertex till we exhaust all vertices of the diagram.

It is important to note that the trees we obtain have white vertices with at most one outgoing edge and black vertices at most two. Besides, we add extra outgoing edges, called black leaves, to the black vertices so that they have exactly two outgoing edges. Then, the tree we obtain are such that below any white vertex, there is always less white vertices than black leaves. This property is natural from the Schwinger-Dyson point of view: The black leaves represent the sources J and the white vertices the differential operator $\mathcal{F}_{\text{loop}}$, which vanishes if there are not enough sources. Conversely, given any tree with this property, we can reconstruct all the Feynman diagrams that appear in the corresponding power of the Schwinger-Dyson equation by relating each white vertex to a black leaf located below. This tree structure drawn on the Feynman diagrams may be used to define another Hopf algebra structure by performing admissible cuts on the tree. Of course, the correspondence between Feynman diagrams and trees is not one to one: To a tree correspond several diagrams and to a diagram several trees.

For planar diagrams, the ambiguity in the association of a rooted tree to a Feynman diagram can be disposed with by deciding to always move counterclockwise around the diagram as we choose the distinguished internal lines, starting from the distinguished external line. This way we associate a unique tree to each planar diagram. To reconstruct the Feynman diagram from the tree, we link every white vertex (starting with the lower ones) which is not a terminal one to the next black leaf above in the hierarchy of vertices, with a link always moving clockwise around the diagram. This way we have constructed a bijection between planar ϕ^3 diagrams

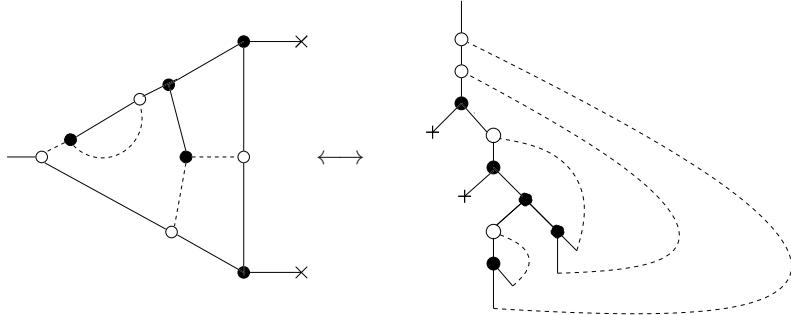


FIGURE 13. A planar ϕ^3 diagram and its reconstruction from a tree

with a distinguished external leg and planar trees with univalent (white) and bivalent (black) vertices fulfilling the additional condition that the number of white vertices located above a given white vertex must be no greater than the number of free edges located below. This is illustrated on a simple example on figure 13. Let us note that this bijection is very similar to the one given by P. Di Francesco in this lecture [19].

The previous bijection is also related to the counting of planar diagrams performed by 't Hooft in [27]. Indeed, the sum of those trees with weight 1 is a geometric series and solves the functional fixed point equation

$$(204) \quad F(x) = x + gF^2(x) + gh \frac{F(x) - F(0)}{x}.$$

Note that here we work with planar trees and count as different two trees that differ by an automorphism. This is why we do not divide by the symmetry factor in the geometric series. F is the generating function for trivalent planar diagrams with a distinguished external leg

$$(205) \quad F(x) = \sum_{m,n,p} a_{m,n,p} g^m h^n x^p,$$

with $a_{m,n,p}$ the number of such diagrams with m vertices, n loops and p external legs (not including the distinguished one). Because of the previous bijection. The first term in the planar Schwinger-Dyson equation (204) generates the black vertices and the second one the white ones. Besides, the subtraction of $F(0)$ ensures that there is a black leaf related to the white vertex (otherwise $F(x) - F(0)$ vanishes) and the division by x removes this leaf since it is connected to the vertex we have added to the Feynman diagram. Because of the previous bijection between trees and diagrams, $a_{m,n,p}$ also equals the number of connected planar ϕ^3 Feynman diagrams with a distinguished external leg, m vertices, n loops and p external legs in addition to the distinguished one. For example, up to order 3 in g , we find

$$(206) \quad F(x) = x + g(x^2 + h) + g^2(2x^3 + 3hx) + g^3(5x^4 + 10hx^2 + 4h^2) + \dots$$

If we set $h = 0$, we obtain the Catalan numbers $1, 2, 5, \dots$ that count the number of planar binary trees.

5. Conclusion and outlook

In this talk, we have presented, in a hopefully pedagogical way, some perturbative algebraic methods common to non linear equations and effective action computations. It is conveniently summarized by the following table.

| | |
|--|---|
| rooted trees | Feynman diagrams |
| non linear analysis | perturbative path integrals |
| fixed point equation | renormalization group equation |
| powers of non linear operators $X^t(x)$ | background field technique $A^\gamma(S)$ |
| $x' = (\text{id} - X)^{-1}(x) = \sum_t \frac{X^t(x)}{S_t}$ | $S'[\phi] = \log \int [D\chi] e^{-\frac{1}{2} x \cdot A \cdot x + S[\phi + \chi]}$ $= \sum_\gamma \frac{A^\gamma(S)}{S_\gamma}[\phi]$ |
| composition | successive integrations |

The examples presented here are very simple and can also be treated efficiently using other techniques. However, we think that Hopf algebraic techniques may prove to be useful when dealing with more complicated situations, in particular in the following subjects.

- **Multiscale renormalization**

In perturbative QFT, we recover a finite theory if we choose a cut-off dependent bare coupling constant and expand in powers of the low energy renormalized coupling constant. Nevertheless, the expansion is not satisfying since renormalons generally leads to a divergent series. For asymptotically free theories, the divergence of the perturbative perturbative series can be cured by expanding not in a single renormalized coupling, but in a scale dependent one [11]. The corresponding couplings are obtained by performing the renormalization not in a single step but by successive effective action computations. Therefore, we expect our techniques to be of interest in the construction of the effective expansion in the scale dependent coupling.

- **Graph and matroid polynomials**

Graph and matroid polynomials form a venerable subject where convolution products based on Hopf algebras have already proved to be useful (see [25]). It may be interesting to further develop the methods presented in the derivation the universality of the Tutte polynomial and its relation to the q -state Potts. In particular, this may be useful to investigate properties of the recently proposed multivariate generalizations of the Tutte polynomial [28], or other graph polynomials like the Martin and flow polynomials, which seem to have relations with QFT.

- **Spin foam models of quantum gravity**

Loop quantum gravity (see the book [29] for a general overview) is a tentative quantum theory of gravity that provides a construction of the Hilbert space of the quantum gravitational field using spin networks. Its path integral counterpart makes use of higher dimensional analogues of Feynman diagrams termed spin foams. The latter define a Hopf algebra,

as uncovered by Markopoulou [30], which may be useful to address the question of the Wilsonian renormalization of group field theory [31].

Acknowledgements

We are grateful to F. Girelli who collaborated with us on an early stage of this work. We would also like to thank V. Rivasseau and his group for their *constructive* remarks and M. Aguilar for some comments and a useful reference. Finally, let us also thank K. Ebrahimi-Fard, M. Marcolli and W. van Suijlekom, who organized the conference "Combinatorics and Physics" at the MPIM in Bonn in march 2007. Work partially supported by a EU Marie Curie fellowship EIF-025947 QGNC.

References

- [1] A. Connes and D. Kreimer, "Hopf algebras, renormalization and noncommutative geometry", *Commun. Math. Phys.* **199** (1998) 203 [arXiv:hep-th/9808042]
- [2] A. Connes and D. Kreimer, "Renormalization in quantum field theory and the Riemann-Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem", *Commun. Math. Phys.* **210** (2000) 249 [arXiv:hep-th/9912092]
- [3] A. Connes and D. Kreimer, "Renormalization in quantum field theory and the Riemann-Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group", *Commun. Math. Phys.* **216** (2001) 215 [arXiv:hep-th/0003188]
- [4] A. Connes and M. Marcolli, *Noncommutative Geometry, Quantum Fields and Motives*, preliminary version available at <http://www.alainconnes.org/en/downloads.php>
- [5] K. Ebrahimi-Fard and D. Kreimer, "Hopf algebra approach to Feynman diagram calculations" *J. Phys. A* **38** (2005) R385 [arXiv:hep-th/0510202]
- [6] A. Zee, *Quantum field theory in a nutshell*, Princeton Univ. Pr. (2003)
- [7] J. Zinn-Justin, *Quantum field theory and critical phenomena*, 4th edition, Oxford Univ. Pr (2002)
- [8] V. Rivasseau, "An Introduction To Renormalization", In *Duplantier, B. (ed.) et al.: Poincare Seminar 2002* 139-177, also available at <http://www.bourbaphy.fr/Rivasseau.pdf>
- [9] J.C. Butcher, *The numerical analysis of ordinary differential equations*, John Wiley & sons (1987)
- [10] J. Polchinski, "Renormalization And Effective Lagrangians", *Nucl. Phys. B* **231** (1984) 269
- [11] V. Rivasseau, *From perturbative to constructive renormalization*, Princeton Univ. Pr. (1991)
- [12] Y.G. Sinai *Probability theory: An introductory course*, Springer (1991)
- [13] G. Jona-Lasinio, 'Renormalization group and probability theory", *Phys. Rept.* **352** (2001) 439
- [14] C. Brouder, "Runge-Kutta methods and renormalization" *Eur. Phys. J. C* **12** (2000) 521 [arXiv:hep-th/9904014]
- [15] P. Cartier, "A review of Hopf algebras and their applications to combinatorics", lecture at this conference, see also "A primer of Hopf algebras", IHES preprint M/06/40
- [16] E. Hairer and G. Wanner, "On the Butcher group and general multi-value methods", *Computing* **13**, 1-15 (1974)
- [17] D. Kreimer, "On the Hopf algebra structure of perturbative quantum field theories" *Adv. Theor. Math. Phys.* **2** (1998) 303 [arXiv:q-alg/9707029]
- [18] R. P. Stanley, *Enumerative combinatorics Vol. 1*, Cambridge Univ. Pr. (1997)
- [19] P. Di Francesco, "Integrable Combinatorics", lecture at this conference, see also "2D quantum gravity, matrix models and graph combinatorics" [arXiv:math-ph/0406013]
- [20] F. Girelli, T. Krajewski and P. Martinetti, "An algebraic Birkhoff decomposition for the continuous renormalization group", *J. Math. Phys.* **45** (2004) 4679 [arXiv:hep-th/0401157]
- [21] T. R. Hurd, "A renormalization group proof of perturbative renormalizability", *Commun. Math. Phys.* **124** (1989) 153
- [22] M. Le Bellac *Quantum and statistical field theory*, Clarendon (1991)
- [23] V.I. Arnol'd *Geometrical methods in the theory of ordinary differential equations* Grundlehren der Mathematischen Wissenschaften, Springer-Verlag (1983)
- [24] J. Ellis-Monaghan and C. Merino, "Graph polynomials and their applications I: The Tutte polynomial", [arXiv:0803.3079]

- [25] W. Kook, V. Reiner and D. Stanton, "A Convolution Formula for the Tutte Polynomial", *J. Comb. Theory* **B76** (1999) 297 [arXiv:math/9712232]
- [26] P. Cvitanovic "Feynman Diagrammatics: Six Easy Pieces" lecture at this conference, see also *Field Theory* (1983), available at <http://ChaosBook.org/FieldTheory/>
- [27] G. 't Hooft, "Counting planar diagrams with various restrictions", *Nucl. Phys. B* **538** (1999) 389 [arXiv:hep-th/9808113]
- [28] A. Sokal "The multivariate Tutte polynomial (alias Potts model) for graphs and matroids", *Surveys in combinatorics 2005*, 173–226, London Math. Soc. Lecture Note Ser., 327, Cambridge Univ. Pr. (2005)
- [29] C. Rovelli, *Quantum Gravity*, Cambridge Univ. Pr. (2004)
- [30] F. Markopoulou, "Coarse graining in spin foam models", *Class. Quant. Grav.* **20** (2003) 777 [arXiv:gr-qc/0203036].
- [31] D. Oriti, "Group Field Theories, or the combinatorial and group-theoretic path to Quantum Gravity and quantum simplicial geometry" talk at this conference, see also "A combinatorial and field theoretic path to quantum gravity: the new challenges of group field theory" arXiv:0709.4157 [hep-th].

CENTRE DE PHYSIQUE THÉORIQUE, UNITÉ MIXE DE RECHERCHE (UMR) 6207 DU CNRS ET DES UNIVERSITÉS AIX-MARSEILLE 1 ET 2, SUD TOULON-VAR, LABORATOIRE AFFILIÉ À LA FRUMAM (FR 2291) AND CNRS-LUMINY, CASE 907, 13288 MARSEILLE CEDEX 9, FRANCE

E-mail address: `krajew@cpt.univ-mrs.fr`

INSTITUT FÜR THEORETISCHE PHYSIK, GEORG-AUGUST UNIVERSITÄT GÖTTINGEN, FRIEDRICH-HUND-PLATZ 1, 37077 GÖTTINGEN, DEUTSCHLAND

E-mail address: `martinetti@theorie.physik.uni-goettingen.de`

Algebraic analysis of non-renormalization theorems in supersymmetric field theories

E. Kraus

ABSTRACT. The extension of coupling constants to space time dependent fields, the local couplings, makes possible to prove the various non-renormalization theorems of supersymmetric field theories entirely by an algebraic analysis. In particular, for super-Yang-Mills theories the construction implies a one-loop non-renormalization theorem of the gauge coupling in the classical context. However, renormalization in presence of the local gauge coupling is peculiar due to an anomalous breaking of supersymmetry in one-loop order and it turns out that the anomaly induces the 2-loop order of the gauge β -function. It is seen that in $N=2$ supersymmetric theories the supersymmetry anomaly is canceled and the gauge beta-function has only a one loop term as expected from the classical analysis.

1. Introduction

The extension of coupling constants to space-time dependent external fields, i.e. to local couplings, has been an important tool in renormalized perturbation theory for a long time [1]. It has been moved in the center of interest again from a string point of view since the couplings there are dynamical fields and enter quite naturally as external fields the effective field theories derived from string theories. It has been seen that local couplings in combination with holomorphicity imply the non-renormalization theorems of supersymmetric field theories [2, 3, 4, 5].

Most of the results have been derived in the framework of Wilsonian renormalization and by using the Wilsonian effective action. We have now started a rigorous and scheme-independent construction of supersymmetric field theories with local couplings. In the Wess–Zumino model [6], in SQED [7] and in softly broken SQED [8] it was shown that local couplings allow the derivation of the non-renormalization theorems [9, 10] as well as the generalized non-renormalization theorem [11] from renormalization properties of the extended model in an algebraic context.

Most interesting is the extension to super-Yang-Mills theories [12, 13]. There the classical analysis applies in the same way as for SQED, but in the procedure of renormalization one finds a new anomaly of supersymmetry in one-loop order [12]. Contrary to the Adler-Bardeen anomaly the supersymmetry anomaly is a variation

Talk presented at the Mini-Workshop on renormalization at the Max Planck Institute for Mathematics Bonn, 15–16 December, 2006.

under BRS-symmetry, but the respective counterterm depends on the logarithm of the coupling. Thus its coefficient is scheme independent and is not introduced in the procedure of renormalization.

Evaluating the symmetry identities one can show, that the anomaly of supersymmetry is induced by the renormalization of the topological term [13]. In addition we demonstrate that in all orders except for one loop the renormalization of the topological term determines the renormalization of the gauge coupling constant. As an application we derive the closed expression of the gauge β function in terms of its one-loop coefficient and the anomaly coefficient and derive therefrom one-loop finiteness of $N=2$ supersymmetric Yang-Mills-theories.

2. Supersymmetric Yang-Mills theories

We consider super-Yang-Mills theories with a simple gauge group as for example $SU(N)$ in the Wess-Zumino gauge. In the Wess-Zumino gauge the vector multiplet consists of the gauge fields $A^\mu = A_a^\mu \tau_a$, the gaugino fields $\lambda^\alpha = \lambda_a^\alpha \tau_a$ and their complex conjugate fields $\bar{\lambda}^{\dot{\alpha}}$ and the auxiliary fields $D = D_a \tau_a$. For constant coupling the action

$$(1) \quad \Gamma_{\text{SYM}} = \text{Tr} \int d^4x \left(-\frac{1}{4g^2} G^{\mu\nu}(gA) G_{\mu\nu}(gA) + i\lambda^\alpha \sigma_{\alpha\dot{\alpha}}^\mu D_\mu \bar{\lambda}^{\dot{\alpha}} + \frac{1}{8} D^2 \right),$$

with

$$(2) \quad \begin{aligned} G^{\mu\nu}(A) &= \partial^\mu A^\nu - \partial^\nu A^\mu + i[A^\mu, A^\nu], \\ D^\mu \lambda &= \partial^\mu \lambda - ig[A^\mu, \lambda], \end{aligned}$$

is invariant under non-abelian gauge transformations and supersymmetry transformations.

$$(3) \quad \begin{aligned} \delta_\alpha A^\mu &= i\sigma_{\alpha\dot{\alpha}}^\mu \bar{\lambda}^{\dot{\alpha}}, & \bar{\delta}_{\dot{\alpha}} A^\mu &= -i\lambda^\alpha \sigma_{\alpha\dot{\alpha}}^\mu, \\ \delta_\alpha \lambda_\beta &= -\frac{i}{2}(\epsilon_{\alpha\beta} D - \sigma_{\alpha\beta}^{\mu\nu} G_{\mu\nu}), & \bar{\delta}_{\dot{\alpha}} \lambda_\alpha &= 0, \\ \delta_\alpha \bar{\lambda}^{\dot{\alpha}} &= 0, & \bar{\delta}_{\dot{\alpha}} \lambda_{\dot{\beta}} &= -\frac{i}{2}(\epsilon_{\dot{\alpha}\dot{\beta}} D - \bar{\sigma}_{\dot{\alpha}\dot{\beta}}^{\mu\nu} G_{\mu\nu}), \\ \delta_\alpha D &= 2\sigma_{\alpha\dot{\alpha}}^\mu D_\mu \bar{\lambda}^{\dot{\alpha}}, & \bar{\delta}_{\dot{\alpha}} D &= 2D_\mu \lambda^\alpha \sigma_{\alpha\dot{\alpha}}^\mu. \end{aligned}$$

In the Wess-Zumino gauge [14, 15] the algebra of supersymmetry transformations does not close on translations but involves an additional field dependent gauge transformation:

$$(4) \quad \begin{aligned} \{\delta_\alpha, \bar{\delta}_{\dot{\alpha}}\} &= 2i\sigma_{\alpha\dot{\alpha}}^\mu (\partial^\mu + \delta_{A^\mu}^{\text{gauge}}), \\ \{\delta_\alpha, \delta_\beta\} &= \{\bar{\delta}_{\dot{\alpha}}, \bar{\delta}_{\dot{\beta}}\} = 0. \end{aligned}$$

Thus, on gauge invariant field monomials the supersymmetry algebra takes their usual form, and it is possible to classify the Lagrangians into the usual $N = 1$ multiplets.

The Lagrangians of supersymmetric field theories are the F or D components of a $N = 1$ multiplet. In particular one finds that the Lagrangian of the super-Yang-Mills action is the highest component of a chiral and an antichiral multiplet. Using a superfield notation in the chiral representation the chiral Lagrangian multiplet

\mathcal{L}_{SYM} is given by

$$(5) \quad \mathcal{L}_{\text{SYM}} = -\frac{1}{2}g^2 \text{Tr} \lambda^\alpha \lambda_\alpha + \Lambda^\alpha \theta_\alpha + \theta^2 L_{\text{SYM}} ,$$

with the chiral super-Yang-Mills Lagrangian L_{SYM}

$$(6) \quad \begin{aligned} L_{\text{SYM}} = & \text{Tr} \left(-\frac{1}{4} G^{\mu\nu}(gA) G_{\mu\nu}(gA) + ig \lambda^\alpha \sigma_{\alpha\dot{\alpha}}^\mu D_\mu(g\bar{\lambda}^{\dot{\alpha}}) + \frac{1}{8} g^2 D^2 \right. \\ & \left. - \frac{i}{8} \epsilon^{\mu\nu\rho\sigma} G_{\mu\nu}(gA) G_{\rho\sigma}(gA) \right) . \end{aligned}$$

By complex conjugation one obtains the respective antichiral multiplet $\bar{\mathcal{L}}_{\text{SYM}}$.

The crucial point for the present considerations is the fact that the topological term $\text{Tr} G\bar{G}$ appears in the supersymmetric Lagrangians (6) and is related to the kinetic term $-\frac{1}{4}\text{Tr}(GG)$ via supersymmetry. In quantum field theory $\text{Tr} G\bar{G}$ is the anomaly term of axial symmetry and it is well-known, that its higher order corrections are unambiguously determined from gauge invariance. For this reason one expects the same effect for the kinetic term, which describes the renormalization of the gauge coupling. In ordinary perturbation theory these relations are not found, since the topological term disappears from the action by integration. However, with a local gauge coupling one is able to include the complete Lagrangian with the topological term into the action and the improved renormalization properties become apparent.

We will now extend the coupling constant g to an external field $g(x)$ [7, 12]. For maintaining at the same time supersymmetry the coupling has to be extended to a supermultiplet. For this purpose we introduce a chiral and an antichiral field multiplet η and $\bar{\eta}$:

$$\eta = \eta + \theta\chi + \theta^2 f , \bar{\eta} = \bar{\eta} + \bar{\theta}\bar{\chi} + \bar{\theta}^2 \bar{f} ,$$

which couple to the Lagrangian multiplet of the Super-Yang-Mills action:

$$(7) \quad \begin{aligned} \Gamma_{\text{SYM}} = & -\frac{1}{4} \int dS \eta \mathcal{L}_{\text{SYM}} - \frac{1}{4} \int d\bar{S} \bar{\eta} \bar{\mathcal{L}}_{\text{SYM}} \\ = & \int d^4x \left(\eta L_{\text{SYM}} - \frac{1}{2} \chi^\alpha \Lambda_\alpha - \frac{1}{2} fg^2 \text{Tr} \lambda^\alpha \lambda_\alpha + \text{c.c.} \right) . \end{aligned}$$

We identify the real part of η with the inverse of the square of the local coupling:

$$(8) \quad \eta + \bar{\eta} = \frac{1}{g^2(x)} ,$$

and observe that the imaginary part of η couples to the topological term. Thus, it takes the role of a space time dependent Θ angle

$$(9) \quad \eta - \bar{\eta} = 2i\Theta .$$

Hence, dependence on the superfield η and $\bar{\eta}$ is governed in loop order l by a topological formula

$$(10) \quad N_g \Gamma^{(l)} = (N_A + N_\lambda + N_D + 2(l-1)) \Gamma^{(l)} .$$

and by the identity:

$$(11) \quad \int d^4x \left(\frac{\delta}{\delta\eta} - \frac{\delta}{\delta\bar{\eta}} \right) \Gamma = -i \int d^4x \frac{\delta}{\delta\Theta} \Gamma = 0 .$$

3. Invariant counterterms

Invariant counterterms are the local field monomials which are invariant under the classical symmetries and hold as counterterms to the non-local loop diagrams in loop order l . As such they are in one-to-one correspondence with the possible UV divergences of the theory.

Constructing the invariant counterterms to Γ_{SYM} with local coupling we find from gauge invariance and supersymmetry:

$$\begin{aligned}
 \Gamma_{\text{ct,phys}}^{(l)} &= z_{\text{YM}}^{(l)} \left(-\frac{1}{8} \int dS \, \boldsymbol{\eta}^{1-l} \mathcal{L}_{\text{SYM}} - \frac{1}{8} \int d\bar{S} \, \bar{\boldsymbol{\eta}}^{1-l} \bar{\mathcal{L}}_{\text{SYM}} \right) \\
 &= z_{\text{YM}}^{(l)} \int d^4x \left(-\frac{1}{4} (2g^2)^{l-1} G^{\mu\nu}(gA) G_{\mu\nu}(gA) \right. \\
 &\quad \left. - \frac{1}{8} (2g^2)^l (l-1) \Theta G^{\mu\nu}(gA) \tilde{G}_{\mu\nu}(gA) + \dots \right). \tag{12}
 \end{aligned}$$

For local coupling the Θ angle couples to a total derivative only for $l = 0$. Hence, using the identity (11) one finds that these counterterms as well as the respective UV divergences are excluded in all loop orders except for one loop:

$$z_{\text{YM}}^{(l)} = 0 \quad \text{for } l \geq 2, \tag{13}$$

i.e. the Θ angle is not renormalized and determines the renormalization of the coupling via supersymmetry. The resulting counterterm action is the holomorphic effective action which has been stated already in [11]:

$$\Gamma_{\text{eff}} = -\frac{1}{8} \int dS \, (\boldsymbol{\eta} + z_{\text{YM}}^{(1)}) \mathcal{L}_{\text{SYM}} + \text{c.c.} \tag{14}$$

It is obvious from eq. (12) that the one-loop order is special. Here the renormalization of the coupling is not related to the renormalization of the Θ angle, and one has an arbitrary counterterm $z_{\text{YM}}^{(1)}$ and a corresponding UV divergence. However, as we will show in the next section, the quantum corrections to $\text{Tr } G\tilde{G}$ induce an anomaly of supersymmetry in one-loop order. It makes quantization of super-Yang-Mills theories non-trivial such that quantum results cannot be obtained from an effective action or by using multiplicative renormalization.

4. The anomalous breaking of supersymmetry

To quantize supersymmetric field theories in the Wess-Zumino gauge one includes the gauge transformations, supersymmetry transformations and translations into the nilpotent BRS operator [16, 17]:

$$(15) \quad \mathbf{s}\phi = \delta_c^{\text{gauge}} + \epsilon^\alpha \delta_\alpha + \bar{\delta}_{\dot{\alpha}} \epsilon^{\dot{\alpha}} - i\omega^\nu \delta_\nu^T .$$

The fields $c(x)$ are the usual Faddeev-Popov ghosts, ϵ^α , $\bar{\epsilon}^{\dot{\alpha}}$ and ω^ν are constant ghosts of supersymmetry and translations, respectively.

As for usual gauge theories BRS transformations are encoded in the Slavnov-Taylor identity:

$$(16) \quad \mathcal{S}(\Gamma_{\text{cl}}) = 0$$

with

$$(17) \quad \Gamma_{\text{cl}} = \Gamma_{\text{SYM}} + \Gamma_{g.f.} + \Gamma_{\phi\pi} + \Gamma_{ext.f.} .$$

Details of the construction are not relevant for the following considerations, but it is only important to keep in mind that the Slavnov-Taylor identity for supersymmetric actions expresses gauge invariance as well as supersymmetry.

In addition to the Slavnov-Taylor identity the dependence on the external fields η and $\bar{\eta}$ is restricted to all orders by the identity (11) [12].

In the course of renormalization the Slavnov-Taylor identity (16) has to be established for the 1PI Green functions to all orders of perturbation theory. From the quantum action principle one finds that the possible breaking terms are local in one-loop order:

$$(18) \quad \mathcal{S}(\Gamma) = \Delta_{\text{brs}} + \mathcal{O}(\hbar^2) .$$

Algebraic consistency yields the constraint

$$(19) \quad \mathbf{s}_{\Gamma_{\text{cl}}} \Delta_{\text{brs}} = 0 .$$

Gauge invariance can be established as usually, i.e. one has

$$(20) \quad \mathcal{S}(\Gamma) \Big|_{\epsilon, \bar{\epsilon}=0} = \mathcal{O}(\hbar^2) ,$$

and the remaining breaking terms depend on the supersymmetry ghosts ϵ and $\bar{\epsilon}$, and represent as such a breaking of supersymmetry. Having gauge invariance established, the supersymmetry algebra closes on translations. Using the supersymmetry algebra one obtains that the remaining breaking terms of supersymmetry are variations of field monomials with the quantum numbers of the action:

$$(21) \quad \Delta_{\text{brs}} = \mathbf{s}_{\Gamma_{\text{cl}}} \hat{\Gamma}_{\text{ct}} .$$

However, not all of the field monomials in $\hat{\Gamma}_{\text{ct}}$ represent scheme-dependent counter-terms of the usual form. There is one field monomial in $\hat{\Gamma}_{\text{ct}}$, which depends on the

logarithm of the gauge coupling, but whose BRS variation is free of logarithms:

$$\begin{aligned}
 (22) \quad \Delta_{\text{brs}}^{\text{anomaly}} &= s \int d^4x \ln g(x) (L_{\text{SYM}} + \bar{L}_{\text{SYM}}) \\
 &= (\epsilon^\alpha \delta_\alpha + \bar{\epsilon}^{\dot{\alpha}} \bar{\delta}_{\dot{\alpha}}) \int d^4x \ln g(x) (L_{\text{SYM}} + \bar{L}_{\text{SYM}}) \\
 &= \int d^4x \left(i \ln g(x) (\partial_\mu \Lambda^\alpha \sigma_{\alpha\dot{\alpha}}^\mu \bar{\epsilon}^{\dot{\alpha}} - \epsilon^\alpha \sigma_{\alpha\dot{\alpha}}^\mu \partial_\mu \bar{\Lambda}^{\dot{\alpha}}) \right. \\
 &\quad \left. - \frac{1}{2} g^2(x) (\epsilon \chi + \bar{\epsilon} \bar{\chi}) (L_{\text{SYM}} + \bar{L}_{\text{SYM}}) \right).
 \end{aligned}$$

Indeed, due to the total derivative in the first line the breaking $\Delta_{\text{brs}}^{\text{anomaly}}$ is free of logarithms for constant coupling and for any test with respect to the local coupling. Moreover one immediately verifies, that $\Delta_{\text{brs}}^{\text{anomaly}}$ satisfies the topological formula in one-loop order. Therefore $\Delta_{\text{brs}}^{\text{anomaly}}$ satisfies all constraints on the breakings and can appear as a breaking of the Slavnov-Taylor identity in the first order of perturbation theory.

However, being the variation of a field monomial depending on the logarithm of the coupling $\Delta_{\text{brs}}^{\text{anomaly}}$ cannot be induced by divergent one-loop diagrams, which are all power series in the coupling. Thus, the corresponding counterterm is not related to a naive contribution induced in the procedure of subtraction and does not represent a naive redefinition of time-ordered Green functions. Therefore, $\Delta_{\text{brs}}^{\text{anomaly}}$ is an anomalous breaking of supersymmetry in perturbation theory and we remain with

$$(23) \quad \mathcal{S}(\Gamma) = r_\eta^{(1)} \Delta_{\text{brs}}^{\text{anomaly}} + \mathcal{O}(\hbar^2).$$

From its characterization it is straightforward to prove with algebraic methods that the coefficient of the anomaly is gauge and scheme independent [12]. Furthermore, by evaluating the Slavnov-Taylor identity one can find an expression for $r_\eta^{(1)}$ in terms of convergent loop integrals [13]. Using background gauge fields \hat{A}^μ and Feynman gauge $\xi = 1$ the anomaly coefficient is explicitly related to insertions of the topological term and the axial current of gluinos into self energies of background fields:

$$(24) \quad g^2 r_\eta^{(1)} = -\frac{1}{2} \Sigma_{\eta-\bar{\eta}}^{(1)}(p_1, -p_1) \Big|_{\xi=1},$$

where $\Sigma_{\eta-\bar{\eta}}$ is defined by

$$\begin{aligned}
 \Gamma_{\eta-\bar{\eta} \hat{A}_a^\mu \hat{A}_b^\nu}(q, p_1, p_2) &= \left([i \text{Tr}(\partial(g^2 \lambda \sigma \bar{\lambda}) - \frac{1}{4} G^{\mu\nu} \tilde{G}_{\mu\nu}(gA + \hat{A}))] \cdot \Gamma \right)_{\hat{A}_a^\mu \hat{A}_b^\nu}(q, p_1, p_2) \\
 (25) \quad &= i \epsilon^{\mu\nu\rho\sigma} p_{1\rho} p_{2\sigma} \delta_{ab} (-2 + \Sigma_{\eta-\bar{\eta}}(p_1, p_2)).
 \end{aligned}$$

Explicit evaluation of the respective one-loop diagrams yields

$$(26) \quad r_\eta^{(1)} = (-1 + 2) \frac{C(G)}{8\pi^2} = \frac{C(G)}{8\pi^2},$$

where the first term comes from the axial anomaly of gauginos and the second term from the insertion of the topological term.

We want to mention that the anomaly coefficient vanishes in SQED, since one-loop diagrams to $\Gamma_{\eta-\bar{\eta} A^\mu A^\nu}$ do not exist. The anomaly coefficient also vanishes in $N = 2$ theories. For $N = 2$ super-Yang-Mills theories the analysis of the previous

sections holds in the same form, where \mathcal{L}_{SYM} and η are now $N = 2$ chiral multiplets. The algebraic analysis yields an anomaly of the same form as for $N = 1$ theories, however, in the explicit evaluation the anomaly coefficient of $N = 2$ theories vanishes since the two fermionic fields just cancel the contribution arising from the topological term $\text{Tr } G^{\mu\nu} \tilde{G}_{\mu\nu}$ (cf. (25, 26)).

5. Renormalization and the gauge- β -functions

In the framework of perturbation theory the anomaly of supersymmetry cannot be removed by a local counterterm. However, one is able to proceed with algebraic renormalization nevertheless by rewriting the anomalous breaking in the form of a differential operator [12]:

$$(27) \quad \Delta_{\text{brs}}^{\text{anomaly}} = \int d^4x \left(g^6 r_\eta^{(1)} (\epsilon\chi + \bar{\chi}\bar{\epsilon}) \frac{\delta}{\delta g^2} - ir_\eta^{(1)} \partial_\mu \ln g^2 \left((\sigma^\mu \bar{\epsilon})^\alpha \frac{\delta}{\delta \chi^\alpha} + (\epsilon \sigma^\mu)^\dot{\alpha} \frac{\delta}{\delta \bar{\chi}^{\dot{\alpha}}} \right) \right) \Gamma_{\text{cl}}$$

Then one has

$$(28) \quad (\mathcal{S} + r_\eta^{(1)} \delta \mathcal{S}) \Gamma = \mathcal{O}(\hbar^2) .$$

For algebraic consistency one has to require the nilpotency properties of the classical Slavnov-Taylor operator also for the extended operator. Nilpotency determines an algebraic consistent continuation of (27). This continuation is not unique but contains at the same time all redefinitions of the coupling compatible with the formal power series expansion of perturbation theory.

As a result one finds an algebraic consistent Slavnov-Taylor identity in presence of the anomaly:

$$(29) \quad \mathcal{S}^{r_\eta}(\Gamma) = 0 \quad \text{and} \quad \int d^4x \left(\frac{\delta}{\delta \eta} - \frac{\delta}{\delta \bar{\eta}} \right) \Gamma = 0 ,$$

where the anomalous part is of the form:

$$(30) \quad \mathcal{S}^{r_\eta}(\Gamma) = \mathcal{S}(\Gamma) - \int d^4x \left(g^4 \delta F(g^2) (\epsilon\chi + \bar{\chi}\bar{\epsilon}) \frac{\delta}{\delta g^2} + i \frac{\delta F}{1 + \delta F} \partial_\mu g^{-2} \left((\sigma^\mu \bar{\epsilon})^\alpha \frac{\delta}{\delta \chi^\alpha} + (\epsilon \sigma^\mu)^\dot{\alpha} \frac{\delta}{\delta \bar{\chi}^{\dot{\alpha}}} \right) \right) \Gamma ,$$

with

$$(31) \quad \delta F(g^2) = r_\eta^{(1)} g^2 + \mathcal{O}(g^4) .$$

The lowest order term is uniquely fixed by the anomaly, whereas the higher orders in $\delta F(g^2)$ correspond to the scheme-dependent finite redefinitions of the coupling.

The simplest choice for δF is given by

$$(32) \quad \delta F = r_\eta^{(1)} g^2 ,$$

and another choice is provided by

$$(33) \quad \frac{\delta F}{1 + \delta F} = r_\eta^{(1)} g^2 , \quad \text{i.e.} \quad \delta F = \frac{r_\eta^{(1)} g^2}{1 - r_\eta^{(1)} g^2} .$$

As seen below, the latter choice gives the NSZV expression of the gauge β function [18, 11].

Algebraic renormalization with the symmetry operator (29) is performed in the conventional way. In particular one can derive the β functions from an algebraic construction of the renormalization group equation in presence of the local coupling. Starting from the classical expression of the RG equation

$$(34) \quad \kappa \partial_\kappa \Gamma_{\text{cl}} = 0$$

we construct the higher orders of the RG equation by constructing the general basis of symmetric differential operator with the quantum numbers of the action:

$$(35) \quad \mathcal{R} = \kappa \partial_\kappa + \mathcal{O}(\hbar) ,$$

with

$$(36) \quad \mathbf{s}_\Gamma^{r_\eta} \mathcal{R} \Gamma - \mathcal{R} \mathcal{S}^{r_\eta} (\Gamma) = 0 ,$$

$$(37) \quad \left[\int d^4x \left(\frac{\delta}{\delta \eta} - \frac{\delta}{\delta \bar{\eta}} \right), \mathcal{R} \right] = 0 .$$

The general basis for the symmetric differential operators consists of the differential operator of the supercoupling η and $\bar{\eta}$ and several field redefinition operators. The differential operator of the coupling determines the β function of the coupling, whereas the field redefinition operators correspond to the anomalous dimensions of fields.

For the present paper we focus on the operator of the β function and neglect the anomalous dimensions. For proceeding we construct first the RG operator, which is symmetric with respect to the classical Slavnov-Taylor operator, i.e. we set $\delta F = 0$. Using a superspace notation we find:

$$(38) \quad \mathcal{R}_{\text{cl}} = - \sum_{l \geq 1} \hat{\beta}_g^{(l)} \left(\int dS \eta^{-l+1} \frac{\delta}{\delta \eta} + \int d\bar{S} \bar{\eta}^{-l+1} \frac{\delta}{\delta \bar{\eta}} \right) + \dots ,$$

and

$$(39) \quad \mathbf{s}_\Gamma \mathcal{R}_{\text{cl}} \Gamma - \mathcal{R}_{\text{cl}} \mathcal{S}(\Gamma) = 0 .$$

Evaluating then the consistency equation (37) we obtain

$$(40) \quad \hat{\beta}_g^{(l)} = 0 \quad \text{for} \quad l \geq 2 .$$

These restrictions are the same restrictions as we have found for the invariant counterterms of the super-Yang-Mills action (see (12) with (13)) and represent the fact, that for super-Yang-Mills the renormalization of the coupling is governed by the renormalization of the Θ angle in loop orders greater than one.

Finally we extend the one-loop operator \mathcal{R}_{cl} to a symmetric operator with respect to the anomalous ST identity (36) and find by a straightforward calculation

$$(41) \quad \begin{aligned} \mathcal{R} &= \hat{\beta}_g^{(1)} \int d^4x g^3 (1 + \delta F(g^2)) \frac{\delta}{\delta g} + \dots \\ &= \hat{\beta}_g^{(1)} \int d^4x g^3 (1 + r_\eta^{(1)} g^2 + \mathcal{O}(\hbar^2)) \frac{\delta}{\delta g} + \dots . \end{aligned}$$

For constant coupling we find from (41) the closed expression of the gauge β function

$$(42) \quad \beta_g = \hat{\beta}_g^{(1)} g^3 (1 + \delta F(g^2)) = \hat{\beta}_g^{(1)} g^3 (1 + r_\eta^{(1)} g^2 + \mathcal{O}(\hbar^2)) .$$

Thus, the two-loop order is uniquely determined by the anomaly and the one-loop coefficient, whereas higher orders depend on the specific form one has chosen for the

function $\delta F(g^2)$. In particular one has for the minimal choice (32) a pure two-loop β -function and for the NSZV-choice (33) one obtains the NSZV expression [18]

$$(43) \quad \beta_g = \hat{\beta}_g^{(1)} g^3 \frac{1}{1 - r_\eta^{(1)} g^2} .$$

of the gauge β function of pure Super-Yang-Mills theories.

For $N = 2$ super-Yang-Mills the anomaly vanishes and the classical Slavnov–Taylor identity (16) can be extended to all orders. Then the classical renormalization group operator (38) is a symmetric operator to all orders and one gets from (37) a pure one-loop contribution to the gauge β function, i.e.,

$$(44) \quad \mathcal{S}(\Gamma^{N=2}) = 0 \implies \beta_g^{N=2} = \hat{\beta}_g^{(1)} g^3 .$$

6. Conclusions

The extension of the coupling constants to external fields is a crucial step for deriving the non-renormalization theorems of supersymmetry in a scheme-independent way and independent from the usage of superspace methods. For super-Yang-Mills theories the extended model yields the non-renormalization of the coupling beyond one-loop order due to the supersymmetry induced relation of the coupling with the Θ angle.

The non-renormalization of the Θ -angle is the real new result gained by the construction with local coupling. Using gauge invariance and the property that the Θ angle couples to a total derivative in the classical action higher order corrections are uniquely determined by convergent expressions. In this way, local coupling in addition gives a simple proof of the non-renormalization of the Adler–Bardeen anomaly [19, 20, 21] also for non-supersymmetric theories [22].

For supersymmetric theories the non-renormalization of the Θ angle induces not only the non-renormalization theorem of the gauge coupling but also a supersymmetry anomaly in one-loop order. The supersymmetry anomaly is the variation of a gauge-invariant field monomial which depends on the logarithm of the local coupling. As such it cannot be induced in the procedure of regularization, since loop diagrams are power series in the coupling. Hence the anomaly found in super-Yang-Mills with local coupling has the same properties as the Adler–Bardeen anomaly: It is determined by convergent one-loop diagrams and its coefficient is gauge- and scheme-independent.

As an application we have constructed the renormalization group equation and the gauge β function. It was shown, that the non-renormalization of the Θ angle first yields vanishing coefficients for the β function in $l \geq 2$. The supersymmetry anomaly induces the 2-loop term in terms of the one-loop coefficient and the anomaly coefficient. Higher order terms are scheme dependent and are determined by finite redefinitions of the coupling. Hence, the non-holomorphic contributions in the β function of pure super-Yang-Mills theories are generated by the supersymmetry anomaly. Since $N = 2$ super-Yang-Mills theories are not anomalous, one can impose the classical Slavnov–Taylor identity and in this case a pure one-loop gauge β function is found.

The construction can be extended also to the matter part [7, 12], and to softly broken gauge theories [8, 23]. Since the soft breakings are the lowest components of Lagrangian multiplets, they are already included in the supersymmetric model

with local coupling and thus softly broken supersymmetry appears as a natural extension of supersymmetric theories with local coupling.

Acknowledgments

I wish to thank the organizers of the workshop, Matilde Marcolli and Kurusch Ebrahimi-Fard, for the invitation and for the organization of this stimulating meeting between mathematicians and physicists. I am also grateful to R. Flume for many valuable discussions on renormalization and non-renormalization theorems.

References

- [1] N.N. Bogolubov and D.V. Shirkov, *Introduction to the theory of quantized fields*, (Interscience, New York, 1969).
- H. Epstein and V. Glaser, *Adiabatic limit in perturbation theory*, in *Renormalization theory*, ed. G. Velo and A.S. Wightman (Reidel, Dordrecht, 1975).
- [2] M.A. Shifman and A.I. Vainshtein, *Nucl. Phys.* **B359** (1991) 571.
- [3] L.Dixon, V. Kaplunovsky and J. Louis, *Nucl. Phys.* **B355** (1991) 649.
- [4] N. Seiberg, *Phys. Lett.* **B318** (1993) 469.
- [5] V. Kaplunovsky and J. Louis, *Nucl. Phys.* **B422** (1994) 57.
- [6] R. Flume and E. Kraus, *Nucl. Phys.* **B569** (2000) 625.
- [7] E. Kraus and D. Stöckinger, *Nucl. Phys.* **B626** (2002) 73, hep-th/0105028.
- [8] E. Kraus and D. Stöckinger, *Phys. Rev.* **D 65** (2001) 115012; hep-ph/0107061.
- [9] K. Fujikawa and W. Lang, *Nucl. Phys.* **B88** (1975) 61.
- [10] M.T. Grisaru, W. Siegel and M. Rocek, *Nucl. Phys.* **B159** (1979) 429; M.T. Grisaru and W. Siegel, *Nucl. Phys.* **B201** (1982) 292.
- [11] M.A. Shifman and A.I. Vainshtein, *Nucl. Phys.* **B277** (1986) 456.
- [12] E. Kraus, *Nucl. Phys.* **B620** (2002) 55; hep-th/0107239.
- [13] E. Kraus *Phys. Rev. D* **65** (2002) 105003; hep-ph/0110323.
- [14] J. Wess and B. Zumino, *Nucl. Phys.* **B78** (1974) 1.
- [15] B. de Wit and D. Freedman, *Phys. Rev.* **D12** (1975) 2286.
- [16] P.L. White, *Class. Quantum Grav.* **9** (1992) 1663.
- [17] N. Maggiore, O. Piguet and S. Wolf, *Nucl. Phys.* **B458** (1996) 403.
- [18] V. Novikov, M. Shifman, A. Vainshtein and V. Zakharov, *Nucl. Phys.* **B229** (1983) 381; *Phys. Lett.* **B166** (1986) 329.
- [19] S.L. Adler, *Phys. Rev.* **177** (1969) 2426.
W.A. Bardeen, *Phys. Rev.* **184** (1969) 1848.
- [20] J.S. Bell and R. Jackiw, *Nuovo Cim.* **60A** (1969) 47.
- [21] S.L. Adler and W.A. Bardeen, *Phys. Rev.* **182** (1969) 1517.
- [22] E. Kraus, *Anomalies in quantum field theory: Properties and characterization*, Proceedings of the Hesselberg workshop 2002, hep-th/0211084.
- [23] E. Kraus and D. Stöckinger, *Phys. Rev. D* **65** (2002) 105014; hep-ph/0201247.

PHYSIKALISCHES INSTITUT, UNIVERSITÄT BONN, NUSSALLEE 12, D-53115 BONN, GERMANY
E-mail address: `kraus@th.physik.uni-bonn.de`

Not so non-renormalizable gravity*

Dirk Kreimer

ABSTRACT. We review recent progress with the understanding of quantum fields, including ideas [1] how gravity might turn out to be a renormalizable theory after all.

1. Introduction

Renormalizable perturbative quantum field theories are embarrassingly successful in describing observed physics. Whilst their mathematical structure is still a challenge albeit an entertaining one, they are testimony to some of the finest achievements in our understanding of nature. The physical law as far as it is insensitive to the surrounding geometry seems completely described by such theories. Alas, if we incorporate gravity, and want to quantize it, we seem at a loss.

In this talk, we report on some recent work [1] which might give hope. Our main purpose is to review the basic idea and to put it into context.

A renormalizable theory poses a computational problem for a theoretical physicist: even if only a finite number of amplitudes need renormalization, the quantum equations of motion -the DysonSchwinger equations (DSE)- ensure that these amplitudes must be calculated as iterated integrals based on a skeleton expansion for the Green functions.

There is an infinite series of skeletons, of growing computational complexity, and thus a formidable challenge at hand. Order is brought to this situation by the fact that the skeletons can be organized in terms of the underlying Hochschild cohomology of the Hopf algebra of a renormalizable theory. The computational challenge remains though in the analytic determination of the skeletons and their Mellin transforms [2, 3, 4]. This approach, combining the analysis of the renormalization group provided in [5] with the analysis of the mathematical structure of DSE provided in [2, 3, 6, 7, 8], has led to new methods in solving DSE beyond perturbation theory [8, 9].

A nice fact is that internal symmetries can be systematically understood in terms of this Hochschild cohomology: Slavnov–Taylor identities are equivalent to the demand that multiplicative renormalization is compatible with the cohomology structure, leading to the identification of Hopf ideals generated by these very Ward and Slavnov–Taylor identities [4, 10].

*Work supported in parts by grant NSF-DMS/0603781. Author supported by CNRS.

For a non-renormalizable theory the situation is worse: the computational challenge for the theorist is repeated infinitely as there is now an infinite number of amplitudes demanding renormalization, each of them still based on an infinite number of possible skeleton iterations. But the interplay with Hochschild cohomology leads to surprising new insights into this situation, which were discussed at an elementary level for the situation of pure gravity in [1]. Note that recent results on the core Hopf algebra [11] give further evidence to the ideas presented here. See also [12] for a summary.

In summary, our analysis suggests that gravity, regarded as a probability conserving but perturbatively non-renormalizable theory, is renormalizable after all, thanks to the structure of its Dyson–Schwinger equations.

As in [1], we will proceed by a comparison of the structure of a renormalizable theory, quantum electrodynamics in four dimensions, and gravity.

It is the role of the Hochschild cohomology [6] in those two different situations which leads to surprising new insights. We will discuss them at an elementary level for the situation of pure gravity. We also allow, in the spirit of the workshops where this material was presented, for the freedom to muse about conceptual consequences at the end.

Acknowledgments. This short contribution is based on talks given in Leipzig (*Recent Developments in QFT*, MPI Math. in the Sciences, Leipzig, July 20-22 2007, B. Fauser, J. Tolksdorf, E. Zeidler, eds.) and Bonn (*Conference on Combinatorics and Physics*, MPI Math., Bonn, March 19-23 2007, K. Ebrahimi-Fard, M. Marcolli, W. van Suijlekom, eds.). It is a pleasure to thank all the organizers for their efforts and hospitality.

2. The structure of Dyson–Schwinger Equations in QED_4

2.1. The Green functions. Quantum electrodynamics in four dimensions of space-time (QED_4) is described in its short-distance behaviour by four Green functions

$$(1) \quad G^{\bar{\psi}\gamma\cdot\partial\psi}, G^{m\bar{\psi}\psi}, G^{\bar{\psi}\gamma\cdot A\psi}, G^{\frac{1}{4}F^2},$$

corresponding to the four monomials in its Lagrangian

$$(2) \quad L = \bar{\psi}\gamma\cdot\partial\psi - \bar{\psi}m\psi - \bar{\psi}\gamma\cdot A\psi - \frac{1}{4}F^2.$$

Here, $G^i = G^i(\alpha, L)$, with α the fine structure constant and $L = \ln q^2/\mu^2$, with subtractions done at $q^2 = \mu^2$ (the MOM scheme in a physicists parlance), so that quantum corrections are constrained to vanish when the kinematical variable q^2 takes the value μ^2 .

We hence work from now in said MOM scheme, subtract at $q^2 = \mu^2$, project the vertex function to its scalar formfactor $G^{\bar{\psi}\gamma\cdot\partial\psi}$ with UV divergences evaluated at zero photon momentum. Similarly the other Green functions are normalized as to be the multiplicative quantum corrections to the tree level monomials above, in momentum space.

In perturbation theory, the degree of divergence of a graph Γ with f external fermion lines and m external photon lines in D dimensions is

$$(3) \quad \omega_D(\Gamma) = \frac{3}{2}f + m - D - (D - 4)(|\Gamma| - 1) \Rightarrow \omega_4(\Gamma) = \frac{3}{2}f + m - 4.$$

This is independent of the loop number for QED₄, $D = 4$, and is a sole function of the number and type of external legs. $\omega_D(\Gamma)$ determines the number of derivatives with respect to masses or external momenta needed to render a graph logarithmically divergent, and hence identifies the top-level residues which drive the iteration of Feynman integrals according to the quantum equations of motion [3].

We define these four Green functions as an evaluation by renormalized Feynman rules (see for example [12] for a review) of a series of one-particle irreducible (1PI) Feynman graphs $\Gamma \in \mathcal{FG}_i$. These series are determined as a fixpoint of the following system in Hochschild cohomology.

$$(4) \quad X^{\bar{\psi}\gamma \cdot \partial\psi} = 1 - \sum_{k=1}^{\infty} \alpha^k B_+^{\bar{\psi}\gamma \cdot \partial\psi, k} (X^{\bar{\psi}\gamma \cdot \partial\psi} Q^{2k}(\alpha)),$$

$$(5) \quad X^{\bar{\psi}\gamma \cdot A\psi} = 1 + \sum_{k=1}^{\infty} \alpha^k B_+^{\bar{\psi}\gamma \cdot A\psi, k} (X^{\bar{\psi}\gamma \cdot A\psi} Q^{2k}(\alpha)),$$

$$(6) \quad X^{\bar{\psi}m\psi} = 1 - \sum_{k=1}^{\infty} \alpha^k B_+^{\bar{\psi}m\psi, k} (X^{\bar{\psi}m\psi} Q^{2k}(\alpha)),$$

$$(7) \quad X^{\frac{1}{4}F^2} = 1 - \sum_{k=1}^{\infty} \alpha^k B_+^{\frac{1}{4}F^2, k} (X^{\frac{1}{4}F^2} Q^{2k}(\alpha)).$$

Here,

$$(8) \quad B_+^{i,k} = \sum_{|\gamma|=k, \Delta'(\gamma)=0, \gamma \in \mathcal{FG}_i} B_+^{\gamma}, \quad \forall i \in \mathcal{R}_{\text{QED}},$$

a sum over all Hopf algebra primitive graphs with given loop number k and contributing to superficially divergent amplitude i , and

$$(9) \quad B_+^{\gamma}(X) = \sum_{\Gamma \in \langle \Gamma \rangle} \frac{\text{bij}(\gamma, X, \Gamma)}{|X|_{\vee}} \frac{1}{\text{maxf}(\Gamma)} \frac{1}{(\gamma|X)} \Gamma,$$

where $\text{maxf}(\Gamma)$ is the number of maximal forests of Γ , $|X|_{\gamma}$ is the number of distinct graphs obtainable by permuting edges of X , $\text{bij}(\gamma, X, \Gamma)$ is the number of bijections of external edges of X with an insertion place in γ such that the result is Γ , and finally $(\gamma|X)$ is the number of insertion places for X in γ [4], and

$$(10) \quad \mathcal{R}_{\text{QED}} = \{\bar{\psi}\gamma \cdot \partial\psi, \bar{\psi}\gamma \cdot A\psi, m\bar{\psi}\psi, \frac{1}{4}F^2\}.$$

Also, we let

$$(11) \quad Q = \frac{X^{\bar{\psi}\gamma \cdot A\psi}}{X^{\bar{\psi}\gamma \cdot \partial\psi} \sqrt{X^{\frac{1}{4}F^2}}}.$$

The resulting maps $B_+^{i,K}$ are Hochschild closed

$$(12) \quad bB_+^{i,K} = 0,$$

in the sense of [2]. We have in fact

$$(13) \quad \Delta(B_+^{\gamma}(X)) = \sum_{\Gamma} n_{\Gamma, X, \gamma} \Gamma$$

where $n_{\Gamma, X, \gamma}$ can be determined from (9,12).

Furthermore, one can choose a basis of primitives γ [3] such that their Mellin transforms $M_\gamma(\rho)$ have the form

$$(14) \quad M_\gamma(\rho) = \int \iota_\gamma(k_i; q) \Big|_{q^2 = \mu^2} \prod_{s=1}^{|\gamma|} \frac{[k_s^2/\mu^2]^{-\rho/|\gamma|} d^4 k_i}{(2\pi)^4} \text{ for } 1 > \Re(\rho) > 0,$$

where the integrand ι_γ is a function of internal momenta k_i and an external momentum q , subtracted at $q^2 = \mu^2$.

The Dyson–Schwinger equations then take the form

$$(15) \quad G_R^i(\alpha, L) = 1 \pm \lim_{\rho \rightarrow 0} \left[\sum_k \alpha^k \sum_{|\gamma|=k} G_r^i(\alpha, \partial_\rho) \mathcal{Q}(\alpha, \partial_\rho) M_\gamma(\rho) \left[\left(\frac{q^2}{\mu^2} \right)^{-\rho} - 1 \right] \right],$$

where

$$(16) \quad \Phi_R(X^i) = G_R^i(\alpha, L),$$

and

$$(17) \quad \Phi_R(Q) = \mathcal{Q}(\alpha, L),$$

is the invariant charge, all calculated with renormalized Feynman rules in the MOM scheme. Note that such subtractions on skeleton kernels not only provide a means to investigate non-perturbative aspects of Green functions [3], but are also well-founded mathematically [13].

2.2. Gauge theoretic aspects. Using Ward identities, we can reduce the set $\mathcal{R}_{\text{QED}} = \{\bar{\psi} \gamma \cdot \partial \psi, m \bar{\psi} \psi, \bar{\psi} \gamma \cdot A \psi, \frac{1}{4} F^2\}$ to three elements upon identifying $G^{\bar{\psi} \gamma \cdot \partial \psi} = G^{\bar{\psi} \gamma \cdot A \psi}$. Using the Baker–Johnson–Willey gauge [14] we can furthermore trivialize

$$(18) \quad G^{\bar{\psi} \gamma \cdot \partial \psi} = G^{\bar{\psi} \gamma \cdot A \psi} = 1.$$

Using their work again [15], we have that $m \bar{\psi} \psi$ can be ignored in \mathcal{R}_{QED} .

We are hence left with the determination of a single gauge-independent Green function $G^{\frac{1}{4} F^2}$ which in the MOM scheme takes the form

$$(19) \quad G^{\frac{1}{4} F^2}(\alpha, L) = 1 - \sum_{k=1}^{\infty} \gamma_k(\alpha) L^k,$$

and the renormalization group determines [7]

$$(20) \quad \gamma_k(\alpha) = \frac{1}{k} \gamma_1(\alpha) (1 - \alpha \partial_\alpha) \gamma_{k-1}(\alpha).$$

Here, $\gamma_1(\alpha) = 2\psi(\alpha)/\alpha$, where $\psi(\alpha)$ is the MOM scheme β -function of QED, which is indeed half of the anomalous dimension γ_1 of the photon field in that scheme.

One can show that $\gamma_1(\alpha)$ as a perturbative series ($\gamma_1(\alpha) = \sum_{j=1}^{\infty} \gamma_{1,j} \alpha^j$) is Gevrey–1 and that the series $\sum_{j=1}^{\infty} \gamma_{1,j} \alpha^j / j!$ has a finite radius of convergence, with a bound involving the lowest order contribution of the β -function and the one-instanton action [3].

Furthermore, $\gamma_1(\alpha)$ fulfills [3]

$$(21) \quad \gamma_1(\alpha) = P(\alpha) - \gamma_1(\alpha) (1 - \alpha \partial_\alpha) \gamma_1(\alpha),$$

an equation which has been studied in detail recently [8], with more of its analytic structure to be exhibited there. In particular, the presence or absence of Landau

poles beyond perturbation theory was clarified in terms of concrete conditions on the asymptotics of $P(\alpha)$ [8]. In this equation, $P(\alpha)$ is obtained from the primitives of the Hopf algebra

$$(22) \quad P(\alpha) = \sum_{\gamma} \alpha^{|\gamma|} \lim_{\rho \rightarrow 0} \rho M_{\gamma}(\rho),$$

and $P(\alpha)$ is known perturbatively as a fifth order polynomial [16] and its asymptotics have been conjectured long ago [17].

This finishes our summary of QED_4 as a typical renormalizable theory.

2.3. Non-abelian gauge theory. The above approach to Green functions remains valid for a non-abelian gauge theory with the definition of a single invariant charge $\mathcal{Q}(\alpha, L)$ being the crucial requirement. This can be consistently done, [4], upon recognizing that the celebrated Slavnov–Taylor identities for the couplings fulfill

$$(23) \quad \frac{S_R^{\phi}(X^{\bar{\psi}\gamma \cdot A\psi})}{S_R^{\phi}(X^{\bar{\psi}\gamma \cdot \partial\psi})} = \frac{S_R^{\phi}(X^{AA\partial A})}{S_R^{\phi}(X^{\partial A\partial A})} = \frac{S_R^{\phi}(X^{AAAA})}{S_R^{\phi}(X^{AA\partial A})} = \frac{S_R^{\phi}(X^{\bar{\phi}A \cdot \partial\phi})}{S_R^{\phi}(X^{\bar{\phi}\square\phi})},$$

for the set of amplitudes

$$(24) \quad \mathcal{R}_{\text{QCD}} = \{DADA, \bar{\psi}\gamma \cdot \partial\psi, \bar{\phi}\square\phi, AADA, AAAA, \bar{\phi}A \cdot \partial\phi, \bar{\psi}\gamma \cdot A\psi\},$$

needing renormalization in QCD.

This allows to define a Hochschild cohomology on the sum of graphs at a given loop order, and hence to obtain multiplicative renormalization in this language from the resulting coideals in the Hopf algebra [4, 10].

Note that the structure of the sub-Hopf algebras underlying this approach [2, 7] implies that the elements $X^i(\alpha)$ close under the coproduct. A general classification of related sub-Hopf algebras has been recently obtained by Loic Foissy [18]. He considers only the case that the lowest order Hochschild cocycle is present in the combinatorial Dyson–Schwinger equations, but his study is rather complete when augmented by the results of [6]. Indeed, this allows to incorporate the next-to-leading order cocycles without any change to the structure of the theory.

3. Gravity

We consider pure gravity understood as a theory based on a graviton propagator and n -graviton couplings as vertices. A fuller discussion incorporating ghosts and matter fields is referred to future work.

3.1. Summary of results of [1].

COROLLARY 1. *Let $|\Gamma| = k$. Then $\omega(\Gamma) = -2(|\Gamma| + 1)$.*

This is a significant change from the behavior of a renormalizable theory: in the renormalizable case, each graph contributing to the same amplitude i has the same powercounting degree regardless of the loop number. Here, we have the dual situation: the loop number determines the powercounting degree, regardless of the amplitude. The following theorem is the main result of [1]. To understand it, note that we first reduce every Green function to an overall logarithmic one, by taking suitable derivatives with respect to external momenta. In the language of Feynman graphs, these derivatives become dots which mark the graphs, and Hopf algebras for such marked graphs are established as before.

THEOREM 2. *The set $d_{\omega(\Gamma)}$ contains no primitive element beyond one loop.*

The set $d_{\omega(\Gamma)}$ is determined as a set of such dotted graphs, with dots representing $\omega(\Gamma)$, those derivatives with respect to masses or external momenta such that the corresponding integrand ι_Γ is overall log-divergent. Whilst in a renormalizable theory, we find for each amplitude in the finite set \mathcal{R} primitives at each loop order in $d_{\omega(\Gamma)}$, here we have an infinite set \mathcal{R} , but only one primitive in it.

PROPOSITION 3. *The relations*

$$(25) \quad \frac{X^{n+1}}{X^n} = \frac{X^n}{X^{n-1}}, \quad n \geq 3,$$

define a sub-Hopf algebra with Hochschild closed one-cocycles $B_+^{1,n}$.

Here, X^n is the sum of all graphs with n external graviton lines. One indeed finds that the combinatorial Dyson–Schwinger equations for gravity provide a sub-Hopf algebra upon requiring these relations, in straightforward generalization of the situation in a non-abelian gauge theory.

Note that for gravity, the Hopf algebra of renormalization agrees with the core Hopf algebra [11]. The BCFW recursion hence obtains a natural role in this approach. It remains some analytic work to be done in the future: understand the off-shell consequences of these combinatorial facts, to achieve an analytic statement concerning renormalizability of gravity.

What is encouraging is that the Feynman rules upon inclusion of ghosts and matter fields still equate the core with the renormalization Hopf algebra, and maintain hence the expected connection to multi-leg BCFW recursions. A detailed discussion can hopefully be given in the future.

3.2. Comments.

3.2.1. *Gauss–Bonnet.* The Gauss–Bonnet theorem ensures here, in the form

$$(26) \quad 0 = \int_{\mathbb{M}} \sqrt{g} (R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma} - 4R_{\alpha\beta} R^{\alpha\beta} + R^2)$$

the vanishing of the one-loop renormalization constants. This does not imply the vanishing of the two-loop renormalization constants as their one-loop subdivergences are off-shell. But it implies that the two-loop counter term has only a first order pole by the scattering type formula, in agreement with the vanishing of $\phi_{\text{off-shell}}(\gamma)\phi_{\text{on-shell}}(\Gamma/\gamma)$. Here, γ , Γ/γ is the decomposition of Γ into one-loop graphs and $\phi_{\text{on/off-shell}}$ denotes suitable Feynman rules.

3.2.2. *Two-loop counterterm.* Also, the universality of the two-loop counterterm suggests that indeed

$$(27) \quad \frac{Z^{\text{gr}_{n+1}}}{Z^{\text{gr}_n}} = \frac{Z^{\text{gr}_n}}{Z^{\text{gr}_{n-1}}}, \quad \text{with } Z^{\text{gr}_n} = S_R^\phi(X^n),$$

holds for off-shell counterterms. In particular, if we compute in a space of constant curvature and conformally reduced gravity which maintains many striking features of asymptotic safe gravity [19, 20], the above identities should hold for suitably defined characters: indeed, in such circumstances we can renormalize using a graviton propagator which is effectively massive with the mass $\sqrt{R/6}$ provided by the constant curvature R , and hence can renormalize at zero external momentum. Using the KLT relations [21], this reduces the above identities to a (cumbersome) combinatorial exercise on one-loop graphs to be worked out in the future.

Continuing this line of thought one expects that a single quantity, the β function of gravity, exhibits short-distance singularities. If this expectation bears out, it certainly is in nice conceptually agreement with the expectation that in theories where gravity has a vanishing β function, gravity is indeed a finite theory [22].

3.2.3. *Other instances of gravity powercounting.* The appearance of Feynman rules such that the powercounting of vertex amplitudes in \mathcal{R}_V cancels the powercounting of propagator amplitudes in \mathcal{R}_E , $\mathcal{R} = \mathcal{R}_V \cup \mathcal{R}_E$, is not restricted to gravity. It indeed appears for example also in the field theoretic description of Brownian fluids and glass possibly, which were recently described at tree-level as a field theory [23]. The dynamics beyond tree level will involve renormalization with powercounting properties similar to the present discussion.

References

- [1] D. Kreimer, *A remark on quantum gravity*, Annals Phys. **323** (2008) 49 [arXiv:0705.3897 [hep-th]].
- [2] D. Kreimer, *Dyson–Schwinger Equations: From Hopf algebras to Number Theory*, in *Universality and Renormalization*, I. Binder, D. Kreimer, eds., Fields Inst Comm. **50** (2007) 225, AMS.
- [3] D. Kreimer and K. Yeats, *An etude in non-linear Dyson–Schwinger equations*, Nucl. Phys. Proc. Suppl. **160** (2006) 116 [arXiv:hep-th/0605096].
- [4] D. Kreimer, *Anatomy of a gauge theory*, Annals Phys. **321** (2006) 2757 [arXiv:hep-th/0509135].
- [5] A. Connes and D. Kreimer, *Renormalization in quantum field theory and the Riemann–Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group*, Commun. Math. Phys. **216** (2001) 215 [arXiv:hep-th/0003188].
- [6] C. Bergbauer and D. Kreimer, *Hopf algebras in renormalization theory: Locality and Dyson–Schwinger equations from Hochschild cohomology*, IRMA Lect. Math. Theor. Phys. **10** (2006) 133 [arXiv:hep-th/0506190].
- [7] D. Kreimer and K. Yeats, *Recursion and growth estimates in renormalizable quantum field theory*, Commun. Math. Phys. **279** (2008) 401; [arXiv:hep-th/0612179].
- [8] D. Kreimer, G. Van Baalen, D. Uminsky, K. Yeats, *The QED beta-function from global solutions to Dyson–Schwinger equations*, Annals of Physics (2008) accepted ms, arXiv:0805.0826 [hep-th]; see also Karen Yeats homepage at <http://math.bu.edu/people/kayeats/papers/picturetalkshort4up.pdf>
- [9] D. Kreimer, G. Van Baalen, D. Uminsky, K. Yeats, *The QCD beta-function from global solutions to Dyson–Schwinger equations*, submitted, arXiv:0906.1754 [hep-th].
- [10] W. D. van Suijlekom, *Renormalization of gauge fields: A Hopf algebra approach*, Commun. Math. Phys. **276** (2007) 773 [arXiv:hep-th/0610137].
- [11] D. Kreimer and W. D. van Suijlekom, *Recursive relations in the core Hopf algebra*, Nucl. Phys. **B 820** [FS] (2009) 682–693; arXiv:0903.2849 [hep-th].
- [12] D. Kreimer, *Algebra for quantum fields*, to appear in Clay Math. Inst. Proceedings, workshop *Motives, Quantum Field Theory, and Pseudodifferential Operators*, Boston University – June 2–13, 2008 arXiv:0906.1851 [hep-th].
- [13] S. Bloch and D. Kreimer, *Mixed Hodge Structures and Renormalization in Physics*, arXiv:0804.4399 [hep-th].
- [14] K. Johnson, M. Baker and R. Willey, *Selfenergy Of The Electron*, Phys. Rev. **136** (1964) B1111.
- [15] M. Baker and K. Johnson, *Asymptotic form of the electron propagator and the selfmass of the electron*, Phys. Rev. D **3** (1971) 2516.
- [16] S. G. Gorishnii, A. L. Kataev, S. A. Larin and L. R. Surguladze, *The Analytical four loop corrections to the QED Beta function in the MS scheme and to the QED psi function: Total reevaluation*, Phys. Lett. B **256** (1991) 81.
- [17] C. Itzykson, G. Parisi and J. B. Zuber, *Asymptotic Estimates In Quantum Electrodynamics*, Phys. Rev. D **16** (1977) 996;

- R. Balian, C. Itzykson, G. Parisi and J. B. Zuber, *Asymptotic Estimates In Quantum Electrodynamics. 2*, Phys. Rev. D **17** (1978) 1041.
- [18] L. Foissy, *Faa di Bruno subalgebras of the Hopf algebra of planar trees from combinatorial Dyson–Schwinger equations*, arXiv:0707.1204 [math.RA].
- [19] M. Reuter, *Nonperturbative Evolution Equation for Quantum Gravity*, Phys. Rev. D **57** (1998) 971 [arXiv:hep-th/9605030];
 O. Lauscher and M. Reuter, *Ultraviolet fixed point and generalized flow equation of quantum gravity*, Phys. Rev. D **65** (2002) 025013 [arXiv:hep-th/0108040];
 M. Niedermaier, *The asymptotic safety scenario in quantum gravity: An introduction*, arXiv:gr-qc/0610018.
- [20] M. Reuter, H. Weyer, *Background Independence and Asymptotic Safety in Conformally Reduced Gravity*, arXiv:0801.3287 [hep-th].
- [21] Z. Bern, D. C. Dunbar and T. Shimada, *String based methods in perturbative gravity*, Phys. Lett. B **312** (1993) 277 [arXiv:hep-th/9307001];
 Z. Bern, L. J. Dixon, D. C. Dunbar, M. Perelstein and J. S. Rozowsky, *On the relationship between Yang-Mills theory and gravity and its implication for ultraviolet divergences*, Nucl. Phys. B **530** (1998) 401 [arXiv:hep-th/9802162];
 Z. Bern and A. K. Grant, *Perturbative gravity from QCD amplitudes*, Phys. Lett. B **457** (1999) 23 [arXiv:hep-th/9904026].
- [22] L. Dixon, *Is $N=8$ supergravity finite?*, available via
<http://www.slac.stanford.edu/lance/Neq8.ppt>.
- [23] A. Velenich, C. Chamon, L. F. Cugliandolo and D. Kreimer, *On the Brownian gas: a field theory with a Poissonian ground state*, J. Phys. A: Math. Theor. **41** 235002 (28pp), arXiv:0802.3212 [cond-mat.stat-mech].

KREIMER@IHES.FR, IHES, 35 RTE. DE CHARTRES, 91440 BURES-SUR-YVETTE, FRANCE AND BOSTON U.

Renormalised multiple zeta values which respect quasi-shuffle relations

Dominique Manchon

This note is a condensed report of a joint work with Sylvie PAYCHA [**MP2**], where details and proofs can be found. We propose a definition of multiple zeta values with integer arguments of any sign, which fulfill the quasi-shuffle relations. They are derived from a character Φ of some quasi-shuffle Hopf algebra with values in the meromorphic functions of a complex variable z , obtained from a suitable regularisation procedure: following Connes and Kreimer, the renormalised values are given by the positive part of the Birkhoff decomposition of Φ evaluated at $z = 0$. The multiple zeta values thus defined are moreover rational at nonpositive arguments.

Acknowledgements

I thank Kurusch Ebrahimi-Fard and Michael E. Hoffman for crucial references and details about quasi-shuffle products, and Herbert Gangl for checking the table of double zeta values at nonpositive arguments with a PARI routine.

Many thanks to Li Guo for various interesting discussions and for bringing to our attention and commenting for us his own work with Bin Zhang ([**GZ**], [**GZ2**]) which also uses a Birkhoff factorisation to renormalise multiple zeta functions. Comments sent by Shlomo Sternberg and Hoang Ngoc Minh are also appreciated.

I have greatly enjoyed the nice and stimulating atmosphere of the conference, and I would like to thank Sylvie Paycha for her careful reading of this note before it reached its final form. Finally I would like to thank the referee for useful remarks and suggestions.

1. The quasi-shuffle Hopf algebra

We recall here the definition of the quasi-shuffle Hopf algebra and the shuffle Hopf algebra built on a commutative (not necessary unital) algebra \mathcal{A} , as well as the explicit construction by M. Hoffman ([**H2**]) of an isomorphism between these two Hopf algebras.

DEFINITION 1. *Let $k, l, r \in \mathbb{N}$ with $k + l - r > 0$. A (k, l) -quasi-shuffle of type r is a surjective map π from $\{1, \dots, k + l\}$ onto $\{1, \dots, k + l - r\}$ such that*

$\pi(1) < \dots < \pi(k)$ and $\pi(k+1) < \dots < \pi(k+l)$. We shall denote by $\text{mix sh}(k, l; r)$ the set of (k, l) -quasi-shuffles of type r . The elements of $\text{mix sh}(k, l; 0)$ are the ordinary (k, l) -shuffles. Quasi-shuffles are also called mixable shuffles or stuffles. We denote by $\text{mix sh}(k, l)$ the set of (k, l) -quasi-shuffles (of any type).

Let (\mathcal{A}, \bullet) a commutative (not necessarily unital) algebra. Let Δ be the deconcatenation coproduct on $\mathcal{T}(\mathcal{A}) = \bigoplus_{k \geq 0} \mathcal{A}^{\otimes k}$, let \star_\bullet the product on $\mathcal{T}(\mathcal{A})$ defined by:

$$(v_1 \otimes \dots \otimes v_k) \star_\bullet (v_{k+1} \otimes \dots \otimes v_{k+l}) = \sum_{\pi \in \text{mix sh}(k, l)} w_1^\pi \otimes \dots \otimes w_{k+l-r}^\pi,$$

with :

$$w_j^\pi = \prod_{i \in \{1, \dots, k+l\}, \pi(i)=j} v_i.$$

(the product above is the product \bullet of \mathcal{A} , and contains only one or two terms).

REMARK 1. When the multiplication \bullet of the algebra \mathcal{A} is set to 0, the quasi-shuffle product \star_\bullet reduces to the ordinary shuffle product III .

THEOREM 1. (M. Hoffman, [H2] theorems 3.1 and 3.3)

- $(\mathcal{T}(\mathcal{A}), \star_\bullet, \Delta)$ is a commutative connected filtered Hopf algebra.
- There is an isomorphism of Hopf algebras :

$$\exp : (\mathcal{T}(\mathcal{A}), \text{III}, \Delta) \xrightarrow{\sim} (\mathcal{T}(\mathcal{A}), \star_\bullet, \Delta).$$

M. Hoffman gives a detailed proof in a slightly more restricted context [H2], which can easily be adapted in full generality (see also [EG]). Hoffman's isomorphism is explicitly built as follows: let $\mathcal{P}(n)$ be the set of compositions of the integer n , i.e. the set of sequences $I = (i_1, \dots, i_k)$ of positive integers such that $i_1 + \dots + i_k = n$. For any $u = v_1 \otimes \dots \otimes v_n \in \mathcal{T}(\mathcal{A})$ and any composition $I = (i_1, \dots, i_k)$ of n we set:

$$I[u] := (v_1 \bullet \dots \bullet v_{i_1}) \otimes (v_{i_1+1} \bullet \dots \bullet v_{i_1+i_2}) \otimes \dots \otimes (v_{i_1+\dots+i_{k-1}+1} \bullet \dots \bullet v_n).$$

We then further define:

$$\exp u = \sum_{I=(i_1, \dots, i_k) \in \mathcal{P}(n)} \frac{1}{i_1! \dots i_k!} I[u].$$

Moreover ([H2], Lemma 2.4), the inverse log of \exp is given by :

$$\log u = \sum_{I=(i_1, \dots, i_k) \in \mathcal{P}(n)} \frac{(-1)^{n-k}}{i_1 \dots i_k} I[u].$$

For example for $v_1, v_2, v_3 \in \mathcal{A}$ we have :

$$\begin{aligned} \exp v_1 &= v_1 \quad , \quad \log v_1 = v_1, \\ \exp(v_1 \otimes v_2) &= v_1 \otimes v_2 + \frac{1}{2}v_1 \bullet v_2 \quad , \quad \log(v_1 \otimes v_2) = v_1 \otimes v_2 - \frac{1}{2}v_1 \bullet v_2, \\ \exp(v_1 \otimes v_2 \otimes v_3) &= v_1 \otimes v_2 \otimes v_3 \quad + \quad \frac{1}{2}(v_1 \bullet v_2 \otimes v_3 + v_1 \otimes v_2 \bullet v_3) \\ &\quad + \frac{1}{6}v_1 \bullet v_2 \bullet v_3, \\ \log(v_1 \otimes v_2 \otimes v_3) &= v_1 \otimes v_2 \otimes v_3 \quad - \quad \frac{1}{2}(v_1 \bullet v_2 \otimes v_3 + v_1 \otimes v_2 \bullet v_3) \\ &\quad + \frac{1}{3}v_1 \bullet v_2 \bullet v_3. \end{aligned}$$

We now choose as commutative algebra \mathcal{A} the space of sequences spanned by $\sigma_s : n \mapsto n^{-s}$, $s \in \mathbb{C}$. This algebra can also be seen as the space of functions on $[1, +\infty[$ spanned by the functions $\sigma_s : t \mapsto t^{-s}$, $s \in \mathbb{Z}$ (we identify by a small abuse of notations the function σ_s and its restriction to the positive integers). The product \bullet on \mathcal{A} will be chosen as the ordinary commutative product of functions, or its opposite.

2. Chen sums and multiple zeta functions

We introduce the *truncated Chen sums*, which are linear forms on $\mathcal{H} = T(\mathcal{A})$ defined by:

$$(1) \quad \sum_{\leq}^{N, \text{Chen}} f_1 \otimes \cdots \otimes f_k := \sum_{1 \leq n_k \leq \cdots \leq n_1 \leq N} f_1(n_1) \cdots f_k(n_k)$$

in its “weak inequality version”, and by:

$$(2) \quad \sum_{<}^{N, \text{Chen}} f_1 \otimes \cdots \otimes f_k := \sum_{1 \leq n_k < \cdots < n_1 < N} f_1(n_1) \cdots f_k(n_k)$$

in its “strict inequality version” for any $f_1, \dots, f_k \in \mathcal{A}$. Provided it exists we call the limit when $N \rightarrow +\infty$ a *Chen sum*. The multiple zeta functions (in both “weak inequality” and “strict inequality” versions) are defined for $\text{Re } s_1 > 1$ and $\text{Re } s_i \geq 1$ for $i = 2, \dots, k$ by the following convergent Chen sums:

$$(3) \quad \bar{\zeta}(s_1, \dots, s_k) := \sum_{\leq}^{\text{Chen}} \sigma_{s_1} \otimes \cdots \otimes \sigma_{s_k} = \sum_{1 \leq n_k \leq n_{k-1} \leq \cdots \leq n_1} n_k^{-s_k} \cdots n_1^{-s_1}$$

$$(4) \quad \zeta(s_1, \dots, s_k) := \sum_{<}^{\text{Chen}} \sigma_{s_1} \otimes \cdots \otimes \sigma_{s_k} = \sum_{1 \leq n_k < n_{k-1} < \cdots < n_1} n_k^{-s_k} \cdots n_1^{-s_1}.$$

We have the following relations between both versions (see [H]):

$$(5) \quad \begin{aligned} \bar{\zeta}(a_1, \dots, a_k) &= \sum_{I=(i_1, \dots, i_r) \in \mathcal{P}(k)} \zeta(b_1^I, \dots, b_r^I), \\ \zeta(a_1, \dots, a_k) &= \sum_{I=(i_1, \dots, i_r) \in \mathcal{P}(k)} (-1)^{k-r} \bar{\zeta}(b_1^I, \dots, b_r^I) \end{aligned}$$

with $b_s^I := a_{i_1+\dots+i_{s-1}+1} + \dots + a_{i_1+\dots+i_s}$. The quasi-shuffle relations between multiple zeta functions can be presented as follows:

THEOREM 2. (1) Consider the commutative algebra (\mathcal{A}, \bullet) where \bullet is the opposite of the ordinary product. For any $f_1, \dots, f_k \in \mathcal{A}$, for any $N \in \mathbb{N}$, the truncated Chen sums fulfill the following relations:

$$(6) \quad \sum_{\leq}^{N, \text{Chen}} (f_1 \otimes \dots \otimes f_k) \star_{\bullet} (g_1 \otimes \dots \otimes g_l) = \left(\sum_{\leq}^{N, \text{Chen}} f_1 \otimes \dots \otimes f_k \right) \left(\sum_{\leq}^{N, \text{Chen}} g_1 \otimes \dots \otimes g_l \right)$$

(2) Whenever the Chen sums converge as $N \rightarrow \infty$, in the limit we have:

$$(7) \quad \sum_{\leq}^{\text{Chen}} (f_1 \otimes \dots \otimes f_k) \star_{\bullet} (g_1 \otimes \dots \otimes g_l) = \left(\sum_{\leq}^{\text{Chen}} f_1 \otimes \dots \otimes f_k \right) \left(\sum_{\leq}^{\text{Chen}} g_1 \otimes \dots \otimes g_l \right)$$

(3) The same statements hold with the strict inequality version provided \bullet is now the ordinary product without minus sign.

Proof:

(1) The domain:

$$P_{k,l} := \{n_1 > \dots > n_k \geq 1\} \times \{n_{k+1} > \dots > n_{k+l} \geq 1\} \subset (\mathbb{N} - \{0\})^{k+l}$$

is partitioned into:

$$P_{k,l} = \coprod_{\pi \in \text{mix sh}(k,l)} P_{\pi},$$

where the domain P_{π} is defined by:

$$P_{\pi} = \{(n_1, \dots, n_{k+l}) / n_{\pi_m} > n_{\pi_p} \text{ if } m > p \text{ and } \pi_m \neq \pi_p, \text{ and } n_m = n_p \text{ if } \pi_m = \pi_p\}.$$

As we must replace strict inequalities by weak ones, let us consider the “closures”

$$\overline{P_{\pi}} := \{(n_1, \dots, n_{k+l}) / n_{\pi_m} \geq n_{\pi_p} \text{ if } m \geq p \text{ and } n_m = n_p \text{ if } \pi_m = \pi_p\}.$$

which then overlap. By the inclusion-exclusion principle we have:

$$(8) \quad \overline{P_{k,l}} = \coprod_{0 \leq r \leq \min(k,l)} (-1)^r \coprod_{\pi \in \text{mix sh}(k,l;r)} \overline{P_{\pi}},$$

where we have set:

$$\overline{P_{k,l}} := \{n_1 \geq \dots \geq n_k \geq 1\} \times \{n_{k+1} \geq \dots \geq n_{k+l} \geq 1\} \subset (\mathbb{N} - \{0\})^{k+l}$$

Each term in equation (8) must be added if r is even, and removed if r is odd. Considering the summation of $f_1 \otimes \dots \otimes f_{k+l}$ over each $\overline{P_{\pi}}$, this decomposition immediately yields the equality:

$$\begin{aligned} & \left(\sum_{1 \leq n_k \leq \dots \leq n_1 \leq N} f_1(n_1) \cdots f_i(n_k) \right) \\ & \times \left(\sum_{1 \leq n_{k+l} \leq \dots \leq n_{k+1} \leq N} f_{k+1}(n_{k+1}) \cdots f_{k+l}(n_{k+l}) \right) = \sum_{\leq}^{N, \text{Chen}} \sum_{\pi \in \text{mix sh}(k,l)} f^{\pi}, \end{aligned}$$

where $f^\pi = f_1^\pi \otimes \cdots \otimes f_{k+l-r}^\pi$ is the tensor product of expressions defined by:

$$f_j^\pi = \bigotimes_{i \in \{1, \dots, k+l\}, \pi(i)=j} f_i.$$

The quasi-shuffle relations (6) are then a reformulation of equality (??) using the commutative algebra (\mathcal{A}, \bullet) .

- (2) Taking the limit as $N \rightarrow \infty$ provides the last statement of the theorem.
- (3) The proof is similar, using the domains P_π rather than the “closures” $\overline{P_\pi}$. As there are no overlappings the signs disappear in Formula (8).

□

3. Regularised Chen sums

We introduce on the algebra \mathcal{A} the *Riesz regularisation procedure* which associates to any element of \mathcal{A} an holomorphic family of elements of \mathcal{A} . It is defined by:

$$(9) \quad \mathcal{R}(\sigma_s)(z) = \sigma_{s+z}.$$

The regularisation \mathcal{R} naturally extends to a morphism on the tensor algebra $T(\mathcal{A})$ via:

$$(10) \quad \tilde{\mathcal{R}}(f_1 \otimes \cdots \otimes f_k) := \mathcal{R}(f_1) \otimes \cdots \otimes \mathcal{R}(f_k).$$

This regularisation is also compatible with the shuffle product. Twisting it with the Hoffman isomorphism \exp gives us a regularisation procedure $\tilde{\mathcal{R}}^* = \exp \circ \tilde{\mathcal{R}} \circ \log$ compatible with the quasi-shuffle product: indeed, using the multiplicativity of \mathcal{R} w.r.to the shuffle product we have:

$$\begin{aligned} \tilde{\mathcal{R}}^*(\sigma \star_\bullet \tau) &= \exp \circ \tilde{\mathcal{R}} \circ \log(\sigma \star_\bullet \tau) \\ &= \exp \circ \tilde{\mathcal{R}}(\log(\sigma) \amalg \log(\tau)) \\ &= \exp \left(\tilde{\mathcal{R}} \circ \log(\sigma) \amalg \tilde{\mathcal{R}} \circ \log(\tau) \right) \\ &= \left(\exp \circ \tilde{\mathcal{R}} \circ \log(\sigma) \right) \star_\bullet \left(\exp \circ \tilde{\mathcal{R}} \circ \log(\tau) \right) \\ &= \tilde{\mathcal{R}}^*(\sigma) \star_\bullet \tilde{\mathcal{R}}^*(\tau). \end{aligned}$$

THEOREM 3. *For any $\sigma \in \mathcal{H} = T(\mathcal{A})$ the Chen sums $\sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma)(z)$ and*

$\sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^(\sigma)(z)$ are holomorphic for $\text{Re } z$ sufficiently large, and extend to the whole complex plane as meromorphic functions of the variable z .*

Sketch of proof: The proof is carried out by induction on the depth of the iterated sum, i.e. the degree of σ in the tensor algebra. The degree 1 case is nothing more than the analytic continuation of the Riemann Zeta function, which can be set up by means of the Euler-MacLaurin formula (see [MP2]). The result in any degree relies on similar properties for *Chen integrals* (see [MP2] and also [MP]), the bridge between Chen integrals and Chen sums being provided by an iterated use of the Euler-MacLaurin formula.

In fact this result is proven in [MP2] in the more general context of classical pseudo-differential symbols on \mathbb{R}^d , allowing higher-dimensional analogues of multiple zeta functions. We briefly report on this point in Section 7. \square

The space of meromorphic functions will be denoted by $\mathcal{M}(\mathbb{C})$. We refer to the meromorphic functions obtained by means of Theorem 3 as *regularised Chen sums*, which we denote by $\sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma)(z)$ and $\sum_{<}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma)(z)$ respectively. By virtue of the compatibility of the regularisation procedure $\tilde{\mathcal{R}}^*$ with the quasi-shuffle product, and by analytic continuation of the quasi-shuffle relations, we have:

PROPOSITION 1. : *The map*

$$\begin{aligned} \mathcal{A} &\rightarrow \mathcal{M}(\mathbb{C}) \\ \sigma &\mapsto \sum_{N \rightarrow +\infty} \sum_{k=1}^N \sigma_z(k) \end{aligned}$$

extends to multiplicative maps

$$\begin{aligned} \Psi^{\mathcal{R}} (\text{ resp. } \Psi'^{\mathcal{R}}) : (\mathcal{T}(\mathcal{A}), \star_{\bullet}) &\rightarrow \mathcal{M}(\mathbb{C}) \\ \sigma &\mapsto \sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma)(z) \quad (\text{ resp. } \sum_{<}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma)(z)) \end{aligned}$$

where \bullet stands for the ordinary product \cdot for $\Psi'^{\mathcal{R}}$, and \star stands for the opposite of the ordinary product for $\Psi^{\mathcal{R}}$. In other words,

$$\begin{aligned} \Psi^{\mathcal{R}}(\sigma_1 \otimes \cdots \otimes \sigma_k)(z) &:= \sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma_1 \otimes \cdots \otimes \sigma_k)(z) \\ \Psi'^{\mathcal{R}}(\sigma_1 \otimes \cdots \otimes \sigma_k)(z) &:= \sum_{<}^{\text{Chen}} \tilde{\mathcal{R}}^*(\sigma_1 \otimes \cdots \otimes \sigma_k)(z) \end{aligned}$$

satisfy the quasi-shuffle relations:

$$\begin{aligned} \Psi^{\mathcal{R}}(\sigma \star_{\bullet} \tau) &= \Psi^{\mathcal{R}}(\sigma) \cdot \Psi^{\mathcal{R}}(\tau), \\ \Psi'^{\mathcal{R}}(\sigma \star_{\bullet} \tau) &= \Psi'^{\mathcal{R}}(\sigma) \cdot \Psi'^{\mathcal{R}}(\tau) \end{aligned}$$

which hold as equalities of meromorphic functions, and where \bullet stands for $\mp \cdot$ as above.

REMARK 2. The results of this paragraph remain valid modulo a small restriction if we replace the Riesz regularisation procedure by more general regularisations: for example, for any holomorphic function $\lambda : \mathbb{C} \rightarrow \mathbb{C}$ one can use the regularisation \mathcal{R}_λ defined by:

$$(11) \quad \mathcal{R}_\lambda(\sigma_s)(z) = \sigma_{s+\lambda(z)}.$$

One can show that the corresponding regularised Chen sums admit an analytic continuation to $\{z \in \mathbb{C}, \lambda'(z) \neq 0\}$. A particular example of this situation is given by $\lambda(z) = \lambda_\mu := z + \mu z^2$ where μ is a constant. The corresponding regularisation $\mathcal{R}_{\lambda_\mu}$

will be denoted by \mathcal{R}_μ . The corresponding regularised Chen sums are then meromorphic on $\mathbb{C} - \{-1/2\mu\}$. In particular they are meromorphic in a neighbourhood of 0.

4. Renormalised Chen sums

The regularised Chen sums introduced in the previous paragraph define a character of the quasi-shuffle Hopf algebra $\mathcal{H} = (T(\mathcal{A}), \star_\bullet)$ (where $\bullet = \mp$ is minus or plus the ordinary product of functions). We can now apply the Birkhoff-Connes-Kreimer decomposition to it:

THEOREM 4. *The map*

$$\begin{aligned} \mathcal{A} &\rightarrow \mathbb{C} \\ \sigma &\mapsto \text{fp} \sum_{z=0} \mathcal{R}(\sigma)(z), \end{aligned}$$

where $\text{fp}_{z=0}$ stands for the constant term in the Laurent expansion at $z = 0$, extends to a multiplicative map:

$$\begin{aligned} \psi^{\mathcal{R}} (\text{ resp. } \psi'^{\mathcal{R}}) : (\mathcal{T}(\mathcal{A}), \star_\bullet), &\longrightarrow \mathbb{C} \\ \sigma &\mapsto \Psi_+^{\mathcal{R}}(\sigma)(0) \quad (\text{ resp. } \Psi_+'^{\mathcal{R}}(\sigma)(0)) \end{aligned}$$

defined from the Birkhoff decomposition of $\Psi^{\mathcal{R}}(\sigma)(z) = \sum_{\leq}^{\text{Chen}} \tilde{\mathcal{R}}^\star(\sigma)(z)$ (respectively $\Psi'^{\mathcal{R}}(\sigma)(z) = \sum_{<}^{\text{Chen}} \tilde{\mathcal{R}}^\star(\sigma)(z)$) with respect to the minimal subtraction scheme. It coincides with the ordinary Chen sums $\sum_{\leq}^{\text{Chen}}$ (resp. $\sum_{<}^{\text{Chen}}$) when the latter converge. Here \bullet stands for the product \mp as in Proposition 1.

Proof: Let us prove the weak inequality case: The Birkhoff decomposition [CK], [M] for the minimal subtraction scheme reads:

$$\Psi^{\mathcal{R}} = (\Psi_-^{\mathcal{R}})^{*^{-1}} * \Psi_+^{\mathcal{R}}$$

with $\bullet = -\cdot$. Since $\Psi_+^{\mathcal{R}}$ is multiplicative on $(\mathcal{T}(\mathcal{A}), \star_\bullet)$, it obeys the quasi-shuffle relation:

$$(12) \quad \Psi_+^{\mathcal{R}}(\sigma \star_\bullet \tau)(z) = \Psi_+^{\mathcal{R}}(\sigma)(z) \Psi_+^{\mathcal{R}}(\tau)(z)$$

which holds as an equality of meromorphic functions holomorphic at $z = 0$. Setting

$$\psi^{\mathcal{R}} := \Psi_+^{\mathcal{R}}(0),$$

and applying the quasi-shuffle relations (12) at $z = 0$ yields

$$(13) \quad \psi^{\mathcal{R}}(\sigma \star_\bullet \tau) = \psi^{\mathcal{R}}(\sigma) \psi^{\mathcal{R}}(\tau).$$

The tensor products $\sigma = \sigma_{s_1} \otimes \cdots \otimes \sigma_{s_k}$ where $\text{Re } s_1 > 1$ and $\text{Re } s_i \geq 1, i = 2, \dots, k$ span a right co-ideal of $\mathcal{T}(\mathcal{A})$. The restriction of $\Psi^{\mathcal{R}}$ to this right co-ideal takes values in functions which are holomorphic at $z = 0$. In that case it follows by construction that:

$$\psi^{\mathcal{R}}(\sigma) = \Psi_+^{\mathcal{R}}(\sigma)(0) = \Psi^{\mathcal{R}}(\sigma)(0) = \sum_{\leq}^{\text{Chen}} \sigma.$$

The strict inequality case can be derived similarly setting $\bullet = \cdot$. \square

On the grounds of this result we set the following definition:

DEFINITION 2. *For any $\sigma \in \mathcal{T}(\mathcal{A})$, the renormalised Chen sums of σ (in both weak and strict inequality versions) are defined by:*

$$\sum_{\leq}^{\text{Chen}, \mathcal{R}} \sigma := \psi^{\mathcal{R}}(\sigma), \quad \sum_{<}^{\text{Chen}, \mathcal{R}} \sigma := \psi'^{\mathcal{R}}(\sigma).$$

By (13) we have:

$$(14) \quad \sum_{\leq}^{\text{Chen}, \mathcal{R}} \sigma *_{\bullet} \tau = \sum_{\leq}^{\text{Chen}, \mathcal{R}} \sigma \sum_{\leq}^{\text{Chen}, \mathcal{R}} \tau$$

for $\bullet = -\cdot$, and similarly for the strict inequality versions with $\bullet = \cdot$.

5. Quasi-shuffle relations for renormalised multiple zeta functions

Recall that multiple zeta functions $\bar{\zeta}(s_1, \dots, s_k)$ and $\zeta(s_1, \dots, s_k)$ defined above converge whenever $\operatorname{Re} s_1 > 1$ and $\operatorname{Re} s_2 \geq 1, \dots, \operatorname{Re} s_k \geq 1$ in which case they obey quasi-shuffle relations. In this section, we implement the renormalisation procedure described in the previous paragraph to extend them to other integer values of s_i while preserving the quasi-shuffle relations.

Let W be the \mathbb{R} -vector space freely spanned by sequences (u_1, \dots, u_k) of real numbers. Let us define the quasi-shuffle product on W by:

$$(15) \quad (u_1, \dots, u_k) \star (u_{k+1}, \dots, u_{k+l}) = \sum_{0 \leq r \leq \min(k, l)} (-1)^r \sum_{\pi \in \text{mix sh}(k, l; r)} (u_1^{\pi}, \dots, u_{k+l-r}^{\pi}),$$

with:

$$u_j^{\pi} = \sum_{i \in \{1, \dots, k+l\}, \pi(i)=j} u_i.$$

(the sum above contains only one or two terms). Define a map $u \mapsto \sigma_u$ from W to $\mathcal{T}(\mathcal{A})$ by:

$$\sigma_{(u_1, \dots, u_k)} := \sigma_{u_1} \otimes \dots \otimes \sigma_{u_k}.$$

Then

$$\sigma_u *_{\bullet} \sigma_v = \sigma_{u \star v}.$$

The same holds with $\bullet = \cdot$ provided we drop the signs $(-1)^r$ in equation (15) defining the quasi-shuffle product on W .

We now apply the results of the previous section to extend the usual quasi-shuffle relations to non converging sums. Given a real number μ , we consider the holomorphic regularisation \mathcal{R}_{μ} defined in Remark 2.

THEOREM 5. *For any real number μ , and any $s_1, \dots, s_k \in \mathbb{C}$ the renormalised multiple zeta values*

$$\bar{\zeta}^{\mu}(s_1, \dots, s_k) := \psi^{\mathcal{R}_{\mu}}(s_1, \dots, s_k), \quad \zeta^{\mu}(s_1, \dots, s_k) := \psi'^{\mathcal{R}_{\mu}}(s_1, \dots, s_k)$$

with $\psi^{\mathcal{R}_{\mu}}, \psi'^{\mathcal{R}_{\mu}}$ as in Definition 2, have the following properties:

(1) *They verify the quasi-shuffle relations:*

$$(16) \quad \bar{\zeta}^\mu(u \star v) = \bar{\zeta}^\mu(u) \bar{\zeta}^\mu(v)$$

when the quasi-shuffle product \star is defined by (15), and:

$$(17) \quad \zeta^\mu(u \star v) = \zeta^\mu(u) \zeta^\mu(v)$$

when the quasi-shuffle product \star is defined by (15) with signs $(-1)^r$ removed.

$$(2) \quad \bar{\zeta}^\mu(s) = \zeta^\mu(s) = \text{fp}_{z=0} \sum k^{-s-f_\mu(z)} \text{ for any } s \in \mathbb{C},$$

(3) *for any positive integer k , and whenever $\text{Re } s_1 > 1$ and $\text{Re } s_j \geq 1$ for $2 \leq j \leq k$,*

$$\begin{aligned} \bar{\zeta}^\mu(s_1, \dots, s_k) &= \sum_{0 < s_1 \leq \dots \leq s_k} n_1^{-s_1} \dots n_k^{-s_k} = \bar{\zeta}(s_1, \dots, s_k), \\ \zeta^\mu(s_1, \dots, s_k) &= \sum_{0 < s_1 < \dots < s_k} n_1^{-s_1} \dots n_k^{-s_k} = \zeta(s_1, \dots, s_k) \end{aligned}$$

independently of μ .

Proof: The proof follows from Theorem 4 applied to \mathcal{R}^μ instead of the Riesz regularisation \mathcal{R} . \square

This is the unique extension of ordinary multiple zeta functions to integer arguments $s_i \geq 1$ which verifies the three properties of Theorem 5. Indeed, the expressions $\bar{\zeta}^\mu(s_1, \dots, s_k)$ and $\zeta^\mu(s_1, \dots, s_k)$ converge whenever $s_1 > 1$ since by assumption, all the s_i are no smaller than 1. When they converge, they obey the quasi-shuffle relations (16) and (17) respectively. The uniqueness of the extension to the case $s_1 = 1$ then follows by induction on the length k from the quasi-shuffle relations (16) which “push” the leading term $s_1 = 1$ whenever it arises, away from the first position and therefore expresses divergent expressions in terms of convergent expressions. This provides the uniqueness of the extension of Riemann multiple zeta functions to regularised multiple zeta functions satisfying quasi-shuffle relations, once the value θ at the argument 1 is imposed; see [H], [W], [Z]. Here it is the parameter μ that plays the role of this constant θ .

More precisely we can easily compute, e.g. by means of the Euler-MacLaurin formula:

$$\begin{aligned} \theta = \zeta^\mu(1) &= \text{fp}_{z=0} \text{fp}_{N \rightarrow +\infty} \sum_{k=1}^N k^{-1-z-\mu z^2} \\ &= - \text{fp}_{z=0} \frac{1}{z} \frac{1}{1+\mu z} (N^{-z-\mu z^2} - 1) \\ &= -\mu. \end{aligned}$$

The other multiple zeta values with 1 on the left are then computed by means of the quasi-shuffle relations, for example:

$$\zeta^\mu(1, 2) = -\mu \zeta(2) - \zeta(2, 1) - \zeta(3).$$

6. Multiple zeta values at nonpositive arguments

Let us sum up the steps of the procedure we implemented to get renormalised values of multiple zeta functions $\bar{\zeta}^\mu(s_1, \dots, s_k)$ which obey quasi-shuffle relations.

- (1) Using a holomorphic regularisation \mathcal{R}^μ we built meromorphic maps

$$z \mapsto \sum_{\leq}^{\text{Chen}} \widetilde{\mathcal{R}^\mu}^*(\sigma_{s_1} \otimes \dots \otimes \sigma_{s_k})$$

which obey the quasi-shuffle relations as identities between meromorphic functions.

- (2) The poles of these meromorphic functions might add superfluous contributions to the finite part and thereby spoil the quasi-shuffle relations when taking finite parts. A Birkhoff factorisation at $z = 0$ was therefore implemented to take care of these extra terms by introducing counterterms. The resulting renormalised multiple zeta values $\bar{\zeta}^\mu(s_1, \dots, s_k)$ indeed obey the quasi-shuffle relations.

Since the meromorphic maps $\sum_{\leq}^{\text{Chen}} \widetilde{\mathcal{R}^\mu}^*(\sigma_{s_1} \otimes \dots \otimes \sigma_{s_k})$ obtained from the holomorphic perturbation turn out to be holomorphic for non negative integers s_i , Birkhoff factorisation is superfluous¹ and will only lead to the same finite parts corresponding to ordinary limits as $z \rightarrow 0$:

$$\bar{\zeta}^\mu(s_1, \dots, s_k) = \text{fp}_{z=0} \widetilde{\mathcal{R}^\mu}^*(\sigma_{s_1} \otimes \dots \otimes \sigma_{s_k}) = \lim_{z \rightarrow 0} \widetilde{\mathcal{R}^\mu}^*(\sigma_{s_1} \otimes \dots \otimes \sigma_{s_k}).$$

We show in [MP2] that for any nonpositive integers s_1, \dots, s_k the multiple zeta values $\bar{\zeta}(s_1, \dots, s_k)$ obtained from Riesz regularisation \mathcal{R} (i.e. for $\mu = 0$) are rational numbers.

The treatment of the strict inequality version $\zeta(s_1, \dots, s_k)$ is completely similar, but we stick to the weak inequality version as it is more easy to handle for an iterated use of the Euler-MacLaurin formula, which is the key ingredient to prove the rationality by induction on the depth. Along the same lines we also outline an algorithm to compute effectively the renormalised multiple zeta values at nonpositive arguments obtained with the Riesz regularisation. For double zeta values at nonpositive arguments we obtain the following formula:

$$(18) \quad \zeta(-a, -b) = \frac{1}{b+1} \sum_{s=0}^{b+1} \binom{b+1}{s} B_s \zeta(-a-b+s-1) + \zeta(-a) \zeta(-b) + (-1)^{a+1} \frac{a!b!}{2(a+b+2)!} B_{a+b+2}.$$

In terms of Bernoulli numbers, this is equivalent to:

$$(19) \quad \zeta(-a, -b) = \frac{1}{b+1} \sum_{s=0}^{b+1} \frac{(-1)^{a+b-s+1}}{a+b-s} \binom{b+1}{s} B_s B_{a+b+2-s} + \frac{(-1)^{a+b}}{(a+1)(b+1)} B_{a+1} B_{b+1} + (-1)^{a+1} \frac{a!b!}{2(a+b+2)!} B_{a+b+2}.$$

¹We refer to [M]; here the tensor algebra is built over the algebra $\mathcal{A}^+ := \{t \mapsto t^s, s \geq 0\}$.

One can then establish a table of values $\zeta(-a, -b)$ for $a, b \in \{0, \dots, 6\}$:

| $\zeta(-a, -b)$ | $a = 0$ | $a = 1$ | $a = 2$ | $a = 3$ | $a = 4$ | $a = 5$ | $a = 6$ |
|-----------------|-------------------|----------------------|-----------------------|---------------------|----------------------------|-----------------------|---------------------------|
| $b = 0$ | $\frac{3}{8}$ | $\frac{1}{12}$ | $\frac{7}{720}$ | $-\frac{1}{120}$ | $-\frac{11}{2520}$ | $\frac{1}{252}$ | $\frac{1}{224}$ |
| $b = 1$ | $\frac{1}{24}$ | $\frac{1}{288}$ | $-\frac{1}{240}$ | $-\frac{19}{10080}$ | $\frac{1}{504}$ | $\frac{41}{20160}$ | $-\frac{1}{480}$ |
| $b = 2$ | $-\frac{7}{720}$ | $-\frac{1}{240}$ | 0 | $\frac{1}{504}$ | $\frac{113}{151200}$ | $-\frac{1}{480}$ | $-\frac{307}{166320}$ |
| $b = 3$ | $-\frac{1}{240}$ | $\frac{1}{840}$ | $\frac{1}{504}$ | $\frac{1}{28800}$ | $-\frac{1}{480}$ | $-\frac{281}{332640}$ | $\frac{1}{264}$ |
| $b = 4$ | $\frac{11}{2520}$ | $\frac{1}{504}$ | $-\frac{113}{151200}$ | $-\frac{1}{480}$ | 0 | $\frac{1}{264}$ | $\frac{117977}{75675600}$ |
| $b = 5$ | $\frac{1}{504}$ | $-\frac{103}{60480}$ | $-\frac{1}{480}$ | $\frac{1}{1232}$ | $\frac{1}{264}$ | $\frac{1}{127008}$ | $-\frac{691}{65520}$ |
| $b = 6$ | $-\frac{1}{224}$ | $-\frac{1}{480}$ | $\frac{307}{166320}$ | $\frac{1}{264}$ | $-\frac{117977}{75675600}$ | $-\frac{691}{65520}$ | 0 |

REMARK 3. The formula for $\zeta(-a, -b)$ coincides with the proposal of **[AET]** at the end of their paper (denoted by $\zeta^*(-b, -a)$ with their notations). This coincidence does not hold in depth ≥ 3 because of the incidence of the Hoffman isomorphism Exp **[MP2]**. If $a + b$ is odd and $b \neq 0$, all terms in equation (18) vanish except the one with $s = 1$. This yields:

$$(20) \quad \zeta(-a, -b) = -\frac{1}{2}\zeta(-a - b),$$

a fact also established by L. Guo and B. Zhang by different means **[GZ]**. We also have for odd a :

$$(21) \quad \zeta(-a, 0) = -\zeta(-a).$$

The values for $\zeta(-a, -a)$ are forced by the quasi-shuffle relations, and hence coincide with those computed in **[GZ]**. Values of $\zeta(-a, -b)$ disagree however for $a + b$ even and a, b distinct. This shows that the quasi-shuffle relations are not sufficient to fully determine multiple zeta values at nonpositive arguments.

7. A higher-dimensional analogue

Considering the d -dimensional vector space \mathbb{R}^d endowed with the supremum norm $|(x_1, \dots, x_d)| := \sup_i |x_i|$ one can consider the d -dimensional analogue of multiple zeta functions:

$$(22) \quad \zeta_d(s_1, \dots, s_k) := \sum_{p_1, \dots, p_k \in \mathbb{Z}^d, 0 < |p_k| < \dots < |p_1|} |p_1|^{-s_1} \cdots |p_k|^{-s_k}.$$

The sum above converges whenever $\operatorname{Re} s_1 > d$ and $\operatorname{Re} s_j \geq d$, $j = 2, \dots, k$. We proved in [MP2], using as commutative algebra \mathcal{A} an algebra of radial pseudo-differential symbols endowed with a suitable product \bullet , that:

- (1) Higher-dimensional multiple zeta functions ζ_d can be renormalised along the same lines as above. The higher-dimensional analogues of the quasi-shuffle relations are reflected in the fact that this procedure gives a character of the associated quasi-shuffle Hopf algebra.
- (2) Riesz-Regularised renormalised higher-dimensional multiple zeta values at nonpositive arguments are again rational numbers.
- (3) Higher-dimensional renormalised multiple zeta fuctions are linear combinations (with positive integer coefficients) of “usual” renormalised multiple zeta fuctions. As an example we have:

$$\begin{aligned} \zeta_1(s) &= 2\zeta(s), \\ \zeta_2(s) &= 8\zeta(s, 0) + 8\zeta(s), \\ \zeta_3(s) &= 48\zeta(s, 0, 0) + 56\zeta(s, 0) + 26\zeta(s). \end{aligned}$$

References

- [BGV] N. Berline, E. Getzler, M. Vergne, *Heat kernels and Dirac operators*, Grundlehren der mathematischen Wissenschaften 298, Springer Verlag (1992)
- [AET] S. Akiyama, S. Egami, Y. Tanigawa, *Analytic continuation of multiple zeta functions and their values at nonpositive integers*, Acta Arithm. 98, 107-116 (2001).
- [C1] P. Cartier, *Fonctions polylogarithmes, nombres polyzeta et groupes pro-unipotents*, Séminaire Bourbaki, Astérisque n. **282** (2002) 137-173
- [C2] P. Cartier, *An introduction to zeta functions*, in “From number theory to physics”, ed. M. Walschmidt, P. Moussa, J.-M. Luck, C. Itzykson Springer Verlag (1992).
- [CM] A. Connes, M. Marcolli, *From Physics to Number theory via Noncommutative Geometry* Frontiers in number theory, physics, and geometry. I, 269-347, Springer, Berlin, 2006
- [CEMP] C. Costermans, H.N. Minh, J.Y. Enjalbert, M. Petitot, *Structure and asymptotic expansion of multiple harmonic sums*, ISSAC Proceedings, Beijing, 24-27 July 2005.
- [CK] A. Connes, D. Kreimer, *Hopf algebras, Renormalisation and Noncommutative Geometry*, Comm. Math. Phys. **199** (1988) 203-242 .
- [EG] K. Ebrahimi-Fard, L. Guo, *Mixable Shuffles, Quasi-shuffles and Hopf Algebras* J. Algebraic Combin. 24, no. 1 (2006) 83-101
- [EGK] K. Ebrahimi-Fard, L. Guo, D. Kreimer, *Integrable renormalization I: the ladder case*, J. Math. Phys. **45**, (2004) 3758-3769; *Integrable renormalization II: the general case*, Ann. Inst. H. Poincaré **6** (2004) 369-395.
- [EGGV] K. Ebrahimi-Fard, J. Gracia-Bondia, L. Guo, J. Varilly, *Combinatorics of renormalization as matrix calculus*, Phys. Lett. B. **19** (2006), 552-558.
- [ENR] M. Espie, J-Ch. Novelli, G. Racinet, *Formal computations about multiple zeta values*, IRMA Lect. Math. Theor. Phys. **3**, Berlin (2003).
- [GZ] L. Guo and B. Zhang, *Renormalization of multiple zeta values*, J. Algebra (to appear), arxiv:math.NT/0606076 (2006).
- [GZ2] L. Guo and B. Zhang, *Differential Birkhoff decomposition and the renormalization of multiple zeta values*, J. Number Theory (to appear) arxiv:0710.0432 (2007).

- [GSW] V. Guillemin, S. Sternberg, J. Weitsman, *The Ehrhart function for symbols*, Surveys in Differential Geometry. Special memorial volume dedicated to S. S. Chern (2006). arXiv:math.CO/0601714v1.
- [Ha] G. Hardy, *Divergent series*, Oxford University Press, 1967.
- [H] M. Hoffman, *Multiple harmonic series*, Pacific J. Math. **152** (1992) 275–290; *The algebra of multiple harmonic series*, Journ. Algebra **194** (1997) 477–495, *The Hopf algebra structure of multiple harmonic sums*, arXiv:math.QA/0406589 (2004).
- [H2] M. Hoffman, *Quasi-shuffle products*, J. Algebraic Combin. **11**, 49-68 (2000)
- [Ka] Ch. Kassel, *Quantum groups*, Graduate Texts in Mathematics, **155**, Springer-Verlag, New York, 1995.
- [K] D. Kreimer, *Chen's iterated integral represents the operator product expansion*, Adv. Theo. Math. Phys. **3** (1999)
- [KSW1] Y. Karshon, S. Sternberg, J. Weitsman, *The Euler-Maclaurin formula for simple integral polytopes*, Proc. Natl. Acad. Sci. USA **100**, no.2 (2003). 426-433.
- [KSW2] Y. Karshon, S. Sternberg, J. Weitsman, *Euler-MacLaurin with remainder for a simple integral polytope*, Duke Math. Journ. **130**, no.3 (2005) 401-434.
- [KV] M. Kontsevich, S. Vishik, *Determinants of elliptic pseudo-differential operators*, Max Planck Institut preprint, 1994.
- [L] M. Lesch, *On the non commutative residue for pseudo-differential operators with log-polyhomogeneous symbols*, Ann. of Global Anal. and Geom. **17** (1998) 151–187.
- [M] D. Manchon, *Hopf algebras, from basics to applications to renormalisation*, Rencontres Mathématiques de Glanon, and arXiv:math. QA/0408405.
- [Mi] H.N. Minh, *Des propriétés structurelles des polylogarithmes aux aspects algorithmiques des sommes harmoniques multiples*, Notes de l'exposé du 12/01/06 au groupe de travail “Polylogarithmes et Polyzeta”, Université Paris VII.
- [MMP] Y. Maeda, D. Manchon, S. Paycha, *Stokes' formulae on classical symbol valued forms and applications*, preprint arXiv:math.DG/0510454 (2005).
- [MP] D. Manchon, S. Paycha, *Shuffle relations for regularised integrals of symbols* Comm. Math. Phys **270** (2007) 13-31
- [MP2] D. Manchon, S. Paycha, *Chen sums of symbols and renormalised multiple zeta values*, submitted, arxiv:math/0702135 (2007).
- [P] S. Paycha, *From heat-operators to anomalies; a walk through various regularization techniques in mathematics and physics*, Emmy Nöther Lectures, Göttingen, 2003 (<http://www.math.uni-goettingen.de>); *Anomalies and regularisation techniques in mathematics an physics*, Lecture Notes, Preprint, Colombia, 2003 (<http://www.lma.univ-bpclermont.fr/~paycha/publications/html>).
- [PS] S. Paycha, S. Scott, *A Laurent expansion for regularised integrals of holomorphic symbols*, Geom. and Funct. Anal. (2) **17** (2007) 491-536
- [W] M. Waldschmidt, *Valeurs zeta mutiples. Une introduction*. Journal de Théorie des Nombres de Bordeaux **12** (2000) 581–595.
- [Z] D. Zagier, *Values of zeta functions and their applications*, in “First European Congress of Mathematics”, Vol. **II**, 497–512, Birkhäuser, 1994; *Multizeta values* (manuscript).
- [Zh] J. Zhao, *Analytic continuation of multiple zeta functions*, Proc. Amer. Math. Soc. 128, 1275–1283 (2000).
- [Zh2] J. Zhao, *renormalization of multiple q -zeta values*, arXiv:math.NT/0612093 (2006)
- [Zu] V.V. Zudilin, *Algebraic relations for multiple zeta values*, Russ. Math. Surveys **58:1** (2003) 3–32.

This page intentionally left blank

Formulas for the Connes-Moscovici Hopf algebra

Frédéric Menous

ABSTRACT. We give explicit formulas for the coproduct and the antipode in the Connes-Moscovici Hopf algebra \mathcal{H}_{CM} . To do so, we first restrict ourselves to a sub-Hopf algebra $\mathcal{H}_{\text{CM}}^1$ containing the nontrivial elements, namely those for which the coproduct and the antipode are nontrivial. There are two ways to obtain explicit formulas. On one hand, the algebra $\mathcal{H}_{\text{CM}}^1$ is isomorphic to the Faà di Bruno Hopf algebra of coordinates on the group of identity-tangent diffeomorphism and computations become easy using substitution automorphisms rather than diffeomorphisms. On the other hand, the algebra $\mathcal{H}_{\text{CM}}^1$ is isomorphic to a sub-Hopf algebra of the classical shuffle Hopf algebra which appears naturally in resummation theory, in the framework of formal and analytic conjugacy of vector fields. Using the very simple structure of the shuffle Hopf algebra, we derive once again explicit formulas for the coproduct and the antipode in $\mathcal{H}_{\text{CM}}^1$.

1. Introduction.

The Connes-Moscovici Hopf algebra \mathcal{H}_{CM} was introduced in [5] in the context of noncommutative geometry. Because of its relation with the Lie algebra of formal vector fields, it was also proved in [5] that its subalgebra $\mathcal{H}_{\text{CM}}^1$ is isomorphic to the Faà di Bruno Hopf algebra of coordinates of identity-tangent diffeomorphisms (see [5],[10]). In the past years, it appeared that this Hopf algebra was strongly related to the Hopf algebras of trees (see [2]) or graphs (see [3],[4]) underlying perturbative renormalization in quantum field theory.

Our aim is to give explicit formulas (announced in [11]) for the coproduct and the antipode in $\mathcal{H}_{\text{CM}}^1$, since only recursive formulas seem to be known.

We remind in section 2 the definition of the Connes-Moscovici Hopf algebra, as well as its properties and links with the Faà di Bruno Hopf algebra and identity-tangent diffeomorphisms (for details, see [5],[10]). The formulas are given in section 3. We present in this section a proof based on the isomorphism between identity-tangent diffeomorphisms and substitution automorphisms which are easier to handle in the computations. These manipulations on substitution automorphisms are very common in J. Ecalle's work on the formal classification of differential equations, vector fields, diffeomorphism... (see [6],[7],[8],[9]). In fact, another proof was sketched in [11], based on Jean Ecalle's "Mould Calculus" that, roughly speaking,

relies on the computation of diffeomorphisms as characters on a shuffle Hopf algebra (see section 4). Sections 5 and 6 give the outlines of this proof which is based on a Hopf morphism from $\mathcal{H}^1 \subset \mathcal{H}_{\text{CM}}$ in a shuffle Hopf algebra.

2. Connes-Moscovici and Faà di Bruno Hopf algebras.

2.1. The Connes-Moscovici Hopf algebra. The Connes-Moscovici Hopf algebra \mathcal{H}_{CM} defined in [5] is the enveloping algebra of the Lie algebra which is the linear span of $Y, X, \delta_n, n \geq 1$ with the relations,

$$(1) \quad [X, Y] = X, [Y, \delta_n] = n\delta_n, [\delta_n, \delta_m] = 0, [X, \delta_n] = \delta_{n+1}$$

for all $m, n \geq 1$. The coproduct Δ in \mathcal{H}_{CM} is defined by

$$(2) \quad \Delta(Y) = Y \otimes 1 + 1 \otimes Y, \Delta(X) = X \otimes 1 + 1 \otimes X + \delta_1 \otimes Y, \Delta(\delta_1) = \delta_1 \otimes 1 + 1 \otimes \delta_1$$

where $\Delta(\delta_n)$ is defined recursively, using equation 1 and the identity

$$(3) \quad \forall h_1, h_2 \in \mathcal{H}_{\text{CM}}, \quad \Delta(h_1 h_2) = \Delta(h_1) \Delta(h_2)$$

The coproduct of X and Y is given, whereas the coproduct of δ_n is nontrivial. Nonetheless, the algebra generated by $\{\delta_n, n \geq 1\}$ is a graded sub-Hopf algebra $\mathcal{H}_{\text{CM}}^1 \subset \mathcal{H}_{\text{CM}}$ where the graduation is defined by

$$(4) \quad \text{gr}(\delta_{n_1} \dots \delta_{n_s}) = n_1 + \dots + n_s$$

As mentioned in [5], the Hopf algebra $\mathcal{H}_{\text{CM}}^1$ is strongly linked to Faà di Bruno Hopf algebra.

2.2. The Faà di Bruno Hopf algebra. Let us consider the group of formal identity tangent diffeomorphisms :

$$G_2 = \{f(x) = x + \sum_{n \geq 1} f_n x^{n+1} \in \mathbf{R}[[x]]\}$$

with, by convention, the product $\mu : G_2 \times G_2 \rightarrow G_2$:

$$\mu(f, g) = g \circ f$$

For $n \geq 0$, the functionals on G_2 defined by

$$a_n(f) = \frac{1}{(n+1)!} (\partial_x^{n+1} f)(0) = f_n \quad a_n : G_2 \rightarrow \mathbf{R}$$

are called de Faà di Bruno coordinates on the group G_2 and $a_0 = 1$ being the unit, they generates a graded unital commutative algebra

$$\mathcal{H}_{\text{FdB}} = \mathbf{R}[a_1, \dots, a_n, \dots] \quad (\text{gr}(a_n) = n)$$

Moreover, the action of these functionals on a product in G_2 defines a coproduct on \mathcal{H}_{FdB} that turns to be a graded connected Hopf algebra (see [10] for details). For $n \geq 0$, the coproduct is defined by

$$(5) \quad a_n \circ \mu = m \circ \Delta(a_n)$$

where m is the usual multiplication in \mathbf{R} , and the antipode reads

$$S \circ a_n = a_n \circ \text{rec}$$

where $\text{rec}(\varphi) = \varphi^{-1}$ is the composition inverse of φ .

For example if $f(x) = x + \sum_{n \geq 1} f_n x^{n+1}$ and $g(x) = x + \sum_{n \geq 1} g_n x^{n+1}$ then if $h = \mu(f, g) = g \circ f$ and $h(x) = x + \sum_{n \geq 1} h_n x^{n+1}$,

$$\begin{aligned} a_0(h) &= 1 = a_0(f)a_0(g) \rightarrow \Delta a_0 = a_0 \otimes a_0 \\ a_1(h) &= f_1 + h_1 \rightarrow \Delta a_1 = a_1 \otimes a_0 + a_0 \otimes a_1 \\ a_2(h) &= f_2 + f_1 g_1 + g_2 \rightarrow \Delta a_2 = a_2 \otimes a_0 + a_1 \otimes a_1 + a_0 \otimes a_2 \end{aligned}$$

As proved in [5] and [10], there exists a Hopf isomorphism between \mathcal{H}_{FdB} and $\mathcal{H}_{\text{CM}}^1$.

2.3. Connes-Moscovici coordinates. Following [5] one can define new functionals on G_2 by $\gamma_0 = a_0 = 1$ (unit) and for $n \geq 1$,

$$\gamma_n(f) = (\partial_x^n \log(f'))(0)$$

These functionals, which may be called the Connes-Moscovici coordinates on G_2 , freely generates the Faà di Bruno Hopf algebra :

$$\mathcal{H}_{\text{FdB}} = \mathbb{R}[a_1, \dots, a_n, \dots] = \mathbb{R}[\gamma_1, \dots, \gamma_n, \dots] \quad \text{gr}(a_n) = \text{gr}(\gamma_n) = n$$

and their coproduct is given by the formula 5. Now, see [5], [2] :

THEOREM 1. *The map Θ defined by $\Theta(\delta_n) = \gamma_n$ is a graded Hopf algebra isomorphism between \mathcal{H}_{FdB} and $\mathcal{H}_{\text{CM}}^1$*

This means that the coproduct and the antipode in $\mathcal{H}_{\text{CM}}^1$ can be rather computed in \mathcal{H}_{FdB} , using the map Θ (see section 3.3 for details on this isomorphism). Unfortunately, if the coproduct and the antipode is well-known for the functionals a_n , using the Faà di Bruno formulas for the composition and the inverse of diffeomorphisms in G_2 , it seems that formulas for the γ_n cannot be easily derived. In order to do so, we will either work with substitution automorphism which are easier to handle than diffeomorphisms (see section 3), or identify \mathcal{H}_{FdB} as a sub-Hopf algebra of a shuffle Hopf algebra and use mould calculus (see sections 4, 5, 6).

3. Formulas in $\mathcal{H}_{\text{CM}}^1$.

3.1. Notations. In the sequel we note

$$\mathcal{N} = \{\mathbf{n} = (n_1, \dots, n_s) \in (\mathbb{N}^*)^s, \quad s \geq 1\}$$

For $\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}$,

$$\|\mathbf{n}\| = n_1 + \dots + n_s, \quad l(\mathbf{n}) = s$$

and, if $n \geq 1$,

$$\mathcal{N}_n = \{\mathbf{n} \in \mathcal{N} \quad ; \quad \|\mathbf{n}\| = n\}$$

For a tuple $\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}$, we note $\mathbf{n}! = n_1! \dots n_s!$. Moreover, $\text{Split}(\mathbf{n})$ is the subset of $\bigcup_{t \geq 1} \mathcal{N}^t$ such that $(\mathbf{n}^1, \dots, \mathbf{n}^t) \in \text{Split}(\mathbf{n})$ if and only if the concatenation of $(\mathbf{n}^1, \dots, \mathbf{n}^t)$ is equal to \mathbf{n} :

$$(6) \quad \text{Split}(\mathbf{n}) = \{(\mathbf{n}^1, \dots, \mathbf{n}^t) \in \mathcal{N}^t, \quad \mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}\}$$

In summation formulas, we will use the fact that

$$(7) \quad \bigcup_{\mathbf{n} \in \mathcal{N}_n} \text{Split}(\mathbf{n}) = \bigcup_{\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}_n} \mathcal{N}_{n_1} \times \dots \times \mathcal{N}_{n_s}$$

so that if f is a function on \mathcal{N} and g is a function on $\bigcup_{t \geq 1} \mathcal{N}^t$, for $n \geq 1$,

$$(8) \quad \begin{aligned} \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_n} \sum_{\mathbf{m}^1 \in \mathcal{N}_{n_1}} f(\mathbf{n}) g(\mathbf{m}^1, \dots, \mathbf{m}^s) = \\ \vdots \\ \sum_{\mathbf{m}^s \in \mathcal{N}_{n_s}} \sum_{\mathbf{n} \in \mathcal{N}_n \atop \mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} f(\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|) g(\mathbf{m}^1, \dots, \mathbf{m}^s) \end{aligned}$$

where $\sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}}$ is the sum over $\text{Split}(\mathbf{n})$.

Finally, for $(\mathbf{n}^1, \dots, \mathbf{n}^t) \in \mathcal{N}^t$ ($t \geq 1$), we set:

$$(9) \quad A(\mathbf{n}^1, \dots, \mathbf{n}^t) = \frac{1}{l(\mathbf{n}^1)! \dots l(\mathbf{n}^t)!} \prod_{i=1}^t \frac{1}{\|\mathbf{n}^i\| + 1}$$

and, for $k \geq 1$,

$$(10) \quad B_k(\mathbf{n}^1, \dots, \mathbf{n}^t) = C_k^{l(\mathbf{n}^t)} \prod_{i=1}^{t-1} C_{\|\mathbf{n}^{i+1}\| + \dots + \|\mathbf{n}^t\| + k}^{l(\mathbf{n}^i)}$$

3.2. Main formulas. We will now prove the following formulas :

THEOREM 2. For $n \geq 1$,

$$(11) \quad \begin{aligned} \Delta(\delta_n) &= \delta_n \otimes 1 + 1 \otimes \delta_n \\ &+ \sum_{\substack{(n_1, \dots, n_{s+1}) \in \mathcal{N}_n \\ s \geq 1}} \frac{n!}{n_1! \dots n_{s+1}!} \alpha_{n_{s+1}}^{n_1, \dots, n_s} \delta_{n_1} \dots \delta_{n_s} \otimes \delta_{n_{s+1}} \end{aligned}$$

where, for $\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}$ ($l(\mathbf{n}) = s$) and $m \geq 1$,

$$(12) \quad \alpha_m^{\mathbf{n}} = \sum_{t=1}^{l(\mathbf{n})} C_m^t \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} A(\mathbf{n}^1, \dots, \mathbf{n}^t)$$

with the convention $C_m^t = \frac{m!}{t!(m-t)!} = 0$ if $t > m$.

For the antipode S :

THEOREM 3. For $n \geq 1$,

$$(13) \quad S(\delta_n) = \sum_{\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}_n} \frac{n!}{n_1! \dots n_s!} \beta^{n_1, \dots, n_s} \delta_{n_1} \dots \delta_{n_s}$$

with $\beta^{n_1} = -1$ and, if $\mathbf{n} = (n_1, \dots, n_{s+1}) \in \mathcal{N}$ ($s \geq 1$),

$$(14) \quad \beta^{n_1, \dots, n_s, n_{s+1}} = \sum_{t=1}^s \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} U_{n_{s+1}}^{\|\mathbf{n}^1\|, \dots, \|\mathbf{n}^t\|} A(\mathbf{n}^1, \dots, \mathbf{n}^t)$$

where, if $\mathbf{m} = (m_1, \dots, m_t) \in \mathcal{N}/\{\emptyset\}$ and $k \geq 1$,

$$(15) \quad U_k^{\mathbf{m}} = \sum_{i=1}^{l(\mathbf{m})} (-1)^{i-1} \sum_{\mathbf{m}^1 \dots \mathbf{m}^i = \mathbf{m}} B_k(\mathbf{m}^1, \dots, \mathbf{m}^i)$$

We will now give a first proof of these formulas. To do so, we will work in the group of substitution automorphisms, isomorphic to G_2 , in which the computations are much more easy to perform.

Note that these formulas were first conjectured and then proved using a Hopf morphism between \mathcal{H}_{CM}^1 and a shuffle Hopf algebra noted $sh(N^*)$ which also relies on substitution automorphisms. We will come back later on this morphism and the afferent proofs.

Before introducing substitution automorphisms, let us look back at the correspondence between FdB coordinates and the CM coordinates on G_2 (see theorem 1).

3.3. Coordinates on G_2 . Let $\varphi(x) = x + \sum_{n \geq 1} \varphi_n x^{n+1} \in G_2$. We have for

$n \geq 1$:

$$(16) \quad a_n(\varphi) = \varphi_n, \quad \gamma_n(\varphi) = (\partial_x^n \log(\varphi'))(0) = f_n$$

If $f(x) = \sum_{n \geq 1} \frac{f_n}{n!} x^n$, then

$$(17) \quad f(x) = \log(\varphi'(x)) \quad \varphi(x) = \int_0^x e^{f(t)} dt$$

For any sequence $(u_n)_{n \geq 1}$, we note

$$(18) \quad \forall \mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}, \quad u_{\mathbf{n}} = u_{n_1} \dots u_{n_s}$$

Using equation 17, we get easily that

$$(19) \quad \begin{aligned} f(x) &= \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}} \frac{(-1)^{l(\mathbf{n})}}{l(\mathbf{n})} (n_1 + 1) \dots (n_s + 1) \varphi_{\mathbf{n}} x^{\|\mathbf{n}\|} \\ \varphi(x) &= x + \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}} \frac{1}{l(\mathbf{n})! n!} \frac{f_{\mathbf{n}}}{\|\mathbf{n}\| + 1} x^{\|\mathbf{n}\| + 1} \end{aligned}$$

and these formulas establish the correspondence between FdB and CM coordinates on G_2 . In order to prove theorems 2 and 3, we need to understand how these coordinates read on φ^{-1} and $\mu(\varphi, \psi) = \psi \circ \varphi$ ($\varphi, \psi \in G_2$). To do so, we will rather work with substitution automorphisms than with diffeomorphism.

3.4. Taylor expansions and substitution automorphisms.

DEFINITION 1. Let \tilde{G}_2 be the set of linear maps from $\mathbf{R}[[x]]$ to $\mathbf{R}[[x]]$ such that

1. For $F \in \tilde{G}_2$, the image $F(x)$ by F of the series x is in G_2 .
2. For any two series A and B in $\mathbf{R}[[x]]$, we have

$$(20) \quad F(A \cdot B) = F(A) \cdot F(B)$$

The elements of \tilde{G}_2 are called substitution automorphisms and

THEOREM 4. \tilde{G}_2 is a group for the composition and the map :

$$\begin{aligned} \tau : \tilde{G}_2 &\rightarrow G_2 \\ F &\mapsto \varphi(x) = F(x) \end{aligned}$$

defines an isomorphism between the groups \tilde{G}_2 and G_2 . Moreover, for $A \in \mathbf{R}[[x]]$,

$$(21) \quad F(A) = A \circ \tau(F)$$

PROOF. If $F \in \tilde{G}_2$, then, thanks to equation 20, for $k \geq 0$,

$$(22) \quad F(x^k) = (F(x))^k = (\tau(F)(x))^k = (\varphi(x))^k$$

thus, for $A(x) = \sum_{k \geq 0} A_k x^k \in \mathbf{R}[[x]]$,

$$(23) \quad \begin{aligned} F(A)(x) &= F\left(\sum_{k \geq 0} A_k x^k\right) \\ &= \sum_{k \geq 0} A_k F(x^k) \\ &= \sum_{k \geq 0} A_k (\varphi(x))^k \\ &= A \circ (\varphi(F))(x) \end{aligned}$$

This proves that τ is injective and for any $\varphi \in G_2$ the map

$$\begin{aligned} F &: \mathbf{R}[[x]] \rightarrow \mathbf{R}[[x]] \\ A &\mapsto A \circ \varphi \end{aligned}$$

is a substitution automorphism of \tilde{G}_2 such that $\tau(F) = \varphi$. The map τ is a bijection. Now, for F and G in \tilde{G}_2 ,

$$(24) \quad \tau(F \circ G)(x) = F(G(x)) = \tau(G) \circ \tau(F)(x) = \mu(\tau(F), \tau(G))(x)$$

and if $H = \tau^{-1}((\tau(F))^{-1})$ then $F \circ H = H \circ F = \text{Id}$. This ends the proof. \square

Using Taylor expansion, we also get formulas for $\tau^{-1}(\varphi)$, $\varphi \in G_2$,

PROPOSITION 1. Let $\varphi(x) = x + \sum_{n \geq 1} \varphi_n x^{n+1} \in G_2$ and $F = \tau^{-1}(\varphi)$, then

$$(25) \quad F = \text{Id} + \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}} \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} x^{\|\mathbf{n}\| + l(\mathbf{n})} \partial_x^{l(\mathbf{n})}$$

This also means that F can be decomposed in homogeneous components :

$$(26) \quad F = \text{Id} + \sum_{n \geq 1} F_n \quad , \quad F_n = \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_n} \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} x^{\|\mathbf{n}\| + l(\mathbf{n})} \partial_x^{l(\mathbf{n})}$$

such that

$$(27) \quad \forall n \geq 1, \quad \forall k \geq 1, \quad \exists c \in \mathbf{R}, \quad F_n(x^k) = c x^{n+k}$$

PROOF. If $\varphi(x) = x + \sum_{n \geq 1} \varphi_n x^{n+1} = x + \bar{\varphi}(x) \in G_2$, then, if $F = \tau^{-1}(\varphi)$, then for $A \in \mathbf{R}[[x]]$,

$$\begin{aligned} F(A)(x) &= A(x + \bar{\varphi}(x)) \\ &= A(x) + \sum_{s \geq 1} \frac{(\bar{\varphi}(x))^s}{s!} A^{(s)}(x) \\ &= A(x) + \sum_{s \geq 1} \sum_{n_1 \geq 1, \dots, n_s \geq 1} \frac{1}{s!} \varphi_{n_1} \dots \varphi_{n_s} x^{n_1 + \dots + n_s + s} A^{(s)}(x) \\ &= \left(\text{Id} + \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}} \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} x^{\|\mathbf{n}\| + l(\mathbf{n})} \partial_x^{l(\mathbf{n})} \right) (A(x)) \end{aligned}$$

\square

The automorphism F can be seen as a differential operator acting on $\mathbf{R}[[x]]$ and from now on we note multiplicatively the action of such operators :

$$(28) \quad F \cdot \varphi = F(\varphi)$$

As this will be of some use later, let us give the following formula : If $\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}$ and $k \geq 1$,

$$(29) \quad \begin{aligned} F_{\mathbf{n}} \cdot x^k &= F_{n_1} \dots F_{n_s} \cdot x^k \\ &= \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} \left(\frac{\varphi_{\mathbf{m}^1} x^{n_1 + l(\mathbf{m}^1)}}{l(\mathbf{m}^1)!} \partial_x^{l(\mathbf{m}^1)} \right) \dots \left(\frac{\varphi_{\mathbf{m}^s} x^{n_s + l(\mathbf{m}^s)}}{l(\mathbf{m}^s)!} \partial_x^{l(\mathbf{m}^s)} \right) \cdot x^k \\ &= \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} B_k(\mathbf{m}^1, \dots, \mathbf{m}^s) \varphi_{\mathbf{m}^1} \dots \varphi_{\mathbf{m}^s} x^{\|\mathbf{n}\| + k} \end{aligned}$$

where

$$B_k(\mathbf{m}^1, \dots, \mathbf{m}^s) = C_k^{l(\mathbf{m}^s)} \prod_{i=1}^{s-1} C_{\|\mathbf{m}^{i+1}\| + \dots + \|\mathbf{m}^s\| + k}^{l(\mathbf{m}^i)}$$

With these results one can already derive formulas for the FdB coordinates on G_2 .

3.5. Formulas in \mathcal{H}_{FdB} .

We recover the usual formulas :

PROPOSITION 2. *We have for $n \geq 1$,*

$$(30) \quad \Delta(a_n) = a_n \otimes 1 + 1 \otimes a_n + \sum_{k=1}^{n-1} \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_k} C_{n-k+1}^{l(\mathbf{n})} a_{\mathbf{n}} \otimes a_{n-k}$$

and

$$(31) \quad S(a_n) = \sum_{\mathbf{n} \in \mathcal{N}_n} \left(\sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} (-1)^s B_1(\mathbf{m}^1, \dots, \mathbf{m}^s) \right) a_{\mathbf{n}}$$

PROOF. Let $\varphi(x) = x + \sum_{n \geq 1} \varphi_n x^{n+1}$ and $\psi(x) = x + \sum_{n \geq 1} \psi_n x^{n+1}$ two elements of G_2 and $\eta = \mu(\varphi, \psi) = \psi \circ \varphi$ with

$$(32) \quad \eta(x) = x + \sum_{n \geq 1} \eta_n x^{n+1}$$

If F , G and H are the substitution automorphisms corresponding to φ , ψ and η , then $H = F \circ G$:

$$(33) \quad \begin{aligned} H &= \text{Id} + \sum_{n \geq 1} H_n \\ &= \left(\text{Id} + \sum_{n \geq 1} F_n \right) \left(\text{Id} + \sum_{n \geq 1} G_n \right) \\ &= \text{Id} + \sum_{n \geq 1} \sum_{k=0}^n F_k G_{n-k} \quad (F_0 = G_0 = \text{Id}) \end{aligned}$$

But for $l \geq 1$, $G_l(x) = \psi_l x^{l+1}$ and then, for $k \geq 1$,

$$\begin{aligned}
 F_k G_l x &= \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_k} \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} x^{\|\mathbf{n}\|+l(\mathbf{n})} \partial_x^{l(\mathbf{n})} (\psi_l x^{l+1}) \\
 (34) \quad &= \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_k} \psi_l \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} \frac{(l+1)!}{(l+1-l(\mathbf{n}))!} x^{\|\mathbf{n}\|+l+1} \\
 &= \left(\sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_k} C_{l+1}^{l(\mathbf{n})} \varphi_{\mathbf{n}} \psi_l \right) x^{k+l+1}
 \end{aligned}$$

and then, for $n \geq 1$,

$$(35) \quad \eta_n = \varphi_n + \psi_n + \sum_{k=1}^{n-1} \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_k} C_{n-k+1}^{l(\mathbf{n})} \varphi_{\mathbf{n}} \psi_{n-k}$$

If now $\tilde{\varphi} = \varphi^{-1}$ and $\tilde{F} = \tau^{-1}(\tilde{\varphi})$, then, as $\tilde{F}F = \text{Id}$ we get

$$(36) \quad \tilde{F} = \text{Id} + \sum_{s \geq 1} (-1)^s F_{n_1} \dots F_{n_s} = \text{Id} + \sum_{\mathbf{n} \in \mathcal{N}} (-1)^{l(\mathbf{n})} F_{\mathbf{n}}$$

but for $\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}$,

$$(37) \quad F_{\mathbf{n}}(x) = \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} B_1(\mathbf{m}^1, \dots, \mathbf{m}^s) \varphi_{\mathbf{m}^1} \dots \varphi_{\mathbf{m}^s} x^{\|\mathbf{n}\|+1}$$

Now

$$(38) \quad \tilde{\varphi}_n = \sum_{\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}_n} (-1)^s \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} B_1(\mathbf{m}^1, \dots, \mathbf{m}^s) \varphi_{\mathbf{m}^1} \dots \varphi_{\mathbf{m}^s}$$

and this gives the attempted result. □

Using the same ideas, we will finally prove theorems 2 and 3

3.6. Proof of Theorems 2 and 3. As before, let $\varphi(x) = x + \sum_{n \geq 1} \varphi_n x^{n+1}$

and $\psi(x) = x + \sum_{n \geq 1} \psi_n x^{n+1}$ two elements of G_2 and $\eta = \mu(\varphi, \psi) = \psi \circ \varphi$ with

$$(39) \quad \eta(x) = x + \sum_{n \geq 1} \eta_n x^{n+1}$$

If

$$\begin{aligned}
 f(x) &= \log(\varphi'(x)) = \sum_{n \geq 1} \frac{f_n}{n!} x^n \quad (f_n = \gamma_n(\varphi)) \\
 (40) \quad g(x) &= \log(\psi'(x)) = \sum_{n \geq 1} \frac{g_n}{n!} x^n \quad (g_n = \gamma_n(\psi)) \\
 h(x) &= \log(\eta'(x)) = \sum_{n \geq 1} \frac{h_n}{n!} x^n \quad (h_n = \gamma_n(\eta))
 \end{aligned}$$

then

$$\begin{aligned}
 h(x) &= \log((\psi \circ \varphi)'(x)) \\
 (41) \quad &= \log(\varphi'(x) \cdot \psi'(\varphi(x))) \\
 &= \log(\varphi'(x)) + (\log \psi') \circ \varphi(x) \\
 &= f(x) + F(g)(x)
 \end{aligned}$$

where F is the substitution automorphism associated to φ . We remind that $F = \text{Id} + \sum_{n \geq 1} F_n$. Because of equation 19,

$$\begin{aligned}
 (42) \quad F_n &= \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_n} \frac{1}{l(\mathbf{n})!} \varphi_{\mathbf{n}} x^{\|\mathbf{n}\| + l(\mathbf{n})} \partial_x^{l(\mathbf{n})} \\
 &= \sum_{\mathbf{n}=(n_1, \dots, n_s) \in \mathcal{N}_n} \frac{1}{l(\mathbf{n})!} \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} \frac{A(\mathbf{m}^1, \dots, \mathbf{m}^s) f_{\mathbf{m}^1} \dots f_{\mathbf{m}^s}}{\mathbf{m}^1! \dots \mathbf{m}^s!} x^{\|\mathbf{n}\| + s} \partial_x^s \\
 &= \sum_{\mathbf{n} \in \mathcal{N}_n} \frac{f_{\mathbf{n}}}{\mathbf{n}!} \sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} A(\mathbf{m}^1, \dots, \mathbf{m}^s) \frac{1}{s!} x^{n+s} \partial_x^s
 \end{aligned}$$

But for $k \geq 1$,

$$(43) \quad F_n \left(\frac{g_k}{k!} x^k \right) = \sum_{\mathbf{n} \in \mathcal{N}_n} \frac{f_{\mathbf{n}} g_k}{\mathbf{n}! k!} \sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} A(\mathbf{m}^1, \dots, \mathbf{m}^s) C_k^s x^{n+k}$$

and we obtain immediately the formula for the coproduct.

Let now $\tilde{\varphi} = \varphi^{-1}$ and

$$(44) \quad \tilde{f}(x) = \log(\tilde{\varphi}'(x)) = \sum_{n \geq 1} \frac{\tilde{f}_n}{n!} x^n \quad (\tilde{f}_n = \gamma_n(\tilde{\varphi}))$$

Since $\tilde{\varphi} \circ \varphi(x) = x$,

$$(45) \quad 0 = \log((\tilde{\varphi} \circ \varphi)'(x)) = f(x) + F \cdot \tilde{f}(x)$$

thus

$$(46) \quad \tilde{f}(x) = -\tilde{F} \cdot f(x) = -f(x) - \sum_{\mathbf{n} \in \mathcal{N}} (-1)^{l(\mathbf{n})} F_{\mathbf{n}}(f)(x)$$

But, once again,

$$\begin{aligned}
 (47) \quad f_{\mathbf{n}, k}(x) &= \sum_{\mathbf{n} \in \mathcal{N}_n} (-1)^{l(\mathbf{n})} F_{\mathbf{n}} \left(\frac{f_k}{k!} x^k \right) \\
 &= \sum_{\mathbf{n} \in \mathcal{N}_n} (-1)^{l(\mathbf{n})} \frac{f_k}{k!} \sum_{\substack{\mathbf{m}^i \in \mathcal{N}_{n_i} \\ 1 \leq i \leq s}} B_k(\mathbf{m}^1, \dots, \mathbf{m}^s) \varphi_{\mathbf{m}^1} \dots \varphi_{\mathbf{m}^s} x^{\|\mathbf{n}\| + k} \\
 &= \sum_{\mathbf{n} \in \mathcal{N}_n} \sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} (-1)^s B_k(\mathbf{m}^1, \dots, \mathbf{m}^s) \varphi_{\mathbf{n}} \frac{f_k}{k!} x^{\|\mathbf{n}\| + k} \\
 &= - \sum_{\mathbf{n} \in \mathcal{N}_n} U_k(\mathbf{n}) \varphi_{\mathbf{n}} \frac{f_k}{k!} x^{\|\mathbf{n}\| + k}
 \end{aligned}$$

Now, replacing $\varphi_{\mathbf{n}}$ as in equation 42,

$$(48) \quad f_{\mathbf{n}, k}(x) = - \sum_{\mathbf{n} \in \mathcal{N}_n} \frac{f_{\mathbf{n}} f_k}{\mathbf{n}! k!} \sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} A(\mathbf{m}^1, \dots, \mathbf{m}^s) U_k(\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|) x^{\|\mathbf{n}\| + k}$$

thus, for $l \geq 1$,

$$\tilde{f}_l = -f_l + \sum_{n=1}^{l-1} \sum_{\mathbf{n} \in \mathcal{N}_n} \frac{l! f_{\mathbf{n}} f_{l-n}}{n!(l-n)!} \sum_{\mathbf{m}^1 \dots \mathbf{m}^s = \mathbf{n}} A(\mathbf{m}^1, \dots, \mathbf{m}^s) U_{l-n}(\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|)$$

and this gives immediately the attempted formula.

This ends the proofs for our formulas but, as we said before, another proof can be derived from Jean Ecalle's "mould calculus" whose core is the use of a Hopf algebra morphism between \mathcal{H}_{FdB} and a shuffle Hopf algebra, in order to simplify the computations.

4. Mould calculus and the shuffle Hopf algebra $\text{sh}(\mathbf{N}^*)$.

Let \mathcal{A}^1 the Lie algebra of formal vector fields generated by the derivations

$$(49) \quad \forall n \geq 1, \quad B_n = x^{n+1} \partial_x$$

Its enveloping algebra $\mathcal{U}(\mathcal{A}^1)$ is a graded Hopf algebra and, see [5], the Hopf algebra $\mathcal{H}_{\text{CM}}^1$ is the graded dual of $\mathcal{U}(\mathcal{A}^1)$. Note that this dual is well-defined as the graded components of $\mathcal{U}(\mathcal{A}^1)$ are vector spaces of finite dimension. If $G(\mathcal{A}^1) \subset \mathcal{U}(\mathcal{A}^1)$ is the group of the group-like elements of $\mathcal{U}(\mathcal{A}^1)$, this is exactly the group of substitution automorphism describe above and it is isomorphic to the group G_2

$$(50) \quad \forall F \in G(\mathcal{A}^1), \forall f \in \mathbf{R}[[x]] \quad F.f = f \circ \varphi, \quad \varphi \in G_2$$

In other terms, $G(\mathcal{A}^1) = \tilde{G}_2$.

We won't give details here but an important idea in mould calculus is to try to compute some substitutions automorphism, with the help of the derivations B_n , as if these derivations were free of any relations.

This suggests to introduce, by analogy with \mathcal{A}^1 , the graded free Lie algebra A^1 generated by a set of primitive elements X_n , $n \geq 1$,

$$(51) \quad \Delta(X_n) = X_n \otimes 1 + 1 \otimes X_n$$

The enveloping algebra $\mathcal{U}(A^1)$ is a Hopf algebra which is also called the concatenation Hopf algebra in combinatorics (see [12]). If the unity is $X_\emptyset = 1$ (\emptyset is the empty sequence), then an element \mathbf{U} of $\mathcal{U}(A^1)$ can be written

$$(52) \quad \begin{aligned} \mathbf{U} &= U^\emptyset X_\emptyset + \sum_{s \geq 1} \sum_{n_1, \dots, n_s \geq 1} U^{n_1, \dots, n_s} X_{n_1} \dots X_{n_s} \\ &= U^\emptyset X_\emptyset + \sum_{s \geq 1} \sum_{n_1, \dots, n_s \geq 1} U^{n_1, \dots, n_s} X_{n_1, \dots, n_s} \\ &= \sum U^\bullet X_\bullet \end{aligned}$$

where the collection of coefficients U^\bullet is called a *mould*. The structure of the enveloping algebra $\mathcal{U}(A^1)$ can be described as follows : the product is given by

$$(53) \quad \forall \mathbf{m}, \mathbf{n} \in \mathcal{N}, \quad X_{\mathbf{m}} X_{\mathbf{n}} = X_{\mathbf{m}\mathbf{n}} \quad (\text{concatenation}),$$

the coproduct is

$$(54) \quad \Delta(X_{\mathbf{n}}) = \sum_{\mathbf{n}^1, \mathbf{n}^2} \text{sh}_{\mathbf{n}}^{\mathbf{n}^1, \mathbf{n}^2} X_{\mathbf{n}^1} \otimes X_{\mathbf{n}^2}$$

where $\text{sh}_{\mathbf{n}}^{\mathbf{n}^1, \mathbf{n}^2}$ is the number of shuffling of the sequences $\mathbf{n}^1, \mathbf{n}^2$ that gives \mathbf{n} . Finally, the antipode S is defined by

$$(55) \quad S(X_{n_1, \dots, n_s}) = (-1)^s X_{n_s, \dots, n_1}$$

Once again one can define the group $G(A^1)$ and if $\mathbf{F} \in G(A^1)$ then

$$(56) \quad \mathbf{F} = \sum_{\mathbf{n} \in \mathcal{N} \cup \{\emptyset\}} F^{\mathbf{n}} X_{\mathbf{n}}$$

where the mould F^\bullet is *symmetral* : $F^\emptyset = 1$ and

$$(57) \quad \forall \mathbf{n}^1, \mathbf{n}^2, \quad F^{\mathbf{n}^1} F^{\mathbf{n}^2} = \sum_{\mathbf{n}} \text{sh}_{\mathbf{n}}^{\mathbf{n}^1, \mathbf{n}^2} F^{\mathbf{n}}$$

Moreover, if \mathbf{G} is the group inverse of \mathbf{F} , then its associated mould is given by the formulas

$$G^{n_1, \dots, n_s} = (-1)^s F^{n_s, \dots, n_1}$$

Thanks to the graduation on $\mathcal{U}(A^1)$, its dual H^1 is a Hopf algebra, the Hopf algebra of coordinates on $G(A^1)$ and, if the dual basis of $\{X_{\mathbf{n}}, \quad \mathbf{n} \in \mathcal{N}\}$ is $\{Z^{\mathbf{n}}, \quad \mathbf{n} \in \mathcal{N}\}$ then the product in H^1 is defined by :

$$(58) \quad \forall \mathbf{n}^1, \mathbf{n}^2, \quad Z^{\mathbf{n}^1} Z^{\mathbf{n}^2} = \sum_{\mathbf{n}} \text{sh}_{\mathbf{n}}^{\mathbf{n}^1, \mathbf{n}^2} Z^{\mathbf{n}}$$

The coproduct is :

$$(59) \quad \Delta(Z^{\mathbf{n}}) = Z^{\mathbf{n}} \otimes 1 + 1 \otimes Z^{\mathbf{n}} + \sum_{\mathbf{n}^1 \mathbf{n}^2 = \mathbf{n}} Z^{\mathbf{n}^1} \otimes Z^{\mathbf{n}^2}$$

where $\mathbf{n}^1 \mathbf{n}^2$ is the concatenation of the two nonempty sequences \mathbf{n}^1 and \mathbf{n}^2 and $Z^\emptyset = 1$ is the unity. Finally, the antipode is given by

$$(60) \quad S(Z^{n_1, \dots, n_s}) = (-1)^s Z^{n_s, \dots, n_1}$$

The structure of H^1 (coproduct, antipode, ...) is fully explicit. This will be of great use since there exists a surjective morphism from A^1 on \mathcal{A}^1 that induces an injective morphism from $\mathcal{H}_{\text{CM}}^1$ into H^1 . In other words, $\mathcal{H}_{\text{CM}}^1$ can be identified to a sub-Hopf algebra of H^1 and, as everything is explicit in H^1 , one can derive formulas for the coproduct and the antipode in $\mathcal{H}_{\text{CM}}^1$.

5. Morphisms.

The linear map defined by $\rho(X_n) = B_n = x^{n+1} \partial_x$ obviously determines a morphism from A^1 (resp. $\mathcal{U}(A^1)$, resp. $G(A^1)$) on \mathcal{A}^1 (resp. $\mathcal{U}(\mathcal{A}^1)$, resp. $G(\mathcal{A}^1) \simeq G_2$) and it is surjective. By duality, it induces a morphism ρ^* from \mathcal{H}^1 to H^1 by

$$(61) \quad \forall \gamma \in \mathcal{H}^1, \quad \rho^*(\gamma) = \gamma \circ \rho$$

and, since ρ is surjective, ρ^* is injective : $\mathcal{H}_{\text{CM}}^1$ is isomorphic to the sub-Hopf algebra $\rho^*(\mathcal{H}_{\text{CM}}^1) \subset H^1$. Using this injective morphism, we define

$$(62) \quad \forall n \geq 1, \quad \Gamma_n = \rho^*(\gamma_n)$$

and $\rho^*(\mathcal{H}_{\text{CM}}^1)$ is then the Hopf algebra generated by the Γ_n . In order to get formulas in $\mathcal{H}_{\text{CM}}^1$, we will use the algebra $\rho^*(\mathcal{H}_{\text{CM}}^1)$ and express the Γ_n in terms of the $Z^{\mathbf{n}}$:

THEOREM 5. For $n \geq 1$,

$$(63) \quad \begin{aligned} \Gamma_n &= n! \sum_{\mathbf{n} \in \mathcal{N}_n} \sum_{t=1}^{l(\mathbf{n})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} Z^{\mathbf{n}^1} \dots Z^{\mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= n! \sum_{\mathbf{n} \in \mathcal{N}_n} Q^{\mathbf{n}} Z^{\mathbf{n}} \end{aligned}$$

where $S^{n_1, \dots, n_s} = \prod_{i=1}^s (n_i + n_{i+1} + \dots + n_s + 1) = \prod_{i=1}^s (\hat{n}_i + 1)$, $Q^{n_1} = (n_1 + 1)$ and $Q^{n_1, \dots, n_s} = (n_s + 1) \prod_{i=2}^s \hat{n}_i$ if $s \geq 2$.

Let $\mathbf{F} = X_\emptyset + \sum_{\mathbf{n} \in \mathcal{N}} F^{\mathbf{n}} X_{\mathbf{n}} \in G(A^1)$. If $F = \rho(\mathbf{F}) \in G(\mathcal{A}^1)$, then

$$(64) \quad \Gamma_n(\mathbf{F}) = \gamma_n(F) = \gamma_n(\varphi) = (\partial_x^n \log(\varphi')(x))_{x=0}$$

where $\varphi \in G_2$ is defined by :

$$(65) \quad \varphi(x) = \rho(\mathbf{F}).x = F.x = x + \sum_{(n_1, \dots, n_s) \in \mathcal{N}} F^{n_1, \dots, n_s} B_{n_1} \dots B_{n_s} x$$

Then

$$(66) \quad \varphi'(x) = 1 + \sum_{(n_1, \dots, n_s) \in \mathcal{N}} F^{n_1, \dots, n_s} S^{n_1, \dots, n_s} x^{n_1 + \dots + n_s}$$

Using the logarithm and derivation, one easily gets the formula

$$(67) \quad \Gamma_n(\mathbf{F}) = n! \sum_{\mathbf{n} \in \mathcal{N}_n} \sum_{t=1}^{l(\mathbf{n})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} F^{\mathbf{n}^1} \dots F^{\mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t}$$

We prove the second part of the formula in section 6, using the fact that F^\bullet is symmetrical. As

$$(68) \quad \Gamma_n(\mathbf{F}) = n! \sum_{\mathbf{n} = (n_1, \dots, n_s) \in \mathcal{N}_n} F^{\mathbf{n}} Q^{\mathbf{n}}$$

and $Z^{\mathbf{n}} \cdot \mathbf{F} = F^{\mathbf{n}}$, theorem 5 will be proved.

For $n \geq 1$, $\rho^*(\delta_n) = \Gamma_n \in H^1$, and, since the coproduct and the antipode are explicit in H^1 , we will once again obtain the formulas given in theorems 2 and 3.

6. A second proof for theorems 2 and 3.

6.1. Proof of theorem 5. Extending the notion of shuffling, for $t \geq 1$, if $\mathbf{m}^1, \dots, \mathbf{m}^t, \mathbf{m}$ are $t+1$ sequences, let $\text{sh}_{\mathbf{m}}^{\mathbf{m}^1, \dots, \mathbf{m}^t}$ be the number of ways to obtain the sequence \mathbf{m} by shuffling the sequences $\mathbf{m}^1, \dots, \mathbf{m}^t$. Then, using equation 67,

$$(69) \quad \begin{aligned} \frac{1}{n!} \Gamma_n &= \sum_{\mathbf{n} \in \mathcal{N}_n} \sum_{t=1}^{l(\mathbf{n})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} Z^{\mathbf{n}^1} \dots Z^{\mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= \sum_{\mathbf{n} \in \mathcal{N}_n} \sum_{t=1}^{l(\mathbf{n})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1 \dots \mathbf{n}^t = \mathbf{n}} \left(\sum_{\mathbf{m}} \text{sh}_{\mathbf{m}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} Z^{\mathbf{m}} \right) S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= \sum_{\mathbf{m} \in \mathcal{N}_n} \left(Z^{\mathbf{m}} \sum_{t=1}^{l(\mathbf{m})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t \in \mathcal{N}} \text{sh}_{\mathbf{m}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \right) \end{aligned}$$

For a given sequence $\mathbf{m} \in \mathcal{N}$, let

$$(70) \quad Q^{\mathbf{m}} = \sum_{t=1}^{l(\mathbf{m})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t \in \mathcal{N}} \text{sh}_{\mathbf{m}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t}$$

it remains to prove that, if $\mathbf{m} = (m_1, \dots, m_s)$ then $Q^{m_1, \dots, m_s} = (m_s + 1) \prod_{i=2}^s \hat{m}_i$ with $Q^{m_1} = (m_1 + 1)$. We prove this formula by induction on $l(\mathbf{m})$.

If $l(\mathbf{m}) = 1$, then $\mathbf{m} = (m_1)$ and

$$(71) \quad Q^{m_1} = \frac{(-1)^0}{1} \sum_{\mathbf{n}^1 \in \mathcal{N}} \text{sh}_{\mathbf{m}}^{\mathbf{n}^1} S^{\mathbf{n}^1} = S^{m_1} = m_1 + 1$$

If $l(\mathbf{m}) = s \geq 2$, then let $\mathbf{m} = (m_1, \dots, m_s)$ and $\mathbf{p} = (m_2, \dots, m_s)$. For any sequence $\mathbf{n} = (n_1, \dots, n_k)$, we note $m_1 \mathbf{n} = (m_1, n_1, \dots, n_k)$. If a shuffling of $t \geq 1$ sequences $\mathbf{n}^1, \dots, \mathbf{n}^t$ gives \mathbf{m} then

- Either there exists $1 \leq i \leq t$ such that $\mathbf{n}^i = (m_1)$ (but then $t \geq 2$), then $S^{\mathbf{n}^i} = m_1 + 1$ and, omitting $\mathbf{n}^i = (m_1)$, the corresponding shuffling of the $t - 1$ remaining sequences gives \mathbf{p} .
- Either there exists $1 \leq i \leq t$ such that $\mathbf{n}^i = m_1 \tilde{\mathbf{n}}^i$ ($\tilde{\mathbf{n}}^i \neq \emptyset$) (necessarily, $t < l(\mathbf{m})$), then $S^{\mathbf{n}^i} = (m_1 + \|\tilde{\mathbf{n}}^i\| + 1) S^{\tilde{\mathbf{n}}^i}$ and, replacing \mathbf{n}^i by $\tilde{\mathbf{n}}^i$, the corresponding shuffling of the t sequences gives \mathbf{p} .

This means that :

$$(72) \quad \begin{aligned} Q^{\mathbf{m}} &= \sum_{t=1}^{l(\mathbf{m})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t} \text{sh}_{\mathbf{m}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= \sum_{t=1}^{l(\mathbf{m})-1} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t} \text{sh}_{\mathbf{p}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} \sum_{i=1}^t S^{\mathbf{n}^1} \dots S^{m_1 \mathbf{n}^i} \dots S^{\mathbf{n}^t} + \\ &\quad \sum_{t=2}^{l(\mathbf{m})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^{t-1}} \text{sh}_{\mathbf{p}}^{\mathbf{n}^1, \dots, \mathbf{n}^{t-1}} \sum_{i=0}^{t-1} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^i} S^{m_1} S^{\mathbf{n}^{i+1}} \dots S^{\mathbf{n}^{t-1}} \end{aligned}$$

but as $S^{m_1} = m_1 + 1$ and $S^{m_1 \mathbf{n}^i} = (m_1 + \|\mathbf{n}^i\| + 1) S^{\mathbf{n}^i}$, after a reindexation of the second term we get :

$$(73) \quad \begin{aligned} Q^{\mathbf{m}} &= \sum_{t=1}^{l(\mathbf{m})-1} \frac{(-1)^{t-1}}{t} (t(m_1 + 1) + \|\mathbf{p}\|) \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t} \text{sh}_{\mathbf{p}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &\quad + \sum_{t=1}^{l(\mathbf{m})-1} \frac{(-1)^t}{t+1} (t+1)(m_1 + 1) \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t} \text{sh}_{\mathbf{p}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= \|\mathbf{p}\| \sum_{t=1}^{l(\mathbf{p})} \frac{(-1)^{t-1}}{t} \sum_{\mathbf{n}^1, \dots, \mathbf{n}^t} \text{sh}_{\mathbf{p}}^{\mathbf{n}^1, \dots, \mathbf{n}^t} S^{\mathbf{n}^1} \dots S^{\mathbf{n}^t} \\ &= \|\mathbf{p}\| Q^{\mathbf{p}} \end{aligned}$$

And it obviously gives the right formula for $Q^{\mathbf{m}}$.

6.2. Proof of theorem 2. Using the above formula we have

$$\begin{aligned}
\Delta\Gamma_n &= n! \sum_{\mathbf{m} \in \mathcal{N}_n} Q^{\mathbf{m}}(\Delta Z^{\mathbf{m}}) \\
&= n! \sum_{\mathbf{m} \in \mathcal{N}_n} Q^{\mathbf{m}} \left(Z^{\mathbf{m}} \otimes 1 + 1 \otimes Z^{\mathbf{m}} + \sum_{\mathbf{p}, \mathbf{q} = \mathbf{m}} Z^{\mathbf{p}} \otimes Z^{\mathbf{q}} \right) \\
(74) \quad &= \left(n! \sum_{\mathbf{m} \in \mathcal{N}_n} Q^{\mathbf{m}} Z^{\mathbf{m}} \right) \otimes 1 + 1 \otimes \left(n! \sum_{\mathbf{m} \in \mathcal{N}_n} Q^{\mathbf{m}} Z^{\mathbf{m}} \right) \\
&\quad + n! \sum_{\mathbf{m} \in \mathcal{N}_n} \sum_{\mathbf{p}, \mathbf{q} = \mathbf{m}} Q^{\mathbf{m}} Z^{\mathbf{p}} \otimes Z^{\mathbf{q}} \\
&= \Gamma_n \otimes 1 + 1 \otimes \Gamma_n + n! \sum_{\mathbf{m} \in \mathcal{N}_n} \sum_{\mathbf{p}, \mathbf{q} = \mathbf{m}} Q^{\mathbf{m}} Z^{\mathbf{p}} \otimes Z^{\mathbf{q}} \\
&= \Gamma_n \otimes 1 + 1 \otimes \Gamma_n + \tilde{\Delta}\Gamma_n
\end{aligned}$$

Now if $\mathbf{p}, \mathbf{q} = \mathbf{m} = (m_1, \dots, m_s)$ with $\mathbf{p}, \mathbf{q} \in \mathcal{N}$ ($s \geq 2$), then

$$(75) \quad \frac{Q^{\mathbf{m}}}{Q^{\mathbf{q}}} = \frac{(m_s + 1) \prod_{i=2}^s \hat{m}_i}{(m_s + 1) \prod_{i=l(\mathbf{p})+2}^s \hat{m}_i} = \prod_{i=2}^{l(\mathbf{p})+1} \hat{m}_i = \prod_{i=2}^{l(\mathbf{p})+1} (\hat{p}_i + \|\mathbf{q}\|) = R_{\|\mathbf{q}\|}^{\mathbf{p}}$$

with the convention that if $i = l(\mathbf{p}) + 1$, then $\hat{p}_i = 0$. As this coefficient only depends on \mathbf{p} and $\|\mathbf{q}\|$,

$$\begin{aligned}
\tilde{\Delta}\Gamma_n &= n! \sum_{\mathbf{m} \in \mathcal{N}_n} \sum_{\mathbf{p}, \mathbf{q} = \mathbf{m}} Q^{\mathbf{m}} Z^{\mathbf{p}} \otimes Z^{\mathbf{q}} \\
&= n! \sum_{\mathbf{m} \in \mathcal{N}_n} \sum_{\mathbf{p}, \mathbf{q} = \mathbf{m}} (R_{\|\mathbf{q}\|}^{\mathbf{p}} Z^{\mathbf{p}}) \otimes Q^{\mathbf{q}} Z^{\mathbf{q}} \\
(76) \quad &= n! \sum_{k=1}^{n-1} \left(\sum_{\mathbf{p} \in \mathcal{N}_{n-k}} R_{\|\mathbf{q}\|}^{\mathbf{p}} Z^{\mathbf{p}} \right) \otimes \left(\sum_{\mathbf{q} \in \mathcal{N}_k} Q^{\mathbf{q}} Z^{\mathbf{q}} \right) \\
&= \sum_{k=1}^{n-1} \left(\frac{n!}{k!} \sum_{\mathbf{p} \in \mathcal{N}_k} R_{\|\mathbf{q}\|}^{\mathbf{p}} Z^{\mathbf{p}} \right) \otimes \Gamma_k \\
&= \sum_{k=1}^{n-1} P_k^n \otimes \Gamma_k
\end{aligned}$$

and it remains to prove that, for $n \geq 1$ and $1 \leq k \leq n-1$,

$$(77) \quad P_k^n = \sum_{\substack{(n_1, \dots, n_s) \in \mathcal{N} \\ n_1 + \dots + n_s = n-k, s \geq 1}} \frac{n!}{n_1! \dots n_s! k!} \alpha_k^{n_1, \dots, n_s} \Gamma_{n_1} \dots \Gamma_{n_s}$$

with

$$(78) \quad \alpha_k^{\mathbf{n}} = \sum_{t=1}^{l(\mathbf{n})} C_k^t \sum_{\substack{\mathbf{n}^1, \dots, \mathbf{n}^t = \mathbf{n} \\ \mathbf{n}^i \neq \emptyset}} \frac{1}{l(\mathbf{n}^1)! \dots l(\mathbf{n}^t)!} \prod_{i=1}^t \frac{1}{\|\mathbf{n}^i\| + 1}$$

Using theorem 5 and the shuffle product, we get

$$\begin{aligned}
 (79) \quad \tilde{P}_k^n &\stackrel{\text{def}}{=} \sum_{(n_1, \dots, n_s) \in \mathcal{N}_{n-k}} \frac{n!}{n_1! \dots n_s! k!} \alpha_k^{n_1, \dots, n_s} \Gamma_{n_1} \dots \Gamma_{n_s} \\
 &= \frac{n!}{k!} \sum_{(n_1, \dots, n_s) \in \mathcal{N}_{n-k}} \alpha_k^{n_1, \dots, n_s} \sum_{\mathbf{m}^i \in \mathcal{N}_{n_i}} Q^{\mathbf{m}^1} \dots Q^{\mathbf{m}^s} Z^{\mathbf{m}^1} \dots Z^{\mathbf{m}^s} \\
 &= \frac{n!}{k!} \sum_{\mathbf{p} \in \mathcal{N}_{n-k}} Z^{\mathbf{p}} \sum_{s \geq 1} \sum_{\mathbf{m}^1, \dots, \mathbf{m}^s} \alpha_k^{\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|} \text{sh}_{\mathbf{p}}^{\mathbf{m}^1, \dots, \mathbf{m}^s} Q^{\mathbf{m}^1} \dots Q^{\mathbf{m}^s}
 \end{aligned}$$

It remains to prove that for a given $\mathbf{p} \in \mathcal{N}_{n-k}$, we have

$$\tilde{R}_k^{\mathbf{p}} = \sum_{s=1}^{l(\mathbf{p})} \sum_{\mathbf{m}^1, \dots, \mathbf{m}^s} \alpha_k^{\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|} \text{sh}_{\mathbf{p}}^{\mathbf{m}^1, \dots, \mathbf{m}^s} Q^{\mathbf{m}^1} \dots Q^{\mathbf{m}^s} = R_k^{\mathbf{p}} = \prod_{i=2}^{l(\mathbf{p})+1} (\hat{p}_i + k)$$

The end of the proof follows the same lines as the previous one. If $l(\mathbf{p}) = 1$ then $R_k^{p_1} = k$ and

$$(80) \quad \tilde{R}_k^{p_1} = \alpha_k^{p_1} Q^{p_1} = C_k^1 \frac{1}{l(\mathbf{p})!} \frac{1}{p_1 + 1} (p_1 + 1) = k$$

and if $l(\mathbf{p}) \geq 2$, as $\mathbf{p} = p_1 \mathbf{q}$, the same careful examination of the position of p_1 in a shuffle of $\mathbf{m}^1, \dots, \mathbf{m}^s$ giving \mathbf{p} , combined with the fact that $Q^{p_1} = (p_1 + 1)$ and $Q^{p_1 \mathbf{m}^i} = \|\mathbf{m}^i\| Q^{\mathbf{m}^i}$, yields

$$(81) \quad \tilde{R}_k^{\mathbf{p}} = \sum_{s=1}^{l(\mathbf{p})-1} \sum_{\mathbf{m}^1, \dots, \mathbf{m}^s} \text{sh}_{\mathbf{q}}^{\mathbf{m}^1, \dots, \mathbf{m}^s} Q^{\mathbf{m}^1} \dots Q^{\mathbf{m}^s} V_{k, p_1}^{\|\mathbf{m}^1\|, \dots, \|\mathbf{m}^s\|}$$

where

$$(82) \quad V_{k, p_1}^{n_1, \dots, n_s} = \sum_{i=1}^s n_i \alpha_k^{n_1, \dots, n_i + p_1, \dots, n_s} + (p_1 + 1) \sum_{i=0}^s \alpha_k^{n_1, \dots, n_i, p_1, n_{i+1}, \dots, n_s}$$

A careful computation leads to

$$(83) \quad V_{k, p_1}^{n_1, \dots, n_s} = (n_1 + \dots + n_s + k) \alpha_k^{n_1, \dots, n_s}$$

Now by induction, if $\mathbf{p} = p_1 \mathbf{q}$, we easily get that

$$(84) \quad \tilde{R}_k^{\mathbf{p}} = (\|\mathbf{q}\| + k) \tilde{R}_k^{\mathbf{q}} = (\|\mathbf{q}\| + k) R_k^{\mathbf{q}} = R_k^{\mathbf{p}}$$

We leave the similar proof of theorem 3 to the reader : the ideas are the same, noticing that

$$\begin{aligned}
 S(\Gamma_n) &= \sum_{(n_1, \dots, n_s) \in \mathcal{N}_n} Q^{n_1, \dots, n_s} S(Z^{n_1, \dots, n_s}) \\
 &= \sum_{(n_1, \dots, n_s) \in \mathcal{N}_n} (-1)^s Q^{n_1, \dots, n_s} Z^{n_s, \dots, n_1}
 \end{aligned}$$

7. Tables and conclusion.

For the coproduct, the following table gives the value of $\frac{n!}{n_1! \dots n_{s+1}!} \alpha_{n_{s+1}}^{n_1, \dots, n_s}$ for a given sequence (n_1, \dots, n_{s+1}) :

| | | | |
|--------------------|--------------------|-----------------------|--------------------|
| $(1, 1) = 1$ | | | |
| $(1, 2) = 3$ | $(2, 1) = 1$ | $(1, 1, 1) = 1$ | |
| $(1, 3) = 6$ | $(2, 2) = 4$ | $(3, 1) = 1$ | $(1, 1, 2) = 7$ |
| $(1, 2, 1) = 3/2$ | $(2, 1, 1) = 3/2$ | $(1, 1, 1, 1) = 1$ | |
| $(1, 4) = 10$ | $(2, 3) = 10$ | $(3, 2) = 5$ | $(4, 1) = 1$ |
| $(1, 1, 3) = 25$ | $(1, 3, 1) = 2$ | $(3, 1, 1) = 2$ | $(1, 2, 2) = 25/2$ |
| $(2, 1, 2) = 25/2$ | $(2, 2, 1) = 3$ | $(1, 1, 1, 2) = 15$ | $(1, 1, 2, 1) = 2$ |
| $(1, 2, 1, 1) = 2$ | $(2, 1, 1, 1) = 2$ | $(1, 1, 1, 1, 1) = 1$ | |

This gives

$$\begin{aligned}
 \tilde{\Delta}\Gamma_1 &= 0 \\
 \tilde{\Delta}\Gamma_2 &= \Gamma_1 \otimes \Gamma_1 \\
 \tilde{\Delta}\Gamma_3 &= (\Gamma_2 + \Gamma_1^2) \otimes \Gamma_1 + 3\Gamma_1 \otimes \Gamma_2 \\
 \tilde{\Delta}\Gamma_4 &= (\Gamma_3 + 3\Gamma_1\Gamma_2 + \Gamma_1^3) \otimes \Gamma_1 + (4\Gamma_2 + 7\Gamma_1^2) \otimes \Gamma_2 + 6\Gamma_1 \otimes \Gamma_3 \\
 \tilde{\Delta}\Gamma_5 &= (\Gamma_4 + 4\Gamma_1\Gamma_3 + 3\Gamma_2^2 + 6\Gamma_1^2\Gamma_2 + \Gamma_1^4) \otimes \Gamma_1 \\
 &\quad + (5\Gamma_3 + 25\Gamma_1\Gamma_2 + 15\Gamma_1^3) \otimes \Gamma_2 + (10\Gamma_2 + 25\Gamma_1^2) \otimes \Gamma_3 + 10\Gamma_1 \otimes \Gamma_4
 \end{aligned}$$

For the antipode, the following table gives the value of $\frac{(n_1 + \dots + n_s)!}{n_1! \dots n_s!} \beta^{n_1, \dots, n_s}$ for a given sequence (n_1, \dots, n_s) :

| | | | |
|-------------------------|---------------------|-----------------------|-----------------------|
| $(1) = -1$ | | | |
| $(2) = -1$ | $(1, 1) = 1$ | | |
| $(3) = -1$ | $(1, 2) = 3$ | $(2, 1) = 1$ | $(1, 1, 1) = -2$ |
| $(4) = -1$ | $(1, 3) = 6$ | $(2, 2) = 4$ | $(3, 1) = 1$ |
| $(1, 1, 2) = -11$ | $(1, 2, 1) = -9/2$ | $(2, 1, 1) = -5/2$ | $(1, 1, 1, 1) = 6$ |
| $(5) = -1$ | $(1, 4) = 10$ | $(2, 3) = 10$ | $(3, 2) = 5$ |
| $(4, 1) = 1$ | $(1, 1, 3) = -35$ | $(1, 3, 1) = -8$ | $(3, 1, 1) = -3$ |
| $(1, 2, 2) = -55/2$ | $(2, 1, 2) = -35/2$ | $(2, 2, 1) = -7$ | |
| $(1, 1, 1, 2) = 50$ | $(1, 1, 2, 1) = 22$ | $(1, 2, 1, 1) = 29/2$ | $(2, 1, 1, 1) = 19/2$ |
| $(1, 1, 1, 1, 1) = -24$ | | | |

This gives :

$$\begin{aligned}
 S(\Gamma_1) &= -\Gamma_1 \\
 S(\Gamma_2) &= -\Gamma_2 + \Gamma_1^2 \\
 S(\Gamma_3) &= -\Gamma_3 + 4\Gamma_1\Gamma_2 - 2\Gamma_1^3 \\
 S(\Gamma_4) &= -\Gamma_4 + 7\Gamma_1\Gamma_3 + 4\Gamma_2^2 - 18\Gamma_1^2\Gamma_2 + 6\Gamma_1^4 \\
 S(\Gamma_5) &= -\Gamma_5 + 11\Gamma_1\Gamma_4 + 15\Gamma_2\Gamma_3 - 46\Gamma_1^2\Gamma_3 - 52\Gamma_1\Gamma_2^2 + 96\Gamma_1^3\Gamma_2 - 24\Gamma_1^5
 \end{aligned}$$

This is the attempted result but the formulas in proposition 2, theorem 2 and 3 are not unique because \mathcal{H}_{CM}^1 is commutative and, in the computations, it is much more "simple" to consider that the algebra generated by the δ_n is somehow non-commutative. This situation calls for further investigations, since the coefficients appearing in proposition 2 for the Faà di Bruno coordinates seem to arise in the study of a noncommutative version of diffeomorphisms (see [1]).

References

- [1] Christian Brouder, Alessandra Frabetti, and Christian Krattenthaler. Non-commutative Hopf algebra of formal diffeomorphisms. *Adv. Math.*, 200(2):479–524, 2006.

- [2] Alain Connes and Dirk Kreimer. Hopf algebras, renormalization and noncommutative geometry. In *Quantum field theory: perspective and prospective (Les Houches, 1998)*, volume 530 of *NATO Sci. Ser. C Math. Phys. Sci.*, pages 59–108. Kluwer Acad. Publ., Dordrecht, 1999.
- [3] Alain Connes and Dirk Kreimer. Renormalization in quantum field theory and the Riemann-Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem. *Commun. Math. Phys.*, 210(1):249–273, 2000.
- [4] Alain Connes and Dirk Kreimer. Renormalization in quantum field theory and the Riemann-Hilbert problem. II: The β -function, diffeomorphisms and the renormalization group. *Commun. Math. Phys.*, 216(1):215–241, 2001.
- [5] Alain Connes and Henri Moscovici. Hopf algebras, cyclic cohomology and the transverse index theorem. *Commun. Math. Phys.*, 198(1):199–246, 1998.
- [6] Jean Écalle. *Les fonctions résurgentes. Tome I*, volume 5 of *Publications Mathématiques d'Orsay 81 [Mathematical Publications of Orsay 81]*. Université de Paris-Sud Département de Mathématique, Orsay, 1981. Les algèbres de fonctions résurgentes. [The algebras of resurgent functions], With an English foreword.
- [7] Jean Écalle. *Les fonctions résurgentes. Tome II*, volume 6 of *Publications Mathématiques d'Orsay 81 [Mathematical Publications of Orsay 81]*. Université de Paris-Sud Département de Mathématique, Orsay, 1981. Les fonctions résurgentes appliquées à l'itération. [Resurgent functions applied to iteration].
- [8] Jean Écalle. *Les fonctions résurgentes. Tome III*, volume 85 of *Publications Mathématiques d'Orsay [Mathematical Publications of Orsay]*. Université de Paris-Sud, Département de Mathématiques, Orsay, 1985. L'équation du pont et la classification analytique des objets locaux. [The bridge equation and analytic classification of local objects].
- [9] Jean Écalle. Singularités non abordables par la géométrie. *Ann. Inst. Fourier (Grenoble)*, 42(1-2):73–164, 1992.
- [10] Héctor Figueroa and José M. Gracia-Bondia. Combinatorial Hopf algebras in quantum field theory. I. *Rev. Math. Phys.*, 17(8):881–976, 2005.
- [11] Frédéric Menous. Formulas for the Connes-Moscovici Hopf algebra. *C. R. Math. Acad. Sci. Paris*, 341(2):75–78, 2005.
- [12] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, New York, 1993. , Oxford Science Publications.

UNIV PARIS-SUD, LABORATOIRE DE MATHÉMATIQUES D'ORSAY, ORSAY CEDEX, F-91405;
CNRS, ORSAY CEDEX, F-91405

E-mail address: `Frederic.Menous@math.u-psud.fr`

This page intentionally left blank

Hopf algebras and the combinatorics of connected graphs in quantum field theory

Ângela Mestre and Robert Oeckl

ABSTRACT. In this talk, we are concerned with the formulation and understanding of the combinatorics of time-ordered n -point functions in terms of the Hopf algebra of field operators. Mathematically, this problem can be formulated as one in combinatorics or graph theory. It consists in finding a recursive algorithm that generates all connected graphs in their Hopf algebraic representation. This representation can be used directly and efficiently in evaluating Feynman graphs as contributions to the n -point functions.

Recently, it was realized that the Hopf algebra structure of the algebra of field operators $S(V)$ (with the normal or with the time-ordered product) can be fruitfully exploited. In particular, the Laplace Hopf algebra created by Rota *et al.* [4, 7] was generalized to provide an algebraic tool for combinatorial problems of quantum field theory [1]. In addition, Sweedler's Hopf algebra cohomology [13] and the Drinfel'd twist [5] were used to show that different products of the algebra of field operators are related by Drinfel'd twists and that interactions correspond to 2-cocycles [2]. Subsequently, the time-ordered Hopf algebra of field operators was used to define an algebraic representation of graphs [11, 12]. That is, every graph with v vertices was associated with an element of $S(V)^{\otimes v}$, the v -fold tensor product of $S(V)$. This representation allowed to derive simple algebraic relations between complete, connected and 1-particle irreducible (1PI) n -point functions [11], and to express a connected n -point function in terms of its loop order contributions [12]. The basic structure is an algorithm to recursively generate all weighted connected graphs. That is, each graph is generated together with a scalar that will correspond to the inverse of its symmetry factor. This algorithm is amenable to direct implementation and allows efficient calculations of connected graphs as well as their values as Feynman graphs. Also, all results apply both to bosonic and fermionic fields. Note, however, that no type of renormalization procedure is taken into account. In this sense the computed n -point functions may be considered as bare ones. This Hopf algebraic approach to express the connected Green functions in terms of the 1PI ones, was recently generalized to many-body physics [3].

This paper reviews the main results of [11, 12]. Section 1 recalls the Hopf algebra structure of the algebra of quantum field operators. Section 2 describes the Hopf algebraic representation of graphs and gives some examples. Section 3 sketches the algorithmic construction of connected graphs as well as the interpretation in

terms of Feynman graphs and n -point functions. Section 4 focuses on expressing the relation between connected and 1PI n -point functions in a completely algebraic language.

1. Field operator algebra as a Hopf algebra

In the following, we are concerned with a generic perturbative quantum field theory with fields $\phi(x)$ as operator valued distributions with adequate test functions. Here, x represents a label that specifies spacetime coordinates, particle type, Minkowski indices, spin, colour, etc. For more information on the basic field operators, we refer the reader to any standard textbook on quantum field theory such as [8]. Note, however, that their essential property used in this paper, is that they are (labeled) elements of a vector space. Moreover, while our notation suggests a field theory with a single scalar field, this is just a convenience. Although our results are general, we limit ourselves in the following exposition to a purely bosonic theory for simplicity. A more extensive discussion including fermionic fields, can be found in Section VI of [11].

We briefly recall the Hopf algebra structure of the time-ordered field operator algebra. Let V denote the vector space of finite linear combinations of elementary field operators $\phi(x)$. Let $S(V) = \bigoplus_{v=0}^{\infty} V^v$ denote the free, unital, commutative algebra generated by all time-ordered products of field operators¹. In particular, $S(V)$ is also a Hopf algebra [6, 9, 10]. Therefore, $S(V)$ is equipped with a linear map $\Delta : S(V) \rightarrow S(V) \otimes S(V)$, called coproduct, which is coassociative: $(\Delta \otimes \text{id}) \circ \Delta = (\text{id} \otimes \Delta) \circ \Delta$. This is defined on V by $\Delta(\mathbf{1}) = \mathbf{1} \otimes \mathbf{1}$, $\Delta(\phi(x)) = \phi(x) \otimes \mathbf{1} + \mathbf{1} \otimes \phi(x)$, and extended to the whole of $S(V)$ due to the compatibility with the product. The coproduct may be interpreted as an operation to split a product of field operators into two parts in all possible ways. Moreover, on $S(V)$ the counit $\epsilon : S(V) \rightarrow \mathbb{C}$ is defined by $\epsilon(\mathbf{1}) = 1$ and $\epsilon(\phi(x_1) \cdots \phi(x_n)) = 0$ for $n > 0$. The characterizing property of the counit is the equality $(\epsilon \otimes \text{id}) \circ \Delta = \text{id} = (\text{id} \otimes \epsilon) \circ \Delta$. Finally, to meet the requirements of a Hopf algebra, there is an antipode map $S : S(V) \rightarrow S(V)$ defined on $S(V)$, by $S(\phi(x_1) \cdots \phi(x_n)) = (-1)^n \phi(x_1) \cdots \phi(x_n)$.

We recall that the various types of n -point functions are vacuum expectation values of n time-ordered field operators. By $G^{(n)}$, $G_c^{(n)}$, $G_{1\text{PI}}^{(n)}$ and $\hat{G}_{1\text{PI}}^{(n)}$, we denote complete, connected, 1PI and modified 1PI² n -point functions, respectively. Moreover, by \mathcal{V} , we denote the vertex functions, i.e., the interaction vertices of each of the terms (other than the Feynman propagator) in the perturbative expansion of the connected Green functions. The ensemble of time-ordered n -point functions of a given type, determine maps $S(V) \rightarrow \mathbb{C}$:

$$\begin{aligned} \rho(\phi(x_1) \cdots \phi(x_n)) &:= G^{(n)}(x_1, \dots, x_n), \\ \sigma(\phi(x_1) \cdots \phi(x_n)) &:= G_c^{(n)}(x_1, \dots, x_n), \\ \tau(\phi(x_1) \cdots \phi(x_n)) &:= G_{1\text{PI}}^{(n)}(x_1, \dots, x_n), \end{aligned}$$

¹For convenience, in $S(V)$, we identify the time-ordered product with the free commutative product.

²Recall that the connected n -point functions G_c are expressible in terms of $\hat{G}_{1\text{PI}}$ as a sum over all tree graphs whose vertices have valence at least three, by associating a connected propagator $G_c^{(2)}$ to every edge.

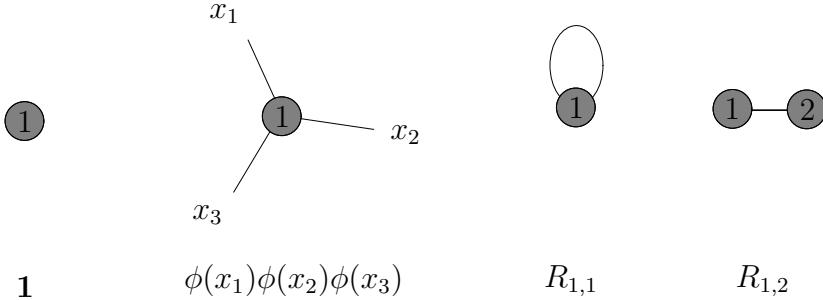


FIGURE 1. Examples of the correspondence between vertices and edges, and elements of $S(V)^{\otimes v}$.

$$\begin{aligned}\hat{\tau}(\phi(x_1) \cdots \phi(x_n)) &:= \hat{G}_{\text{1PI}}^{(n)}(x_1, \dots, x_n), \\ \nu(\phi(x_1) \cdots \phi(x_n)) &:= \mathcal{V}(x_1, \dots, x_n).\end{aligned}$$

The assumption that all 1-point functions vanish means that $\rho(\phi(x)) = \sigma(\phi(x)) = \tau(\phi(x)) = \hat{\tau}(\phi(x)) = 0$. Moreover, the 0-point functions read as $\rho(\mathbf{1}) = 1$, $\sigma(\mathbf{1}) = \tau(\mathbf{1}) = \hat{\tau}(\mathbf{1}) = 0$. Besides, the 2-point function $\hat{\tau}(\phi(x)\phi(y))$ vanishes by construction.

2. A Hopf algebraic representation of graphs

We study the correspondence between graphs and elements of $S(V)^{\otimes v}$ given in [11, 12].

We recall that the Feynman propagator G_F is a Green's function for the free Schrödinger equation. Its inverse (under convolution) G_F^{-1} is determined by the following equation

$$\int dy G_F(x, y) G_F^{-1}(y, z) = \delta(x, z).$$

We now define the following elements of $S(V)^{\otimes v}$ ³:

$$(1) \quad R_{i,j} := \int dx dy G_F^{-1}(x, y) (\mathbf{1}^{\otimes i-1} \otimes \phi(x) \otimes \mathbf{1}^{\otimes j-i-1} \otimes \phi(y) \otimes \mathbf{1}^{\otimes v-j}),$$

where the field operators $\phi(x)$ and $\phi(y)$ are inserted at the positions i and j , respectively, with $i \neq j$; and

$$(2) \quad R_{i,i} := \int dx dy G_F^{-1}(x, y) (\mathbf{1}^{\otimes i-1} \otimes \phi(x)\phi(y) \otimes \mathbf{1}^{\otimes v-i}).$$

The elements $R_{i,i}$ and $R_{i,j}$ are employed in establishing a correspondence between graphs with v vertices and certain elements of $S(V)^{\otimes v}$, the v -fold tensor product of $S(V)$. Namely,

³The integrations in equations (1) and (2) are part of the definitions of $R_{i,j}$ and $R_{i,i}$ to simplify the notation in the following. Indeed, only the integrands are actually elements of $S(V)^{\otimes v}$. To be precise, the integrations over all internal variables would be included later on, on the rhs of the formulas of Corollary 2 and of Theorem 5.

$$R_{1,2} \cdot (\phi(x_1)\phi(x_2) \otimes \phi(x_3)\phi(x_4))$$

$$R_{1,2}^2 \cdot (\phi(x_1)\phi(x_2) \otimes \phi(x_3)\phi(x_4))$$

FIGURE 2. Examples of the algebraic representation of graphs in terms of elements of $S(V)^{\otimes v}$.

- a tensor factor in the i^{th} position corresponds to a vertex numbered i , with $i = 1, \dots, v$;
- a product $\phi(x_1) \cdots \phi(x_n)$ in a given tensor factor corresponds to external edges (of the associated vertex) whose end points are labeled by x_1, \dots, x_n ;
- the element $R_{i,j} \in S(V)^{\otimes v}$ defined by (1), corresponds to an internal edge connecting the vertices i and j ;
- the element $R_{i,i} \in S(V)^{\otimes v}$ for $1 \leq i \leq v$ defined by (2), corresponds to an internal edge connecting the vertex i to itself (i.e., a self-loop).

Figure 1 shows examples of the correspondence between vertices, external edges and internal edges, and elements of $S(V)^{\otimes v}$.

Combining several internal edges and their products with external edges by multiplying the respective expressions in $S(V)^{\otimes v}$, allows to build arbitrary graphs with v vertices. Figure 2 shows some examples. Moreover, applying the vertex functions ν to each tensor factor yields precisely the value of the respective graph as a Feynman graph.

Usually Feynman graphs involve edges of different types depending on particle species, e.g. straight for fermions, wiggly for bosons, etc. Thus, suppose that there are m different fields ϕ^a , with $a = 1, \dots, m$, interacting. Each of these fields is associated with an edge of certain kind. In the present context, this means that edges are represented by distinct elements $R_{i,j}^a, R_{i,i}^a \in S(V)^{\otimes v}$, with $a = 1, \dots, m$, given by

$$R_{i,j}^a := \int dx dy G_F^{a-1}(x,y) (\mathbf{1}^{\otimes i-1} \otimes \phi^a(x) \otimes \mathbf{1}^{\otimes j-i-1} \otimes \phi^a(y) \otimes \mathbf{1}^{\otimes v-j}),$$

and

$$R_{i,i}^a := \int dx dy G_F^{a-1}(x,y) (\mathbf{1}^{\otimes i-1} \otimes \phi^a(x) \phi^a(y) \otimes \mathbf{1}^{\otimes v-i}),$$

respectively. Therefore, the elements $R_{i,j}$ and $R_{i,i}$ that we consider read explicitly as $R_{i,j} = \sum_{a=1}^m R_{i,j}^a$ and $R_{i,i} = \sum_{a=1}^m R_{i,i}^a$, respectively.

A fundamental property of the algebraic representation is that the ordering of the tensor factors of $S(V)^{\otimes v}$ induces an ordering of the vertices of the graphs. However, when applying $\nu^{\otimes v}$ the ordering is “forgotten”. Indeed, it is not relevant for the interpretation of graphs as Feynman graphs, but only plays a role at the

level of their algebraic representation. Moreover, usually the elements of $\mathbb{S}(V)^{\otimes v}$ yield as linear combinations of expressions corresponding to graphs. In this context, we call the scalar multiplying the expression for a given graph the *weight* of the graph. Clearly, if we are interested in unordered graphs, the weight of such a graph is the sum of the weights of all vertex ordered graphs that correspond to it upon forgetting the vertex order.

3. Generating connected graphs via Hopf algebra

Statement of result. We state the main result of [12]. This may be described as an algorithm to recursively generate all connected graphs. In particular, each graph is produced together with a weight factor given by the inverse of its symmetry factor. All graphs are generated in the Hopf algebraic representation introduced in Section 2. This allows their direct evaluation as Feynman graphs.

We emphasize that the discussion here applies only to bare n -point functions for renormalization is outside the scope of the present work. Moreover, the results are valid for any quantum field theory.

We may view the elements $R_{i,j}$ and $R_{i,i}$, given by formulas (1) and (2), respectively, as operators on $\mathbb{S}(V)^{\otimes v}$ by multiplication. They are used to define the following linear maps:

- $T_i : \mathbb{S}(V)^{\otimes v} \rightarrow \mathbb{S}(V)^{\otimes v}$, with $1 \leq i \leq v$, as the operator $R_{i,i}$ together with the factor $1/2$:

$$T_i := \frac{1}{2} R_{i,i},$$

- $Q_i : \mathbb{S}(V)^{\otimes v} \rightarrow \mathbb{S}(V)^{\otimes v+1}$, with $1 \leq i \leq v$, given by the composition of $R_{i,i+1}$ with the coproduct applied to the i^{th} component of $\mathbb{S}(V)^{\otimes v}$, i.e., $\Delta_i := \text{id}^{\otimes i-1} \otimes \Delta \otimes \text{id}^{\otimes v-i} : \mathbb{S}(V)^{\otimes v} \rightarrow \mathbb{S}(V)^{\otimes v+1}$, together with the factor $1/2$:

$$Q_i := \frac{1}{2} R_{i,i+1} \circ \Delta_i.$$

The map T_i endows the vertex i of a vertex ordered graph with a self-loop. The map Q_i splits the vertex i into two new vertices, numbered i and $i+1$, distributes the ends of edges ending on the split vertex between the two new ones in all possible ways and connects the two new vertices with an edge.

The maps T_i increase both the loop and edge numbers of a graph by one unit, leaving the vertex number invariant. Also, the maps Q_i increase both the vertex and edge numbers by one unit, leaving the loop number invariant.

Clearly, both the maps T_i and Q_i produce connected graphs from connected ones, so that the following theorem holds.

THEOREM 1. *Let $l, n \geq 0$, $v \geq 1$ denote integers and let the set of maps $\Omega^{l,v} : \mathbb{S}(V) \rightarrow \mathbb{S}(V)^{\otimes v}$ be defined recursively as follows:*

$$(3) \quad \begin{aligned} \Omega^{0,1} &:= \text{id}, \\ \Omega^{l,v} &:= \frac{1}{l+v-1} \left(\sum_{i=1}^{v-1} Q_i \circ \Omega^{l,v-1} + \sum_{i=1}^v T_i \circ \Omega^{l-1,v} \right). \end{aligned}$$

Then, for fixed values of l, v, n and operator labels x_1, \dots, x_n , $\Omega^{l,v}(\phi(x_1) \cdots \phi(x_n))$ corresponds to the weighted sum over all connected graphs with l loops, v vertices

and n external edges whose end points are labeled by x_1, \dots, x_n , each with weight the inverse of its symmetry factor.

In the recursion equation above the Q and T summands do not appear when $v = 1$ or when $l = 0$, respectively. The proof of Theorem 1 proceeds by induction on the number of internal edges $e = l + v - 1$ [12]. Moreover, formula (3) is an example of a double recursion. Therefore, its algorithmic implementation is that of any recurrence which makes two calls to itself, such as the defining recurrence of the binomial coefficients.

Now, we turn to the interpretation in terms of Feynman graphs and n -point functions. Denote the l -loop and v -vertex contribution to the ensemble σ of connected n -point functions by $\sigma^{l,v}$. The l -loop order contribution σ^l to σ and σ itself, are given by

$$\sigma^l = \sum_{v=0}^{\infty} \sigma^{l,v}, \quad \sigma = \sum_{l=0}^{\infty} \sigma^l.$$

There is only one contribution with zero vertex number. This is the Feynman propagator contributing to the 2-point function. Hence, $\sigma^{l,v}$ is zero if $v = 0$ and $l \neq 0$, while $\sigma^{0,0}$ is non-zero only on $V \otimes V$ and coincides there with the Feynman propagator. All non-zero vertex number contributions are captured by the following corollary.

COROLLARY 2. *For $v \geq 1$:*

$$\sigma^{l,v} = \nu^{\otimes v} \circ \Omega^{l,v}.$$

Alternative recursion formula. A key feature of $\Omega^{l,v}$ is that of satisfying an alternative recursion relation. This has the advantage over (3), that it may be translated directly into a recursion relation of the resulting n -point functions $\sigma^{l,v}$, related via Corollary 2.

PROPOSITION 3. *Let $v \geq 1$ and $l \geq 0$, but not $v = 1$ and $l = 0$. Then,*

$$(4) \quad \Omega^{l,v} = \frac{1}{l+v-1} \left(\Omega^{l-1,v} \circ T + \sum_{a=0}^l \sum_{b=1}^{v-1} (\Omega^{a,b} \otimes \Omega^{l-a,v-b}) \circ Q \right).$$

The first summand does not contribute if $l = 0$, while the second does not contribute if $v = 1$.

Proposition 3 is proved by induction on the number of internal edges $e = l + v - 1$ [12]. Formula (4) has a straightforward interpretation in terms of sums over weighted graphs following the correspondence of Section 2. Namely, the formula states that the weighted sum over graphs with l loops and v vertices is given by a sum of two terms divided by e . The first term is the sum over all weighted graphs with $l - 1$ loops and v vertices which have an extra internal edge attached, its end points being connected to vertices in all possible ways. The second term is a sum over all ordered pairs of weighted graphs with total number of vertices equal to v and total number of loops equal to l , connected in all possible ways with an internal edge.

Combining this result with Corollary 2 yields the corresponding recursion formula for $\sigma^{l,v}$.

COROLLARY 4. *Let $v \geq 1$ and $l \geq 0$, but not $v = 1$ and $l = 0$. Then,*

$$\sigma^{l,v} = \frac{1}{l+v-1} \left(\sigma^{l-1,v} \circ T + \sum_{a=0}^l \sum_{b=1}^{v-1} (\sigma^{a,b} \otimes \sigma^{l-a,v-b}) \circ Q \right).$$

It is understood that the first summand does not contribute if $l = 0$, while the second does not contribute if $v = 1$.

4. Connected and 1PI n -point functions

Now, we turn attention to simple algebraic expressions for the relation between connected and 1PI n -point functions following from Theorem 1 and Corollary 4 [11].

THEOREM 5. *The connected n -point functions σ may be expressed in terms of the 1PI ones τ through the formula*

$$\sigma = \sigma^0 + \sum_{v=1}^{\infty} \sigma^v, \quad \text{with} \quad \sigma^v := \tau^{\otimes v} \circ \Omega^{0,v},$$

where σ^0 is non-zero only on $V \otimes V$ and coincides there with the Feynman propagator.

Since the 0-point and 1-point 1PI functions are zero, Theorem 5 holds as long as all tree graphs with v vertices, n external edges whose end points are labeled by x_1, \dots, x_n , and the property that each vertex has valence at least two, have no non-trivial symmetries. In other words, all tree graphs with the aforesaid properties are required to occur in $\Omega^{0,v}(\phi(x_1) \cdots \phi(x_n))$ with exactly weight 1. This is ensured by the following lemma.

LEMMA 6. *Consider a tree graph γ , all of whose vertices have valence at least two. Then, γ has no non-trivial symmetries.*

In particular, by Corollary 4, the σ^v satisfy the following recursion formula.

PROPOSITION 7. *σ^v may be determined recursively via $\sigma^1 = \tau$ and with the recursion equation for $v \geq 2$,*

$$\sigma^v = \frac{1}{v-1} \sum_{i=1}^{v-1} (\sigma^i \otimes \sigma^{v-i}) \circ Q.$$

We notice that all results immediately carry over to the relation between connected and modified 1PI n -functions. To this end, we modify the definition of R (and consequently that of $\Omega^{0,v}$) by replacing G_F^{-1} with $G_c^{(2)-1}$ in formula (1). Recall that for the modified 1PI functions $\hat{G}_{1\text{PI}}$ not only 0- and 1-point functions vanish, but also 2-point functions. This implies that only trees contribute which have the property that all their vertices have valence at least three. Actually, for a given number of external edges, there are only finitely many such trees. Therefore, a connected function yields a finite sum over tree graphs with modified 1PI functions as vertices, for each set of external edges, i.e., for each element of $S(V)$ to which it is applied.

Acknowledgments. We would like to thank the organizers of the Conference on Combinatorics and Physics for their kind invitation. The research was supported in part by the Czech Ministry of Education, Youth and Sports within the project LC06002.

References

- [1] Ch. Brouder, *A quantum field algebra*, arXiv: math-ph/0201033 (2002).
- [2] Ch. Brouder, A. Frabetti, B. Fauser, and R. Oeckl, *Quantum field theory and Hopf algebra cohomology*, J. Phys. **A** **37** (2004), 5895–5927.
- [3] Ch. Brouder, A. Frabetti, and F. Patras, *One-particle irreducibility with initial correlations*, arXiv.org: 0803.3747 (2008).
- [4] P. Doubilet, G.-C. Rota, and J.A. Stein, *On the foundations of combinatorial theories IX: Combinatorial methods in invariant theory*, Stud. Appl. Math. **53** (1974), 185–216.
- [5] G. Drinfel'd, *Quasi-Hopf algebras*, Leningrad Math. J. **1** (1990), 1419–57.
- [6] D. Eisenbud, *Commutative algebra with a view toward algebraic geometry*, Springer, Berlin, 1995.
- [7] F.D. Grosshans, G.-C. Rota, and J.A. Stein, *Invariant theory and superalgebras*, Amer. Math. Soc., Providence, 1987.
- [8] C. Itzykson and J.-B. Zuber, *Quantum field theory*, McGraw-Hill, New York, 1980.
- [9] C. Kassel, *Quantum groups*, Springer-Verlag, New York, Inc, 1995.
- [10] J.-L. Loday, *Cyclic homology*, Springer, Berlin, 1998.
- [11] Â. Mestre and R. Oeckl, *Combinatorics of n -point functions via Hopf algebra in quantum field theory*, J. Math. Phys. **47** (2006), 052301.
- [12] ———, *Generating loop graphs via Hopf algebra in quantum field theory*, J. Math. Phys. **47** (2006), 122302.
- [13] M. E. Sweedler, *Cohomology of algebras over Hopf algebras*, Trans. Am. Math. Soc. **133** (1968), 205–239.

INSTITUT DE MINÉRALOGIE ET DE PHYSIQUE DES MILIEUX CONDENSÉS, UNIVERSITÉ PIERRE ET MARIE CURIE, CAMPUS BOUCICAUT, 140 RUE DE LOURMEL, F-75015 PARIS, FRANCE; EMAIL: MESTRE@IMPMC.JUSSIEU.FR, AND DOPPLER INSTITUTE & DEPARTMENT OF MATHEMATICS,, FN-SPE, CZECH TECHNICAL UNIVERSITY,, TROJANOVA 13, 120 00 PRAHA 2,, CZECH REPUBLIC,, ROBERT OECKL EMAIL: ROBERT@MATMOR.UNAM.MX, INSTITUTO DE MATEMÁTICAS, UNAM, CAMPUS MORELIA,, C. P. 58190, MORELIA, MICHOACÁN, MEXICO,

Hopf algebras of formal diffeomorphisms and numerical integration on manifolds

Alexander Lundervold and Hans Munthe-Kaas

ABSTRACT. B-series originated from the work of John Butcher in the 1960s as a tool to analyze numerical integration of differential equations, in particular Runge–Kutta methods. Connections to renormalization theory in perturbative quantum field theory have been established in recent years. The algebraic structure of classical Runge–Kutta methods is described by the Connes–Kreimer Hopf algebra.

Lie–Butcher series are generalizations of B-series that are aimed at studying Lie-group integrators for differential equations evolving on manifolds. Lie group integrators are based on general Lie group actions on a manifold, and classical Runge–Kutta integrators appear in this setting as the special case of \mathbb{R}^n acting upon itself by translations. Lie–Butcher theory combines classical B-series on \mathbb{R}^n with Lie-series on manifolds. The underlying Hopf algebra H_N combines the Connes–Kreimer Hopf algebra with the shuffle Hopf algebra of free Lie algebras.

Aimed at a general mathematical audience, we give an introduction to Hopf algebraic structures and their relationship to structures appearing in numerical analysis. In particular, we explore the close connection between Lie series, time-dependent Lie series and Lie–Butcher series for diffeomorphisms on manifolds. The role of the Euler and Dynkin idempotents in numerical analysis is discussed. A non-commutative version of a Faà di Bruno bialgebra is introduced, and the relation to non-commutative Bell polynomials is explored.

1. Outline

The main point of this paper is to explore algebraic structures underlying groups of formal diffeomorphisms on manifolds. The focus is on some important mathematical structures appearing in numerical integration on manifolds that are likely to find applications also in other areas of mathematics. The relationship between classical Lie series on manifolds, time-dependent Lie series and Lie–Butcher series is explained in detail. We develop the algebraic structures introduced in [38, 39, 40, 42, 45, 44, 3], and in particular explore connections between Hopf- and Lie algebras, differential geometry and analysis of numerical integration on manifolds. The paper does not go into a detailed study of applications of these algebraic structures in numerical analysis, but we sketch briefly several of the many applications in this field.

The introductory Chapters 2 and 3 give an overview of well-known results. Chapter 2 contains a brief introduction to numerical integration and algebraic structures appearing in numerical analysis, both classical methods on \mathbb{R}^n and Lie group methods generalizing to manifolds. Chapter 3 presents a brief introduction to Hopf algebraic structures.

Chapter 4 contains more new and recent material. It details the algebraic structures of Lie–Butcher theory, and discusses the interplay between algebraic and differential geometric points of view. In particular, we want to emphasize the strong connections between the algebraic theory of Lie series, time-dependent Lie series and Lie–Butcher series. Chapter 4 therefore starts with a discussion of classical Lie series and pullback formulas on manifolds, continuing with an exploration of some less known time-dependent pullback formulas. We will explain the relevance of the Euler and Dynkin idempotents in numerics and introduce a non-commutative *Dynkin–Faà di Bruno* bialgebra, related to non-commutative Bell polynomials appearing in various contexts in earlier works: numerical analysis [39], control theory [37] and quantization [31]. This Dynkin–Faà di Bruno bialgebra is related to, but different from, the Hopf algebras explored by Brouder et. al. in [6].

In the final part of Chapter 4, we turn to Lie–Butcher series. We explore backward error analysis and the substitution law in the setting of algebras of non-commuting frames on manifolds. Although we will not give detailed expositions and applications of these subjects, we hope that this presentation will systematize the theory and open the topics for further research.

2. Introduction to numerical integrators and their analysis

Let \mathcal{M} be a manifold and $F: \mathcal{M} \rightarrow T\mathcal{M}$ a vector field. By the flow of an autonomous vector field F we mean the diffeomorphism $\Phi_{t,F}: \mathcal{M} \rightarrow \mathcal{M}$, defined for $t \in \mathbb{R}$ such that $\Phi_{s,F} \circ \Phi_{t,F} = \Phi_{s+t,F}$, $\Phi_{0,F} = \text{Id}$ and $\partial/\partial t|_{t=0} \Phi_{t,F}(p) = F(p)$ for all $p \in \mathcal{M}$.

Numerical integration of ODEs is about constructing good numerical approximations to $\Phi_{t,F}$ for a given vector field F . A numerical integration algorithm yields a diffeomorphism $\Psi_{h,F}$, henceforth called the *numerical integrator*. The real parameter h is called the *step size*. For an initial point $y_0 \in \mathcal{M}$, and a chosen step size $h > 0$, the numerical method produces a discrete sequence of solution points $y_i = \Psi_{h,F}(y_{i-1})$, with the goal of arriving at $y_k \approx \Phi_{kh,F}(y_0)$. Note that, unlike the exact flow, numerical integrators are *not* 1-parameter Lie groups in h . In general we have $\Psi_{h,F} \circ \Psi_{s,F} \neq \Psi_{h+s,F}$, and $\Psi_{-h,F} \neq \Psi_{h,F}^{-1}$. Integrators for which the latter identity holds are called *(time-)symmetric* methods. Most integrators satisfy the consistency conditions $\Psi_{0,F} = \text{Id}$ and $\partial/\partial t|_{t=0} \Psi_{t,F}(p) = F(p)$ as well as scaling homogeneity $\Psi_{h,F} = \Psi_{1,hF}$.

Many algebraic aspects of numerical integration are related to the computation of compositions, logarithms and exponentials of numerical integrators. In this introduction we will introduce some basic algebraic structures arising in the analysis of numerical integrators. In particular we will focus on structures that originate from the study of *Lie group integrators*, which are numerical integrators on general manifolds. The resulting theory combines Lie theory with the classical Butcher theory that describes numerical integrators on \mathbb{R}^n .

This first section presents a survey of well known results from numerical analysis. A detailed understanding of this introductory section is not necessary for

reading the rest of the paper, and readers mainly interested in algebraic structures may jump directly to Section 3.

2.1. Classical integrators. In the early 1960s, John Butcher set out to explore the algebraic territory of numerical algorithms for integrating ODEs evolving on vector spaces

$$(2.1) \quad y'(t) = F(y), \quad y \in \mathbb{R}^n, \quad F: \mathbb{R}^n \rightarrow \mathbb{R}^n.$$

In particular he studied the family of Runge–Kutta methods. Given a time step $h \in \mathbb{R}$, these methods advance the solution from $y_0 = y(0)$ to $y_1 \approx y(h)$ as:

```

for  $r = 1 : s$  do
   $Y_r = \sum_{k=1}^s a_{rk} F_k + y_0$ 
   $F_r = hF(Y_r)$ 
end
 $y_1 = \sum_{k=1}^s b_k F_k + y_0.$ 

```

This basic step is iterated: $y_0 \mapsto y_1 \mapsto \dots \mapsto y_n$, with constant or variable step sizes h , until the final solution $y_n \approx y(t_n)$ is reached. The coefficients a_{rk} and b_k for $r, k \in \{1, \dots, s\}$ define a particular s -stage RK method.

A goal of numerical analysis is to characterize coefficients a_{rk} and b_k that yield ‘good’ methods (when applied to a given class of differential equations). The view of what a good integration method is has, however, evolved over the last decades. Traditionally, *order theory* and *stability* were the most important properties to consider. A numerical integrator is of order p if the first $p + 1$ terms of the Taylor expansion of the analytical solution agrees with the first $p + 1$ terms of the numerical method (developed in the parameter h). Requiring a certain order results in algebraic conditions, called *order conditions*, on the coefficients of the method.

Many numerical methods for solving the equation (2.1) can be studied by using *B-series* (see e.g. [25]), introduced by Hairer and Wanner in 1974 [26]. A B-series is a (formal) series indexed over the set of rooted trees T , and can for a vector field F be written as

$$(2.2) \quad B_h(a)(y) = a(\mathbb{I})y + \sum_{\tau \in T} \frac{h^{|\tau|}}{\sigma(\tau)} a(\tau) \mathcal{F}_F(\tau)(y).$$

Here a is a map $a: T \rightarrow \mathbb{R}$, \mathbb{I} is the empty tree, $|\tau|$ is the number of vertices of τ (the *order* of the tree τ) and σ is a certain symmetry factor. The map $\mathcal{F}_F(\tau): \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the *elementary differential* of the tree τ and is given recursively as follows:

$$(2.3) \quad \mathcal{F}_F(\tau) = F^{(m)} (\mathcal{F}_F(\tau_1)(y), \dots, \mathcal{F}_F(\tau_m)(y)) (y),$$

where $\tau = B^+(\tau_1, \dots, \tau_m)$ is the tree constructed by adding a common root to the subtrees $\tau_1 \dots \tau_m$, and $F^{(m)}$ is the m th derivative of the vector field.

One way in which B-series can be applied to the study of numerical methods is to order theory. For example, the order conditions for Runge–Kutta methods can easily be obtained by writing the method as a B-series and then comparing the coefficients of this series with the exact solution written as a B-series (see e.g. [25, Chapter. III.1.2]).

The composition of Runge–Kutta methods is also of great interest, and this leads to the study of the composition of B-series. A series $B_{hF}(a)$ is inserted into another series $B_{hF}(b)$, which gives the B-series $B_{hF}(a)(B_{hF}(b)(y)) = B_{hF}(a \cdot b)(y)$. The resulting product $a \cdot b$ gives rise to a group, called the *Butcher group* [9, 16].

Butcher realized early on that the set of Runge–Kutta methods forms a group, and characterized algebraically the composition and inverse in this group. Much later, this group was identified with the character group of the Connes–Kreimer Hopf algebra [17, 19, 5].

In recent years the importance of preserving various geometric properties of the underlying continuous dynamical system has become better understood. The research topic *Geometric Numerical Integration* [25] emphasizes this view. Geometric integration algorithms have been successfully developed for various classes of differential equations, such as volume preserving flows, Hamiltonian equations, systems with first integrals and equations evolving on manifolds. An important tool for investigating the geometrical properties of a numerical integrator is through *backward error analysis*. For a given numerical method $\Psi_{h,F}$, we seek a series expansion of a modified vector field $(h, F) \mapsto \tilde{F}_h$ such that the numerical solution equals¹ the analytical flow of the modified vector field: $\Psi_{h,F} = \Phi_{t,\tilde{F}_h} \Big|_{t=h}$.

This is computed as a formal logarithm $\tilde{F}_h = \text{Log}(\Psi_{h,F})$, which in Hopf algebraic language is expressed by the *Eulerian idempotent* (Section 3.3.2).

Still another idea, which has been developed in [14], is to ask for a series development of a modified vector field \bar{F}_h such that when the numerical method is applied to \bar{F}_h , the exact analytical solution is produced: $\Psi_{h,\bar{F}_h} = \Phi_{h,F}$. This has been taken much further in recent work [15, 10]. The algebraic operation $(h, F) \mapsto \bar{F}_h$ is commonly referred to as a *substitution law*. The Hopf algebra of the substitution law is introduced in [10].

The theory of B-series is often a very important component in a numerical analyst’s toolbox, and is used to study all of the above: order theory, backward error analysis, modified vector fields and structure preserving properties of numerical integrators.

2.2. Lie group integrators. Numerical Lie group integrators for ODEs is a generalization of numerical integration of ODEs from the classical setting of equations on \mathbb{R}^n to differential equations on manifolds. See [28] for an extensive survey.

2.2.1. Exponential Euler method. Let \mathcal{M} denote a manifold, $\mathcal{X}\mathcal{M}$ its Lie algebra of vector fields with the Jacobi bracket and $\text{Diff}(\mathcal{M})$ the group of diffeomorphisms on \mathcal{M} . Let $\exp : \mathcal{X}\mathcal{M} \rightarrow \text{Diff}(\mathcal{M})$ denote the flow operator. We want to numerically integrate an ODE on \mathcal{M} given as

$$(2.4) \quad y'(t) = F(y), \quad y(0) = y_0 \quad \text{for } F \in \mathcal{X}\mathcal{M},$$

with the analytical solution $y(t) = \exp(tF) \cdot y_0$. Here $\exp(tF) \cdot y_0$ denotes the evaluation of the diffeomorphism $\exp(tF)$ at $y_0 \in \mathcal{M}$.

Assumption 2.1. *The fundamental assumption for numerical Lie group integrators is the existence of a subalgebra $\mathfrak{g} \subset \mathcal{X}\mathcal{M}$ such that*

- All vector fields $V \in \mathfrak{g}$ can be exponentiated exactly.
- The Lie algebra \mathfrak{g} defines a frame on $T\mathcal{M}$, i.e. \mathfrak{g} spans the tangent space $T_p\mathcal{M}$ at all points $p \in \mathcal{M}$. In other words, the action generated by \mathfrak{g} is transitive on \mathcal{M} .

¹The series for \tilde{F}_h is a formal series which usually does not converge. By truncating the series at an optimal point we find a modified equation which is exponentially close to the numerical solution, see [25]. In this paper we deal only with formal series, and convergence is not considered.

The vector fields in \mathfrak{g} are called the *frozen vector fields*. Due to the frame assumption, we can always express the vector field F and the ODE (2.4) in terms of frozen vector fields via a function $f : \mathcal{M} \rightarrow \mathfrak{g}$ as $F(y) = f(y) \cdot y$, where $f(y) \in \mathcal{X}\mathcal{M}$ and $f(y) \cdot y$ denotes evaluation of this vector field in y . Thus (2.4) can be written in the form

$$(2.5) \quad y'(t) = f(y) \cdot y, \quad y(0) = y_0 \in \mathcal{M}.$$

In the case where \mathfrak{g} forms a basis for $T_y\mathcal{M}$, the function $f(y)$ is uniquely defined. In more general situations, \mathfrak{g} is an overdetermined frame for $T_y\mathcal{M}$, and there is a freedom in the choice of f . This is called a choice of isotropy, and is of major importance for the quality of the numerical integrator.

With the equation written as (2.5), we can present the simplest of all Lie group integrators: the *exponential Euler method*. Given a time step $h \in \mathbb{R}$, the method advances the solution from $y_0 = y(0)$ to $y_1 \approx y(h)$ as:

Algorithm 2.2 (Exponential Euler).

$$y_1 = \exp(hf(y_0)) \cdot y_0.$$

In each step the solution is advanced $y_k \mapsto y_{k+1}$ by integrating the frozen vector field equation

$$y'(t) = f(y_k) \cdot y, \quad y(0) = y_k$$

from $t = 0$ to $t = h$. We will in the sequel present methods of higher order and with superior qualities compared to this simple scheme. The main theme of the paper is the algebraic structures arising from the numerical analysis of such integration schemes.

2.2.2. Choosing a good action. In practice it is of importance that $\exp : \mathfrak{g} \rightarrow \text{Diff}(\mathcal{M})$ can be computed fast, and furthermore that the given vector field $F(y)$ is locally well approximated by $f(y_0) \cdot y$. Exactly what ‘well’ means depends on what we want to achieve. In many situations we can choose \mathfrak{g} so that certain first integrals of the original system are exactly preserved by the frozen flows. Choosing \mathfrak{g} and $f : \mathcal{M} \rightarrow \mathfrak{g}$ is in many ways similar to choosing a *preconditioner* in iterative methods for solving linear equations: we want a good approximation which is easy to compute.

A simple choice of \mathfrak{g} is obtained by embedding $\mathcal{M} \subset \mathbb{R}^N$ and choosing $\mathfrak{g} = \mathbb{R}^N$ as the (commutative) algebra generated by $\{\partial/\partial x_j\}_{j=1}^N$, i.e., the constant vector fields on \mathbb{R}^N . Since the vector fields are constant, we have $f(y_0) \cdot y = f(y_0)$, so the function f simply becomes $f(y_0) = F(y_0) \in \mathbb{R}^N$, all commutators in \mathfrak{g} vanish and the exponential on \mathfrak{g} is $\exp(V) \cdot p = V + p$ for $V, p \in \mathbb{R}^N$. In this case all Lie group integrators will reduce to classical integrators, e.g. exponential Euler becomes the classical Euler $y_1 = hF(y_0) + y_0$.

The other extreme is $\mathfrak{g} = \mathcal{X}\mathcal{M}$ and $f(y) = F$ for all y , in which case exponential Euler yields the analytical solution exactly. However, the computation of the exponential on \mathfrak{g} is just as difficult as solving the original equation. We seek efficient choices in between these two extremes.

In many cases \mathfrak{g} is given as the infinitesimal generators of a (e.g. left) Lie group action on \mathcal{M} . For example, consider the sphere $\mathcal{M} = S^2$ acted upon from left by the group $G = \text{SO}(3)$ of orthogonal 3×3 matrices, whose Lie algebra $\mathfrak{so}(3)$ consists of skew 3×3 matrices. Any matrix $V \in \mathfrak{so}(3)$ is uniquely identified

with the infinitesimal generator² $\xi_V \in \mathcal{XM}$, given by matrix-vector multiplication $\xi_V \cdot y = Vy$ for $y \in S^2$. Therefore (2.5) becomes $y'(t) = V(y)y$, where $V(y)$ is a skew symmetric matrix. The exponentiation is related to the matrix exponential $\text{exp}(V)$ as $\exp(\xi_V) \cdot y = \text{exp}(V)y$.

Another important example of group actions arise in the solution of *isospectral differential equations*, where $\text{GL}(n)$ acts on $\mathfrak{gl}(n)$ by the *adjoint action* (similarity transform) $A \cdot Y = AYA^{-1}$ for $A \in \text{GL}(n)$, $Y \in \mathfrak{gl}(n)$. In these problems $\mathcal{M} \subset \mathfrak{gl}(n)$ is one of the (isospectral) orbits of the action. For this action (2.5) acquires the isospectral form $Y'(t) = [B(Y), Y]$ for some $B(Y) \in \mathfrak{gl}(n)$. Since the action is a similarity transform it is guaranteed that all the eigenvalues of $Y(t)$ are preserved also by the numerical integrator.

Yet another example, which occurs in *Lie–Poisson problems* in computational mechanics, is the *coadjoint action* of a Lie group on the dual of its Lie algebra, \mathfrak{g}^* . In this case $\mathcal{M} \subset \mathfrak{g}^*$ is a coadjoint orbit. Using this action we can guarantee that the numerical Lie group integrator exactly preserves the Casimirs of the continuous system.

For other problems it may be advantageous to choose \mathfrak{g} by simplifying the original equation to a family of integrable equations. An example is the computation of the motion of charged particles in a magnetic field. The solution in the case of constant magnetic fields is given by helical motions around the field lines. The corresponding Lie algebra yields fast and accurate Lie group integrators for the full problem of non-constant magnetic fields. Another example is integration of a spinning top, where we obtain simpler equations by considering the direction of gravity as being constant in body coordinates. In both these problems, the action preserves important first integrals of the system. A third example is integration of stiff equations on \mathbb{R}^n , where an integrable Lie algebra is obtained by considering all affine linear vector fields. This connects the theory of Lie group integrators with the so-called *exponential integrators*. See [28] for details.

2.2.3. Higher order methods. Most Lie group methods for integrating (2.5) are built from linear operations and commutators in \mathfrak{g} and computation of flows of frozen vector fields (exponentials). Runge–Kutta type methods with basic motions expressed in terms of an exponential of a sum of elements in \mathfrak{g} are commonly referred to as RKM methods [28], as in the following example:

Algorithm 2.3 (4th order RKM from [39]).

$$\begin{aligned}
 Y_1 &= y_0 & F_1 &= hf(Y_1) \\
 Y_2 &= \exp\left(\frac{1}{2}F_1\right) \cdot y_0 & F_2 &= hf(Y_2) \\
 Y_3 &= \exp\left(\frac{1}{2}F_2 + \frac{1}{24}[F_1, F_2]\right) \cdot y_0 & F_3 &= hf(Y_3) \\
 Y_4 &= \exp\left(F_3 + \frac{1}{6}[F_1, F_3]\right) \cdot y_0 & F_4 &= hf(Y_4) \\
 V &= \frac{1}{6}F_1 + \frac{1}{3}(F_2 + F_3) + \frac{1}{6}F_4 & I &= \frac{1}{8}F_1 + \frac{1}{12}(F_2 + F_3) - \frac{1}{24}F_4 \\
 y_1 &= \exp(V + [I, V]) \cdot y_0.
 \end{aligned}$$

Methods where the basic motions are products of exponentials of simple elements in \mathfrak{g} are called *Crouch–Grossman methods* [18, 45].

²Recall that the identification of the Lie algebra of a left group action with the infinitesimal generator in \mathcal{XM} is an anti-homomorphism, $[\xi_V, \xi_W] = -\xi_{[V, W]}$. In this paper the brackets are Jacobi brackets on \mathcal{XM} , and some signs may differ when compared to cited papers.

Algorithm 2.4 (3rd order Crouch–Grossman method from [45]).

$$\begin{aligned} Y_1 &= y_0 & F_1 &= hf(Y_1) \\ Y_2 &= \exp\left(\frac{3}{4}F_1\right) \cdot y_0 & F_2 &= hf(Y_2) \\ Y_3 &= \exp\left(\frac{119}{216}F_2\right) \cdot \exp\left(\frac{17}{108}F_1\right) \cdot y_0 & F_3 &= hf(Y_3) \\ y_1 &= \exp\left(\frac{13}{51}F_3\right) \cdot \exp\left(-\frac{2}{3}F_2\right) \cdot \exp\left(\frac{24}{17}F_3\right) y_0. \end{aligned}$$

More recently methods have been developed which combine exponentials of sums and products of exponentials, as in the *commutator free Lie group methods* [13]. An example is:

Algorithm 2.5 (4th order commutator free method from [13]).

$$\begin{aligned} Y_1 &= y_0 & F_1 &= hf(y_0) \\ Y_2 &= \exp\left(\frac{1}{2}F_1\right) \cdot y_0 & F_2 &= hf(Y_2) \\ Y_3 &= \exp\left(\frac{1}{2}F_2\right) \cdot y_0 & F_3 &= hf(Y_3) \\ Y_4 &= \exp\left(-\frac{1}{2}F_1 + F_3\right) \cdot Y_2 & F_4 &= hf(Y_4) \\ y_1 &= \exp\left(\frac{1}{4}F_1 + \frac{1}{6}(F_2 + F_3) - \frac{1}{12}F_4\right) \cdot \exp\left(-\frac{1}{12}F_1 + \frac{1}{6}(F_2 + F_3) + \frac{1}{4}F_4\right) \cdot y_0. \end{aligned}$$

For equations of *Lie type*, $y'(t) = f(t) \cdot y$, numerical methods based on Magnus and Fer expansions have been developed in [29, 27], and the algebraic theory has recently been developed further in [21].

To study order conditions, backward error analysis and structure preservation of such methods, it is important to understand B-series in a general setting of group actions on manifolds. A first attempt at combining Lie and B-series in a common mathematical framework appeared in [38, 39]. Hopf algebraic aspects have been explored further in [40, 3, 42].

3. Hopf algebras

This section gives a short collection of some facts and properties of Hopf algebras that we will use in this work. For a more thorough introduction, see e.g. [1], [35], [48], [30], [11].

3.1. Basic definitions. Let k be a field containing \mathbb{Q} .

Definition 3.1. A k -algebra A consists of a k -vector space A together with two maps $\mu : A \otimes A \rightarrow A$ and $\eta : k \rightarrow A$, called the *product* and the *unit* of A , such that:

- (i) μ is associative, i.e. $\mu \circ (I \otimes \mu) = \mu \circ (\mu \otimes I)$,
- (ii) the composites $A \cong A \otimes k \xrightarrow{I \otimes \eta} A \otimes A \xrightarrow{\mu} A$ and $A \cong A \otimes k \xrightarrow{\eta \otimes I} A \otimes A \xrightarrow{\mu} A$ both equal I .

Here I denotes the identity map.

An algebra A is called commutative if $\mu \circ \tau = \mu$, where τ is the *flip* map $\tau(a_1 \otimes a_2) = a_2 \otimes a_1$.

Definition 3.2. A k -coalgebra C is a k -vector space equipped with two maps $\Delta : C \rightarrow C \otimes C$ and $\epsilon : C \rightarrow k$, the *coproduct* and the *counit*, such that:

- (i) Δ is coassociative, i.e. $(\Delta \otimes I) \circ \Delta = (I \otimes \Delta) \circ \Delta$,
- (ii) the composites $C \xrightarrow{\Delta} C \otimes C \xrightarrow{I \otimes \epsilon} C \otimes k \cong C$ and $C \xrightarrow{\Delta} C \otimes C \xrightarrow{\epsilon \otimes I} k \otimes C \cong C$ both equal I .

A coalgebra C is called cocommutative if $\tau \circ \Delta = \Delta$.

Definition 3.3. A *bialgebra* H over k is a k -vector space equipped with both an algebra (H, μ, η) and a coalgebra structure (H, Δ, ϵ) , such that the coproduct $\Delta : H \rightarrow H \otimes H$ and the counit $\epsilon : H \rightarrow k$ are algebra morphisms. These compatibility conditions can be expressed in terms of the following commutative diagrams³, where τ denotes the *flip* operation $\tau(h_1, h_2) = (h_2, h_1)$:

$$\begin{array}{ccc}
 H^{\otimes 4} & \xrightarrow{I \otimes \tau \otimes I} & H^{\otimes 4} \\
 \Delta \otimes \Delta \uparrow & & \downarrow \mu \otimes \mu \\
 H \otimes H & \xrightarrow{\mu} & H \xrightarrow{\Delta} H \otimes H
 \end{array}
 \qquad
 \begin{array}{ccc}
 H \otimes H & \xrightarrow{\epsilon \otimes \epsilon} & k \otimes k \\
 \mu \downarrow & & \downarrow \cong \\
 H & \xrightarrow{\epsilon} & k
 \end{array}$$

A bialgebra H is called commutative if it is commutative as an algebra, and cocommutative if it is cocommutative as a coalgebra.

Remark 3.4. There is symmetry in the definition of a bialgebra. Rather than requiring the coalgebra structure to respect the algebra structure in the above sense, we could have switched the role of the two structures. To complete the symmetry, we could in addition reverse the arrows in the two diagrams above. This would result in an equivalent definition.

Grading. Let H be a *graded* k -vector space, i.e. $H = \bigoplus_{n \geq 0} H_n$. There is a notion of a graded bialgebra, obtained by requiring the following of the algebra and coalgebra structure, respectively:

- (i) $\mu(H_p, H_q) \subset H_{p+q}$
- (ii) $\Delta(H_n) \subset \bigoplus_{p+q=n} H_p \otimes H_q$.

The grading of an algebra H gives rise to the *grading operator* $Y : H \rightarrow H$ given by

$$Y : h \mapsto \sum_{k \geq 0} kh_k,$$

where $h = \sum_{n \geq 0} h_n \in \bigoplus_{n \geq 0} H_n$. A graded bialgebra $H = \bigoplus_{n \geq 0} H_n$ is called *connected* if $H_0 = k$.

Proposition 3.5 ([35]). *Let H be a connected, graded bialgebra. Then, for any $x \in H_n$, $n \geq 0$, we have:*

$$\Delta x = 1 \otimes x + x \otimes 1 + \tilde{\Delta}x, \text{ where } \tilde{\Delta}x \in \bigoplus_{p+q=n, p,q>0} H_p \otimes H_q.$$

We will often use the *Sweedler notation* for the coproduct:

$$\Delta x = \sum_{(x)} x_{(1)} \otimes x_{(2)} \quad \text{and} \quad \tilde{\Delta}x = \sum_{(x)} x' \otimes x''.$$

³All diagrams were created using Paul Taylor's diagram package, available from <http://www.paultaylor.eu/diagrams/>

Definition 3.6. A *Hopf algebra* is a bialgebra $(H, \mu, \eta, \Delta, \epsilon)$ together with an antihomomorphism S on H , called the *antipode*, with the property given by the commutativity of the following diagram:

$$\begin{array}{ccccc}
 & H \otimes H & \xrightarrow{S \otimes 1} & H \otimes H & \\
 \swarrow \triangleright & & & & \searrow \triangleleft \\
 H & \xrightarrow{\epsilon} & k & \xrightarrow{\eta} & H \\
 \searrow \triangleleft & & & & \swarrow \triangleright \\
 & H \otimes H & \xrightarrow{1 \otimes S} & H \otimes H &
 \end{array}$$

A Hopf algebra is graded if it is graded as a bialgebra and the antipode satisfies $S(H_n) \subset H_n$. If a bialgebra is graded and connected, then it is automatically a graded Hopf algebra:

Proposition 3.7 ([35]). *Any connected graded bialgebra is a Hopf algebra. The antipode S is given recursively by $S(1) = 1$ and*

$$S(x) = -x - \sum_{(x)} S(x')x''$$

for $x \in \ker \epsilon$.

3.2. Examples: The concatenation and shuffle Hopf algebras. Recurring in the sequel are Hopf algebras built from letters in an alphabet. We follow the notation of Reutenauer [47]. Consider a finite or infinite alphabet of letters $\mathcal{A} = \{a, b, c, \dots\}$. We write \mathcal{A}^* for the collection of all empty or non-empty words over \mathcal{A} , where \mathbb{I} is the empty word. Let $k\langle\mathcal{A}\rangle$ be the k -algebra of non-commutative polynomials in \mathcal{A} . A polynomial $P \in k\langle\mathcal{A}\rangle$ will be written as a sum

$$P = \sum_{\omega \in \mathcal{A}^*} (P, \omega) \omega,$$

where $(P, \omega) \in k$ is non-zero only for a finite number of ω . Let $P, Q \in k\langle\mathcal{A}\rangle$. The product of P and Q , written as PQ , has coefficients

$$(PQ, \omega) = \sum_{\omega=uv} (P, u)(Q, v).$$

The k -linear dual space denoted $k\langle\langle\mathcal{A}\rangle\rangle := \text{Hom}_k(k\langle\mathcal{A}\rangle, k)$ is identified with all infinite k -linear combinations of words. An $\alpha \in k\langle\langle\mathcal{A}\rangle\rangle$ can be written as an infinite series

$$\alpha = \sum_{\omega \in \mathcal{A}^*} (\alpha, \omega) \omega,$$

where $(\alpha, \omega) \equiv \alpha(\omega) \in k$ and (\cdot, \cdot) is the dual pairing defined such that words in \mathcal{A}^* are orthogonal, $(\omega_1, \omega_2) = \delta_{\omega_1, \omega_2}$ for all $\omega_1, \omega_2 \in \mathcal{A}^*$.

We define two different associative products on $k\langle\mathcal{A}\rangle$. The *concatenation product* $\omega_1, \omega_2 \mapsto \omega_1\omega_2$ obtained by concatenation of words and the *shuffle product* $\omega_1, \omega_2 \mapsto \omega_1 \sqcup \omega_2$ obtained by linearly combining all possible *shuffles* of the two words i.e. combinations where the letters within each word are not internally permuted:

$$abc \sqcup de = abcde + abdce + adbce + dabce + abdec + adbec + dabec + adebc + daebc + deabc.$$

The shuffle product can be defined recursively as

$$(a\omega_1) \sqcup (b\omega_2) = a(\omega_1 \sqcup b\omega_2) + b(a\omega_1 \sqcup \omega_2),$$

where $a, b \in \mathcal{A}$ and $\omega_1, \omega_2 \in \mathcal{A}^*$. The unit of both concatenation and shuffle is the empty word \mathbb{I} .

By dualization of these products we obtain the *deconcatenation* and the *deshuffle* coproducts. The deconcatenation coproduct $\Delta_d: k\langle\mathcal{A}\rangle \rightarrow k\langle\mathcal{A}\rangle \otimes k\langle\mathcal{A}\rangle$ is defined for $\omega = a_1 a_2 \cdots a_k \in \mathcal{A}^*$ as:

$$(3.1) \quad \Delta_d(\omega) = \sum_{i=1}^k a_1 \cdots a_i \otimes a_{i+1} \cdots a_k.$$

This coproduct is the dual of the concatenation product, so for any $P, Q \in k\langle\mathcal{A}\rangle$

$$(PQ, \omega) = (P \otimes Q, \Delta_d(\omega)) = \sum_{(\omega)_{\Delta_d}} (P, \omega_{(1)})(Q, \omega_{(2)}).$$

The deshuffle product $\Delta_{\sqcup}: k\langle\mathcal{A}\rangle \rightarrow k\langle\mathcal{A}\rangle \otimes k\langle\mathcal{A}\rangle$ is similarly defined such that

$$(P \sqcup Q, \omega) = (P \otimes Q, \Delta_{\sqcup}(\omega)) = \sum_{(\omega)_{\Delta_{\sqcup}}} (P, \omega_{(1)})(Q, \omega_{(2)}).$$

The two coproducts can also be characterized by requiring that the letters in the alphabet \mathcal{A} are primitive, i.e. that $\Delta(a) = 1 \otimes a + a \otimes 1$ for $a \in \mathcal{A}$, and then extending Δ to be a homomorphism with respect to either of the two products on $k\langle\mathcal{A}\rangle$. We refer to [47] for explicit presentations of the deshuffle coproduct.

We remark that the vector space $k\langle\mathcal{A}\rangle$ can now be turned into Hopf algebras in two different ways. The cocommutative *concatenation Hopf algebra* is obtained by taking the concatenation as product and the deshuffle as coproduct. The commutative *shuffle Hopf algebra* $\mathcal{H}_{\text{Sh}}(\mathcal{A})$ is obtained by taking the shuffle as product and the deconcatenation as coproduct. Both these Hopf algebras share the same antipode:

$$(3.2) \quad S(a_1 a_2 \cdots a_k) = (-1)^k a_k a_{k-1} \cdots a_1,$$

and in both cases the unit and counit is given by $\eta(\mathbb{I}) = \mathbb{I}$ and $\epsilon(\mathbb{I}) = 1$, $\epsilon(\omega) = 0$ for all $\omega \in \mathcal{A}^* \setminus \mathbb{I}$. We write \mathcal{H}_{Sh} when \mathcal{A} is understood.

The vector space $k\langle\mathcal{A}\rangle$ can be identified with the vector space underlying the tensor algebra $T(V)$ on the vector space V generated by the alphabet \mathcal{A} . The two algebra structures (concatenation and shuffling) correspond to the usual algebra structures given to the tensor algebra $T(V)$ and the tensor coalgebra $T^c(V)$, respectively.

3.3. Characters and endomorphisms. This section is based on [20]. See also [35], [7] and [46].

Let $(H, \mu, \eta, \Delta, \epsilon)$ be a graded bialgebra, and (A, \cdot, η_A) an algebra. The set $\text{Hom}_k(H, A)$ of linear maps from H to A sending $\eta_H(1) =: 1_H$ to $\eta_A(1) =: 1_A$ has an algebra structure given by the *convolution product*:

$$\alpha * \beta = \mu_A \circ (\alpha \otimes \beta) \circ \Delta.$$

The convolutional unit is the composition of the counit of H and the unit of A : $\delta := \eta_A \circ \epsilon$. The convolution can be written using the Sweedler notation:

$$\alpha * \beta = \sum_{(x)} \alpha(x_{(1)}) \cdot \beta(x_{(2)}),$$

from which we find $\alpha * \delta = \delta * \alpha = \alpha$. The *unital algebra morphisms* from H to A consists of all $\alpha \in \text{Hom}_k(H, A)$ such that $\alpha(1_H) = 1_A$ and $\alpha(\mu(h, h')) = \alpha(h) \cdot \alpha(h')$ for all $h, h' \in H$.

Proposition 3.8 ([35]). *Let H be a graded Hopf algebra and A a commutative algebra. The set $\text{Hom}_{Alg}(H, A)$ of unital algebra morphisms from H to A equipped with the convolution product, forms a group, $G(H, A)$, called the group of A -valued characters of H . The inverse of an element α is given by*

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

where S is the antipode of H .

In the special case $A = k$, we get the group of characters of H , written as $G(H) := G(H, k)$. The grading on H splits the group of A -valued characters into graded components:

$$G(H, A) \cong \prod_{n \geq 0} \text{Hom}_{Alg}(H_n, A).$$

This is not a graded vector space, but rather the completion of one (see e.g. [20]), but we will still refer to it as a graded vector space. The restriction of a character $\alpha : H \rightarrow A$ to the degree n component H_n of H will be denoted by α_n .

3.3.1. Infinitesimal characters, the exponential and the logarithm. The *infinitesimal A -valued characters*, written $\mathfrak{g}(H, A)$ are the linear maps α from H to A such that:

$$\alpha(\mu(h, h')) = \alpha(h) \cdot \delta(h') + \delta(h) \cdot \alpha(h'),$$

where $\delta = \eta_A \circ \epsilon$. This is a Lie algebra under the bracket induced by the convolution product: $[\alpha, \beta] = \alpha * \beta - \beta * \alpha$. In the special case where $A = k$ we write $\mathfrak{g}(H)$ for $\mathfrak{g}(H, k)$.

The characters and the infinitesimal characters are connected via the *exponential* and the *logarithmic* map. For $\alpha \in \text{Hom}_k(H, A)$, the exponential and logarithm with respect to convolution are given by the formal series:

$$\begin{aligned} \exp^*(\alpha) &= \sum_{n \geq 0} \frac{1}{n!} \alpha^{*n} \\ \log^*(\delta + \alpha) &= \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \alpha^{*n}. \end{aligned}$$

If H is graded and connected, and if $\alpha(1) = 0$, where $1 \in k = H_0$, then $\alpha^{*k} = \alpha * \dots * \alpha = 0$ on H_n for $n < k$, and therefore both these sums are finite when restricted to H_n . The maps \exp^* and \log^* give a bijection between $G(H, A)$ and $\mathfrak{g}(H, A)$.

Example 3.9. Let \mathcal{H}_{Sh} denote the shuffle algebra over \mathcal{A} . Consider the dual space $k\langle\langle \mathcal{A} \rangle\rangle$ equipped with the convolution product

$$(\alpha * \beta, \omega) = \sum_{(\omega)_{\Delta_d}} (\alpha, \omega_{(1)}) (\beta, \omega_{(2)}) = \sum_{\omega=\omega_1\omega_2} (\alpha, \omega_1) (\beta, \omega_2).$$

Note that convolution is just concatenation of series $\alpha * \beta = \alpha\beta$. The characters and infinitesimal characters $\mathfrak{g}(\mathcal{H}_{\text{Sh}}), G(\mathcal{H}_{\text{Sh}}) \subset k\langle\langle \mathcal{A} \rangle\rangle$ are given as

$$\begin{aligned}\mathfrak{g}(\mathcal{H}_{\text{Sh}}) &= \{ \alpha \in k\langle\langle \mathcal{A} \rangle\rangle \mid \alpha(\mathbb{I}) = 0 \text{ and } \alpha(\omega_1 \sqcup \omega_2) = 0 \text{ for all } \omega_1, \omega_2 \in \mathcal{A}^* \setminus \mathbb{I} \} \\ G(\mathcal{H}_{\text{Sh}}) &= \{ \alpha \in k\langle\langle \mathcal{A} \rangle\rangle \mid \alpha(\mathbb{I}) = 1 \text{ and } \alpha(\omega_1 \sqcup \omega_2) = \alpha(\omega_1)\alpha(\omega_2) \ \forall \omega_1, \omega_2 \in \mathcal{A}^* \setminus \mathbb{I} \}.\end{aligned}$$

The convolutional unit δ is given as $(\delta, \mathbb{I}) = 1$ and $(\delta, \omega) = 0$ for all $\omega \in \mathcal{A} \setminus \mathbb{I}$. The logarithm of $\alpha \in G(\mathcal{H}_{\text{Sh}})$ can be computed as $\log(\alpha) = \sum_{n>0} \frac{(-1)^{n-1}}{n} (\alpha - \delta)^{*n}$. For any $\omega \in \mathcal{A}^*$ we find that $(\log(\alpha), \omega)$ is given by a finite sum expressed in terms of the Eulerian idempotent.

3.3.2. Eulerian idempotent. Let H be a commutative, connected and graded Hopf algebra. Consider $\text{End}_k(H) = \text{Hom}_k(H, H)$ equipped with the convolution product $*$. Let $\text{Id} \in \text{End}_k(H)$ be the identity endomorphism and $\delta = \eta \circ \epsilon \in \text{End}_k(H)$ the unit of convolution.

Definition 3.10 ([32]). The Eulerian idempotent $e \in \text{End}(H)$ is given by the formal power series

$$e := \log^*(\text{Id}) = J - \frac{J^{*2}}{2} + \frac{J^{*3}}{3} + \cdots + (-1)^{i+1} \frac{J^{*i}}{i} + \cdots,$$

where $J = \text{Id} - \delta$.

Proposition 3.11 ([32]). For any commutative graded Hopf algebra H , the element $e \in \text{End}_k(H)$ defined above is an idempotent: $e \circ e = e$.

The practical importance of the Eulerian idempotent in numerical analysis arises in backward error analysis, where the following lemma provides a computational formula for the logarithm:

Proposition 3.12. For $\alpha \in G(H)$ and $h \in H$, we have

$$\log^*(\alpha)(h) = \alpha(e(h)).$$

In other words, the logarithm can be written as right composition with the eulerian idempotent:

$$\log^* = _ \circ e : G(H) \rightarrow \mathfrak{g}(H).$$

The result follows from the following computation, which uses that α is a homomorphism:

$$((\alpha - \delta)^{*l}, \omega) = \mu_k^l \circ (\alpha \otimes \cdots \otimes \alpha) \circ \tilde{\Delta}^l \omega = \alpha \circ \mu_H^l \circ \tilde{\Delta}^l \omega = \alpha \circ J^{*l} \omega,$$

where $(-)^l$ denotes l -fold application.

3.3.3. The graded Dynkin operator. There is another bijection between the infinitesimal characters and the characters in any commutative graded Hopf algebra H , described in [20]. The bijection is given in terms of the *Dynkin operator* $D : H \rightarrow H$.

Classically, the Dynkin operator is a map $D : k\langle\mathcal{A}\rangle \rightarrow \text{Lie}(\mathcal{A})$, where $\text{Lie}(\mathcal{A}) = \mathfrak{g}(\mathcal{H}_{\text{Sh}}) \cap k\langle\mathcal{A}\rangle$ are the Lie polynomials. The classical Dynkin operator is given by left-to-right bracketing:

$$D(a_1 \dots a_n) = [\dots [[a_1, a_2], a_3], \dots, a_n], \quad \text{where } [a_i, a_j] = a_i a_j - a_j a_i.$$

Letting $Y(\omega) = \#(\omega)\omega$ denote grading operator, where $\#(\omega)$ is word length, it is known that the *Dynkin idempotent*, given as $Y^{-1}D$, is an idempotent projection on the subspace of Lie polynomials. As in [20], the Dynkin operator can be written

as the convolution of the antipode S and the grading operator $D = S * Y$. This description can be generalized to any graded, connected and commutative Hopf algebra H :

Definition 3.13. Let H be a graded, commutative and connected Hopf algebra with grading operator $Y : H \rightarrow H$. The *Dynkin operator* is the map $D : H \rightarrow H$ given as

$$D := S * Y.$$

Lemma 3.14 ([20]). *The Dynkin operator is a H -valued infinitesimal character of H .*

Theorem 3.15 ([20]). *Right composition with the Dynkin operator induces a bijection between $G(H)$ and $\mathfrak{g}(H)$:*

$$- \circ D : G(H) \rightarrow \mathfrak{g}(H).$$

The inverse is given by $\Gamma : \mathfrak{g}(H) \rightarrow G(H)$ as

$$(3.3) \quad \Gamma(\alpha) = \sum_n \sum_{\substack{k_1 + \dots + k_l = n, \\ k_1, \dots, k_l > 0}} \frac{\alpha_{k_1} * \dots * \alpha_{k_l}}{k_1(k_1 + k_2) \dots (k_1 + \dots + k_l)},$$

where $\alpha_k = \alpha|_{H_k}$.

Later we will apply the Dynkin operator and its inverse in the setting of a shuffle algebra $\mathcal{H}_{\text{Sh}}(\text{OT})$, where OT is an alphabet of all ordered rooted trees, and the grading $|\tau|$ of $\tau \in \text{OT}$ counts the nodes in the tree.

4. Algebras of formal diffeomorphisms on manifolds

The main goal of this section is to arrive at Lie–Butcher series and the underlying Hopf algebra \mathcal{H}_N . This Hopf algebra contains the Connes–Kreimer Hopf algebra as a subalgebra and is also closely related to \mathcal{H}_{Sh} . To emphasize the natural connection between Lie–Butcher series, \mathcal{H}_N and more classical Lie series, we start with a discussions of Lie series (autonomous and non-autonomous).

4.1. Autonomous Lie series. In this section we review the well-known theory of Lie series on manifolds and the corresponding Hopf algebraic structures of the free Lie algebra. The algebraic theory is detailed in [47, 11] and for the analytical theory we refer to [2].

Let F be a vector field on a manifold \mathcal{M} and $\Phi_{t,F} : \mathcal{M} \rightarrow \mathcal{M}$ its flow. Let $\psi : \mathcal{M} \rightarrow E$ be a section of a vector bundle over \mathcal{M} , and let $\Phi_{t,F}^* \psi$ denote the pullback. For the applications later in this paper we will only consider trivial bundles, in which case we write $\psi : \mathcal{M} \rightarrow \mathcal{V}$ for some vector space \mathcal{V} and define pullback as composition $\Phi_{t,F}^* \psi = \psi \circ \Phi_{t,F}$. The *Lie derivative* of ψ is defined as

$$(4.1) \quad F[\psi] = \left. \frac{\partial}{\partial t} \right|_{t=0} \Phi_{t,F}^* \psi.$$

Composition of Lie derivatives defines an associative, non-commutative product of vector fields $F, G \mapsto FG$, where vector fields are first order differential operators. The product FG is the second order differential operator $(FG)[\psi] = F[G[\psi]]$ etc. We let \mathbb{I} denote the 0th order identity operator $\mathbb{I}[\psi] = \psi$. The linear span of all differential operators of all orders forms the universal enveloping algebra $U(\mathcal{X}\mathcal{M})$.

The basic pullback formula is ([2]):

$$(4.2) \quad \frac{\partial}{\partial t} \Phi_{t,F}^* \psi = \Phi_{t,F}^*(F[\psi]).$$

Iterating this we find $\partial^n/\partial t^n|_{t=0} \Phi_{t,F}^* \psi = F[F[\cdots [\psi]]] := F^n[\psi]$, and hence follows the (Taylor)–Lie form of a *pullback series*:

$$\Phi_{t,F}^*[\psi] = \sum_{j=0}^{\infty} \frac{t^j}{j!} F^j[\psi] := \text{Exp}(tF)[\psi].$$

Fundamental questions are: Which series in $U(\mathcal{X}\mathcal{M})$ represent vector fields and which represent pullback series? How do we algebraically characterize compositions and the inverse of pullback series? How do we understand the Exp map taking vector fields to their pullback series, and what about the inverse Log operation? These questions are elegantly answered in terms of the shuffle Hopf algebra. We will detail these issues, and see that the same structures reappear in the discussion of B-series later.

An algebraic abstraction of Lie series starts with fixing a (finite or infinite) alphabet \mathcal{A} and a map $\nu: \mathcal{A} \rightarrow \mathcal{X}\mathcal{M}$ assigning each letter to a vector field. As in Example 3.2 we let $\mathbb{R}\langle\langle\mathcal{A}\rangle\rangle$ denote all finite \mathbb{R} -linear combinations of words built from \mathcal{A} and \mathcal{H}_{Sh} the shuffle algebra. The map ν can be uniquely extended to a linear $\mathcal{F}_\nu: \mathbb{R}\langle\langle\mathcal{A}\rangle\rangle \rightarrow U(\mathcal{X}\mathcal{M})$ as a concatenation homomorphism:

$$\begin{aligned} \mathcal{F}_\nu(\mathbb{I}) &= \mathbb{I}, \\ \mathcal{F}_\nu(a) &= \nu(a) \quad \text{for all letters } a \in \mathcal{A}, \\ \mathcal{F}_\nu(\omega_1\omega_2) &= \mathcal{F}_\nu(\omega_1)\mathcal{F}_\nu(\omega_2) \quad \text{for all words } \omega_1, \omega_2 \in \mathcal{A}^*. \end{aligned}$$

We extend \mathcal{F}_ν to a map \mathcal{B}_t taking an infinite series $\alpha \in \mathbb{R}\langle\langle\mathcal{A}\rangle\rangle$ to an infinite formal series $\mathcal{B}_t(\alpha) \in U(\mathcal{X}\mathcal{M})^*$, defined for $t \in \mathbb{R}$ as follows: Consider the alphabet \mathcal{A} with a grading $|a| \in \mathbb{N}^+$ for all $a \in \mathcal{A}$. This extends to \mathcal{H}_{Sh} as $|\omega| = |a_1| + \dots + |a_k|$ for all $\omega = a_1 \dots a_k \in \mathcal{A}^*$, $|\mathbb{I}| = 0$, thus \mathcal{H}_{Sh} becomes a graded connected Hopf algebra. Given the grading we define

$$(4.3) \quad \mathcal{B}_t(\alpha) = \sum_{\omega \in \mathcal{A}^*} t^{|\omega|} \alpha(\omega) \mathcal{F}_\nu(\omega).$$

Consider $\mathcal{H}_{\text{Sh}}^* = \mathbb{R}\langle\langle\mathcal{A}\rangle\rangle$ with the convolution $\alpha * \beta = \alpha\beta$ as in Example 3.9. By construction \mathcal{B}_t is a convolution homomorphism,

$$\mathcal{B}_t(\alpha * \beta) = \mathcal{B}_t(\alpha)\mathcal{B}_t(\beta).$$

For a real valued infinitesimal character $\alpha \in \mathfrak{g}(\mathcal{H}_{\text{Sh}})$, and a fixed $t = h$, $\mathcal{B}_h(\alpha)$ is a formal vector field on \mathcal{M} . For a real valued character $\beta \in G(\mathcal{H}_{\text{Sh}})$, $\mathcal{B}_h(\beta)$ represents a formal diffeomorphism Φ_h on \mathcal{M} via the pullback series

$$\mathcal{B}_h(\beta)[\psi] = \psi \circ \Phi_h \quad \text{for } \psi: \mathcal{M} \rightarrow \mathbb{R}.$$

Note, however, that pullbacks compose contravariantly with respect to composition of diffeomorphisms:

$$\mathcal{B}_h(\beta_1 * \beta_2)[\psi] = \mathcal{B}_h(\beta_1)\mathcal{B}_h(\beta_2)[\psi] = \psi \circ \Phi_2 \circ \Phi_1.$$

To summarize: Composition of diffeomorphisms is modelled by convolution in $G(\mathcal{H}_{\text{Sh}})$ (in opposite order), the inverse of a diffeomorphism is computed by right

composing with the antipode, the convolutional exponential maps to the exponential of Lie series and the logarithm is computed by composing with the Eulerian idempotent.

$$\begin{aligned}\mathcal{B}_h(\beta \circ S)\mathcal{B}_h(\beta) &= \mathbb{I} \quad \text{for } \beta \in G(\mathcal{H}_{\text{Sh}}) \\ \mathcal{B}_h(\exp^*(\alpha)) &= \text{Exp}(\mathcal{B}_h(\alpha)) \quad \text{for } \alpha \in \mathfrak{g}(\mathcal{H}_{\text{Sh}}) \\ \mathcal{B}_h(\beta) &= \text{Exp}(\mathcal{B}_h(\beta \circ e)) \quad \text{for } \beta \in G(\mathcal{H}_{\text{Sh}}).\end{aligned}$$

In the next section we discuss flows of non-autonomous equations, and we will see that right composition with the Dynkin idempotent represents algebraically the operation of finding a non-autonomous vector field corresponding to a diffeomorphism on a manifold.

Remark 4.1. In [41] the Lie algebra of infinitesimal characters $\mathfrak{g}(\mathcal{H}_{\text{Sh}})$ is studied as a graded free Lie algebra. An explicit formula for the dimension of the homogeneous components $\mathfrak{g}_k = \mathfrak{g}(\mathcal{H}_{\text{Sh}})|_k$ is derived for general gradings. This is very useful for the study of the complexity of Lie group integrators.

4.2. Time-dependent Lie series. The classical Faà di Bruno Hopf algebra models the composition of formal diffeomorphisms on \mathbb{R} ([23], [22], [24]). We will see that this has a natural generalization to compositions of time-dependent flows on manifolds. We introduce a *Dynkin–Faà di Bruno bialgebra* describing the composition of flows of time-dependent vector fields on a coarse level that considers only the grading of the terms in the t -expansion of the time-dependent vector fields.

4.2.1. Non-commutative Bell polynomials and Dynkin–Faà di Bruno bi-algebra. Let $\mathcal{I} = \{d_j\}_{j=1}^\infty$ be an infinite alphabet in 1–1 correspondence with \mathbb{N}^+ , and consider the free associative algebra $\mathcal{D} = \mathbb{R}\langle\mathcal{I}\rangle$ with the grading given by $|d_j| = j$ and $|d_{j_1} \cdots d_{j_k}| = j_1 + \cdots + j_k$. Let $\partial: \mathcal{D} \rightarrow \mathcal{D}$ be the derivation given by $\partial(d_i) = d_{i+1}$, linearity and the Leibniz rule $\partial(\omega_1\omega_2) = \partial(\omega_1)\omega_2 + \omega_1\partial(\omega_2)$ for all $\omega_1, \omega_2 \in \mathcal{I}^*$. We let $\#(\omega)$ denote the length of the word ω .

Definition 4.2. The non-commutative Bell polynomials $B_n \equiv B_n(d_1, \dots, d_n) \in \mathbb{R}\langle\mathcal{I}\rangle$ are defined by the recursion

$$\begin{aligned}B_0 &= \mathbb{I} \\ B_n &= (d_1 + \partial)B_{n-1} = (d_1 + \partial)^n \mathbb{I} \quad \text{for } n > 0.\end{aligned}$$

The first of these are given as

$$\begin{aligned}B_0 &= \mathbb{I} \\ B_1 &= d_1 \\ B_2 &= d_1^2 + d_2 \\ B_3 &= d_1^3 + 2d_1d_2 + d_2d_1 + d_3 \\ B_4 &= d_1^4 + 3d_1^2d_2 + 2d_1d_2d_1 + d_2d_1^2 + 3d_1d_3 + d_3d_1 + 3d_2d_2 + d_4.\end{aligned}$$

The polynomials B_n are introduced in [38, 39] to explain the Butcher order theory of Runge–Kutta methods in a manifold context, and generalize to certain classes of numerical integrators on manifolds.

Remark 4.3. Additional insight to the Bell polynomials are obtained by considering the free associative algebra generated by two symbols d_1 and ∂ , defining

$$d_i := [\partial, d_{i-1}] = \partial d_{i-1} - d_{i-1}\partial \quad \text{for } i > 1.$$

We find by induction that $(d_1 + \partial)^n$ satisfies the binomial relation

$$(4.4) \quad (d_1 + \partial)^n = \sum_{k=0}^n \binom{n}{k} B_k(d_1, \dots, d_k) \partial^{n-k},$$

which yields the formula

$$(4.5) \quad \exp(d_1 + \partial) = \sum_{m=0}^{\infty} \frac{B_m(d_1, \dots, d_m)}{m!} \exp(\partial),$$

and also the recursion

$$(4.6) \quad B_{n+1}(d_1, \dots, d_{n+1}) = \sum_{k=0}^n \binom{n}{k} B_k(d_1, \dots, d_k) d_{n-k+1} \quad \text{for } n > 0.$$

The non-commutative *partial Bell polynomials* $B_{n,k} \equiv B_{n,k}(d_1, \dots, d_{n-k+1})$ are defined as the part of B_n consisting of the words ω of length $\#(\omega) = k > 0$, e.g. $B_{4,3} = 3d_1^2 d_2 + 2d_1 d_2 d_1 + d_2 d_1^2$. Thus

$$B_n = \sum_{k=1}^n B_{n,k}.$$

A bit of combinatorics yields an explicit formula:

$$(4.7) \quad B_{n,k} = \sum_{\substack{\omega \in \mathcal{I}^* \\ |\omega|=n, \#(\omega)=k}} \kappa(\omega) \binom{n}{\omega} \omega,$$

where for $\omega = d_{j_1} d_{j_2} \cdots d_{j_k}$

$$\binom{n}{\omega} \equiv \binom{n}{|d_{j_1}|, |d_{j_2}|, \dots, |d_{j_k}|} := \frac{n!}{j_1! j_2! \cdots j_k!}$$

are the multinomial coefficients and the coefficients $\kappa(\omega)$ are defined as

$$(4.8) \quad \kappa(\omega) \equiv \kappa(|d_{j_1}|, |d_{j_2}|, \dots, |d_{j_k}|) := \frac{j_1 j_2 \cdots j_k}{j_1(j_1 + j_2) \cdots (j_1 + j_2 + \cdots + j_k)}.$$

The coefficients κ form a partition of unity on the symmetric group S_k ,

$$\sum_{\sigma \in S_k} \kappa(\sigma(\omega)) = 1,$$

where $\sigma(\omega)$ denotes a permutation of the letters in ω . E.g. $\kappa(1, 2) + \kappa(2, 1) = \frac{2}{3} + \frac{1}{3} = 1$.

It is often useful to employ polynomials Q_n and $Q_{n,k}$ related to B_n and $B_{n,k}$ by the following rescaling:

$$(4.9) \quad Q_{n,k}(d_1, \dots, d_{n-k+1}) = \frac{1}{n!} B_{n,k}(1!d_1, \dots, j!d_j, \dots) = \sum_{|\omega|=n, \#(\omega)=k} \kappa(\omega) \omega$$

$$(4.10) \quad Q_n(d_1, \dots, d_n) = \sum_{k=1}^n Q_{n,k}(d_1, \dots, d_{n-k+1})$$

$$(4.11) \quad Q_0 := \mathbb{I}.$$

Note that B_n and $B_{n,k}$ become the classical Bell- and partial Bell polynomials when the product in $\mathbb{R}\langle\mathcal{I}\rangle$ is commutative, i.e. in the free commutative algebra on \mathcal{I} . A non-commutative Faà di Bruno Hopf algebra is studied in [6]. However,

their definition differs from the present by defining the polynomials $Q_{n,k}$ without the factor κ that associates different factors to different permutations of a word (adding up to 1 over all permutations).

These Bell polynomials are closely related to the graded Dynkin operator on a connected graded Hopf algebra H . For $\alpha \in H^*$, define a graded algebra homomorphism $d_i \mapsto d_i(\alpha) : \mathcal{D} \rightarrow H^*$ as

$$(4.12) \quad d_i(\alpha) = \alpha_i = \alpha|_{H_i}, \quad d_i d_j(\alpha) = \alpha_i * \alpha_j.$$

Proposition 4.4. *The operator defined as*

$$(4.13) \quad Q(\alpha) = \sum_{n=0}^{\infty} Q_n(\alpha),$$

is a bijection from infinitesimal characters to characters $Q : \mathfrak{g}(H) \rightarrow G(H)$ with inverse given by right composition with the Dynkin idempotent $Y^{-1} \circ D$,

$$(4.14) \quad Q^{-1}(\beta) = \beta \circ Y^{-1} \circ D,$$

*where Y is the grading operator on H and $D = S * Y$ is the graded Dynkin operator.*

PROOF. For $\alpha \in \mathfrak{g}(H)$ we have

$$(4.15) \quad \Gamma(\alpha \circ Y) = \sum_{n=0}^{\infty} \sum_{j_1 + \dots + j_k = n} \frac{j_1 j_2 \cdots j_k}{j_1(j_1 + j_2) \cdots (j_1 + \dots + j_k)} \alpha_{j_1} * \dots * \alpha_{j_k} = Q(\alpha),$$

thus the result follows from Theorem 3.15. \square

The non-commutative Dynkin–Faà di Bruno bialgebra \mathcal{D} is obtained by taking the algebra structure of \mathcal{D} and defining the coproduct $\Delta_{\mathcal{D}}$ as

$$(4.16) \quad \Delta_{\mathcal{D}}(\mathbb{I}) = \mathbb{I} \otimes \mathbb{I}$$

$$(4.17) \quad \Delta_{\mathcal{D}}(d_n) = \sum_{k=1}^n B_{n,k} \otimes d_k.$$

This extends to all of \mathcal{D} by the product rule $\Delta_{\mathcal{D}}(d_i d_j) = \Delta_{\mathcal{D}}(d_i) \Delta_{\mathcal{D}}(d_j)$. Thus, e.g.

$$\begin{aligned} \Delta_{\mathcal{D}}(d_1) &= d_1 \otimes d_1 \\ \Delta_{\mathcal{D}}(d_2) &= d_1^2 \otimes d_2 + d_2 \otimes d_1 \\ \Delta_{\mathcal{D}}(d_1 d_2) &= d_1^3 \otimes d_1 d_2 + d_1 d_2 \otimes d_1^2. \end{aligned}$$

Note that the coproduct is *not* graded by $|\cdot|$, thus Proposition 3.7 does not hold for \mathcal{D} . By a lengthy (but not enlightening) induction argument we can prove:

Lemma 4.5. *The coproduct of the partial Bell polynomials are given as*

$$(4.18) \quad \Delta_{\mathcal{D}}(B_{n,k}) = \sum_{\ell=1}^n B_{n,\ell} \otimes B_{\ell,k}.$$

Note that $B_{n,1} = d_n$, thus (4.17) is a special case of (4.18). Summing the partial $B_{n,k}$ over k , we find the coproduct of the full Bell polynomials:

$$\Delta_{\mathcal{D}}(B_n) = \sum_{k=1}^n B_{n,k} \otimes B_k.$$

Using Lemma 4.5 and the fact that $B_{n,k} = 0$ for $k > n$, one can easily show that \mathcal{D} is a bialgebra.

Proposition 4.6. $\mathcal{D} = \mathbb{R}\langle\mathcal{I}\rangle$ with the non-commutative concatenation product and the coproduct $\Delta_{\mathcal{D}}$ form a bialgebra \mathcal{D} which is neither commutative nor cocommutative.

4.2.2. *Pullback along time-dependent flows.* Let $F_t = \sum_{j=0}^{\infty} F_{j+1} \frac{t^j}{j!}$ be a time-dependent vector field on \mathcal{M} where $F_j = F_t^{(j-1)} \Big|_{t=0}$. Let Φ_{t,F_t} be the solution operator of the corresponding non-autonomous equation, such that

$$y(t) = \Phi_{t,F_t} y_0 \quad \text{solves} \quad y'(t) = F_t(y(t)), \quad y(0) = y_0.$$

Note that Φ_{t,F_t} is not a 1-parameter subgroup of diffeomorphisms in t .

Lemma 4.7. [38] *The n -th time derivative of the pullback of a (time-independent) function ψ along the time-dependent flow Φ_{t,F_t} is given as*

$$(4.19) \quad \frac{\partial^n}{\partial t^n} \Phi_{t,F_t}^* \psi = B_n(F_t)[\psi],$$

where $B_n(F_t)$ is the image of B_n under the homomorphism from \mathcal{D} to $U(\mathcal{X}\mathcal{M})$ given by $d_i \mapsto F_t^{(i-1)}$. In particular

$$(4.20) \quad \frac{\partial^n}{\partial t^n} \Big|_{t=0} \Phi_{t,F_t}^* \psi = B_n(F_1, \dots, F_n)[\psi].$$

PROOF. The non-autonomous vector field F_t on \mathcal{M} corresponds to the autonomous field $F_t + \partial/\partial t$ on $\mathcal{M} \times \mathbb{R}$, thus (4.2) yields

$$\frac{\partial}{\partial t} \Phi_{t,F_t}^* \psi = \Phi_{t,F_t}^* ((F_t + \partial/\partial t)[\psi]) \Rightarrow \frac{\partial^n}{\partial t^n} \Phi_{t,F_t}^* \psi = \Phi_{t,F_t}^* ((F_t + \partial/\partial t)^n[\psi]).$$

Consider the homomorphism induced from $d_1 \mapsto F_t$ and $\partial \mapsto \partial/\partial t$, thus $d_i \mapsto F_t^{(i-1)}$. Equation (4.19) follows directly from Definition 4.2. At $t = 0$ we have $d_i \mapsto F_i$, thus (4.20). \square

Remark 4.8. Note that (4.5) yields a space-time split formula for pullback which is valid also for pullback of a time-dependent function ψ_t . The pullback for $t \in [0, h]$ developed at $t = 0$ becomes

$$\begin{aligned} \Phi_{h,F_t}^* \psi_t &= \exp \left(h(F_t + \frac{\partial}{\partial t}) \right) \Big|_{t=0} [\psi_t] = \sum_{n=0}^{\infty} \frac{h^n}{n!} B_n(F_t) \exp(h \frac{\partial}{\partial t}) \Big|_{t=0} [\psi_t] \\ &= \sum_{n=0}^{\infty} \frac{h^n}{n!} B_n(F_1, \dots, F_n)[\psi_h]. \end{aligned}$$

The Dynkin idempotent relates pullback series with their corresponding time-dependent vector fields. Let \mathcal{A} be an arbitrary alphabet with a grading $|\cdot|: \mathcal{A} \rightarrow \mathbb{N}^+$, let $\mathcal{H}_{\text{Sh}} = \mathcal{H}_{\text{Sh}}(\mathcal{A})$ be the corresponding graded shuffle algebra and let $\mathcal{B}_t(\alpha)$ be as in (4.3).

Proposition 4.9. *Let $\alpha \in \mathfrak{g}(\mathcal{H}_{\text{Sh}})$ and $\beta = Q(\alpha) \in G(\mathcal{H}_{\text{Sh}})$ be related by the graded Dynkin idempotent as in Proposition 4.4. Define the time-dependent vector field*

$$F_t = \frac{\partial}{\partial t} \mathcal{B}_t(\alpha).$$

Then pullback of a time-independent ψ along the time-dependent flow Φ_{t,F_t} is given as

$$(4.21) \quad \Phi_{t,F_t}^* \psi = \mathcal{B}_t(\beta)[\psi].$$

PROOF. We have $F_t = \sum_{j=0}^{\infty} F_{j+1} \frac{t^j}{j!}$ where $F_j = \mathcal{F}_\nu(j! \alpha_j)$. Developing the Taylor series of $\Phi_{t,F_t}^* \psi$ at $t = 0$ we get from (4.20)

$$\Phi_{t,F_t}^* \psi = \sum_{n=0}^{\infty} \frac{t^n}{n!} B_n(F_1, \dots, F_n)[\psi].$$

Thus

$$\frac{1}{n!} B_n(F_1, \dots, F_n) = \mathcal{F}\left(\frac{1}{n!} B_n(1! \alpha_1, \dots, n! \alpha_n)\right) = \mathcal{F}(Q_n(\alpha_1, \dots, \alpha_n)).$$

Using (4.15) we obtain the result. \square

4.3. Lie–Butcher theory. Pullback formulas such as (4.21) relate the time derivatives of F_t with the spatial derivatives of a function ψ . We have captured the algebraic structure of the temporal derivations through the Dynkin idempotent $Y^{-1} \circ D: G(\mathcal{H}_{\text{Sh}}) \rightarrow \mathfrak{g}(\mathcal{H}_{\text{Sh}})$ and its inverse $\Gamma \circ Y: \mathfrak{g}(\mathcal{H}_{\text{Sh}}) \rightarrow G(\mathcal{H}_{\text{Sh}})$. However, the spatial Lie derivation $\mathcal{B}_t(\beta)[\psi]$ cannot be algebraically characterized within this structure. In order to do this, we need to refine the Hopf algebra \mathcal{H}_{Sh} . On the manifold M , we obtain a refined version of $U(\mathcal{X}\mathcal{M})$ by expanding differential operators in terms of a non-commuting frame on $\mathcal{X}\mathcal{M}$. If the manifold is \mathbb{R}^n and the frame is the standard commutative coordinate frame, the construction yields the classical Butcher formulation and the Connes–Kreimer Hopf algebra [4]. More generally we obtain a Hopf algebra \mathcal{H}_N , built on forests of planar trees, which contains the Connes–Kreimer algebra as a subalgebra. In \mathcal{H}_N we can represent Lie derivation in terms of tree graftings.

4.3.1. Differential operators in $U(\mathcal{X}\mathcal{M})$ expanded in a non-commuting frame. Let $\mathcal{X}\mathcal{M}$ denote the Lie algebra of all vector fields on \mathcal{M} and let $\mathfrak{g} \subset \mathcal{X}\mathcal{M}$ be a transitive Lie subalgebra, in the sense that \mathfrak{g} everywhere spans $T\mathcal{M}$. This means that \mathfrak{g} defines a frame on the tangent bundle. We do not assume that the frame forms a basis. In general $\dim(\mathfrak{g}) \geq \dim(\mathcal{M})$, and in case of strict inequality we have a non-trivial isotropy subgroup at any point.

Let $U(\mathfrak{g})$ denote the universal enveloping algebra of \mathfrak{g} . We let $\mathfrak{g}^{\mathcal{M}}$ and $U(\mathfrak{g})^{\mathcal{M}}$ denote maps from \mathcal{M} to \mathfrak{g} and from \mathcal{M} to $U(\mathfrak{g})$. Since \mathfrak{g} is assumed to be transitive, we can represent any vector field $F \in \mathcal{X}\mathcal{M}$ with a function $f \in \mathfrak{g}^{\mathcal{M}}$ as in Section 2.2.1. Similarly, any higher order differential operator in $U(\mathcal{X}\mathcal{M})$ can be represented as a function in $U(\mathfrak{g})^{\mathcal{M}}$. We have the natural inclusion $\mathfrak{g} \subset \mathfrak{g}^{\mathcal{M}}$ and $U(\mathfrak{g}) \subset U(\mathfrak{g})^{\mathcal{M}}$ as constant maps, called frozen vector fields and higher order differential operators. We identify $U(\mathfrak{g})^{\mathcal{M}}$ with sections of the trivial vector bundle $\mathcal{M} \otimes U(\mathfrak{g}) \rightarrow \mathcal{M}$, and for a diffeomorphism $\Phi: \mathcal{M} \rightarrow \mathcal{M}$ we define pullback of $f \in U(\mathfrak{g})^{\mathcal{M}}$ as $\Phi^* f = f \circ \Phi \in U(\mathfrak{g})^{\mathcal{M}}$. Pullback in this bundle defines a parallel transport which gives rise to a flat connection with torsion. For $f, g \in U(\mathfrak{g})^{\mathcal{M}}$ we define the connection $f[g] \in U(\mathfrak{g})^{\mathcal{M}}$ pointwise from the Lie derivative as

$$f[g](p) = (f(p)[g])(p), \quad p \in \mathcal{M}.$$

Similarly, the concatenation in $U(\mathfrak{g})$ is extended pointwise to a concatenation product $fg \in U(\mathfrak{g})^{\mathcal{M}}$ as

$$(fg)(p) = f(p)g(p), \quad p \in \mathcal{M}.$$

This is called the *frozen composition* of f and g . We can also compose f and g as non-frozen differential operators $f \bullet g \in U(\mathfrak{g})^{\mathcal{M}}$:

$$(f \bullet g)[h] = f[g[h]], \quad \text{for all } h \in U(\mathfrak{g})^{\mathcal{M}}.$$

This is identical to the composition in $U(\mathcal{X}\mathcal{M})$, which in Section 4.1 was written as $F, G \mapsto FG$ for $F, G \in \mathcal{X}\mathcal{M}$.

It might be illustrative to write out the operations explicitly in terms of a basis $\{\partial_k\}_{k=1}^n$ of (non-commuting) vector fields spanning \mathfrak{g} . Writing $f, g \in \mathfrak{g}^{\mathcal{M}}$ in terms of the frame as $f = \sum_k f_k \partial_k$ and $g = \sum_\ell g_\ell \partial_\ell$ for $f_k, g_\ell \in \mathbb{R}^{\mathcal{M}}$, we have

$$\begin{aligned} fg &= \sum_{k, \ell} f_k g_\ell \partial_k \partial_\ell \\ f[g] &= \sum_{k, \ell} f_k \partial_k [g_\ell] \partial_\ell \\ f \bullet g &= \sum_{k, \ell} f_k \partial_k [g_\ell] \partial_\ell + \sum_{k, \ell} f_k g_\ell \partial_k \partial_\ell. \end{aligned}$$

The connection $f[g]$, the frozen composition fg and nonfrozen composition $f \bullet g$ are related as:

Lemma 4.10. *Let $f \in \mathfrak{g}^{\mathcal{M}}$ and $g, h \in U(\mathfrak{g})^{\mathcal{M}}$. Then we have*

$$\begin{aligned} \mathbb{I}[g] &= g \\ f[gh] &= f[g]h + g(f[h]), \quad (\text{Leibniz}) \\ (f \bullet g)[h] &:= f[g[h]] = (fg)[h] + (f[g])[h], \end{aligned}$$

where $\mathbb{I} \in U(\mathfrak{g})^{\mathcal{M}}$ is the constant identity map.

The proof is given in [42]. Note the difference between fg and $f \bullet g$. In the concatenation the value of g is frozen to $g(p)$ before the differentiation with f is done, whereas in the latter case the spatial variation of g is seen by the differentiation using f . Interestingly, the work of Cayley from 1857 [12] starts with the same result for vector fields expanded in the commuting frame $\partial/\partial x_i$.

From this lemma we may compute the torsion and curvature of the connection. Let $f, g \in \mathfrak{g}^{\mathcal{M}}$. We henceforth let $[f, g]_\bullet := f \bullet g - g \bullet f$ denote the Jacobi bracket and $[f, g] = fg - gf$ the *frozen bracket*. The frozen bracket is computed pointwise from the bracket in \mathfrak{g} as $[f, g](p) = [f(p), g(p)]_{\mathfrak{g}}$. Writing the connection as $\nabla_f g := f[g]$, we find

$$\begin{aligned} T(f, g) &= \nabla_f g - \nabla_g f - [f, g]_\bullet = gf - fg = -[f, g] \\ R(f, g)h &= \nabla_f \nabla_g h - \nabla_g \nabla_f h - \nabla_{[f, g]_\bullet} h = 0. \end{aligned}$$

Note that if \mathfrak{g} is commutative, then $[f, g] = 0$ and the connection is both flat and torsion free. In this case $f[g]$ is a pre-Lie product generating the Jacobi bracket: $f[g] - g[f] = [f, g]_\bullet$, but in general $f[g] - g[f] = [f, g]_\bullet - [f, g]$.

The product $f \bullet g$ is associative, and thus $U(\mathfrak{g})^{\mathcal{M}}$ with the binary operations $f, g \mapsto f[g]$ and $f, g \mapsto f \bullet g$ forms a unital dipterous algebra [33], however, it has more structure than this. Following [42] we define:

Definition 4.11. Let $\mathcal{A} = \mathbb{I} \oplus \overline{\mathcal{A}}$ be a unital associative algebra with product $f, g \mapsto fg$, and also equipped with a non-associative composition $f, g \mapsto f[g]: \mathcal{A} \times \mathcal{A} \rightarrow \mathcal{A}$. Let $D(\mathcal{A})$ denote all $f \in \mathcal{A}$ such that $f[\cdot]$ is a derivation:

$$D(\mathcal{A}) = \{ f \in \mathcal{A} \mid f[gh] = (f[g])h + g(f[h]) \}.$$

We assume that $D(\mathcal{A})$ generates $\overline{\mathcal{A}}$. We call \mathcal{A} a D-algebra if for any derivation $f \in D(\mathcal{A})$ and any $g, h \in \mathcal{A}$ we have

$$(4.22) \quad g[f] \in D(\mathcal{A})$$

$$(4.23) \quad \mathbb{I}[g] = g$$

$$(4.24) \quad f[g[h]] = (fg)[h] + (f[g])[h].$$

Definition 4.12. A D-algebra homomorphism is a map $\mathcal{F}: \mathcal{A} \rightarrow \mathcal{A}'$ between D-algebras such that $\mathcal{F}(D(\mathcal{A})) \subset D(\mathcal{A}')$ and for all $g, h \in \mathcal{A}$ we have

$$(4.25) \quad \mathcal{F}(\mathbb{I}) = \mathbb{I}$$

$$(4.26) \quad \mathcal{F}(gh) = \mathcal{F}(g)\mathcal{F}(h)$$

$$(4.27) \quad \mathcal{F}(g([h])) = \mathcal{F}(g)[\mathcal{F}(h)].$$

4.3.2. *The free D-algebra and elementary differentials.* The following definitions are detailed in [42]. Let OT denote the alphabet of all ordered (planar) rooted trees:

$$\text{OT} = \{\bullet, \bullet\bullet, \bullet\bullet\bullet, \bullet\bullet\bullet\bullet, \bullet\bullet\bullet\bullet\bullet, \bullet\bullet\bullet\bullet\bullet\bullet, \bullet\bullet\bullet\bullet\bullet\bullet\bullet, \dots\}.$$

More generally, we consider decorated ordered rooted trees, where \mathcal{C} is a (finite or infinite) set of colors. Decorated trees are trees with a color from \mathcal{C} assigned to each node. As above, we let OT^* denote words of trees (forests), let \mathbb{I} be the empty word and let $\omega_1, \omega_2 \mapsto \omega_1\omega_2$ denote concatenation for $\omega_1, \omega_2 \in \text{OT}^*$. Identifying $\mathcal{C} \subset \text{OT}$ with 1-node trees, we can recursively build all words in OT^* from \mathcal{C} by concatenation and adding roots. For $c \in \mathcal{C}$ and $\omega \in \text{OT}^*$, define $B_c^+(\omega) \in \text{OT}$ as the tree with branches ω and root c . Often we will be interested in the case where $\mathcal{C} = \{\bullet\}$, just one color.

As above, let $\mathbb{R}\langle\text{OT}\rangle$ denote real polynomials (finite \mathbb{R} -linear combinations of words) and $\mathbb{R}\langle\langle\text{OT}\rangle\rangle$ the dual space of infinite series, such as

$$\begin{aligned} \alpha = \alpha(\mathbb{I})\mathbb{I} + \alpha(\bullet)\bullet + \alpha(\bullet\bullet)\bullet\bullet + \alpha(\bullet\bullet\bullet)\bullet\bullet\bullet + \alpha(\bullet\bullet\bullet\bullet)\bullet\bullet\bullet\bullet \\ + \alpha(\bullet\bullet\bullet\bullet\bullet)\bullet\bullet\bullet\bullet\bullet + \alpha(\bullet\bullet\bullet\bullet\bullet\bullet)\bullet\bullet\bullet\bullet\bullet + \alpha(\bullet\bullet\bullet\bullet\bullet\bullet\bullet)\bullet\bullet\bullet\bullet\bullet + \alpha(\bullet\bullet\bullet\bullet\bullet\bullet\bullet\bullet)\bullet\bullet\bullet\bullet\bullet + \dots \end{aligned}$$

On $\mathbb{R}\langle\text{OT}\rangle$ we define *left grafting* $(\cdot)[\cdot]: \mathbb{R}\langle\text{OT}\rangle \times \mathbb{R}\langle\text{OT}\rangle \rightarrow \mathbb{R}\langle\text{OT}\rangle$ by extending the following definition for trees by linearity. For all $c \in \mathcal{C}$, all $\tau \in \text{OT}$ and all $\omega, \omega' \in \text{OT}^*$ we define:

$$\begin{aligned} \omega[c] &= B_c^+(\omega) \\ \mathbb{I}[\omega] &= \omega \\ \tau[\omega\omega'] &= \tau[\omega]\omega' + \omega(\tau[\omega']) \\ \tau[\omega[\omega']] &= (\tau\omega)[\omega'] + (\tau[\omega])[\omega']. \end{aligned}$$

Compare this with Lemma 4.10. The left grafting $\tau[\omega]$ is obtained by attaching τ in all possible ways from the left to the vertices of ω , and $(\tau\tau')[\omega]$ is obtained by

attaching from the left first τ' and then τ on all nodes of ω :

$$\begin{aligned} \mathbb{O} \left[\begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} \right] &= \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} \\ \mathbb{O} \left[\begin{smallmatrix} \bullet & \bullet \\ \bullet & \bullet \end{smallmatrix} \right] &= \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix} + \begin{smallmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{smallmatrix}. \end{aligned}$$

We henceforth let $|\omega|$ denote the grading counting the total number of nodes, i.e. $|c| = 1$ for all $c \in \mathcal{C}$, $|\omega\omega'| = |\omega| + |\omega'|$ and $|\omega[\omega']| = |\omega| + |\omega'|$.

Proposition 4.13. *Let OT be planar trees decorated with colors \mathcal{C} . Consider $\mathcal{N} = \mathbb{R}\langle\text{OT}\rangle$ with concatenation $\omega, \omega' \mapsto \omega\omega'$, left grafting $\omega, \omega' \mapsto \omega[\omega']$ and unit \mathbb{I} as defined above. \mathcal{N} is a free D -algebra over \mathcal{C} , such that for any D -algebra \mathcal{A} and any map $\nu: \mathcal{C} \rightarrow D(\mathcal{A})$ there exists a unique D -algebra homomorphism map $\mathcal{F}_\nu: \mathcal{N} \rightarrow \mathcal{A}$ such that $\mathcal{F}_\nu(c) = \nu(c)$ for all $c \in \mathcal{C}$.*

$$\begin{array}{ccc} \mathcal{C} & \hookrightarrow & \mathcal{N} \\ \nu \downarrow & & \downarrow \exists! \mathcal{F}_\nu \\ D(\mathcal{A}) & \hookrightarrow & \mathcal{A} \end{array}$$

Definition 4.14. We define the *ordered Grossman–Larson*⁴ product on \mathcal{N} for all $\omega, \omega' \in \text{OT}^*$ as

$$\omega \bullet \omega' = B^-(\omega[B^+(\omega')]).$$

I.e. we add a root to ω' , graft on ω and finally remove the root again.

Proposition 4.15. *The GL-product is associative and, for all $n, n', n'' \in \mathcal{N}$, satisfies*

$$(4.28) \quad n[n'[n'']] = (n \bullet n')[n'']$$

$$(4.29) \quad \mathcal{F}_\nu(n \bullet n') = \mathcal{F}_\nu(n) \bullet \mathcal{F}_\nu(n').$$

Remark 4.16. The classical setting of Cayley, Merson and Butcher is the case where $\mathcal{M} = \mathbb{R}^n$ and $\mathfrak{g} = \{\partial/\partial x_i\} \subset \mathcal{X}\mathcal{M}$ is the standard commutative coordinate frame. The construction of Section 4.3.1 produces $U(\mathfrak{g})^{\mathcal{M}}$ as a D -algebra where the concatenation is commutative. The connection is now flat and torsionless, and $f[g]$ becomes a pre-Lie product. The images of the trees $\mathcal{F}(\tau)$, for $\tau \in \text{OT}$, are called the *elementary differentials* in Butcher’s theory (see [8]). These are explicitly given in (2.3). The images of the forests $\mathcal{F}(\omega)$, for $\omega \in \text{OT}^*$, are called *elementary differential operators* in Merson’s theory (see [36]).

4.3.3. A generalized Connes–Kreimer Hopf algebra of planar trees. We recall from [42] the definition of the Hopf algebra \mathcal{H}_N . On the vector space $\mathbb{R}\langle\text{OT}\rangle$ we define the shuffle product \sqcup , and we define the coproduct Δ_N as the dual of the ordered GL product, such that

$$(4.30) \quad (\alpha \bullet \beta)(\omega) = \sum_{(\omega)_{\Delta_N}} \alpha(\omega_{(1)})\beta(\omega_{(2)}) \quad \text{for all } \alpha, \beta \in \mathbb{R}\langle\langle\text{OT}\rangle\rangle.$$

The motivation for this construction is the representation of $U(\mathcal{X}\mathcal{M})$ in terms of a frame $\mathfrak{g} \subset \mathcal{X}\mathcal{M}$ as $U(\mathfrak{g})^{\mathcal{M}}$. The shuffle product is the correct product to characterize which series in $\mathbb{R}\langle\langle\text{OT}\rangle\rangle$ represent vector fields on \mathcal{M} and which represent

⁴The GL product is usually defined in a similar way over non-planar trees.

diffeomorphisms. The composition in $U(\mathcal{X}\mathcal{M})$ appears as the product \bullet on $U(\mathfrak{g})^{\mathcal{M}}$, thus with the coproduct Δ_N the convolution on $\mathbb{R}\langle\langle\text{OT}\rangle\rangle$ represents composition in $U(\mathcal{X}\mathcal{M})$.

It remains to give a precise characterization of Δ_N and the antipode in \mathcal{H}_N . As in the Connes–Kreimer case, both Δ_N and the antipode can be defined directly in terms of admissible cuts or in a recursive fashion. Recursively Δ_N is given as

$$(4.31) \quad \begin{aligned} \Delta_N(\mathbb{I}) &= \mathbb{I} \otimes \mathbb{I}, \\ \Delta_N(\omega\tau) &= \omega\tau \otimes \mathbb{I} + \Delta_N(\omega) \sqcup \cdot (I \otimes B_c^+) \Delta_N(\omega_1), \end{aligned}$$

where $\tau = B_c^+(\omega_1) \in \text{OT}$, where $\omega, \omega_1 \in \text{OT}^*$ and where $\sqcup \cdot$ denotes shuffle on the left and concatenation on the right: $(\omega_1 \otimes \tau_1) \sqcup \cdot (\omega_2 \otimes \tau_2) = (\omega_1 \sqcup \omega_2) \otimes (\tau_1 \tau_2)$. The direct formula is

$$(4.32) \quad \Delta_N(\omega) = \sum_{\ell \in \text{FALC}(\omega)} P^\ell(\omega) \otimes R^\ell(\omega),$$

where FALC denotes *Full Admissible Left Cuts*, $P^\ell(\omega)$ is the shuffle of all the cut off parts, and $R^\ell(\omega)$ is the remaining part containing the root (see [42]). Calculations of the coproduct for forests up to order 4 can be found in Table 1.

Theorem 4.17. *Let \mathcal{H}_N be the vector space $\mathcal{N} = \mathbb{R}\langle\langle\text{OT}\rangle\rangle$ with the operations*

$$\begin{aligned} \text{product: } \mu_N(a \otimes b) &= a \sqcup b, \\ \text{coproduct: } \Delta_N &, \\ \text{unit: } u_N(1) &= \mathbb{I}, \\ \text{counit: } e_N(\omega) &= \begin{cases} 1, & \text{if } \omega = \mathbb{I}, \\ 0, & \text{else.} \end{cases} \end{aligned}$$

Then \mathcal{H}_N is a Hopf algebra with an antipode S_N given by the recursion

$$(4.33) \quad \begin{aligned} S_N(\mathbb{I}) &= \mathbb{I}, \\ S_N(\omega\tau) &= -\mu_N((S_N \otimes I)(\Delta_N(\omega) \sqcup \cdot (I \otimes B_i^+) \Delta_N(\omega_1))), \end{aligned}$$

where $\tau = B_i^+(\omega_1) \in \text{OT}$ and $\omega, \omega_1 \in \text{OT}^$.*

4.3.4. Lie–Butcher series and flows on manifolds. The set of maps $U(\mathfrak{g})^{\mathcal{M}}$ from \mathcal{M} to $U(\mathfrak{g})$ is a D-algebra where the derivations are the vector fields $\mathfrak{g}^{\mathcal{M}}$. Thus, given a set of colors \mathcal{C} and a map $\nu: \mathcal{C} \rightarrow \mathfrak{g}^{\mathcal{M}}$ there exists a unique map $\mathcal{F}_\nu: \mathcal{N} \rightarrow U(\mathfrak{g})^{\mathcal{M}}$ such that for all $c \in \mathcal{C}$ and all $g, h \in \mathcal{N}$ we have

$$(4.34) \quad \mathcal{F}_\nu(c) = \nu(c)$$

$$(4.35) \quad \mathcal{F}_\nu(\mathbb{I}) = \mathbb{I}$$

$$(4.36) \quad \mathcal{F}_\nu(gh) = \mathcal{F}_\nu(g)\mathcal{F}_\nu(h)$$

$$(4.37) \quad \mathcal{F}_\nu(g[h]) = \mathcal{F}_\nu(g)[\mathcal{F}_\nu(h)]$$

$$(4.38) \quad \mathcal{F}_\nu(g \bullet h) = \mathcal{F}_\nu(g) \bullet \mathcal{F}_\nu(h).$$

$$(4.39)$$

Definition 4.18. For an infinite series $\alpha \in \mathcal{N}^* = \mathbb{R}\langle\langle\text{OT}\rangle\rangle$ a Lie–Butcher series is a formal series in $U(\mathfrak{g})^{\mathcal{M}}$ defined as

$$\mathcal{B}_t(\alpha) = \sum_{\omega \in \text{OT}^*} t^{|\omega|} \alpha(\omega) \mathcal{F}_\nu(\omega).$$

Note that \mathcal{N} can be turned into a Hopf algebra two different ways: either as \mathcal{H}_{Sh} with product \sqcup and deconcatenation coproduct Δ_d , or as \mathcal{H}_N with the same product \sqcup , but where the coproduct Δ_N is the dual of the ordered GL product. This gives rise to two different convolutions on \mathcal{N}^* , the frozen composition $\alpha, \beta \mapsto \alpha\beta$ in Example 3.9, and the non-frozen composition $\alpha, \beta \mapsto \alpha \bullet \beta$ as in (4.30). Since the product is the same, we have that the characters and the infinitesimal characters are the same as vector spaces

$$\begin{aligned}\mathfrak{g}(\mathcal{H}_{\text{Sh}}) &= \mathfrak{g}(\mathcal{H}_N) = \{ \alpha \in \mathcal{N} \mid \alpha(\mathbb{I}) = 0, \alpha(\omega \sqcup \omega') = 0 \ \forall \omega, \omega' \in \text{OT}^* \setminus \mathbb{I} \} \\ G(\mathcal{H}_{\text{Sh}}) &= G(\mathcal{H}_N) = \{ \alpha \in \mathcal{N} \mid \alpha(\mathbb{I}) = 1, \alpha(\omega \sqcup \omega') = \alpha(\omega)\alpha(\omega') \ \forall \omega, \omega' \in \text{OT}^* \}.\end{aligned}$$

However, the exponential, logarithm, Dynkin and Eulerian idempotents, as well as the antipode depend on whether they are based on \mathcal{H}_{Sh} or \mathcal{H}_N . Which to use in practice depends on which operation we want to express on the manifold. Recall that *frozen elements* of $U(\mathfrak{g})^{\mathcal{M}}$ are constant functions $g: \mathcal{M} \rightarrow U(\mathfrak{g})$. If g is frozen then $f[g] = 0$ for all f , and hence $f \bullet g = fg$. The subalgebra of frozen vector fields therefore reduces to \mathcal{H}_{Sh} .

We summarize the basic properties of LB-series: \mathcal{B}_t sends infinitesimal characters to (formal) vector fields on \mathcal{M} and characters to pullback series representing formal diffeomorphisms on \mathcal{M} . LB-series preserve both frozen and non-frozen composition and sends left grafting to the connection on $U(\mathfrak{g})^{\mathcal{M}}$.

$$\begin{aligned}\mathcal{B}_t(\alpha\beta) &= \mathcal{B}_t(\alpha)\mathcal{B}_t(\beta) \\ \mathcal{B}_t(\alpha \bullet \beta) &= \mathcal{B}_t(\alpha) \bullet \mathcal{B}_t(\beta) \\ \mathcal{B}_t(\alpha[\beta]) &= \mathcal{B}_t(\alpha)[\mathcal{B}_t(\beta)].\end{aligned}$$

Note that if $\alpha \in G(\mathcal{H}_N)$, then $\alpha[\beta]$ represents algebraically the pullback (parallel transport) of β along the flow of α . On the manifold

$$\mathcal{B}_h(\alpha[\beta])(y_0) = \mathcal{B}_h(\alpha)[\mathcal{B}_h(\beta)](y_0) = \mathcal{B}_h(\beta)(\Phi(y_0)),$$

where Φ is the diffeomorphism represented by $\alpha \in G(\mathcal{H}_N)$ at $t = h$. Since the connection is flat, the pullback depends only on the endpoint $\Phi(y_0)$ and not on the actual path.

There are (at least) three ways to represent a flow $y_0 \mapsto y_t = \Phi_t(y_0)$ on \mathcal{M} , using LB-series:

- (1) In terms of pullback series. Find $\alpha \in G(\mathcal{H}_N)$ such that

$$(4.40) \quad \psi(y(t)) = \mathcal{B}_t(\alpha)[\psi](y_0) \quad \text{for any } \psi \in U(\mathfrak{g})^{\mathcal{M}}.$$

This representation is used in the analysis of Crouch–Grossman methods by Owren and Marthinsen [45]. In the classical setting, this is called a *S*-series [43].

- (2) In terms of an autonomous differential equation. Find $\beta \in \mathfrak{g}(\mathcal{H}_N)$ such that $y(t)$ solves

$$(4.41) \quad y'(t) = \mathcal{B}_h(\beta)(y(t)) \cdot y(t).$$

In the classical setting, this is called backward error analysis. In the Lie group setting, this formulation has, however, never been investigated in detail (but it should!).

- (3) In terms of a non-autonomous equation of *Lie type* (time dependent frozen vector field). Find $\gamma \in \mathfrak{g}(\mathcal{H}_{\text{Sh}})$ such that $y(t)$ solves

$$(4.42) \quad y'(t) = \frac{\partial}{\partial t} \mathcal{B}_t(\gamma)(y_0) \cdot y(t).$$

This representation is used in [38, 39]. In the classical setting this is (almost) the standard definition of B -series. The connection with the classical B -series is discussed below.

The algebraic relationship between α , β and γ is given as follows:

| | |
|--|--|
| $\beta = \alpha \circ e$ | e is Euler idempotent in \mathcal{H}_N . |
| $\alpha = \exp^\bullet(\beta)$ | Exponential wrt. GL-product |
| $\gamma = \alpha \circ Y^{-1} \circ D$ | Dynkin idempotent in $\mathcal{H}_{\text{Sh}}(\text{OT})$. |
| $\alpha = Q(\gamma)$ | Q -operator (4.13) in $\mathcal{H}_{\text{Sh}}(\text{OT})$. |

Example 4.19. Two examples are of particular interest; the exact solution and exponential Euler method. In both cases we consider $y'(t) = f(y) \cdot y$, where $\mathcal{C} = \{\bullet\}$ and $\nu(\bullet) = f$.

The exponential Euler method is particularly simple. Since each step of the method follows the flow the frozen vector field $f(y_n) \in \mathfrak{g}$, the Type 3 LB-series for Exponential Euler must be given by

$\gamma_{\text{Euler}} =$ •

just as in the classical setting⁵.

Type 3 LB-series for the exact solution can be derived in various ways. Theorem 2.2 in [39] derives the exact solution as the solution of

$$y' = f_t \cdot y, \quad y(0) = y_0.$$

where $f_t = f(y(t)) \in \mathfrak{g}$ is the pullback of f along the time dependent flow of f_t .

Letting

$f_t = \frac{\partial}{\partial t} \mathcal{B}_t(\gamma)$ we obtain

$$Y \circ \gamma = Q(\gamma)[\bullet] \Rightarrow \gamma = Y^{-1} \circ B^+(Q(\gamma)).$$

Note that this is reminiscent of a so-called combinatorial Dyson–Schwinger equation [24]. Solving by iteration yields

Remarkably, the LB-series of the exact solution is just a combination of trees, and not commutators of trees. Thus in Type 3 LB-series developments of numerical integrators, commutators of trees must be zero up to the order of the method.

⁵The classical presentation is $\gamma = \mathbb{I} + \bullet$, when the B-series is given in the form (2.2).

Composition and inverse is simplest for pullback series, Type 1. For series of Type 3, we map to Type 1, compose (or invert) and map back again. If $\gamma, \tilde{\gamma}$ are series of Type 3, then the basic operations are done as:

$$(4.43) \quad \text{Composition: } \gamma, \tilde{\gamma} \mapsto (Q(\gamma) \bullet Q(\tilde{\gamma})) \circ Y^{-1} \circ D$$

$$(4.44) \quad \text{Inverse: } \gamma^{-1} = Q(\gamma) \circ S \circ Y^{-1} \circ D$$

$$(4.45) \quad \text{Backward error: } \text{Log}_3(\gamma) := Q(\gamma) \circ e.$$

4.3.5. *Relations to classical B-series.* The relation between classical B-series and LB-series is detailed in [42]. Classical B-series are expressed in terms of linear combinations of non-planar trees T , resulting in the Connes–Kreimer Hopf algebra \mathcal{H}_C built from non-planar trees [4]. In the classical setting the connection is torsion-free, and concatenation is commutative. Therefore $\mathfrak{g}(\mathcal{H}_C) = \text{span}(T)$. That is, $\mathfrak{g}(\mathcal{H}_C)$ is just linear combinations of trees. This fact is the reason why many discussions in the classical setting can avoid series involving forests of trees (words in T^*). Also the difference between series of Type 1 and Type 3 is not emphasized in many papers. Since the coefficients κ of the Q -polynomials add up to one under symmetrization, we find in the classical setting that

$$Q(\alpha)(\omega) = \alpha(\tau_1)\alpha(\tau_2)\cdots\alpha(\tau_k)\omega,$$

for $\omega = B^+(\tau_1\tau_2\cdots\tau_k)$, so formulas involving pullbacks are often expressed directly from B-series (Type 3) using the Q -polynomials in this form. Our claim that classical B-series fits best into series of Type 3 is based on the trivial observation that the curve $y_t = \mathcal{B}_t(\alpha)(y)$ in (4.42) solves a differential equation with a time dependent frozen vector field given as

$$y(t) = \frac{\partial}{\partial t} \sum_{\tau \in T} \frac{t^{|\tau|}}{\sigma(\tau)} \mathcal{F}(\tau).$$

One can ask why the symmetrization $\sigma(\tau)$ is natural to include in the classical setting, but not in the LB-series setting. To explain the relationship between the two theories we define a symmetrization operator:

Definition 4.20. The symmetrization operator $\Omega : \mathcal{N} \rightarrow \mathcal{N}$ is defined for $\omega \in \text{OT}^*$ and $\tau \in \text{OT}$ as

$$\begin{aligned} \Omega(\mathbb{I}) &= \mathbb{I}, \\ \Omega(\omega\tau) &= \Omega(\omega) \sqcup \Omega(\tau), \\ \Omega(B_i^+(\omega)) &= B_i^+(\Omega(\omega)). \end{aligned}$$

The shuffle product permutes the trees in a forest in all possible ways, and the symmetrization of a tree is a recursive splitting in sums over all permutations of the branches. The symmetrization defines an equivalence relation on OT^* , that is

$$\Omega(\omega_1) = \Omega(\omega_2) \iff \omega_1 \sim \omega_2.$$

Let $\iota : \mathcal{H}_C \rightarrow \mathcal{H}_N$ be an inclusion where a tree is identified with one of its equivalent planar trees. In [42] we show that $\tilde{\Omega} = \Omega \circ \iota : \mathcal{H}_C \rightarrow \mathcal{H}_N$ is a Hopf algebra isomorphism onto its image, i.e. \mathcal{H}_C is a proper subalgebra of \mathcal{H}_N . The adjoint map $\tilde{\Omega}^* : \mathcal{H}_N^* \rightarrow \mathcal{H}_C^*$ is given as

$$\tilde{\Omega}^*(\alpha)(\omega) = \sigma(\omega) \sum_{\omega' \sim \omega} \alpha(\omega').$$

The tree symmetrization $\sigma(\omega)$ enters exactly such that the LB-series as given in (4.18) maps to the classical B-series in (2.2).

4.4. Substitution law for LB-series. The so-called *substitution law* for B-series [14] can without much difficulty be generalized to LB series. Consider \mathcal{N} as a D-algebra where the derivations are the Lie polynomials $D(\mathcal{N}) = \mathfrak{g}(\mathcal{H}_N) \cap \mathcal{N}$. By the universality property of \mathcal{N} , we know that for any map $a: \mathcal{C} \rightarrow D(\mathcal{N})$ there exists a unique D-algebra homomorphism $\mathcal{F}_a: \mathcal{N} \rightarrow \mathcal{N}$ such that $\mathcal{F}_a(c) = a(c)$ for all $a \in \mathcal{C}$. This is called the substitution law.

Definition 4.21. For any map $a: \mathcal{C} \rightarrow D(\mathcal{N})$ there exists a unique D-algebra homomorphism $a\star: \mathcal{N} \rightarrow \mathcal{N}$ such that $a(c) = a\star c$ for all $c \in \mathcal{C}$. The map $a\star$ is called a -substitution⁶.

$$\begin{array}{ccc} \mathcal{C} & \hookrightarrow & \mathcal{N} \\ a \downarrow & & \downarrow a\star \\ D(\mathcal{N}) & \hookrightarrow & \mathcal{N}. \end{array}$$

The properties of this substitution law, together with applications of it, will be studied in a forthcoming paper ([34]). We just mention that many of the useful properties of the substitution law follow immediately from the fact that $a\star: \mathcal{N} \rightarrow \mathcal{N}$ is a homomorphism. For example, for all $n, n' \in \mathcal{N}$ we have:

$$\begin{aligned} a\star \mathbb{I} &= \mathbb{I} \\ a\star (nn') &= (a\star n)(a\star n') \\ a\star (n[n']) &= (a\star n)[a\star n'] \\ a\star (n \bullet n') &= (a\star n) \bullet (a\star n') \end{aligned}$$

5. Final remarks and outlook

Inspired by problems in numerical analysis we have discussed various algebraic structures arising in the study of formal diffeomorphisms on manifolds. We have seen that the Connes–Kreimer Hopf algebra naturally extends from commutative frames on \mathbb{R}^n to non-commutative frames on general manifolds. In particular we have presented the Dynkin and Euler operators and non-commutative Faà di Bruno type bialgebras in this generalized setting.

The formalism in this paper has many applications in numerical analysis, and analysis of Lie group integrators in particular. However, the underlying structures are general constructions with possible applications in other fields, such as geometric control theory and sub-Riemannian geometry. Connections to stochastic differential equations on manifolds is an other topic which is worth investigating further.

Acknowledgements

The authors would like to thank Alessandra Frabetti, Dominique Manchon, Gilles Vilmart and Will Wright for interesting discussions on topics of this paper. In particular we would like to thank Kurusch Ebrahimi-Fard for his support and

⁶In most applications we want to substitute infinite series and extend $a\star$ to a homomorphism $a\star: \mathcal{N}^* \rightarrow \mathcal{N}^*$. The extension to infinite substitution is straightforward because of the grading, we omit details.

useful remarks in the writing process. His enthusiasm and inclusive spirit have been of crucial importance for the completion of this paper.

References

- [1] E. Abe. *Hopf Algebras*. Cambridge University Press, 1980.
- [2] R. Abraham, J. E. Marsden, and T. Ratiu. *Manifolds, Tensor Analysis, and Applications*. AMS 75. Springer-Verlag, Second edition, 1988.
- [3] H. Berland and B. Owren. Algebraic structures on ordered rooted trees and their significance to Lie group integrators. *Group theory and numerical analysis*, 39:49–63, 2005.
- [4] C. Brouder. Runge-Kutta methods and renormalization. *The European Physical Journal C-Particles and Fields*, 12(3):521–534, 2000.
- [5] C. Brouder. Trees, renormalization and differential equations. *BIT*, 44(3):425–438, 2004.
- [6] C. Brouder, A. Frabetti, and C. Krattenthaler. Non-commutative Hopf algebra of formal diffeomorphisms. *Advances in Mathematics*, 200(2):479–524, 2006.
- [7] E. Burgunder. Eulerian idempotent and Kashiwara-Vergne conjecture. 58(4):1153–1184, 2008.
- [8] J. C. Butcher. Coefficients for the study of Runge-Kutta integration processes. *J. Austral. Math. Soc.*, 3:185–201, 1963.
- [9] J. C. Butcher. An algebraic theory of integration methods. *Math. Comp.*, 26:79–106, 1972.
- [10] D. Calaque, K. Ebrahimi-Fard, and D. Manchon. Two interacting Hopf algebras of trees. *To appear in Adv. Appl. Math.*, 2009, math.CO/0806.2238v3.
- [11] P. Cartier. A primer of Hopf algebras. In *Frontiers in number theory, physics, and geometry*, volume II, pages 537–615. Springer, Berlin, 2007.
- [12] A. Cayley. On the theory of the analytical forms called trees. *Philos. Mag.*, 13(19):4–9, 1857.
- [13] E. Celledoni, A. Marthinsen, and B. Owren. Commutator-free Lie group methods. *Future Generation Computer Systems*, 19(3):341–352, 2003.
- [14] P. Chartier, E. Hairer, and G. Vilmart. A substitution law for B-series vector fields. *INRIA report*, (5498), 2005.
- [15] P. Chartier, E. Hairer, and G. Vilmart. Numerical integrators based on modified differential equations. *Mathematics of Computation*, 76(260):1941, 2007.
- [16] P. Chartier and A. Murua. An algebraic theory of order. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43(4):607–630, 2009.
- [17] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Communications in Mathematical Physics*, 199(1):203–242, 1998.
- [18] P. E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *J. Nonlinear Sci.*, 3:1–33, 1993.
- [19] A. Dür. *Möbius functions, incidence algebras and power series representations*, volume 1202 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986.
- [20] K. Ebrahimi-Fard, J.M. Gracia-Bondía, and F. Patras. A Lie Theoretic Approach to Renormalization. *Communications in Mathematical Physics*, 276(2):519–549, 2007.
- [21] K. Ebrahimi-Fard and D. Manchon. A Magnus-and Fer-type formula in dendriform algebras. *Foundations of Computational Mathematics*, 9:1–22, 2009, math.CO/07070607v3.
- [22] H. Figueroa and J.M Gracia-Bondía. Combinatorial Hopf algebras in quantum field theory I. *Rev.Math.Phys.*, 17:881, 2005, hep-th/0408145v3.
- [23] H. Figueroa, J.M. Gracia-Bondía, and J.C. Varilly. Faa di Bruno Hopf algebras. *Preprint*, 2005, math.CO/0508337.
- [24] L. Foissy. Faà di Bruno subalgebras of the Hopf algebra of planar trees from combinatorial Dyson–Schwinger equations. *Advances in Mathematics*, 218(1):136–162, 2008, 0707.1204v2.
- [25] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration*. Springer-Verlag, second edition, 2006.
- [26] E. Hairer and G. Wanner. On the Butcher group and general multi-value methods. *Computing (Arch. Elektron. Rechnen)*, 13(1):1–15, 1974.
- [27] A. Iserles, A. Marthinsen, and S.P. Nørsett. On the implementation of the method of Magnus series for linear differential equations. *BIT Numerical Mathematics*, 39(2):281–304, 1999.
- [28] A. Iserles, H.Z. Munthe-Kaas, S.P. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica*, 9:215–365, 2000.

- [29] A. Iserles and S.P. Nørsett. On the solution of linear differential equations in Lie groups. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 357(1754):983–1019, 1999.
- [30] C. Kassel. *Quantum groups*. Springer-Verlag, 1995.
- [31] R. Lenczewski. A noncommutative limit theorem for homogeneous correlations. *Studia Mathematica*, 129(3), 1998.
- [32] J.L. Loday. *Cyclic Homology*. Springer-Verlag, second edition, 1997.
- [33] J.L. Loday and M.O. Ronco. Combinatorial Hopf algebras. *Clay Mathematics Proceedings*, 12:347–384, 2010, math.CO/0508337.
- [34] A. Lundervold and H.Z. Munthe-Kaas. Backward error analysis and the substitution law for Lie group integrators. *Preprint*, 2010.
- [35] D. Manchon. Hopf algebras, from basics to applications to renormalization. *Preprint*, 2006, math.QA/0408405v2.
- [36] R. H. Merson. An operational method for the study of integration processes. In *Proc. Conf., Data Processing & Automatic Computing Machines*, pages 110–1–11025, 1957.
- [37] S. Monaco, D. Normand-Cyrot, and C. Califano. From chronological calculus to exponential representations of continuous and discrete-time dynamics: a Lie-algebraic approach. *IEEE Transactions on Automatic Control*, 52(12):2227–2241, 2007.
- [38] H. Munthe-Kaas. Lie–Butcher theory for Runge–Kutta methods. *BIT*, 35(4):572–587, 1995.
- [39] H. Munthe-Kaas. Runge–Kutta methods on Lie groups. *BIT*, 38(1):92–111, 1998.
- [40] H. Munthe-Kaas and S. Krogstad. On enumeration problems in Lie–Butcher theory. *Future Generation Computer Systems*, 19(7):1197–1205, 2003.
- [41] H. Munthe-Kaas and B. Owren. Computations in a free Lie algebra. *R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357(1754):957–981, 1999.
- [42] H.Z. Munthe-Kaas and W. Wright. On the Hopf algebraic structure of Lie group integrators. *Found. Comput. Math.*, 8(2):227 – 257, 2008, math/0603023v1.
- [43] A. Murua. Formal series and numerical integrators, Part I: Systems of ODEs and symplectic integrators. *Applied numerical mathematics*, 29(2):221–251, 1999.
- [44] B. Owren. Order conditions for commutator-free Lie group methods. *Journal of Physics A–Mathematical and General*, 39(19):5585–5600, 2006.
- [45] B. Owren and A. Marthinsen. Runge–Kutta methods adapted to manifolds and based on rigid frames. *BIT*, 39(1):116–142, 1999.
- [46] F. Patras. On Dynkin and Klyachko idempotents in graded bialgebras. *Advances in Applied Mathematics*, 28(3/4):560–579, 2002.
- [47] C. Reutenauer. *Free Lie algebras*. Oxford University Press, 1993.
- [48] M. E. Sweedler. *Hopf algebras*. Mathematics Lecture Note Series. W. A. Benjamin, Inc., New York, 1969.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BERGEN, JOHANNES BRUNSGATE 12, N-5008
BERGEN, NORWAY

E-mail address: alexander.lundervold;hans.munthe-kaas@math.uib.no

TABLE 1. Examples of the coproduct Δ_N , defined in (4.31).

A combinatorial and field theoretic path to quantum gravity: the new challenges of group field theory

D. Oriti

ABSTRACT. Group field theories are a new type of field theories over group manifolds and a generalization of matrix models, that have recently attracted much interest in quantum gravity research. They represent a development of and a possible link between different approaches such as loop quantum gravity and simplicial quantum gravity. After a brief introduction to the GFT formalism we put forward a long but still far from exhaustive list of open issues that this line of research faces, and that could be represent interesting challenges for mathematicians and mathematical physicists alike.

1. Introduction

This article (as our talk at the workshop) has one main goal: to draw attention on one recent approach to quantum gravity, named group field theory, and on some of the many outstanding open issues of this approach, in particular those that can be (in our humble opinion) both mathematically attractive for those mathematicians and mathematical physicists interested in combinatorics and quantum field theory, as well as physically crucial from a quantum gravity perspective. The hope and the expectation (as stressed during the workshop) are therefore that more mathematicians and mathematical physicists will join quantum gravity theorists in the analysis and development of the group field theory formalism, for mutual benefit and amusement, as well as for speeding up progress towards a complete theory of quantum gravity.

Group field theories [1, 2, 3] can be understood on the one hand as quantum field theories on particular group manifolds, which are combinatorially non-local in a sense to be clarified below, and on the other hand as a generalization of matrix models of 2d quantum gravity. This generalization takes place at two levels: in the added group structure and thus in the number of dynamical degrees of freedom (GFTs are true quantum field theories in contrast to the quantum mechanical matrix models) and in the combinatorial structure of the Feynman diagrams of the theory, arising in perturbative expansion (general n -complexes as opposed to 2d ones). Indeed one can also see GFTs as obtained by adding group structure to so-called tensor models (see e.g. [7]), in turn a purely combinatorial generalization of matrix models. A much richer framework follows, much of it being still

unexplored from a mathematical as well as physical perspective. The basic interpretative framework of group field theory goes as follows: the GFT field, in models (aiming at) describing D-dimensional quantum gravity, is interpreted as a second quantized (D-1)-simplex, with (D-2)-faces of the same labelled by group theoretic data, interpreted as (pre-)geometric elementary quantities, or discrete quantum gravity variables. Equivalently, the same data can be associated to the links of a topologically dual graph, and the field is then seen as the second quantization of a spin network functional [4]. This means that GFTs can be seen equivalently as a second quantized formulation of spin network dynamics or as a field theory *of* simplicial geometry.

Already at this point one reason of interest in GFTs becomes apparent: GFTs can potentially represent a common framework for different current approaches to quantum gravity, in particular canonical loop quantum gravity[4] and simplicial quantum gravity formalisms, namely quantum Regge calculus [5] and (causal) dynamical triangulations [6], because the same mathematical structures that characterize these approaches also enter necessarily and in very similar fashion in the GFT framework. Let us be slightly more explicit here, even though all of this will become clear later on when the GFT formalism will be described in more detail. The connection with loop quantum gravity arises first of all because GFT boundary states are given by (open or closed) spin network states, i.e. graphs labelled by group representations or group elements, which are indeed the kinematical quantum states of gravity as discovered by loop quantum gravity. Also, when the GFT partition function is expanded in Feynman diagrams, they turn out to be given by spin foams, i.e. labelled 2-complexes that first arose in the loop quantum gravity context to represent the histories of spin network states, and the Feynman amplitudes that are associated to them are nothing else than spin foam models, i.e. combinatorial and algebraic sum over histories first introduced to encode the dynamics of loop quantum gravity states. The same Feynman diagrams, as we shall see, identify simplicial complexes to which the GFT assigns geometric data, weighted by amplitudes that can be derived from or related to path integrals for simplicial gravity on the given complex. Thus, for given Feynman diagram, the GFT provides a model quantization of gravity in the spirit and language of quantum Regge calculus, while, for given assignment of field degrees of freedom (i.e. fixing the geometric data), the same GFT provides a definition of the dynamics of quantum geometry via a sum over triangulations (the perturbative expansion of the same GFT) in the same spirit of the dynamical triangulations approach.

We refer the reader to the literature (especially [1]) for a more extensive discussion on this. Here we stress only that, this set of relations between different approaches to quantum gravity, within the GFT formalism, suggests that any future development and improved understanding of any aspect of GFTs is likely to have implications and an impact in all of them. In this sense, at the very least, GFTs may play a crucial role in current quantum gravity research, in our opinion.

2. The GFT formalism

We now proceed to introduce the GFT formalism, in its main features. Our treatment is going to be rather sketchy, and we refer once more to the literature, in particular the reviews [1, 2, 3], for a more complete and detailed treatment and a more extensive list of references.

2.1. Kinematics. We start from a field taken to be a \mathbb{C} -valued function of D group elements, for a generic group G , one for each of the D boundary (D -2)-faces of the (D -1)-simplex that the field ϕ represents:

$$\phi(g_1, g_2, \dots, g_D) : G^{\times D} \rightarrow \mathbb{C}.$$

We can identify the ordering of the arguments of the field with a choice of orientation for the (D -1)-simplex it represents, and we then encode the orientation properties of the corresponding (D -1)-simplex in the complex structure by requiring invariance of the field under even permutations σ of its arguments (that do not change the orientation) and trading odd permutations of them with complex conjugation of the field. Given this symmetry, then, a more precise definition of the field is: $\phi(g_1, \dots, g_D) \equiv \sum_{\sigma} \phi(g_{\sigma(1)}, \dots, g_{\sigma(D)})$. Other symmetry properties can also be considered.

An additional symmetry that is usually imposed on the field (but this is again model-dependent, of course) is the invariance under diagonal action of the group G on the D arguments of the field: $\phi(g_1, \dots, g_D) = \phi(g_1 g, \dots, g_D g)$. This is the simplicial counterpart of the Lorentz gauge invariance of continuum and discrete first order gravity actions, and it has also the geometric interpretation, at the simplicial level, of requiring the D faces of a (D -1)-simplex to close to form the close S^{D-2} surface representing its boundary. This reduces the number of degrees of freedom from those represented by the group elements associated to the $(D-2)$ faces, but at the same time intertwines them in a rather non-trivial way, as it is clear by the resulting more complicated geometry and topology of configuration space.

As in any other field theory, a momentum representation for the field and its dynamics is also available and is obtained by harmonic analysis on the group manifold G . The field can be expanded in modes as:

$$\phi(g_i) = \sum_{J_i, \Lambda, k_i} \phi_{k_i}^{J_i \Lambda} \left(\prod_i D_{k_i l_i}^{J_i}(g_i) \right) C_{l_1 \dots l_D}^{J_1 \dots J_D \Lambda},$$

with the J 's labelling representations of G , the k 's vector indices in the representation spaces, and the C 's being intertwiners of the group G . We have labelled an orthonormal basis of intertwiners by an extra parameter Λ (depending on the group chosen and on the dimension D , this may actually be a shorthand notation for a set of parameters). That this decomposition is possible is not guaranteed in general (as the harmonic analysis of non-compact groups, for example, may not be under control), but it is in fact true for all the known quantum gravity GFT models, which are based on the Lorentz group or on extensions of it. A geometric interpretation of the field variables is obtained either looking at the Feynman amplitudes for the GFT at hand, in turn usually obtained from a discretization of some continuum gravity action, or from the direct (1st) quantization of simplicial structures, e.g. by geometric quantization methods. The group variables are then seen to represent parallel transport of a (gravity) connection along elementary paths dual to the (D -2)-faces, and the representations J are usually put in correspondence with the volumes of the same (D -2)-faces, the details of this correspondence depending, however, on the specific model.

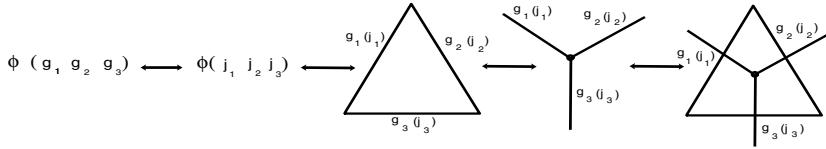


FIGURE 1. For the $D = 3$ case, the association of a field with a 2-simplex, a triangle, or equivalently its dual vertex, and of its arguments with the edges (1-faces) of the triangle, or equivalently with the links incident to the vertex, together with the consequent labelling of the by group-theoretic variables.

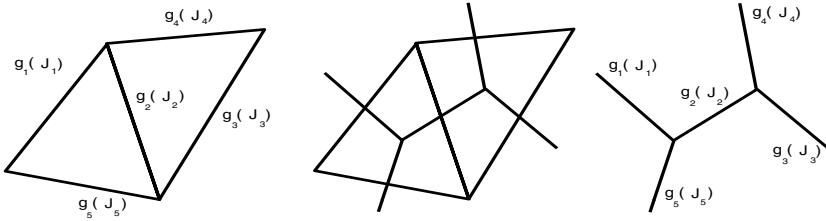


FIGURE 2. A ‘2-particle state’(again, in the $D=3$ example)

Just as one identifies a single field (or the corresponding 1st quantized wave function) with a single $(D-1)$ -simplex, a simplicial space built out of N such $(D-1)$ -simplices is described by the tensor product of N such wave functions, with suitable constraints implementing the fact that some of their $(D-2)$ -faces are identified. For example, a state describing two $(D-1)$ -simplices glued along one common $(D-2)$ -face would be represented by: $\phi_{k_1 k_2 \dots k_D}^{J_1 J_2 \dots J_D \Lambda} \phi_{\tilde{k}_1 \tilde{k}_2 \dots \tilde{k}_D}^{\tilde{J}_1 \tilde{J}_2 \dots \tilde{J}_D \tilde{\Lambda}}$, where the gluing is along the face labelled by the representation J_2 , and effected by the contraction of the corresponding vector indices (of course, states corresponding to disjoint $(D-1)$ -simplices are also allowed).

We see that states of the theory are then labelled, in momentum space, by *spin networks* of the group $G[4]$. The corresponding second quantized theory is a GFT. Spin networks are also relevant for the construction of GFT observables. These are given [3] by gauge invariant functionals of the GFT field, and can be constructed in momentum space using spin networks according to the formula:

$$O_{\Psi=(\gamma, j_e, i_v)}(\phi) = \left(\prod_{(ij)} \int dg_{ij} dg_{ji} \right) \Psi_{(\gamma, j_e, i_v)}(g_{ij} g_{ji}^{-1}) \prod_i \phi(g_{ij}),$$

where $\Psi_{(\gamma, j_e, i_v)}(g)$ identifies a spin network functional [4] for the spin network labelled by a graph γ with representations j_e associated to its edges and intertwiners i_v associated to its vertices, and g_{ij} are group elements associated to the edges (ij) of γ that meet at the vertex i .

2.2. Dynamics. On the basis of the above kinematical structure, one aims at defining a field theory for describing the interaction of fundamental building blocks of space (($D-1$)-simplices or spin network vertices), and in which a typical interaction process will be characterized by a D -dimensional simplicial complex. In

the dual picture, the same will be represented as a spin foam (labelled 2-complex). This is the straightforward generalization of the way in which 2d discretized surfaces emerge from the interaction of matrices (graphically, segments), or ordinary Feynman graphs emerge from the interaction of point particles. A *discrete* spacetime emerge then from the theory as a virtual construct, a possible interaction process among the quanta of the theory.

In order for this to be realized, the classical field action in group field theories has to be chosen appropriately. In this choice lies the main peculiarity of GFTs with respect to ordinary field theories. This action, in configuration space, has the general QFT structure:

$$(1) \quad S_D(\phi, \lambda) = \frac{1}{2} \left(\prod_{i=1}^D \int dg_i d\tilde{g}_i \right) \phi(g_i) \mathcal{K}(g_i \tilde{g}_i^{-1}) \phi(\tilde{g}_i) \\ + \frac{\lambda}{(D+1)!} \left(\prod_{i \neq j=1}^{D+1} \int dg_{ij} \right) \phi(g_{1j}) \dots \phi(g_{D+1j}) \mathcal{V}(g_{ij} g_{ji}^{-1}),$$

and it is of course the choice of kinetic and interaction functions \mathcal{K} and \mathcal{V} that define the specific model considered. Obviously, the same action can be written in momentum space after harmonic decomposition on the group manifold.

The mentioned peculiarity is in the combinatorial structure of the pairing of field arguments in the kinetic and vertex terms, as well as from their degree as polynomials in the field. The interaction term describes the interaction of D+1 (D-1)-simplices (the fundamental ‘atoms of space’) to form a D-simplex (‘the fundamental virtual ‘atom of spacetime’) by gluing along their (D-2)-faces (arguments of the fields), that are *pairwise* linked by the interaction vertex. The nature of this interaction is specified by the choice of function \mathcal{V} . The (quadratic) kinetic term involves two fields each representing a given (D-1)-simplex seen from one of the two D-simplices (interaction vertices) sharing it, so that the choice of kinetic functions \mathcal{K} specifies how the information and therefore the geometric degrees of freedom corresponding to their D (D-2)-faces are propagated from one vertex of interaction (fundamental spacetime event) to another. Let us mention here that one can consider generalizations of the above combinatorial structure, for example defining vertex functions whose combinatorics corresponds to the gluing of (D-1)-simplices to form different sorts of D-dimensional complexes (e.g. hypercubes etc). We are not going to elaborate further on this.

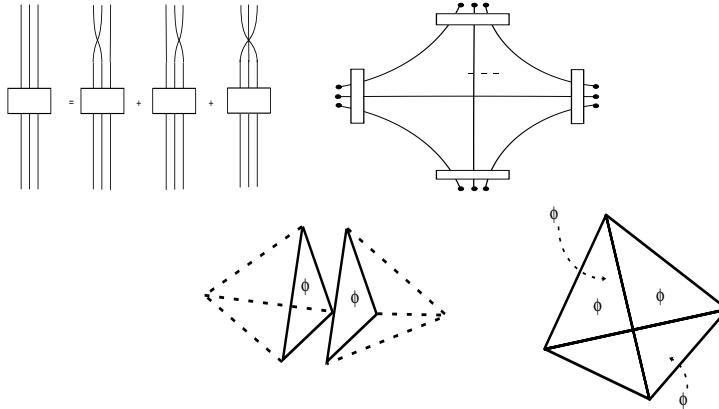
Because of the peculiar way in which field arguments are paired in the interaction term, we may consider GFTs as *combinatorially non-local field theories* (as opposed to theories in which the nonlocality is originated by higher derivatives terms, for example). Before detailing more the interesting structure of Feynman diagrams resulting from this combinatorial nonlocality, let us stress once more that it is basically in it that lies the peculiarity of GFTs as field theories. Indeed, as for the rest, we have an almost ordinary field theory, in that we can rely on a fixed background metric structure, given by the invariant Killing-Cartan metric on the group manifold, a fixed topology, given again by the topology of the group manifold, the usual splitting between kinetic (quadratic) and interaction (higher order) term in the action, that will later allow for a straightforward perturbative expansion, and the usual conjugate pictures of configuration and momentum space. This allows us

to use all usual QFT techniques and language in the analysis of GFTs, and thus of quantum gravity, even though we remain in a background independent (in the physical sense of ‘spacetime independent’) context. The importance of this, in a non-perturbative quantum gravity framework, should not be underestimated, we think.

Let us now turn to the quantum dynamics. Most of the research in this area has concerned the perturbative aspects of this dynamics around the no-particle state, the complete vacuum, and the main guide for model building have been, up to now, only the properties of the resulting Feynman amplitudes. This relevant Feynman expansion is:

$$Z = \int \mathcal{D}\phi e^{-S[\phi]} = \sum_{\Gamma} \frac{\lambda^{N_v(\Gamma)}}{sym[\Gamma]} Z(\Gamma),$$

where N_v is the number of interaction vertices v in the Feynman diagram Γ , $sym[\Gamma]$ is the number of automorphisms of Γ and $Z(\Gamma)$ the corresponding Feynman amplitude. Each edge of the Feynman graph is made of D strands, one for each argument of the field¹ and each one is then re-routed at the interaction vertex, with the combinatorial structure of an D -simplex, following the pairing of field arguments in the vertex operator. Diagrammatically:



Each strand in an edge of the Feynman diagram goes through several vertices, coming back where it started, for closed Feynman diagrams, and therefore identifies a 2-cell (for open graphs, it may end up on the boundary, identifying then an open 2-cell). Each Feynman diagram Γ is then a collection of 2-cells (faces), edges and vertices, i.e. a 2-complex, that, because of the chosen combinatorics for the arguments of the field in the action, is topologically dual to a D -dimensional simplicial complex. Notice that the resulting 2-cells can be glued (i.e. can share edges) in all sorts of ways, forming for example “bubbles”, i.e. closed 3-cells.

¹One could write explicitly down and keep track of each of the spacetime coordinates on which the field depends also in ordinary QFT in Minkowski space, but it would be a pedantic and rather useless complication. The need to keep track of each argument of the field, in a GFT context, comes from the mentioned nonlocality of the interaction term, which is at the origin of the combinatorially nontrivial structure of the resulting Feynman diagrams.

No restriction on the topology of the resulting diagram/complex is imposed, a priori, in their construction, so the resulting complexes/triangulations can have arbitrary topology. Each resulting 2-complex or triangulation corresponds to a particular *scattering process* of the fundamental building blocks of space, i.e. $(D-1)$ -simplices. Each line of propagation, made as we said out of D strands, is labelled, on top of the group/representation data, by a permutation of $(1, \dots, D)$, representing the labelling of the field variables, and all these data are summed over in the construction of the Feynman expansion. The sum over permutations affects directly the combinatorics of the allowed gluings of vertices with propagators. The above choice of permutation symmetry for the field (with the orientation encoded in the complex structure) implies that only *even* permutation appear as labellings of propagation lines. In turn, this ensures that only orientable complexes are generated in the Feynman expansion of the field theory (see [8] for a more detailed treatment). The more restrictive choice of invariance of the field under *any* permutation of its arguments results, for example, in the presence of non-orientable complexes as well in the Feynman expansion.

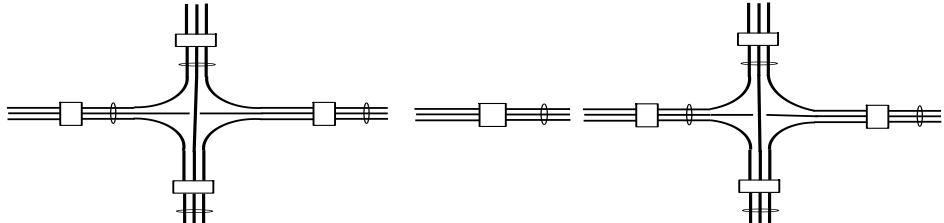


FIGURE 3. The gluing of vertices of interaction through propagators, again in the $D=3$ example. The rectangles represent the additional integrations imposing gauge invariance under the action of G , while the ellipses represent the implicit sum over permutations of the (labels of the) strands to be glued.

As said, each strand in a propagation line carries a field variable, i.e. a group element in configuration space or a representation label in momentum space. After the closure of the strand to form a 2-cell in a closed diagram, the same representation label ends up being associated to this 2-cell. Therefore in momentum space each Feynman graph is given by a spin foam (a 2-complex with faces labelled by representation variables), and each Feynman amplitude (a complex function of the representation labels, obtained by contracting vertex amplitudes with propagator functions) by a so-called spin foam model [9]:

$$Z(\Gamma) = \sum_{J_f} \prod_f A_f(J_f) \prod_e A_e(J_{f|e}) \prod_v A_v(J_{f|v}),$$

where we have highlighted the fact that the amplitudes can be factori Given the mentioned geometric interpretation of the representation variables (edge lengths, areas, etc) (see [9, 1]), each of these Feynman amplitudes corresponds to a definition of a sum-over-histories for discrete quantum gravity on the specific triangulation dual to the Feynman graph, although the quantum amplitudes for each geometric configuration are not necessarily given by the exponential of a discrete gravity action.

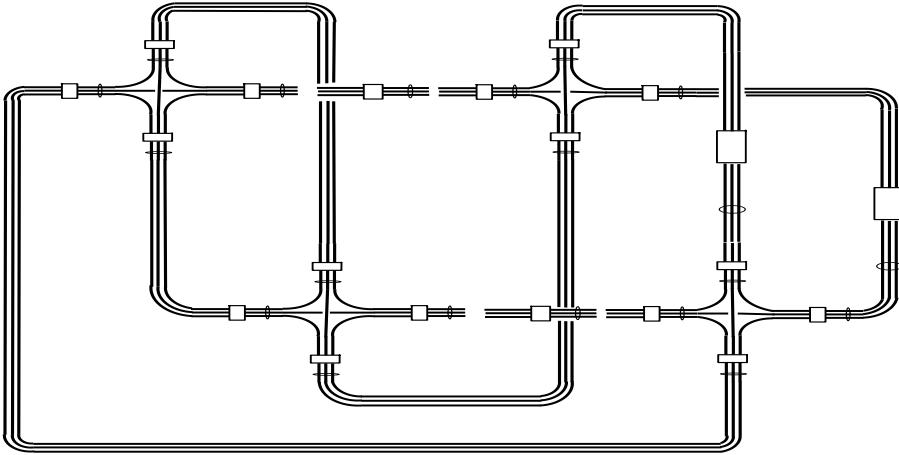


FIGURE 4. An example, in $D=3$, of a closed GFT Feynman diagram, with 4 vertices and 8 propagators.

One can show that the inverse is also true: any local spin foam model can be obtained from a GFT perturbative expansion [11, 3] (even though this does not imply that the reconstruction of the underlying GFT action from the knowledge of the spin foam amplitudes is immediate nor easy). This implies on the one hand that the GFT approach incorporates the spin foam in its perturbative aspects, and on the other hand that it goes potentially far beyond it, since there is of course much more in a QFT than its perturbative expansion. The sum over Feynman graphs gives then a sum over spin foams, and equivalently a sum over triangulations, augmented by a sum over algebraic data (group elements or representations) with a geometric interpretation, assigned to each triangulation. This perturbative expansion of the partition function also allows for a perturbative evaluation of expectation values of GFT observables, as in ordinary QFT. In particular, the transition amplitude (probability amplitude for a certain scattering process) between certain boundary data represented by two spin networks, of arbitrary combinatorial complexity, can be expressed as the expectation value of the field operators having the same combinatorial structure of the two spin networks [3, 1].

$$\langle \Psi_1 | \Psi_2 \rangle = \int \mathcal{D}\phi O_{\Psi_1} O_{\Psi_2} e^{-S(\phi)} = \sum_{\Gamma / \partial\Gamma = \gamma_{\Psi_1} \cup \gamma_{\Psi_2}} \frac{\lambda^N}{\text{sym}[\Gamma]} Z(\Gamma)$$

where the sum involves only 2-complexes (spin foams) with boundary given by the two spin networks chosen.

The above perturbative expansion involves thus two types of sums: one is the sum over geometric data (group elements or representations of G) entering the definition of the Feynman amplitudes as the GFT analogue of the integral over momenta or positions of usual QFT; the other is the overall sum over Feynman diagrams. We stress again that, in absence of additional restrictions being imposed

on the GFT, the last sum includes a sum over all triangulations for a given topology and a sum over all topologies².

2.3. Examples. We now give a few examples of specific GFT models, again referring to the literature for more details.

2.3.1. *D=2 and matrix models.* The easiest example is a straightforward generalization of matrix models for 2d quantum gravity to a GFT, obtained by adding group structure to them, but keeping the same combinatorics, and it is given:

(2)

$$S[\phi] = \int_G dg_1 dg_2 \frac{1}{2} \phi(g_1, g_2) \phi(g_1, g_2) + \frac{\lambda}{3!} \int dg_1 dg_2 dg_3 \phi(g_1, g_2) \phi(g_1, g_3) \phi(g_2, g_3)$$

where G is a generic compact group, say $SU(2)$, and the symmetries mentioned above are imposed on the field ϕ implying, in this case: $\phi(g_1, g_2) = \tilde{\phi}(g_1 g_2^{-1})$. The relation with matrix models is apparent in momentum space, expanding the field in representations j of G to give:

$$(3) \quad S[\tilde{\phi}] = \sum_j \text{dim}(j) \left(\frac{1}{2} \text{tr}(\tilde{\phi}_j^2) + \frac{\lambda}{3!} \text{tr}(\tilde{\phi}_j^3) \right)$$

where the field modes $\tilde{\phi}_j$ are indeed matrices with dimension $\text{dim}(j)$. Thus, one gets a sum of matrix models actions of increasing dimensions. Alternatively, one can see the above as the action for a single matrix model in which the dimension of the matrices has been turned from a parameter into a dynamical variable.

2.3.2. *D=3 and topological BF theory.* Another much studied example (actually the best understood one) is the group field theory, first proposed by D. Boulatov, for topological BF theory in 3d, which in turn is closely related to 3-dimensional quantum gravity in first order formalism.

Kinetic and vertex terms are then chosen as follows:

$$(4) \quad \mathcal{K}(g_i, \tilde{g}_i) = \int_G dg \prod_i \delta(g_i \tilde{g}_i^{-1} g), \quad \mathcal{V}(g_{ij}, g_{ji}) = \prod_i \int_G dg_i \prod_{i < j} \delta(g_i g_{ij} g_{ji}^{-1} g_j^{-1}),$$

where the integrals impose the gauge invariance under the action of G . The choice of $G = SO(3)$ or $G = SO(2, 1)$ provides a quantization of 3d gravity in the Euclidean and Minkowskian signatures, respectively, and the Feynman amplitudes are given by the so-called Ponzano-Regge spin foam model. The choice of the quantum group $SU(2)_q$ gives the Turaev-Viro topological invariant, conjectured to correspond, when q is a root of unity, to 3d gravity with positive cosmological

²There is no algorithmic procedure that allows to distinguish topologies in $D \geq 3$, so we cannot partition this sum into the two sub-sums mentioned. What ensures us that all topologies are present in the full sum is, however, that any simplicial complex is obtained by an appropriate gluing and face identification of fundamental simplicial building blocks, and that all possible such gluings and identifications are included in the GFT Feynman expansion.

constant. For the $SU(2)$ case, and for real field, the action is then:

$$(5) \quad S[\phi] = \prod_i \int_{SU(2)} \phi(g_1, g_2, g_3) \phi(g_1, g_2, g_3) \\ + \frac{\lambda}{4!} \prod_{i=1}^6 \int_{SU(2)} dg_i \phi(g_1, g_2, g_3) \phi(g_3, g_4, g_5) \phi(g_5, g_2, g_6) \phi(g_6, g_4, g_1),$$

and the Feynman amplitudes are:

$$Z(\Gamma) = \left(\prod_{e^* \in \Gamma} \int dg_{e^*} \right) \prod_{f^*} \delta \left(\prod_{e^* \in \partial f^*} g_{e^*} \right)$$

where e^* are dual edges of the 2-complex and f^* its 2-cells. This is the same quantity one obtains from a path integral quantization of discretized BF theory on the triangulation dual to Γ , confirming the above-given interpretation.

Lots is known about the last model, including the appropriate gauge fixing procedure of its Feynman amplitudes, the coupling of matter fields, etc. and we refer once more to the literature for more details, starting from [10].

2.3.3. *Models with additional structure.* One may then consider models with additional structure and more involved kinetic and vertex operators. One example are the generalized models of [12], based, in the $D=3$ case, on a complex field on $(G \times \mathbb{R})^{\times 3}$, but with a G -only diagonal invariance, with $G = SU(2)$ or $G = SL(2, \mathbb{R})$ in the Riemannian and Lorentzian cases respectively. The kinetic and vertex operators are then:

$$(6) \quad \mathcal{K}(g_i, s_i, \tilde{g}_i, \tilde{s}_i) = \sum_{\sigma} \int dg \prod_i (i\partial_{s_i} + \square_i) \delta(g_i g \tilde{g}_{\sigma(i)}^{-1}) \delta(s_i - \tilde{s}_{\sigma(i)}^{-1}) \\ \mathcal{V}(g_{ij}, s_{ij}) = \prod_{i \neq j} \delta(g_{ij} g_{ji}^{-1}) \delta(s_{ij} - s_{ji})$$

where $g_i \in G$, $s_i \in \mathbb{R}$ and \square is the Laplace-Beltrami operator on G .

One more example of a similar type, a sort of ‘relativistic’ upgrading of the previous one, and that is currently being studied and developed [13], as the possible explicit common ground for loop quantum gravity and simplicial quantum gravity, uses instead a complex field over $(G \times \mathbb{R}^3)^{\times 3}$ and the kinetic and vertex terms:

$$(7) \quad \mathcal{K}(g_i, x_i, \tilde{g}_i, \tilde{x}_i) = \sum_{\sigma} \int dg \prod_i (\Delta_i + \square_i) \delta(g_i g \tilde{g}_{\sigma(i)}^{-1}) \delta(x_i - \tilde{x}_{\sigma(i)}^{-1}) \\ \mathcal{V}(g_{ij}, x_{ij}) = \prod_{i \neq j} \delta(g_{ij} g_{ji}^{-1}) \delta(x_{ij} - x_{ji})$$

where $g_i \in G$, $x_i \in \mathbb{R}^3$, Δ is now the Laplace-Beltrami on \mathbb{R}^3 and \square is again the Laplace-Beltrami on G . The interest in this last model lies in the fact that the corresponding Feynman amplitudes have *exactly* the form of path integrals for simplicial quantum gravity, i.e. quantum Regge calculus, in 1st order form. Similar models exist in higher dimension ($D \geq 3$)[13].

3. A (far from comprehensive) list of open issues

We now outline, in a rather schematic way, a list of open issues within the group field theory approach, that are in our opinion of considerable mathematical interest as well as crucial for a better physical understanding of this class of models. We stress once more that this list is, by all means, far from exhaustive, even from a purely mathematical perspective. At the same time, we also stress that it is only for the stated aims of this article that we do not review or discuss more what, instead, is known about GFTs, and the many results already obtained in this approach. For an account of them, we refer to the cited literature [1, 2, 3].

3.1. Classical solutions and their relevance. GFTs were historically developed from the perspective of either loop quantum gravity and spin foam models or dynamical triangulations, and thus the focus of attention has always been their perturbative structure and the properties of their Feynman amplitudes. However, changing slightly the perspective on them, and seeing them just as field theories (of a peculiar type), the first thing one would think of analyzing is their classical structure. In particular, one would start from their classical equations of motion. The importance of these equations from the GFT point of view are obvious: they define the classical dynamics of the field theory, they identify of classical background configurations around which to expand in a semi-classical perturbative definition of the path integral, etc. Their relevance from a quantum gravity perspective is easily understood if one recalls the interpretation of GFTs as second quantized theories of simplicial geometry or of spin network states. Then the GFT classical action should encode the full 1st quantized dynamics, and the classical equations of motion should correspond to the full 1st quantized wave function equations. In the specific case of gravity, then, solving the GFT classical equations means identifying non-trivial quantum gravity wave functions satisfying **all** the quantum gravity constraints, an important and still unachieved goal of canonical quantum gravity, even in the modern loop quantized formalism.

The classical equations of motion following from the general form of the action 1 (considering for simplicity a real field) are:

$$\prod_j \int d\tilde{g}_{1j} \mathcal{K}(g_{1j} \tilde{g}_{1j}^{-1}) \phi(\tilde{g}_{1j}) + \frac{\lambda}{D!} \left(\prod_{j=1}^{D+1} \int dg_{2j} \dots dg_{D+1j} \right) \phi(g_{2j}) \dots \phi(g_{D+1j}) \mathcal{V}(\{g_{ij}\}) = 0.$$

For a generic differential operator \mathcal{K} , they are then rather complicated non-linear equations of integro-differential type, with the main complications coming once more from the particular pairing of variables in the interaction term. No detailed analysis of these equations in any specific GFT model nor of their solutions has been carried out to date.

Solving these equations in complete generality, obtaining a complete classification of their exact solutions, is probably beyond reach, even in low D . However, two more modest goals are probably within reach, and both would be of importance.

1) Identify at least *some* exact solutions of the full GFT equations, for some currently studied GFT model, and analyze their mathematical structure and physical meaning. Most likely, this has to be done first in symmetry reduced cases, and, on top of identifying some particularly simple solution of a given GFT model, it would be of considerable interest to develop a general theory of symmetry reduction for GFTs, to be then applied to the various models.

2) Develop an exact formalism for obtaining *approximate* solutions to the above field equations, e.g. in perturbation expansion in the coupling parameter λ , or in some other parameter that could allow to take into account both the non-linearity of the equations and the non-local nature of the field coupling in a perturbative way.

Indeed, the classical solutions of the (local and linear) free field theory equations

$$\left(\prod_j \int d\tilde{g}_j \right) \mathcal{K}(g_j \tilde{g}_j^{-1}) \phi(\tilde{g}_j) = 0$$

are rather easily obtained, but at the same time are not very interesting physically, for most known models.

Let us now give one example of a GFT equation of motion, to clarify some of the above issues. Take the, arguably simplest, GFT model of equation (4); the corresponding equations are:

$$\begin{aligned} \int d\tilde{g} \phi(g_1 \tilde{g}, g_2 \tilde{g}, g_3 \tilde{g}) + \lambda \prod_{i=1}^3 \int d\tilde{g}_i \prod_{j=4}^6 \int dg_j \phi(g_3 \tilde{g}_1, g_4 \tilde{g}_1, g_5 \tilde{g}_1) \\ \times \phi(g_5 \tilde{g}_2, g_6 \tilde{g}_2, g_2 \tilde{g}_2) \phi(g_6 \tilde{g}_3, g_4 \tilde{g}_3, g_1 \tilde{g}_3) = 0 \end{aligned}$$

where we have included additional integrals implementing the gauge invariance. We do not know the general solution of this equation, nor any exact method for obtaining it or any exact procedure for approximating it. We do know some exact simplified (families of) solutions [14], though, which are the following:

$$\phi_f(g_1, g_2, g_3) = \sqrt{\frac{3!}{-\lambda}} \int dg \delta(g_1 g) f(g_2 g) \delta(g_3 g),$$

parametrized by a real function f on $SU(2)$. More solutions of the same type can be constructed, even in higher dimensions.

However, there is a lot yet to be clarified concerning the physical meaning of these simple known solutions, and how their mathematical structure implements it. We refer in particular to the issue of topology change. Let us clarify. As discussed above, the GFT of equation (4) provides a quantization of topological BF theory in 3d, which takes into account topology changing configurations, as it is clear in perturbative expansion. Accordingly, *any* classical solution should be interpreted as the wave function for a *flat* geometry and some prescribed dynamics for the *topological* degrees of freedom. How a flat geometry is implemented in the above solutions, and which prescription for the dynamics for space topology is encoded in them, as well as how the two ingredients are intertwined, is mathematically yet to be understood. Same applies to other known simple solutions.

The same issues, from the purely mathematical ones to those concerning the physical interpretation, are of course even more complex and interesting at the

same time in the $D = 4$ case, for non-topological models, where geometry itself has a highly non-trivial dynamics. For example, the model of equation (7), which presents already a set of non-trivial free field solutions (e.g. (the group analogue of) plane waves on both \mathbb{R}^3 and G).

3.2. Classical and quantum Hamiltonian analysis. Still sticking to a purely formal field theoretic perspective on GFTs, i.e. keeping aside momentarily the quantum gravity interpretation of the same, one more basic aspect of them is both undeveloped, and mathematically challenging: their canonical/Hamiltonian formulation. A general formalism for the canonical analysis of GFTs at both classical and quantum levels has been recently proposed [16], and applied to the model of equation (6), but it represents, especially at the mathematical level, only a first step.

What are the main difficulties involved in this analysis for GFTs?

Once more, they stem from the peculiar combinatorial structure of the GFT action, but also from the need, resulting from the quantum gravity interpretation, to deal with all of the field arguments on equal footing, since they propagate almost independently from each other (due again to the combinatorics in the vertex).

Let us be more explicit. Consider the action with kinetic and vertex term of equation (6), and restrict the attention to the kinetic term only (determining the symplectic structure of the theory and thus providing the basis for the hamiltonian analysis):

$$S = \left[\prod_i \int_G dg_i \int_{\mathbb{R}} ds_i \right] \phi^\dagger(g_1, s_1; \dots; g_D, s_D) \prod_i (i\partial_{s_i} + \square_i) \phi(g_1, s_1; \dots; g_D, s_D) + h.c.$$

for generic group G (Riemannian or Lorentzian).

The kinetic term has the structure of a product of differential operators, each acting independently on one of the D (sets of) arguments of the field. Each of them is a Schroedinger-like operator with “Hamiltonian” \square . This suggests that one should consider the variables s_i as “time” variables, to be used in a GFT generalization of the usual time+space splitting of the configuration space coordinates, with the group elements treated instead as “space”. This is reasonable (and indeed the correct thing to do), but it implies we have a field theory with D “times”, all to be treated on equal footing. In turn this immediately implies that the naive phase space has coordinates $(\phi, \pi_i = \frac{\delta L}{\delta \partial_{s_i} \phi})$, with one field and D conjugate momenta. Clearly, a generalization of usual Hamiltonian mechanics is needed. One approach, and the one chosen in [16], is to use the DeDonder-Weyl generalized Hamiltonian mechanics, as developed at both the classical and quantum level by Kanatchikov [15], as a starting point and to adapt it to the peculiar GFT setting. The framework chosen is thus that of *polysymplectic (or polymomentum) mechanics* [15], and we refer to the literature for more mathematical details on this beautiful formalism.

The general idea of how to adapt this general formalism to the GFT case is the following. One starts from a “covariant” definition of momenta, hamiltonian density, Poisson brackets, etc treating all “time variables” on equal footing at first, i.e. when defining densities. Then one defines the ‘scalar’ quantities referring to each ‘time direction’ (to be turned into operators at the quantum level), including a set of D Hamiltonians, by integration over appropriate hypersurfaces in $(G \times \mathbb{R})^{\times D}$, so that each Hamiltonian refers to a single time direction, but at the same time all

time directions are treated equally but independently, corresponding to an equal but independent evolution (propagation) of the corresponding degrees of freedom. A similar procedure is adopted for other canonical quantities, e.g. Poisson brackets, scalar products etc.

Let us sketch one example of such procedure, for the case $D = 2$ and the above choice of GFT action, outlining only the definition of the Hamiltonians, in order to convey at least the flavour of the formalism. We refer to [16] for the full treatment of this case as well as for a full, but still in many ways preliminary, description of the GFT Hamiltonian analysis, while we refer to [15] for the complete exposition of the state of the art in polymomentum Hamiltonian mechanics.

We start from the naive phase space $(\phi, \phi^\dagger, \pi_\phi^i = \frac{\delta L}{\delta \partial_{s_i} \phi}, \pi_{\phi^\dagger}^i = \frac{\delta L}{\delta \partial_{s_i} \phi^\dagger})$ (we call it ‘naive’ simply because in the polymomentum formalism phase space variables are actually differential forms on spacetime, while the π above are not), with the product structure of the kinetic term resulting in a peculiar expression for the momenta, e.g. $\pi_\phi^1 = (-i\partial_2 + \square_2)\phi^\dagger$, and define the DeDonder-Weyl Hamiltonian density (summation over repeated indices understood):

$$\mathcal{H}_{DW} = \pi_\phi^i \partial_{s_i} \phi + \pi_{\phi^\dagger}^i \partial_{s_i} \phi^\dagger - L = 2\pi_{\phi^\dagger}^1 \pi_\phi^2 + i\pi_\phi^1 \square_1 \phi + i\pi_\phi^2 \square_2 \phi + h.c..$$

One then proceeds to re-write it as a sum of two contributions, each uniquely associated to a single time parameter: $\mathcal{H}_{DW} = \mathcal{H}_1 + \mathcal{H}_2$, with $\mathcal{H}_i = \pi_{\phi^\dagger}^1 \pi_\phi^2 + i\pi_\phi^i \square_i \phi + h.c.$

The Hamiltonians governing the ‘time evolution’ with respect to the different time directions identified by each variable s_i are then defined by integration over independent hypersurfaces, each orthogonal to a different time direction, e.g. $H_1 = \int ds_2 dg_2 \mathcal{H}_1$. Each H_i results in being independent of time s_i , as one would expect.

One can then proceed, after suitable decomposition in modes of fields and momenta, the definition of (a GFT-adapted version of) the covariant Poisson brackets, etc, to the canonical quantization of the theory, with the definition of a Fock structure on the space of states. In this way, one is able to make precise the intuition of GFTs being a dynamical theory of creation and annihilation of fundamental quanta of space. We refer once more to [16] for the results of this analysis, among which we mention only the interesting interplay between statistics and group structure in GFTs, that causes fields on a Riemannian group G to be quantized as bosons, and fields on a Lorentzian group G to be quantized, necessarily, as fermions, in order to preserve positivity of the Hamiltonians H_i . However, this result may depend rather crucially on the specific choice of kinetic term, and so be model-dependent, but we believe that this shows how the issue of field statistics in GFTs deserves to be further analyzed in more precise and general terms.

This is only one of the many features of the polysymplectic formalism for field theories, in general, and of its GFT incarnation, in particular, that need to be developed and clarified at both the mathematical and physical level.

They include, at the classical level, and for the general polysymplectic formalism: the formulation of a covariant Hamilton-Jacobi theory corresponding to it, and the study of its implications for quantum field theory; an analysis of the relationship between the DeDonder-Weyl polysymplectic field theory and other known covariant extension of ordinary Hamiltonian field theory (see again [15] for a discussion); a

complete analysis of the exact reduction of formalism and results of polymomentum Hamiltonian field theory to the usual single-time Hamiltonian formalism; more technically, there are many unsolved issues concerning the algebraic structure induced on the space of (horizontal) differential forms constituting the generalized phase space by Kanatchikov's definition of the (graded) Poisson bracket for them (that seems to be a generalized type of Gerstenhaber algebra); also, the general theory of conservation laws in polymomentum mechanics, with the corresponding definition of conserved currents and charges, the generalization of Noether theorem, etc, is to be developed in more detail and its physical consequences for known theories (and here we refer to ordinary field theories as well as to GFTs) have to be analyzed, also given the great physical importance attached to them.

At the quantum level, the state of the art is even more full of open issues, given that work on this has really just started: for the quantization of field theories based on the full DeDonder-Weyl-Kanatchikov Hamiltonian theory we refer to [15], and subsequent work by the same author, for the first steps; for the GFT-adapted formalism, the same consideration apply, so we refer to [16]. We only mention three open problems specifically related to the GFT approach. The first is the general issue of conserved quantities and symmetries for GFTs; this is crucial for a better understanding of the appropriate scalar product in the space of fields, or equivalently for the GFT definition of the kinematical inner product for canonical 1st quantized quantum gravity wave functions from the perspective of GFTs, and at the same time would provide the basis for the analysis of symmetries at the quantum level, and thus the consequent gauge fixing of Feynman amplitudes, i.e. spin foam models; we will discuss this issue more in the following. The second is related to the first: what is the GFT analogue of the notion of anti-particles? Being strictly intertwined with the complex structure of the field, one would expect it to be linked with the orientation of the simplices corresponding to the fundamental quanta of the GFT, and this would match some insights coming from recent developments in spin foam models; however, the whole issue is far from clear, and probably is best addressed within an Hamiltonian context, like the one outlined above. The third is the canonical derivation of GFT quantum propagators, i.e. 2-point functions; in particular, one would like to put on more solid grounds the choice of propagator made in the construction of the spin foam models/Feynman amplitudes corresponding to the models of equations (7) and (6); clearly, in light of the above discussion, this means providing, among other things, a suitable definition, in an Hamiltonian setting, of a “multi-time-ordering” of field operators.

3.3. Symmetries. We have mentioned above the open issue of symmetries in group field theories. We now expand on it, trying to clarify what is known and what is not known.

All the past work on symmetries in group field theories has proceeded in a rather awkward way, from a field theoretic perspective. Consider the model of equation (4), which is the only one of which we understand reasonably well the symmetries and their gauge fixing at the level of Feynman amplitudes. As said, the same Feynman amplitudes can be derived from a discretization of the continuum action and path integral of topological BF theory in 3 dimensions. One can [10] identify the discrete analogue of the continuum symmetries of BF theory: the translational and local Lorentz symmetry, at the level of each GFT Feynman diagram/simplicial complex, then devise the appropriate gauge fixing procedure of the spin foam model/GFT

Feynman amplitude, to get rid of the redundant gauge degrees of freedom, and the corresponding Faddeev-Popov determinant, to obtain in the end fully gauge-fixed and finite GFT Feynman amplitude.

Now, first of all, this has been done, to date, -only- for this specific GFT model, and not much is known about the relevant symmetries of other models. Second, even in this case, from a GFT perspective there is still much to be understood. In particular, while the Lorentz symmetry has a clear GFT origin in the invariance of the field under the diagonal action of the group G on its 3 arguments, the GFT origin of the translation symmetry remains mysterious³.

On top of this, as said, it is rather obvious that this is a cumbersome way of proceeding, from a purely field theory perspective: one would study directly the symmetries of the GFT action that originates the Feynman amplitudes under consideration, and derive the relevant perturbative identities between Feynman amplitudes and n -point function, rather than try to guess what these symmetries are from an analysis of the individual Feynman amplitudes, which are effectively path integrals for the corresponding single and multi-particle theories.

This is indeed what needs to be done and understood in all mathematical and physical details. As mentioned, the natural starting point would be the Hamiltonian analysis of symmetries for the various GFTs, and the derivation of the corresponding conserved currents and charges. But one can as well remain at the Lagrangian level and be concerned only with the perturbative expansion of GFTs in terms of spin foam models; in this case, one should identify the various symmetry transformations of the fields in the GFT action, for any specific model considered, and derive from them the corresponding Ward identities for the n -point functions. These would be exactly the identities between spin foam amplitudes that can (and were) in some cases discovered by direct analysis of the classical discrete theory from which the same amplitudes, can be derived. Both the general formalism for deriving Ward identities and explicit examples of symmetry analysis for specific GFT models are yet to be developed.

Notice that this is far from a trivial task, even at the classical level, as one needs to work out a suitable generalization of Noether theorem, adapted to GFTs, that overcomes the difficulties posed by presence of several time variables (leading necessarily to a polymomentum formalism), and by the product form of the general kinetic operator, as well as by the higher-order in the “spacetime” derivatives that results from it, even for simple choices of the kinetic term.

One more reason why this would be of interest is that, as it was stressed forcefully by P. Cvitanovic in his lectures at the workshop, the presence of symmetries and of the consequent Ward identities in a field theory affects greatly, among other things, the growth rate and convergence properties of the corresponding perturbative series, and thus in the GFT case it can tell us a lot about the mathematical structure and properties of the sum over topologies, as well as over geometries, implicit in their expansion in Feynman diagrams.

3.4. Combinatorial structure of the Feynman diagrams. Let us now move on, indeed, to the discussion of the open problems concerning the Feynman diagrams themselves. These have to do with their combinatorial structure mainly,

³And a better understanding of this translation symmetry in GFTs would be of great relevance from a quantum gravity point of view since it has been shown [10] that it is strictly related to the discrete Bianchi identities on the simplicial complex dual to the GFT Feynman diagram.

and can be motivated by the need, and at the same time the chance that group field theories offer, to re-phrase quantum gravity questions in purely (quantum) field theoretic terms, and to tackle them with (quantum) field theoretic tools.

To start with there is a urgent need to clarify the general combinatorial structure of GFT Feynman diagrams, by developing the basic concepts of ordinary quantum field theory in this new and rather peculiar context. What are straightforward questions in QFT, become a bit less straightforward ones in GFTs. For example: what are the 1-particle irreducible diagrams, now that lines of propagation are actually formed by several strands (equivalently, what is the generalization of 1PI to 2-complexes)? and what are their properties? what is the (approximate) form of the 1-loop effective action for the various GFTs? what is the exact combinatorial content of the Dyson-Schwinger equations in this setting?

If these are still rather simple questions (although may involve technical or formal complications), as soon as we try to unravel in more explicit term the combinatorial structure of GFT Feynman diagrams with respect to their dual picture as simplicial complexes, things start to complicate considerably.

Even if every 2-complex arising as a GFT Feynman diagram can be understood as topologically dual to a D-dimensional simplicial complex, this would not be, in general, a simplicial *manifold*. In fact, the data attached to GFT Feynman diagrams, nor the feynmanological rules for their construction, do not constrain the neighbourhoods of simplices of dimensions from (D-3) downwards to be spheres. This implies that in the general case, the resulting simplicial complex, obtained by gluing D-simplices along their (D-1)-faces, would correspond to a *pseudo-manifold*, i.e. to a manifold with *conical singularities* [8]. The issue does not arise in D=2, where all GFT Feynman diagrams (combinatorially the same as those of simple matrix models) are dual to simplicial manifolds, if the orientation condition is satisfied. Neglecting for the moment the issue of whether this is a problem from a physical perspective⁴, or whether on the contrary it is possible to give some physical meaning to these singularities, one remains with the task of analyzing these configurations from a purely mathematical point of view. A precise set of conditions under which the GFT Feynman diagrams correspond to manifolds is identified and discussed at length in [8], both at the level of simplicial complexes and of the corresponding dual 2-complexes, in D=2,3,4. All the relevant conditions can be checked algorithmically on any given Feynman graph. However, we feel it would be good to build upon the analysis of [8], in two main directions: 1) try to identify a suitable reformulation of the found conditions, or an alternative but equivalent set of conditions, that would make the quantum field theoretic interpretation and role of the pseudo-manifold configurations more transparent; 2) if and once this can be done, find how to impose these conditions at the GFT level (a sort of superselection rules?) or construct suitably constrained GFT, that would generate only manifold-like complexes in their Feynman expansion. This last task makes sense, of course, only if the first does not prove that, because of their field theoretic interpretation, pseudo-manifold configurations are in fact needed for consistency at the quantum level (in the same sense, for example, as loop diagrams are in ordinary field theory). It may also turn out that non-manifold-like configurations can not be removed but

⁴As pointed out by one of the participants during the workshop, after all, we do not know whether or not spacetime *really* possesses conical singularities in the microscopic regime....

are instead suppressed in certain sectors of the theory, in specific models, as for example happens in some 3d tensor models [17].

Another wide landscape of interesting questions opens up when considering the topological structure of the GFT Feynman diagrams. As discussed above, the sum over GFT Feynman diagrams includes a sum over all simplicial topologies as well as over all simplicial decompositions of the same topology. In $D=2$ different topologies are weighted by a single topological invariant, the Euler characteristics, and one can then try to express the GFT Feynman amplitudes as a function of this topological parameter, and then identify the sector of the theory or of the parameter space in which, say, the trivial topology dominates. In matrix models, indeed, where the Feynman amplitudes have a simpler form (as discussed above, they correspond to the truncation of the simple GFT 3 to a fixed representation J), one can easily show that non-trivial topologies are suppressed in the limit of infinite matrix dimension (large J). No analogue of this result in the full $D=2$ GFTs is known. In $D \geq 3$ topologies are not classified so the situation is much more intricate. A first guess at how non-trivial topologies enter the GFT Feynman expansion is that they only appear beyond tree level, i.e. in the quantum regime; in fact, in matrix models the expansion in the matrix dimension can also be understood as a loop expansion and confirms the above expectation. This was confirmed also for generic GFTs in [3] by analyzing the Dyson-Schwinger equations: at tree level only diagrams of trivial topology appear, and the order of loops can then be related to the number of handles in the simplicial complex dual to the Feynman diagram. Moreover, by suitable re-scaling of the field, the number of loops could be related to (a power of) the GFT coupling constant, that thus acquires a possible interpretation as the parameter governing topology change; we refer to [3] for more details⁵.

The number of handles is, however, just one of the quantities that can characterize simplicial topology, and a more in dept analysis of the relation between the topological properties of the simplicial complexes appearing in the Feynman expansion and their field theoretic interpretation would be very much welcome. And more generally, one would also like to be able to identify more clearly the dependence of the GFT Feynman amplitudes (for generic dimension, and for non-trivial models) on the topology of the underlying diagram. This would help greatly in the attempt to gain control over the sum over topologies implicit in the GFT perturbative expansion.

For example, just as one can relate handles and quantum loops, is there a QFT interpretation of other known topological invariants or for other topological properties of simplicial complexes?

The model of equation (4), for example, generates Feynman amplitudes that are topological invariants themselves [10], meaning that they evaluate to the same number (after appropriate gauge fixing and regularization) on any element of equivalence class of simplicial complexes related by Pachner moves, while they still provide a different amplitude for topologically inequivalent simplicial complexes. Can one identify the field theoretic feature of this model (and of its higher dimensional equivalents, whose amplitudes share the same property) that can be seen as the origin of this property of its amplitudes? If this can be done, can we use this

⁵Much remains to be understood concerning this correspondence, anyway, as one can easily find high order (in λ) diagrams which still correspond to trivial topology.

new insight to construct new topological invariants of simplicial complexes by field theoretic means?

Even remaining at the level of trivial topology, e.g. say we just consider diagrams with the cylindrical topology $\Sigma \times \mathbb{R}$, there are further structures that one would like to identify in the diagrams and characterize in field-theoretic terms. In the modern version of the dynamical triangulations approach [6], namely *causal dynamical triangulations*, one finds that imposing additional combinatorial restrictions to the simplicial complexes summed over in the definition of the quantum gravity path integral, the continuum properties of the same are drastically improved and one can recover many of the wanted properties of a continuum spacetime from its purely combinatorial quantum definition, e.g. its spectral dimension [6]. These additional restrictions include the presence of a fixed foliation of the triangulations and the absence of branching (baby universe) configurations with respect to this foliation. Given that the GFT partition function can be re-expressed, in perturbative expansion, as a sum over triangulations, weighted by Feynman amplitudes, what is the field theory interpretation of these combinatorial restriction? or, can one identify specific GFT models or general properties of GFT Feynman amplitudes that would lead to a strong suppression, if not the absence, of configurations not satisfying these restrictions?

3.5. Renormalization. We now come to the important issue of renormalization of group field theories, the importance of which we certainly do not need to stress, also given that all known GFT models, at present, are *defined* by their perturbative expansion.

Once more, this is almost completely unexplored territory. The whole question of perturbative renormalization of the various GFT models has not been analyzed in any detail, nor any general scheme for performing the perturbative renormalization of GFTs has been developed. Once more, it is the peculiar combinatorial structure of the GFT Feynman diagrams, and the non-local pairing of variables in the vertex term in the action, that makes the whole issue of divergences and renormalization at the same time very intricate and challenging. For any given Feynman diagram, and after gauge fixing, the sum over geometric data has two potential sources of divergences: depending on the kinetic term (propagator) one chooses, one can have a potential divergence for each loop, i.e. each 2-cell, but also one has a potential divergence for every ‘bubble’ of the GFT Feynman diagram, i.e. for every closed surface of it identified by a collection of 2-cells glued together along common edges. This is, in a sense, the true GFT analogue of loop divergences of usual QFT. In addition to this ‘double-layer’ structure of potential divergences, there are other combinatorial peculiarities of GFTs that make the conventional wisdom (including, for example, simple counting of the degree of divergence of a diagram) less easily applicable, and a brand new scheme of analysis badly needed. One example is the fact that a closer look at the way degrees of freedom are propagated within a Feynman diagram reveals a sort of ‘propagation-only’ structure, in the sense that also within vertices of interaction the field theory degrees of freedom are simply re-routed, i.e. propagated, and there is no coincidence at the same “point” of more than two field variables. This may mean that a regularization of the GFT propagators may suffice to cure loop divergences. However, also this needs a more careful

study. Certainly, due to this combinatorial peculiarities, usual results on the non-renormalizability of ϕ^n theories for high enough n have to be at least reconsidered, to check that they still apply in the GFT framework.

In fact, whether the GFT amplitudes are divergent or not depends on the specific model, even for the same polynomial order in the interaction term. For example, the model of equation (4) turns out to be finite after gauge fixing, while the most natural definition of the group field theory for the Barrett-Crane spin foam model for 4d gravity [9], for example, presents indeed bubble divergences, and the Perez-Rovelli modification of it [9], producing a different version of the same model based on a GFT of the same order and same combinatorics, but different symmetries, possesses again *finite* Feynman amplitudes, i.e. it is *perturbatively finite* without the need for any regularization, even before any gauge fixing.

Once more, one has to develop thus a general theory of perturbative renormalization for GFTs, that would allow to unravel first the combinatorial structure of GFT divergences, and then to regularize them away. On the one hand, conventional regularization and renormalization techniques would be the first thing to try as they look like the most efficient way of performing explicit computations and extract physical results from a field theory; indeed, the development and use of their GFT-adapted analogues would be of great value. On the other hand, the Hopf algebra approach to renormalization [18] have proven to be especially suitable for capturing and elucidating the combinatorial structure of Feynman diagrams and of their divergences; therefore it is a natural guess and hope that one can apply similar techniques to GFTs, where usual tools may be less powerful exactly because of the combinatorial intricacies of the GFT Feynman diagrams (some of which have been highlighted in the previous section).

That the Hopf algebra techniques developed by Kreimer and Connes, among many others, can be the correct language to tackle the issue of GFT renormalization is suggested also by recent work on the renormalization of spin foam models, in particular by the work of [19]. Here the conceptual setup and the perspective on spin foam models were rather different from the one presented in the present article, and group field theories were really not part of the picture. Instead, spin foam models were studied as background independent discrete quantum gravity path integrals (or statistical mechanical state sum models), and the task was to develop a coarse graining and renormalization procedure, a background independent analogue of the usual one used in lattice statistical mechanics and statistical field theory, that would bypass the difficulties coming from the absence of a fixed lattice geometry (e.g. fixed lattice spacing) and could be used for lattices of arbitrary combinatorial structure, as spin foams. The most natural language was found to be, in fact, that of Hopf algebra renormalization and a *Hopf algebra of spin foams* (more precisely, a Hopf algebra of *partitioned* spin foams) was defined, together with a Hopf algebra of (parenthesized) spin foam weights (amplitudes), based on the identification of appropriate *subfoams*. As in the Connes-Kreimer approach, it is the antipode of the two algebras that plays a crucial role in the definition of renormalization group transformations. Indeed, Markopoulou went on defining exact and approximate block transformations on spin foams, based on the discovered Hopf algebra structure, and in particular on its antipode, and suggested that this could be the correct starting point for the analysis of renormalization of known spin foam models.

This approach acquires a new light and, in a sense, a further justification from a group field theory perspective⁶. Recall that spin foams are nothing more than the Feynman diagrams of group field theories, and spin foam models are their Feynman amplitudes. The definition of a renormalization procedure for spin foam models would therefore amount to a definite prescription for perturbative renormalization of group field theories. From this standpoint, then, the fact that the Kreimer-Connes Hopf algebra approach to renormalization, originally developed for perturbative renormalization of QFTs, can be adapted to spin foam models seems only natural. The tasks however are: to identify the Kreimer-Connes algebra of Feynman diagrams for GFTs and compare it with the algebra of spin foams as defined by Markopoulou (as a first step, this involves comparing GFT 1PIs to the subfoam structure proposed in [19]); go on identifying the corresponding Hopf algebra of GFT perturbative renormalization, and again compare it with the one of [19]; clarify the role of GFT gauge symmetries within it; apply it to specific GFT models. Given the large number of interesting mathematical results coming out of recent work on QFT Hopf algebra renormalization, we expect further work along these lines in the GFT framework to be particularly exciting and rewarding, also considering on the one hand the dual simplicial spacetime interpretation of GFT diagrams, and on the other hand the relevance of any result so obtained for quantum gravity⁷.

In the context of renormalization, let us also mention that it would be of great interest to go beyond the perturbative level and develop an exact Wilsonian renormalization group analysis of group field theories. On top of providing important information on the renormalizability of GFTs, it would represent an even more (with respect to the perturbative renormalization of their Feynman diagrams/spin foam models) powerful tool for the study of their continuum limit and relation with classical General Relativity.

Another related issue, concerning the perturbative expansion of GFTs in Feynman diagrams/spin foam models, is that of summability of the perturbative series. Of course, even though, from the interpretation of it as a discretized version of a quantum gravity path integral, one may hope it to be finite, this is not such a reasonable expectation from a purely field theory perspective, which is the one we would like to advocate here. However, summability is a more limited but perfectly reasonable hope for a QFT perturbative series. Notice that such a property would still be rather remarkable, even though only if achieved in a physically meaningful way, from a quantum gravity perspective, given that it amounts to the possibility of gaining (non-perturbative) control over a sum over *all simplicial geometries and all simplicial topologies*. Also this issue has been largely unexplored, both in general terms and for specific GFT models. The only (important) result in this direction obtained so far concerned the model of equation (4) and is the following[20]: a simple modification of the GFT action of equation (4) gives a model, which with a fixed cutoff, is Borel summable relatively to a coupling constant taken sufficiently small, depending on the cutoff. The modification amounts to adding another vertex term of the same order to the original one, to give:

⁶This was indeed noted already in [19]

⁷In fact such an analysis would represent one direct way of tackling the issue of the continuum limit of spin foam models and of their relation with continuum Einstein's gravity.

$$+ \frac{\lambda}{4!} \prod_{i=1}^6 \int_{SU(2)} dg_i \left[\phi(g_1, g_2, g_3) \phi(g_3, g_4, g_5) \phi(g_5, g_2, g_6) \phi(g_6, g_4, g_1) \right. \\ \left. + \delta \phi(g_1, g_2, g_3) \phi(g_3, g_4, g_5) \phi(g_4, g_2, g_6) \phi(g_6, g_5, g_1) \right],$$

with $|\delta| < 1$.

The new term corresponds simply to a slightly different recoupling of the group/representation variables at each vertex of interaction, and geometrically to the only other possible way of gluing 4 triangles to form a closed surface, i.e. to form a “pillow” instead of a regular tetrahedron. In turn this “pillow” configurations are equivalent to two tetrahedra glued together along -two- common triangles, instead of one⁸. Even if the above modification of the Boulatov GFT model has no clear physical interpretation yet from the quantum gravity point of view, e.g. in terms of gravity coupled to some sort of matter, it is indeed a very mild modification, and most importantly one that one would expect to be forced upon us by renormalization group-type of argument, that usually require us to include in the action of our field theory all possible terms that are compatible with the symmetries. What the other terms would be and what their effect on the perturbative series is another interesting open issue, together with the (more physically important) possibility of obtaining a similar result for any of the known GFT models of quantum gravity in $D = 4$.

4. Conclusions

In this article, we have introduced the group field theory formalism, trying to clarify its relevance for quantum gravity research, as well as key aspects and peculiarities of GFT models, defining combinatorially non-local field theories on the one hand, and a generalization of matrix models and field theories of random surfaces on the other. Most importantly, we have presented and discussed many outstanding open issues and unanswered questions that we thought could be of special interest to mathematical physicists and mathematicians working on quantum field theories in general, and in particular those interested in the related perturbative combinatorics and renormalization⁹. Our expectation is that GFTs can represent for mathematicians and mathematical physicists a very valuable playground and toolbox, and an important source of insights and, obviously, amusement, also in light of the immense corpus of mathematical results, tools and applications that has already developed from work on matrix models and random surfaces (and which is by all means still growing). Adding to this the mentioned relevance of all of these open issues for the construction of a satisfactory theory of quantum gravity, the urge to join the current efforts of many theoretical physicists working in this area should be, we hope, unrestrained.....

⁸It would be interesting to understand in more detail the field theoretic interpretation of these configurations.

⁹The selection, of course, has been operated from the particular (and certainly limited) perspective of a theoretical physicist working on quantum gravity, and this has affected also the style of the discussion. Being unavoidable, we do not feel the need to apologize for this.

Acknowledgements

We would like to thank the organizers of the Conference on Combinatorics and Physics for their kind invitation and for having made it a very enjoyable and profitable event.

References

- [1] D. Oriti, in *Approaches to Quantum Gravity: toward a new understanding of space, time and matter*, Cambridge University Press, Cambridge (2008), D. Oriti ed., gr-qc/0607032;
- [2] D. Oriti, in *Quantum Gravity*, B. Fausser, J. Tolksdorf and E. Zeidler, eds., Birkhaeuser, Basel (2007), gr-qc/0512103;
- [3] L. Freidel, Int. J. Theor. Phys. 44, 1769 (2005), hep-th/0505016;
- [4] T. Thiemann, gr-qc/0110034;
- [5] R. Williams, in *Approaches to Quantum Gravity: toward a new understanding of space, time and matter*, Cambridge University Press, Cambridge (2008), D. Oriti ed.
- [6] R. Loll, Lect. Notes Phys. 631, 137-171 (2003), hep-th/0212340; J. Ambjorn, J. Jurkiewicz, R. Loll, in *Approaches to Quantum Gravity: toward a new understanding of space, time and matter*, Cambridge University Press, Cambridge (2008), D. Oriti ed., hep-th/0604212;
- [7] J. Ambjorn, B. Durhuus, T. Jonsson, Mod. Phys. Lett. A 6, 1133-1146 (1991)
- [8] R. De Pietri, C. Petronio, J. Math. Phys. 41, 6671 (2000), gr-qc/0004045;
- [9] D. Oriti, Rept. Prog. Phys. 64, 1489 (2001), gr-qc/0106091; A. Perez, Class. Quant. Grav. 20, R43 (2003), gr-qc/0301113;
- [10] L. Freidel, D. Louapre, Class. Quant. Grav. 21, 5685-5726 (2004), hep-th/0401076
- [11] M. Reisenberger, C. Rovelli, Class. Quant. Grav. 18, 121 (2001), gr-qc/0002095;
- [12] D. Oriti, Phys. Rev. D 73, 061502 (2006), gr-qc/0512069;
- [13] D. Oriti, T. Tlas, *A new class of causal group field theories for 1st order discrete quantum gravity*, to appear;
- [14] W. Fairbairn, E. Livine, gr-qc/0702125
- [15] I. V. Kanatchikov, Rept. Math. Phys. 41, 49-90 (1998), hep-th/9709229; I. V. Kanatchikov, Rept. Math. Phys. 43, 157-170 (1999), hep-th/9810165;
- [16] D. Oriti, J. Ryan, *Hamiltonian analysis of Group Field Theories and the Fock structure of quantum space*, to appear
- [17] M. Gross, Nucl. Phys. Proc. Suppl. 25 A, 144-149 (1992)
- [18] D. Kreimer, Adv. Theor. Math. Phys. 2, 303-334 (1998), q-alg/9707029
- [19] F. Markopoulou, Class. Quant. Grav. 20, 777-800 (2003), gr-qc/0203036
- [20] L. Freidel, D. Louapre, Phys. Rev. D 68, 104004 (2003), hep-th/0211026

E-mail address: d.oriti@phys.uu.nl

INSTITUTE FOR THEORETICAL PHYSICS AND SPINOZA INSTITUTE, UTRECHT UNIVERSITY, LEUVENLAAN 4, 3584 TD Utrecht, THE NETHERLANDS

This page intentionally left blank

Noncommutative formal Taylor expansions and second quantised regularised traces

Sylvie Paycha

ABSTRACT. In this paper which extends results of [P], we first derive formal “non commutative” Taylor expansions:

$$\Phi(x+a) = \sum_{n=0}^{\infty} \sum_{k_1, \dots, k_n=0}^{\infty} \frac{\Phi^{(k_1+\dots+k_n+n)}(x)}{(k_1+1) \cdots (k_1 + \dots + k_n + n)} \frac{\text{ad}_x^{k_1}(a) \cdots \text{ad}_x^{k_n}(a)}{k_1! \cdots k_n!}$$

where $\text{ad}_x(a) = [x, a]$. We then transpose these to a pseudodifferential operator setup at the cost of replacing identities by identities modulo “smoothing operators”. Such formulae can be implemented to define second quantised versions of the usual ζ -regularised traces on classical pseudodifferential operators. In contrast to the latter, second quantised regularised traces are local and give rise to second quantised weighted traces which can be expressed as a residue of a quantised logarithm.

Acknowledgements

I would like to thank Frédéric Patras for enlightening discussions around algebraic issues underlying this paper as well as the referee for useful comments.

Introduction

Locality in the Atiyah-Singer index theorem is very much related to supersymmetry, the supertrace $\text{str}(e^{-t\Delta})$ of a heat operator acting on sections of some \mathbb{Z}_2 -graded bundle boiling down to a local expression.

We argue that the locality in the index theorem in noncommutative geometry of Connes and Moscovici [CM] does not only rely on supersymmetry. The local index theorem in noncommutative geometry not only involves (super-) traces but also linear extensions of these to tensor products of operators, which we refer to as *second quantised (super-) traces*, the ordinary (super-) trace corresponding to the zero-th quantised level. When transposed to the classical pseudodifferential setup, second quantised traces give rise to local terms which can be expressed as linear combinations of noncommutative residues, independently of whether they are super traces or not. At higher quantised level, it is therefore not so much the supersymmetry which plays a role in locality issues as the “quantised” nature of these higher traces which is responsible for their locality via that of the noncommutative residue.

Underlying the construction of these quantised traces are polarised non commutative extensions of ordinary Taylor expansions, which we describe in some detail in the first part of the paper, whereas the second and third part of the paper are dedicated to second quantised traces.

In Section 1, we consider a simple noncommutative setup, that of a free algebra of formal power series in several variables. We extend the usual Taylor formula $\Phi(x + a) = \sum_{n=0}^{\infty} \frac{\Phi^{(n)}(x)}{n!} a^n$ to a noncommutative analog $\Phi(x + a) = \sum_{n=0}^{\infty} \Phi_n(x)(a^{\otimes n})$, where Φ_n is defined by (see (13) in Theorem 1):

$$\Phi_n(x) : a_1 \otimes \cdots \otimes a_n \mapsto \sum_{|k|=0}^{\infty} \frac{\Phi^{(|k|+n)}(x)}{(k_1+1)(k_1+k_2+2) \cdots (|k|+n)} \frac{\text{ad}_x^{k_1}(a_1) \cdots \text{ad}_x^{k_n}(a_n)}{k_1! \cdots k_n!}$$

where we have set $|k| := k_1 + \cdots + k_n$. The maps $\Phi_n(x)$ are interpreted as level n second quantisations of the original map $\Phi(x)$.

They can be seen as polarised non commutative extensions of ordinary Taylor expansions in various respects. When applied to monomials $\Phi(x) = x^K$, such a non-commutative formal Taylor expansion provides noncommutative generalisations of binomial formulae. When applied to exponential maps, it leads to Baker-Campbell-Hausdorff type formulae (see Proposition 2) as well as to Chen integrals reminiscent of Duhamel type formula (see Theorem 2).

Interestingly, similar coefficients to those arising in these noncommutative Taylor expansions occur in the formal expression derived by Connes and Marcolli [**CMA**] for a universal singular frame on principal bundles as well as in the formula for right composition by the Dynkin operator derived by Ebrahimi-Fard, Gracia-Bondia and Patras in [**EGP**].

In Section 2 which is based on [**P**], we work instead on the algebra of classical pseudodifferential operators acting on smooth sections of a finite rank vector bundle over a closed manifold M , replacing the formal variable x by a reference classical elliptic pseudodifferential operator Q and a by a classical pseudo-differential operator A . Similar noncommutative Taylor expansions hold “modulo smoothing operators” which involve second quantisations of ordinary maps (see Theorem 3) on classical pseudodifferential operators.

A second quantisation of complex power maps $\Phi : Q \mapsto Q^z$ leads to second quantised zeta regularised traces using a ζ -function approach which we describe in Section 3. We prove the locality of their finite parts (at some integer p) by expressing them as a linear combination of noncommutative residues (see Theorem 4). As explained in Remark 10, this is very similar to Higson’s approach [**H**] in deriving the residue cocycle of Connes and Moscovici [**CM**].

Unlike ordinary regularised traces, their quantised counterparts obtained by extracting finite parts at $p = 0$ of the above meromorphic extensions, are local as finite linear combinations of noncommutative residues (see Theorem 4) and can actually be written as a residue of a quantised logarithm (see Theorem 5). Whereas ordinary weighted traces can be written as residues of logarithms only in very special cases such as the case of differential operators (cfr. (39)), by contrast quantised weighted traces on tensor products of *any* classical pseudodifferential operators can be written as residues of quantised logarithms (cfr. (40)).

1. Noncommutative Taylor expansions on formal power series

1.1. From a commutative to a noncommutative setup. We first give a general idea how ordinary Taylor expansions in the commutative setup can be generalised to a noncommutative one before going into the technicalities in the next paragraph.

We follow the notations and presentation of [MP]. Let $\mathcal{A}_{\text{fin}} := \mathbb{C}[x]$ be the algebra of complex polynomials in the variable x . Seen as the space of continuous complex valued functions on \mathbb{N}_0 with compact support, its topological dual is the commutative algebra $\mathcal{A} = \mathbb{C}[x]$ of formal power series in the variable x whose elements we write:

$$a = \sum_{n=0}^{\infty} \alpha_n x^n, \quad \alpha_n \in \mathbb{C},$$

equipped with the weakest topology which makes the projections $a \mapsto \alpha_n$ continuous. \mathcal{A} is a Fréchet space with topology given by a family of semi-norms $|a|_n = \sup_{k \leq n} |\alpha_k|$ indexed by non negative integers n .

Given an analytic function $\phi(z) = \phi_0 + \phi_1 z + \phi_2 z^2 + \dots$ at 0, we have $\phi_n = \frac{n!}{2i\pi} \int_C z^{-1-n} \phi(z) dz$ along some contour C around $z = 0$. Since¹ $(\lambda - x)^{-1} = \sum_{n=0}^{\infty} \lambda^{-(n+1)} x^n$ for any $\lambda \in \mathbb{C}^*$, it follows that the formal power series $\Phi(x) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \phi_n x^n$ can be written in terms of a Cauchy integral in \mathcal{A} :

$$(1) \quad \Phi(x) = \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi(\lambda) d\lambda := \sum_{n=0}^{\infty} \left(\int_C \lambda^{-(n+1)} \phi(\lambda) d\lambda \right) x^n$$

defined in terms of the complex valued Cauchy integral $\int_C \lambda^{-(n+1)} \phi(\lambda) d\lambda$. Applying this to $\phi^{(k)}$ yields:

$$(2) \quad \Phi^{(k)}(x) = \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi^{(k)}(\lambda) d\lambda := \sum_{n=0}^{\infty} \left(\int_C \lambda^{-(n+1)} \phi^{(k)}(\lambda) d\lambda \right) x^n.$$

The convergence of the Cauchy integrals in \mathcal{A} is ensured by the convergence of the corresponding Cauchy integrals $\int_C \lambda^{-(n+1)} \phi(\lambda) d\lambda$ and $\int_C \lambda^{-(n+1)} \phi^{(k)}(\lambda) d\lambda$ in \mathbb{C} .

Repeated integration by parts in the complex valued Cauchy formula combined with the fact that $(\lambda - x)^{-(k+1)} = \sum_{n=0}^{\infty} \frac{(k+n)!}{k! n!} \lambda^{-(k+n+1)} x^n$ for any $\lambda \in \mathbb{C}^*$ leads to an integration by parts formula in \mathcal{A} :

$$\begin{aligned} \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi^{(k)}(\lambda) d\lambda &= \sum_{n=0}^{\infty} \left(\frac{1}{2i\pi} \int_C \lambda^{-(n+1)} \phi^{(k)}(\lambda) d\lambda \right) x^n \\ &= \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} \left(\frac{1}{2i\pi} \int_C \lambda^{-1-n-k} \phi(\lambda) d\lambda \right) x^n \\ &=: k! \frac{1}{2i\pi} \int_C (\lambda - x)^{-(k+1)} \phi(\lambda) d\lambda, \end{aligned}$$

¹The generator x of the formal power series algebra has its spectrum concentrated at 0. The variable x plays a dual role here, as a generator of an algebra and as argument of a function; this is reminiscent of the role played by the operator D in noncommutative geometry, which is both the generator of the algebra of abstract pseudodifferential operators and the variable in the functional calculus.

with the last Cauchy integral in \mathcal{A} defined by $\frac{1}{2i\pi} \int_C \lambda^{-(k+n+1)} \phi(\lambda) d\lambda$ in \mathbb{C} via the last identity, so that we can write:

$$(3) \quad \frac{\Phi^{(k)}(x)}{k!} = \frac{1}{2i\pi} \int_C (\lambda - x)^{-(k+1)} \phi(\lambda) d\lambda.$$

We now perturb x by translation $x \mapsto x + a$ for some given element a in the polynomial algebra $\mathcal{A}_{\text{fin}} \subset \mathcal{A}$. The perturbed resolvent is a formal power series in \mathcal{A} :

$$(\lambda - (x + a))^{-1} = \sum_{k=0}^{\infty} (\lambda - x)^{-(k+1)} a^k = \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \lambda^{-(k+n+1)} \frac{(k+n)!}{n! k!} x^n a^k$$

since only a finite number of terms of the infinite sums contribute to the coefficient of a given product $x_{k_1} \cdots x_{k_i}$ of formal variables in X .

With the same notations as above, we can define a Cauchy integral in \mathcal{A} involving the perturbed resolvent, in terms of a formal power series:

$$\begin{aligned} \frac{1}{2i\pi} \int_C (\lambda - (x + a))^{-1} \phi(\lambda) d\lambda &:= \frac{1}{2i\pi} \sum_{k=0}^{\infty} \left(\int_C (\lambda - x)^{-(k+1)} \phi(\lambda) d\lambda \right) a^k \\ &= \frac{1}{2i\pi} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{k!} \left(\frac{1}{2i\pi} \int_C \lambda^{-(n+1)} \phi^{(k)}(\lambda) d\lambda \right) x^n a^k \end{aligned}$$

involving Cauchy integrals $\frac{1}{2i\pi} \int_C \lambda^{-(n+1)} \phi^{(k)}(\lambda) d\lambda$ which converge in \mathbb{C} .

From (3) we infer the following Taylor expansion² at x :

$$\begin{aligned} \Phi(x + a) &:= \frac{1}{2i\pi} \int_C (\lambda - (x + a))^{-1} \phi(\lambda) d\lambda \\ (4) \quad &= \phi(x) + \sum_{k=1}^{\infty} \frac{\Phi^{(k)}(x)}{k!} a^k \quad \forall a \in \mathcal{A}_{\text{fin}}. \end{aligned}$$

If Φ is a monomial $x \mapsto x^K$, the expansion is finite and (4) yields back the well-known binomial formulae $(x + a)^K = \sum_{k=0}^K C_K^k x^{K-k} a^k$.

In these notes we are concerned with a noncommutative generalisation of this setup in the spirit of [CM]; the commutative algebra of formal power series in one variable is replaced by the noncommutative free algebra \mathcal{A} of formal power series in several variables chosen in a finite set $X = \{x_1, \dots, x_L\}$, i.e. elements of the form:

$$a = \alpha_0 + \alpha_k x_k + \alpha_{k_1 k_2} x_{k_1} x_{k_2} + \cdots, \quad \alpha_{k_1 \dots k_i} \in \mathbb{C},$$

equipped with the weakest topology which makes the projections $a \mapsto \alpha_{k_1 \dots k_i}$ continuous. Let \mathcal{A}_{fin} denote the set of formal power series with all but a finite number of non vanishing terms.

Clearly, the case $L = 1$ gives back the above commutative setup.

As before, an analytic function $\phi(z) = \phi_0 + \phi_1 z + \phi_2 z^2 + \cdots$ at 0 induces a formal power series $\Phi(x) = \phi_0 + \phi_1 x + \phi_2 x^2 + \cdots$ in any variable x in X which can be written³ in terms of a Cauchy integral $\Phi(x) = \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi(\lambda) d\lambda$ along

²We assume that the contour C can be “enlarged” so as to circle around the spectrum of $a + x$.

³As before, each generator x of the formal power series algebra has its spectrum concentrated at 0.

some contour C around $z = 0$ and similarly for the n -th derivative, $\Phi^{(n)}(x) = \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi^{(n)}(\lambda) d\lambda$.

However, due to the noncommutativity, perturbations $x \mapsto x + a$ for some fixed $a \in \mathcal{A}_{\text{fin}}$ lead to a more complicated formula for the Taylor expansion. These now involve iterated commutators $\text{ad}_x^k(a) := [x, [\dots, [x, a]]]$ (with k left brackets), k in \mathbb{N} and where we have set $[a, b] = ab - ba$.

We first introduce maps $\Phi_n(x) : \mathcal{A}_{\text{fin}}^{\otimes n} \rightarrow \mathcal{A}$ (see (13) in Theorem 1) defined in terms of derivatives $\Phi^{(l)}(x)$ introduced in (2)

$$(5) \quad \Phi_n(x)(a_1 \otimes \dots \otimes a_n) := \sum_{|k|=0}^{\infty} \frac{\Phi^{(|k|+n)}(x)}{(k_1 + 1)(k_1 + k_2 + 2) \dots (k_1 + \dots + k_n + n)} \frac{\text{ad}_x^{k_1}(a_1) \dots \text{ad}_x^{k_n}(a_n)}{k_1! \dots k_n!},$$

where we have set $|k| = k_1 + \dots + k_n$. Such an expression yields a formal power series in \mathcal{A} for any x in X ; indeed, since the derivatives $\Phi^{(|k|+n)}(x) := \sum_{p=0}^{\infty} \phi_p^{(|k|+n)} x^p$ are defined by formal power series in x , for given a_1, \dots, a_n in \mathcal{A}_{fin} , only a finite number of terms in the resulting infinite sums:

$$\Phi_n(x)(a_1 \otimes \dots \otimes a_n) := \sum_{|k|=0}^{\infty} \sum_{p=0}^{\infty} \frac{\phi_p^{(|k|+n)}}{(k_1 + 1)(k_1 + k_2 + 2) \dots (k_1 + \dots + k_n + n)} \frac{x^p \text{ad}_x^{k_1}(a_1) \dots \text{ad}_x^{k_n}(a_n)}{k_1! \dots k_n!},$$

actually contribute to the coefficient of a given product $x_{j_1} \dots x_{j_i}$ of variables in X . If $L = 1$, only the term $|k| = 0$ remains so that

$$\Phi_n(x)(a_1 \otimes \dots \otimes a_n) = \frac{\Phi^{(n)}(x)}{n!} a_1 \dots a_n,$$

which is a formal power series in x .

With these notations, we show that (see Corollary 1)

$$(6) \quad \Phi(x + a) = \phi(x) + \sum_{n=1}^{\infty} \Phi_n(x)(a^{\otimes n}),$$

which yields back (4) for $a_1 = \dots = a_n = a$, so that formula (6) can be seen as a “noncommutative” Taylor expansion.

When applied to monomials $\Phi : x \mapsto x^K$, the expansion (5) only has a finite number of non vanishing terms and gives rise to

$$(x + a)^K = \sum_{|k|+n=0}^K \frac{\text{ad}_x^{k_1}(a) \dots \text{ad}_x^{k_n}(a)}{(k_1 + 1)(k_1 + k_2 + 2) \dots (k_1 + \dots + k_n + n)} \frac{K!}{(K - (|k| + n))!} \frac{x^{K - (|k| + n)}}{k_1! \dots k_n!}$$

where $|k| := k_1 + \dots + k_n$ and for any a in \mathcal{A}_{fin} .

If $L = 1$, only the term $|k| = 0$ survives and we get back the usual binomial formula; thus (??) can be seen as a noncommutative version of the usual binomial formulae.

Applying (6) to the exponential map $\phi(z) = e^z$ gives rise to the following Baker-Campbell-Hausdorff type formula (see formula (15)):

$$e^{x+a} e^{-x} = \sum_{n=0}^N \sum_{|k|=0}^{\infty} \frac{\text{ad}_x^{k_1}(a) \cdots \text{ad}_x^{k_n}(a)}{(k_1+1)(k_1+k_2+2) \cdots (k_1+k_2+\cdots+k_n+n) k_1! \cdots k_n!}$$

where we have set $e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}$ and taken a in \mathcal{A}_{fin} .

1.2. Second quantised resolvents. Let us now make these statements precise. As before, \mathcal{A} is the free algebra of formal power series generated by a finite subset X , \mathcal{A}_{fin} the set of those formal power series which have all but a finite number of non vanishing terms. \mathcal{A} is noncommutative unless X reduces to one single element.

A perturbation of the resolvent $x \mapsto r(\lambda, x) = (\lambda - x)^{-1}$ (as before x is chosen in X) by a translation $x \mapsto x + a$ with a in \mathcal{A}_{fin} reads:

$$\begin{aligned} r(\lambda, x + a) &= (\lambda - (x + a))^{-1} \\ (7) \quad &= (\lambda - x)^{-1} + (\lambda - x)^{-1} a (\lambda - (x + a))^{-1} \\ &= \sum_{n=0}^{\infty} (\lambda - x)^{-1} a (\lambda - x)^{-1} a (\lambda - x)^{-1} \cdots (\lambda - x)^{-1} a (\lambda - x)^{-1} \\ &\quad \text{with } a \text{ occurring } n \text{ times in the } n\text{-th term.} \end{aligned}$$

This perturbed resolvent is a formal power series in \mathcal{A} ; indeed,

$$(8) \quad r(\lambda, x + a) = \sum_{n=0}^{\infty} \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \lambda^{-(|k|+n)} x^{k_1} a x^{k_2} a \cdots x^{k_n-1} a x^{k_n},$$

where we have set $|k| = k_1 + \cdots + k_n$, so that only a finite number of terms contribute to the coefficient of a given product $x_{j_1} x_{j_2} \cdots x_{j_i}$ of formal variables in X .

Let $\mathcal{T}(\mathcal{A}_{\text{fin}}) := \bigoplus_{n=0}^{\infty} \mathcal{T}_n(\mathcal{A}_{\text{fin}})$ be the tensor algebra generated by \mathcal{A}_{fin} , where we have set $\mathcal{T}_n(\mathcal{A}_{\text{fin}}) := \otimes^n \mathcal{A}_{\text{fin}}$.

Motivated by (7) we set the following definitions.

DEFINITION 1. *The second quantised resolvent $r_{\bullet}(\lambda, x) : \mathcal{T}(\mathcal{A}_{\text{fin}}) \rightarrow \mathcal{A}$ at a point x in X is defined on $\mathcal{T}_n(\mathcal{A}_{\text{fin}})$ by*

$$\begin{aligned} r_n(\lambda, x) : \quad \mathcal{T}_n(\mathcal{A}_{\text{fin}}) &\rightarrow \mathcal{A} \\ a_1 \otimes \cdots \otimes a_n &\mapsto r(\lambda, x) a_1 r(\lambda, x) a_2 \cdots a_{n-1} r(\lambda, x) a_n r(\lambda, x). \end{aligned}$$

REMARK 1. *For any formal variable x in X , the expression*

$$r(\lambda, x) a_1 r(\lambda, x) a_2 \cdots a_{n-1} r(\lambda, x) a_n r(\lambda, x)$$

is a formal power series in \mathcal{A} for similar reasons for which $r(\lambda, x + a)$ was, for any a in \mathcal{A}_{fin} , as can be seen replacing the j -th occurrence of a in its formal series expansion by a_j in (8).

With these notations we have:

$$(9) \quad r(\lambda, x + a) = r(\lambda, x) + \sum_{n=1}^{\infty} r_n(\lambda, x) (a^{\otimes n}).$$

Let us state a useful technical lemma which we take from [H] (see Lemma 4.20).

LEMMA 1. *Let λ be a non zero complex number and h a positive integer. The bracket $[(\lambda - x)^{-h}, a]$ of the h -th power of the resolvent with any element a in \mathcal{A}_{fin} reads:*

$$(10) \quad [(\lambda - x)^{-h}, a] = \sum_{k=1}^{\infty} \frac{(h+k-1)!}{(h-1)! k!} \text{ad}_x^k(a) (\lambda - x)^{-h-k}.$$

Proof: We first observe that the result holds for $h = 1$, for multiplying the identity $a(\lambda - x) = (\lambda - x)a + [x, a]$ on either side by the resolvent $(\lambda - x)^{-1}$ and iterating this procedure with x replaced by iterated brackets $\text{ad}_x^k(a)$, $k \in \mathbb{N}$, we get

$$(11) \quad \begin{aligned} [(\lambda - x)^{-1}, a] &= (\lambda - x)^{-1} [x, a] (\lambda - x)^{-1} \\ &= \text{ad}_x(a) (\lambda - x)^{-2} + (\lambda - x)^{-1} \text{ad}_x(a) (\lambda - x)^{-2} \\ &= \sum_{k=1}^{\infty} \text{ad}_x^k(a) (\lambda - x)^{-(k+1)}. \end{aligned}$$

The expression $\sum_{k=1}^{\infty} \text{ad}_x^k(a) (\lambda - x)^{-(k+1)}$ defines a formal power series in \mathcal{A} since

$$\sum_{k=1}^{\infty} \text{ad}_x^k(a) (\lambda - x)^{-(k+1)} = \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \frac{(k+n)!}{n! k!} \lambda^{-(k+n+1)} \text{ad}_x^k(a) x^n,$$

so that only a finite number of terms contribute to the coefficient of any given product $x_{j_1} \cdots x_{j_i}$ of formal variables in X .

Differentiating⁴ $h - 1$ times identity (11) with respect to λ yields (10). \square

PROPOSITION 1. *The second quantised resolvent applied to the tensor product of any a_1, \dots, a_n in \mathcal{A}_{fin} reads:*

$$\begin{aligned} r_n(\lambda, x)(a_1 \otimes \cdots \otimes a_n) &= \sum_{|k|=0}^{\infty} \frac{(k_1 + \cdots + k_n + n - 1)!}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 1) \cdots (k_1 + \cdots + k_{n-1} + n - 1)} \\ &\quad \times a_1^{(k_1)} a_2^{(k_2)} \cdots a_n^{(k_n)} (\lambda - x)^{-|k|-n-1}, \end{aligned}$$

where we have set $|k| = k_1 + \cdots + k_n$.

⁴I thank the referee for pointing this short proof out to me.

Proof: Applying Lemma 1 to $h_1 = 1, h_2 = 2+k_1, \dots, h_i = i+k_1+k_2+\dots+k_{i-1}$, we have

$$\begin{aligned}
& r_n(\lambda, x)(a_1 \otimes \dots \otimes a_n) \\
&= (\lambda - x)^{-1} a_1 (\lambda - x)^{-1} \dots (\lambda - x)^{-1} a_n (\lambda - x)^{-1} \\
&= \sum_{k_1=0}^{\infty} a_1^{(k_1)} (\lambda - x)^{-2-k_1} a_2 \dots (\lambda - x)^{-1} a_n (\lambda - x)^{-1} \\
&= \sum_{k_1=0}^{\infty} a_1^{(k_1)} \sum_{k_2=0}^{\infty} \frac{(k_1+k_2+1)!}{(k_1+1)!k_2!} a_2^{(k_2)} (\lambda - x)^{-3-k_1-k_2} \dots (\lambda - x)^{-1} a_n (\lambda - x)^{-1} \\
&= \sum_{k_1=0}^{\infty} a_1^{(k_1)} \sum_{k_2=0}^{\infty} \frac{(k_1+k_2+1)!}{(k_1+1)!k_2!} a_2^{(k_2)} \sum_{k_3=0}^{\infty} \frac{(k_1+k_2+k_3+2)!}{(k_1+k_2+2)!k_3!} \\
&\quad \cdot a_3^{(k_3)} (\lambda - x)^{-4-k_1-k_2-k_3} a_4 \dots (\lambda - x)^{-1} a_n (\lambda - x)^{-1} \\
&= \sum_{|k|=0}^{\infty} \frac{(k_1+k_2+1)! \dots (k_1+k_2+\dots+k_n+n-1)!}{k_2! \dots k_n! (k_1+1)! (k_1+k_2+2)! \dots (k_1+\dots+k_{n-1}+n-1)!} \\
&\quad \cdot a_1^{(k_1)} a_2^{(k_2)} \dots a_n^{(k_n)} (\lambda - x)^{-|k|-n-1} \\
&= \sum_{|k|=0}^{\infty} \frac{(k_1+\dots+k_n+n-1)!}{k_1! \dots k_n! (k_1+1) (k_1+k_2+1) \dots (k_1+\dots+k_{n-1}+n-1)} \\
&\quad \cdot a_1^{(k_1)} a_2^{(k_2)} \dots a_n^{(k_n)} (\lambda - x)^{-|k|-n-1}.
\end{aligned}$$

1.3. Second quantised functionals. Second quantised functionals are then defined via Cauchy integrals. As in the commutative case (see Section 1.1), for any formal variable x in X and any analytic function $\phi(z) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}(0)}{n!} x^n$ at $z = 0$, the formal power series $\Phi(x) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}(0)}{n!} x^n$ in any formal variable x in X can be written in terms of a Cauchy integral $\Phi(x) = \frac{1}{2i\pi} \int_C (\lambda - x)^{-1} \phi(\lambda) d\lambda$ in \mathcal{A} and its p -th derivative reads:

$$(12) \quad \Phi^{(p)}(x) = \frac{p!}{2i\pi} \int_C (\lambda - x)^{-p-1} \phi(\lambda) d\lambda.$$

Combining (12) with Proposition 1 we can set the following definition which leads to a noncommutative Taylor type formula.

THEOREM 1. *To an analytic function $\phi(z)$ at $z = 0$ and to an element x in X we assign a map $\Phi_{\bullet}(x) : \mathcal{T}(\mathcal{A}_{\text{fin}}) \rightarrow \mathcal{A}$, called second quantisation of $\Phi(x)$ defined on $\mathcal{T}_n(\mathcal{A}_{\text{fin}})$ by*

$$\begin{aligned}
\Phi_n(x) : \quad \mathcal{T}_n(\mathcal{A}_{\text{fin}}) &\rightarrow \mathcal{A} \\
a_1 \otimes \dots \otimes a_n &\mapsto \frac{1}{2i\pi} \int_C r_n(\lambda, x)(a_1 \otimes \dots \otimes a_n) \phi(\lambda) d\lambda
\end{aligned}$$

where we have set:

$$\begin{aligned}
& \frac{1}{2i\pi} \int_C r_n(\lambda, x)(a_1 \otimes \cdots \otimes a_n) \phi(\lambda) d\lambda \\
:= & \sum_{|k|=0}^{\infty} \frac{(k_1 + \cdots + k_n + n - 1)!}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 1) \cdots (k_1 + \cdots + k_{n-1} + n - 1)} \\
& \cdot a_1^{(k_1)} a_2^{(k_2)} \cdots a_n^{(k_n)} \left(\frac{1}{2i\pi} \int_C (\lambda - x)^{-|k|-n-1} \phi(\lambda) d\lambda \right),
\end{aligned}$$

and the following noncommutative Taylor type formula holds:

$$(13) \quad \Phi_n(x)(a_1 \otimes \cdots \otimes a_n) = \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1 + 1) \cdots (k_1 + \cdots + k_n + n)} \Phi^{(|k|+n)}(x)$$

where we have set $\delta_k(a) := \frac{\text{ad}_x^k(a)}{k!}$.

Proof: The map Φ_n is well defined since (13) indeed defines a formal power series in \mathcal{A} and formula (13) easily follows from (12) combined with Proposition 1:

$$\begin{aligned}
& \Phi_n(x)(a_1 \otimes \cdots \otimes a_n) = \\
= & \sum_{|k|=0}^{\infty} \frac{(k_1 + \cdots + k_n + n - 1)!}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_{n-1} + n - 1)} \\
& \cdot a_1^{(k_1)} a_2^{(k_2)} \cdots a_n^{(k_n)} \frac{1}{2i\pi} \int_C \phi(\lambda)(\lambda - x)^{-|k|-n-1} d\lambda \\
= & \sum_{|k|=0}^{\infty} \frac{1}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_{n-1} + n)} \\
& \cdot a_1^{(k_1)} a_2^{(k_2)} \cdots a_n^{(k_n)} \frac{1}{2i\pi} \int_C \phi^{(|k|+n)}(\lambda)(\lambda - x)^{-1} d\lambda \\
= & \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_n + n)} \Phi^{(|k|+n)}(x).
\end{aligned}$$

□

COROLLARY 1. Let $\Phi(x)$ be defined from an analytic function ϕ as in (1) for some $x \in X$. Then for a in \mathcal{A}_{fin}

$$(14) \quad \Phi(x + a) = \Phi(x) + \sum_{n=1}^{\infty} \Phi_n(x)(a^{\otimes n}).$$

Proof: Using (9) we can write:

$$\begin{aligned}
\Phi(x + a) &= \frac{1}{2i\pi} \int_C \phi(\lambda)(\lambda - (x + a))^{-1} d\lambda \\
&= \frac{1}{2i\pi} \int_C \phi(\lambda) r(\lambda, x + a) d\lambda \\
&= \frac{1}{2i\pi} \int_C \phi(\lambda) r(\lambda, x) d\lambda + \sum_{n=1}^{\infty} \frac{1}{2i\pi} \int_C \phi(\lambda) r_n(\lambda, x)(a^{\otimes n}) d\lambda \\
&= \Phi(x) + \sum_{n=1}^{\infty} \Phi_n(x)(a^{\otimes n}).
\end{aligned}$$

□

1.4. The second quantised exponential map. Applying the above constructions to the analytic function $\exp : z \mapsto e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$ we define for $x \in X$ maps $\text{Exp}_n(x) : \mathcal{T}(\mathcal{A}_{\text{fin}}) \rightarrow \mathcal{A}$ by

$$\text{Exp}_n(x)(a_1 \otimes \cdots \otimes a_n) = \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1+1)(k_1+k_2+2) \cdots (k_1 + \cdots + k_n + n)} e^x,$$

where as before we have set $\delta_k(a) = \frac{\text{ad}_x^k(a)}{k!}$.

Theorem 1 leads to a Baker-Campbell-Hausdorff type formula:

PROPOSITION 2. *For $x \in X$ and $a \in \mathcal{A}_{\text{fin}}$,*

$$(15) \quad e^{x+a} e^{-x} = 1 + \sum_{n=1}^{\infty} \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a) \cdots \delta_{k_n}(a)}{(k_1+1)(k_1+k_2+2) \cdots (k_1 + \cdots + k_n + n)}.$$

Proof: This follows from (14) applied to $\Phi(x) = e^x$. □

We further extend the maps

$$\begin{aligned} \gamma_t(x) : \mathcal{A} &\rightarrow \mathcal{A} \\ a &\mapsto e^{-t x} a e^{t x} \end{aligned}$$

parametrised by $t \in \mathbb{R}$ to maps $\Gamma_t(x) : \mathcal{T}(\mathcal{A}) \rightarrow \mathcal{A}$ on the tensor algebra given on $\mathcal{T}_n(\mathcal{A})$ by

$$\begin{aligned} \Gamma_{n,t}(x) : \mathcal{T}_n(\mathcal{A}_{\text{fin}}) &\rightarrow \mathcal{A} \\ a_1 \otimes \cdots \otimes a_n &\mapsto \text{Exp}_n(-t x)(a_1 \otimes \cdots \otimes a_n) e^{t x} \end{aligned}$$

LEMMA 2. *For any fixed $x \in X$ and any $a \in \mathcal{A}_{\text{fin}}$*

$$\gamma_t(x)(a) = a + \sum_{k=1}^{\infty} (-t)^k \delta_k(a) = \sum_{k=0}^{\infty} (-t)^k \delta_k(a),$$

with $\delta_0(a) := a$.

Proof: This follows from $\Phi(x)(a) = \sum_{k=0}^{\infty} \frac{\text{ad}_x^k(a)}{k!} \Phi^{(k)}(x)$ applied to $\Phi_t(x) = e^{-tx}$. □

Iterating this lemma yields the following description of $\Gamma_{n,t}(x)$ which involves Chen integrals.

THEOREM 2. *For any fixed $x \in X$, any $a_1, \dots, a_n \in \mathcal{A}_{\text{fin}}$ and any positive t*

$$\begin{aligned} &\Gamma_{n,t}(x)(a_1 \otimes \cdots \otimes a_n) \\ &= \int_{0 \leq u_1 \leq \cdots \leq u_n \leq t} \gamma_{u_1}(a_1) \gamma_{u_2}(a_2) \cdots \gamma_{u_n}(a_n) du_1 \cdots du_n \\ &= \int_{t_1 + \cdots + t_{n+1} = t} e^{-t_1 x} a_1 e^{-t_2 x} a_2 \cdots e^{-t_n x} a_n e^{-(t_{n+1} - t)x} dt_1 \cdots dt_n \\ &= (-1)^n \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1+1)(k_1+k_2+2) \cdots (k_1 + \cdots + k_n + n)} (-t)^{|k|+n}. \end{aligned}$$

Proof: Setting $u_1 = t_1, u_2 = t_1 + t_2, \dots, u_n = t_1 + \dots + t_n$ and iterating Lemma 2 we write

$$\begin{aligned}
& \int_{t_1+\dots+t_{n+1}=t} e^{-t_1 x} a_1 e^{-t_2 x} a_2 \cdots e^{-t_n x} a_n e^{-t_{n+1} x - t x} dt_1 \cdots dt_n \\
&= \int_{t_1+\dots+t_{n+1}=t} \gamma_{t_1}(a_1) \gamma_{t_1+t_2}(a_2) \cdots \gamma_{t_1+\dots+t_n}(a_n) dt_1 \cdots dt_n \\
&= \int_{0 \leq u_1 \leq \dots \leq u_n \leq t} \gamma_{u_1}(a_1) \gamma_{u_2}(a_2) \cdots \gamma_{u_n}(a_n) du_1 \cdots du_n \\
&= \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \left(\int_{0 \leq u_1 \leq \dots \leq u_n \leq t} (-u_1)^{k_1} \cdots (-u_n)^{k_n} du_1 \cdots du_n \right) \alpha_{k_1}(a_1) \cdots \alpha_{k_n}(a_n) \\
&= \sum_{|k|=0}^{\infty} (-1)^{|k|} \left(\int_{0 \leq u_1 \leq \dots \leq u_n \leq t} u_1^{k_1} \cdots u_n^{k_n} du_1 \cdots du_n \right) \delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n) \\
&= (-1)^n \sum_{|k|=0}^{\infty} (-t)^{|k|+n} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1+1)(k_1+k_2+2) \cdots (k_1+\dots+k_n+n)} \\
&= (-1)^n \sum_{|k|=0}^{\infty} \frac{\delta_{k_1}(a_1) \cdots \delta_{k_n}(a_n)}{(k_1+1)(k_1+k_2+2) \cdots (k_1+\dots+k_n+n)} \partial_x^{|k|+n} (e^{-t x}) e^{t x} \\
&= (-1)^n \exp_n(-t x) (a_1 \otimes \cdots \otimes a_n) e^{t x} \\
&= \Gamma_{n,t}(x) (a_1 \otimes \cdots \otimes a_n).
\end{aligned}$$

□

2. Quantised functionals on classical pseudodifferential operators

Noncommutative formal Taylor expansions on the algebra of formal power series can be generalised to the algebra of classical pseudodifferential operators on a closed manifold at the cost of replacing identities by identities “modulo smoothing operators”. These lead to quantised regularised traces which we then express in terms of noncommutative residues of classical pseudodifferential operators.

2.1. The noncommutative residue and the canonical trace. We first briefly recall some basic notions concerning symbols and pseudodifferential operators and fix the corresponding notations. Classical references for the polyhomogeneous symbol calculus are e.g. [Ho], [Se2], [Sh], [Ta].

Given an open subset U of \mathbb{R}^n and an auxiliary (finite-dimensional) normed vector space V , symbols on U with coefficients in $\text{End}(V)$ consist of those functions $\sigma(x, \xi)$ in $C^\infty(T^*U, \text{End}(V))$ such that $D_x^\mu D_\xi^\nu \sigma(x, \xi) = O((1 + |\xi|)^{r-|\nu|})$ for some real number r and all multi-indices μ, ν , uniformly in ξ , and, on compact subsets of U , uniformly in x . A smoothing symbol is one which fulfills this condition for any real number r . A *classical* (polyhomogeneous) symbol of complex order a means a function $\sigma(x, \xi)$ in $C^\infty(T^*U, \text{End}(V))$ such that for each positive integer N and each integer $0 \leq j \leq N-1$ there exist

- σ_{a-j} in $C^\infty(T^*U, \text{End}(V))$ which is positively homogeneous in ξ of degree $a-j$ for $|\xi| \geq 1$, so $\sigma_{a-j}(x, t\xi) = t^{a-j} \sigma_{a-j}(x, \xi)$ for $t \geq 1, |\xi| \geq 1$,
- and a symbol $\sigma_{(N)}$ in $C^\infty(T^*U, \text{End}(V))$ for which the real number r can be chosen equal to $\text{Re}(a) - N$,

such that

$$(16) \quad \sigma(x, \xi) = \sum_{j=0}^{N-1} \sigma_{\alpha-j}(x, \xi) + \sigma_{(N)}(x, \xi) \quad \forall (x, \xi) \in T^*U.$$

We then write $\sigma(x, \xi) \sim \sum_{j=0}^{\infty} \sigma_{\alpha-j}(x, \xi)$. Let $CS^a(U, V)$ denote the subset of classical symbols of order a . The star product of two classical symbols $\sigma_i \in CS^{a_i}(U, V)$, $i = 1, 2$

$$\sigma_1 \star \sigma_2(x, \xi) \sim \sum_{\alpha \in \mathbb{N}^n} \frac{(-i)^{|\alpha|}}{\alpha!} \partial_\xi^\alpha \sigma_1(x, \xi) \partial_x^\alpha \sigma_2(x, \xi)$$

where \sim stands for “identity modulo smoothing symbols”, does not generally lie in $CS^{a_1+a_2}(U, V)$; however it does if $a_1 - a_2$ is an integer. Let $CS(U, V)$ denote the algebra generated by all classical symbols on U with coefficients in $\text{End}(V)$.

The local residue density only involves the $-n$ homogeneous component of the symbol.

DEFINITION 2. *Given an open subset $U \subset \mathbb{R}^n$ and a point $x \in U$, the local noncommutative residue is defined for σ in $CS^a(U, V)$ by*

$$\text{res}_x(\sigma) = \int_{S_x^* U} \text{tr}(\sigma_{-n}(x, \xi)) \, d_S \xi,$$

where $d_S \xi := (2\pi)^{-n} d_S \xi$ with $d_S \xi$ the sphere measure on $S_x^* U$, the unit sphere in the cotangent space $T_x^* U$ to U at point x and where tr stands for the matrix trace.

In contrast, the following extension of the ordinary Lebesgue integral involves the whole symbol.

DEFINITION 3. *The finite-part integral of $\sigma(x, \cdot)$ with σ in $CS^\alpha(U, V)$ and x in U is defined as the constant term in the asymptotic expansion of $\int_{B_x^*(0, R)} \sigma(x, \xi) \, d\xi$ when R tends to ∞ :*

$$(17) \quad \int_{T_x^* U} \sigma(x, \xi) \, d\xi := \text{fp}_{R \rightarrow \infty} \int_{B_x^*(0, R)} \text{tr}(\sigma(x, \xi)) \, d\xi,$$

where as before, tr stands for the matrix trace. Here $B_x^*(0, R)$ is the cotangent ball centered at 0 of radius R in $T_x^* U$ and $d\xi = \frac{1}{(2\pi)^n} d\xi$ the normalised Lebesgue measure on $T_x^* U$.

In the sequel, E denotes a smooth Hermitian vector bundle based on some closed Riemannian manifold M modelled on some vector space V .

A pseudodifferential operator (PDO), which for a given atlas on M has a classical symbol with coefficients in $\text{End}(V)$ in the local coordinates defined by each chart is called *classical*. Let $C\ell^\alpha(M, E)$ denote the subset of operators in $C\ell(M, E)$ of order α ; the product of two operators $A_i \in C\ell^{\alpha_i}(M, E)$, $i = 1, 2$ does not generally lie in $C\ell^{\alpha_1+\alpha_2}(M, E)$; it does if $\alpha_1 - \alpha_2$ is an integer. Let $C\ell(M, E)$ denote the algebra of all classical PDOs acting on $C^\infty(M, E)$ generated by $\bigcup_{\alpha \in \mathbb{C}} C\ell^\alpha(M, E)$. Guillemin [Gu] and Wodzicki [Wo] showed the following remarkable property.

PROPOSITION 3. *Let A in $C\ell^\alpha(M, E)$ be a classical PDO represented in a local coordinate chart U by $\sigma(A)$ in $CS^a(U, V)$. Then $\text{res}_x(\sigma(A)) \, dx$ determines a global density on M which defines the projectively unique trace on $C\ell(M, E)$:*

$$(18) \quad \text{res}(A) := \int_M \text{res}_x(\sigma(A)) \, dx = \int_M dx \int_{S_x^* M} \text{tr}_x((\sigma(A))_{-n}(x, \xi)) \, d_S \xi,$$

known as the noncommutative residue (also called the Guillemin-Wodzicki residue). Here tr_x denotes the fibrewise trace over $x \in M$.

The terminology refers to the trace property; if the manifold M is connected and has dimension larger than 1, then up to a scalar multiple, equation (18) defines on $\mathcal{C}\ell(M, E)$ the unique linear functional vanishing on commutators $[A, B] = AB - BA$

$$\text{res}([A, B]) = 0, \quad \forall A, B \in \mathcal{C}\ell(M, E).$$

It also follows from its definition that the residue trace vanishes on operators of order $< -n$ and on non-integer order operators.

On the other hand, it was observed by Kontsevich and Vishik [KV] that on a closed n -dimensional manifold, the usual L^2 -trace on Ψ DOs of real order $< -n$ extends to a functional on the set $\mathcal{C}\ell^{\mathbb{C} \setminus \mathbb{Z}}(M, E) := \cup_{a \in \mathbb{C} \setminus \mathbb{Z}} \mathcal{C}\ell^a(M, E)$ of Ψ DOs of non-integer order and vanishes on commutators of non-integer order. Whenever a classical pseudodifferential operator A has non-integer order, so has its symbol $\sigma(A)$ and

$$(19) \quad \text{TR}_x(A) dx = \left(\int_{T_x^* U} \text{tr}_x(\sigma(A)(x, \xi)) d\xi \right) dx$$

defines a global density. The canonical trace can therefore be defined without ambiguity on $\mathcal{C}\ell^{\mathbb{C} \setminus \mathbb{Z}}(M, E)$.

DEFINITION 4. For a pseudodifferential operator A in $\mathcal{C}\ell^{\mathbb{C} \setminus \mathbb{Z}}(M, E)$ the canonical trace is defined by

$$\text{TR}(A) := \int_M \text{TR}_x(A) dx = \int_M dx \int_{T_x^* U} \text{tr}_x(\sigma(A)(x, \xi)) d\xi.$$

It coincides with the usual trace on Ψ DOs of order $< -n$.

On commutators the canonical trace has the following vanishing property, providing some justification for the fact that it is called a trace.

PROPOSITION 4. Let A be an operator in $\mathcal{C}\ell^a(M, E)$, and B an operator in $\mathcal{C}\ell^b(M, E)$. If $a + b \notin [-n, \infty) \cap \mathbb{Z}$, then the canonical trace is defined on the commutator $[A, B] = AB - BA$ and is equal to zero,

$$\text{TR}([A, B]) = 0.$$

The canonical trace is canonical in so far as it is the unique [MSS] linear form on $\mathcal{C}\ell^{\mathbb{C} \setminus \mathbb{Z}}(M, E)$ (meaning by this that it takes linear combinations in $\mathcal{C}\ell^{\mathbb{C} \setminus \mathbb{Z}}(M, E)$ to linear combinations of their images) which coincides with the L^2 -trace tr on smoothing operators.

Let us finally recall an important formula based on results by Guillemin [Gu] and Wodzicki [Wo] which were later generalised and popularised by Kontsevich and Vishik in [KV]. Let $A(z)$ be a holomorphic family in $\mathcal{C}\ell(M, E)$ of affine order $\alpha(z)$ such that $\alpha'(0) \neq 0$, which coincides with $A \in \mathcal{C}\ell(M, E)$ at zero. The map $z \mapsto \text{TR}(A(z))$ is meromorphic with a discrete set of simple poles and its complex residue at $z = 0$ relates to the noncommutative residue of A :

$$(20) \quad \text{Res}_{z=0} \text{TR}(A(z)) = -\frac{1}{\alpha'(0)} \text{res}(A).$$

In particular, for any admissible operator Q of positive order q , the map $z \mapsto \text{TR}(A Q^{-z})$ is meromorphic with a discrete set of poles and

$$(21) \quad \text{Res}_{z=0} \text{TR}(A Q^{-z}) = \frac{\text{res}(A)}{q}.$$

We call the finite part at $z = 0$, namely the constant term in the Laurent expansion, the Q -weighted trace of A :

$$(22) \quad \text{Tr}^Q(A) = \text{fp}_{z=0} \text{TR}(A Q^{-z}).$$

Unlike the complex residue, the constant term cannot generally be expressed in terms of noncommutative residues. However, when A is a differential operator, the constant term in the Laurent expansion is also a residue [PS]:

$$(23) \quad \text{fp}_{z=0} \text{TR}(A(0)) = -\frac{\text{res}(A'(0))}{\alpha'(0)}.$$

In particular,

$$(24) \quad \text{Tr}^Q(A) = -\frac{\text{res}(\log Q)}{q}.$$

2.2. Cauchy integrals on PDOs. We recall how Cauchy integrals can be defined on elliptic operators and establish commutation formulae similar to the formal commutation relations established in Section 1.

Recall that an operator R in $\text{Cl}(M, E)$ of order r acting on smooth sections of the vector bundle $\pi : E \rightarrow M$ induces a bounded map from $H^s(M, E)$ to $H^{s-\text{Re}(r)}(M, E)$ for any $s \in \mathbb{R}$. For an operator R in $\text{Cl}(M, E)$ and two real numbers s, t we define the operator norms (whenever they are finite)

$$(25) \quad \|R\|_{s,t} = \sup_{u \neq 0} \frac{\|Ru\|_t}{\|u\|_s}, \quad \|R\|^{(s)} := \|R\|_{s,s}.$$

DEFINITION 5. We call admissible with spectral cut ϕ an operator $Q \in \text{Cl}(M, E)$ with leading symbol $\sigma_L(Q)$ that has no eigenvalue on the ray $L_\phi = \{re^{i\phi}, r \geq 0\}$ in which case it is elliptic, and such that the spectrum of Q does not meet the ray $\{re^{i\phi}, r \geq 0\}$, thereby requiring its invertibility⁵.

We consider the resolvent

$$R(\lambda, Q) = (\lambda - Q)^{-1}$$

associated with the operator Q . It follows from the theory of elliptic operators on closed manifolds that, for an admissible operator Q , in a neighborhood of infinity (see [G] Lemma 1.7.3) we have:

$$(26) \quad \|R(\lambda, Q)\|^{(s)} = O(|\lambda|^{-1}),$$

i.e.

$$\forall s \in \mathbb{R}, \quad \exists C_s \in \mathbb{R}^+ \text{ such that } \|R(\lambda, Q)\|^{(s)} \leq C_s |\lambda|^{-1}.$$

Let us introduce some notation which will be useful for what follows.

⁵This assumption could be weakened to include non invertible operators at the cost of working instead with $Q + \pi_Q$, where π_Q stands for the orthogonal projection onto the kernel of Q , whenever invertibility is necessary.

DEFINITION 6. For A in $Cl(M, E)$, we set $A_Q^{(0)} = A$ and for any positive integer j :

$$A_Q^{(j)} := \text{ad}_Q^j(A), \quad \text{where} \quad \text{ad}_Q(B) = [Q, B],$$

so that $A_Q^{(j+1)} = \text{ad}_Q(A^{(j)}) = [Q, A^{(j)}]$.

REMARK 2. We shall often drop the subscript Q , writing $A^{(j)}$ instead of $A_Q^{(j)}$. Note that if Q has scalar leading symbol, Q has order q and A order a then $A^{(k)}$ has order $a + k(q - 1)$.

Following notations of [H], for an operator T in $Cl(M, E)$ and operators $T_k, k \in \mathbb{N}$ in $Cl(M, E)$ whose orders have decreasing real part as k increases, we write

$$(27) \quad T \simeq \sum_{k \geq 0} T_k \iff \forall N \in \mathbb{N}, \exists K(N) \in \mathbb{N} \text{ such that } T - \sum_{k=0}^{K(N)} T_k \in Cl^{-N}(M, E),$$

which can be understood as convergence of $\sum_{k \geq 0} T_k$ to T “up to a smoothing operator”. If T_k has order t_k and $t_k - t_{k'}$ lies in \mathbb{Z} for any two indices k and k' , then T_k lies in $Cl^{t_0}(M, E)$ for any non negative integer k ; convergence of $\sum_{k \geq 0} T_k$ in $Cl^{t_0}(M, E)$ equipped with its natural Fréchet topology⁶ then implies convergence up to a smoothing operator. For the sequel, it is also important to note that convergence in the Fréchet topology of $Cl^a(M, E)$ for a given complex number a , implies convergence in the $\|\cdot\|_{s,s-\text{Re}(a)}$ Sobolev norms defined in (25).

We henceforth assume that Q has scalar leading symbol, an assumption which will ensure convergence of infinite series in $Cl^a(M, E)$ for a given complex number a .

The following result that we quote from [H] (see Lemma 4.20 and the proof of Proposition 4.14), is a cornerstone to prove the existence of the quantised weighted traces as well as their locality.

PROPOSITION 5. Let Q in $Cl(M, E)$ be an admissible elliptic operator with positive integer order and scalar leading symbol. For any complex number λ outside the spectrum of Q , and for any operator A in $Cl(M, E)$ and any non negative integer h

$$(28) \quad [(\lambda - Q)^{-h}, A] \simeq \sum_{k=1}^{\infty} \frac{(h+k-1)!}{(h-1)!k!} A^{(k)} (\lambda - Q)^{-h-k}.$$

⁶The Fréchet structure on $Cl^a(M, E)$ for any complex number a is inherited from a Fréchet topology on the corresponding space of symbols. Given a trivialising local chart over an open subset U of M , one equips the set $CS^a(U, V) = CS^a(U) \otimes \text{End}(V)$ of classical symbols of order a on an open subset U of \mathbb{R}^n with coefficients in $\text{End}(V)$ (V is equipped with a norm $\|\cdot\|$) with a Fréchet structure induced by the following countable family of semi-norms labelled by multiindices α, β and integers $j \geq 0, m, N$ (see [H]):

$$\begin{aligned} & \sup_{x \in K_m, \xi \in \mathbb{R}^n} (1 + |\xi|)^{-\text{Re}(a) + |\beta|} \|\partial_x^\alpha \partial_\xi^\beta \sigma(x, \xi)\|; \\ & \sup_{x \in K_m, \xi \in \mathbb{R}^n} |\xi|^{-\text{Re}(a) + N + |\beta|} \left\| \partial_x^\alpha \partial_\xi^\beta \left(\sigma - \sum_{j=0}^{N-1} \psi(\xi) \sigma_{a-j} \right)(x, \xi) \right\|; \\ & \sup_{x \in K_m, |\xi|=1} \|\partial_x^\alpha \partial_\xi^\beta \sigma_{a-j}(x, \xi)\|, \end{aligned}$$

where $K_m, m \in \mathbb{N}$ is a countable sequence of compact sets in U such that $\cup_{m \in \mathbb{N}} K_m = U$.

REMARK 3. Note the analogy between (28) and (10). Here we observe that since Q has integer order, $\lambda - Q$ has order equal to the order q of Q (which is positive by assumption) so that the operators $A^{(k)}(\lambda - Q)^{-h-k}$ have orders $a + k(q-1) - (h+k)q = a - hq - k$ with decreasing real part as the index k increases.

Proof: The proof goes as that of Lemma 1 in Section 1. The case $h = 1$ follows from iterating the following identities

$$\begin{aligned} [(\lambda - Q)^{-1}, A] &= (\lambda - Q)^{-1} [Q, A] (\lambda - Q)^{-1} \\ &= [(\lambda - Q)^{-1}, [Q, A]] (\lambda - Q)^{-1} + [Q, A] (\lambda - Q)^{-2} \\ &= [Q, A] (\lambda - Q)^{-2} + A^{(2)} (\lambda - Q)^{-3} + [(\lambda - Q)^{-1}, A^{(2)}] (\lambda - Q)^{-1} \\ &= \dots \\ &\simeq \sum_{k \geq 1} A^{(k)} (\lambda - Q)^{-(k+1)}. \end{aligned}$$

The general case $h > 1$ then follows by differentiating $h-1$ times with respect to λ . \square

For an admissible operator Q in $C\ell(M, E)$ with spectral cut ϕ , we define a contour $C_\phi = C_{1,\phi,r} \cup C_{2,\phi,r} \cup C_{3,\phi,r}$. The positive real number r is chosen sufficiently small and $C_{1,\theta,r} = \{\lambda = |\lambda|e^{i\phi} \mid +\infty > |\lambda| \geq r\}$, $C_{2,\phi,r} = \{\lambda = re^{i\psi} \mid \phi \geq \psi \geq \phi - 2\pi\}$ and $C_{3,\phi,r} = \{\lambda = |\lambda|e^{i(\phi-2\pi)} \mid r \leq |\lambda| < +\infty\}$.

Let f be a complex valued continuous function on the contour C_ϕ around the spectrum of Q and such that $\rho \mapsto \frac{f(\rho e^{i\phi})}{\rho}$ lies in $L^1([1, \infty[)$.

Equation (26) and the assumption on f imply that the Cauchy integral

$$f(Q) = \frac{1}{2i\pi} \int_{C_\phi} f(\lambda) R(\lambda, Q) d\lambda$$

converges in each Sobolev norm $\|\cdot\|^{(s)}$ defined in (25).

REMARK 4. In general, the definition of $f(Q)$ depends on the choice of ϕ and should therefore carry a subscript ϕ writing $f_\phi(Q)$; when there is no ambiguity in the choice of spectral cut, we may omit it in order to simplify notations.

For any positive integer k , $f^{(k)}(Q)$ is defined by a Cauchy integral:

$$(29) \quad f^{(k)}(Q) = \frac{k!}{2i\pi} \int_{C_\phi} f(\lambda) (\lambda - Q)^{-(k+1)} d\lambda$$

which converges in the $\|\cdot\|^{(s)}$ norms for all $s \in \mathbb{R}$.

Since Q has scalar leading symbol, if $f^{(k)}(Q)$ lies in $C\ell(M, E)$, then the operators $\text{ad}_Q^k(A) f^{(k)}(Q)$ have order $k(q-1) + \text{ord}(f^{(k)}(Q))$ where as before, $\text{ord}(A)$ denotes the order of A .

PROPOSITION 6. Under the assumptions of Proposition 5 and provided the real part of $k(q-1) + \text{ord}(f^{(k)}(Q))$ decreases by integer steps as k increases, then for any $A \in C\ell(M, E)$ the operator $\text{ad}_Q^k(A) f^{(k)}(Q)$ in $C\ell(M, E)$ has order with real part decreasing as the index k grows and

$$(30) \quad [f(Q), A] \simeq \sum_{k=1}^{\infty} \frac{\text{ad}_Q^k(A)}{k!} f^{(k)}(Q).$$

Proof: By Proposition 5 we have

$$\begin{aligned}
[f(Q), A] &= \frac{1}{2i\pi} \int_{C_\phi} f(\lambda) [R(\lambda, Q), A] d\lambda \\
&\simeq \sum_{k=1}^{\infty} \text{ad}_Q^k(A) \frac{1}{2i\pi} \int_{C_\phi} f(\lambda) R(\lambda, Q)^{k+1} d\lambda \\
&\simeq \sum_{k=1}^{\infty} \frac{\text{ad}_Q^k(A)}{k!} \frac{1}{2i\pi} \int_{C_\phi} f^{(k)}(\lambda) R(\lambda, Q) d\lambda \\
&\simeq \sum_{k=0}^{\infty} \frac{\text{ad}_Q^k(A)}{k!} f^{(k)}(Q).
\end{aligned}$$

□

EXAMPLE 1. The function $f(z) = \lambda^{-z}$ with $\text{Re}(z) > 0$ satisfies the integrability assumption and the corresponding Cauchy integral leads to complex powers

$$(31) \quad Q_\phi^{-z} = \frac{i}{2\pi} \int_{C_\phi} \lambda^{-z} R(\lambda, Q) d\lambda$$

which converge in norm $\|\cdot\|^{(s)}$ for any $s \in \mathbb{R}$.

As usual, for $k \in \mathbb{N}$ the complex power Q^{-z} is then extended to the half plane $\text{Re}(z) > -k$ via the relation

$$Q^k Q_\phi^{-z-k} = Q_\phi^{-z}$$

With the notations of Proposition 6, we have $k(q-1) + o(f^{(k)}(Q)) = k(q-1) - (z+k)q = -zq - k$ and equation (30) reads:

$$[Q_\phi^{-z}, A] \simeq \sum_{k=1}^{\infty} (-1)^k \frac{\Gamma(z+k)}{\Gamma(z)} \frac{\text{ad}_Q^k(A)}{k!} Q_\phi^{-z-k}.$$

EXAMPLE 2. The function $f(\lambda) = e^{-t\lambda}$ on \mathbb{R}^+ also has the required integrability property on C_π . Thus, if Q is a non negative self-adjoint operator, we can define the corresponding heat operator by a Cauchy integral

$$e^{-tQ} = \frac{1}{2i\pi} \int_{C_0} e^{-\lambda t} R(\lambda, Q) d\lambda$$

for any $t > 0$.

Since $f^{(k)}(Q)$ is smoothing, the proposition does not quite apply directly. However one can check in a similar manner that:

$$[e^{-tQ}, A] \sim \sum_{k=1}^{\infty} (-t)^k \frac{\text{ad}_Q^k(A)}{k!} e^{-tQ}$$

where we have extended the relation \simeq to

$$T \sim \sum_{k=0}^{\infty} T_k \iff \exists C \in C\ell(M, E) \quad \text{s.t.} \quad T C^{-1} \simeq \sum_{k=0}^{\infty} T_k C^{-1}$$

with $C = e^{-tQ}$ here.

2.3. Quantised resolvents on PDO chains. Let Q in $C\ell(M, E)$ be an admissible operator with positive integer order q and spectral cut L_ϕ . The quantised resolvent is obtained from a polarised version of the perturbative expansion of $R(\lambda, Q + A) = (\lambda - (Q + A))^{-1}$ of the resolvent.

LEMMA 3. *For any A in $C\ell^0(M, E)$ we have*

$$(\lambda - (Q + A))^{-1} \simeq (\lambda - Q)^{-1} + \sum_{n=1}^{\infty} (\lambda - Q)^{-1} A (\lambda - Q)^{-1} \cdots A (\lambda - Q)^{-1} A (\lambda - Q)^{-1}.$$

REMARK 5. *This can be compared with (7). Here we observe that since the operator Q has positive integer order q , $\lambda - Q$ has order q and the operator $(\lambda - Q)^{-1} A (\lambda - Q)^{-1} \cdots A (\lambda - Q)^{-1} A (\lambda - Q)^{-1}$ corresponding to the n -th term in the sum has order $-nq$ which decreases with n .*

Proof: We write $(\lambda - Q)(\lambda - Q - A)^{-1} = (\lambda - Q - A)(\lambda - Q - A)^{-1} + A(\lambda - Q + A)^{-1} = 1 + A(\lambda - Q - A)^{-1}$, from which it follows that

$$(\lambda - Q - A)^{-1} = (\lambda - Q)^{-1} + (\lambda - Q)^{-1} A (\lambda - Q - A)^{-1}.$$

By induction we get

$$\begin{aligned} (\lambda - Q - A)^{-1} &= (\lambda - Q)^{-1} + \sum_{n=1}^N (\lambda - Q)^{-1} A (\lambda - Q)^{-1} A \cdots A (\lambda - Q)^{-1} A (\lambda - Q)^{-1} \\ &\quad + R_N(Q, A, \lambda), \end{aligned}$$

where R_N stands for the remainder term. The result follows since the remainder term is an operator whose order has decreasing real part as the index k increases, as a consequence of the above remark. \square

DEFINITION 7. *The quantised resolvent $R_\bullet(\lambda, Q)$ of Q applies to the space of chains $C\ell_\bullet(M, E) = \bigotimes^{\bullet+1} C\ell(M, E)$ built from the algebra $C\ell(M, E)$ and coincides with the ordinary resolvent on $C_0(M, E) = C\ell(M, E)$. It is defined on $C\ell_n(M, E) = \bigotimes^{n+1} C\ell(M, E)$ for any positive integer n by⁷*

$$\begin{aligned} R_n(\lambda, Q) : \quad C_{n-1}(M, E) &\rightarrow C\ell(M, E) \\ A_1 \otimes \cdots \otimes A_n &\mapsto R(\lambda, Q) A_1 \cdots R(\lambda, Q) A_n R(\lambda, Q). \end{aligned}$$

The following proposition shows that $R_n(\lambda, Q)(A_1 \otimes \cdots \otimes A_n)$ lies in $C\ell(M, E)$.

PROPOSITION 7. *For any $A_1, \dots, A_n \in C\ell(M, E)$, for any positive integer n ,*

$$\begin{aligned} &R_n(\lambda, Q)(A_1 \otimes \cdots \otimes A_n) \\ &\simeq \sum_{|k| \geq 0} c(k_1, \dots, k_n) A_1^{(k_1)} A_2^{(k_2)} \cdots A_n^{(k_n)} (\lambda - Q)^{-(|k|+n+1)}. \end{aligned}$$

where for any multiindex $k = (k_1, \dots, k_n)$ we have set $c(k_1) = 1$ if $n = 1$ and

$$(32) \quad c(k) := c(k_1, \dots, k_n) = \frac{\Gamma(k_1 + \cdots + k_n + n)}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_{n-1} + n - 1)}.$$

if $n \geq 2$.

⁷Since this sum restricts to a finite sum on each of the $C_n(M, E)$, we do not need to worry about convergence issues at this stage, which is why we do not restrict to zero order operators as in the previous lemma.

REMARK 6. This is similar to Proposition 1 in Section 1. Here, we observe that since Q has scalar symbol, the product $A_1^{(k_1)} A_2^{(k_2)} \cdots A_n^{(k_n)}$ has order $|a| + |k|(q-1)$ where $|a|$ is the total order of the product $A_1 \cdots A_n$ so that $A_1^{(k_1)} A_2^{(k_2)} \cdots A_n^{(k_n)} (\lambda - Q)^{-|k|-n-1}$ has order $|a| + |k|(q-1) - q(|k| + n + 1) = |a| - |k| - q(n+1)$, whose real part decreases as $|k|$ grows.

Proof of the proposition: Applying Proposition 5 to $h_i = k_1 + \cdots + k_i + i + 1$ with $i = 1, \dots, n-1$ we find

$$\begin{aligned}
& R_n(\lambda, Q)(A_1 \otimes \cdots \otimes A_n) \\
&= (\lambda - Q)^{-1} A_1 \cdots (\lambda - Q)^{-1} A_n (\lambda - Q)^{-1} \\
&\simeq \sum_{k_1 \geq 0} A_1^{(k_1)} (\lambda - Q)^{-2-k_1} A_2 \cdots (\lambda - Q)^{-1} A_n (\lambda - Q)^{-1} \\
&\simeq \sum_{k_1 \geq 0} A_1^{(k_1)} \sum_{k_2 \geq 0} \frac{(k_1 + k_2 + 1)!}{(k_1 + 1)! k_2!} A_2^{(k_2)} (\lambda - Q)^{-3-k_1-k_2} \cdots A_n (\lambda - Q)^{-1} \\
&\simeq \sum_{k_1 \geq 0} A_1^{(k_1)} \sum_{k_2 \geq 0} \frac{(k_1 + k_2 + 1)!}{(k_1 + 1)! k_2!} A_2^{(k_2)} \sum_{k_3 \geq 0} \frac{(k_1 + k_2 + k_3 + 2)!}{(k_1 + k_2 + 2)! k_3!} \\
&\quad A_3^{(k_3)} (\lambda - Q)^{-4-k_1-k_2-k_3} A_4 \cdots (\lambda - Q)^{-1} A_n (\lambda - Q)^{-1} \\
&\simeq \sum_{|k| \geq 0} \frac{(k_1 + k_2 + 1)! \cdots (k_1 + k_2 + \cdots + k_n + n - 1)!}{k_1! \cdots k_n! (k_1 + 1) (k_1 + k_2 + 2)! \cdots (k_1 + \cdots + k_{n-1} + n - 1)!} \\
&\quad \cdot A_1^{(k_1)} A_2^{(k_2)} \cdots A_n^{(k_n)} (\lambda - Q)^{-|k|-n-1} \\
&\simeq \sum_{|k| \geq 0} c(k_1, \dots, k_n) A_1^{(k_1)} A_2^{(k_2)} \cdots A_n^{(k_n)} (\lambda - Q)^{-|k|-n-1}.
\end{aligned}$$

□

2.4. Quantised maps on Ψ DO chains. Let Q in $C\ell(M, E)$ be an admissible operator with positive integer order q , scalar leading symbol and spectral cut L_ϕ . Using the quantised resolvent of Q , one can define quantised maps as in Theorem 1.

PROPOSITION 8. Let f be a complex valued function defined on the contour $C_\phi \subset \mathbb{C}$ around the spectrum of Q and such that $\rho \mapsto \frac{f(\rho e^{i\phi})}{\rho^j}$ lies in $L^1([1, \infty[)$ for any $j \in \mathbb{N}$. Then, for any positive integer n and for any A_1, \dots, A_n in $C\ell(M, E)$, the Cauchy integral

$$\begin{aligned}
& \frac{1}{2i\pi} \int_{C_\phi} f(\lambda) R_n(\lambda, Q)(A_1, \dots, A_n) d\lambda \\
&= \frac{1}{2i\pi} \int f(\lambda) R(\lambda, Q) A_1 \cdots A_{n-1} R(\lambda, Q) A_n R(\lambda, Q) d\lambda
\end{aligned}$$

converges in any Sobolev operator norm defined in (25).

Proof: Since $\|AB\|^{(s)} \leq \|A\|^{(s)} \|B\|^{(s)}$, for any $A_1, \dots, A_n \in \mathcal{C}\ell(M, E)$ of order a_1, \dots, a_n respectively,

$$\begin{aligned} \|R(\lambda, Q)A_1 \cdots R(\lambda, Q)A_n R(\lambda, Q)\|^{(s)} &\leq \prod_{i=1}^n \|A_i\|^{(s)} \left(\|R(\lambda, Q)\|^{(s)} \right)^{n+1} \\ &= O(\lambda^{-(n+1)}) \end{aligned}$$

Under the assumptions on f , the convergence in any Sobolev norm $\|\cdot\|^{(s)}$, $s \in \mathbb{R}$ of the Cauchy integral then follows. \square

We therefore set the following definition:

DEFINITION 8. Let f be a complex valued function defined on the contour C_ϕ around the spectrum of Q and such that $\rho \mapsto \frac{f(\rho e^{i\phi})}{\rho^j}$ lies in $L^1([1, \infty[)$ for any $j \in \mathbb{N}$. The quantised Cauchy integral is defined by:

$$(33) \quad f_\bullet(Q) := \frac{1}{2i\pi} \int_{C_\phi} f(\lambda) R_\bullet(\lambda, Q) d\lambda,$$

where the expression $R_\bullet(\lambda, Q)$ is the quantised resolvent. In particular, $f_0(Q) = f(Q)$ and for any positive integer n we have

$$\begin{aligned} f_n(Q)(A_1 \otimes \cdots \otimes A_n) &= \frac{1}{2i\pi} \int_{C_\phi} d\lambda f(\lambda) R_n(\lambda, Q) (A_1 \otimes \cdots \otimes A_n) \\ &= \frac{1}{2i\pi} \int_{C_\phi} d\lambda f(\lambda) R(\lambda, Q) A_1 \cdots R(\lambda, Q) A_n R(\lambda, Q). \end{aligned}$$

for all $A_1 \otimes \cdots \otimes A_n$ in $\mathcal{C}_n(M, E)$.

The following result shows that the operator $f_n(Q)(A_1 \otimes \cdots \otimes A_n)$ lies in $\mathcal{C}\ell(M, E)$.

THEOREM 3. Let f be as in the above definition. Provided $f^{(|k|+n)}(Q)$ lies in $\mathcal{C}\ell(M, E)$ and the real part of $|k|(q-1) + \text{ord}(f^{(|k|+n)}(Q))$ decreases as $|k|$ increases, where $\text{ord}(A)$ denotes the order of A , then for any A_1, \dots, A_n in $\mathcal{C}\ell(M, E)$ the operator $f_\bullet(Q)(A_1 \otimes \cdots \otimes A_n)$ lies in $\mathcal{C}\ell(M, E)$ and reads:

$$\begin{aligned} &f_\bullet(Q)(A_1 \otimes \cdots \otimes A_n) \\ &\simeq \sum_{|k| \geq 0} \frac{c(k)}{(|k|+n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} f^{(|k|+n)}(Q) \\ (34) \quad &\simeq \sum_{|k| \geq 0} \frac{\delta_{k_1}(A_1) \cdots \delta_{k_n}(A_n)}{(k_1+1)(k_1+k_2+2) \cdots (k_1 + \cdots + k_n + n)} f^{(|k|+n)}(Q), \end{aligned}$$

where $f^{(l)}(Q)$ was defined in (29) and $c(k)$ in (32). Here $\delta_k(A) := \frac{A^{(k)}}{k!}$.

REMARK 7. Note the analogy between (34) and (13). Here, by assumption the order $a_1 + \cdots + a_n - |k|(q-1) + o(f^{(|k|+n)}(Q))$ of $A_1^{(k_1)} \cdots A_n^{(k_n)} f^{(|k|+n)}(Q)$ has decreasing real part as $|k|$ grows.

Proof of the theorem: Combining Proposition 7 with (29) applied to $|k| + n$ instead of k yields:

$$\begin{aligned} f_n(Q)(A_1 \otimes \cdots \otimes A_n) &= \frac{1}{2i\pi} \int f(\lambda) R_n(\lambda, Q)(A_1 \otimes \cdots \otimes A_n) d\lambda \\ &\simeq \sum_{|k| \geq 0} c(k) A_1^{(k_1)} \cdots A_n^{(k_n)} \frac{1}{2i\pi} \int f(\lambda) (\lambda - Q)^{-|k|-n-1} d\lambda \\ &\simeq \sum_{|k| \geq 0} \frac{c(k)}{(|k| + n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} f^{(|k|+n)}(Q). \end{aligned}$$

It follows that $f_\bullet(Q)$ induces a map:

$$\begin{aligned} \mathcal{C}_\bullet(M, E) &\rightarrow C\ell(M, E) \\ A_0 \otimes \cdots \otimes A_n &\mapsto A_0 f_n(Q)(A_1 \otimes \cdots \otimes A_n). \end{aligned}$$

We now apply Theorem 3 to $f(\lambda) = \lambda^{-z}$ for which the assumptions of the theorem hold so that Q_\bullet^{-z} is well defined.

COROLLARY 2. *For any $A_1, \dots, A_n \in C\ell(M, E)$ and for any positive integer n we have:*

$$\begin{aligned} &Q_n^{-z} (A_1 \otimes \cdots \otimes A_n) \\ &\simeq \sum_{|k| \geq 0} \frac{(-1)^{|k|+n} c(k) \Gamma(z + |k| + n)}{\Gamma(z)(|k| + n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-|k|-n}. \end{aligned}$$

REMARK 8. *With the notations of [H] (cfr. formula (4.1)), we have*

$$(35) \quad \langle A_0, \dots, A_n \rangle_z = (-1)^n \Gamma(z) \text{TR} (A_0 Q_\bullet^{-z} (A_0 \otimes \cdots \otimes A_n)).$$

Proof of the corollary: When $f(\lambda) = \lambda^{-z}$, the real part of $|k|(q-1) + \text{ord}(f^{(|k|+n)}(Q)) = -|k| - q(z+n)$ decreases as $|k|$ increases and the quantised complex power of Q reads

$$\begin{aligned} &Q_n^{-z} (A_1 \otimes \cdots \otimes A_n) \\ &\simeq \sum_{k \geq 0} \frac{c(k)(-z)(-z-1) \cdots (-z-(|k|+n-1))}{(|k|+n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \\ &\simeq \sum_{|k| \geq 0} \frac{(-1)^{|k|+n} c(k) (z+|k|+n-1) \cdots (z+1)z}{(|k|+n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \\ &\simeq \sum_{|k| \geq 0} \frac{(-1)^{|k|+n} c(k) \Gamma(z+|k|+n)}{\Gamma(z)(|k|+n)!} A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)}. \end{aligned}$$

□

3. Second quantised regularised traces

We now define quantised regularised traces similar to the ones that arise in the context of the index theorem in non commutative geometry. Unlike ordinary regularised traces, they lead to local expressions involving noncommutative residues. As we pointed out in the introduction, the locality arises here from the fact that we are considering higher tensor products (indexed by n) of operators.

3.1. Second quantised zeta regularisation. Let Q in $C\ell(M, E)$ be invertible elliptic with scalar leading symbol and positive integer order q . Zeta regularisation $\zeta^{Q, \zeta}$, which sends an operator A in $C\ell(M, E)$ to a holomorphic map $z \mapsto A Q^{-z}$ in $C\ell(M, E)$, can be quantised to a map $\mathcal{R}_\bullet^{Q, \zeta}$ which sends a chain $A_0 \otimes A_1 \otimes \cdots \otimes A_n$ in $\mathcal{C}_\bullet(M, E)$ to the map

$$z \mapsto \mathcal{R}_n^{Q, \zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z) := A_0 Q_n^{-z} (A_1 \otimes \cdots \otimes A_n)$$

which lies in $C\ell(M, E)$.

THEOREM 4. *Let Q in $C\ell(M, E)$ be an admissible operator with scalar leading symbol and positive integer order q .*

For any positive integer n , for any $K > 0$ and any A_0, \dots, A_n in $C\ell(M, E)$, the expression

$$\begin{aligned} \mathcal{R}_n^{Q, \zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z) - \sum_{|k|=0}^K \frac{c(k)(-1)^{|k|+n}}{(|k|+n)!} \frac{\Gamma(z+|k|+n)}{\Gamma(z)} \\ \times A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \end{aligned}$$

is trace-class for $\text{Re}(z) > \frac{d-K+|a|}{q} - n$.

The map $z \mapsto \text{Tr}(\mathcal{R}_n^{Q, \zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z))$ extends to a meromorphic map $\text{Tr}(\mathcal{R}_n^{Q, \zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n))|^{mer}$ defined on the whole plane with a discrete set of simple poles

$$\mathcal{P}_n = \left\{ \frac{|a| + d - K}{q} - n, \quad K \in \mathbb{N} \right\}.$$

Its finite part at $z = p$ is a local expression given by a linear combination of non-commutative residues:

$$(36) \quad \begin{aligned} & \text{fp}_{z=p} (\text{Tr}(\mathcal{R}_n^{Q, \zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n))|^{mer}) \\ &= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} + p - n} \frac{(-1)^{|k|+n} c(k)}{q (|k|+n)} \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right) \end{aligned}$$

$$(37) \quad \begin{aligned} &= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} + p - n} \frac{(-1)^{|k|+n}}{q k_1! \cdots k_n!} \frac{\Gamma(k_1 + \cdots + k_n + n)}{(k_1 + 1) \cdots (k_1 + \cdots + k_n + n)} \\ & \quad \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right), \end{aligned}$$

where as before,

$$c(k) = \frac{\Gamma(k_1 + \cdots + k_n + n)}{k_1! \cdots k_n! (k_1 + 1) (k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_{n-1} + n - 1)}.$$

Here $|a| = a_1 + \cdots + a_n$ stands for the sum of the orders a_i of the A_i 's and d is the dimension of M .

REMARK 9. *Since Q has scalar leading symbol, the operator $A^{(k)}$ has order $a+k(q-1)$ where a is the order of A and the operator $A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)+p}$ has order $|a| + |k| (q-1) - (|k|+n)q + pq = |a| - |k| + (p-n)q$ the real part of which decreases as $|k|$ increases. Since the noncommutative residue vanishes for operators of order $< -d$ where d is the underlying dimension of the manifold M ,*

only a finite number ($|k| \leq |a| + (p - n)q + d$) of non vanishing residues arise in (36).

Proof of the theorem: The proof relies on (21). By Corollary 2 we have

$$\begin{aligned} & \mathcal{R}_n^{Q,\zeta} (A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z) \\ & \simeq \sum_{|k| \geq 0} c(k) (-1)^{|k|+n} \frac{\Gamma(z + |k| + n)}{\Gamma(z) (|k| + n)!} A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)}. \end{aligned}$$

The operator $A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)-z}$ of order $|a| - |k| - nq - qz$ is trace-class for $\operatorname{Re}(z) > \frac{|a|-|k|+d}{q} - n$. Since the operators $A_i^{(k_i)}$ are classical, so is the operator $A := A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)}$, and the map $z \mapsto \operatorname{Tr}(A Q^{-z})$ which is holomorphic on this half plane, extends to a meromorphic function:

$$z \mapsto \operatorname{TR} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)-z} \right)$$

on the whole complex plane with a discrete set of simple poles

$$\mathcal{P}_{n,k} = \left\{ \frac{|a| + d - K}{q} - n, \quad K \in \mathbb{N} \cap [|k|, \infty[\right\}.$$

For any positive integer K the map

$$\begin{aligned} z \mapsto \operatorname{Tr} \left(\mathcal{R}_n^{Q,\zeta} (A_0 \otimes A_1 \otimes \cdots \otimes A_n) \right. \\ \left. - \sum_{|k|=0}^K \frac{c(k) (-1)^{|k|+n} \Gamma(z + |k| + n)}{\Gamma(z) (|k| + n)!} A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \right) \end{aligned}$$

is holomorphic on the half-plane $\operatorname{Re}(z) > \frac{d-K+|a|}{q} - n$. Hence, the map which to z assigns

$$\begin{aligned} & \sum_{|k|=0}^K c(k) (-1)^{|k|+n} \frac{\Gamma(z + |k| + n)}{\Gamma(z) (|k| + n)!} \operatorname{TR} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \right) \\ & + \operatorname{Tr} \left(\mathcal{R}_n^{Q,\zeta} (A_0 \otimes A_1 \otimes \cdots \otimes A_n) \right. \\ & \left. - \sum_{|k|=0}^K c(k) (-1)^{|k|+n} \frac{\Gamma(z + |k| + n)}{\Gamma(z) (|k| + n)!} A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \right) \end{aligned}$$

defines a meromorphic extension of $\operatorname{Tr}(\mathcal{R}_n^{Q,\zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z))$ to the half-plane $\operatorname{Re}(z) > \frac{d-K+|a|}{q} - n$ with simple poles in the set

$$\mathcal{P}_n = \left\{ \frac{|a| + d - K}{q} - n, \quad K \in \mathbb{N} \right\}.$$

These patch up to a meromorphic extension $\operatorname{Tr}(\mathcal{R}_n^{Q,\zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n))|_{\operatorname{mer}}(z)$ of $\operatorname{Tr}(\mathcal{R}_n^{Q,\zeta}(A_0 \otimes A_1 \otimes \cdots \otimes A_n)(z))$ to the whole complex plane with the same discrete set of simple poles \mathcal{P}_n .

Let us compute its finite part at $z = p$ for some integer p .
For any positive integer K ,

$$\frac{\Gamma(u + K)}{\Gamma(u)} \sim_{u \rightarrow 0} \Gamma(K) u$$

so that for $n > p$

$$\begin{aligned} \frac{\Gamma(z - p + |k| + n)}{\Gamma(z - p)} &= \frac{\Gamma(z - p + |k| + n)(z - 1) \cdots (z - p)}{\Gamma(z)} \\ &\sim_{z \rightarrow 0} (-1)^p p! \Gamma(|k| + n - p) z. \end{aligned}$$

Hence,

$$\begin{aligned} &\text{fp}_{z=p} \left(\frac{\Gamma(z + |k| + n)}{\Gamma(z) (|k| + n)!} \text{TR} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-(|k|+n)} \right) \right) \\ &\text{fp}_{z=0} \left(\frac{\Gamma(z - p + |k| + n)}{\Gamma(z - p) (|k| + n)!} \text{TR} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z+p-(|k|+n)} \right) \right) \\ &= \frac{(-1)^p p! \Gamma(|k| + n - p)}{(|k| + n)!} \text{Res}_{z=0} \text{TR} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-z-|k|-n+p} \right) \\ &= \frac{(-1)^p p! \Gamma(|k| + n - p)}{q (|k| + n)!} \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right), \end{aligned}$$

as a result of (21) applied to $A = A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p}$.

For $|k| > \frac{d + \text{Re}(|a|)}{q} + p - n$, the operators $A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p}$ have order $|a| - |k| + (p - n)q$ with real part $< -d$ so that $\text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right)$ vanish. It follows that

$$\begin{aligned} &\text{fp}_{z=p} \left(\text{Tr} \left(\mathcal{R}_n^{Q, \zeta} (A_0 \otimes A_1 \otimes \cdots \otimes A_n) \right) |^{\text{mer}}(z) \right) \\ &= \text{fp}_{z=0} \left(\text{Tr} \left(\mathcal{R}_n^{Q, \zeta} (A_0 \otimes A_1 \otimes \cdots \otimes A_n) \right) |^{\text{mer}}(z - p) \right) \\ &= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} + p - n} \frac{(-1)^{|k|+n+p} p!}{q} \frac{c(k) \Gamma(|k| + n - p)}{q (|k| + n)!} \\ &\quad \times \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right), \end{aligned}$$

which is a local expression as a finite linear combination of noncommutative residues.
□

REMARK 10. *With the notations of [H] (see (35) formula (4.1)) we set*

$$\langle A_0 \otimes \cdots \otimes A_n \rangle_z = (-1)^n \frac{\Gamma(z)}{2i\pi} \text{Tr} \left(\int \lambda^{-z} A_0 (\lambda - Q)^{-1} \cdots A_n (\lambda - Q)^{-1} d\lambda \right),$$

so that (36) reads

$$\begin{aligned}
& \text{Res}_{z=0} (\langle A_0 \otimes \cdots \otimes A_n \rangle_{z-p}) \\
&= (-1)^n \text{Res}_{z=0} (\Gamma(z-p) \text{Tr} (\mathcal{R}_n^{Q,\zeta} (A_0 \otimes \cdots \otimes A_n)(z-p))) \\
&= (-1)^n \text{Res}_{z=0} \left(\frac{\Gamma(z)}{(z-1) \cdots (z-p)} \text{Tr} (\mathcal{R}_n^{Q,\zeta} (A_0 \otimes \cdots \otimes A_n)(z-p)) \right) \\
&= \frac{(-1)^{n+p}}{p!} \text{fp}_{z=0} \text{Tr} (\mathcal{R}_n^{Q,\zeta} (A_0 \otimes \cdots \otimes A_n)(z-p)) \\
&= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} + p - n} \frac{(-1)^{|k|} c(k) \Gamma(|k| + n - p)}{q(|k| + n)!} \\
&\quad \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p} \right) \\
&= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} + p - n} \frac{(-1)^{|k|} c(k) \Gamma(|k| + n - p)}{(|k| + n)!} \\
&\quad \times \text{Res}_{z=0} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p-z} \right),
\end{aligned}$$

where we have used (21) applied to the operator $A = A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n+p}$ in the last line. For $p = n/2$ this reads:

$$\begin{aligned}
& \text{Res}_{z=0} (\langle A_0 \otimes \cdots \otimes A_n \rangle_{z-p}) \\
&= \sum_{0 \leq |k| \leq \frac{\text{Re}(|a|)+d}{q} - n/2} c(n, k) \text{Res}_{z=0} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n/2-z} \right)
\end{aligned}$$

which compares with Theorem 5.6 in [H] where we have set:

$$(38) \quad c(n, k) := \frac{(-1)^{|k|}}{k_1! \cdots k_n!} \frac{\Gamma(|k| + n/2)}{(k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_n + n)}.$$

3.2. Quantised weighted traces as residues of quantised logarithms.
On the grounds of the results of Theorem 4, extracting finite parts at $p = 0$ we can set the following definition which extends the notion of Q -weighted trace recalled in formula (22).

DEFINITION 9. Let Q in $C\ell(M, E)$ be an admissible operator with scalar leading symbol and positive integer order q , n a positive integer n and any A_0, \dots, A_n operators in $C\ell(M, E)$. We call the finite part at $z = 0$ of the meromorphic extension

$$\text{Tr}_n^Q (A_0 \otimes \cdots \otimes A_n) := \text{fp}_{z=0} \text{Tr} (\mathcal{R}_n^\zeta (A_0 \otimes A_1 \otimes \cdots \otimes A_n)|^{mer}(z))$$

the quantised Q -weighted trace of $A_0 \otimes \cdots \otimes A_n \in \mathcal{C}_n(M, E)$.

The following corollary is a straightforward consequence of Theorem 4.

COROLLARY 3. Given operators $A_0, \dots, A_n \in C\ell(M, E)$ the second quantised Q -weighted trace of $A_0 \otimes A_1 \otimes \cdots \otimes A_n$ is a local expression given by a linear

combination of noncommutative residues:

$$\begin{aligned} & \text{Tr}_n^Q (A_0 \otimes A_1 \otimes \cdots \otimes A_n) \\ = & \sum_{0 \leq |k| \leq \frac{\text{Re}(\|a\|) + d}{q} - n} \frac{(-1)^{|k|+n} c(k)}{q(|k|+n)} \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-|k|-n} \right). \end{aligned}$$

This local feature is specific to quantised weighted traces since an ordinary weighted trace is not generally expected to be local.

The various residue terms in the sum on the r.h.s can be combined to form a residue of a quantised operator, thus generalising the fact that the Q -weighted trace of a differential operator A is local as the residue of a logarithm (see formula (24) taken from [PS]):

$$(39) \quad \text{Tr}^Q (A) = -\frac{1}{q} \text{res} (A \log Q).$$

Interestingly, the following generalisation holds for any operators A_0, A_1, \dots, A_n in $C\ell(M, E)$ as long as $n \geq 1$, whereas when $n = 0$ we used the assumption that $A = A_0$ is differential to derive (39).

THEOREM 5. *Given an admissible operator $Q \in C\ell(M, E)$ with scalar leading symbol and positive integer order q , for any positive integer n the second quantised Q -weighted trace of $A_0 \otimes \cdots \otimes A_n$ in $\mathcal{C}_n(M, E)$ is the residue of a quantised logarithm:*

$$(40) \quad \text{Tr}^Q (A_0 \otimes A_1 \otimes \cdots \otimes A_n) = -\frac{1}{q} \text{res} (A_0 \log_{\bullet} Q (A_1 \otimes \cdots \otimes A_n)).$$

Proof: Applying (34) to $f(\lambda) = \log \lambda$ we have:

$$\begin{aligned} & \log_{\bullet} (Q) (A_1 \otimes \cdots \otimes A_n) \\ \simeq & \sum_{|k| \geq 0} \frac{\delta_{k_1}(A_1) \cdots \delta_{k_n}(A_n)}{(k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_n + n)} \\ & \times (-1)^{|k|+n-1} (|k| + n - 1)! Q^{-(|k|+n)}. \end{aligned}$$

Multiplying on the left by A_0 and implementing the noncommutative residue yields:

$$\begin{aligned} & \text{res} (A_0 \log_{\bullet} (Q) (A_1 \otimes \cdots \otimes A_n)) \\ = & \sum_{|k| \geq 0} \frac{(-1)^{|k|+n-1} (|k| + n - 1)!}{k_1! \cdots k_n! (k_1 + 1)(k_1 + k_2 + 2) \cdots (k_1 + \cdots + k_n + n)} \\ & \times \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)} \right). \\ = & \sum_{|k| \geq 0} \frac{(-1)^{|k|+n-1} c(k)}{k_1 + \cdots + k_n + n} \text{res} \left(A_0 A_1^{(k_1)} \cdots A_n^{(k_n)} Q^{-(|k|+n)} \right). \end{aligned}$$

This combined with Corollary 3 yields (40). \square

REMARK 11. *We could not extend the proof to provide a similar description of the finite part at $p > 0$ for, unlike in the case $p = 0$, the higher derivatives of*

$f(\lambda) = \lambda^p \log \lambda$:

$$f^{(j)}(\lambda) = \partial_x|_{x=p} (x(x-1) \cdots (x-j+1) \lambda^{x-j}) =$$

$$\frac{p!}{(p-j)!} \left(\sum_{i=0}^{j-1} \frac{1}{p-i} + \log \lambda \right) \lambda^{p-j}$$

involve the logarithm. Hence, Theorem 3 is not applicable here. Indeed inserting these higher derivatives in (34) applied to $f(\lambda) = \lambda^p \log \lambda$ would yield:

$$f_{\bullet}(Q) (A_1 \otimes \cdots \otimes A_n)$$

$$\simeq \sum_{|k| \geq 0} \frac{\delta_{k_1}(A_1) \cdots \delta_{k_n}(A_n)}{(k_1+1) \cdots (k_1 + \cdots + k_n + n)} \frac{p!}{(p - (|k| + n))!}$$

$$\times \left(\sum_{i=0}^{(|k|+n)-1} \frac{1}{p-i} + \log Q \right) Q^{p-(|k|+n)},$$

which is built up from expressions that also involve the logarithm of Q and therefore have a priori ill-defined residues.

References

- [APS] M. Atiyah, V. Patodi, I. Singer, *Spectral asymmetry and Riemannian geometry, I*, Math. Proc. Cambridge Phil. Soc. **77**, (1975) 43–69
- [CMa] A. Connes, M. Marcolli, **Noncommutative Geometry, Quantum Fields and Motives**, Colloquium Publications, Amer. Math. Soc. (2008)
- [CM] A. Connes, H. Moscovici, *The local index formula in noncommutative geometry*, Geom. Funct. Anal. **5** (2) (1995) 174–243
- [EGP] K. Ebrahimi-Fard, J. Gracia-Bondia, F. Patras, *A Lie theoretic approach to renormalization*, Comm. Math. Phys. **276** (2007) 519–549
- [G] P. Gilkey, **Invariance theory, the heat equation and the Atiyah-Singer index theorem**, Studies in Advanced Mathematics 1995
- [Gu] V. Guillemin, *A new proof of Weyl's formula on the asymptotic distribution of eigenvalues*, Adv. Math. **55**, (1985) 131–160
- [H] N. Higson, *The residue index theorem of Connes and Moscovici*, Clay Mathematics Proceedings (2004)(<http://www.math.psu.edu/higson/ResearchPapers.html>)
- [Ho] L. Hörmander, **The analysis of linear PDO, III**, Springer-Verlag Heidelberg 1985
- [KV] M. Kontsevich, S. Vishik, *Determinants of elliptic pseudo-differential operators*, Max Planck Institut Preprint, 1994
- [Le] M. Lesch, *On the noncommutative residue for pseudo-differential operators with log-polyhomogeneous symbols*, Ann. Global Anal. Geom., **17** (1999) 151–187
- [MP] B. Mielnik, J. Plebanski, *Combinatorial approach to Baker-Campbell-Hausdorff exponents*, Annales de l'I.H.P., section A , tome 12 n. 3 (1970) 215–254
- [MSS] L. Maniccia, E. Schrohe, J. Seiler, *Uniqueness of the Kontsevich-Vishik trace* Proc. Amer. Math. Soc. **136** no. 2, (2008) 747–752
- [P] S. Paycha, *(Second) quantised resolvents and regularised traces*, J. Geom. Phys. **57** (2007) 1345–1369
- [PS] S. Paycha, S. Scott, *A Laurent expansion for regularised integrals of holomorphic symbols*, Geom. Funct. Anal., Geom. Funct. Anal. **17** (2007) 491–536
- [Sel] R.T. Seeley, *Complex powers of an elliptic operator, Singular integrals*, Proc. Symp. Pure Math., Chicago, Amer. Math. Soc., Providence (1966) 288–307
- [Se2] R.T. Seeley, *Topics in pseudodifferential operators*. CIME Conference on Pseudodifferential Operators 1968, Edizioni Cremonese, Roma, 169–305 (1969).
- [Sh] M.A. Shubin, **Pseudodifferential operators and spectral theory**, Springer Verlag 1987.
- [Ta] M. Taylor, **Pseudodifferential operators**, Princeton University Press 1981

- [Wo] M. Wodzicki, *Non commutative residue* in Lecture Notes in Math. **1283**, Springer Verlag 1987; Spectral asymmetry and noncommutative residue (in Russian), Thesis, Steklov Institute (former) Soviet Academy of Sciences, Moscow, 1984.

Motives: an introductory survey for physicists

Abhijnan Rej
(with an appendix by Matilde Marcolli)

ABSTRACT. We survey certain accessible aspects of Grothendieck's theory of motives in arithmetic algebraic geometry for mathematical physicists, focussing on areas that have recently found applications in quantum field theory. An appendix (by Matilde Marcolli) sketches further connections between motivic theory and theoretical physics.

post hoc, ergo ante hoc

– Umberto Eco, *Interpretation and overinterpretation*

CONTENTS

| | |
|---|-----|
| 1. Introduction | 377 |
| 2. The Grothendieck ring | 379 |
| 3. The Tannakian formalism | 384 |
| 4. Weil cohomology | 390 |
| 5. Classical motives | 394 |
| 6. Mixed motives | 398 |
| 7. Motivic measures and zeta functions | 403 |
| Appendix A. Motivic ideas in physics (by M. Marcolli) | 407 |
| References | 413 |

1. Introduction

This survey paper is based on lectures given by the author at Boston University, the Max Planck Institute in Bonn, at Durham University and at the Indian Statistical Institute and the S.N. Bose National Center for the Basic Sciences in Kolkata and the Indian Institute of Technology in Mumbai. The purpose of these

1991 *Mathematics Subject Classification.* 19E15, 14C25, 81T18, 18F30.

Key words and phrases. motives, K -theory, Grothendieck ring, algebraic cycles, motivic zeta functions, Feynman graphs, tensor categories, gauge theory, dualities, mirror symmetry.

Part of this work was done while the author was employed by the Clay Mathematics Institute as a Junior Research Scholar.

introductory notes are to familiarize an audience of physicists with some of the algebraic and algebro-geometric background upon which Grothendieck's theory of motives of algebraic varieties relies. There have been many recent developments in the interactions between high energy physics and motives, mostly within the framework of perturbative quantum field theory and the evaluation of Feynman diagrams as periods of algebraic varieties, though motives are beginning to play an important role in other branches of theoretical physics, such as string theory, especially through the recent interactions with the Langlands program, and through the theory of BPS states. We focus here mostly on the quantum field theoretic applications when we need to outline examples that are of relevance to physicists. The appendix to the paper, written by Matilde Marcolli, sketches several examples of how the main ideas involved in the theory of motives, algebraic cycles, periods, Hodge structures, K-theory, have in fact been already involved in many different ways in theoretical physics, from condensed matter physics to mirror symmetry. Most of the paper focuses on the mathematical background.

We describe the Grothendieck ring of varieties and its properties, since this is where most of the explicit computations of motives associated to Feynman integrals are taking place. We then discuss the Tannakian formalism, because of the important role that Tannakian categories and their Galois groups play in the theory of perturbative renormalization after the work of Connes–Marcolli. We then describe the background cohomological notions underlying the construction of the categories of pure motives, namely the notion of Weil cohomology, and the crucial role of algebraic cycles in the theory of motives. It is in fact mixed motives and not the easier pure motives are involved in the application to quantum field theory, due to the fact that the projective hypersurfaces associated to Feynman graphs are highly singular, as well as to the fact that relative cohomologies are involved since the integration computing the period computation that gives the Feynman integral is defined by an integration over a domain with boundary. We will not cover in this survey the construction of the category of mixed motives, as this is a technically very challenging subject, which is beyond what we are able to cover in this introduction. However, the most important thing to keep in mind about mixed motives is that they form *triangulated* categories, rather than abelian categories, except in very special cases like mixed Tate motives over a number field, where it is known that an abelian category can be constructed out of the triangulated category from a procedure known as the heart of a t-structure. For our purposes here, we will review, as an introduction to the topic of mixed motives, some notions about triangulated categories, Bloch–Ogus cohomologies, and mixed Hodge structures. We end by reviewing some the notion of motivic zeta function and some facts about motivic integration. Although this last topic has not yet found direct applications to quantum field theory, there are indications that it may come to play a role in the subject.

A word about the references: the list is much longer than the list of works cited in the main body of the text. This is deliberate since the list is meant to also serve as a guide to further reading.

Acknowledgements. These notes, as noted in the introduction, form a part of my lectures on motives at various institutions in the US, Germany, the UK and India during 2005–2009, in particular, my talk at the Bonn workshop on renormalization in December 2006. I thank the organizers (Ebrahimi-Fard, van Suijlekom and

Marcolli) for a stimulating meeting and for their infinite patience in waiting for this written contribution towards the proceedings of that meeting.

2. The Grothendieck ring

2.1. Definition of the Grothendieck groups and rings. Recall the definition of the usual Grothendieck group of vector bundles on a smooth variety X :

Definition 2.1 (exercise 6.10 of [36]). $K_0(X)$ is defined as the quotient of the free abelian group generated by all vector bundles (= locally free sheaves) on X by the subgroup generated by expressions $\mathcal{F} - \mathcal{F}' - \mathcal{F}''$ whenever there is an exact sequence of vector bundles $0 \rightarrow \mathcal{F}' \rightarrow \mathcal{F} \rightarrow \mathcal{F}'' \rightarrow 0$. The group $K_0(X)$ can be given a ring structure via the tensor product.

We remark that isomorphism classes of vector bundles form an abelian monoid under direct sum; in fact, Grothendieck groups can be defined in a much more general way because of the following universal property:

Proposition 2.2. *Let M be an abelian monoid. There exists an abelian group $K(M)$ and $\gamma : M \rightarrow K(M)$ a monoid homomorphism satisfying the following universal property: If $f : M \rightarrow A$ is a homomorphism into an abelian group A , then there exists a unique homomorphism of abelian groups $f_* : K(M) \rightarrow A$ such that the following diagram commutes:*

$$\begin{array}{ccc} M & \xrightarrow{\gamma} & K(M) \\ f \downarrow & & \downarrow f_* \\ A & \xlongequal{\quad} & A \end{array}$$

The proof of proposition 2.2 is very simple: construct a free abelian group F generated by M and let $[x]$ be a generator of F corresponding to the element $x \in M$. Denote by B the subgroup of F generated by elements of the form $[x+y] - [x] - [y]$ and set $K(M)$ to be the quotient of F by B . Letting γ to be the injection of M into F and composing with the canonical map $F \rightarrow F/B$ shows that γ satisfies the universal property.

One can show that projective modules over a ring A give rise to a Grothendieck group denoted as $K(A)$ (p. 138 of [42]). This is simply done by noting that isomorphism classes of (finite) projective A -modules form a monoid (again under the operation of direct sum) and taking the subgroup B to be generated by elements of the form $[P \oplus Q] - [P] - [Q]$ for finite projective modules P and Q . We can refine $K(A)$ by imposing the following equivalence: P is equivalent to $Q \iff$ there exists free modules F, F' such that $[P \oplus F] = [Q \oplus F']$. Taking the quotient of $K(A)$ with respect to this equivalence relation gives us $K_0(A)$.

The explicit determination of K_0 (and its “higher dimensional” analogues) for a given ring A is the rich subject of algebraic K -theory. (A standard reference for algebraic K -theory is [54]; see also the work-in-progress [60].) For the simplest case when A is a field, we note that $K_0(A) \simeq \mathbb{Z}$; this immediately follows from the definition and the trivial fact that modules over a field are vector spaces. As a non-trivial example, we have the following

Example 2.3 (example 2.1.4 of [60]). Let A be a semisimple ring with n simple modules. Then $K_0(A) \simeq \mathbb{Z}^n$.

Through the theorem of Serre-Swan which demonstrates that the categories of vector bundles and finite projective modules are equivalent, we see that definition 2.1 arise very naturally from proposition 2.2. Furthermore, the generality of proposition 2.2 enables us construct Grothendieck groups and rings of (isomorphism classes of) other objects as well.

Let \mathbf{Var}_k the category of quasi-projective varieties over a field k . Let X be an object of \mathbf{Var}_k and $Y \hookrightarrow X$ a closed subvariety. Central to our purposes would be the following

Definition 2.4. The Grothendieck ring $K(\mathbf{Var}_k)$ is the quotient of the free abelian group of isomorphism classes of objects in \mathbf{Var}_k by the subgroup generated the expressions $[X] - [Y] - [X \setminus Y]$. The ring structure given by fiber product: $[X] \cdot [Y] := [X \times_k Y]$.

In fact one can generalize definition 2.4 to define the Grothendieck ring of a *symmetric monoidal category*. Recall the following

Definition 2.5 (defintion 5.1 of chapter II of [60]). A category \mathbf{C} is called *symmetric monoidal* if there is a functor $\otimes : \mathbf{C} \rightarrow \mathbf{C}$ and a distinguished $\mathbb{I} \in \text{Obj}(\mathbf{C})$ such that the following are isomorphisms for all $S, T, U \in \text{Obj}(\mathbf{C})$:

$$\begin{aligned} \mathbb{I} \otimes S &\cong S \\ S \otimes \mathbb{I} &\cong S \\ S \otimes (T \otimes U) &\cong (S \otimes T) \otimes U \\ S \otimes T &\cong T \otimes S. \end{aligned}$$

Furthermore one requires the isomorphisms above to be *coherent*, a technical condition that guarantees that one can write expressions like $S_1 \otimes \dots \otimes S_n$ without paranthesis without ambiguity¹. (See any book on category theory for the precise definitions.)

Remark 2.6. The category \mathbf{Var}_k is symmetric monoidal with the functor \otimes given by fiber product of varieties and the unit object \mathbb{I} given by $\mathbb{A}^0 = \text{point}$.

An important example of a symmetric monoidal category is furnished by the category of finite dimensional complex representations of a finite group G (with the morphisms given by intertwiners between representations), denoted as $\mathbf{Rep}_{\mathbb{C}}(G)$.

Example 2.7 (example 5.2.3 of [60]). $\mathbf{Rep}_{\mathbb{C}}(G)$ is symmetric monoidal under direct sums of representations. Furthermore, $K_0(\mathbf{Rep}_{\mathbb{C}}(G)) \simeq R(G)$ generated by irreducible representations and where $R(G)$ is the representation ring of G .

Note that for more general G (say an affine group scheme over an arbitrary base), the category $\mathbf{Rep}_{\mathbb{C}}(G)$ has much more structure than being merely symmetric monoidal, namely it is a Tannakian category. We will turn to this rich subject in section 3.

2.2. The Grothendieck ring of varieties $K(\mathbf{Var}_k)$. Recall definition 2.4. The Grothendieck ring $K(\mathbf{Var}_k)$ has the following properties (cf. the review [56]):

- (1) If X is a variety and U, V locally closed subvarieties in X , then

$$[U \cup V] + [U \cap V] = [U] + [V].$$

¹Note that we will sometimes call a symmetric monoidal category an “ACU \otimes -category” and refer to the functor \otimes as the *tensor functor*.

- (2) If X is the disjoint union of locally closed subvarieties X_1, \dots, X_n , then $[X] = \sum_{i=1}^n [X_i]$.
- (3) Let C be a constructible subset of a variety X (that is, C is the disjoint union of locally closed subsets of X .) Then C has a class in $K(\mathbf{Var}_k)$.

In section 6 of [44], Marcolli proposes a category of Feynman motivic sheaves which involves viewing the Kirchhoff polynomial as a morphism $\Psi_\Gamma : \mathbb{A}^n \setminus X_\Gamma \rightarrow \mathbb{G}_m$ where X_Γ is viewed as an *affine* hypersurface obtained by setting $\Psi_\Gamma = 0$. (Here n is the number of edges of Γ and \mathbb{G}_m is the multiplicative group.) As a setting for such relative questions, we introduce the following:

Definition 2.8 (Bittner, 2.1.1 of [56]). Let S be a variety over k . The ring $K(\mathbf{Var}_S)$ is the free abelian group generated by isomorphism classes $[X]_S$ (where X is a variety over S) modulo the relation $[X]_S = [X \setminus Y]_S + [Y]_S$ where $Y \subset X$ is a closed subvariety. The ring structure is induced by the fiber product of varieties.

Remark 2.9. We note the following properties of $K(\mathbf{Var}_S)$:

- (1) $K(\mathbf{Var}_S)$ is a $K(\mathbf{Var}_k)$ -module.
- (2) There is an bilinear associative exterior product

$$K(\mathbf{Var}_S) \times K(\mathbf{Var}_T) \xrightarrow{\boxtimes} K(\mathbf{Var}_{S \times_k T}).$$

- (3) Let $f : S \rightarrow S'$ be a morphism of varieties. It induces $f_! : K(\mathbf{Var}_S) \rightarrow K(\mathbf{Var}_{S'})$ and $f^* : K(\mathbf{Var}_{S'}) \rightarrow K(\mathbf{Var}_S)$ that are functorial with respect to \boxtimes : let $g : T \rightarrow T'$ be another morphism of varieties. Then $(f \times g)_!(A \boxtimes B) = f_!(A) \boxtimes g_!(B)$ and $(f \times g)^*(A \boxtimes B) = f^*(A) \boxtimes g^*(B)$.

Note that in case of $S = \text{Spec } k$, the definition 2.8 is the same as definition 2.4.

Theorem 2.10 ([6], from [56]). *Let k a field of characteristic zero. Then $K(\mathbf{Var}_k)$ is generated by smooth varieties.*

The idea behind the proof is this (proof of proposition 2.1.2 of [56]): Set $d := \dim X$ and let $X \hookrightarrow X'$ for a complete variety X' . Writing $[X] + [Z] = [X']$ for some Z with $\dim Z \leq d-1$ and using Hironaka's theorem, we get $[\overline{X}] = [X'] - ([C] - [E])$ where C is the smooth center of the blowup \overline{X} and E its exceptional divisor with $\dim C, \dim E \leq d-1$. We can write any arbitrary X as the disjoint union of smooth varieties in this way by inducting on d .

In $K(\mathbf{Var}_k)$, there are two “distinguished” classes, the class of a point $[\mathbb{A}^0]$ and the class of the affine line $\mathbb{A}^1 = \text{Spec } k[x]$ denoted as \mathbb{L} . The standard cell decomposition of the projective space in terms of flags (see p.194 of [31]) can be lifted to $K(\mathbf{Var}_k)$ in terms of these classes:

$$\begin{aligned} [\mathbb{P}^n] &= 1 + \sum_{i=0}^{n-1} \mathbb{L}^i &= \frac{1 - \mathbb{L}^{n+1}}{1 - \mathbb{L}} \\ &= \frac{(1 + \mathbb{T})^{n+1} - 1}{\mathbb{T}} \end{aligned}$$

where $\mathbb{T} := [\mathbb{A}^1] - [\mathbb{A}^0]$ the class of the torus \mathbb{G}_m .

In order to prove this decomposition, note that $[\mathbb{P}_k^1 \setminus \mathbb{A}_k^1] = [\mathbb{P}_k^1] - [\mathbb{A}_k^1]$ since $\mathbb{A}_k^1 \subset \mathbb{P}_k^1$ is open with ∞ as the compliment. Induction gives the result since $\mathbb{P}_k^{n+1} \setminus \mathbb{A}_k^{n+1} \simeq \mathbb{P}_k^n$.

One can formally invert \mathbb{L} in $K(\mathbf{Var}_k)$ to get the *Tate motive* denoted as $\mathbb{Q}(1)$. By $\mathbb{Q}(n)$ one means $\mathbb{Q}(1) \cdots \mathbb{Q}(1)$ (n -times) and $n \in \mathbb{Z}$ is called the *twist* of the Tate motive. One also sets as a matter of notation (as of now!) that $\mathbb{Q}(-1) := \mathbb{L}$.

A rather useful fact in the context of blowups of graph hypersurfaces is the following

Proposition 2.11 ([56]). *Let $f : X \rightarrow Y$ be a proper morphism of smooth varieties which is a blowup with the smooth center $Z \subset Y$ of codimension d . Then*

$$[f^{-1}(Z)] = [Z][\mathbb{P}^{d-1}].$$

As an example of decomposition of varieties into Lefschetz motives, we take a family of hypersurfaces in \mathbb{P}^{n-1}

$$X_{\Gamma_n} = \{(t_1, \dots, t_n) | \Psi_{\Gamma_n} = t_1 \cdots t_n \left(\frac{1}{t_1} + \cdots + \frac{1}{t_n} \right) = 0\}.$$

The polynomial Ψ_{Γ_n} is the Kirchoff polynomial attached to a graph with two vertices and n number of parallel edges between them (the *banana graphs*); it arises in quantum field theory in the denominator of the Schwinger parametrized integral attached to the same graph. Then we have

Example 2.12 (Aluffi-Marcolli, theorem 3.10 of [1]). The class associated to the banana graph hypersurfaces X_{Γ_n} is given by

$$\begin{aligned} [X_{\Gamma_n}] &= \frac{(1 + \mathbb{T})^n - 1}{\mathbb{T}} - \frac{\mathbb{T}^n - (-1)^n}{\mathbb{T} + 1} - n\mathbb{T}^{n-2} \\ &= \frac{\mathbb{L}^n - 1}{\mathbb{L} - 1} - \frac{(\mathbb{L} - 1)^n - (-1)^n}{\mathbb{L}} - n(\mathbb{L} - 1)^{n-2}. \end{aligned}$$

There is an alternative characterization of $K(\mathbf{Var}_k)$ due to Bittner, which is also sometimes useful in the context of Feynman graphs and motives.

Definition 2.13 (definition 2.3 of [56]). $K^{\text{bl}}(\mathbf{Var}_k)$ is the abelian group generated by smooth complete varieties, subject to the conditions

- (1) $[\emptyset] = 0$ and
- (2) $[\text{Bl}_Y X] = [X] - [Y] + [E]$ where Y is a smooth and complete subvariety of X and $\text{Bl}_Y X$ is the blowup of X along Y with the exceptional divisor of the blowup E .

We need to show that the definitions 2.4 and 2.13 do coincide, following [56]. To do this we need to show that there exists a map $K(\mathbf{Var}_k) \rightarrow K^{\text{bl}}(\mathbf{Var}_k)$ which is an isomorphism. This is constructed as an induced map from a map e on varieties satisfying $e(X \setminus Y) = e(X) - e(Y)$. The map e is constructed in the following way: Let \overline{X} be the smooth completion of a smooth connected variety X and $D = X \setminus \overline{X}$ a normal crossing divisor. Let us define the map $e(X) = \sum (-1)^l [D^{(l)}]_{\text{bl}}$ with $D^{(l)}$ denoting the disjoint union of l -fold intersections of irreducible components of D . ($[-]_{\text{bl}}$ denotes a class in $K^{\text{bl}}(\mathbf{Var}_k)$.) The fact that $e(X)$ is independent of the choice of completion \overline{X} follows from the weak factorization theorem. The fact that $e(X) = e(X \setminus Y) + e(Y)$ is proved by choosing a smooth and complete $X \subset \overline{X}$ such that $D = \overline{X} \setminus X$ is simple normal crossing and the closure \overline{Y} in X is smooth and normal-crossing with divisor D . ($D \cap Y$ is simple normal crossings divisor in Y .)

Yet another useful result, which can be useful in applications in the context of Feynman graphs, is the following.

Proposition 2.14 ([56]). *Let X be a smooth connected variety, $Y \subset X$ a smooth connected subvariety of X of codimension d . Let E be the exceptional divisor of the blowup $\text{Bl}_Y X$ of X along Y . Then*

$$[\text{Bl}_Y X] = [X] + \mathbb{L}[E] - \mathbb{L}^d[Y].$$

A consequence of proposition 2.14 is the notion of *Tate twist*: there is a ring involution

$$\mathbb{L} \mapsto \mathbb{L}^{-1} = \text{ and } [X] \mapsto \mathbb{L}^{-\dim X}[X] = \mathbb{Q}(1)^{\dim X}[X].$$

This is the true meaning of the Tate twist alluded above. In cohomological calculations, the product of $\mathbb{Q}(1)^{\dim X}$ and $[X]$ is replaced by a tensor product of a one dimensional rational vector space (“working mod torsion”) raised to the dimension of the variety whose cohomology is being computed and a piece of the cohomology. In fact this correspondence between classes in $K(\mathbf{Var}_k)$ and cohomology is one important aspect of the theory of motives.

The reader may wonder what $K(\mathbf{Var}_k)$ explicitly is and how does it depend on k . The fact of the matter is that this is a very hard question though certain things are known. For example, we have

Theorem 2.15 (Poonen, theorem 1 of [53]). *Let k be a field of characteristic zero. Then $K(\mathbf{Var}_k)$ is not an integral domain.*

In fact when $k = \mathbb{C}$, we have

Theorem 2.16 (Larsen–Lunts, theorem 2.3 of [43]). *Let I be the ideal generated by the affine line, i.e., $I = \mathbb{L}$ and denote by SB the monoid of classes of stable birational complex varieties (the monoid structure coming from fiber products of varieties.) Then*

$$K(\mathbf{Var}_k)/I \simeq \mathbb{Z}[SB].$$

(Two varieties X and Y are *stably birational* if $X \times \mathbb{P}^m$ is birational to $Y \times \mathbb{P}^n$ for $m, n \geq 0$.)

In Sahasrabudhe’s thesis [56], theorem 2.16 is proved using Bittner’s definition 2.13.

The main obstruction in obtaining results like theorems 2.15 and 2.16 in characteristic p is that those results depend crucially on the resolution of singularities and the weak factorization of birational morphisms, statements that are not yet known to hold true in characteristic p .

2.3. A Grothendieck ring for supermanifolds. One can generalize the construction of the Grothendieck ring of \mathbf{Var}_k in a quite straightforward way for the category of *complex supermanifolds*. (See [49] for a standard introduction to supermanifolds.)

Definition 2.17 (definition 2.2 of [46]). *Let $\mathbf{SVar}_{\mathbb{C}}$ be the category of complex supermanifolds. Let $K(\mathbf{SVar}_{\mathbb{C}})$ denote the free abelian group generated by the isomorphism classes of objects $\mathcal{X} \in \text{Obj } (\mathbf{SVar}_{\mathbb{C}})$ subject to the following relations. Let $F : \mathcal{Y} \hookrightarrow \mathcal{X}$ be a closed embedding of supermanifolds. Then*

$$[\mathcal{X}] = [\mathcal{Y}] + [\mathcal{X} \setminus \mathcal{Y}],$$

where $\mathcal{X} \setminus \mathcal{Y}$ is the supermanifold

$$\mathcal{X} \setminus \mathcal{Y} = (X \setminus Y, \mathcal{A}_X|_{X \setminus Y}).$$

The standard notation $\mathcal{A}_X|_{X \setminus Y}$ denotes restriction of the sheaf of supercommutative rings on \mathcal{X} to the complement. We can relate $K(\mathbf{SVar}_{\mathbb{C}})$ to $K(\mathbf{Var}_{\mathbb{C}})$ through

Proposition 2.18 (corollary 2.4 of [46]). *The Grothendieck ring $K(\mathbf{SVar}_{\mathbb{C}})$ of supervarieties is a polynomial ring over the Grothendieck ring of ordinary varieties of the form*

$$K(\mathbf{SVar}_{\mathbb{C}}) = K(\mathbf{V}_{\mathbb{C}})[T],$$

where $T = [\mathbb{A}^{0|1}]$ is the class of the affine superspace of dimension $(0, 1)$.

The notion of birational and stable birational equivalence can be generalized also in a straightforward way to supermanifolds: two supermanifolds \mathcal{X} and \mathcal{Y} are said to be birationally equivalent if there are superprojective spaces $\mathbb{P}^{n|m}$ and $\mathbb{P}^{r|s}$ such that $\mathcal{X} \times \mathbb{P}^{n|m}$ is birationally equivalent to $\mathcal{Y} \times \mathbb{P}^{r|s}$. Denote by $\mathbb{Z}[SSB]$ the monoidal ring of such supermanifolds. Then we have a result similar to theorem 2.16.

Proposition 2.19 (corollary 2.5 of [46]). *There is an isomorphism*

$$K(\mathbf{SVar}_{\mathbb{C}})/I \simeq \mathbb{Z}[SSB],$$

where I is the ideal generated by the classes $[\mathbb{A}^{1|0}]$ and $[\mathbb{A}^{0|1}]$.

The main application of the Grothendieck rings constructed in this section is in defining a large class of topological and arithmetic invariants as measures on these rings. We take up this subject in section 7.

3. The Tannakian formalism

3.1. Categorical notions. Most of the material in this section is based on [22] and Breen's survey of Saveedra-Rivano's thesis under Grothendieck [55]. In a nutshell, the main idea behind a Tannakian category is to equip an abelian category with a natural functor to vector spaces such that a fiber over each object in this category is a finite dimensional vector space.

Let (\mathbf{C}, \otimes) be a symmetric monoidal category.

Definition 3.1 (Internal homs). Let $X, Y \in \text{Obj } (\mathbf{C})$ and consider the functor

$$\begin{aligned} F : \mathbf{C} &\rightarrow \mathbf{Set}, \\ T &\mapsto \text{Hom}(T \otimes X, Y). \end{aligned}$$

If $F = \text{Hom}(-, K)$ where K is some object in \mathbf{C} , then define $\underline{\text{Hom}}(X, Y) =: K$.

Remark 3.2. Explicitly, $\text{Hom}(T \otimes X, Y) = \text{Hom}(T, \underline{\text{Hom}}(X, Y))$. This spells out the fact that the functor F is *representable*.

Example 3.3. Consider the category of all R -modules \mathbf{Mod}_R . This is an obvious tensor category with $\mathbb{I} = R$. In \mathbf{Mod}_R , $\underline{\text{Hom}}(X, Y) = \text{Hom}_{\mathbf{Mod}_R}(X, Y)$. To see this, define the functor F of definition 3.1 to be $T \mapsto \text{Hom}(T \otimes X, Y)$. Hom-tensor adjointness states $\text{Hom}_{\mathbf{Mod}_R}(T \otimes X, Y) = \text{Hom}_{\mathbf{Mod}_R}(T, \text{Hom}_{\mathbf{Mod}_R}(X, Y))$ and hence $F = \text{Hom}_{\mathbf{Mod}_R}(-, \text{Hom}_{\mathbf{Mod}_R}(X, Y))$ which, by definition, proves that $\underline{\text{Hom}}(X, Y) = \text{Hom}_{\mathbf{Mod}_R}(X, Y)$.

Let $\text{ev}_{X,Y} : \underline{\text{Hom}}(X, Y) \otimes X \rightarrow Y$ be the morphism corresponding to $\text{id}_{\underline{\text{Hom}}(X, Y)}$. To see that this correspondence does make sense, take $T = \underline{\text{Hom}}(X, Y)$ in definition 3.1 to get

$$\text{Hom}(\underline{\text{Hom}}(X, Y) \otimes X, Y) = \text{Hom}(\underline{\text{Hom}}(X, Y), \underline{\text{Hom}}(X, Y)).$$

Define the dual of an object X as $\hat{X} := \underline{\text{Hom}}(X, \mathbb{I})$. For example in \mathbf{Mod}_R we have the following diagram

$$\begin{array}{ccc} \underline{\text{Hom}}(X, Y) \otimes X & \xlongequal{\quad} & \text{Hom}(X, Y) \otimes X \\ \text{ev}_{X,Y} \downarrow & & \downarrow \text{ev}_{X,Y} \\ Y & \xlongequal{\quad} & Y \end{array}$$

and $\text{ev}_{X,Y}$ is given by $f \otimes x = f(x)$, the usual evaluation map in \mathbf{Mod}_R . This shows that the abstract notion of evaluation makes sense.

Definition 3.4 (Reflexive objects). Suppose we are given the following diagram

$$\begin{array}{ccc} \underline{\text{Hom}}(X, Y) & \xlongequal{\quad} & \underline{\text{Hom}}(X, Y) \\ \psi \uparrow & & \uparrow \text{ev}_{X,\mathbb{I}} \\ X \otimes \hat{X} & \xrightarrow{\text{ev}_{X,\mathbb{I}} \circ \psi} & \mathbb{I} \end{array}$$

where X is an object in some tensor category. If $\text{ev}_{X,\mathbb{I}} \circ \psi$ is an isomorphism, we say that X is reflexive. If all objects in a tensor category are reflexive, we call the category reflexive.

Remark 3.5. We note that the above definition does correspond to our usual notion of reflexiveness: from the definition of internal hom,

$$\text{Hom}(X \otimes \hat{X}, \mathbb{I}) = \text{Hom}(X, \underline{\text{Hom}}(\hat{X}, \mathbb{I})).$$

If $\text{ev}_{X,\mathbb{I}} \circ \psi$ is an isomorphism, we get a correspondence

$$\text{ev}_{X,\mathbb{I}} \circ \psi \longleftrightarrow X \xrightarrow{\sim} \hat{X}$$

with $\text{ev}_{X,\mathbb{I}} \circ \psi \in \text{Hom}(X \otimes \hat{X}, \mathbb{I})$ and $X \xrightarrow{\sim} \hat{X} \in \text{Hom}(X, \underline{\text{Hom}}(\hat{X}, \mathbb{I}))$.

Of course, not all categories are reflexive. For example, a one-line argument shows that \mathbf{Mod}_R is not. (Take $R = \mathbb{Z}$ and consider $\mathbb{Z}/2\mathbb{Z}$.)

Putting all of these notions together, we have

Definition 3.6 (Rigid tensor category). A tensor category \mathbf{C} is said to be *rigid* if

- (1) $\underline{\text{Hom}}(X, Y)$ exists for every $X, Y \in \text{Obj}(\mathbf{C})$.
- (2) Functoriality of internal homs: All natural maps

$$\bigotimes_{i \in I} \underline{\text{Hom}}(X_i, Y_i) \longrightarrow \underline{\text{Hom}}\left(\bigotimes_{i \in I} X_i, \bigotimes_{i \in I} Y_i\right)$$

are isomorphisms.

- (3) (\mathbf{C}, \otimes) is a reflexive category.

We need a last assumption for the main definition, namely that a tensor category be *abelian*. We quickly review the definition: A category \mathbf{A} is *abelian* [27] if

- $\text{Hom}(X, Y)$ is an abelian group for all $x, y \in \text{Obj}(\mathbf{A})$ and composition of morphisms is biadditive.

- There exists a zero object 0 such that $\text{Hom}(0, 0) = \emptyset$.
- For all $X, Y \in \text{Obj}(\mathbf{A})$, there exists an object $Z \in \text{Obj}(\mathbf{A})$ and morphisms $\iota_X : X \rightarrow Z$, $\iota_Y : Y \rightarrow Z$, $\pi_X : Z \rightarrow X$ and $\pi_Y : Z \rightarrow Y$ with $\pi_X \circ \iota_X = \text{id}_X$, $\pi_Y \circ \iota_Y = \text{id}_Y$ and $\iota_X \circ \pi_X + \iota_Y \circ \pi_Y = \text{id}_Z$ and $\pi_X \circ \iota_Y = \pi_Y \circ \iota_X = 0$.
- For any $f \in \text{Hom}(X, Y)$ there is a sequence $K \xrightarrow{k} X \xrightarrow{i} I \xrightarrow{j} Y \xrightarrow{c}$ where K (resp. C) is the kernel of f (resp. cokernel of f) and such that $j \circ i = f$ and I is both kernel of c and cokernel of k .

With this last piece at hand, we have

Definition 3.7 (Neutral Tannakian category). Let k be a field of arbitrary characteristic. A neutral Tannakian category over k is a abelian rigid tensor category \mathbf{C} with a k -linear exact faithful functor called the *fiber functor* $\omega : \mathbf{C} \rightarrow \mathbf{Vect}_k$ with \mathbf{Vect}_k the category of k -vector spaces².

To fix notation: a neutral Tannakian category would be denoted as a tuple $(\mathbf{C}, \otimes, \omega)$ where $\omega(X \otimes Y) = \omega(X) \otimes \omega(Y)$ from the definition 3.7. Faithfulness, exactness and k -linearity are the usual notions.

3.2. The structure of $\mathbf{Rep}_k(G)$. Let us come back to example 2.7 with G an *affine group scheme* instead, defined over k a field of arbitrary characteristic. Let V be a finite dimensional vector space over k . A representation of G is a morphism of affine schemes $G \times V \rightarrow V$, denoted as ρ . The category $\mathbf{Rep}_k(G)$ is defined in the following way.

- Objects of $\mathbf{Rep}_k(G)$ are representations (ρ, V) .
- Morphisms between two representations are intertwiners. Let (ρ_1, V_1) and (ρ_2, V_2) be two representations. $\text{Hom}_{\mathbf{Rep}_k(G)}((\rho_1, V_1), (\rho_2, V_2))$ consists of maps $f : V_1 \rightarrow V_2$ such that for every $g \in G$, the following diagram commutes:

$$\begin{array}{ccc} G \times V_1 & \xrightarrow{\rho_1} & V_1 \\ \text{id}_G \times f \downarrow & & \downarrow f \\ G \times V_2 & \xrightarrow{\rho_2} & V_2 \end{array}$$

The tensor structure is given by $\rho_1 \otimes \rho_2 : G \rightarrow \text{Aut}(V_1 \otimes_k V_2)$, the unit object $\mathbb{1} : G \rightarrow \text{Aut}(\mathbf{1})$ where $\mathbf{1}$ is the one-dimensional vector space with trivial G -action. (A more general construction of $\mathbf{Rep}_k(G)$ over R , a commutative ring with identity, involves V being a projective R -module of finite rank.)

By a general well-known argument, $\mathbf{Rep}_k(G)$ is equivalent to the category of commutative (though not necessarily cocommutative) Hopf algebras³.

The main theorem along the lines of the classical Pontryagin duality is

Theorem 3.8 (Tannaka-Krein theorem). $\mathbf{Rep}_k(G)$ is a neutral Tannakian category with ω given by the forgetful functor. Furthermore $\text{Aut}^\otimes(\omega) \simeq G$.

The notation $\text{Aut}^\otimes(\omega)$ denotes the natural transformations of the functor ω to itself preserving the tensor structure.

²Some authors, for example Andre [4], demand an additional $\mathbb{Z}/2\mathbb{Z}$ -grading on \mathbf{Vect}_k .

³This is a fact of vital importance in the Connes-Kreimer and Connes-Marcolli theories of renormalization as a Riemann-Hilbert problem.

| Category | Unit | Dual |
|---|----------------------------|---|
| \mathbf{Vect}_k | $\mathbf{1}$ | dual vector space \hat{V} |
| $\mathbf{Rep}_k(G)$ for G affine group scheme | k | \hat{V} with induced G -action |
| Flat \mathbb{C} -vector bundles on X | $\underline{\mathbb{C}}_X$ | $\hat{\mathcal{E}} = \text{Hom}(\mathcal{E}, \underline{\mathbb{C}}_X)$ |
| Connections on \mathbb{G}_m | \mathbb{G}_a | $(\hat{\mathcal{E}}, \hat{\nabla})$ |

TABLE 1. Examples of rigid tensor categories

3.3. An equivalent definition. There is a definition of a neutral Tannakian category, due to Deligne [22], which is equivalent to definition 3.7. This definition has the benefit of being more explicit. I present it for the benefit of the reader who may want to gain a shift in perspective.

Let (\mathbf{C}, \otimes) be a symmetric monoidal category over a field k with a unit object \mathbb{I} . (The k -linearity of the functor \otimes is understood.) Furthermore *assume* (\mathbf{C}, \otimes) to be abelian.

Definition 3.9 (Rigid tensor category, alternate). The category $(\mathbf{C}, \otimes, \mathbb{I})$ is *rigid* if

- (1) $\text{End}(\mathbb{I}) = \text{Hom}(\mathbb{I}, \mathbb{I}) \simeq k$,
- (2) For all $X \in \text{Obj}(\mathbf{C})$ there exist objects $\hat{X} \in \text{Obj}(\mathbf{C})$ and morphisms $\delta : \mathbb{I} \rightarrow \hat{X} \otimes X$ and $\text{ev} : X \otimes \hat{X} \rightarrow \mathbb{I}$ such that

$$\begin{aligned} X &\xrightarrow{\text{id} \otimes \delta} X \otimes \hat{X} \otimes X \xrightarrow{\text{ev} \otimes \text{id}} X, \\ \hat{X} &\xrightarrow{\delta \otimes \text{id}} \hat{X} \otimes X \otimes \hat{X} \xrightarrow{\text{id} \otimes \text{ev}} \hat{X}. \end{aligned}$$

We now *define* the internal hom to be $\underline{\text{Hom}}(X, Y) := \hat{X} \otimes Y$ for all $X, Y \in \text{Obj}(\mathbf{C})$. One can verify, upon making the identification $\text{Hom}(X, Y) = \text{Hom}(X, Y \otimes \mathbb{I})$ that this coincide with definition 3.1. The dual is a functor $\mathbf{C} \rightarrow \mathbf{C}$ given by $(\hat{-}) = \underline{\text{Hom}}(-, \mathbb{I})$. It exists if and only if $\underline{\text{Hom}}(X, -)$ exists and $\underline{\text{Hom}}(X, \mathbb{I}) \otimes Y \xrightarrow{\sim} \underline{\text{Hom}}(X, Y)$. Deligne's definition of a neutral Tannakian category is the same as definition 3.7, namely, an abelian rigid tensor category with a fiber functor $\omega : \mathbf{C} \rightarrow \mathbf{Vect}_k$, the functor being exact, faithful, k -linear and preserving the tensor structure.

The main theorem, in any case, is a stronger formulation of theorem 3.8, variously attributed to Deligne, Grothendieck and Saavedra-Rivano:

Theorem 3.10. *Let \mathbf{T} be a neutral Tannakian category over a field k . Then $\text{Aut}^\otimes(\omega)$ is an affine group scheme over k and we have the equivalence of categories*

$$\mathbf{T} \simeq \mathbf{Rep}_k(\text{Aut}^\otimes(\omega)),$$

where $\mathbf{Rep}_k(-)$ is the category of k -linear representations. Furthermore, if $\mathbf{T} \simeq \mathbf{Rep}_k(G)$ for some affine group scheme G , then there exists an exact, faithful, k -linear tensor functor ω such that

$$G \xrightarrow{\sim} \text{Aut}^\otimes(\omega).$$

The functor ω is representable, following the remark 3.2. Ultimately, the rigidity features guarantee the group structure. In fact, the definition of ω in terms of “functor-of-points” suggests we look at ω with \mathbf{Set} as the target category; this

provides an interpretation of the action of the group $\text{Aut}^\otimes(\omega)$ as a fundamental group, a theme we take up next.

3.4. Fundamental groups of schemes. Tannakian categories can also be used as a language to understand a beautiful theory developed by Grothendieck, Artin and others, of the relationship between fundamental groups of étale coverings of a scheme and Galois groups. The canonical reference for this material is SGA I [32].

Definition 3.11 (Étale covering of a scheme). Let Y be a scheme over k . An étale covering $Y \rightarrow X$ is an affine morphism given by $A \rightarrow B$ with $X = \text{Spec } A$ and $Y = \text{Spec } B$, A and B algebras, and such that

- (1) B is flat over A ,
- (2) $\text{Der}_A B = 0$ and
- (3) B is finite over A .

Example 3.12. Let X be a smooth variety over \mathbb{C} . Saying that $Y \rightarrow X$ is an étale covering over X is the same as $Y_{\mathbb{C}} \rightarrow X_{\mathbb{C}}$ is a covering in the usual sense.

Example 3.13. Let $X = \text{Spec } k$. Then we have the equivalence of sets (in fact of categories!):

$$\{\text{connected étale coverings of } X\} \simeq \{\text{finite separable field extensions of } k\}.$$

For a given and fixed scheme X , étale coverings form a category with morphisms $Y \rightarrow Y'$ and such that the following diagram

$$\begin{array}{ccc} Y & \longrightarrow & Y' \\ \downarrow & & \downarrow \\ X & \xlongequal{\quad} & X \end{array}$$

commutes for étale coverings $Y \rightarrow X$ and $Y' \rightarrow X$. Denote this category as EtCov_X . Given a geometric point $\bar{x} \in X$, define a functor

$$\begin{aligned} \omega : \text{EtCov}_X &\longrightarrow \text{Set}, \\ (Y \xrightarrow{\pi} X) &\mapsto \pi^{-1}(\bar{x}). \end{aligned}$$

Remark 3.14. In case of the universal covering $\tilde{X} \rightarrow X$ has the property that $\text{Hom}(\tilde{X}, X) = \omega(Y)$ does not exist but

$$\omega(Y) = \varprojlim_i \text{Hom}(X_i, Y),$$

This motivates our next definition.

Definition 3.15 (Fundamental group).

$$\pi_1(X, \bar{x}) := \varprojlim_i \text{Aut}_X(X_i).$$

Example 3.16. Let X be a complex variety. Then $\pi_1(X) = \pi_1(\widehat{X_{\mathbb{C}}}, x_{\mathbb{C}})$.

More interestingly,

Example 3.17. Let $X = \text{Spec } k$. Then $\pi_1(X) = \text{Gal}(\bar{k}/k) =: G_k$.

In fact, we have, following definition 3.15:

Theorem 3.18 (Grothendieck). *The category of finite étale schemes over k is equivalent to the category of finite sets with continuous G_k -action.*

In fact, in a Tannakian category with the fiber functor ω , the theorem 3.18 can be understood as saying that the fundamental group acts as automorphisms of ω . To see this, view the category of flat vector bundles over a scheme X as being equivalent to the category of representations of $\pi_1(X, x)$. Furthermore, the category is recovered by studying the category of representations of $\pi_1(X, x)$ using theorem 3.10.

I remark that theorem 3.18 is a special case of the general *Grothendieck-Galois correspondence* which states that the category of finite étale k -schemes is equivalent to the category of finite sets with continuous G_k -action.

3.5. The function-sheaf correspondence. The function-sheaf correspondence is a profound application of constructibility to obtain all “interesting” functions on a space over an arbitrary base ring in terms of certain sheaves on it. As such, it also connects with the *Geometric Langlands program*. In this subsection I scratch the surface, closely following the notes of Sug Woo Shin [58].

Fix a prime l and let K be a finite extension of \mathbb{Q}_l . Denote by \mathcal{O}_K the ring of integers of K . Let X be a connected scheme over k with $\text{char } k \neq l$. Let X_{et} be an étale site of X . Call the étale sheaf \mathcal{G} *locally constant* if $f|_U$ is locally constant for an étale covering U of X . The sheaf G is said to be *constructible* if X is constructible (i.e. can be written as disjoint union of locally closed subschemes of X) and \mathcal{G} is such that it defines a locally constant sheaf, finite on each strata of X .

Definition 3.19 (locally constant l -adic sheaf). An l -adic sheaf on X_{et} is a projective system $(\mathcal{F}_n)_{n \in \mathbb{N}}$ of constructible sheaves \mathcal{F}_n such that $\mathcal{F}_{n+1} \rightarrow \mathcal{F}_n$ induces an isomorphism $\mathcal{F}_{n+1} \otimes (\mathbb{Z}/l^n) \simeq \mathcal{F}_n$. An l -adic sheaf is *locally constant* if each \mathcal{F}_n is so.

One defines a *category of K -sheaves* as a category whose objects are constructible \mathcal{O}_K -sheaves (= constructible sheaves of \mathcal{O}_K -modules) and with

$$\text{Hom}_{K\text{-sheaves}}(\mathcal{F} \otimes K, \mathcal{G} \otimes K) := \text{Hom}_{\mathcal{O}_K\text{-sheaves}}(\mathcal{F}, \mathcal{G}) \otimes_{\mathcal{O}_K} K.$$

By taking a limit over l of the categories of K -sheaves, we obtain the category of $\overline{\mathbb{Q}}_l$ -sheaves. Locally constant $\overline{\mathbb{Q}}_l$ -sheaves are limits of locally constant \mathcal{O}_K -sheaves.

Definition 3.20 (l -adic local system). A locally constant $\overline{\mathbb{Q}}_l$ -sheaf is called an l -adic local system.

Let $\bar{x} \in X$ be a geometric point. Let \mathcal{E} be an l -adic local system on X . The stalk of \mathcal{E} is defined as the direct limit of stalks of K -sheaves. These are, in turn, defined in the following way: for \mathcal{F} a \mathcal{O}_K -sheaf on X , the stalk at \bar{x} is $\varprojlim_n (\mathcal{F}_n)_{\bar{x}}$ where $(\mathcal{F}_n)_{\bar{x}}$ are the usual stalks on the étale site. Denote the stalk of \mathcal{E} at \bar{x} as $\mathcal{E}_{\bar{x}}$. We say that the local system \mathcal{E} has *rank* r if $\mathcal{E}_{\bar{x}}$ contains r -copies of $\overline{\mathbb{Q}}_l$.

Theorem 3.21. *Let $\bar{x} \in X$ be a geometric point of a finite connected scheme X . The functor $\mathcal{F} \rightarrow \mathcal{F}_{\bar{x}}$ makes the category of l -adic local systems on X equivalent to the category of continuous representations of $\pi_1(X, \bar{x})$ on finite dimensional $\overline{\mathbb{Q}}_l$ vector spaces.*

Compare this with theorem 3.18: in both cases, action of the automorphisms of the fiber functor are interpreted in terms of “nice” (= étale!) geometric categories.

We end this discussion by noting a special case of the function-sheaf correspondence: let H be a (connected, separated) commutative group scheme of finite type over \mathbb{F}_{p^n} with $p \neq l$ and let \mathcal{E} be an l -adic local system on H . Let m^* denote the pull-back of the multiplication map $m : H \times_{\mathbb{F}_{p^n}} H \rightarrow H$. We call \mathcal{E} an l -adic *character sheaf* if $m^*(\mathcal{E}) \simeq p_1^*(\mathcal{E}) \otimes p_2^*(\mathcal{E})$. Here p_1 and p_2 are the projection morphisms $H \times_{\mathbb{F}_q} H \rightarrow H$.

Theorem 3.22 (Function-sheaf correspondence). *We have a natural bijection of sets*

$$\mathrm{Hom}_{\mathbf{AbGp}}(H(\mathbb{F}_q), \overline{\mathbb{Q}}_l) \leftrightarrow \{l\text{-adic character sheaves on } H\}.$$

Remark 3.23. A basic fact: any commutative group scheme over a field k is always an extension of an abelian variety by an affine group k -scheme (= a commutative k -Hopf algebra.) Therefore, we lose nothing by following the notation and thinking of H as an abelian variety. For example, the reader may think of H as the Picard group of a curve, to fix ideas.

The proof of theorem 3.22 uses the functoriality of the action of the absolute Frobenius⁴ on the stalks of l -adic local system \mathcal{E} (cf. [58] for a nice sketch of the proof.)

A speculation: Upon a fixed identification $\overline{\mathbb{Q}}_l \xrightarrow{\sim} \mathbb{C}$, one way to think of theorem 3.22 is⁵: all character maps ϕ of a given commutative Hopf algebra \mathbf{H} arise as appropriate l -adic character sheaves on \mathbf{H} . It remains a very interesting project to actually implement this philosophy when \mathbf{H} is the Connes-Kreimer Hopf algebra of Feynman graphs⁶ ([17] [18]) and ϕ a Feynman rule that assigns a generally divergent projective integral to a given Feynman graph and understanding its relationship with the Arapura motivic sheaves and renormalization as outlined in Marcolli’s paper [44].

4. Weil cohomology

This section presents the notion of a Weil cohomology, a set of functorial properties any good cohomology theory of smooth projective varieties should satisfy. The section very closely follows the first parts of Kleiman’s article in [39].

4.1. Main definition. Let k (resp. K) be fields of arbitrary characteristic (resp. zero characteristic.) Let \mathbf{SmProj}_k be the category of smooth projective schemes over k with smooth scheme maps as morphisms. Let \mathbf{GrVect} be the category of \mathbb{Z} -graded anticommutative K -algebras with K -algebra maps that preserve gradings as morphisms.

⁴Let X be a (connected) scheme over \mathbb{F}_{p^n} . By the “absolute Frobenius” we mean an \mathbb{F}_{p^n} -morphism which is the identity on X as a topological space and is the map $x \mapsto x^p$ on the structure sheaf \mathcal{O}_X .

⁵There are several subtleties involving base change and defining the absolute Frobenius, so the reader should take this statement with a grain of salt.

⁶In the Connes-Kreimer theory, the coproduct of \mathbf{H} gives a recursive formula for the factorization of loops in the prounipotent complex Lie group $G(\mathbb{C}) := \mathrm{Hom}(\mathbf{H}, \mathbb{C})$. The Birkhoff factorization of ϕ as algebra homomorphisms, needed for the counterterms, satisfy the Rota-Baxter identity.

Suppose $V \in \text{Obj}(\mathbf{GrVect})$. Therefore $V := \bigoplus_{i \in \mathbb{Z}} V_i$. Set $V_0 = K$. A Weil cohomology is a contravariant functor

$$H^* : \mathbf{SmProj}_k \longrightarrow \mathbf{GrVect}$$

with $H^*(X) = \bigoplus_{n \in \mathbb{Z}} H^n(X)$ that satisfies the following properties:

4.1.1. *Finiteness.* Each $H^i(X)$ has finite dimension and $H^i(X) = 0$ unless $0 \leq i \leq \dim X$.

4.1.2. *Poincaré duality.* There are two equivalent versions. The first version goes as follows: Let $r := \dim X$. For each such X , there is an isomorphism

$$H^{2r}(X) \xrightarrow{\sim} K$$

and a nondegenerate pairing

$$H^i(X) \times H^{2r-i}(X) \xrightarrow{\sim} K.$$

The second version of the Poincaré duality goes as follows:

$$\widehat{H^i(X)} \simeq H^{2r-i}(X),$$

where $\widehat{H^i(X)} = \text{Hom}(H^i(X), K)$.

4.1.3. *Künneth formula.* Let $X \times_k Y$ be the fiber product of two objects X and Y of \mathbf{SmProj}_k . Consider the following diagram of projections

$$\begin{array}{ccc} X \times_k Y & \xrightarrow{\pi_X} & X \\ \pi_Y \downarrow & & \\ Y & & \end{array}$$

The projections induce the isomorphism

$$H^*(X) \otimes_K H^*(Y) \simeq H^*(X \otimes Y).$$

NB: since we are working over fields, the usual “Tor” term is absent.

4.1.4. *Cycle maps.* Let $C^i(X)$ denote the free abelian group generated by closed irreducible subschemes of X of codimension i . In a Weil cohomology, there should be a group homomorphism

$$\gamma_X^i : C^i(X) \longrightarrow H^{2i}(X)$$

satisfying the following properties:

- (1) [Functoriality] There are two functorial conditions on cycle maps. The first (“pull back”) goes as follows: Let $r := \dim X$. Let $f : X \rightarrow Y$ be a morphism of \mathbf{SmProj}_k . Let $C^i(X)$ and $C^i(Y)$ denote the group of cycles of codimension i on X and Y respectively. The following diagram commutes:

$$\begin{array}{ccc} C^i(Y) & \xrightarrow{f^*} & C^i(X) \\ \gamma_Y^i \downarrow & & \downarrow \gamma_X^i \\ H^{2i}(Y) & \xrightarrow{f^*} & H^{2i}(X) \end{array}$$

(The top map is a pull-back on cycles.)

The second functorial condition (“push forward”) goes as follows: Let $s := \dim Y$. Now consider the pull back map defined above: if

$$H^{2i}(Y) \xrightarrow{f^*} H^{2i}(X),$$

then by Poincaré duality, we obtain the push-forward map

$$H^{2r-2i}(X) \xrightarrow{f_*} H^{2s-2i}(Y)$$

such that the following diagram commutes:

$$\begin{array}{ccc} C^{r-i}(X) & \xrightarrow{f_*} & C^{s-i}(X) \\ \gamma_X^{r-i} \downarrow & & \downarrow \gamma_Y^{s-i} \\ H^{2r-2i}(X) & \xrightarrow{f_*} & H^{2s-2i}(Y) \end{array}$$

(The map on the top denotes push-forward on cycles.)

- (2) [Multiplicativity] The map $\gamma_{X \times_k Y}^{i+j} = \gamma_X^i(Z) \otimes_K \gamma_Y^j(W)$. We see, by applying the definition of the cycle map and the Künneth formula that

$$\begin{aligned} Z \times_k W \in C^{i+j}(X \times_k Y) & \xrightarrow{\gamma_{X \times_k Y}^{i+j}} H^{2(i+j)}(X \times_k Y) \\ & \xrightarrow{\sim} \bigoplus_i H^{2(i+j)-l} H^{2(i+j)-l}(X) \otimes_K H^l(Y) \end{aligned}$$

so the multiplicativity axiom makes sense.

- (3) [Calibration] Let P be a point. Then the cycle map $\gamma_P : C^0 \rightarrow H^0(P)$ is the same as the inclusion of \mathbb{Z} in K .

4.1.5. Lefschetz theorems. There are two versions of the Lefschetz theorems, one ‘weak’ and the other ‘strong’. Let $h : W \rightarrow X$ be the inclusion of a smooth hyperplane section W of some smooth projective scheme X of dimension r .

- (1) [Weak Lefschetz] The induced map $h^* : H^i(X) \rightarrow H^i(W)$ is an isomorphism for $i \leq r-2$ and an injection for $i = r-1$.
(2) [Hard Lefschetz] Define the *Lefschetz operator*

$$\begin{aligned} L : H^i(X) & \rightarrow H^{i+2}(X) \\ Lx & \mapsto x \cdot \gamma_X^1(W). \end{aligned}$$

Then for $i \leq r$, $L^{r-i} : H^i(X) \rightarrow H^{2r-i}(X)$ is an isomorphism.

In summary: a Weil cohomology theory over k is a contravariant functor from the category **SmProj** $_k$ to the category **GrVect** with a nondegenerate pairing satisfying a Poincaré duality, with a Künneth formula, with cycle maps that are functorial with respect to pull-backs and push-forwards and satisfying the Lefschetz theorems.

4.2. Some properties of cycle maps. There are two important properties of cycle maps in the context of pure motives.

4.2.1. Intersection pairing and cup products. Let Z and W be two properly intersecting cycles in X and let $Z \cap W$ denote the intersection pairing. Let $\Delta : X \rightarrow X \times X$ be the diagonal map. Then $\gamma_X(Z \cap W) = \gamma_{X \times X} \circ \Delta^*(Z \times W)$ where Δ^* denotes the pull-back of $Z \times W$. We knew from the functoriality of cycle maps that this equals $\Delta^* \circ \gamma_{X \times X}(Z \times W) = \Delta^*(\gamma_X(Z) \otimes \gamma_X(W))$ by the multiplicativity of the cycle maps. So we get

$$\Delta^*(\gamma_X(Z) \otimes \gamma_X(W)) = \gamma_X(Z) \cdot \gamma_X(W)$$

by applying pull-back again.

4.2.2. *Correspondences as operators.* Let $r := \dim X$. Observe from the definitions that

$$\begin{aligned} H^i(X \times Y) &= H^i(X) \otimes H^i(Y) \\ &= \widetilde{H^{2r-i}(X)} \otimes H^i(Y) \\ &= \text{Hom}(H^{2r-i}(X), K) \otimes H^i(Y) \\ &= \text{Hom}(H^{2r-i}(X), H^i(Y)). \end{aligned}$$

Therefore, we can view each element of $H^i(X \times Y)$ as an *operator* from $H^{2r-i}(X)$ to $H^i(Y)$. Such operators are called *correspondences* for the obvious reasons of section 5.

4.3. The Standard Conjectures. The original formulation goes back to Grothendieck's Tata lecture [33]. I follow Murre's lecture [50].

Rescaling the indices in the hard Lefschetz theorem, the isomorphism can be equivalently written as $L^{r-i} : H^{r-j}(X) \rightarrow H^{r+j}$ for $r = \dim X$. Using the map L we can define a unique linear operator Λ which makes the following: For $0 \leq j \leq r-2$ the diagram

$$\begin{array}{ccc} H^{r-j}(X) & \xrightarrow{L^j} & H^{r+j}(X) \\ \Lambda \downarrow & & \downarrow L \\ H^{r-j-2}(X) & \xrightarrow{L^{j+2}} & H^{r+j+2}(X) \end{array}$$

commutes, and for $2 \leq j \leq r$

$$\begin{array}{ccc} H^{r-j+2}(X) & \xrightarrow{L^{j-2}} & H^{r+j-2}(X) \\ L \uparrow & & \uparrow \Lambda \\ H^{r-j}(X) & \xrightarrow{L^j} & H^{r+j+2}(X) \end{array}$$

the diagram commutes. We have

Conjecture 4.1 (B(X)). *The operator Λ is algebraic. That is*

$$\Lambda(Z) = \gamma_{X \times X}^*(Z) \text{ for all } Z \in \text{CH}^i(X \times X) \otimes \mathbb{Q}.$$

By the hard Lefschetz theorem, we have the isomorphisms $L^{r-i} : H^i(X) \rightarrow H^{2r-i}(X)$. Let

$$P^i(X) := \ker\{L^{r-i-1} : H^i(X) \rightarrow H^{2r-i+2}(X)\}$$

be the set of *primitive elements* of $H^i(X)$. Let $x, y \in C^i(X) \cap P^{2i}(X)$ for $i \leq r/2$.

Conjecture 4.2 (Hdg(X)). *Consider the pairing*

$$(x, y) \mapsto (-1)^i \langle L^{r-2i}(x), y \rangle$$

where $\langle -, - \rangle$ denotes the cup product. This pairing is positive definite.

It is known that conjecture 4.2 is true for étale cohomology in characteristic zero. (The fact that étale cohomology is a Weil cohomology and that, in particular, the hard and weak Lefschetz theorems hold, requires some work to show.)

Conjectures 4.1 and 4.2 taken together imply the Weil conjectures, among other things. One cares about the standard conjectures because they capture something intrinsic about the cycles independent of the Weil cohomology that the cycle map maps to. For example, we have the following, due to Grothendieck:

Theorem 4.3. *Assume conjectures 4.1 and 4.2. Then the Betti numbers $\dim H^i(X)$ is independent on the choice of X .*

5. Classical motives

As noted in the introduction, one of the central goal of theory of motives is *linearization* of the category of algebraic varieties (or schemes) over an arbitrary base field (or ring). In fact, what we are after is obtaining something stronger than additivity (which is the “coarsest” form of linearization)– we want an abelianization of the category of algebraic varieties! One way to abelianize the category of varieties is to replace morphisms in this category by *correspondences*.

5.1. Correspondences of curves. By a curve, we mean a nonsingular variety of dimension 1. Let us start with the case of correspondences of curves over the complex numbers (classical theory due to Castelnuovo and Severi).

Definition 5.1 (section 2.5 of [31]). Let C and C' be two curves. A *correspondence* of degree d between C and C' is a holomorphic map

$$T : C \rightarrow C', p \mapsto T(p)$$

where $T(p)$ a divisor of degree d on C . Furthermore, the *curve of correspondence* between C and C' is given by the curve

$$D = \{(p, q) | q \in T(p)\} \subset C \times C'.$$

Let $D \subset C \times C'$ be a curve. Then the correspondence associated to D is defined by

$$T(p) = \pi_p^*(D) \in \text{Div}(C')$$

where $\pi_p^* : C' \rightarrow C \times C'$ is given by $q \mapsto (p, q)$.

Remark 5.2. The “converse” part of definition (5.1) states that we should view correspondences as formal linear combinations of all 0-dimensional subvarieties on C' .

5.2. Equivalence relations on algebraic cycles. All through we will work with smooth projective varieties over a field k . We denote the category of such objects as \mathbf{SmProj}_k . Furthermore, most of the time, we’d restrict ourselves to such varieties that are also irreducible. Much of the material in this section (and the next) is from the excellent reviews by Murre [50] and Scholl [57]. Recall

Definition 5.3. An *algebraic cycle* Z on X of codimension i is a formal linear combination of closed irreducible subvarieties Z_α of codimension i ; thus,

$$Z = \sum_{\alpha} n_{\alpha} Z_{\alpha}, n_{\alpha} \in \mathbb{Z}.$$

Algebraic cycles (or cycles, for short) of a given codimension i in a given smooth projective X form an abelian group with respect to formal sums. Denote such a group as $\mathcal{Z}^i(X)$. As it stands, $\mathcal{Z}^i(X)$ is very “big”, so we want to obtain a smaller group by taking a quotient of $\mathcal{Z}^i(X)$ by equivalence relations on cycles.

There are several equivalence relations that one can impose on cycles, but we’d restrict ourselves to the following.

- (1) (Rational equivalence) Let Z be a cycle of a fixed codimension i on X irreducible of dimension d . We say Z is rationally equivalent to zero and write $Z \sim_{\text{rat}} 0$ if there exists (Y_α, f_α) with Y_α irreducible of codimension $i-1$ and f_α a rational function on Y_α such that $\sum_\alpha v_{Y_\alpha}(f_\alpha)Y_\alpha = Z$. (Here v denotes the valuation of the rational function.)
- (2) (Homological equivalence) Once again, let Z be a cycle of codimension i . Let $H_B^*(X, \mathbb{Q})$ be the Betti cohomology of X . Then there exists a map (in fact a group homomorphism!) $Z \mapsto [Z]$ where $[Z]$ is a cohomology class in $H_B^{2i}(X, \mathbb{Q})$. We say that Z is homologically equivalent to zero and write $Z \sim_{\text{hom}} 0$ if the image of Z under this map vanishes. Furthermore, cycle maps are general gadgets that associate i -cycles to $2i$ -cohomology classes in any good cohomology theory (not just Betti) and satisfying certain functoriality conditions, a notion formalized by Weil cohomology, see 4.1. Homological equivalence is a notion that is to hold true in *any* (and perhaps more strongly in *all*) of these good cohomology theories.
- (3) (Numerical equivalence) Let Z as in above. We say Z is numerically equivalent to zero and write $Z \sim_{\text{num}} 0$ if the intersection number $\#(Z \cap Z') = 0$ for all Z' cycles of the same codimension as Z . (To spell things out a bit more: the definition says that $\#(Z_\alpha \cap Z'_\beta) = 0$ for all α, β and where $Z = \sum_\alpha n_\alpha Z_\alpha$ and $Z' = \sum_\beta n_\beta Z'_\beta$.)

It can be seen that the relations given above are indeed equivalence relations. From now on, we would write \sim to mean any of three relations above unless we fix the equivalence relation, in which case we would specify it as such.

Definition 5.4. Fix a smooth projective variety X . The group

$$C_{\sim}^i(X) := \mathcal{Z}^i(X) / \mathcal{Z}_{\sim}^i(X)$$

where $\mathcal{Z}_{\sim}^i(X)$ denotes the subgroup of cycles identified under the equivalence relation \sim . When X is irreducible and of dimension d , $C_{\sim}(X) := \bigoplus_{i=0}^d C_{\sim}^i(X)$. When $\sim = \sim_{\text{hom}}$, $C_{\sim}^i(X)$ is called the *Chow group* and is denoted as $CH^i(X)$. Notation: we write $C_{\sim}(X)_{\mathbb{Q}}$ to mean $C_{\sim}(X) \otimes \mathbb{Q}$.

Remark 5.5. Notice that (1) is just a generalization of the notion of divisors on curves and that of linear equivalence. The relation (3) is trickier because the intersection pairing $Z \cap Z'$ may not be always defined. In this case, one resorts to various *moving lemmas*. Perhaps the most problematic is the equivalence relation (2) since the definition is based on a choice of a cohomology theory; however assuming one of the implications of the Standard conjectures along with Jannsen's result help solving this problem (see theorem 5.13.)

An important observation is

$$C_{\sim_{\text{rat}}}^i(X) \subseteq C_{\sim_{\text{hom}}}^i(X) \subseteq C_{\sim_{\text{num}}}^i(X) \subset C_{\sim}^i(X)$$

going from the finest to the coarsest equivalence relations on cycles. An extremely important conjecture in the theory of classical motives is

Conjecture 5.6 (Fundamental Conjecture D(X)).

$$C_{\sim_{\text{hom}}}^i(X) = C_{\sim_{\text{num}}}^i(X).$$

The conjecture 5.6 follows from the Standard Conjectures of Grothendieck B(X) (conjecture 4.1) and Hdg(X) (conjecture 4.2).

5.3. Pure motives. Just like in the case of curves (see definition (5.1)), we view correspondences on a smooth projective variety as formal linear combinations of closed irreducible subvarieties, i.e, as elements in *group of correspondences* $C_{\sim}(X \times Y)_{\mathbb{Q}} := \text{Corr}_{\sim}(X, Y)$.

Let $f \in \text{Corr}_{\sim}(X, Y)$ and $g \in \text{Corr}_{\sim}(Y, Z)$. The composition of correspondences is given by

$$g \bullet f = \text{pr}_{XZ}((f \times Z) \cdot (X \times g)).$$

A *emphprojector* $p \in \text{Corr}_{\sim}^0(X, X)$ is an idempotent $p \bullet p = p$.

Definition 5.7 (the category of effective pure motives). Fix an equivalence relation on cycles \sim . The category of effective pure motives over a field k is denoted as $\mathbf{Mot}_{\sim}^+(k)$ has

- Objects: (X, p) where $X \in \text{Obj}(\mathbf{SmProj}_k)$ and p a projector.
- Morphisms: Let $M = (X, p)$ and $N = (Y, q)$. Then $\text{Hom}_{\mathbf{Mot}_{\sim}^+(k)}(M, N) := q \bullet \text{Corr}_{\sim}^0(X, Y) \bullet p$.

Elements of $\text{Hom}_{\mathbf{Mot}_{\sim}^+(k)}(M, N)$ are simply the composition of correspondences $X \rightarrow X \rightarrow Y \rightarrow Y$. By modifying the definition 5.7 to allow for Tate twists, we get the category of virtual pure motives over k .

Definition 5.8 (the category of virtual pure motives). The category of virtual pure motives over a field k is denoted as $\mathbf{Mot}_{\sim}(k)$ has

- Objects: (X, p, m) where X and p as above and $m = \dim X$.
- Morphisms: Let $M = (X, p, m)$ and $N = (Y, q, n)$ (and $n = \dim Y$). Then $\text{Hom}_{\mathbf{Mot}_{\sim}(k)}(M, N) := q \bullet \text{Corr}_{\sim}^{n-m}(X, Y) \bullet p$.

Some trivial and distinguished pure motives: $\mathbf{0} := (\text{Spec } k, \text{id}, 0)$, the motive of a point, $\mathbf{L} := (\text{Spec } k, \text{id}, -1)$ the Lefschetz motive and $\mathbf{T} := (\text{Spec } k, \text{id}, -1)$ the Tate motive.

5.4. The motives functor. Recall that if $\psi : Y \rightarrow X$ is a scheme morphism, then the *graph* of ψ , $\Gamma_{\psi} := (\psi \times \text{id}_Y) \circ \Delta_Y$ where id_Y and Δ_Y are respectively the identity and diagonal maps on Y . That is, the graph is simply the composition of maps in the sequence

$$Y \xrightarrow{\Delta_Y} Y \times Y \xrightarrow{\psi \times \text{id}_Y} X \times Y.$$

The transpose Γ_{ψ}^t is obtained by exchanging the factors of $X \times Y$.

Definition 5.9. The functor

$$\mathbf{m}_{\sim} : \mathbf{SmProj}_k^{op} \rightarrow \mathbf{Mot}_{\sim}(k)$$

is defined in the following way: $\mathbf{m}_{\sim}(X) := (X, \Delta_X, 0)$ and $\mathbf{m}_{\sim}(\psi) = \Gamma_{\psi}^t : \mathbf{m}_{\sim}(Y) \rightarrow \mathbf{m}_{\sim}(X)$ and where Γ_{ϕ}^t is the transpose of the graph of $\phi : X \rightarrow Y$ a morphism of smooth projective schemes.

Let $e \in X$ be a point. Take $\pi_0 = e \times X$ and $\pi_{2d} = X \times e$ where $d := \dim X$ and X is irreducible. Write $\mathbf{m}_{\sim}^0(X) := (X, \pi_0, 0)$ and $\mathbf{m}_{\sim}^{2d}(X) := (X, \pi_{2d}, 0)$. We can show that

$$\mathbf{m}_{\sim}^{2d}(X) \simeq (\text{Spec } k, \text{id}, -d)$$

and use this to show the fundamental fact

$$\mathbf{L} \simeq \mathbf{m}_{\sim}^2(\mathbb{P}^1, \mathbb{P}^1 \times e, 0)$$

where \mathbb{P}^1 is the projective line over k and e a point in \mathbb{P}^1 .

The category of (virtual) pure motives $\mathbf{Mot}_{\sim}(k)$ has the following properties.

- $\mathbf{Mot}_{\sim}(k)$ is an additive category. That is, $\text{Hom}_{\mathbf{Mot}_{\sim}(k)}(M, N)$ are abelian groups and \oplus exists: $M \oplus N := (X \amalg Y, p \amalg q, m)$ for $\dim X = \dim Y = m$.
- $\mathbf{Mot}_{\sim}(k)$ is a pseudoabelian category; it is additive with a well-defined image of p .
- There is a tensor structure on $\mathbf{Mot}_{\sim}(k)$:

$$M \otimes N := (X \times Y, p \times q, m + n).$$

- There is a multiplicative structure on $\mathbf{Mot}_{\sim}(k)$:

$$m_X : \mathbf{m}_{\sim}(X) \otimes \mathbf{m}_{\sim}(X) \simeq \mathbf{m}_{\sim}(X \times X) \xrightarrow{\mathbf{m}_{\sim}(\Delta)} \mathbf{m}_{\sim}(X).$$

Let k be a field and \bar{k} be its algebraic closure. Fix an equivalence relation \sim on the cycles of X . Let M be a virtual pure motives of a smooth projective variety X , i.e., $M = \mathbf{m}_{\sim}(X)$.

Definition 5.10 (Realization of a pure motive). Define the realization of a motive M as the functor

$$\mathbf{Mot}_{\sim}(k) \xrightarrow{\text{real}} \mathbf{GrVect} \mathbf{m}_{\sim}(X) \mapsto H^*(X_{\bar{k}}, \mathbb{Q})$$

for a Weil cohomology functor $H^*(-, \mathbb{Q})$. We say that a motive can be realized if there exists at least one Weil cohomology for which the functor **real** is exact and faithful for all adequate relations \sim .

The reader may set $H^*(X_{\bar{k}}, \mathbb{Q})$ as algebraic de Rham cohomology as to fix ideas.

Grothendieck's original conception of motives was such that the following diagram of functors commute *for all adequate equivalence relations \sim and all Weil cohomologies $H^*(X, \mathbb{Q})$* :

$$\begin{array}{ccc} \mathbf{Mot}_{\sim}(k) & \xlongequal{\quad} & \mathbf{Mot}_{\sim}(k) \\ \mathbf{m}_{\sim}(-) \uparrow & & \downarrow \text{real}(-) \\ \mathbf{SmProj}_k^{op} & \xrightarrow{H^*(-, \mathbb{Q})} & \mathbf{GrVect} \end{array}$$

There are three categories of motives of particular interest to us: the category of *Grothendieck motives* when the equivalence relation is homological equivalence $\mathbf{Mot}_{\text{hom}}(k)$, the category of *Chow motives* when the equivalence relation is homological equivalence $\mathbf{Mot}_{\text{rat}}(k)$ and the category of *numerical motives* when the equivalence relation is numerical equivalence $\mathbf{Mot}_{\text{num}}(k)$. We have the following “fundamental theorem of pure motives”:

Theorem 5.11. *$\mathbf{Mot}_{\text{hom}}(k)$ can be realized for all Weil cohomologies and is neutral Tannakian with the fiber functor **real**.*

This is the reason why number theorists often use the term “motive of X ” when they mean “the l -adic cohomology of X with \mathbb{Q}_l -coefficients and a continuous G_k action on it”. (The action of G_k is through the motivic Galois group of X , cf. section 5.5.)

Recall the yoga of Tannakian formalism of section 3. Deligne defines the dimension of an object in a rigid tensor category \mathbf{C} in the following way:

Definition 5.12 (dimension). Let \mathbb{I} be the unit object in \mathbf{C} . Let $M \in \text{Obj}(\mathbf{C})$ be an arbitrary object. Let $f \in \text{End } M$ and trace of f on M is defined by the composition of maps

$$\mathbb{I} \xrightarrow{\delta} \widehat{M} \otimes M \xrightarrow{t} M \otimes \widehat{M} \xrightarrow{\text{ev}} \mathbb{I}$$

and denote as $\text{tr } f_M$. (The map t exchanges the factors in $\widehat{M} \otimes M$.) The dimension of M is $\underline{\dim} M := \text{tr } 1_M$.

A result of Deligne show that dimension is always positive: \mathbf{C} rigid tensor \iff for all $M \in \text{Obj}(\mathbf{C})$, $\underline{\dim} M \in \mathbb{N}$.

Furthermore for $\mathbf{Mot}_{\text{hom}}(k)$, Deligne shows that this definition of dimension coincides (through the Weil conjectures) with our usual notion of cohomological dimension:

$$\underline{\dim} M = \dim H^*(X, \mathbb{Q}) := \bigoplus_i \dim H^i(X, \mathbb{Q})$$

where $H^i(X, \mathbb{Q})$ is a realization of the motive M . The Standard Conjectures imply that $\dim M$ is independent of the choice of the Weil cohomology.

Theorem 5.13 (Jannsen). *The category $\mathbf{Mot}_{\text{num}}(k)$ is semisimple and abelian.*

5.5. Motivic Galois groups. The Tannakian formalism along with theorem 5.11 gives

$$\mathbf{Rep}_k(\text{Aut}^{\otimes} H_B^*(-, \mathbb{Q})) \xrightarrow{\sim} \mathbf{Mot}_{\text{hom}}(k) \xrightarrow{\text{real}(-)} \mathbf{GrVect} \circlearrowleft G_{\text{mot}}$$

where $G_{\text{mot}} := \text{Aut}^{\otimes} H_B^*(-, \mathbb{Q})$ is the *motivic Galois group*. It acts on the image of the realization functor and is a proalgebraic group. (It is proreductive assuming Jannsen's theorem 5.13 and the Conjecture D(X) 5.6.)

If $\mathbf{Mot}_{\sim}(k)$ is generated by $\mathbf{m}_{\sim}(\text{Spec } E)$ where E is a finite extension of k then $G_{\text{mot}} = G_k$. If $\mathbf{Mot}_{\sim}(k)$ is generated by the Lefschetz motive \mathbf{L} , then $G_{\text{mot}} = \mathbb{G}_m$.

6. Mixed motives

Pure motives are motives of smooth projective varieties. In most physical applications of the theory of motives, most notably in the case of the motives of hypersurfaces associated to Feynman graphs in perturbative quantum field theory, working with smooth projective varieties is too restrictive. In fact, one knows that the projective hypersurfaces obtained from the parametric form of Feynman integrals are typically singular, and this already leaves the world of pure motives. Moreover, since the parametric Feynman integral (see the appendix) is computed over a domain of integration with boundary, what is involved from the motivic point of view is really a relative cohomology of the hypersurface complement, relative to a normal crossings divisor that contains the boundary of the domain of integration. This is a second reason why it is mixed motives and not pure motives that are involved, since these are the natural environment where long exact cohomology sequences and relative cohomologies live. Unfortunately, from the mathematical point of view, the theory of mixed motives is far more complicated than that of pure motives. At present, one only has a triangulated category of mixed motives, with various equivalent constructions due to Voevodsky, Levine, and Hanamura. Only in very special cases, such as mixed Tate motives over a number field, it is possible to construct and abelian category. While we are not going to give any details on the construction of the triangulated category of mixed motives, we recall the relevant cohomological properties and some properties of their analytic

counterpart, mixed Hodge structures. We also recall some preliminary notions about triangulated categories to aid the reader in her further study.

6.1. Derived and triangulated categories.⁷ This material is taken from Dimca [24]. Let \mathbf{Ab} be an abelian category. Let $C(\mathbf{Ab})$ be the category of complexes in \mathbf{Ab} . (The reader unfamiliar with homological algebra should immediately set \mathbf{Ab} to be the category of R -modules \mathbf{Mod}_R .) $C(\mathbf{Ab})$ contains three important full subcategories $C^\bullet(\mathbf{Ab})$:

(1) $C^+(\mathbf{Ab})$ of complexes bounded on the left:

$$\cdots \longrightarrow 0 \longrightarrow \cdots \longrightarrow A^{-1} \longrightarrow A^0 \longrightarrow \cdots$$

(2) $C^-(\mathbf{Ab})$ of complexes bounded on the right:

$$\cdots \longrightarrow A^0 \longrightarrow A^1 \longrightarrow \cdots \longrightarrow 0 \longrightarrow \cdots$$

(3) $C^b(\mathbf{Ab})$ the full subcategory of complexes bounded both on the right and the left.

Let X^\bullet, Y^\bullet be two complexes in $C^\bullet(\mathbf{Ab})$. Call the morphism $u : X^\bullet \rightarrow Y^\bullet$ a *quasi-isomorphism* if at the level of cohomology $H^k(u) : H^k(X^\bullet) \rightarrow H^k(Y^\bullet)$ is an isomorphism for all k .

Define a *shift automorphism* on complexes $T : C^\bullet \rightarrow C^\bullet$ as

$$\begin{aligned} (X[n])^r &:= X^{n+r}, \\ d_{T(X^\bullet)}^s &:= -d_{X^\bullet}^{s+1} \end{aligned}$$

for the complex

$$A^\bullet : \cdots \longrightarrow A^{m-1} \xrightarrow{d^{m-1}} A^m \xrightarrow{d^m} A^{m+1} \xrightarrow{d^{m+1}} \cdots$$

Let $u : X^\bullet \rightarrow Y^\bullet$ be a morphism of complexes in $C^\bullet(\mathbf{Ab})$. The *mapping cone* of the morphism u is the complex in $C^\bullet(\mathbf{Ab})$ given by

$$C_u^\bullet = Y^\bullet \oplus (X^\bullet[1]).$$

This gives rise to the *standard triangle* for a morphism u

$$T_u : X^\bullet \xrightarrow{u} Y^\bullet \xrightarrow{q} C_u^\bullet \xrightarrow{p} X^\bullet[1],$$

where q (resp. p) is the inclusion (resp. projection) morphisms. The standard triangle gives rise to long exact sequences in cohomology.

An important construction is that of a *homotopic* category⁸. This is an additive category $K^\bullet(\mathbf{Ab})$ with

- $\text{Obj}(K^\bullet(\mathbf{Ab})) = \text{Obj}(C^\bullet(\mathbf{Ab}))$,
- $\text{Hom}_{K^\bullet(\mathbf{Ab})}(X^\bullet, Y^\bullet) = \text{Hom}(X^\bullet, Y^\bullet) / \sim$ where \sim is homotopy equivalence: $u \sim v \implies H^k(u) = H^k(v)$.

A family of triangles in $K^\bullet(\mathbf{Ab})$ is *distinguished* if they are isomorphic to a standard triangle for some morphism u .

In a homotopic category $K^\bullet(\mathbf{Ab})$, distinguished triangles satisfy a list of four properties referred to as TR1–TR4 in the literature. We will not repeat them here (see, for example, proposition 1.2.4 of [24]). It suffices to say that TR1–TR4

⁷Because of the rather dry and technical nature of the precise definitions, I have tried to give a more general idea about triangulated and derived categories; the following are a collection of concepts as opposed to a list of formal definitions and properties.

⁸The reader is invited to compare this with the definition of a Grothendieck group in 2.2.

guarantees nice functorial properties of distinguished triangles, compatible with homotopy.

A *triangulated category* is an additive category \mathcal{A} with a shift self-equivalence T , with $X[1] = TX$ and with a collection of distinguished triangles \mathcal{T} that satisfy TR1–TR4 in the original definition due to Verdier. A *derived category* $D^\bullet(\mathbf{Ab})$ of an abelian category \mathbf{Ab} is a triangulated category obtained from $K^\bullet(\mathbf{Ab})$ by localization with respect to the multiplicative system of quasi-isomorphisms in $K^\bullet(\mathbf{Ab})$.

6.2. Bloch–Ogus cohomology. The material for this section is taken from the seminal paper of Bloch–Ogus [10]. Bloch–Ogus cohomology is the mixed motives counterpart of Weil cohomology in the pure motives case, and as such, is a universal cohomology theory for schemes of a much general type.

Let \mathbf{Sch}_k be the category of schemes of finite type over the field k . By \mathbf{Sch}_k^* I mean the category with

- Objects: $Y \hookrightarrow X$ closed immersions in $X \in \text{Obj}(\mathbf{Sch}_k)$ and
- Morphisms: $\text{Hom}_{\mathbf{Sch}_k^*}((Y \hookrightarrow X), (Y' \hookrightarrow X'))$ being the following commutative cartesian squares:

$$\begin{array}{ccc} Y & \xrightarrow{\subseteq} & X \\ f_Y \downarrow & & \downarrow f_X \\ Y' & \xrightarrow{\subseteq} & X' \end{array}$$

Definition 6.1 (cohomology with supports). A *twisted cohomology theory with supports* is a sequence of contravariant functors

$$\begin{aligned} H^i : \mathbf{Sch}_k^* &\longrightarrow \mathbf{AbGp} \\ (Y \hookrightarrow X) &\mapsto \bigoplus_i H_Y^i(X, n), \end{aligned}$$

satisfying the following:

- (1) For $Z \subseteq Y \subseteq X$, there is a long exact sequence

$$\cdots \longrightarrow H_Z^i(X, n) \longrightarrow H_Y^i(X, n) \longrightarrow H_{Y \setminus Z}^i(X \setminus Z, n) \longrightarrow H_Z^{i+1}(X, n) \longrightarrow \cdots.$$

- (2) Let

$$\begin{aligned} f : (Y \hookrightarrow X) &\longrightarrow (Y' \hookrightarrow X'), \\ g : (Z \hookrightarrow X) &\longrightarrow (Z' \hookrightarrow X') \end{aligned}$$

and $k : (Y \setminus Z \hookrightarrow X \setminus Z) \rightarrow (Y' \setminus Z' \hookrightarrow X' \setminus Z')$ be the induced arrow.

Let $h : (Z \hookrightarrow X) \rightarrow (Z' \hookrightarrow X')$. Then the following diagram commutes:

$$\begin{array}{ccccccc} \rightarrow H_Z^i(X, n) & \longrightarrow & H_Y^i(X, n) & \longrightarrow & H_{Y \setminus Z}^i(X \setminus Z, n) & \longrightarrow & H_Z^{i+1}(X, n) \rightarrow \\ H^*(h) \uparrow & & H^*(f) \uparrow & & H^*(k) \uparrow & & H^*(g) \uparrow \\ \rightarrow H_{Z'}^i(X', n) & \longrightarrow & H_{Y'}^i(X', n) & \longrightarrow & H_{Y' \setminus Z'}^i(X' \setminus Z', n) & \longrightarrow & H_{Z'}^{i+1}(X', n) \rightarrow \end{array}$$

- (3) [Excision] Let $(Z \hookrightarrow X) \in \text{Obj}(\mathbf{Sch}_k^*)$ and $(U \hookrightarrow X)$ be open in X containing Z . Then

$$H_Z^i(X, n) \xrightarrow{\sim} H_Z^i(U, n).$$

Dually we have

Definition 6.2. [Homology with supports] A *twisted homology with supports* is a sequence of covariant functors

$$H_* : \mathbf{Sch}_{k*} \longrightarrow \mathbf{AbGp}$$

where the category \mathbf{Sch}_{k*} has objects of \mathbf{Sch}_k and the morphisms are proper morphisms of \mathbf{Sch}_k satisfying:

- (1) H_* is a presheaf in étale topology. If $\alpha : X' \rightarrow X$ is an étale morphism there exists morphisms $\alpha^* : H_i(X, n) \rightarrow H_i(X', n)$.
- (2) Let $\alpha : Y' \rightarrow Y$ and $\beta : X' \rightarrow X$ be étale. Let $f : X \rightarrow Y$ and $g : X' \rightarrow Y'$ be proper and consider the following cartesian square:

$$\begin{array}{ccc} X' & \xrightarrow{\beta} & X \\ g \downarrow & & \downarrow f \\ Y' & \xrightarrow{\alpha} & X' \end{array}$$

Then the following square commutes:

$$\begin{array}{ccc} H_i(X', n) & \xleftarrow{\beta^*} & H_i(X, n) \\ H_i(g, n) \downarrow & & \downarrow H_i(f, n) \\ H_i(Y', n) & \xleftarrow{\alpha^*} & H_i(Y, n) \end{array}$$

- (3) Let $i : Y \hookrightarrow X$ be a closed immersion and $\alpha : X \setminus Y \hookrightarrow X$ be an open immersion. Then there exists a long exact sequence

$$\cdots \rightarrow H_i(Y, n) \xrightarrow{i^*} H_i(X, n) \xrightarrow{\alpha^*} H_i(X \setminus Y, n) \rightarrow H_{i-1}(Y, n) \rightarrow \cdots.$$

- (4) Let $f : X' \rightarrow X$ be proper and $Z = f(Z')$ for $Z \hookrightarrow X$. Let $\alpha : X' \setminus f^{-1}(Z) \hookrightarrow X' \setminus Z'$. Then the following square commutes:

$$\begin{array}{ccccccc} \rightarrow H_i(Z', n) & \longrightarrow & H_i(X', n) & \longrightarrow & H_i(X' \setminus Z', n) & \longrightarrow & H_{i-1}(Z', n) \rightarrow \\ f_* \downarrow & & f_* \downarrow & & f_* \alpha^* \downarrow & & f_* \downarrow \\ \rightarrow H_i(Z, n) & \longrightarrow & H_i(X, n) & \longrightarrow & H_i(X \setminus Z, n) & \longrightarrow & H_{i-1}(Z, n) \rightarrow \end{array}$$

Definition 6.3 (Poincaré duality with supports). A *Poincaré duality theory for schemes of finite type with supports* is a twisted cohomology theory H^* with

- (1) For all $Y \hookrightarrow X \in \text{Obj}(\mathbf{Sch}_k^*)$ there is a pairing

$$H_i(X, n) \times H_Y^j(X, n) \rightarrow H_{i+j}(Y, m+n).$$

- (2) If $Y \hookrightarrow X \in \text{Obj}(\mathbf{Sch}_k^*)$ and $(\beta \hookrightarrow \alpha) : (Y' \hookrightarrow X') \rightarrow (Y \hookrightarrow X)$ an étale morphism in \mathbf{Sch}_k^* , then for $a \in H_Y^j(X, n)$ and $z \in H_i(X, m)$,

$$\alpha^*(a) \cap \alpha^*(Z) = \beta^*(a \cap z).$$

- (3) [Projection] Let $f : (Y_1 \hookrightarrow X_1) \rightarrow (Y_2 \hookrightarrow X_2)$ proper. Then for $a \in H_{Y_2}^i(X_2, n)$ and $z \in H_i(X_1, m)$,

$$H^i(f_X)(z) \cap a = H_i(f_Y)(z \cap H^i(f)(a)).$$

- (4) [Fundamental class] Let $X \in \text{Obj}(\mathbf{Sch}_k)$ be irreducible and of dimension d . There exists a global section η_X of $H_{2d}(X, d)$ such if $\alpha : X' \rightarrow X$ is étale, then $\alpha^* \eta_X = \eta_{X'}$.

- (5) [Poincaré duality] Let $X \in \text{Obj}(\mathbf{Sch}_k)$ be smooth and of dimension d . Let $Y \hookrightarrow X$ be a closed immersion. Then

$$H_Y^{2d-i}(X, d-n) \xrightarrow{\cap \eta_X} H_i(Y, n)$$

is an isomorphism.

The main theorem of [10] is

Theorem 6.4. *Given a Poincaré duality theory, an étale morphism f_X and the commutative square*

$$\begin{array}{ccc} Z' & \longrightarrow & X' \\ f_Z \downarrow & & \downarrow f_X \\ Z & \xrightarrow{\subset} & X \end{array}$$

the following diagram commutes:

$$\begin{array}{ccc} H_{Z'}^i(X', n) & \xrightarrow{\cap \eta_{X'}} & H_{2d-i}(Z', d-n) \\ H^*(f) \uparrow & & \uparrow f_Z^* \\ H_Z^i(X, n) & \xrightarrow{\cap \eta_X} & H_{2d-i}(Z, d-n) \end{array}$$

6.3. Hodge structures. Let us quickly review the basic definitions of Hodge theory.

Definition 6.5 (Pure Hodge structure). A pure Hodge structure of weight m on a finite dimensional vector space V is a decreasing filtration

$$\cdots \subset F^{p+1}V_{\mathbb{C}} \subset F^pV_{\mathbb{C}} \subset \cdots$$

with $V_{\mathbb{C}} := V \otimes_{\mathbb{R}} \mathbb{C}$ and satisfying the *Hodge decomposition*

$$V_{\mathbb{C}} = \bigoplus_{p+q=m} V^{p,q},$$

where $V^{p,q} = F^p \cap \overline{F^q V_{\mathbb{C}}}$ and $\overline{}$ denotes the conjugate filtration.

Hodge structures form a category **Hodge** with morphisms of vector spaces compatible with the filtration F making up the hom set. The category **Hodge** is a tensor category with the evident tensor product. Furthermore, by taking formal differences of Hodge structures $[H] - [H']$, we obtain the Grothendieck ring of Hodge structures $K(\mathbf{Hodge})$.

Let $A \subset \mathbb{R}$. An A -mixed Hodge structure consists of the following data:

- (1) An A -module of finite type V_A .
- (2) An increasing filtration called the *weight filtration*

$$\cdots \subset W_n \subset W_{n+1} \subset \cdots$$

of $A \otimes \mathbb{Q}$ -module $V_A \otimes \mathbb{Q}$.

- (3) A decreasing filtration called the *Hodge filtration*

$$\cdots F^{p+1}V_{\mathbb{C}} \subset F^pV_{\mathbb{C}} \subset \cdots$$

where $V_{\mathbb{C}} \otimes V_A \otimes \mathbb{C}$.

- (4) A *graded weight j factor* $\text{gr}_j^W(V_A) := (W_j/W_{j-1}) \otimes \mathbb{C}$ with a pure Hodge structure induced by the filtration F and \overline{F} on $V_{\mathbb{C}}$.

Mixed Hodge structures too form a category **MHodge** with morphisms $V_A \rightarrow V_{A'}$ A -module homomorphisms that are compatible with the Hodge and weight filtrations. The following theorem is of great importance:

Theorem 6.6 (Deligne). *The category **MHodge** is abelian with kernels and cokernels with induced filtrations.*

6.4. Mixed Tate motives over a number field. Mixed Tate motives can be defined as the triangulated subcategory of the triangulated category of mixed motives generated by the Tate objects $\mathbb{Q}(n)$. Over a number field, however, it is possible to obtain a nicer category in the following way.

Let k be a number field. Let $\mathbb{Q}(1)$ be the pure Tate motive. The category of mixed Tate motives over k , denoted as **MTMot** _{k} is constructed in the following way. Consider the simple objects $\mathbb{Q}(n)$ and assume that $\mathbb{Q}(a)$ and $\mathbb{Q}(b)$ are isomorphic for $a \neq b$ and that any simple object of **MTMot** _{k} is isomorphic to some $\mathbb{Q}(\cdot)$. Also consider the groups

$$\mathrm{Ext}_{\mathbf{MTMot}_k}^i(\mathbb{Q}(0), \mathbb{Q}(n))$$

and assume that they vanish for $i > 1$. An important result of Borel identifies these extensions:

Theorem 6.7 (Borel). *$\mathrm{Ext}_{\mathbf{MTMot}_k}^1(\mathbb{Q}(0), \mathbb{Q}(n)) \simeq K_{2n-1}(k) \otimes \mathbb{Q}$ where $K_{2n-1}(k)$ is the Quillen K -theory of the field k .*

Now $K_{2n-1}(k) \otimes \mathbb{Q} = H_{\bullet}(GL(n, k))$, the linear group with entries in k , so the reader may think of the Quillen K -theory in this case in terms of this homological identification. We have the following fundamental theorem:

Theorem 6.8 (Deligne–Goncharov). *The category **MTMot** _{k} for a number field k is a Tannakian category with objects $\mathbb{Q}(n)$ and the extensions described above. Furthermore, the Hodge realization functor $\mathbf{MTMot}_k \rightarrow \mathbf{MHodge}$ is exact and faithful.*

7. Motivic measures and zeta functions

The material for this section is based on [23] and [37].

7.1. Overview of basics. The *generalized Euler characteristic* χ associates to each object in **Var** _{k} an element in a fixed commutative ring R such that $\chi(X) = \chi(Y)$ for $X \simeq Y$ and $\chi(X) = \chi(Y) + \chi(X \setminus Y)$ for a closed subvariety $Y \subset X$. The product in R is given by the fibered product of varieties: $\chi(X \times Y) = \chi(X)\chi(Y)$. In fact the Euler characteristic is the canonical example of a generalized Euler characteristic.

Konstevich's original motivation behind inventing motivic measures was to prove the following theorem

Theorem 7.1 (Kontsevich). *Let X and Y be two birationally equivalent Calabi-Yau manifolds. Then X and Y have the same Hodge numbers.*

These types of questions are of enormous importance for duality questions in string theory, namely in mirror symmetry. (There is a beautiful theory around this called the *geometric McKay correspondence* which I completely omit from the discussion.)

An example, following the proof of this theorem of Kontsevich, of a generalized Euler characteristic, comes from Hodge numbers of a complex manifold X . Recall the definition of mixed Hodge structures. For a Hodge structure V , we can define the class of V in $K(\mathbf{Hodge})$ in terms of the graded m -factors in the weight filtration: $[V] := \sum_m [\mathrm{gr}_m^W(V)]$. Define the *Hodge-Deligne polynomial* as

$$\chi_h(X) := \sum_i [H_c^i(X, \mathbb{Q})] \in K(\mathbf{Hodge})$$

where $H_c^i(X, \mathbb{Q})$ is the i -th cohomology of X with compact support. If $Y \subset X$ is locally closed, then the Hodge characteristic is compatible with the exact Gysin sequence:

$$\cdots \longrightarrow H_c^r(X \setminus Y, \mathbb{Q}) \longrightarrow H_c^r(X, \mathbb{Q}) \longrightarrow H_c^r(Y, \mathbb{Q}) \longrightarrow H_c^{r+1}(X \setminus Y, \mathbb{Q}) \longrightarrow \cdots.$$

That is

$$\chi_h(X) = \chi_h(Y) + \chi_h(X \setminus Y).$$

Remark 7.2. For X the affine line \mathbb{A}_k^1 , $H_c^r(\mathbb{A}_k^1, \mathbb{Q}) = 0$ for all $r \neq 2$ and $H_c^2(\mathbb{A}_k^1, \mathbb{Q})$ is one-dimensional of Hodge type $(1, 1)$. So $\chi_h(\mathbb{A}_k^1)$ is invertible. This is the Lefschetz motive \mathbb{L} .

We have the following two ring homomorphisms

- [Hodge characteristic]

$$\begin{aligned} \chi_h : K(\mathbf{Hodge}) &\rightarrow \mathbb{Z}[u, u^{-1}, v, v^{-1}] \\ H^i(X, \mathbb{Q}) &\mapsto \sum_{p+q=i} \dim(H^{p,q}(X, \mathbb{Q})) u^p v^q \end{aligned}$$

- [Weight characteristic]

$$\begin{aligned} \chi_{wt} : K(\mathbf{Hodge}) &\rightarrow \mathbb{Z}[w, w^{-1}] \\ u, v &\mapsto w. \end{aligned}$$

Here $H^{p,q}(X, \mathbb{Q})$ is the (p, q) -th piece of $H_c^i(X, \mathbb{Q})$. Evaluating the weight characteristic at 1 gives us the topological Euler characteristic.

7.2. Equivariant Grothendieck ring. Denote by $\hat{\mu}$ the projective limit of the affine scheme of roots of unity μ_n :

$$\hat{\mu} := \varprojlim_n \mu_n = \varprojlim_n \mathrm{Spec} \ k[x]/(x^n - 1)$$

Let X be an S -variety. A *good μ_n -action on X* is a group action $\mu_n \times X \rightarrow X$ such that each orbit is contained in an affine subvariety of X . (A *good $\hat{\mu}$ -action* factors through a good μ_n -action for some n .)

Definition 7.3 (equivariant Grothendieck ring). The equivariant Grothendieck ring of varieties $K^{\hat{\mu}}(\mathbf{Var}_{X_0})$ is a an abelian group generated by $[X, \hat{\mu}]_S$ for X an S -variety with good $\hat{\mu}$ -action with the relations

- (1) $[X, \hat{\mu}]_S = [Y, \hat{\mu}]_S$ for $X \simeq Y$ as S -varieties with good $\hat{\mu}$ -action.
- (2) $[X, \hat{\mu}]_S = [Y, \hat{\mu}]_S + [X \setminus Y, \hat{\mu}]_S$ if Y is closed in X and the $\hat{\mu}$ -action on Y induced by $\hat{\mu}$ -action on X .
- (3) [monodromy] $[X \times V, \hat{\mu}]_S = [\mathbb{A}_k^n, \hat{\mu}]_S$ where V is the n -dimensional affine space with a good $\hat{\mu}$ -action and \mathbb{A}_k^n is the affine space with a trivial $\hat{\mu}$ -action.

Definition 7.4 (equivariant Euler characteristic). The *equivariant Euler characteristic* is a ring homomorphism

$$\begin{aligned}\chi_{\text{top}}(-, \alpha) : K^{\hat{\mu}}(\mathbf{Var}_{X_0})[\mathbb{L}^{-1}] &\longrightarrow Z \\ (-) &\mapsto \sum_{q \geq 0} (-1)^q \dim H^q(-, \mathbb{C})_{\alpha}\end{aligned}$$

where $H^q(-, \mathbb{C})_{\alpha} \subset H^*(-, \mathbb{C})$ on which there is a good $\hat{\mu}$ -action through multiplication by α .

7.3. Arc spaces. Let k be a field of characteristic zero and let X be a variety over k . For all natural numbers n , we define the *arc space* $\mathcal{L}_n(X)$ as an algebraic variety over k whose k -rational points are $K[t]/t^{n+1}$ -rational points for some $k \subset K$. We set $\mathcal{L}(X) = \varprojlim_n \mathcal{L}_n(X)$. Note that $\mathcal{L}_0(X) = X$ and $\mathcal{L}_1(X) = TX$, the (Zariski) tangent space of X . We call the K -rational points of $\mathcal{L}_n(X)$ K -arcs of X or arcs for short. There are the structure morphisms

$$\begin{aligned}\pi_n : \mathcal{L}(X) &\rightarrow \mathcal{L}_n(X), \\ \pi_n^m : \mathcal{L}_m(X) &\rightarrow \mathcal{L}_n(X).\end{aligned}$$

The *origin* of an arc γ is $\pi_0(\gamma)$.

A “cleaner” (but equivalent) definition of the space of arcs is the following:

Definition 7.5 (Space of arcs). Let X be a variety over k (which, I remind the reader is a separated scheme of finite type over k .) Denote by $\mathcal{L}(X)$ the *scheme of germs of arcs on X* . It is defined as a scheme over k such that, for any extension $k \subset K$ there is a natural bijection

$$\mathcal{L}(X)(K) \simeq \text{Hom}_{\mathbf{Sch}_k}(\text{Spec } k[[t]], K)$$

The scheme $\mathcal{L}(X) := \varprojlim_n \mathcal{L}_n(X)$ in the category of schemes $\mathcal{L}_n(X)$ representing the functor

$$R \mapsto \text{Hom}_{\mathbf{Sch}_k}(\text{Spec } R[t]/t^{n+1}, X)$$

defined on the category of K -algebras.

Not much is known about the space $\mathcal{L}(X)$. An important result is the following:

Proposition 7.6 (Kolchin). *Let X be an integral scheme. Then $\mathcal{L}(X)$ is irreducible.*

7.4. The Nash Problem. Arcs were first introduced by Nash in connection with singularities: let P be a singular point on X and set $\mathcal{L}_{\{P\}}(X) := \pi_0^{-1}(P)$ of arcs with origin P . He studied the space $\mathcal{N}_{\{P\}}(X) \subset \mathcal{L}_{\{P\}}(X)$ of arcs not contained in the singular locus $\text{Sing } X$. Let me briefly sketch Nash’s beautiful idea, following Loeser’s Trieste lectures [38].

Let $Y \xrightarrow{\rho} X$ be a resolution of singularities of a scheme X . Recall that this means: Y is smooth, ρ is proper and ρ induces an isomorphism $Y \setminus \rho^{-1}(\text{Sing } X) \simeq X \setminus \text{Sing } X$. We say that the resolution of singularities is *divisorial* if the locus E where ρ is *not* a local isomorphism (called the *exceptional set*) is a divisor in Y . In such cases, we call the locus an *exceptional divisor*. Let $Y' \xrightarrow{\rho'} X$ be another proper birational morphism (and assume that X is normal) and let p be the generic point

of E . We say that E appears in ρ' if $\rho'^{-1} \circ \rho$ is a local isomorphism at p . The exceptional divisor E is called an *essential component* of a resolution of X if E appears in every divisorial resolution of X . Denote the set of essential components as $\mathcal{C}_{\{P\}}(X)$.

Theorem 7.7 (Nash). *The mapping $\mathcal{C}_{\{P\}}(X) \xrightarrow{\nu} \mathcal{N}_{\{P\}}(X)$ is injective.*

Basically, what this theorem tells is that every irreducible component of $\mathcal{L}(X)$ through a singularity P corresponds to an exceptional divisor that occurs on every resolution. The *Nash Problem* asks

Question 7.8. For what X is the map ν a bijection?

An explicit nonexample was provided by Ishii-Kollar when X is a 4-dimensional hypersurface singularity $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{A}_k^5 \mid x_1^3 + x_2^3 + x_3^3 + x_4^3 + x_5^6 = 0\}$ which has 1 irreducible family of arcs but 2 essential components. (The base k is obviously not allowed to be of characteristic 2 or 3.)

7.5. Motivic zeta functions. Let t be a *fixed* coordinate on \mathbb{A}_k^1 and let $n \geq 1$ be an integer. A morphism $f : X \rightarrow \mathbb{A}_k^1$ induces a morphism $f_n : \mathcal{L}_n(X) \rightarrow \mathcal{L}_n(\mathbb{A}_k^1)$. Any $\alpha \in \mathcal{L}(\mathbb{A}_k^1)$ (resp. $\alpha \in \mathcal{L}_n(\mathbb{A}_k^1)$) gives a power series $\alpha(t) \in K[[t]]$ (resp. $\alpha(t) \in K[[t]]/t^{n+1}$) by definition. Define a map $\text{ord}_t : \mathcal{L}(\mathbb{A}_k^1) \rightarrow \mathbb{Z}_{\geq 0} \cup \{\infty\}$ as $\text{ord}_t(\alpha) := \max_{t^e \mid \alpha(t)} \{e\}$.

Definition 7.9. $\mathcal{X}_n := \{\phi \in \mathcal{L}_n(X) \mid \text{ord}_t f_n(\phi) = n\}$.

The variety \mathcal{X}_n is a locally closed subvariety of $\mathcal{L}_n(X)$. Let X_0 be a variety obtained by setting $f = 0$. The variety \mathcal{X}_n is an X_0 -variety through the structure maps π_0^n . Define the morphism

$$\bar{f}_n : \mathcal{X}_n \rightarrow \mathbb{G}_m \text{ given by } \phi \mapsto \text{coefficient of } t_n \text{ in } f_n(\phi)$$

and set $\mathcal{X}_{n,1} := \bar{f}_n^{-1}(1)$. The following characterization of \mathcal{X}_n allows us to pass on to étale settings “without reductions at bad primes”.

Proposition 7.10. *As $\mathbb{G}_m \times X_0$ -variety, \mathcal{X}_n is a quotient of $\mathcal{X}_{n,1} \times \mathbb{G}_m$ by the μ_n -action $a(\phi, b) = (a\phi, a^{-1}b)$.*

Recall the definition of the equivariant Grothendieck ring $K^{\hat{\mu}}(\mathbf{Var}_{X_0})$ of section 7.2.

Definition 7.11 (Denef-Loeser, Looijenga). The *motivic zeta function* of a morphism $f : X \rightarrow \mathbb{A}_k^1$ over $K^{\hat{\mu}}(\mathbf{Var}_{X_0})[\mathbb{L}^{-1}]$ is defined by the sum

$$Z(T) = \sum_{n \geq 1} [\mathcal{X}_{n+1}, \hat{\mu}]_{X_0} \mathbb{L}^{-nd} T^n$$

where $[-, \hat{\mu}]_{X_0}$ is a class in $K^{\hat{\mu}}(\mathbf{Var}_{X_0})$ and $d = \dim \mathcal{X}_{n+1}$.

Theorem 7.12 (Denef-Loeser). *The motivic zeta function $Z(T)$ in definition 7.11 is a rational function.*

Appendix A. Motivic ideas in physics (by M.Marcolli)

The theory of motives was introduced to mathematicians by Grothendieck as a universal cohomology theory for algebraic varieties. At present, while a lot of progress happened since its origin, it is still a field in very rapid development, with a wealth of intriguing conjectures still unsolved and new unexpected applications being found. Among these, motives recently made their appearance in the world of theoretical physics, a latecomer with respect to more traditional mathematical tools adopted by the physics community, but one that is likely to lead to a wider range of future applications. We attempt here to convince our audience that, in various forms, some of the main ingredients of the theory of motives have already penetrated the horizon of the high-energy physics community. Roughly speaking, the theory of motives encompasses an *analytic* side, which is based on Hodge theory and mixed Hodge structures, a *geometric* side based on algebraic varieties and algebraic cycles, and an *algebraic* side coming from algebraic K -theory. This appendix will recall briefly some contexts familiar to theoretical physics where these notions have come play an important role. We will try to focus not only on the setting of perturbative scalar quantum field theories and the parametric representation of Feynman integrals, which is one of the main interactions between physics and motives of interest today, but also on other occurrences of some of the main motivic ideas in other aspects of theoretical physics.

A.1. The dawn of motives in physics. One can trace the origin of motivic ideas in physics to the theory of electromagnetism and the “cohomological” interpretation of the Maxwell equations in terms of differential forms. In fact, electromagnetism is the historic origin of Hodge theory, which in turn is one of the essential building blocks of the theory of motives. Naturally, the current ideas on mixed Hodge structures are a long distance away from the Maxwell equations, but as we discuss briefly below, ideas related to Hodge theory continued to play a role in physics beyond electromagnetism, most notably in the more recent context of mirror symmetry. On the other hand, periods of algebraic varieties, another important aspect of the theory of motives, have made repeated appearances in the context of theoretical physics, from their role in residues of Feynman graphs, which is being extensively investigated today, to the modeling of spectral density functions in condensed matter physics.

A.2. Hodge theory: from electromagnetism to mirror symmetry. An important way in which some motivic ideas make contact with the world of physics is through Hodge theory. In fact, the very origin of Hodge theory is in the formulation of Maxwell’s equations of electromagnetism. In terms of electric and magnetic fields, these are usually written in the form

$$(1) \quad \left\{ \begin{array}{l} \nabla \times E = \frac{\partial B}{\partial t} \\ \nabla \times B = -\frac{\partial E}{\partial t} \\ \text{div } E = 0 \\ \text{div } B = 0, \end{array} \right.$$

with $E = (E_x, E_y, E_z)$ and $B = (B_x, B_y, B_z)$. One can better assemble the electric and magnetic field in a single entity, described by a 2-form

$$F = (E_x dx + E_y dy + E_z dz) \wedge dt + B_x dy \wedge dz - B_y dx \wedge dz + B_z dx \wedge dy.$$

The advantage of this formulation is that it transforms the Maxwell equations (1) into the much more appealing and conceptually simple form

$$(2) \quad \begin{cases} dF = 0 \\ d^*F = 0, \end{cases}$$

where d is the deRham differential with adjoint d^* . See the beautiful paper of Raoul Bott [14] for a more detailed discussion of these and other examples of geometrization of physics.

This formulation of the Maxwell equations leads immediately to considerations on the Hodge decomposition of differential forms $\omega = d\alpha + d^*\beta + \gamma$, with γ harmonic, satisfying $d\gamma = d^*\gamma = 0$ as above. The identification of deRham cohomology with harmonic forms, in turn, relates then to the Hodge decomposition of the cohomology of compact Kähler manifolds

$$H^k(X, \mathbb{C}) = \bigoplus_{p+q=k} H^{p,q}(X),$$

which is the origin of the notion of Hodge structures as the “analytic side” of the theory of motives of algebraic varieties.

In this more refined form, Hodge theory made a spectacular return into the world of theoretical physics with the discovery of mirror symmetry. This originated with the observation that certain Calabi-Yau threefolds, which serve as compactified dimensions in string theory, form mirror pairs when their σ -models determine the same superconformal field theory. In particular, it was observed that two such mirror varieties X and X' , while geometrically very different, have a very simple “mirror relation” between their Hodge numbers

$$(3) \quad h^{p,q}(X) = h^{3-p,q}(X'),$$

where $h^{p,q}(X) = \dim H^{p,q}(X)$. From this initial observation the field of mirror symmetry rapidly developed into an extremely interesting and elaborate branch of contemporary mathematical physics, where several sophisticated ideas from algebraic geometry and arithmetic geometry, including various aspects of the theory of motives such as algebraic cycles, came to play an increasingly important role.

A.3. Algebraic cycles and homological mirror symmetry. The main way in which algebraic cycles came to play a role in the more recent developments of homological mirror symmetry (see [40]) is through the fact that they define objects in the bounded derived category $D^b(X)$ of coherent sheaves. For a variety of general type that is realized as a complete intersection in a toric variety, homological mirror symmetry predicts that the derived category $D^b(X)$ can be realized as a subcategory of the Fukaya–Seidel category $F(Y, W)$ of a mirror Landau–Ginzburg model specified by a symplectic manifold Y with a superpotential W . As shown in [40], using this approach one can transform questions about algebraic cycles into corresponding questions in the mirror environment, typically in terms of the information associated to the singularities of W . This point of view may prove very useful from the motivic viewpoint. In fact, one of the main source of difficulties in the theory of motives is the fact that questions about algebraic cycles are extremely difficult to approach, which is the reason why Grothendieck’s “standard conjectures”

on motives remain unsolved. One can, potentially, at least in some cases, employ the homological mirror symmetry viewpoint to relate such questions, as in [40], to constructions involving vanishing cycles and singularity theory in the mirror Landau-Ginzburg model.

A.4. Algebraic varieties and periods: Fermi surfaces. Periods are a very interesting class of complex numbers, which are in general not algebraic, but which can be obtained via a geometric procedure where all the objects involved are themselves algebraic. More precisely, these are numbers that are obtained by integrating an algebraic differential form over a cycle in an algebraic variety. For the purpose of our setting here, we can assume that the varieties themselves are defined over \mathbb{Q} (or more generally a finite extension, a number field) and so are the algebraic differential forms, while the domain of integration is a cycle defined again over the same field. One also allows for the case where the domain of integration has boundary, in which case it is defined by inequalities, as a semi-algebraic set, again over the same field, and in this case the integration of a differential form correspond to a pairing with a relative cohomology group. We'll see this in more explicit examples later. There are relations between periods coming from change of variables in the integrals that compute them and from the Stokes formula. The properties, known and conjectural, of the algebra of periods are analyzed in detail in the paper of Kontsevich–Zagier [41]. The type of numbers that one obtains as periods can be used as numerical signatures of the motivic complexity of the algebraic variety. For example, multiple zeta values are the tell-tale sign of the presence of a particular class of motives, the mixed Tate motive, about which more later.

A first beautiful example of the role of periods of algebraic varieties arises in the context of Fermi surfaces in condensed matter, see [29], [30], [52]. One considers electrons moving in a lattice, in the independent electron approximation, where one can describe the dynamics as a single particle model with an effective potential. The energy levels of the system are obtained from the Schrödinger equation

$$(4) \quad (-\Delta + V)\psi = \lambda\psi, \quad \text{with} \quad \psi(x + \gamma) = e^{i\langle k, \gamma \rangle} \psi(x),$$

for a fixed $k \in \mathbb{R}^d$ and for all $\gamma \in \Gamma$, the lattice of ions. The j -th eigenvalue $\lambda_j(k)$, for given crystal momentum k , defines the j -th band in the spectrum.

The Fermi surface separates in the space of crystal momenta the occupied and non-occupied states at zero temperature. One typically replaces the spectral problem (4) with a discretized version, with the Laplacian replaced by a Harper operator (discrete Laplacian) on a lattice, $H = \sum_j (T_j + T_j^*)$, summed over the generators of the lattice, with T_j the shift operators,

$$(5) \quad (-H + V - \lambda)\psi = 0, \quad \text{with} \quad T_j^{a_j} \psi = \xi_j \psi,$$

with $T_j^{a_j}$ the shifts corresponding to the lattice $\Gamma = \bigoplus_{j=1}^d \mathbb{Z}a_j$. In these geometric terms, one then describes the Fermi surface in terms of the Bloch variety

$$(6) \quad B(V) = \{(\xi, \lambda) \in (\mathbb{C}^*)^d \times \mathbb{C} \mid \exists \psi \text{ solution of (5)}\}.$$

The Fermi surface F_λ is obtained from the Bloch variety as the fiber $F_\lambda = \pi^{-1}(\lambda)$ of the projection $\pi : B(V) \rightarrow \mathbb{C}$. This is an algebraic variety, from which one can read physical properties of the system. One of the main results of [29] is the fact

that the *density of states* $\rho(\lambda)$ of the system can be computed as a *period* of this algebraic variety in the form

$$(7) \quad \frac{d\rho}{d\lambda} = \frac{1}{(2\pi)^d a_1 \cdots a_d} \int_{F_\lambda \cap (S^1)^d} \omega_\lambda,$$

where the domain of integration are the real points of the Fermi surface F_λ and the algebraic differential form is defined by the relation

$$\pi^*(d\lambda) \wedge \omega_\lambda = \frac{d\xi_1}{\xi_1} \wedge \cdots \wedge \frac{d\xi_d}{\xi_d}.$$

Although the algebro-geometric properties of the Bloch varieties and Fermi surfaces have been studied extensively in [29], there has been so far no explicitly motivic analysis of this family of varieties and of the corresponding periods. A notable exception is the paper of Steinstra [59], which gives a motivic interpretation of the period in terms of Deligne cohomology and Mahler measures, in the case of zero potential. One can expect that a more extensive analysis of the motivic properties of Bloch varieties and Fermi surfaces will lead to interesting results, in the same vein and with similar techniques as those adopted in the study of the motivic aspects of Feynman integrals described here below.

A.5. The parametric form of Feynman integrals. Another example of the role in physics of periods of algebraic varieties, which is presently being widely investigated, is that of parametric Feynman integrals of perturbative scalar quantum field theories.

In perturbative quantum field theory, to a scalar field theory with Euclidean Lagrangian of the form

$$(8) \quad \mathcal{L}(\phi) = \int \left(\frac{1}{2} (\partial_\mu \phi)^2 + \frac{m^2}{2} \phi^2 + \mathcal{P}(\phi) \right) dv,$$

with the interaction term given by a polynomial $\mathcal{P}(\phi)$, one assigns a formal perturbative expansion

$$\sum_{\Gamma} \frac{U(\Gamma, p_1, \dots, p_N)}{\#Aut(\Gamma)}$$

parameterized by Feynman graphs Γ . These are finite graphs with internal edges connecting pairs of vertices and external edges attached to a single vertex and carrying external momenta p_i . The vertices are constrained to have valences equal to the powers of the monomials in $\mathcal{P}(\phi)$. The contribution $U(\Gamma, p_1, \dots, p_N)$ is dictated by the Feynman rules of the theory, and can be expressed as a (typically divergent) finite dimensional integral. In the parametric form (or α -parametrization) these integrals can be written in the form

$$(9) \quad U(\Gamma, p_1, \dots, p_N) = \frac{\Gamma(n - \frac{D\ell}{2})}{(4\pi)^{D\ell/2}} \int_{\sigma_n} \frac{P_\Gamma(t, p)^{-n+D\ell/2} \omega_n}{\Psi_\Gamma(t)^{-n+D(\ell+1)/2}},$$

where $n = \#E_{int}(\Gamma)$ and $\ell = b_1(\Gamma)$. The domain of integration is the simplex $\sigma_n = \{t \in \mathbb{R}_+^n \mid \sum_i t_i = 1\}$. The Kirchhoff polynomial (also called first Symanzik polynomial) $\Psi_\Gamma(t)$ is given by

$$(10) \quad \Psi_\Gamma(t) = \sum_{T \subset \Gamma} \prod_{e \notin E(T)} t_e,$$

where the sum is over all the spanning trees (forests) T of the graph Γ and for each spanning tree the product is over all edges of Γ that are not in the tree. The second Symanzik polynomial $P_\Gamma(t, p)$ is similarly defined in terms of cut sets instead of spanning trees, and it depends explicitly on the external momenta of the graph (see [7] §18). The polynomial $\Psi_\Gamma(t)$ is homogeneous of degree $\ell = b_1(\Gamma)$, while, in the massless case, the polynomial $P_\Gamma(t, p)$ is homogeneous of degree $\ell + 1$.

After removing the divergent Gamma factor from (9), one is left with the residue of the Feynman integral, which has then the form of a period, in the sense that it is the integration of an algebraic differential form on a semi-algebraic set in the algebraic variety given by the complement of the hypersurface defined by the vanishing of the graph polynomial in the denominator of the integrand of (9). Saying that the integral is a period, of course, ignores the important issue of divergences: in fact, even after removing the Gamma factor, the remaining integral can still contain divergences coming from the intersections of the domain of integration with the hypersurface. Modulo this issue, which needs to be addressed via a suitable regularization and renormalization procedure, one is considering a period whose value contains information on the motivic nature of the relative cohomology

$$(11) \quad H^{n-1}(\mathbb{P}^{n-1} \setminus X_\Gamma, \Sigma_n \setminus (X_\Gamma \cap \Sigma_n)),$$

where Σ_n is the divisor of coordinate hyperplanes in \mathbb{P}^{n-1} , which contains the boundary of the domain of integration $\partial\sigma_n \subset \Sigma_n$. The frequent appearance of multiple zeta values in Feynman integral computations (see [15]) suggests that the cohomology (11) may be a realization of a mixed Tate motive. However, [5] shows that the graph hypersurfaces X_Γ themselves can be arbitrarily complex in motivic terms: their classes $[X_\Gamma]$ span the Grothendieck ring of varieties. Understanding precise conditions on the graphs under which the Feynman integral gives a period of a mixed Tate motive is presently a very interesting open question in the field.

A.6. Motivic Galois groups in physics. A manifestation of motivic Galois groups in physics arises in the context of the Connes–Kreimer theory of perturbative renormalization. In [19] it was shown that the counterterms in the BPHZ renormalization procedure, described as in Connes–Kreimer in terms of Birkhoff factorization in the affine group scheme dual to the Hopf algebra of Feynman graphs, can be equivalently formulated in terms of solutions to a certain class of differential systems with irregular singularities. This is obtained by writing the terms in the Birkhoff factorization as time ordered exponentials, and then using the fact that

$$\text{Te}^{\int_a^b \alpha(t) dt} := 1 + \sum_{n=1}^{\infty} \int_{a \leq s_1 \leq \dots \leq s_n \leq b} \alpha(s_1) \cdots \alpha(s_n) ds_1 \cdots ds_n$$

is the value $g(b)$ at b of the unique solution $g(t) \in G$ with value $g(a) = 1$ of the differential equation $dg(t) = g(t) \alpha(t) dt$.

The type of singularities are specified by physical conditions, such as the independence of the counterterms on the mass scale. These conditions are expressed geometrically through the notion of G -valued *equisingular connections* on a principal \mathbb{C}^* -bundle B over a disk Δ , where G is the pro-unipotent Lie group of characters of the Connes–Kreimer Hopf algebra of Feynman graphs. In physical terms, the disk Δ gives the complexified dimension of dimensional regularization of the Feynman integrals, while the fiber over $z \in \Delta$ consists of the μ^z , with the parameter μ giving the energy scale of the renormalization group flow. The *equisingularity* condition

is the property that such a connection ω is \mathbb{C}^* -invariant and that its restrictions to sections of the principal bundle that agree at $0 \in \Delta$ are mutually equivalent, in the sense that they are related by a gauge transformation by a G -valued \mathbb{C}^* -invariant map regular in B , hence they have the same type of (irregular) singularity at the origin. The classification of equivalence classes of these differential systems can be done via the Riemann–Hilbert correspondence and differential Galois theory. This means assembling these data, in the form of a category of “flat equisingular vector bundles”, which one can show ([19]) is a Tannakian category. The Tannakian formalism then identifies it with the category of finite dimensional linear representations of a Tannakian Galois group $U^* = U \rtimes \mathbb{G}_m$, where U is pro-unipotent with Lie algebra the free graded Lie algebra with one generator e_{-n} in each degree $n \in \mathbb{N}$. The group U^* is identified (non-canonically) with the motivic Galois group of mixed Tate motives over the cyclotomic ring $\mathbb{Z}[e^{2\pi i/N}]$, for $N = 3$ or $N = 4$, localized at N .

A.7. Algebraic K-theory and conformal field theory. Algebraic K -theory, which one can think of as the more algebraic manifestation of the theory of motives, may appear at first to be the part that is more remote from physics. On the contrary, it has come to play an important role, for example in the context of conformal field theory. We mention here very briefly some of the recent developments in this direction, while we refer the reader to [51] for a more informative overview.

Recall that the dilogarithm is defined as

$$Li_2(z) = \int_z^0 \frac{\log(1-t)}{t} dt = \sum_{n=1}^{\infty} \frac{z^n}{n^2}.$$

It satisfies the functional equation $Li_2(z) + Li_2(1-z) = Li_2(1) - \log(z) \log(1-z)$, where $Li_2(1) = \zeta(2)$, for $\zeta(s)$ the Riemann zeta function. A variant is given by the Rogers dilogarithm

$$L(x) = Li_2(x) + \frac{1}{2} \log(x) \log(1-x).$$

See [61] for a detailed survey on the dilogarithm and its properties.

There is a relation between the torsion elements in the algebraic K -theory group $K_3(\mathbb{C})$ and rational conformally invariant quantum field theories in two dimensions, see [51]. There is, in fact, a map, given by the dilogarithm, from torsion elements in the Bloch group (closely related to the algebraic K -theory) to the central charges and scaling dimensions of the conformal field theories.

This correspondence arises by considering sums of the form

$$(12) \quad \sum_{m \in \mathbb{N}^r} \frac{q^{Q(m)}}{(q)_m},$$

where $(q)_m = (q)_{m_1} \cdots (q)_{m_r}$, with $(q)_{m_i} = (1-q)(1-q^2) \cdots (1-q^{m_i})$ and

$$Q(m) = m^t A m / 2 + b m + h$$

has rational coefficients. Such sums are naturally obtained from considerations involving the partition function of a bosonic rational CFT. In particular, (12) can define a modular function only if all the solutions of the equation

$$(13) \quad \sum_j A_{ij} \log(x_j) = \log(1 - x_i)$$

determine elements of finite order in an extension $\hat{B}(\mathbb{C})$ of the Bloch group, which accounts for the fact that the logarithm is multi-valued. The Rogers dilogarithm gives a natural group homomorphism $(2\pi i)^2 L : \hat{B}(\mathbb{C}) \rightarrow \mathbb{C}/\mathbb{Z}$, which takes values in \mathbb{Q}/\mathbb{Z} on the torsion elements. These values give the conformal dimensions of the fields in the theory.

References

- [1] Paolo Aluffi and Matilde Marcolli, *Feynman motives of banana graphs*, arXiv: 0801.1690v2 [hep-th] (2008).
- [2] ———, *Algebro-geometric Feynman rules*, arXiv:0811.2514v1 [hep-th] (2008).
- [3] ———, *Parametric Feynman integrals and determinant hypersurfaces*, arXiv:0901.2107v1 [math.AG] (2009).
- [4] Yves André, *Une introduction aux motifs (motifs purs, motifs mixtes, périodes)*, Panoramas et Synthèses [Panoramas and Syntheses], vol. 17, Société Mathématique de France, Paris, 2004 (French, with English and French summaries).
- [5] Prakash Belkale and Patrick Brosnan, *Matroids, motives, and a conjecture of Kontsevich*, Duke Math. J. **116** (2003), no. 1, 147–188.
- [6] Franziska Bittner, *The universal Euler characteristic for varieties of characteristic zero*, Compos. Math. **140** (2004), no. 4, 1011–1032.
- [7] James D. Bjorken and Sidney D. Drell, *Relativistic quantum fields*, McGraw-Hill Book Co., New York, 1965.
- [8] Spencer Bloch, *Motives associated to graphs*, Jpn. J. Math. **2** (2007), no. 1, 165–196.
- [9] ———, *Motives associated to sums of graphs*, arXiv:0810.1313v1 [math.AG] (2008).
- [10] Spencer Bloch and Arthur Ogus, *Gersten conjectures and the homology of schemes*, Annals. Sci. ENS **7** (1974).
- [11] Spencer Bloch, Hélène Esnault, and Dirk Kreimer, *On motives associated to graph polynomials*, Comm. Math. Phys. **267** (2006), no. 1, 181–225.
- [12] Spencer Bloch and Dirk Kreimer, *Mixed Hodge structures and renormalization in physics*, Commun.Num.Theor.Phys. **2** (2008), no. 4, 637–718.
- [13] Christian Bogner and Stefan Weinzierl, *Periods and Feynman integrals*, arXiv:0711.4863v2 [hep-th] (2007).
- [14] Raoul Bott, *On some recent interactions between mathematics and physics*, Canad. Math. Bull. **28** (1985), no. 2, 129–164.
- [15] D. J. Broadhurst and D. Kreimer, *Association of multiple zeta values with positive knots via Feynman diagrams up to 9 loops*, Phys. Lett. B **393** (1997), no. 3-4, 403–412.
- [16] Pierre Colmez and Jean-Pierre Serre (eds.), *Correspondance Grothendieck-Serre*, Documents Mathématiques (Paris) [Mathematical Documents (Paris)], 2, Société Mathématique de France, Paris, 2001 (French).
- [17] Alain Connes and Dirk Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem I: The Hopf algebra structure of graphs and main theorem*, Comm. Math. Phys. **210** (2000), 249–273.
- [18] ———, *Renormalization in quantum field theory and the Riemann-Hilbert problem II: The β -function, diffeomorphisms and the renormalization group*, Comm. Math. Phys. **216** (2001), 215–241.
- [19] Alain Connes and Matilde Marcolli, *Renormalization, the Riemann-Hilbert correspondence, and motivic Galois theory*, Frontiers in number theory, physics, and geometry. II, Springer, Berlin, 2007, pp. 617–713.
- [20] ———, *Noncommutative geometry, quantum fields and motives*, American Mathematical Society Colloquium Publications, vol. 55, American Mathematical Society, Providence, RI, 2008.
- [21] Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. **43** (1974), 273–307 (French).
- [22] P. Deligne, *Catégories tannakiennes*, The Grothendieck Festschrift, Vol. II, Progr. Math., vol. 87, Birkhäuser Boston, Boston, MA, 1990, pp. 111–195 (French).
- [23] Jan Denef and Francois Loeser, *Geometry of arc spaces on algebraic varieties*, arXiv: 0006050v1 [math.AG] (2000).

- [24] Alexandru Dimca, *Sheaves in topology*, Springer-Verlag, New York, 2003.
- [25] Dzmitry Doryn, *Cohomology of graph hypersurfaces associated to certain Feynman graphs*, Doctoral thesis, Universität Duisburg-Essen, arXiv:0811.0402v1 [math.AG] (2008).
- [26] William Fulton, *Intersection theory*, Springer-Verlag, New York, 1998. Second edition of the 1984 original.
- [27] S.I. Gelfand and Yu.I. Manin, *Methods of homological algebra*, Springer-Verlag, Berlin, 2003. Springer Monographs in Mathematics.
- [28] H. Gillet and C. Soulé, *Descent, motives and K-theory*, J. Reine Angew. Math. **478** (1996), 127–176.
- [29] D. Gieseker, H. Knörrer, and E. Trubowitz, *The geometry of algebraic Fermi curves*, Perspectives in Mathematics, vol. 14, Academic Press Inc., Boston, MA, 1993.
- [30] ———, *An overview of the geometry of algebraic Fermi curves*, Algebraic geometry: Sundance 1988, 1991, pp. 19–46.
- [31] Phillip Griffiths and Joseph Harris, *Principles of algebraic geometry*, Wiley Classics Library, John Wiley & Sons Inc., New York, 1994. Reprint of the 1978 original.
- [32] Alexander Grothendieck, *Revêtements étalés et groupe fondamental. Fasc. I: Exposés 1 à 5*, Séminaire de Géométrie Algébrique, vol. 1960/61, Institut des Hautes Études Scientifiques, Paris, 1963.
- [33] A. Grothendieck, *Standard conjectures on algebraic cycles*, Algebraic Geometry (Internat. Colloq., Tata Inst. Fund. Res., Bombay, 1968), Oxford Univ. Press, London, 1969, pp. 193–199.
- [34] Alexandre Grothendieck, *Esquisse d'un programme*, Geometric Galois actions, 1, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, pp. 5–48 (French, with French summary). With an English translation on pp. 243–283.
- [35] Thomas C. Hales, *What is motivic measure?*, Bull. Amer. Math. Soc. (N.S.) **42** (2005), no. 2, 119–135 (electronic).
- [36] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.
- [37] Eduard Looijenga, *Motivic measures*, arXiv: 0006220v2 [math.AG] (2000).
- [38] Francois Loeser, *Motivic integration and McKay correspondence*, Lectures at “Trieste School and Conference on Intersection Theory and Moduli” (2002).
- [39] Uwe Jannsen, Steven Kleiman, and Jean-Pierre Serre (eds.), *Motives*, Proceedings of Symposia in Pure Mathematics, vol. 55, American Mathematical Society, Providence, RI, 1994.
- [40] Ludmil Katzarkov, *Homological Mirror Symmetry and Algebraic Cycles*, Homological Mirror Symmetry, 2008, pp. 1–28.
- [41] Maxim Kontsevich and Don Zagier, *Periods*, Mathematics unlimited—2001 and beyond, Springer, Berlin, 2001, pp. 771–808.
- [42] Serge Lang, *Algebra*, 3rd ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002.
- [43] Michael Larsen and Valery A. Lunts, *Motivic measures and stable birational geometry*, Mosc. Math. J. **3** (2003), no. 1, 85–95, 259 (English, with English and Russian summaries).
- [44] Matilde Marcolli, *Motivic renormalization and singularities*, arXiv:0804.4824v2 [math-ph] (2008).
- [45] ———, *Feynman integrals and motives*, Plenary lecture at ECM 2008 (2008).
- [46] Matilde Marcolli and Abhijnan Rej, *Supermanifolds from Feynman graphs*, J. Phys. A **41** (2008), 315402–315423.
- [47] Ju. I. Manin, *Correspondences, motifs and monoidal transformations*, Mat. Sb. (N.S.) **77** (**119**) (1968), 475–507 (Russian).
- [48] Yuri Manin, *Lectures on zeta functions and motives (according to Deninger and Kurokawa)*, Astérisque **228** (1995), 4, 121–163. Columbia University Number Theory Seminar (New York, 1992).
- [49] Yuri I. Manin, *Gauge field theory and complex geometry*, 2nd ed., Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 289, Springer-Verlag, Berlin, 1997. Translated from the 1984 Russian original by N. Koblitz and J. R. King; With an appendix by Sergei Merkulov.
- [50] Jacob P. Murre, *Lecture on motives*, Transcendental aspects of algebraic cycles, London Math. Soc. Lecture Note Ser., vol. 313, Cambridge Univ. Press, Cambridge, 2004, pp. 123–170.

- [51] Werner Nahm, *Conformal field theory and torsion elements of the Bloch group*, Frontiers in number theory, physics, and geometry. II, 2007, pp. 67–132.
- [52] Chris Peters, *Algebraic Fermi curves (after Gieseker, Trubowitz and Knörrer)*, Astérisque **189–190** (1990), Exp. No. 723, 239–258. Séminaire Bourbaki, Vol. 1989/90.
- [53] Bjorn Poonen, *The Grothendieck ring of varieties is not a domain*, Math. Res. Lett. **9** (2002), no. 4, 493–497.
- [54] Jonathan Rosenberg, *Algebraic K-theory and its applications*, Graduate Texts in Mathematics, vol. 147, Springer-Verlag, New York, 1994.
- [55] Neantro Saavedra Rivano, *Catégories Tannakiennes*, Lecture Notes in Mathematics, Vol. 265, Springer-Verlag, Berlin, 1972 (French).
- [56] Neeraja Sahasrabudhe, *Grothendieck ring of varieties*, Master’s Thesis, Université Bordeaux 1 (2007).
- [57] A. J. Scholl, *Classical motives*, Motives (Seattle, WA, 1991), Proc. Sympos. Pure Math., vol. 55, Amer. Math. Soc., Providence, RI, 1994, pp. 163–187.
- [58] Sug Woo Shin, *Grothendieck function-sheaf correspondence*, lecture notes from the Harvard seminar on geometric class field theory (2005).
- [59] Jan Stienstra, *Motives from diffraction*, Algebraic cycles and motives. Vol. 2, 2007, pp. 340–359.
- [60] Charles Weibel, *The K-book: An introduction to algebraic K-theory*, in preparation, chapters I–IV available at <http://www.math.rutgers.edu/~weibel/Kbook.html>.
- [61] Don Zagier, *The dilogarithm function*, Frontiers in number theory, physics, and geometry. II, 2007, pp. 3–65.

E-mail address: abhijnan.rej@gmail.com

URL: <http://abhijnan.rej.googlepages.com>

This page intentionally left blank

Combinatorics and Feynman graphs for gauge theories

Walter D. van Suijlekom

ABSTRACT. We give an overview of the use of combinatorics in renormalization of gauge theories, using the Connes–Kreimer Hopf algebra. We prove some physical results in perturbative quantum gauge theories without relying on the formal manipulations involving path integrals. Instead, we take a perturbative series of Feynman graphs as a starting point. After a careful setup and counting of Feynman graphs, we study the structure of the renormalization Hopf algebra of gauge theories on the level of Green’s functions. This involves Slavnov–Taylor identities, described by Hopf ideals, and Dyson–Schwinger equations, described by Hochschild cocycles [7]. As a new result, we prove the Kreimer’s gauge theory theorem formulated in [14].

1. Introduction

These lecture notes are a compilation and extension of my previous work [20, 18, 19, 21] on renormalization Hopf algebras of gauge theories. It is supposed to serve as a self-contained treatment of the following themes:

- (1) Feynman graphs in (non-abelian) gauge theories and their combinatorics
- (2) Renormalization of gauge theories using Hopf algebras of Feynman graphs
- (3) Dyson–Schwinger equations in terms of Hochschild cocycles.

The last point is based on the articles [7, 2, 14].

The description of (BPHZ)-renormalization in terms of Hopf algebras dates back to [12]. It involved rooted trees that encoded the combinatorial structure of renormalization. The description in terms of Feynman graphs – which are the objects of particular computational use in physics – was established in [4, 5]. Although a lot of progress has been made since then in several directions we will focus on the particular case of gauge theories.

Of course, their Hopf algebraic description was contained in the original papers *loc.cit.* – which after all applied to *any* quantum field theory – but the richer structure due to the presence of a gauge symmetry was explored in this context more recently. Non-abelian gauge theories were dissected in [13], with a prominent role played by the Slavnov–Taylor identities for the couplings and Dyson–Schwinger equations. The compatibility of these Slavnov–Taylor identities with the Hopf algebra was established in [20]. The Dyson–Schwinger equations were shown to be nicely captured by Hochschild cohomology of the Hopf algebra, see [7, 2, 14]. They also formed the starting point – via [15] – for a powerful method to obtain

solutions to these latter equations in [1]. The relation of the Hopf algebra with the gauge symmetry was discussed in [16].

In these lecture notes, we try to describe in a self-contained manner these steps towards an understanding of renormalization of perturbative gauge theories. In Section 2, we start with some background in perturbative quantum field theory. A precise definition of Feynman graphs is given in Section 3, where also the Connes–Kreimer Hopf algebra is defined. After proving some combinatorial identities, we derive at the end of that section formulas for the coproduct on full (1PI) Green’s functions. This we apply in Section 4 to two physical theories, namely quantum electrodynamics and quantum chromodynamics.

Finally, in section 5 we review the approach taken in [7, 2, 14] on Dyson–Schwinger equations in terms of Hochschild cocycles of renormalization Hopf algebras. As a new result, we prove the gauge theory theorem [14, Theorem 5].

2. Crash course in perturbative quantum field theory

We start by giving some background from physics and try to explain the origin of Feynman graphs in the perturbative approach to quantum field theory.

Our starting point is the interpretation of the path integral as a formal series in Feynman graphs. It encodes the probability amplitudes for physical processes. Let us make this more explicit by some examples taken from physics.

EXAMPLE 1. *The interaction of the photon with the electron in quantum electrodynamics (QED) is described by the following formal expansion,*

$$\text{Diagram with a shaded circle} = \text{Diagram with a wiggly line and a straight line} + \text{Diagram with a wiggly line and two straight lines} + \frac{1}{2} \text{Diagram with a wiggly line and three straight lines} + \dots$$

Here all graphs appear that can be built from the vertex that connects a wiggly line (the photon) to two straight lines (the electron). The factor $\frac{1}{2}$ is a symmetry factor (cf. Definition 11 below).

EXAMPLE 2. *The quartic gluon self-interaction in quantum chromodynamics is given by*

$$\text{Diagram with a shaded circle} = \text{Diagram with three gluons} + \text{Diagram with four gluons} + \text{Diagram with a square loop} + \dots$$

This expansion involves the gluon vertex of valence 3 and 4 (wiggly lines), as well as the quark-gluon interaction (involving two straight lines)

We shall call these formal expansions *Green’s functions*. Of course, this name originates from the theory of partial differential equations and the zeroth order terms in the above expansions are in fact Green’s functions in the usual sense. The expansion is then treated as an asymptotic series for a Green’s function for a perturbed differential operator. We use the notation $G \prec$ and $G \times$ for the Green’s function, indicating the external structure of the graphs in the above two expansions, respectively.

From these expansions, physicists can actually derive numbers, giving the probability amplitudes mentioned above. The rules of this game are known as the Feynman rules; we briefly list them for the case of quantum electrodynamics. Feynman rules for non-abelian gauge theories can be found in most standard textbooks on

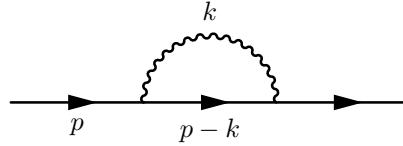
quantum field theory (see for instance [3, Section 2.14]).

Assigning momentum k to each edge of a graph, we have:

$$\begin{aligned} \text{~~~~~} &= \frac{1}{k^2 + i\epsilon} \left(-\delta_{\mu\nu} + \frac{k_\mu k_\nu}{k^2 + i\epsilon} (1 - \xi) \right) \\ \frac{1}{k} &= \frac{1}{\gamma^\mu k_\mu + m} \\ \begin{array}{c} k_3 \\ \diagup \\ k_1 \text{~~~~~} \diagdown \\ k_2 \end{array} &= -ie\gamma^\mu \delta(k_1 + k_2 + k_3) \end{aligned}$$

Here, e is the electron charge, m the electron mass and γ^μ are 4×4 Dirac gamma matrices; they satisfy $\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = -2\delta^{\mu\nu}$. Also, ϵ is an infrared regulator and $\xi \in \mathbb{R}$ is the so-called gauge fixing parameter. In addition to the above assignments, one integrates the above internal momenta k (for each internal edge) over \mathbb{R}^4 .

EXAMPLE 3. Consider the following electron self-energy graph Γ :



According to the Feynman rules, the amplitude for this graph is

$$(1) \quad U(\Gamma) = \int d^4k (e\gamma^\mu) \frac{1}{\gamma^\kappa(p_\kappa - k_\kappa) + m} (e\gamma^\nu) \left(-\frac{\delta_{\mu\nu}}{k^2 + i\epsilon} + \frac{k_\mu k_\nu}{(k^2 + i\epsilon)^2} (1 - \xi) \right)$$

with summation over repeated indices understood.

The alert reader may have noted that the above improper integral is actually not well-defined. This is the typical situation – i.e. happening for most graphs – and are the famous divergences in perturbative quantum field theory. This apparent failure can be resolved, leading eventually to spectacularly accurate predictions in physics.

The theory that proposes a solution to these divergences is called *renormalization*. This process consists of two steps: *regularization* and *subtraction*. Let us give two examples of a regularization prescription.

The first we consider is a *momentum cut-off*. This means that we perform the integral above up to a real parameter Λ . More precisely, we make the replacement

$$\int d^4k \rightsquigarrow \int_{|k| \leq \Lambda} d^4k.$$

Let us consider the type of integrations we would like to perform, in a very simplified form. If the integrand is $(k^2 + m^2)^{-2}$ then

$$\int_{|k| \leq \Lambda} d^4 k \frac{1}{(k^2 + m^2)^2} \sim \log \Lambda.$$

This explains the divergent integrals we encountered above as $\Lambda \rightarrow \infty$, but we now have a control of the divergence. Although the momentum cut-off regularization is simple and physically natural, it is not the best regularization prescription for gauge theories since it breaks the gauge invariance. Nevertheless, it is the starting point for the powerful Wilsonian approach to renormalization, which has been studied in the Hopf algebraic setup as well in [11].

Another regularization prescription is *dimensional regularization*. Instead of integrating in 4 dimensions, one integrates in $4 - z$ dimensions, with z a complex number. Of course, this only makes sense after prescribing some rules for such an integration. The key rule is the following:

$$(2) \quad \int d^D k e^{-\pi \lambda k^2} = \lambda^{-D/2} \quad (D \in \mathbb{C}).$$

This formula clearly holds for D a positive integer, where it is just the Gaussian integral. However, if we *demand* it to hold for any complex D , it turns out to provide a very convenient regularization prescription. So, let us consider once more integration over $(k^2 + m^2)^{-2}$, but now in $4 - z$ dimensions. We write using so-called Schwinger parameters, or, equivalently the Laplace transform,

$$\frac{1}{k^2 + m^2} = \int_{s>0} ds e^{-s(k^2 + m^2)}.$$

Then, using the above Eq. (2) we find that

$$\begin{aligned} \int d^{4-z} k \frac{1}{(k^2 + m^2)^2} &= \int_{s>0} ds \int_{t>0} dt \int d^{4-z} k e^{-(s+t)(k^2 + m^2)} \\ &= \pi^{2-z/2} \int_{s>0} ds \int_{t>0} dt (s+t)^{-2+z/2} e^{-(s+t)m^2} \end{aligned}$$

where we assumed that we could interchange the integrals. If we now change variables to $s = xy$ and $t = (1 - x)y$ we obtain

$$\pi^{2-z/2} \int_0^1 dx \int_{y>0} dy y^{-1+z/2} e^{-ym^2} = \pi^{2-z/2} m^{-z} \Gamma(z/2),$$

with Γ the complex gamma function. It has a pole at $z = 0$, reflecting the divergence before regularization. Again, this gives us control on the divergence.

The second step in the process of renormalization is *subtraction*. For dimensional regularization, we let T be the projection onto the pole part of Laurent series in z , i.e.,

$$T \left[\sum_{n=-\infty}^{\infty} a_n z^n \right] = \sum_{n<0} a_n z^n$$

More generally, we have a projection on the divergent part in the regularizing parameter. This is the origin of the study of Rota-Baxter algebras in the setting of quantum field theories [9]. We will however restrict ourselves to dimensional regularization with this so-called *minimal subtraction*, as it is a well suited renormalization scheme for gauge theories.

For the above graph Γ , we define the *renormalized amplitude* $R(\Gamma)$ by simply subtracting the divergent part, that is, $R(\Gamma) = U(\Gamma) - T[U(\Gamma)]$. Clearly, the result is finite for $z \rightarrow 0$. More generally, a graph Γ might have subgraphs $\gamma \subset \Gamma$ which lead to subdivergences in $U(\Gamma)$. The so-called *BPHZ-procedure* (after its inventors Bogoliubov, Parasiuk, Hepp and Zimmermann) provides a way to deal with those subdivergences in a recursive manner. It gives for the *renormalized amplitude*:

$$(3a) \quad R(\Gamma) = U(\Gamma) + C(\Gamma) + \sum_{\gamma \subset \Gamma} C(\gamma)U(\Gamma/\gamma)$$

where C is the so-called *counterterm* defined recursively by

$$(3b) \quad C(\Gamma) = -T \left[U(\Gamma) + \sum_{\gamma \subset \Gamma} C(\gamma)U(\Gamma/\gamma) \right]$$

The two sums here are over all subgraphs in a certain class; we will make this more precise in the next section.

2.1. Gauge theories. We now focus on a special class of quantum field theories – quantum gauge theories – which are of particular interest for real physical processes. Without going too much into details of what classical gauge field theories are, we focus on the consequences on the quantum side of the presence of a classical gauge symmetry. Such a gauge symmetry acts (locally) on the classical fields, thereby leaving invariant the theory that is defined, say, by a Lagrangian. These so-called gauge transformations form a group, called the *gauge group*. This group is typically infinite dimensional, since it consists of functions on space-time taking values in a Lie group. For quantum electrodynamics this Lie group is abelian and just $U(1)$, for quantum chromodynamics – the theory of gluons and quarks – it is $SU(3)$.

When (perturbatively) quantizing the gauge theory, one is confronted with this extra infinity, in addition to the divergences discussed in the previous subsection. A way to handle it is by *fixing the gauge*, in other words, by choosing an orbit under the action of the gauge group. All this can be made quite precise in the so-called *BRST-formalism*, after its inventors Becchi, Rouet, Stora and Tyutin. Although in this process the gauge symmetry completely disappears, certain identities between Green's functions appear. This is a purely ‘quantum property’ and therefore interesting to study. In addition, being potential identities between full Green's functions, it is interesting with a view towards nonperturbative quantum field theory.

For quantum electrodynamics, the identities are simple and linear in the Green's functions:

$$(4) \quad U(G \rightsquigarrow) = U(G \overleftarrow{}).$$

These are known as *Ward identities* as they were first derived by Ward in [23]. The apparent mismatch between the number of external lines on the left and right-hand-side is resolved because the vertex graphs are considered at *zero momentum transfer*. This means that the momentum on the photon line is evaluated at $p = 0$.

For non-abelian gauge theories such as quantum chromodynamics (QCD), the identities are quadratic in the Green's functions and read:

$$(5) \quad \begin{aligned} U(G \swarrow) U(G \swarrow) &= U(G \times) U(G \dashv); \\ U(G \swarrow) U(G \dashv) &= U(G \times) U(G \cdots); \\ U(G \swarrow) U(G \cdots) &= U(G \times) U(G \cdots). \end{aligned}$$

The dotted and straight line here corresponds to the ghost and quark, respectively. After their inventors, they are called the *Slavnov–Taylor identities* for the couplings [17, 22].

The importance of these identities lie in the fact that they are compatible with renormalization under the condition that gauge invariance is compatible with the regularization procedure. In fact, it turns out that dimensional regularization satisfies this requirement, see for instance [10, Section 13.1]. As a consequence, the Slavnov–Taylor identities hold after replacing U by R or C in the above formula. For instance, in the case of quantum electrodynamics one obtains the identity $Z_1 = Z_2$ that was actually derived by Ward in [23], where $Z_1 = C(G \swarrow)$ and $Z_2 = C(G \dashv)$. For quantum chromodynamics on the other hand, one derives the formulae

$$(6) \quad \frac{Z \swarrow}{Z \dashv \sqrt{Z \cdots}} = \frac{Z \swarrow}{Z \cdots \sqrt{Z \cdots}} = \frac{Z \swarrow}{(Z \cdots)^{3/2}} = \frac{\sqrt{Z \times}}{Z \cdots},$$

where the notation is as above: $Z^r := C(G^r)$. The above formula can be readily obtained from the above Slavnov–Taylor identities (5) after replacing U by C . They are the key to proving renormalizability of non-abelian gauge theories, let us try to sketch this argument.

First of all, the different interactions that are present in the theory can be weighted by a coupling constant. For example, in QCD there are four different interactions: gluon-quark, gluon-ghost, cubic and quartic gluon self-interaction. All of these come with their own coupling constants and gauge invariance (or rather, BRST-invariance) requires them to be identical. In the process of renormalization, the coupling constants are actually not constant and depend on the energy scale. This is the *running of the coupling constant* and is the origin of the renormalization group describing how they change. For QCD, the four coupling constants $g_0, \swarrow, g_0, \dashv, g_0, \cdots, g_0, \times$ are expressed in terms of the original coupling constant g as

$$(7) \quad \begin{aligned} g_0, \swarrow &= \frac{Z \swarrow}{Z \dashv \sqrt{Z \cdots}} g, & g_0, \dashv &= \frac{Z \swarrow}{Z \cdots \sqrt{Z \cdots}} g, \\ g_0, \cdots &= \frac{Z \swarrow}{(Z \cdots)^{3/2}} g, & g_0, \times &= \frac{\sqrt{Z \times}}{Z \cdots} g. \end{aligned}$$

We see that the Slavnov–Taylor identities guarantee that *the four coupling constants remain equal* after renormalization.

The above compatibility of renormalization with the Slavnov–Taylor identities is usually derived using the Zinn-Justin equation (or the more general BV-formalism) relying heavily on path integral techniques. Our goal in the next sections

is to derive this result taking the formal expansion of the Green's functions in Feynman graphs as a starting point. We will work in the setting of the Connes–Kreimer Hopf algebra of renormalization.

3. The Hopf algebra of Feynman graphs

We start with some definitions on Feynman graphs and their symmetries, thereby making precise several properties needed later.

3.1. Feynman graphs. The Feynman graphs we will consider are built from a certain set of edges and vertices R , and following [14] we write $R = R_V \cup R_E$. For example, in massive ϕ^3 -theory, the set R_V contains the bi- and trivalent vertex and R_E the straight line, but more interesting theories such as gauge theories contain different types of edges and vertices (for example involving curly, dotted and straight lines) corresponding to different particles. More precisely, we have the following definition [6].

DEFINITION 4. A Feynman graph Γ is given by a set $\Gamma^{[0]}$ of vertices each of which is an element in R_V and $\Gamma^{[1]}$ of edges in R_E , and maps

$$\partial_j : \Gamma^{[1]} \rightarrow \Gamma^{[0]} \cup \{1, 2, \dots, N\}, \quad j = 0, 1,$$

that are compatible with the type of vertex and edge as parametrized by R_V and R_E , respectively. Moreover, we exclude the case that ∂_0 and ∂_1 are both in $\{1, 2, \dots, N\}$. The set $\{1, 2, \dots, N\}$ labels the external lines, so that $\sum_j \text{card } \partial_j^{-1}(v) = 1$ for all $v \in \{1, \dots, N\}$.

The set of external lines is $\Gamma_{\text{ext}}^{[1]} = \cup_i \partial_i^{-1}\{1, \dots, N\}$ and its complement $\Gamma_{\text{int}}^{[1]}$ in $\Gamma^{[1]}$ is the set of internal lines.

Note that the elements in $\Gamma_{\text{ext}}^{[1]}$ can thus be labeled as e_1, \dots, e_N where $e_k := \cup_i \partial_i^{-1}(k)$ and we understand this labeling as being fixed.

REMARK 5. An equivalent way of looking at the above ‘types’ of edges and vertices is the following, staying more closely to the conventional graph theory. Instead of assigning a certain ‘type’ to each edge, we could just as well color them, with the color now corresponding to a propagating particle in our theory. A vertex then represents an interaction of the particles that correspond to the colors that meet at that vertex. It happened so that nature has chosen only certain interactions between certain particles, so we restrict ourselves to graphs that only involve the vertices with allowed coloring.

If a Feynman graph Γ has two external lines, both corresponding to the same field, we would like to distinguish between propagators and mass terms. In more mathematical terms, since we have vertices of valence two, we would like to indicate whether a graph with two external lines corresponds to such a vertex, or to an edge. A graph Γ with two external lines is dressed by a bullet when it corresponds to a vertex, i.e. we write Γ_\bullet . The above correspondence between Feynman graphs and vertices/edges is given by the *residue* $\text{res}(\Gamma)$. It is defined as the vertex or edge the graph corresponds to after collapsing all its internal points. For example, we have:

$$\text{res} \left(\begin{array}{c} \text{wavy} \\ \text{bullet} \\ \text{wavy} \end{array} \right) = \text{wavy} \quad \text{and} \quad \text{res} \left(\begin{array}{c} \text{bullet} \\ \text{wavy} \\ \text{wavy} \end{array} \right) = \text{---}$$

but if the last graph is dressed with a bullet it is understood to correspond to a valence 2 vertex:

$$\text{res} \left(\begin{array}{c} \text{wiggly loop} \\ \hline \text{straight line} \end{array} \bullet \right) = \bullet \bullet.$$

For the purpose of renormalization, one is mainly interested in one-particle irreducible Feynman graphs with residues in the set R .

DEFINITION 6. *A Feynman graph is called one-particle irreducible (1PI) if it is not a tree and can not be disconnected by removal of a single edge.*

For example, all graphs in this paper are one-particle irreducible, *except* the following which is one-particle reducible:

$$\text{wiggly loop} \text{---} \text{wiggly loop}.$$

3.2. The Hopf algebra structure on graphs. We restrict to the class of 1PI Feynman graphs Γ for which $\text{res}(\Gamma) \in R$ and will denote a generic graph with residue $r \in R$ by Γ^r . If it also has loop number L , we denote it by Γ_L^r .

DEFINITION 7 (Connes–Kreimer [4]). *The Hopf algebra H of Feynman graphs is the free commutative algebra (over \mathbb{C}) generated by all 1PI Feynman graphs with residue in R , with counit $\epsilon(\Gamma) = 0$ unless $\Gamma = \emptyset$, in which case $\epsilon(\emptyset) = 1$, coproduct,*

$$\Delta(\Gamma) = \Gamma \otimes 1 + 1 \otimes \Gamma + \sum_{\gamma \subsetneq \Gamma} \gamma \otimes \Gamma/\gamma,$$

where the sum is over disjoint unions of 1PI subgraphs with residue in R . The quotient Γ/γ is defined to be the graph Γ with the connected components of the subgraph contracted to the corresponding vertex/edge. If a connected component γ' of γ has two external lines, then there are possibly two contributions corresponding to the valence two vertex and the edge; the sum involves the two terms $\gamma_\bullet \otimes \Gamma/(\gamma \rightarrow \bullet)$ and $\gamma \otimes \Gamma/\gamma$.

The antipode is given recursively by,

$$S(\Gamma) = -\Gamma - \sum_{\gamma \subsetneq \Gamma} S(\gamma)\Gamma/\gamma.$$

The above Hopf algebra is an example of a connected graded Hopf algebra: it is graded by the *loop number* $L(\Gamma)$ of a graph Γ . Indeed, one checks that the coproduct (and obviously also the product) satisfy the grading by loop number and H^0 consists of complex multiples of the empty graph, which is the unit in H , so that $H^0 = \mathbb{C}1$. We denote by q_l the projection in H onto H^l .

In addition, there is another grading on this Hopf algebra. It is given by the number of vertices and already appeared in [4]. However, since we consider vertices and edges of different types (wiggly, dotted, straight, *et cetera*), we extend to a multigrading as follows. As in [20], we denote by $m_{\Gamma,r}$ the number of vertices/internal edges of type r appearing in Γ , for $r \in R$. Moreover, let $n_{\gamma,r}$ be the number of connected components of γ with residue r . For each $v \in R_V$ we define a degree d_v by

$$d_v(\Gamma) = m_{\Gamma,v} - n_{\Gamma,v}.$$

The multidegree indexed by R_V is compatible with the Hopf algebra structure as follows easily from the following relation:

$$m_{\Gamma/\gamma,v} = m_{\Gamma,v} - m_{\gamma,v} + n_{\gamma,v},$$

and the fact that $m_{\Gamma\Gamma',v} = m_{\Gamma,v} + m_{\Gamma',v}$, and $n_{\Gamma\Gamma',v} = n_{\Gamma,v} + n_{\Gamma',v}$. This gives a decomposition

$$H = \bigoplus_{n_1, \dots, n_k \in \mathbb{Z}^k} H^{n_1, \dots, n_k},$$

where $k = |R_V|$. We denote by p_{n_1, \dots, n_k} the projection onto H^{n_1, \dots, n_k} .

LEMMA 8. *There is the following relation between the grading by loop number and the multigrading by number of vertices:*

$$\sum_{v \in R_V} (N(v) - 2)d_v = 2L$$

where $N(v)$ is the number of lines attached to v , i.e., $N(v) := |\partial(v)|$.

PROOF. This can be easily proven by induction on the number of internal edges using invariance of the quantity $\sum_v (N(v) - 2)d_v - 2L$ under the adjoint of an edge. \square

3.2.1. *Renormalization as a Birkhoff decomposition.* We now demonstrate how to obtain Equation (3) for the renormalized amplitude and the counterterm for a graph as a Birkhoff decomposition in the group of characters of H . Let us first recall the definition of a Birkhoff decomposition.

We let $l : C \rightarrow G$ be a loop with values in an arbitrary complex Lie group G , defined on a smooth simple curve $C \subset \mathbb{P}_1(\mathbb{C})$. Let C_{\pm} be the two complements of C in $\mathbb{P}_1(\mathbb{C})$, with $\infty \in C_-$. A *Birkhoff decomposition* of l is a factorization of the form

$$l(z) = l_-(z)^{-1}l_+(z); \quad (z \in C),$$

where l_{\pm} are (boundary values of) two holomorphic maps on C_{\pm} , respectively, with values in G . This decomposition gives a natural way to extract finite values from a divergent expression. Indeed, although $l(z)$ might not holomorphically extend to C_+ , $l_+(z)$ is clearly finite as $z \rightarrow 0$.

We now look at the group $G(K) = \text{Hom}_{\mathbb{Q}}(H, K)$ of K -valued characters of a connected graded commutative Hopf algebra H , where K is the field of convergent Laurent series in z .¹ The product, inverse and unit in the group $G(K)$ are defined by the respective equations:

$$\begin{aligned} \phi * \psi(X) &= \langle \phi \otimes \psi, \Delta(X) \rangle, \\ \phi^{-1}(X) &= \phi(S(X)), \\ e(X) &= \epsilon(X), \end{aligned}$$

for $\phi, \psi \in G(K)$. We claim that a map $\phi \in G(K)$ is in one-to-one correspondence with loops l on an infinitesimal circle around $z = 0$ and values in $G(\mathbb{Q}) = \text{Hom}_{\mathbb{Q}}(H, \mathbb{Q})$. Indeed, the correspondence is given by

$$\phi(X)(z) = l(z)(X),$$

¹In the language of algebraic geometry, there is an affine group scheme G represented by H in the category of commutative algebras. In other words, $G = \text{Hom}_{\mathbb{Q}}(H, \cdot)$ and $G(K)$ are the K -points of the group scheme.

and to give a Birkhoff decomposition for l is thus equivalent to giving a factorization $\phi = \phi_-^{-1} * \phi_+$ in $G(K)$. It turns out that for graded connected commutative Hopf algebras such a factorization exists.

THEOREM 9 (Connes–Kreimer [4]). *Let H be a graded connected commutative Hopf algebra. The Birkhoff decomposition of $l : C \rightarrow G$ (given by an algebra map $\phi : H \rightarrow K$) exists and is given dually by*

$$\phi_-(X) = \epsilon(X) - T[m(\phi_- \otimes \phi)(1 \otimes (1 - \epsilon)\Delta(X))]$$

and $\phi_+ = \phi_- * \phi$.

The graded connected property of H assures that the recursive definition of ϕ_- actually makes sense. In the case of the Hopf algebra of Feynman graphs defined above, the factorization takes the following form:

$$\begin{aligned}\phi_-(\Gamma) &= -T \left[\phi(\Gamma) + \sum_{\gamma \subsetneq \Gamma} \phi_-(\gamma)\phi(\Gamma/\gamma) \right] \\ \phi_+(\Gamma) &= \phi(\Gamma) + \phi_-(\Gamma) + \sum_{\gamma \subsetneq \Gamma} \phi_-(\gamma)\phi(\Gamma/\gamma)\end{aligned}$$

The key point is now that the Feynman rules actually define an algebra map $U : H \rightarrow K$ by assigning to each graph Γ the regularized Feynman rules $U(\Gamma)$, which are Laurent series in z . When compared with Equations (3) one concludes that the algebra maps U_+ and U_- in the Birkhoff factorization of U are precisely the renormalized amplitude R and the counterterm C , respectively. Summarizing, we can write the BPHZ-renormalization as the Birkhoff decomposition $U = C^{-1} * R$ of the map $U : H \rightarrow K$ dictated by the Feynman rules.

Although the above construction gives a very nice geometrical description of the process of renormalization, it is a bit unphysical in that the Hopf algebra relies on individual graphs. Rather, as mentioned before, in physics the probability amplitudes are computed from the full expansion of Green's functions. Individual graphs do not correspond to physical processes and therefore a natural question to pose is how the Hopf algebra structure behaves at the level of the Green's functions. Let us make the following definition.

DEFINITION 10. *For a vertex or edge $r \in R$ we define the 1PI Green's function by*

$$(8) \quad G^r = 1 \pm \sum_{\text{res}(\Gamma)=r} \frac{\Gamma}{\text{Sym}(\Gamma)}$$

where the sign is $+$ if r is a vertex and $-$ if it is an edge. Finally, we denote the restriction of the sum to graphs Γ at loop order $L(\Gamma) = L$ by G_L^r .

We are particularly interested in the form of the coproduct on 1PI Green's functions, and more generally, the Hopf algebra structure of Green's functions. Before we derive the coproduct on these Green's functions, we need a significant amount of combinatorial relations and identities.

$$\text{Sym}(\text{---○---}) = 2 \quad \text{Sym}(\text{------}) = 1$$

FIGURE 1. Automorphisms of Feynman graphs respect the type of vertex/edge in R .

3.3. Intermezzo: Counting Feynman graphs. In this subsection we recall some of the combinatorial properties we have derived in [19]. Let us start with the natural notion of graph automorphisms, taking Definition 4 of Feynman graphs above as a starting point.

DEFINITION 11. *An automorphism of a Feynman graph Γ is given by an isomorphism $g^{[0]}$ from $\Gamma^{[0]}$ to itself, and an isomorphism $g^{[1]}$ from $\Gamma^{[1]}$ to itself that is the identity on $\Gamma_{\text{ext}}^{[1]}$ and such that for all $e \in \Gamma^{[1]}$,*

$$(9) \quad \cup_j g^{[0]}(\partial_j(e)) = \cup_j \partial_j(g^{[1]}(e)).$$

Moreover, we require $g^{[0]}$ and $g^{[1]}$ to respect the type of vertex/edge in the set R .

The automorphism group $\text{Aut}(\Gamma)$ of Γ consists of all such automorphisms; its order is called the symmetry factor of Γ and is denoted by $\text{Sym}(\Gamma)$.

Similarly, there is a notion of an isomorphism of two graphs Γ and Γ' as a pair of maps that intertwines the maps ∂_i as in Eq. (9). We remark that we correct in this way for the apparent orientation given by the two maps ∂_0 and ∂_1 and we stress that the fermionic lines are unoriented. We take the complex character of the fermionic fields into account by summing over all possible orientations once we apply the Feynman rules.

Note that for $\Gamma = \prod_i \Gamma_i$ the disjoint union of n graphs, the symmetry factor is given by $\text{Sym}(\Gamma) = n_1! \cdots n_k! \text{Sym}(\Gamma_1) \cdots \text{Sym}(\Gamma_n)$ where n_i are the numbers of isomorphic (with fixed external lines) connected components of Γ' . Equivalently, one has for a 1PI graph Γ' ,

$$(10) \quad \text{Sym}(\Gamma \Gamma') = n_c(\Gamma, \Gamma') \text{Sym}(\Gamma) \text{Sym}(\Gamma'),$$

with $n_c(\Gamma, \Gamma')$ the number of connected components of $\Gamma\Gamma'$ that are isomorphic to Γ' .

The above definition of automorphism differs from the usual notion of graph automorphism (cf. for instance [8]) in that the latter might also permute the elements in $\{1, \dots, N\}$ when understood as external vertices. In the above notation, such an automorphism of Γ would be given by an isomorphism $g^{[0]}$ from $\Gamma^{[0]} \cup \{1, \dots, N\}$ to itself, and an isomorphism $g^{[1]}$ from $\Gamma^{[1]}$ to itself such that Equation (9) holds.

If Γ is a connected Feynman graph with external lines labeled by $\{1, \dots, N\}$, we can construct another graph Γ^σ , by permuting the external lines by an element $\sigma \in S_N$, respecting the type of external lines. The graph Γ^σ is given by the same sets $\Gamma^{[0]}$ and $\Gamma^{[1]}$ but with maps

$$\partial_j^\sigma := \sigma \circ \partial_j : \Gamma^{[1]} \rightarrow \Gamma^{[0]} \cup \{1, \dots, N\}.$$

This permutation affects the labeling of the external lines by $\{1, \dots, N\}$, which explains the terminology permutation of external lines; we write e^σ for the edge in Γ^σ corresponding to an edge $e \in \Gamma^{[1]}$ under the permutation σ .

DEFINITION 12. *A permutation σ of the external lines of Γ is called trivial if there exists an isomorphism between Γ^σ and Γ , leaving the labeling of the external lines fixed.*

The number of non-isomorphic graphs Γ^σ obtained by a permutation σ of the external lines of Γ , is denoted by $|\Gamma|_\vee$ and extended to disconnected graphs by $|\Gamma\Gamma'|_\vee = |\Gamma|_\vee|\Gamma'|_\vee$.

LEMMA 13. *A permutation σ of the external lines of Γ is trivial if and only if there exists an automorphism g of the graph Γ not necessarily leaving the external lines fixed, such that $g^{[0]}|_{\{1, \dots, N\}} = \sigma$.*

PROOF. Firstly, if σ is trivial, there exists an isomorphism $f : \Gamma^\sigma \rightarrow \Gamma$ and the pair $(f^{[0]} \circ \sigma, f^{[1]} \circ \sigma)$ is an automorphism g of Γ (without fixed external vertices), since,

$$\cup_j g^{[0]}(\partial_j(e)) = \cup_j f^{[0]}(\partial_j^\sigma(e^\sigma)) = \cup_j \partial_j(f^{[1]}(e^\sigma)) = \cup_j \partial(g^{[1]}(e)).$$

On the other hand, such an automorphism g is given by two maps $g^{[0]}$ and $g^{[1]}$, where $g^{[0]}$ is the product of two permutations of the disjoint sets $\Gamma^{[0]}$ and $\{1, \dots, N\}$, say $f^{[0]}$ and σ , respectively. Correspondingly, σ acts on $\Gamma_{\text{ext}}^{[1]}$ by permutation, so that also $g^{[1]} = f^{[1]} \circ \sigma$. This factorization gives rise to an isomorphism f from Γ^σ to Γ , which leaves external lines fixed. \square

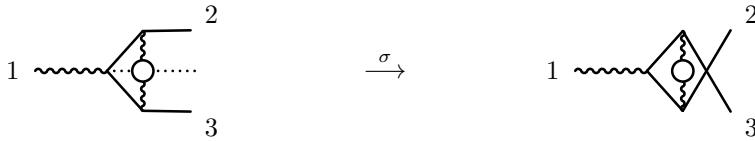


FIGURE 2. The permutation $\sigma = (23)$ of the external lines of the graph Γ is trivial since reflection in the dotted line induces an automorphism g of Γ such that $g^{[0]}|_{\{1,2,3\}} = \sigma$. Moreover, this is the only trivial permutation so that $|\Gamma|_\vee = 3!/2 = 3$

DEFINITION 14. *An insertion place for a (connected) graph γ in Γ is the subset of $\Gamma^{[0]} \cup \Gamma_{\text{int}}^{[1]}$ consisting of vertices/internal edges of the form $r = \text{res}(\gamma)$. It can be extended to disconnected graphs $\gamma = \prod_{i=1}^n \gamma_i$ by giving n -tuples of insertion places for $\gamma_1, \dots, \gamma_n$, thereby allowing several insertions of the connected components with residue r in R_E on the same internal edge in Γ of the form r . The number of insertion places for γ in Γ is denoted by $\Gamma \mid \gamma$.*

An explicit expression for $\Gamma \mid \gamma$ can be obtained as follows [14]. Recall the notation $m_{\Gamma, r}$ for the number of vertices/edges r in $\Gamma^{[0]} \cup \Gamma^{[1]}$, for $r \in R$, and $n_{\gamma, r}$ for the number of connected components of γ with residue r . Since insertion of a vertex graph (i.e. with residue in R_V) on a $v \in \Gamma^{[0]}$ prevents a subsequent insertion at v of a vertex graph with the same residue, whereas insertion of an edge graph

(i.e. with residue in R_E) creates two new edges and hence two insertion places for a subsequent edge graph, we find the following expression,

$$(11) \quad \Gamma | \gamma = \prod_{v \in R_V} n_{\gamma,v}! \binom{m_{\Gamma,v}}{n_{\gamma,v}} \prod_{e \in R_E} n_{\gamma,e}! \binom{m_{\Gamma,e} + n_{\gamma,e} - 1}{n_{\gamma,e}}.$$

Indeed, the binomial coefficients arises for each vertex v since we are choosing $n_{\gamma,v}$ out of $m_{\Gamma,v}$ whereas for an edge e we choose $n_{\gamma,e}$ out of $m_{\Gamma,e}$ with repetition. We extend this definition to empty graphs by defining $\Gamma | \emptyset = \emptyset | \gamma = \emptyset | \emptyset = 1$ for a 1PI graph γ , and $\emptyset | \gamma = 0$ for a disconnected graph γ .

REMARK 15. *Our expression for $\Gamma | \gamma$ differs slightly from the one given in [14] where additional factors of $1/n_{\gamma,r}!$ are present for $r \in R$. It turns out that the above expression appears naturally in the coproduct on 1PI Green's functions (see below).*

A few examples are in place:

$$\text{Diagram 1: } \text{Diagrammatic expression} = \binom{2}{1} = 2 \quad \text{whereas} \quad \text{Diagram 2: } \text{Diagrammatic expression} = 2! \binom{3}{2} = 6.$$

DEFINITION 16. *An insertion of a connected graph γ at the insertion place x in Γ , is given by a bijection between the set $\gamma_{\text{ext}}^{[1]}$ of external lines of γ and the set $\partial^{-1}(x)$. If $x \in \Gamma^{[0]}$, $\partial^{-1}(x)$ denotes the set of lines attached to the vertex x , and if $x \in \Gamma_{\text{ext}}^{[1]}$, $\partial^{-1}(x)$ denotes the set of adjacent edges to any internal point of x . The graph obtained in this way is denoted by $\Gamma \circ_{(x,\phi)} \gamma$.*

Two insertions (x,ϕ) and (x',ϕ') are called equivalent if $x = x'$ and $\phi' = \phi \circ \sigma$ for some trivial permutation σ of the external lines of γ . The set of all insertions of γ in Γ up to equivalence is denoted by $X(\Gamma, \gamma)$; it consists of equivalence classes $[x, \phi]$.

This definition of equivalence relation on insertions is motivated by the fact that $\Gamma \circ_{(x,\phi)} \gamma \simeq \Gamma \circ_{(x',\phi')} \gamma$ whenever $(x,\phi) \sim (x',\phi')$. We extend $X(\Gamma, \gamma)$ to disconnected graphs γ as follows. If $\gamma = \prod_{i=1}^n \gamma_i$ is the disjoint union of n graphs, the set $X(\Gamma, \gamma)$ of insertions of γ in Γ is defined as the set of n -tuples of pairs $([x_1, \phi_1], \dots, [x_n, \phi_n])$, where $[x_1, \phi_1] \in X(\Gamma, \gamma_1)$ and $[x_{k+1}, \phi_{k+1}]$ is an element in $X(\Gamma \circ_{(x_1, \phi_1)} \dots \circ_{(x_k, \phi_k)} \prod_{i=1}^k \gamma_i, \gamma_{k+1})$ which is not part of any of the inserted graphs $\gamma_1, \dots, \gamma_{k-1}$ for $k = 1, \dots, n-1$. The cardinality of $X(\Gamma, \gamma)$ is the number $\Gamma | \gamma$ of insertion places for γ in Γ times the number $|\gamma|_{\vee}$ of non-trivial permutations of the external lines of γ .

We also need the following generalization for the number of insertion places.

DEFINITION 17. *Let Γ, γ, γ' be three (disjoint unions of) 1PI graphs. We define $\Gamma | \gamma | \gamma'$ to be the number of places to insert γ into Γ (say, at x using ϕ) and then subsequently insert γ' in $\Gamma \circ_{(x,\phi)} \gamma$. In other words,*

$$\Gamma | \gamma | \gamma' := \frac{1}{|\gamma|_{\vee}} \sum_{[x,\phi] \in X(\Gamma, \gamma)} \Gamma \circ_{(x,\phi)} \gamma | \gamma'.$$

Moreover, we set $\Gamma | \emptyset | \gamma' = \Gamma | \gamma'$ and $\emptyset | \gamma | \gamma' = \gamma | \gamma'$ if γ is 1PI and $\emptyset | \gamma | \gamma' = 0$ if γ is disconnected.

The factor $1/|\gamma|_\vee$ corrects for the overcounting due to the several (not equivalent) ways to insert γ into Γ at a particular place. Note that automatically $\Gamma \mid \gamma \mid \emptyset = \Gamma \mid \gamma$ and if $\Gamma, \gamma, \gamma' \neq \emptyset$, we have

$$(12) \quad \Gamma \mid \gamma \mid \gamma' = \Gamma \mid \gamma\gamma' + (\Gamma \mid \gamma)(\gamma \mid \gamma').$$

Suppose γ is a 1PI graph. There is a natural action of $\text{Aut}(\Gamma)$ on $X(\Gamma, \gamma)$ given by

$$g \cdot (x, \phi) = (gx, g \circ \phi).$$

One readily checks that this action respects the equivalence relation on insertions, and therefore acts on the equivalence classes $[x, \phi]$. Moreover, an element $g \in \text{Aut}(\Gamma)$ naturally induces an isomorphism $\Gamma \circ_{(x, \phi)} \gamma$ to $\Gamma \circ_{g(x, \phi)} \gamma$.

For an element $[x, \phi]$ in $X(\Gamma, \gamma)$, we denote by $M(x, \phi)$ the number of graphs γ' in $\Gamma \circ_{(\phi, x)} \gamma$ that are images of γ under some element in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$. Moreover, $N(x, \phi)$ denotes the number of orbits $\text{Aut}(\Gamma)[x', \phi']$ such that $\Gamma \circ_{(x', \phi')} \gamma \simeq \Gamma \circ_{(x, \phi)} \gamma$. Both definitions are independent of the choice of a representative (x, ϕ) as well as the choice of the element $[x', \phi']$ in the orbit. Indeed, an element g in $\text{Aut}(\Gamma)$ will induce a natural isomorphism $\Gamma \circ_{(x', \phi')} \gamma \simeq \Gamma \circ_{g(x', \phi')} \gamma$.

LEMMA 18. *Suppose γ is a 1PI graph and let $[x, \phi] \in X(\Gamma, \gamma)$. The length of the orbit $\text{Aut}(\Gamma)[x, \phi]$ is given by*

$$|\text{Aut}(\Gamma)[x, \phi]| = \frac{\text{Sym}(\gamma)\text{Sym}(\Gamma)M(x, \phi)}{\text{Sym}(\Gamma \circ_{(x, \phi)} \gamma)}.$$

PROOF. We use the orbit-stabilizer theorem, stating in this case that the orbit $\text{Aut}(\Gamma)[x, \phi]$ is isomorphic to the left cosets of the stabilizer $\text{Aut}(\Gamma)_{[x, \phi]}$ in $\text{Aut}(\Gamma)$. In particular, we have for its length,

$$|\text{Aut}(\Gamma)[x, \phi]| = [\text{Aut}(\Gamma) : \text{Aut}(\Gamma)_{[x, \phi}]] = \frac{|\text{Aut}(\Gamma)|}{|\text{Aut}(\Gamma)_{[x, \phi]}|}.$$

The order of $\text{Aut}(\Gamma)_{[x, \phi]}$ can be computed as follows. Let $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma$ be the subgroup of $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$ consisting of automorphisms that map γ to itself (but possibly permuting the external lines of γ). There is a short exact sequence of groups

$$1 \rightarrow \text{Aut}(\gamma) \rightarrow \text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma \rightarrow \text{Aut}(\Gamma)_{[x, \phi]} \rightarrow 1.$$

Indeed, the image \tilde{g} inside $\text{Aut}(\Gamma)$ of an element g in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma$ is defined by restricting g to $\Gamma - \{x\}$ and by the identity map on the vertex x . Then, by Lemma 13, \tilde{g} might permute the edges connected to the vertex x but always in a trivial way, since g induces an automorphism of γ not necessarily leaving its external lines fixed. Therefore, $\tilde{g}(x, \phi) = (x, \phi \circ \sigma)$ for some trivial permutation σ of $\gamma_{\text{ext}}^{[1]}$, so that it is an element in the fixed point subgroup $\text{Aut}(\Gamma)_{[x, \phi]}$. Moreover, the kernel of the map that sends such a g to \tilde{g} consists precisely of those elements in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma$ that correspond to the identity on Γ ; in other words, these are automorphisms of γ that leave external lines fixed.

We conclude that the quotient group $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma / \text{Aut}(\gamma)$ is isomorphic to $\text{Aut}(\Gamma)_{[x, \phi]}$. Since $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$ is generated by the elements in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma$ and automorphisms that map γ isomorphically to a subgraph γ' of Γ , we see that

$$|\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)_\gamma| = \frac{|\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)|}{M(x, \phi)}.$$

Combining these results, we conclude that

$$|\text{Aut}(\Gamma)[x, \phi]| = \frac{|\text{Aut}(\gamma)| |\text{Aut}(\Gamma)|}{|\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)|} = \frac{\text{Sym}(\gamma) \text{Sym}(\Gamma) M(x, \phi)}{\text{Sym}(\Gamma \circ_{(x, \phi)} \gamma)}.$$

□

As a final preparation to the next section, we will write the coproduct as a sum of maps Δ_γ , with γ a disjoint union of 1PI graphs (with fixed external lines). They are given by

$$\Delta_\gamma(\Gamma) = \sum_{\gamma' \subset \Gamma, \gamma' \simeq \gamma} \Gamma / \gamma',$$

and defined to be zero if Γ contains no subgraphs isomorphic to γ . In particular, Δ_\emptyset is the identity map, $\Delta_\Gamma(\Gamma) = \emptyset$ and $\Delta_\gamma(\emptyset) = 0$ if $\gamma \neq \emptyset$. However, since only subgraphs isomorphic to γ enter in this formula – hence no reference is made to a particular labeling of the external lines of γ – we have to correct by a factor of $|\gamma|_V$ if we are to sum over all disjoint unions of 1PI graphs *with fixed external lines*,

$$\Delta = \sum_{\gamma} \frac{1}{|\gamma|_V} \gamma \otimes \Delta_\gamma.$$

We recall the following combinatorial factor from [4]; for a given Γ, γ, Γ' , we denote by $n(\Gamma, \gamma, \Gamma')$ the number of subgraphs $\gamma' \simeq \gamma$ in Γ such that $\Gamma / \gamma \simeq \Gamma'$. With this definition, we can write

$$(13) \quad \Delta_\gamma(\Gamma) = \sum_{\Gamma'} n(\Gamma, \gamma, \Gamma') \Gamma',$$

which also yields the following formula for the coproduct,

$$\Delta(\Gamma) = \sum_{\gamma, \Gamma'} \frac{n(\Gamma, \gamma, \Gamma')}{|\gamma|_V} \gamma \otimes \Gamma'.$$

REMARK 19. *From this last formula, one easily derives the Lie bracket on Feynman graphs as defined in [4]. Indeed, one can define a pre-Lie product between 1PI graphs Γ_1, Γ_2 by duality*

$$\langle \Gamma_1 * \Gamma_2, \Gamma \rangle := \langle \Gamma_1 \otimes \Gamma_2, \Delta(\Gamma) \rangle,$$

*with the pairing given by $\langle \Gamma_1, \Gamma_2 \rangle = 1$ if $\Gamma_1 \simeq \Gamma_2$ and zero otherwise. This pre-Lie product defines a Lie bracket by $[\Gamma_1, \Gamma_2] = \Gamma_1 * \Gamma_2 - \Gamma_2 * \Gamma_1$ with $*$ given explicitly by*

$$\Gamma_1 * \Gamma_2 = \sum_{\Gamma} \frac{n(\Gamma, \Gamma_1, \Gamma_2)}{|\Gamma_1|_V} \Gamma.$$

LEMMA 20. *If Γ and γ are nonempty (connected) 1PI graphs, then*

$$n(\Gamma \circ_{(x, \phi)} \gamma, \gamma, \Gamma) = M(x, \phi) N(x, \phi).$$

PROOF. We have to count the number of subgraphs $\gamma' \simeq \gamma$ of $\Gamma \circ_{(x, \phi)} \gamma$ such that there is an isomorphism $(\Gamma \circ_{(x, \phi)} \gamma) / \gamma' \simeq \Gamma$.

This isomorphism can be trivial in the sense that there exists an element in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$ mapping γ' to γ . Otherwise, the existence of such an isomorphism implies that there is an isomorphism $\Gamma \circ_{(x, \phi)} \gamma \simeq \Gamma \circ_{(x', \phi')} \gamma$, with (x', ϕ') the image in Γ of $\text{res}(\gamma')$ in the quotient $(\Gamma \circ_{(x, \phi)} \gamma) / \gamma'$; such an isomorphism maps γ in $\Gamma \circ_{(x, \phi)} \gamma$ to a certain subgraph γ' of Γ . Moreover, $[x, \phi]$ and $[x', \phi']$ are in disjoint

$\text{Aut}(\Gamma)$ -orbits, since if $(x', \phi') = g(x, \phi)$, the isomorphism would be the composition of an element in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$ and an element in $\text{Aut}(\Gamma)$.

We claim that all subgraphs γ' obtained in this way (for disjoint orbits) are all different subgraphs of Γ , and cannot be the image of γ under the action of an element in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$. This would then lead to $M(x, \phi)N(x, \phi)$ many subgraphs γ' of $\Gamma \circ_{(x, \phi)} \gamma$ satisfying $(\Gamma \circ_{(x, \phi)} \gamma)/\gamma' \simeq \Gamma$.

Let $[x, \phi], [x', \phi'], [x'', \phi''] \in X(\Gamma, \gamma)$ be in disjoint orbits and suppose that there are isomorphisms

$$\begin{aligned} g' : \Gamma \circ_{(x', \phi')} \gamma &\rightarrow \Gamma \circ_{(x, \phi)} \gamma, \\ g'' : \Gamma \circ_{(x'', \phi'')} \gamma &\rightarrow \Gamma \circ_{(x, \phi)} \gamma, \end{aligned}$$

mapping γ to subgraphs γ' and γ'' in Γ , respectively. If γ' and γ'' coincide (up to an isomorphism h), then the composition $(g'')^{-1} \circ h \circ g'$ gives an isomorphism from $\Gamma \circ_{(x', \phi')} \gamma$ to $\Gamma \circ_{(x'', \phi'')} \gamma$ mapping γ to itself. It therefore induces an element in $\text{Aut}(\Gamma)$ that sends $[x', \phi']$ to $[x'', \phi'']$, which cannot be true. We conclude that γ' and γ'' are different subgraphs of Γ .

On the other hand, if there is an element ϕ in $\text{Aut}(\Gamma \circ_{(x, \phi)} \gamma)$ that maps γ to such a subgraph $\gamma' \in \Gamma$, the composition $\phi^{-1} \circ g'$ would map $\Gamma \circ_{(x', \phi')} \gamma$ to $\Gamma \circ_{(x, \phi)} \gamma$ isomorphically, sending γ to itself. Again, such a map must be induced by an element in $\text{Aut}(\Gamma)$ mapping $[x, \phi]$ to $[x', \phi']$, contradicting our assumptions. \square

LEMMA 21. *Let γ, γ' be as above. Then,*

$$\Delta_{\gamma\gamma'} = \frac{1}{n_c(\gamma, \gamma')} \Delta_\gamma \Delta_{\gamma'} - \rho_{\gamma\gamma'},$$

where $n_c(\gamma, \gamma')$ is the number of connected components of $\gamma\gamma'$ (cf. Eq. (10)) and $\rho_{\gamma\gamma'}$ is defined by

$$\rho_{\gamma\gamma'} = \sum_{[y, \psi] \in X(\gamma, \gamma')} \frac{\text{Sym}(\gamma \circ_{(y, \psi)} \gamma')}{\text{Sym}(\gamma\gamma')} \Delta_{\gamma \circ_{(y, \psi)} \gamma'}.$$

PROOF. Consider $\Delta_{\gamma\gamma'}(\Gamma)$ on a 1PI graph Γ ; if γ and γ' appear as disjoint subgraphs of Γ , this expression is given by $\Delta_\gamma \Delta_{\gamma'}(\Gamma)$, up to a factor of $n_c(\gamma, \gamma')$ which corrects for the overcounting. Indeed, let $\gamma_1, \dots, \gamma_m$ denote all subgraphs of Γ that are isomorphic to γ . If $m \geq n + 1$, then

$$\begin{aligned} \Delta_{\gamma^{n+1}}(\Gamma) &= \sum_{\substack{\{i_1, \dots, i_{n+1}\} \\ \subset \{1, \dots, m\}}} \frac{1}{(n+1)!} \Gamma/\gamma_{i_1} \cdots \gamma_{i_{n+1}}; \\ \Delta_{\gamma^n} \Delta_\gamma(\Gamma) &= \sum_{i=1}^m \sum_{\substack{\{i_1, \dots, i_n\} \\ \subset \{1, \dots, \hat{i}, \dots, m\}}} \frac{1}{n!} \Gamma/\gamma_i \gamma_{i_1} \cdots \gamma_{i_n}, \end{aligned}$$

leading precisely to the factor $n_c(\gamma^n, \gamma) = n + 1$. On the other hand, if $m < n + 1$, then both terms vanish.

In the case that Γ contains a subgraph $\tilde{\gamma}$ such that $\tilde{\gamma}/\gamma' \simeq \gamma$, we find a discrepancy between the two terms which is given by the following sum,

$$\rho_{\gamma\gamma'}(\Gamma) = \frac{1}{n_c(\gamma, \gamma')} \sum_{\tilde{\gamma} \subset \Gamma, \tilde{\gamma}/\gamma' \simeq \gamma} n(\tilde{\gamma}, \gamma', \gamma) \Gamma/\tilde{\gamma}.$$

Here $n(\tilde{\gamma}, \gamma', \gamma)$ is by definition the number of disjoint subgraphs of $\tilde{\gamma}$ that are isomorphic to γ' and such that $\tilde{\gamma}/\gamma' \simeq \gamma$, which do indeed all contribute to $\Delta_\gamma \Delta_{\gamma'}(\Gamma)$. We replace the above sum by a sum over insertion places of γ' in γ , while correcting for the equivalent insertions. The latter correcting factor is given as the number of elements $[y', \phi'] \in X(\gamma, \gamma')$ such that $\gamma^{\circ(y', \phi')} \simeq \gamma^{\circ(y, \phi)}$. Such an isomorphism can be induced by an element $g \in \text{Aut}(\gamma)$, with $[y', \phi'] = g[y, \phi]$ but leaving γ' untouched, leading to a factor of $|\text{Aut}(\gamma)[y, \phi]|$. The number of isomorphisms $\gamma^{\circ(y', \phi')} \simeq \gamma^{\circ(y, \phi)}$ that are not induced by such an element, is given precisely by the factor $N(y, \phi)$. Thus, on inserting the expression for $n(\gamma^{\circ(y, \phi)} \gamma', \gamma', \gamma)$ derived in Lemma 20, we infer that,

$$\begin{aligned} \rho_{\gamma\gamma'}(\Gamma) &= \frac{1}{n_c(\gamma, \gamma')} \sum_{[y, \psi] \in X(\gamma, \gamma')} \frac{M(y, \psi) N(y, \psi)}{N(y, \psi) |\text{Aut}(\gamma)[y, \psi]|} \Delta_{\gamma^{\circ(y, \psi)} \gamma'}(\Gamma) \\ &= \sum_{[y, \psi] \in X(\gamma, \gamma')} \frac{\text{Sym}(\gamma^{\circ(y, \psi)} \gamma')}{\text{Sym}(\gamma\gamma')} \Delta_{\gamma^{\circ(y, \psi)} \gamma'}(\Gamma), \end{aligned}$$

where we have applied Lemma 18 in going to the second line. We have also used Equation (10) to replace $n_c(\gamma, \gamma') \text{Sym}(\gamma) \text{Sym}(\gamma')$ by $\text{Sym}(\gamma\gamma')$. \square

3.4. The coproduct on Green's functions. In this section we derive a key formula for the coproduct on the Green's function defined above. We start with a technical lemma; recall that $\sum_{\Gamma_L^r}$ denotes the sum over all 1PI graphs with residue r and loop number L .

LEMMA 22 ([20]). *If γ is a 1PI graph of loop order K , then the following holds:*

$$(14) \quad \sum_{\Gamma_L^r} \frac{\Gamma \mid \gamma_0}{|\gamma| \vee \text{Sym}(\Gamma)} \Delta_\gamma(\Gamma) = \sum_{\tilde{\Gamma}_{L-K}^r} \frac{\tilde{\Gamma} \mid \gamma \mid \gamma_0}{\text{Sym}(\gamma) \text{Sym}(\tilde{\Gamma})} \tilde{\Gamma}.$$

PROOF. If $\gamma = \emptyset$, there is nothing to prove, since $\Delta_\gamma(\Gamma) = \Gamma$, $\text{Sym}(\emptyset) = 1$ and $\Gamma \mid \emptyset \mid \gamma_0 \equiv \Gamma \mid \gamma_0$. We claim that the following equality holds for $\gamma, \tilde{\Gamma} \neq \emptyset$,

$$\sum_{\Gamma} \frac{(\Gamma \mid \gamma_0) n(\Gamma, \gamma, \tilde{\Gamma})}{|\gamma| \vee \text{Sym}(\Gamma)} = \sum_{[x, \phi] \in X(\tilde{\Gamma}, \gamma)} \frac{(\tilde{\Gamma}^{\circ(x, \phi)} \gamma \mid \gamma_0) n(\tilde{\Gamma}^{\circ(x, \phi)} \gamma, \gamma, \tilde{\Gamma})}{|\gamma| \vee N(x, \phi) |\text{Aut}(\tilde{\Gamma})[x, \phi]| \text{Sym}(\tilde{\Gamma}^{\circ(x, \phi)} \gamma)}$$

Indeed, one can replace the sum on the left-hand-side over Γ by a sum over insertion places of γ in $\tilde{\Gamma}$ (so that $\Gamma \simeq \tilde{\Gamma}^{\circ(x, \phi)} \gamma$ for some $[x, \phi] \in X(\Gamma, \gamma)$, and also $\text{res}(\tilde{\Gamma}) = \text{res}(\Gamma)$), provided one divides by a combinatorial factor counting the number of equivalent insertions. This factor is given as the number of elements $[x', \phi'] \in X(\Gamma, \gamma)$ such that $\tilde{\Gamma}^{\circ(x', \phi')} \gamma \simeq \tilde{\Gamma}^{\circ(x, \phi)} \gamma$, in which case $\text{Sym}(\tilde{\Gamma}^{\circ(x, \phi)} \gamma) = \text{Sym}(\tilde{\Gamma}^{\circ(x', \phi')} \gamma)$ and also $\tilde{\Gamma}^{\circ(x, \phi)} \gamma \mid \gamma_0 = \tilde{\Gamma}^{\circ(x', \phi')} \gamma \mid \gamma_0$.

Such an isomorphism $\tilde{\Gamma}^{\circ(x', \phi')} \gamma \simeq \tilde{\Gamma}^{\circ(x, \phi)} \gamma$ can be induced by an element in $g \in \text{Aut}(\tilde{\Gamma})$ with $[x', \phi'] = g[x, \phi]$ but leaving γ untouched. This leads to division by the length of the orbit $\text{Aut}(\tilde{\Gamma})[x, \phi]$. Otherwise, an isomorphism from $\tilde{\Gamma}^{\circ(x, \phi)} \gamma$ to $\tilde{\Gamma}^{\circ(x', \phi')} \gamma$ has to map γ to an isomorphic subgraph $\gamma' \subset \tilde{\Gamma}$. In that case, it can not be induced by an element in $\text{Aut}(\tilde{\Gamma})$, leading precisely to the additional factor of $N(x, \phi)$.

Equation (14) now follows directly by inserting the expressions obtained in Lemma 18 and 20 in the above equation and summing over all 1PI graphs $\tilde{\Gamma}$, as in Equation (13). We also noted on the way that by definition

$$\frac{1}{|\gamma|_{\vee}} \sum_{[x,\phi] \in X(\tilde{\Gamma}, \gamma)} \tilde{\Gamma} \circ_{(x,\phi)} \gamma \mid \gamma_0 = \tilde{\Gamma} \mid \gamma \mid \gamma_0.$$

The case $\tilde{\Gamma} = \emptyset$ arises whenever $K = L$ and $\gamma \simeq \Gamma$, in which case the combinatorial factors $\Gamma \mid \gamma_0$ and $\emptyset \mid \gamma \mid \gamma_0$ coincide. \square

PROPOSITION 23 ([20]). *The coproduct takes the following form on the 1PI Green's functions:*

$$\Delta(G_L^r) \equiv \sum_{\Gamma_L^r} \frac{1}{\text{Sym}(\Gamma)} \Delta(\Gamma) = \sum_{K=0}^L \sum_{\gamma_K, \tilde{\Gamma}_{L-K}^r} \frac{\tilde{\Gamma} \mid \gamma}{\text{Sym}(\gamma) \text{Sym}(\tilde{\Gamma})} \gamma \otimes \tilde{\Gamma},$$

where the sums are over all 1PI graphs $\Gamma, \tilde{\Gamma}$ with the indicated residue and loop number, and graphs γ at the indicated loop order that are disjoint unions of 1PI graphs.

PROOF. Since $\Delta = \frac{1}{|\gamma|_{\vee}} \sum_{\gamma} \gamma \otimes \Delta_{\gamma}$, this would follow from the validity of Eq. (14) for γ any disjoint union of 1PI graphs at loop order $K < L$ and an auxiliary graph γ_0 . Indeed, putting $\gamma_0 = \emptyset$ and summing over γ then gives the desired result. We show that this analogue of Equation (14) holds by induction on the number of connected components of γ .

Assume that Equation (14) holds for γ a (non-empty) disjoint union of 1PI graphs of loop order K . We will prove that it also holds for the disjoint union $\gamma\gamma' = \gamma \cup \gamma'$ of it with a non-empty 1PI graph γ' of loop order K' . An application of Lemma 21 yields,

$$(15) \quad \begin{aligned} \frac{\Gamma \mid \gamma_0}{|\gamma\gamma'|_{\vee} \text{Sym}(\Gamma)} \Delta_{\gamma\gamma'}(\Gamma) &= \frac{\Gamma \mid \gamma_0}{n_c(\gamma, \gamma') |\gamma\gamma'|_{\vee} \text{Sym}(\Gamma)} \Delta_{\gamma} \Delta_{\gamma'}(\Gamma) \\ &\quad - \frac{\Gamma \mid \gamma_0}{|\gamma\gamma'|_{\vee} \text{Sym}(\Gamma)} \rho_{\gamma\gamma'}(\Gamma). \end{aligned}$$

Since γ' is a 1PI graph, we can apply Lemma 22 to the first term, which gives for the sum over all graphs Γ_L^r ,

$$\begin{aligned} \frac{1}{n_c(\gamma, \gamma')} \sum_{\Gamma_L^r} \frac{\Gamma \mid \gamma_0}{|\gamma\gamma'|_{\vee} \text{Sym}(\Gamma)} \Delta_{\gamma} \Delta_{\gamma'}(\Gamma) \\ &= \frac{1}{n_c(\gamma, \gamma')} \sum_{\Gamma_{L-K'}^r} \frac{\Gamma \mid \gamma' \mid \gamma_0}{|\gamma|_{\vee} \text{Sym}(\gamma') \text{Sym}(\Gamma)} \Delta_{\gamma}(\Gamma) \\ &= \frac{1}{n_c(\gamma, \gamma')} \sum_{\Gamma_{L-K'}^r} \frac{\Gamma \mid \gamma' \gamma_0 + (\Gamma \mid \gamma')(\gamma' \mid \gamma_0)}{\text{Sym}(\gamma') |\gamma|_{\vee} \text{Sym}(\Gamma)} \Delta_{\gamma}(\Gamma), \end{aligned}$$

using also Equation (12). The induction hypothesis – that is, validity of Eq. (14) in the case of γ – now yields,

$$\frac{1}{n_c(\gamma, \gamma')} \sum_{\Gamma_L^r} \frac{\Gamma \mid \gamma_0}{|\gamma\gamma'|_{\vee} \text{Sym}(\Gamma)} \Delta_{\gamma} \Delta_{\gamma'}(\Gamma) = \sum_{\tilde{\Gamma}_{L-K-K'}^r} \frac{\tilde{\Gamma} \mid \gamma \mid \gamma' \gamma_0 + (\tilde{\Gamma} \mid \gamma \mid \gamma')(\gamma' \mid \gamma_0)}{\text{Sym}(\gamma\gamma') \text{Sym}(\tilde{\Gamma})} \tilde{\Gamma}$$

combining once more the symmetry factors $\text{Sym}(\gamma)$ and $\text{Sym}(\gamma')$ with the help of $n_c(\gamma, \gamma')$. For the second term in Equation (15), we can use the induction hypothesis on $\Delta_{\gamma \circ (y, \psi) \gamma'}$ to show that

$$\sum_{\Gamma_L^r} \frac{\Gamma \mid \gamma_0}{|\gamma\gamma'| \vee \text{Sym}(\Gamma)} \rho_{\gamma\gamma'}(\Gamma) = \sum_{\tilde{\Gamma}_{L-K-K'}^r} \sum_{[y, \psi] \in X(\gamma, \gamma')} \frac{\tilde{\Gamma} \mid \gamma \circ (y, \psi) \gamma' \mid \gamma_0}{|\gamma'| \vee \text{Sym}(\gamma\gamma') \text{Sym}(\tilde{\Gamma})} \tilde{\Gamma},$$

since $|\gamma \circ (y, \psi) \gamma'| \vee = |\gamma| \vee$. We conclude with the following equality,

$$\tilde{\Gamma} \mid \gamma\gamma' \mid \gamma_0 = \tilde{\Gamma} \mid \gamma \mid \gamma'\gamma_0 + (\tilde{\Gamma} \mid \gamma \mid \gamma')(\gamma' \mid \gamma_0) - \frac{1}{|\gamma'| \vee} \sum_{[y, \psi] \in X(\gamma, \gamma')} \tilde{\Gamma} \mid \gamma \circ (y, \psi) \gamma' \mid \gamma_0,$$

which follows easily from Definition 17. Indeed, by definition

$$\begin{aligned} & \tilde{\Gamma} \mid \gamma \mid \gamma'\gamma_0 + (\tilde{\Gamma} \mid \gamma \mid \gamma')(\gamma' \mid \gamma_0) \\ &= \frac{1}{|\gamma\gamma'| \vee} \sum_{[x, \phi] \in X(\tilde{\Gamma}, \gamma)} \sum_{[x', \phi'] \in X(\tilde{\Gamma} \circ (x, \phi) \gamma, \gamma')} (\tilde{\Gamma} \circ (x, \phi) \gamma) \circ (x', \phi') \gamma' \mid \gamma_0, \end{aligned}$$

which counts the number of places to insert $\gamma\gamma'$ and then γ_0 in $\tilde{\Gamma}$. Subtraction of the number of such places with γ' sitting inside γ , leads precisely to the number of places to subsequently insert $\gamma\gamma'$ and γ_0 in $\tilde{\Gamma}$. \square

With the explicit expressions (11) for the number of insertions $\tilde{\Gamma} \mid \gamma$ we can further simplify this expression. Recall the grading d_v by the number of vertices, and the projection p_{n_1, \dots, n_k} onto the vector subspace of H consisting of graphs with n_1 vertices of type v_1 , n_2 vertices of type v_2 , *et cetera*.

LEMMA 24 ([19]).

$$\Delta(G^r) = \sum_{\text{res}(\Gamma)=r} \prod_{v \in R_V} (G^v)^{m_{\Gamma, v}} \prod_{e \in R_E} (G^e)^{-m_{\Gamma, e}} \otimes \frac{\Gamma}{\text{Sym}(\Gamma)}.$$

PROOF. Let us simplify a little and consider a scalar field theory with just one type of vertex and edge, i.e. $R = \{ \prec, - \}$. We consider the sum

$$\sum_{\gamma} \frac{\Gamma \mid \gamma}{\text{Sym}(\gamma)} \gamma = \sum_{\gamma_v} \frac{n_{\gamma, v}!}{\text{Sym}(\gamma_v)} \binom{m_{\Gamma, v}}{n_{\gamma, v}} \gamma_v \sum_{\gamma_e} \frac{n_{\gamma, e}!}{\text{Sym}(\gamma_e)} \binom{m_{\Gamma, e} + n_{\gamma, v} - 1}{n_{\gamma, v}} \gamma_e,$$

naturally split into a sum over vertex and edge graphs. We have also inserted the above combinatorial expression for the number of insertion places. Next, we write $\gamma_v = \gamma'_v \gamma''_v$ and try factorize the sum over γ_v into a sum over γ'_v (connected) and γ''_v . Some care should be taken here regarding the combinatorial factors but let us ignore them for the moment. In fact, if we fix the number of connected components $h^0(\gamma_v)$ of γ_v in the sum to be n_V we can write

$$\sum_{h^0(\gamma_v)=n_V} n_V! \frac{\gamma_v}{\text{Sym}(\gamma_v)} = \sum_{h^0(\gamma_v)=n_V} \left(\sum_{\substack{\gamma'_v, \gamma''_v \\ \gamma'_v \gamma''_v \simeq \gamma_v}} \frac{n_c(\gamma''_v, \gamma'_v) + 1}{n_V} \right) n_V! \frac{\gamma_v}{\text{Sym}(\gamma_v)},$$

with γ'_v a connected graph. Here, we have simply inserted 1,

$$\sum_{\substack{\gamma''_v, \gamma'_v \\ \gamma'_v \gamma''_v \simeq \gamma_v}} \frac{n_c(\gamma''_v, \gamma'_v) + 1}{n_V} = \sum_{\gamma'_v} \frac{n_c(\gamma_v, \gamma'_v)}{n_V} = 1,$$

which follows directly from the definition of $n_c(\gamma_v, \gamma'_v)$ as the number of connected components of γ_v isomorphic to γ'_v . Now, by definition $\text{Sym}(\gamma'_v \gamma''_v) = (n_c(\gamma''_v, \gamma'_v) + 1) \text{Sym}(\gamma'_v) \text{Sym}(\gamma''_v)$ for a connected graph γ'_v so that we obtain for the above sum

$$\sum_{\gamma'_v} \frac{\gamma'_v}{\text{Sym}(\gamma'_v)} \sum_{h^0(\gamma''_v) = n_V - 1} (n_V - 1)! \frac{\gamma''_v}{\text{Sym}(\gamma''_v)} = \dots = (G^v - 1)^{n_V},$$

by applying the same argument n_V times. Recall also the definition of the Green's function G^v from Eq. (8). A similar argument applies to the edge graphs, leading to a contribution $(1 - G^e)^{n_E}$, with n_E the number of connected components of γ_e . When summing over n_V and n_E , taking also into account the combinatorial factors, we obtain:

$$\sum_{n_V=0}^{\infty} \binom{m_{\Gamma,v}}{n_V} (G^v - 1)^{n_V} \sum_{n_E=0}^{\infty} \binom{m_{\Gamma,e} + n_E - 1}{n_E} (1 - G^e)^{n_E} = (G^v)^{m_{\Gamma,v}} (G^e)^{-m_{\Gamma,e}}$$

The extension to the general setting where the set R contains different types of vertices and edges is straightforward. \square

THEOREM 25 ([19]). *If we label our vertices as $R_V = \{v_1, \dots, v_k\}$, then*

$$\Delta(G^r) = \sum_{n_1, \dots, n_k \in \mathbb{Z}} G^r \prod_{i=1}^k \left(\frac{G^{v_i}}{\prod_e (G^e)^{N_e(v_i)/2}} \right)^{n_i} \otimes p_{n_1, \dots, n_k}(G^r).$$

PROOF. An additional counting of the number of edges and numbers of vertices in Γ gives the following relations:

$$2m_{\Gamma,e} + N_e(\text{res}(\Gamma)) = \sum_{v \in R_V} N_e(v) m_{\Gamma,v}$$

where $N_e(r)$ is the number of lines (of type e) attached to $r \in R$. For instance $N_e(\text{electron})$ equals 2 if e is an electron line and 1 if e is a photon line. One checks the above equality by noting that the left-hand-side counts the number of internal half lines plus the external lines which are connected to the vertices that appear at the right-hand-side, taken into account their valence.

With this formula, we can write the expression of Lemma 24 as

$$(16) \quad \Delta(G^r) = \prod_e (G^e)^{N_e(r)/2} \sum_{\text{res}(\Gamma)=r} \prod_v \left(\frac{G^v}{\prod_e (G^e)^{N_e(v)/2}} \right)^{m_{\Gamma,v}} \otimes \frac{\Gamma}{\text{Sym}(\Gamma)}.$$

It now remains to observe that $m_{\Gamma,v} = d_v(\Gamma)$ unless $v = r$ (the residue of Γ) in which case $m_{\Gamma,v} = d_v(\Gamma) + 1$. This yields the extra factor G^r . \square

4. Examples of the Hopf algebra on Green's functions

4.1. Quantum electrodynamics. Let us now apply the above formula to the case of quantum electrodynamics (QED). In (massless) quantum electrodynamics, there is only the vertex of valence three, describing the interaction of the photon with a pair of electrons. There are two types of edges corresponding to the photon (wiggly edge) and the electron (straight edge). Summarizing, we have in the notation of the previous section: $R = R_V \cup R_E$ with

$$R_V = \{ \text{~~~} \};$$

$$R_E = \{ \text{---} , \text{~~~} \}.$$

In particular, this means that in the process of renormalization, only three types of graphs are of importance: the vertex graph, the electron self-energy graph and the vacuum polarization. Correspondingly, we have the three 1PI Green's functions,

$$G^{\text{~~~}} = 1 + \sum_{\Gamma \text{~~~}} \frac{\Gamma}{\text{Sym}(\Gamma)};$$

$$G^e = 1 - \sum_{\Gamma^e} \frac{\Gamma}{\text{Sym}(\Gamma)},$$

with $e = \text{---}, \text{~~~}$.

Since there is only one vertex in QED, we can use Lemma 8 to simplify Theorem 25 above.

PROPOSITION 26 ([18]). *For $r = \text{~~~}, \text{---}$ or ~~~ the following holds*

$$\Delta(G^r) = \sum_{l=0}^{\infty} G^r \left(\frac{G^{\text{~~~}}}{G^e \sqrt{G^{\text{~~~}}}} \right)^{2l} \otimes q_l(G^r)$$

with q_l the projection onto graphs of loop order l . \square

COROLLARY 27. *The elements $q_l(G^{\text{~~~}}) - q_l(G^{\text{---}}) \in H$ for $l = 1, 2, \dots$ generate a Hopf ideal I , i.e.*

$$\Delta(I) \subseteq I \otimes H + H \otimes I, \quad \epsilon(I) = 0, \quad S(I) \subseteq I.$$

PROOF. This follows easily by applying Proposition 26 to the coproduct evaluated on the difference $G^{\text{~~~}} - G^{\text{---}}$, in combination with the recursive definition of the antipode. \square

The identities $G^{\text{~~~}} = G^{\text{---}}$ which hold in the corresponding quotient Hopf algebra H/I have a physical meaning: they are the famous *Ward identities of quantum electrodynamics* [23]. The above claim that they can be implemented on the Hopf algebra of Feynman graphs corresponds to the physical statement that the Ward identities are compatible with renormalization. In fact, we have the following.

PROPOSITION 28. *Suppose the regularized (but unrenormalized) Feynman rules $U : H \rightarrow K$ satisfy the Ward identities. Then the counterterms C and the renormalized Feynman rules R satisfy the Ward identities:*

$$C(G^{\text{~~~}}) = C(G^{\text{---}}); \quad R(G^{\text{~~~}}) = R(G^{\text{---}})$$

Note that the first equation is usually written as $Z_1 = Z_2$ [23].

PROOF. This follows directly from the Birkhoff decomposition (cf. Theorem 9 above) applied to the character group of the graded connected commutative Hopf algebra H/I . \square

4.2. Quantum chromodynamics. We work in the setting of the non-abelian gauge theory *quantum chromodynamics* (QCD). It describes the interaction between quarks (the fermions) via gluons (the gauge bosons).

In contrast with quantum electrodynamics described previously, there are now three vertices of valence three, describing the interaction of the fermion and ghost with the gluon, as well as the cubic gluon self-interaction. In addition, there is the quartic gluon self-interaction. This means that the Feynman graphs are built from the following two sets of vertices and edges:

$$R_V = \{ \text{---} \swarrow \text{---}, \text{---} \cdot \text{---}, \text{---} \text{---} \text{---}, \text{---} \text{---} \text{---} \text{---} \};$$

$$R_E = \{ \text{---}, \dots, \text{---} \text{---} \text{---} \},$$

where the plain, dotted and curly lines represent the quark, ghost and gluon, respectively.

The relation between the number of vertices and loop order is not as simple as in QED, due to the presence of 4 different vertices. Nevertheless, we can explore the structure of the Hopf algebra by introducing the following formal elements corresponding to each vertex $v \in R_V$:

$$X_v := \left(\frac{G^v}{\prod_e (G^e)^{N_e(v)/2}} \right)^{1/N(v)-2}.$$

Theorem 25 in combination with the fact that Δ is an algebra map then yields the following result.

PROPOSITION 29. *The coproduct on X_v ($v \in R_V$) is given by*

$$\Delta(X_v) = \sum_{n_1, \dots, n_4} X_v (X \text{---} \text{---})^{n_1} (X \text{---} \text{---})^{n_2} (X \text{---} \text{---})^{n_3} (X \text{---} \text{---})^{2n_4} \otimes p_{n_1, \dots, n_4}(X_v).$$

COROLLARY 30. *The elements $q_l(X_v) - q_l(X_{v'})$ for any $v, v' \in R_V$ and $l = 1, 2, \dots$ generate a Hopf ideal I in H .*

PROOF. Let I be the algebra ideal in H generated by $q_l(X_v) - q_l(X_{v'})$. In the expression of Proposition (29) for the coproduct on X_v , we can replace each X_v that appears on the first leg of the tensor product by $X_{\tilde{v}}$ for some fixed but arbitrary $\tilde{v} \in R_V$, as long as we add terms in $I \otimes H$ to it. Thus, for all $v \in R_V$ we have

$$\begin{aligned} \Delta(X_v) &= \sum_{n_1, \dots, n_4} (X_{\tilde{v}})^{n_1 + \dots + 2n_4 + 1} \otimes p_{n_1, \dots, n_4}(X_v) + I \otimes H \\ &= \sum_l (X_{\tilde{v}})^{2l+1} \otimes q_l(X_v) + I \otimes H, \end{aligned}$$

where in going to the second line, we have applied Lemma 8 to write $n_1 + \dots + 2n_4 = 2l$. With the first leg of $\Delta(X_v)$ independent of v , we easily obtain that

$$\Delta(X_v) - \Delta(X_{v'}) = \sum_l (X_{\tilde{v}})^{2l+1} \otimes q_l(X_v - X_{v'}) + I \otimes H$$

thus establishing that $\Delta(I) \subset I \otimes H + H \otimes I$. The recursive definition of the antipode allows one to conclude directly that $S(I) \subset I$. \square

Again these identities have a physical meaning, they are the *Slavnov–Taylor identities for the couplings* in quantum chromodynamics. Again, the above fact that they can be imposed on the Hopf algebra is a rephrasing of the physical fact that these identities are compatible with renormalization. More precisely, we have the following

PROPOSITION 31. *Suppose the regularized (but unrenormalized) Feynman rules $U : H \rightarrow K$ satisfy the Slavnov–Taylor identities for the couplings. Then the counterterms C and the renormalized Feynman rules R satisfy the Slavnov–Taylor identities:*

$$\begin{aligned} C(X_{\prec}) &= C(X_{\preccurlyeq}) = C(\prec) = C(\sqrt{\mathbb{X}}) \\ R(X_{\prec}) &= R(X_{\preccurlyeq}) = R(\prec) = R(\sqrt{\mathbb{X}}) \end{aligned}$$

Again, the first equation is typically written in terms of Z -factors. This would lead precisely to Equation (6) above.

PROOF. As in the case of QED, this follows from the Birkhoff decomposition in the character group of the quotient Hopf algebra H/I . \square

Let us end this section with a formula for the coproduct on the element $X := X_v^{1/N(v)-2}$ in H/I :

$$(17) \quad \Delta(X) = \sum_{l=0}^{\infty} X^{2l+1} \otimes q_l(X).$$

5. Dyson–Schwinger equations and Hochschild cohomology

In this section, we will review how Hochschild cohomology fits nicely in the context of renormalization Hopf algebras, following [7, 2] and [14]. In particular, we will relate it to the Dyson–Schwinger equations and prove the so-called gauge theory theorem that was announced in [14].

Let us first recall the definition of Hochschild cohomology for Hopf algebras, – or, more generally, for bialgebras – with values in a bicomodule. This dualizes the definition of Hochschild cohomology for algebras to bialgebras. Let H be a bialgebra and M an H -bicomodule, i.e. there are two cocommuting left and right coactions $\rho_L : M \rightarrow H \otimes M$ and $\rho_R : M \rightarrow M \otimes H$. We denote by $C^n(H, M)$ the space of linear maps $\phi : M \rightarrow H^{\otimes n}$ and define the *Hochschild coboundary map* $b : C^n(H, M) \rightarrow C^{n+1}(H, M)$ by

$$b\phi = (\text{id} \otimes \phi)\rho_L + \sum_{i=1}^n (-1)^i \Delta_i \phi + (-1)^{n+1} (\phi \otimes \text{id})\rho_R.$$

where Δ_i denotes the application of the coproduct on the i 'th factor in $H^{\otimes n}$. Coassociativity implies that b is a differential, i.e. that $b^2 = 0$.

DEFINITION 32. *The Hochschild cohomology $HH^\bullet(H, M)$ of the bialgebra H with values in the H -comodule M is defined as the cohomology of the complex $(C^\bullet(H, M), b)$ defined above.*

We are interested in the particular case that $M = H$ is a comodule over itself, with $\rho_L = \Delta$ but with $\rho_R = (\text{id} \otimes \epsilon)\Delta$. We denote the Hochschild cohomology groups in this case by $HH^\bullet(H, H_\epsilon)$ or simply $HH_\epsilon^\bullet(H)$ as in [7]. Let us consider the case $n = 1$, then $\phi \in HH_\epsilon^1(H)$ means simply that

$$\Delta\phi = (\text{id} \otimes \phi)\Delta + (\phi \otimes \mathbb{I}).$$

where $(\phi \otimes \mathbb{I})(h) \equiv \phi(h) \otimes 1$ for $h \in H$. As was observed in [7] the grafting operator on rooted trees is an example of such a 1-cocycle. We will give an example in the case of the Hopf algebra of Feynman graphs (cf. (2) of Theorem 33 below), following [14].

This starts with the observation that the Green's functions can be dissected as follows [14, Theorem 4]:

$$(18) \quad G^r = \sum_{\gamma \text{ prim}} B_+^\gamma \left(\frac{\prod_{v \in R_V} (G^v)^{m_{\gamma,v}}}{\prod_{e \in R_E} (G^e)^{m_{\gamma,e}}} \right) = \sum_{\gamma \text{ prim}} B_+^\gamma (G^{\text{res}(\gamma)} X^{2l(\gamma)})$$

where B_+^γ is the (normalized) grafting operator that inserts in γ the graphs given as its argument on the appropriate insertion places. The sum is over all primitive graphs γ , i.e. satisfying $\Delta(\gamma) = \gamma \otimes 1 + 1 \otimes \gamma$. It is clear that any graph in G^r is of the form $B_+^\gamma(\Gamma_1 \cdots \Gamma_N)$ for some 1PI graphs $\Gamma_1, \dots, \Gamma_N$ but this decomposition is highly non-unique. In order to correct for the overcounting, the grafting operators have to be normalized appropriately as was done in [14]. We will instead take Eq. (18) as a *definition* of the normalized maps B_+^γ , without explicitly describing this normalization. The sum of the B_+^γ over all primitive 1PI Feynman graphs at a given loop order and with given residue will be denoted by $B_+^{k;r}$, as in *loc. cit.*. More precisely,

$$B_+^{k;r} = \sum_{\substack{\gamma \text{ prim} \\ l(\gamma)=k \\ \text{res}(\gamma)=r}} \frac{1}{\text{Sym}(\gamma)} B_+^\gamma$$

and, of course, $G^r = \sum_{k,r} B_+^{k;r}(X_{k,r})$, where we have denoted $X_{k,r} = G^r X^{2k}$. With this and the formulas of the previous section on QCD, we can prove the *gauge theory theorem* as formulated in [14, Theorem 5]:

THEOREM 33. *Let $\tilde{H} = H/I$ be the Hopf algebra of QCD Feynman graphs (cf. Sect. 4.2) with the Slavnov–Taylor identities for the couplings imposed.*

- (1) $G^r = \sum_{k=0}^{\infty} B_+^{k;r}(X_{k,r})$
- (2) $\Delta(B_+^{k;r}(X_{k,r})) = B_+^{k;r}(X_{k,r}) \otimes \mathbb{I} + (\text{id} \otimes B_+^{k;r})\Delta(X_{k,r})$.
- (3) $\Delta(G^r) = \sum_{j=0}^k \text{Pol}_j^r(G) \otimes G^r_{k-j}$.

where $\text{Pol}_j^r(G)$ is a polynomial in the G_m^r of degree j , determined as the order j term in the loop expansion of $G^r X^{2k-2j}$.

PROOF. The first claim is just the definition of the $B_+^{k;r}$. For (2), we first enhance the result of Proposition 29 to partial sums in G^r over graphs that have ‘primitive residue’ isomorphic to a fixed primitive graph γ . In other words, if $G^{r,\gamma}$ is the part of G^r that sums only over graphs that are obtained by inserting graphs into the primitive graph γ , then

$$\Delta(G^{r,\gamma}) = G^{r,\gamma} \otimes 1 + \sum_{l=1}^{\infty} G^r X^{2l} \otimes q_l(G^{r,\gamma}).$$

Here we have imposed the Slavnov–Taylor identities for the couplings to write this in terms of a single coupling, X . Combing Proposition 29 and Eq. (17) we obtain for the coproduct of $X_{k,r} = G^r X^{2k}$:

$$\Delta(X_{k,r}) = \sum_{l=0}^{\infty} G^r X^{2l+2k} \otimes q_l X_{k,r}.$$

Since $G^{r,\gamma} = B_+^\gamma(X_{k,r})$, it follows by a combination of the above two formula that

$$\Delta(B_+^\gamma(X_{k,r})) = B_+^\gamma(X_{k,r}) \otimes \mathbb{I} + (\text{id} \otimes B_+^\gamma)\Delta(X_{k,r}).$$

and summing over all primitive graphs with residue r at loop order k gives the desired result.

Finally, (2) follows by combining Theorem 25 with Proposition 29, thereby taking into account the Slavnov–Taylor identities. \square

REMARK 34. *We have corrected for the apparent misprint in [14, Eq. (83)].*

In fact, this proves the slightly stronger result that actually every B_+^γ defines a Hochschild 1-cocycle:

PROPOSITION 35. *For γ a primitive graph at loop order k and residue r , we have*

$$\Delta(B_+^\gamma(X_{k,r})) = B_+^\gamma(X_{k,r}) \otimes \mathbb{I} + (\text{id} \otimes B_+^\gamma)\Delta(X_{k,r}).$$

Actually, the above results apply in full generality for any Hopf algebra as defined in Definition 7. However, the meaning of the Hopf ideals as imposing Slavnov–Taylor identities can only be given in the context of a non-abelian gauge theory. Moreover, the above Hochschild cocycles $B_+^{k,r}$ play an important role in that they give quantum equation of motions. These *Dyson–Schwinger equations* are the recursive construction of the 1PI Green’s functions G^r from the lower order Green’s functions in $X_{k,r}$, forming the argument of the B_+ -operations. In fact, Equation (18) are precisely the Dyson–Schwinger equations for quantum chromodynamics.

Acknowledgements. The author would like to thank the Hausdorff Research Institute for Mathematics in Bonn for their hospitality during the summer of 2008. Also, the anonymous referee is thanked for valuable comments.

References

- [1] G. van Baalen, D. Kreimer, D. Uminsky, and K. Yeats. The QED beta-function from global solutions to Dyson–Schwinger equations, arXiv:0805.0826.
- [2] C. Bergbauer and D. Kreimer. Hopf algebras in renormalization theory: Locality and Dyson–Schwinger equations from Hochschild cohomology. *IRMA Lect. Math. Theor. Phys.* 10 (2006) 133–164.
- [3] J. Collins. *Renormalization*. Cambridge University Press, 1984.
- [4] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann–Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem. *Comm. Math. Phys.* 210 (2000) 249–273.
- [5] A. Connes and D. Kreimer. Renormalization in quantum field theory and the Riemann–Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group. *Commun. Math. Phys.* 216 (2001) 215–241.
- [6] A. Connes and M. Marcolli. *Noncommutative Geometry, Quantum Fields and Motives*. AMS, Providence, 2008.
- [7] A. Connes and D. Kreimer. Hopf algebras, renormalization and noncommutative geometry. *Commun. Math. Phys.* 199 (1998) 203–242.

- [8] R. Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.
- [9] K. Ebrahimi-Fard and L. Guo. Rota-Baxter algebras in renormalization of perturbative quantum field theory. *Fields Inst. Commun.* 50 (2007) 47–105.
- [10] G. 't Hooft and M. J. G. Veltman. Diagrammar. *CERN yellow report*. 73 (1973) 1–114.
- [11] T. Krajewski and P. Martinetti. Wilsonian renormalization, differential equations and Hopf algebras. arXiv:0806.4309.
- [12] D. Kreimer. On the Hopf algebra structure of perturbative quantum field theories. *Adv. Theor. Math. Phys.* 2 (1998) 303–334.
- [13] D. Kreimer. Dyson-Schwinger equations: From Hopf algebras to number theory. hep-th/0609004.
- [14] D. Kreimer. Anatomy of a gauge theory. *Ann. Phys.* 321 (2006) 2757–2781.
- [15] D. Kreimer and K. Yeats. An étude in non-linear Dyson-Schwinger equations. *Nuclear Phys. B Proc. Suppl.* 160 (2006) 116–121.
- [16] D. V. Prokhorchenko. Renormalization of gauge theories and the Hopf algebra of diagrams. arXiv:0705.3906 [hep-th].
- [17] A. A. Slavnov. Ward identities in gauge theories. *Theor. Math. Phys.* 10 (1972) 99–107.
- [18] W. D. van Suijlekom. Multiplicative renormalization and Hopf algebras. To appear in *Arithmetic and geometry around quantization*. Eds. Ö. Ceyhan, Yu.-I. Manin and M. Marcolli. Birkhäuser Verlag, Basel 2008. [arXiv:0707.0555].
- [19] W. D. van Suijlekom. Renormalization of gauge fields using Hopf algebras. To appear in *Recent Developments in Quantum Field Theory*. Eds. B. Fauser, J. Tolksdorf and E. Zeidler. Birkhäuser Verlag, Basel 2008. [arXiv:0801.3170].
- [20] W. D. van Suijlekom. Renormalization of gauge fields: A Hopf algebra approach. *Commun. Math. Phys.* 276 (2007) 773–798.
- [21] W. D. van Suijlekom. Representing Feynman graphs on BV-algebras, arXiv:0807.0999
- [22] J. C. Taylor. Ward identities and charge renormalization of the Yang-Mills field. *Nucl. Phys.* B33 (1971) 436–444.
- [23] J. C. Ward. An identity in quantum electrodynamics. *Phys. Rev.* 78 (1950) 182.

INSTITUTE FOR MATHEMATICS, ASTROPHYSICS AND PARTICLE PHYSICS, FACULTY OF SCIENCE,
RADBOUD UNIVERSITY NIJMEGEN, HEYENDAALSEWEG 135, 6525 ED NIJMEGEN, THE NETHERLANDS

E-mail address: `waltervs@math.ru.nl`

Multi-scale Analysis and Non-commutative Field Theory

Fabien Vignes-Tourneret

ABSTRACT. We review the renormalisation group properties of non-commutative quantum field theory.

1. Introduction

General relativity and ordinary differential geometry should be replaced by non-commutative geometry at some point between the currently accessible energies of about 1 - 10 Tev (after starting the Large Hadron Collider (LHC) at CERN) and the Planck scale, which is 10^{15} times higher, where space-time and gravity should be quantized.

This could occur either at the Planck scale or below. Quantum field theory on a non-commutative space-time (NCQFT) could very well be an intermediate theory relevant for physics at energies between the LHC and the Planck scale. It certainly looks intermediate in structure between ordinary quantum field theory on commutative \mathbb{R}^4 and string theory, the current leading candidate for a more fundamental theory including quantized gravity. NCQFT in fact appears as an effective model for certain limits of string theory [1, 2].

In joint work with R. Gurau, J. Magnen and V. Rivasseau [3], using direct space methods, we provided recently a new proof that the Grosse-Wulkenhaar scalar Φ_4^4 theory on the Moyal space \mathbb{R}^4 is renormalisable to all orders in perturbation theory.

The Grosse-Wulkenhaar breakthrough [4, 5] was to realize that the right propagator in non-commutative field theory is not the ordinary commutative propagator, but has to be modified. Grosse and Wulkenhaar were able to compute the corresponding propagator in the so called “matrix base” which transforms the Moyal product into a matrix product. This is a real *tour de force*! They use this representation to prove perturbative renormalisability of the theory up to some estimates which were finally proven in [6].

Our direct space method builds upon the previous works of Filk and Chepelev-Roiban [7, 8]. These works however remained inconclusive [9], since these authors used the right interaction but not the right propagator, hence the problem of ultraviolet/infrared mixing prevented them from obtaining a finite renormalised perturbation series.

We also extend the Grosse-Wulkenhaar results to more general models with covariant derivatives in a fixed magnetic field [10]. Our proof relies on a multiscale analysis analogous to [6] but in direct space.

Non-commutative field theories (for a general review see [11]) deserve a thorough and systematic investigation, not only because they may be relevant for physics beyond the standard model, but also (although this is often less emphasized) because they can describe effective physics in our ordinary standard world but with non-local interactions.

We propose to base physics upon the renormalisability principle, more than any other axiom. Renormalisability means genericity; only renormalisable interactions survive a few RG steps, hence only them should be used to describe generic effective physics of any kind. The search for renormalisability could be the powerful principle on which to orient ourselves in the jungle of all possible non-local interactions.

Renormalisability has also attracted considerable interest in the recent years as a pure mathematical structure. The work of Kreimer and Connes [12, 13, 14] recasts the recursive BPHZ forest formula of perturbative renormalisation in a nice Hopf algebra structure. The renormalisation group ambiguity reminds mathematicians of the Galois group ambiguity for roots of algebraic equations. Finding new renormalisable theories may therefore be important for the future of pure mathematics as well as for physics. Main open conjectures in pure mathematics such as the Riemann hypothesis [15, 16] or the Jacobian conjecture [17] may benefit from the quantum field theory and renormalisation group approach.

A critical goal to enlarge the class of renormalisable non-commutative field theories and to attack the Quantum Hall effect problem is to extend the results of Grosse-Wulkenhaar to Fermionic theories. The simplest theory, the two-dimensional Gross-Neveu model can be shown renormalisable to all orders in their Langmann-Szabo covariant versions, using either the matrix basis [18] or the direct space version developed here [19]. However the x -space version seems the most promising for a complete non-perturbative construction, using Pauli's principle to control the apparent (fake) divergences of perturbation theory.

Finally let us conclude this short introduction by reminding that a very important and difficult goal is to also extend the Grosse-Wulkenhaar breakthrough to gauge theories.

2. Field theory on Moyal space

The recent progresses concerning the renormalisation of non-commutative field theory have been obtained on a very simple non-commutative space namely the Moyal space. From the point of view of quantum field theory, it is certainly the most studied space. Let us start with its precise definition.

2.1. The Moyal space \mathbb{R}_θ^D . Let us define $E = \{x^\mu, \mu \in [1, D]\}$ and $\mathbb{C}\langle E \rangle$ the free algebra generated by E . Let Θ a $D \times D$ non-degenerate skew-symmetric matrix (which requires D even) and I the ideal of $\mathbb{C}\langle E \rangle$ generated by the elements $x^\mu x^\nu - x^\nu x^\mu - i\Theta^{\mu\nu}$. The Moyal algebra \mathcal{A}_Θ is the quotient $\mathbb{C}\langle E \rangle/I$. Each element in \mathcal{A}_Θ is a formal power series in the x^μ 's for which the relation $[x^\mu, x^\nu] = i\Theta^{\mu\nu}$ holds.

Usually, one puts the matrix Θ into its canonical form :

$$(2.1) \quad \Theta = \begin{pmatrix} 0 & \theta_1 & & (0) & \\ -\theta_1 & 0 & & & \\ & & \ddots & & \\ (0) & & & 0 & \theta_{D/2} \\ & & & -\theta_{D/2} & 0 \end{pmatrix}.$$

Sometimes one even set $\theta = \theta_1 = \dots = \theta_{D/2}$. The preceding algebraic definition whereas short and precise may be too abstract to perform real computations. One then needs a more analytical definition. A representation of the algebra \mathcal{A}_Θ is given by some set of functions on \mathbb{R}^d equipped with a non-commutative product: the *Groenwald-Moyal* product. What follows is based on [20].

The Algebra \mathcal{A}_Θ . The Moyal algebra \mathcal{A}_Θ is the linear space of smooth and rapidly decreasing functions $\mathcal{S}(\mathbb{R}^D)$ equipped with the non-commutative product defined by: $\forall f, g \in \mathcal{S}_D := \mathcal{S}(\mathbb{R}^D)$,

$$(2.2) \quad (f \star_\Theta g)(x) = \int_{\mathbb{R}^D} \frac{d^D k}{(2\pi)^D} d^D y f(x + \frac{1}{2}\Theta \cdot k) g(x + y) e^{ik \cdot y}$$

$$(2.3) \quad = \frac{1}{\pi^D |\det \Theta|} \int_{\mathbb{R}^D} d^D y d^D z f(x + y) g(x + z) e^{-2iy\Theta^{-1}z}.$$

This algebra may be considered as the “functions on the Moyal space \mathbb{R}_θ^D ”. In the following we will write $f \star g$ instead of $f \star_\Theta g$ and use : $\forall f, g \in \mathcal{S}_D$,

$$(2.4) \quad (\mathcal{F}f)(x) = \int f(t) e^{-itx} dt$$

for the Fourier transform and

$$(2.5) \quad (f \diamond g)(x) = \int f(x - t) g(t) e^{2ix\Theta^{-1}t} dt$$

for the twisted convolution. As on \mathbb{R}^D , the Fourier transform exchange product and convolution:

$$(2.6) \quad \mathcal{F}(f \star g) = \mathcal{F}(f) \diamond \mathcal{F}(g)$$

$$(2.7) \quad \mathcal{F}(f \diamond g) = \mathcal{F}(f) \star \mathcal{F}(g).$$

One also shows that the Moyal product and the twisted convolution are **associative**:

$$(2.8) \quad ((f \diamond g) \diamond h)(x) = \int f(x - t - s) g(s) h(t) e^{2i(x\Theta^{-1}t + (x-t)\Theta^{-1}s)} ds dt \\ = \int f(u - v) g(v - t) h(t) e^{2i(x\Theta^{-1}v - t\Theta^{-1}v)} dt dv$$

$$(2.9) \quad = (f \diamond (g \diamond h))(x).$$

Using (2.7), we show the associativity of the \star -produit. The complex conjugation is **involutive** in \mathcal{A}_Θ

$$(2.10) \quad \overline{f \star_\Theta g} = \overline{g} \star_\Theta \overline{f}.$$

One also have

$$(2.11) \quad f \star_\Theta g = g \star_{-\Theta} f.$$

Proposition 2.1 (Trace). *For all $f, g \in \mathcal{S}_D$,*

$$(2.12) \quad \int dx (f \star g)(x) = \int dx f(x)g(x) = \int dx (g \star f)(x).$$

PROOF.

$$(2.13) \quad \begin{aligned} \int dx (f \star g)(x) &= \mathcal{F}(f \star g)(0) = (\mathcal{F}f \diamond \mathcal{F}g)(0) \\ &= \int \mathcal{F}f(-t)\mathcal{F}g(t)dt = (\mathcal{F}f * \mathcal{F}g)(0) = \mathcal{F}(fg)(0) \\ &= \int f(x)g(x)dx \end{aligned}$$

where $*$ is the ordinary convolution. \square

In the following sections, we will need lemma 2.2 to compute the interaction terms for the Φ_4^4 and Gross-Neveu models. We write $x \wedge y := 2x\Theta^{-1}y$.

Lemma 2.2. *For all $j \in [1, 2n+1]$, let $f_j \in \mathcal{A}_\Theta$. Then*

$$(2.14) \quad (f_1 \star_\Theta \cdots \star_\Theta f_{2n})(x) = \frac{1}{\pi^{2D} \det^2 \Theta} \int \prod_{j=1}^{2n} dx_j f_j(x_j) e^{-\imath x \wedge \sum_{i=1}^{2n} (-1)^{i+1} x_i} e^{-\imath \varphi_{2n}},$$

$$(2.15) \quad (f_1 \star_\Theta \cdots \star_\Theta f_{2n+1})(x) = \frac{1}{\pi^D \det \Theta} \int \prod_{j=1}^{2n+1} dx_j f_j(x_j) \delta\left(x - \sum_{i=1}^{2n+1} (-1)^{i+1} x_i\right) e^{-\imath \varphi_{2n+1}},$$

$$(2.16) \quad \forall p \in \mathbb{N}, \varphi_p = \sum_{i < j=1}^p (-1)^{i+j+1} x_i \wedge x_j.$$

Corollary 2.3. *For all $j \in [1, 2n+1]$, let $f_j \in \mathcal{A}_\Theta$. Then*

$$(2.17) \quad \int dx (f_1 \star_\Theta \cdots \star_\Theta f_{2n})(x) = \frac{1}{\pi^D \det \Theta} \int \prod_{j=1}^{2n} dx_j f_j(x_j) \delta\left(\sum_{i=1}^{2n} (-1)^{i+1} x_i\right) e^{-\imath \varphi_{2n}},$$

$$(2.18) \quad \int dx (f_1 \star_\Theta \cdots \star_\Theta f_{2n+1})(x) = \frac{1}{\pi^D \det \Theta} \int \prod_{j=1}^{2n+1} dx_j f_j(x_j) e^{-\imath \varphi_{2n+1}},$$

$$(2.19) \quad \forall p \in \mathbb{N}, \varphi_p = \sum_{i < j=1}^p (-1)^{i+j+1} x_i \wedge x_j.$$

The cyclicity of the product, inherited from proposition 2.1 implies: $\forall f, g, h \in \mathcal{S}_D$,

$$(2.20) \quad \langle f \star g, h \rangle = \langle f, g \star h \rangle = \langle g, h \star f \rangle$$

and allows to extend the Moyal algebra by duality into an algebra of tempered distributions.

Extension by Duality. Let us first consider the product of a tempered distribution with a Schwartz-class function. Let $T \in \mathcal{S}'_D$ and $h \in \mathcal{S}_D$. We define $\langle T, h \rangle := T(h)$ and $\langle T^*, h \rangle = \overline{\langle T, \bar{h} \rangle}$.

Definition 2.1. Let $T \in \mathcal{S}'_D$, $f, h \in \mathcal{S}_D$, we define $T \star f$ and $f \star T$ by

$$(2.21) \quad \langle T \star f, h \rangle = \langle T, f \star h \rangle,$$

$$(2.22) \quad \langle f \star T, h \rangle = \langle T, h \star f \rangle.$$

For example, the identity $\mathbb{1}$ as an element of \mathcal{S}'_D is the unity for the \star -product: $\forall f, h \in \mathcal{S}_D$,

$$(2.23) \quad \begin{aligned} \langle \mathbb{1} \star f, h \rangle &= \langle \mathbb{1}, f \star h \rangle \\ &= \int (f \star h)(x) dx = \int f(x) h(x) dx \\ &= \langle f, h \rangle. \end{aligned}$$

We are now ready to define the linear space \mathcal{M} as the intersection of two sub-spaces \mathcal{M}_L and \mathcal{M}_R of \mathcal{S}'_D .

Definition 2.2 (Multipliers algebra).

$$(2.24) \quad \mathcal{M}_L = \{S \in \mathcal{S}'_D : \forall f \in \mathcal{S}_D, S \star f \in \mathcal{S}_D\},$$

$$(2.25) \quad \mathcal{M}_R = \{R \in \mathcal{S}'_D : \forall f \in \mathcal{S}_D, f \star R \in \mathcal{S}_D\},$$

$$(2.26) \quad \mathcal{M} = \mathcal{M}_L \cap \mathcal{M}_R.$$

One can show that \mathcal{M} is an associative \star -algebra. It contains, among others, the identity, the polynomials, the δ distribution and its derivatives. Then the relation

$$(2.27) \quad [x^\mu, x^\nu] = \imath \Theta^{\mu\nu},$$

often given as a definition of the Moyal space, holds in \mathcal{M} (but not in \mathcal{A}_Θ).

2.2. The ϕ^4 -theory on \mathbb{R}_θ^4 . The simplest non-commutative model one may consider is the ϕ^4 -theory on the four-dimensional Moyal space. Its Lagrangian is the usual (commutative) one where the pointwise product is replaced by the Moyal one:

$$(2.28) \quad S[\phi] = \int d^4x \left(-\frac{1}{2} \partial_\mu \phi \star \partial^\mu \phi + \frac{1}{2} m^2 \phi \star \phi + \frac{\lambda}{4} \phi \star \phi \star \phi \star \phi \right)(x).$$

Thanks to the formula (2.3), this action can be explicitly computed. The interaction part is given by the corollary 2.3:

$$(2.29) \quad \begin{aligned} \int dx \phi^{\star 4}(x) &= \int \prod_{i=1}^4 dx_i \phi(x_i) \delta(x_1 - x_2 + x_3 - x_4) e^{\imath \varphi}, \\ \varphi &= \sum_{i < j=1}^4 (-1)^{i+j+1} x_i \wedge x_j. \end{aligned}$$

The main characteristic of the Moyal product is its non-locality. But its non-commutativity implies that the vertex of the model (2.28) is only invariant under cyclic permutation of the fields. This restricted invariance incites to represent the associated Feynman graphs with ribbon graphs. One can then make a clear distinction between planar and non-planar graphs. This will be detailed in section 4.1.

Thanks to the delta function in (2.29), the oscillation may be written in different ways:

(2.30a)

$$\delta(x_1 - x_2 + x_3 - x_4)e^{i\varphi} = \delta(x_1 - x_2 + x_3 - x_4)e^{ix_1 \wedge x_2 + ix_3 \wedge x_4}$$

(2.30b)

$$= \delta(x_1 - x_2 + x_3 - x_4)e^{ix_4 \wedge x_1 + ix_2 \wedge x_3}$$

(2.30c)

$$= \delta(x_1 - x_2 + x_3 - x_4) \exp i(x_1 - x_2) \wedge (x_2 - x_3).$$

The interaction is real and positive^a:

$$(2.31) \quad \begin{aligned} & \int \prod_{i=1}^4 dx_i \phi(x_i) \delta(x_1 - x_2 + x_3 - x_4) e^{i\varphi} \\ &= \int dk \left(\int dx dy \phi(x) \phi(y) e^{ik(x-y) + ix \wedge y} \right)^2 \in \mathbb{R}_+. \end{aligned}$$

It is also translation invariant as shows equation (2.30c).

The property 2.1 implies that the propagator is the usual one: $\hat{C}(p) = 1/(p^2 + m^2)$.

2.3. UV/IR mixing. The non-locality of the \star -product allows to understand the discovery of Minwalla, Van Raamsdonk and Seiberg [21]. They showed that not only the model (2.28) isn't finite in the UV but also it exhibits a new type of divergences making it non-renormalisable. In the article [7], Filk computed the Feynman rules corresponding to (2.28). He showed that the “planar” amplitudes equal the commutative ones whereas the “non-planar” ones give rise to oscillations coupling the internal and external legs. A typical example is the the “non-planar” tadpole:

$$(2.32) \quad \begin{aligned} \text{Diagram: } & \text{A horizontal line with an arrow pointing right, labeled } p. \text{ A vertical oval loop is attached to the line, with an arrow pointing clockwise, labeled } k. \\ &= \frac{\lambda}{12} \int \frac{d^4 k}{(2\pi)^4} \frac{e^{ip_\mu k_\nu \Theta^{\mu\nu}}}{k^2 + m^2} \\ &= \frac{\lambda}{48\pi^2} \sqrt{\frac{m^2}{(\Theta p)^2}} K_1(\sqrt{m^2(\Theta p)^2}) \underset{p \rightarrow 0}{\simeq} p^{-2}. \end{aligned}$$

Remark. In the literature, the graph here above is called non-planar tadpole. Unfortunately it is a misleading denomination. The relevant class of graphs for the study of the Filk oscillations are the truncated graphs, and it is therefore their topology which classifies the different types of oscillations. If one cuts the external legs of the graph of equation (2.32), it is no longer non-planar but has two broken faces, ie its two boundary components bear external points. We refer to section 4.1 for more details.

If $p \neq 0$, this amplitude is finite but, for small p , it diverges like p^{-2} . In other words, if we put an ultraviolet cut-off Λ to the k -integral, the two limits $\Lambda \rightarrow \infty$ and $p \rightarrow 0$ do not commute. This is the UV/IR mixing phenomena. A chain of non-planar tadpoles, inserted in bigger graphs, makes divergent any function (with six points or more for example). But this divergence is not local and can't be absorbed in a mass redefinition. This is what makes the model non-renormalisable. We will

^aAnother way to prove it is from (2.10), $\overline{\phi^{\star 4}} = \phi^{\star 4}$.

see in section 5.3 that the UV/IR mixing results in a coupling of the different scales of the theory. We will also note that we should distinguish different types of mixing.

The UV/IR mixing was studied by several groups. First, Chepelev and Roiban [8] gave a power counting for different scalar models. They were able to identify the divergent graphs and to classify the divergences of the theories thanks to the topological data of the graphs. Then V. Gayral [22] showed that UV/IR mixing is present on all isospectral deformations (they consist in curved generalisations of the Moyal space and of the non-commutative torus). For this, he considered a scalar model (2.28) and discovered contributions to the effective action which diverge when the external momenta vanish. The UV/IR mixing is then a general characteristic of the non-commutative theories, at least on the deformations.

3. Multi-scale Analysis

3.1. Phase-space. Before turning to the solution to the UV/IR mixing problem, we now introduce a very efficient technique called multi-scale analysis. It allows to study carefully and precisely the renormalization group properties of Euclidean field theory.

As explained by K. Wilson [23], renormalization in Euclidean space can be interpreted as a coherence phenomenon among different *scales*. By scale, we mean here length scale. Inspired by this fact we will divide the propagator into slices in order to analyse our model scale by scale.

Our benchmark model throughout this section will be the commutative four-dimensional ϕ^4 theory:

$$(3.1) \quad L[\phi] = \int d^4x \frac{1}{2}\phi(-\Delta + m^2)\phi + \frac{\lambda}{4!}\phi^4$$

We then slice the propagator using the Schwinger trick:

$$(3.2) \quad \hat{C}(p) = \frac{1}{p^2 + m^2} = \int_0^\infty dt e^{-t(p^2 + m^2)},$$

$$(3.3) \quad C(x, y) = \int_0^\infty \frac{dt}{(4\pi t)^2} e^{-\frac{(x-y)^2}{4t} - tm^2} = \sum_{i=0}^\infty C^i(x, y),$$

$$(3.4) \quad C^0 = \int_1^\infty \frac{dt}{(4\pi t)^2} e^{-\frac{(x-y)^2}{4t} - tm^2},$$

$$(3.5) \quad C^i = \int_{M^{-2i}}^{M^{-2(i-1)}} \frac{dt}{(4\pi t)^2} e^{-\frac{(x-y)^2}{4t} - tm^2}, \quad i = 1, \dots, \infty$$

where $M \in \mathbb{R}_+$ is strictly greater than one. This slicing is useful because it allows to get a simple bound on each of the C^i 's:

Lemma 3.1. *For all $i \in \mathbb{N}$, there exists $K, k \in \mathbb{R}_+$ such that*

$$(3.6) \quad C^i(x, y) \leq K M^{2(i+1)} e^{-k M^{i+1} |x-y|}.$$

The preceding lemma allows to understand why we used the propagator to define length scales. In a slice i (that is with a propagator C^i), a “particle’s” displacement is of order M^{-i} .

Let us now consider a graph G (see figure 1 for an example). Its amplitude is given by

$$(3.7) \quad A_G = \int_{\mathbb{R}} \prod_{\nu \in G} dx_{\nu} \prod_{l \in G} C_l = \sum_{i_1, \dots, i_I} \int_{\mathbb{R}} \prod_{\nu \in G} dx_{\nu} \prod_{l \in G} C_l^{i_l} = \sum_{\mu \in \mathbb{N}^{I(G)}} A_G^{\mu}$$

where $I(G)$ is the number of internal lines of G . Let us fix $\mu \in \mathbb{N}^{I(G)}$, A_G^{μ} is a graph whose each internal line bears an integer. This one represents the scale index of the corresponding sliced propagator. The $I(G)$ -tuple μ is called a *scale attribution*. On figure 2 one can see the so-called “phase-space” representation of

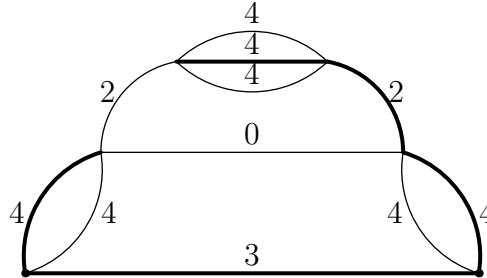


FIGURE 1. A ϕ^4 graph

the scale attribution corresponding to figure 1. The full x -space (here \mathbb{R}^4) runs along the horizontal axis while the vertical one counts the (discrete) length scales. Plain lines are propagators hence join two spatial points. For each scale attribution, each propagator is equipped with both an ultraviolet and an infrared cutoff. Then the integration over the internal vertices (the x_{ν} 's in (3.7)) in A_G^{μ} is convergent. The divergence of the graph is hidden in the sum over the scale attributions. For completely convergent graphs or renormalised ones, the dotted lines in figure 2 represent exponential decay in the scale indices. Let us compute the power-counting of our model and see how such decays appear.

3.2. Power-counting. Let a graph G , $n = n(G)$ the number of its internal vertices and $N = N(G)$ the number of its external points. Let $\mu \in \mathbb{N}^{I(G)}$. The amplitude of the graph G with scale attribution μ is given by

$$(3.8) \quad A_G^{\mu} = \int \prod_{\substack{\nu \in G, \\ \nu \neq \nu_0}} dx_{\nu} \prod_{l \in G} C_l^{i_l}(x_l, y_l).$$

In this last equation, we only integrate over $n-1$ positions because of the translation invariance of the propagators. We are going to use the bound (3.6) to integrate. We need a minimal set of propagators to perform this integration. Graphically this minimal structure is a *tree*. This is a connected subset of $n(G) - 1$ internal lines of G without loop. The bold lines in figure 1 are an example of a tree. We have then the following bound

$$(3.9) \quad A_G^{\mu} \leq K^{2n-N/2} \prod_{l \in G} M^{2(i_l+1)} \int \prod_{\substack{\nu \in G, \\ \nu \neq \nu_0}} dx_{\nu} \prod_{l \in \mathcal{T}} e^{-kM^{i_l+1}|x-y|}$$

Scale Decomposition

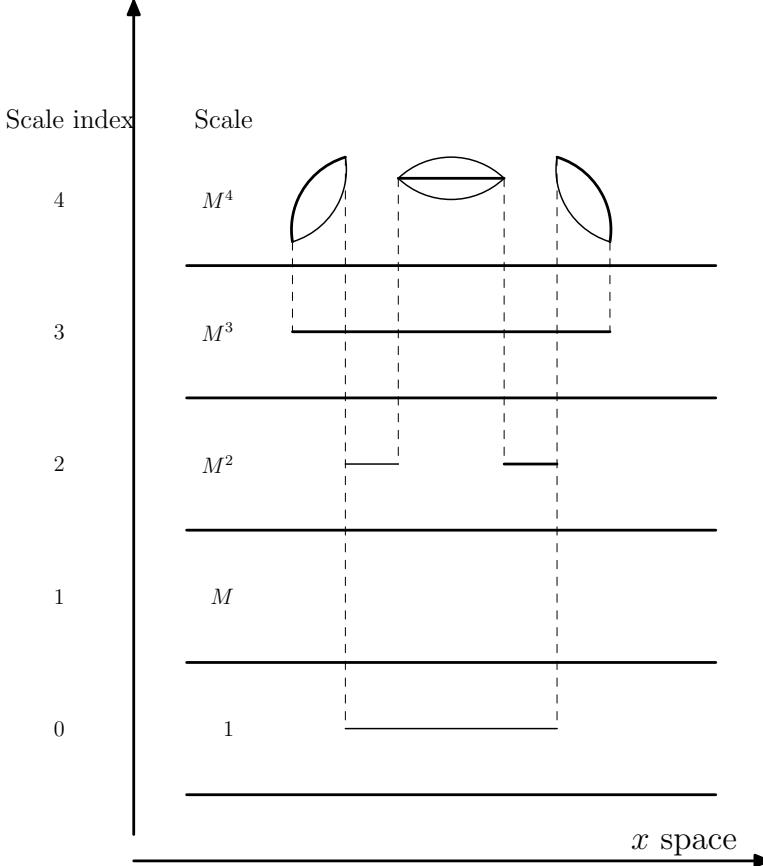


FIGURE 2. Example of scale attribution

where \mathcal{T} stands for the tree in G .

$$(3.10) \quad A_G^\mu \leq K^n \prod_{l \in G} M^{2(i_l+1)} \prod_{l \in \mathcal{T}} M^{4(i_l+1)} = K^n \prod_{l \in G} \prod_{i=0}^{i_l} M^2 \prod_{l \in \mathcal{T}} \prod_{j=0}^{i_l} M^{-4}$$

where the value of the constant K has changed from one equation to another.

We now introduce certain subgraphs of G :

$$\forall i \in \mathbb{N}, G^i := \{l \in G : i_l \geq i\}.$$

Let $i \in \mathbb{N}$, G^i is made of the internal lines of G whose indices are greater than i . In general G^i is not connected and then can be written as the disjoint union of *connected components* G_k^i 's: $G^i = \bigcup_k G_k^i$. Figure 4 shows the various connected components of the graph of figure 1 for different values of the scale parameter.

Equation (3.10) can be rewritten as

$$(3.11) \quad A_G^\mu \leq K^n \prod_{i,k} M^{-\omega(G_k^i)}, \quad \omega(G_k^i) = N(G_k^i) - 4.$$

Let us now consider the example of figure 3 where $i > j$. This is a completely conver-

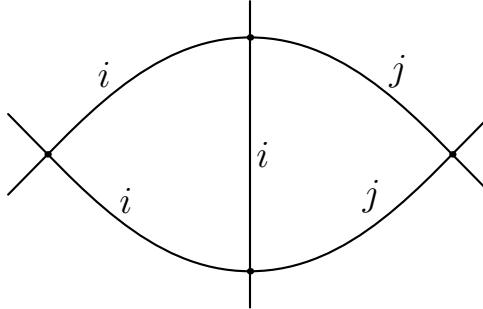


FIGURE 3. Completely convergent graph

gent graph which means that all its connected components have a strictly positive degree of convergence ω . The power-counting (3.11) gives $A_G^\mu \leq K^4 M^{-2(i-j)M^{-2j}}$. We note the exponential decay in the scale indices which allows to perform the sum over $i > j$.

For a quantum field theory to be renormalisable we first need a renormalisable power-counting. This is the case for the commutative ϕ^4 model in four dimensions (3.11): a finite number of Schwinger functions diverges. Second the divergent part of the graphs have to be proportionnal to the terms present in the initial Lagrangian. We will not explain here how this point is checked for our test model in the framework of multi-scale analysis. Nevertheless we will come back to that point in the conclusion when we will try to interpret the results we got for the non-commutative Gross-Neveu model. And before turning to this model, let us have a look at the solution given by H. Grosse and R. Wulkenhaar to the UV/IR mixing problem.

4. The Grosse-Wulkenhaar breakthrough

H. Grosse and R. Wulkenhaar discovered a way to define a renormalisable non-commutative model. We are going to detail their results but the main message is the following. By adding an harmonic term to the Lagrangian (2.28),

(4.1)

$$S[\phi] = \int d^4x \left(-\frac{1}{2} \partial_\mu \phi \star \partial^\mu \phi + \frac{\Omega^2}{2} (\tilde{x}_\mu \phi) \star (\tilde{x}^\mu \phi) + \frac{1}{2} m^2 \phi \star \phi + \frac{\lambda}{4} \phi \star \phi \star \phi \star \phi \right) (x)$$

where $\tilde{x} = 2\Theta^{-1}x$ and the metric is Euclidean, the model, in four dimensions, is renormalisable at all orders of perturbation [5]. We will see that this additional term gives rise to an infrared cut-off and allows to decouple the different scales of the theory. The new model (4.1), we call it Φ_4^4 , do not exhibit any mixing. This result is very important because it opens the way towards other non-commutative field theories. In the following, we will call *vulcanisation*^b the procedure consisting in adding a new term to a Lagrangian of a non-commutative theory in order to

^bTECHNOL. Opration consistant traiter le caoutchouc naturel ou synthtique par addition de soufre, pour en amliorer les proprits mcaniques et la rsistance aux variations de temptrature, Trsor de la Langue Franaise informatis, <http://www.lexilogos.com/>.

Scale decomposition

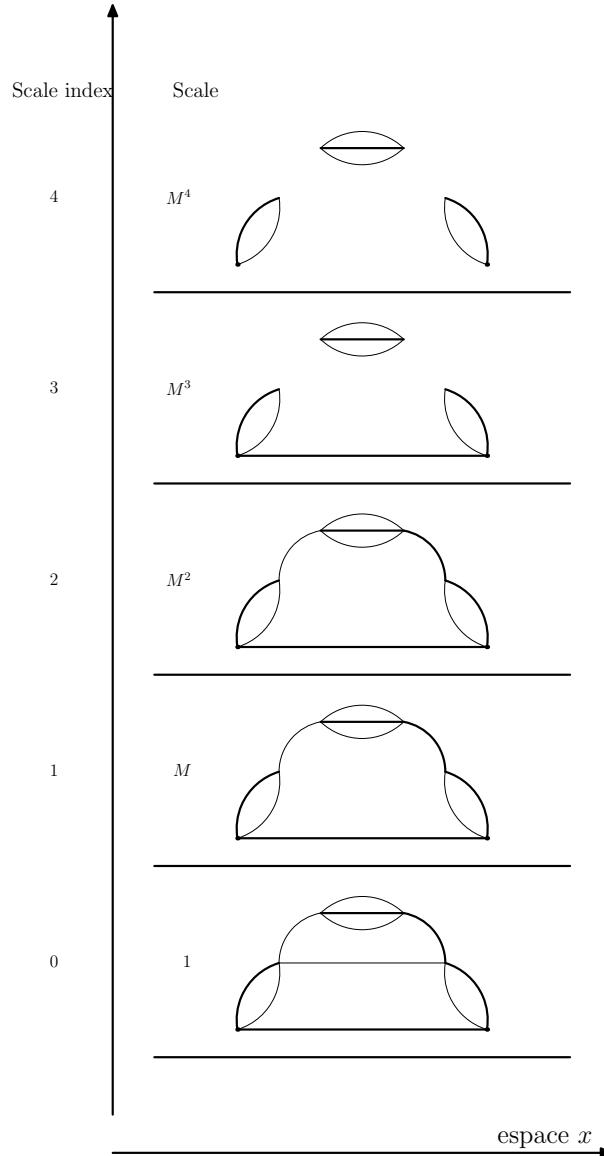


FIGURE 4. Connected components

make it renormalisable.

The propagator C of this Φ^4 theory is the kernel of the inverse operator $-\Delta + \Omega^2 \tilde{x}^2 + m^2$. It is known as the Mehler kernel [24, 18]

$$(4.2) \quad C(x, y) = \frac{\Omega^2}{\theta^2 \pi^2} \int_0^\infty \frac{dt}{\sinh^2(2\tilde{\Omega}t)} e^{-\frac{\tilde{\Omega}}{2} \coth(2\tilde{\Omega}t)(x-y)^2 - \frac{\tilde{\Omega}}{2} \tanh(2\tilde{\Omega}t)(x+y)^2 - m^2 t}.$$

Langmann and Szabo remarked that the quartic interaction with Moyal product is invariant under a duality transformation. It is a symmetry between momentum and direct space. The interaction part of the model (4.1) is (see equation (2.17))

$$(4.3) \quad S_{\text{int}}[\phi] = \int d^4x \frac{\lambda}{4} (\phi \star \phi \star \phi \star \phi)(x)$$

$$(4.4) \quad = \int \prod_{a=1}^4 d^4x_a \phi(x_a) V(x_1, x_2, x_3, x_4)$$

$$(4.5) \quad = \int \prod_{a=1}^4 \frac{d^4p_a}{(2\pi)^4} \hat{\phi}(p_a) \hat{V}(p_1, p_2, p_3, p_4)$$

with

$$V(x_1, x_2, x_3, x_4) = \frac{\lambda}{4} \frac{1}{\pi^4 \det \Theta} \delta(x_1 - x_2 + x_3 - x_4) \cos(2(\Theta^{-1})_{\mu\nu}(x_1^\mu x_2^\nu + x_3^\mu x_4^\nu))$$

$$\hat{V}(p_1, p_2, p_3, p_4) = \frac{\lambda}{4} (2\pi)^4 \delta(p_1 - p_2 + p_3 - p_4) \cos\left(\frac{1}{2} \Theta^{\mu\nu} (p_{1,\mu} p_{2,\nu} + p_{3,\mu} p_{4,\nu})\right)$$

where we used a *cyclic* Fourier transform: $\hat{\phi}(p_a) = \int dx e^{(-1)^a \imath p_a x_a} \phi(x)$. The transformation

$$(4.6) \quad \hat{\phi}(p) \leftrightarrow \pi^2 \sqrt{|\det \Theta|} \phi(x), \quad p_\mu \leftrightarrow \tilde{x}_\mu$$

exchanges (4.4) and (4.5). In addition, the free part of the model (2.28) isn't covariant under this duality. The vulcanisation adds a term to the Lagrangian which restores the symmetry. The theory (4.1) is then covariant under the Langmann-Szabo duality:

$$(4.7) \quad S[\phi; m, \lambda, \Omega] \mapsto \Omega^2 S[\phi; \frac{m}{\Omega}, \frac{\lambda}{\Omega^2}, \frac{1}{\Omega}].$$

Thus, by symmetry, one can always choose the parameter Ω in $[0, 1]$. Let us note that for $\Omega = 1$, the model is invariant.

. The interpretation of that harmonic term is not yet clear. But the vulcanisation procedure already allowed to prove the renormalisability of several other models on Moyal spaces such that ϕ_2^4 [25], $\phi_{2,4}^3$ [26, 27] and the LSZ models [10, 28, 29]. These last are of the type

$$(4.8) \quad S[\phi] = \int d^4x \left(\frac{1}{2} \bar{\phi} \star (-\partial_\mu + \tilde{x}_\mu + m)^2 \phi + \frac{\lambda}{4} \bar{\phi} \star \phi \star \bar{\phi} \star \phi \right)(x).$$

By comparison with (4.1), one notes that here the additional term is formally equivalent to a fixed magnetic background. Deep is the temptation to interpret it as such. This model is invariant under the above duality and is exactly soluble. Let us remark that the complex interaction in (4.8) makes the Langmann-Szabo duality more natural. It doesn't need a cyclic Fourier transform. The ϕ^3 have been studied at $\Omega = 1$ where they also exhibit a soluble structure.

4.1. Topology and power counting. Let a graph G with V vertices and I internal lines. Interactions of quantum field theories on the Moyal space are only invariant under **cyclic permutation** of the incoming/outcoming fields. This restricted invariance replaces the permutation invariance which was present in the case of local interactions.

A good way to keep track of such a reduced invariance is to draw Feynman graphs as ribbon graphs. Moreover there exists a basis for the Schwartz class functions (hereafter called matrix basis) where the Moyal product becomes an ordinary matrix product [25, 20]. This further justifies the ribbon representation.

Let us consider the example of figure 5. Propagators in a ribbon graph are

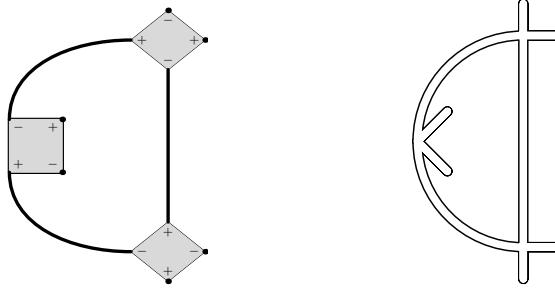


FIGURE 5. A graph with two broken faces, in x -space representation (left) and Ribbon representation (right)

made of double lines. Let us call F the number of faces (loops made of single lines) of a ribbon graph. The graph of figure 5b has $V = 3, I = 3, F = 2$ (after having “closed” the external legs). Each ribbon graph can be drawn on a manifold of genus g . The genus is computed from the Euler characteristic $\chi = F - I + V = 2 - 2g$. If $g = 0$ one has a *planar graph*, otherwise one has a *non-planar graph*. For example, the graph of figure 5b may be drawn on a manifold of genus 0. Note that some of the F faces of a graph may be “broken” by external legs. In our example, both faces are broken. We denote the number of broken faces by B .

Furthermore let N the number of external legs of the graph. For the commutative ϕ^4 model one has the following superficial degree of convergence $\omega = N - 4$. Thus one has to deal only with the renormalization of the two- and four-point functions. In the case of the Grosse-Wulkenhaar model, the situation is different. In [4, 5, 30] it was proven that

$$(4.9) \quad \omega = (N - 4) + 8g + 4(B - 1).$$

Note that, as proven in [19] the power counting of the Gross-Neveu model (lemma 5.3) is more involved but leads to the same conclusion: one has to deal only with the renormalization of the $B = 1$, planar two- and four-point graphs hereafter qualified as *planar regular*.

4.2. Orientation. The power counting (4.9) has been proven in the matrix basis [4, 5]. In this subsection we refer to a work done in x -space [3]. The main advantage of the matrix basis is that it avoids the difficult treatment of the vertex oscillations (see equation (2.29)). But it renders the propagator very complicated. On the contrary in x -space the propagator is nice but the vertices oscillate. From my point of view, the x -space has the major advantage of allowing the comparison between ordinary and non-commutative renormalisation group.

As in section 3, we cut the propagator (4.2) into slices:

$$(4.10) \quad C^i(x, y) = \frac{\Omega^2}{\theta^2 \pi^2} \int_0^\infty \frac{dt}{\sinh^2(2\tilde{\Omega}t)} e^{-\frac{\tilde{\Omega}}{2} \coth(2\tilde{\Omega}t)(x-y)^2 - \frac{\tilde{\Omega}}{2} \tanh(2\tilde{\Omega}t)(x+y)^2 - m^2 t}$$

$$(4.11) \quad C^i(x, y) \leq K M^{2(i+1)} e^{-k M^{i+1} |x-y| - k M^{-i-1} |x+y|}.$$

We see on the bound (4.11) that the harmonic x^2 term added to the Lagrangian gives a propagator which both decays in $x-y$ and $x+y$. Note that the length scales associated to these decay are inverse of each other. This additional decay (compare with (3.6)) allows to control the vertex oscillations and get a renormalisable power counting. We will not detail here how to get a bound on a general graph's amplitude (for a review see [31]) but instead we will comment the result.

Before stating the power counting lemma we get in [3], we need to introduce one more notion.

Definition 4.1. *A graph is **orientable** if its lines can be equipped with arrows alternatively incoming and outgoing around the vertices.*

Figure 6 shows examples of orientable and non-orientable graphs.

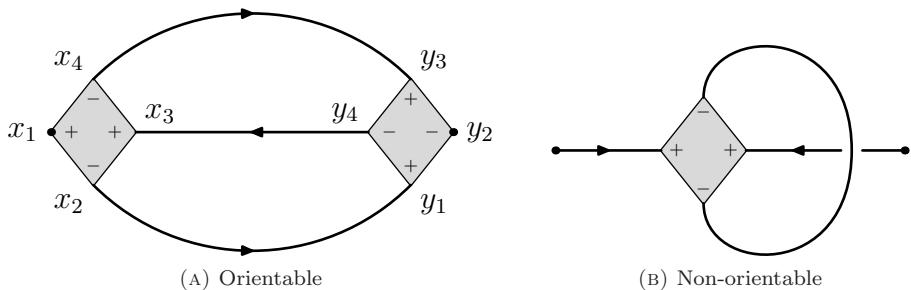


FIGURE 6. Orientable or not

Lemma 4.1. *Let G a connected graph. For all $\Omega \in]0, 1]$, there exists $K \in \mathbb{R}_+$ such that the amputated amplitude A_G^μ integrated over test functions is bounded by*

$$(4.12) \quad |A_G^\mu| \leq K^n \prod_{i,k} M^{-\omega(G_k^i)}$$

$$(4.13) \quad \text{with } \omega(G_k^i) = \begin{cases} N & \text{if } G_k^i \text{ is non-orientable,} \\ & \text{or if } G_k^i \text{ is orientable, } g = 0 \text{ and } B \geq 2, \\ N + 4 & \text{if } G_k^i \text{ is orientable, } g \geq 1, \\ N - 4 & \text{if } G_k^i \text{ is orientable, } g = 0 \text{ and } B = 1. \end{cases}$$

This lemma proves that only the planar regular graphs may diverge but the bound (4.13) is less precise than (4.9). It allowed nevertheless to singularize orientable graphs. Why is orientation a relevant notion? That's a point we will address in the next subsection.

4.3. Locality vs Moyality. The main problem with non-commutative field theory is UV/IR mixing. Technically, for the two-point function, it consists in graphs whose divergent part is not local. Let us consider one of those graphs. If its internal lines are of length M^{-i} with $i \gg 1$ (far UV), its divergence arises when external point are M^i (far IR) away from each other. Such a coupling between scales inverse from each other should not be surprising on a non-commutative space whose commutation relation is $[x^\mu, x^\nu] = i\Theta^{\mu\nu}$.

A crucial aspect of the uplifting from commutative to non-commutative renormalization is that the principle of locality of renormalized interactions of commutative QFT is replaced with a new principle: renormalized interactions have a non-local Moyal vertex form. This is nothing but the analog of the locality phenomenon which occurs in commutative renormalization. Moreover one can prove that the orientable two-point graphs are always local. This implies that the UV/IR mixing phenomenon is entirely due to non-orientable graphs which makes orientation a relevant notion for non-commutative field theory.

One can thus speak of a new type of renormalization group, where the locality is just replaced by ‘‘Moyality’’. The divergent parts of the planar regular two- and four-point graphs with one broken face (the only divergent graphs) are proportional to the (1PI) tree level terms of the perturbative expansion. Let us also argue here that, despite this uplifting, the combinatorial backbone of renormalization theory is almost the same when dealing with commutative or non-commutative QFT. In fact the combinatorics of non-commutative renormalization has been shown to be encoded by a Hopf algebra [32].

5. The non-commutative Gross-Neveu model

Apart from the Φ_4^4 , the modified Bosonic LSZ model [3] and supersymmetric theories [33, 34], we now know several renormalizable non-commutative field theories. Nevertheless they either are super-renormalizable (Φ_2^4 [25]) or (and) studied at a special point in the parameter space where they are solvable ($\Phi_2^3, \Phi_4^3, \Phi_6^3$ [26, 27, 35], the LSZ models [10, 28, 29]). Although only logarithmically divergent for parity reasons, the non-commutative Gross-Neveu model is a just renormalizable quantum field theory as Φ_4^4 . One of its main interesting features is that it can be interpreted as a non-local Fermionic field theory in a constant magnetic background. Then apart from strengthening the ‘‘vulcanization’’ procedure to get renormalizable non-commutative field theories, the Gross-Neveu model may also be useful for the study of the quantum Hall effect. Moreover its commutative counterpart being asymptotically free and exhibiting dynamical mass generation [36, 37, 38], a study of the physics of this model would be interesting.

5.1. The commutative model. Before reviewing the results we get about the non-commutative Gross-Neveu model, let us recall the main features of its commutative counterpart. The Gross-Neveu model on \mathbb{R}^2 is defined by the following action

$$(5.1) \quad S[\bar{\psi}, \psi] = \int dx \left(\bar{\psi} (\imath \not{\partial} + m) \psi + \left(\sum_\alpha \bar{\psi}_\alpha \psi_\alpha \right)^2 \right) (x)$$

where ψ is a two-dimensional spinor, $\not{\partial} = \gamma^\mu \partial_\mu$. The γ^0 and γ^1 matrices obey the Clifford relation $\{\gamma^\mu, \gamma^\nu\} = -2\delta^{\mu\nu}$. The sum over α runs over spin indices (indices of the γ^μ matrices) and also on an N -dimensional color space. If $N \geq 2$ one replaces

ψ by N spinors and the propagator is then considered diagonal in the color space indices.

If $N = 1$ due to locality there exists only one possible interaction:

$$(5.2) \quad \bar{\psi}_\uparrow \psi_\uparrow \bar{\psi}_\downarrow \psi_\downarrow = \frac{1}{2} \left(\sum_\alpha \bar{\psi}_\alpha \psi_\alpha \right)^2 = \frac{1}{4} \bar{\psi} \gamma^\mu \psi \bar{\psi} \gamma_\mu \psi.$$

The last term in (5.2) is known as the Thirring interaction. The model (5.1) is just renormalisable in two dimensions. It is asymptotically free in the UV [36]. For $N = 1$ the β function even vanishes identically.

The model has been rigorously constructed by two different groups [39, 40]. It also exhibits non-perturbative mass generation as studied by D. Gross and A. Neveu [37] and proven by C. Kopper, J. Magnen and V. Rivasseau [38].

5.2. The non-commutative Gross-Neveu model. The non-commutative Gross-Neveu model (GN_Θ^2) is a Fermionic quartically interacting quantum field theory on the Moyal plane \mathbb{R}_θ^2 . The skew-symmetric matrix Θ is

$$(5.3) \quad \Theta = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix}.$$

The action is

$$(5.4) \quad S[\bar{\psi}, \psi] = \int dx \left(\bar{\psi} (-i\cancel{d} + \Omega \cancel{x} + m + \mu \gamma_5) \psi + V_o(\bar{\psi}, \psi) + V_{\text{no}}(\bar{\psi}, \psi) \right) (x)$$

where $\tilde{x} = 2\Theta^{-1}x$, $\gamma_5 = i\gamma^0\gamma^1$ and $V = V_o + V_{\text{no}}$ is the interaction part given hereafter. The μ -term appears at two-loop order. We use a Euclidean metric and the Feynman convention $\cancel{d} = \gamma^\mu a_\mu$. The γ^0 and γ^1 matrices form a two-dimensional representation of the Clifford algebra $\{\gamma^\mu, \gamma^\nu\} = -2\delta^{\mu\nu}$. Let us remark that the γ^μ 's are then skew-Hermitian: $\gamma^{\mu\dagger} = -\gamma^\mu$.

Propagator. The propagator corresponding to the action (5.4) is given by the following lemma:

Lemma 5.1 (Propagator [18]). *The propagator of the Gross-Neveu model is*

$$(5.5) \quad C(x, y) = \int d\mu_C(\bar{\psi}, \psi) \psi(x) \bar{\psi}(y) = (-i\cancel{d} + \Omega \cancel{x} + m)^{-1}(x, y) \\ = \int_0^\infty dt C(t; x, y),$$

$$(5.6) \quad C(t; x, y) = -\frac{\Omega}{\theta\pi} \frac{e^{-tm^2}}{\sinh(2\tilde{\Omega}t)} e^{-\frac{\tilde{\Omega}}{2} \coth(2\tilde{\Omega}t)(x-y)^2 + i\Omega x \wedge y} \\ \times \left\{ i\tilde{\Omega} \coth(2\tilde{\Omega}t)(\cancel{x} - \cancel{y}) + \Omega(\cancel{x} - \cancel{y}) - m \right\} e^{-2i\Omega t \gamma \Theta^{-1} \gamma}$$

with $\tilde{\Omega} = \frac{2\Omega}{\theta}$ et $x \wedge y = 2x\Theta^{-1}y$.

We also have $e^{-2i\Omega t \gamma \Theta^{-1} \gamma} = \cosh(2\tilde{\Omega}t) \mathbb{1}_2 - i\frac{\theta}{2} \sinh(2\tilde{\Omega}t) \gamma \Theta^{-1} \gamma$.

If we want to study a N -color model, we can consider a propagator diagonal in these color indices.

Interactions. Concerning the interaction part V , recall that (see corollary 2.3) for any f_1, f_2, f_3, f_4 in \mathcal{A}_Θ ,

$$(5.7) \quad \int dx (f_1 \star f_2 \star f_3 \star f_4)(x) = \frac{1}{\pi^2 \det \Theta} \int \prod_{j=1}^4 dx_j f_j(x_j) \delta(x_1 - x_2 + x_3 - x_4) e^{-i\varphi},$$

$$(5.8) \quad \varphi = \sum_{i < j=1}^4 (-1)^{i+j+1} x_i \wedge x_j.$$

This product is non-local and only invariant under cyclic permutations of the fields. Then, contrary to the commutative Gross-Neveu model, for which there exists only one spinorial interaction, the GN_Θ^2 model has, at least, six different interactions: the *orientable* ones

$$(5.9a) \quad V_o = \frac{\lambda_1}{4} \int dx (\bar{\psi} \star \psi \star \bar{\psi} \star \psi)(x)$$

$$(5.9b) \quad + \frac{\lambda_2}{4} \int dx (\bar{\psi} \star \gamma^\mu \psi \star \bar{\psi} \star \gamma_\mu \psi)(x)$$

$$(5.9c) \quad + \frac{\lambda_3}{4} \int dx (\bar{\psi} \star \gamma_5 \psi \star \bar{\psi} \star \gamma_5 \psi)(x),$$

where ψ 's and $\bar{\psi}$'s alternate and the *non-orientable* ones

$$(5.10a) \quad V_{no} = \frac{\lambda_4}{4} \int dx (\psi \star \bar{\psi} \star \bar{\psi} \star \psi)(x)$$

$$(5.10b) \quad + \frac{\lambda_5}{4} \int dx (\psi \star \gamma^\mu \bar{\psi} \star \bar{\psi} \star \gamma_\mu \psi)(x)$$

$$(5.10c) \quad + \frac{\lambda_6}{4} \int dx (\psi \star \gamma_5 \bar{\psi} \star \bar{\psi} \star \gamma_5 \psi)(x).$$

All these interactions have the same x kernel thanks to the equation (5.7). The reason for which we call these interactions orientable or not is the following: let us assign outgoing arrows to the ψ 's fields (say) and incoming ones to the $\bar{\psi}$'s. For the so-called orientable interactions, ψ 's and $\bar{\psi}$'s alternate around the vertices. Moreover the propagator only links a ψ to a $\bar{\psi}$. Then thanks to definition 4.1 orientable interactions only create orientable graphs. It implies that all the non-orientable graphs are made of non-orientable interactions. Note that non-orientable interactions also create orientable graphs.

With that setting, we have proven the following theorem:

Theorem 5.2 (BPHZ for GN_Θ^2 [19]). *For all $\Omega \in [0, 1)$, the quantum field theory defined by the action*

$$S[\bar{\psi}, \psi] = \int dx (\bar{\psi} (\imath \partial^\mu + \Omega \vec{x}^\mu + m + \mu \gamma_5) \psi + V_o(\bar{\psi}, \psi))(x)$$

is renormalisable at all orders of perturbation.

This result has been proven with only orientable interactions. In the next subsection we explain why it is difficult to extend with non-orientable graphs. Note also that in subsection 4.3 we reported that UV/IR mixing is due to non-orientable

graphs. This implies that a (non-commutative) field theory producing only orientable graphs is very likely to be renormalisable with the usual inverse Laplacian propagator. A special case is the one we consider in theorem 5.2 where the model is renormalisable even for $\Omega = 0$.

5.3. A renormalisable mixing. In the following when we refer to a propagator from x to y , we call short variable the difference $u = x - y$ and long variable the sum $v = x + y$. Remind that the Φ^4 propagator decays both in u and v , see (4.2). The qualifications for these variables should be clear from (4.11).

The main technical difficulty of the GN_Θ^2 model is the absence of decreasing functions for the long v variables in the propagator replaced by an oscillation, see lemma 5.1 ($x \wedge y = \frac{1}{2}u \wedge v$). Note that these decreasing functions are in principle created by integration over the u variables:

$$(5.11) \quad \int du e^{-\frac{\tilde{\Omega}}{2} \coth(2\tilde{\Omega}t)u^2 + iu \wedge v} = K \tanh(2\tilde{\Omega}t) e^{-k \tanh(2\tilde{\Omega}t)v^2}.$$

But to perform all these Gaussian integrations for a general graph is a difficult task (see [41]) and is in fact not necessary for a BPHZ theorem. We can instead exploit the vertices and propagators oscillations to get rational decreasing functions in some linear combinations of the long v variables. The difficulty is then to prove that all these linear combinations are independent and hence allow to integrate over all the v variables (this is where we stuck for non-orientable graphs). To solve this problem we need the exact expression of the total oscillation in terms of the short and long variables. This consists in a generalization of the Filk's work [7]. This has been done in [19]. Once the oscillations are proven independent, one can just use the same arguments than in the Φ^4 case to compute an upper bound for the power counting:

Lemma 5.3 (Power counting GN_Θ^2). *Let G a connected orientable graph. For all $\Omega \in [0, 1)$, there exists $K \in \mathbb{R}_+$ such that its amputated amplitude A_G integrated over test functions is bounded by*

$$(5.12) \quad |A_G| \leq K^n M^{-\frac{1}{2}\omega(G)}$$

$$(5.13) \quad \text{with } \omega(G) = \begin{cases} N - 4 & \text{if } (N = 2 \text{ or } N \geq 6) \text{ and } g = 0, \\ & \text{if } N = 4, g = 0 \text{ and } B = 1, \\ & \text{if } G \text{ is critical,} \\ N & \text{if } N = 4, g = 0, B = 2 \text{ and } G \text{ non-critical,} \\ N + 4 & \text{if } g \geq 1. \end{cases}$$

As in the non-commutative Φ^4 case, only the planar graphs are divergent. But the behaviour of the graphs with more than one broken face is different. In the multiscale framework, the Feynmann diagrams are endowed with a scale attribution which gives each line a scale index. The only subgraphs we meet in this setting have all their internal scales higher than their external ones. Then a subgraph G of scale i is called *critical* if it has $N = 4, g = 0, B = 2$ and that the two “external” points in the second broken face are only linked by a single line of scale $j < i$. The typical example is the graph of figure 7a. In this case, the subgraph is logarithmically divergent whereas it is convergent in the Φ^4 model. Let us now show roughly how it happens in the case of figure 7a in x -space.

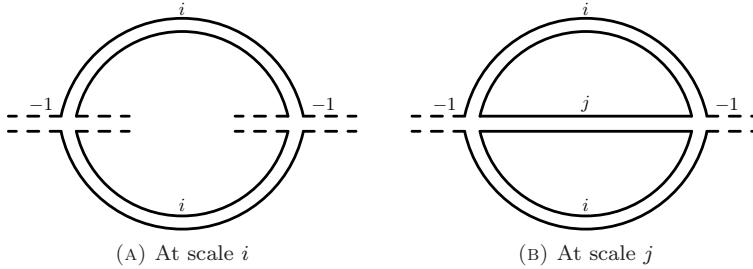


FIGURE 7. Sunset Graph

The same arguments than in the Φ^4 model (see [3] or [31] for a review) prove that the integrations over the internal points of the graph 7a lead to a logarithmical divergence which means that $A_{G^i} \simeq \mathcal{O}(1)$ in the multiscale framework. But there is a remaining oscillation between a long variable of this graph and the external points in the second broken face of the form $v \wedge (x - y)$. But v is of order M^i which leads to a decreasing function implementing $x - y$ of order M^{-i} . If these points are true external ones, they are integrated over test functions of norm 1. Then thanks to the additional decreasing function for $x - y$ we gain a factor M^{-2i} which makes the graph convergent. But if x and y are linked by a single line of scale $j < i$ (as in figure 7b), instead of test functions we have a propagator between x and y . This one behaves like (see lemma 5.1):

$$(5.14) \quad C^j(x, y) \simeq M^j e^{-M^{2j}(x-y)^2 + i x \wedge y}.$$

The integration over $x - y$ instead of giving M^{-2j} gives M^{-2i} thanks to the oscillation $v \wedge (x - y)$. Then we have gained a good factor $M^{-2(i-j)}$. But the oscillation in the propagator $x \wedge y$ now gives $x + y \simeq M^{2i}$ instead of M^{2j} and the integration over $x + y$ cancels the preceding gain. The critical component of figure 7a is logarithmically divergent.

This kind of argument can be repeated and refined for more general graphs to prove that this problem appears only when the external points of the auxiliary broken faces are linked only by a *single* lower line [19]. This phenomenon can be seen as a mixing between scales. Indeed the power counting of a given subgraph now depends on the graphs at lower scales. This was not the case in the commutative realm. Fortunately this mixing doesn't prevent renormalisation. Note that whereas the critical subgraphs are not renormalisable by a vertex-like counterterm, they are regularised by the renormalisation of the two-point function at scale j . The proof of this point relies heavily on the fact that there is only one line of lower scale.

Let us conclude this section by mentionning the flows of the critical models. One very interesting feature of the non-commutative Φ^4 model is the boundedness of its flows and even the vanishing of its beta function for a special value of its bare parameters [42, 43]. Note that its commutative counterpart (the usual ϕ^4 model on \mathbb{R}^4) is asymptotically free in the infrared and has then an unbounded flow. It turns out that the flow of the non-commutative Gross-Neveu model is not regularized by the non-commutativity. The one-loop computation of its beta functions [44] shows that it is asymptotically free in the ultraviolet region as in the commutative case.

6. Discussion

In this section we would like to give some arguments about the interpretation of the additional x^2 term in the Φ^4 propagator. First of all let us consider the partition function, for example, of the commutative ϕ^4 model:

$$(6.1) \quad Z = \int d\mu_C(\phi) \exp \left(-\frac{\lambda}{4} \int d^4x \phi^4(x) \right)$$

where μ_C is the Gaussian measure associated to the covariance (or propagator) $C = (-\Delta)^{-1}$. As we mentionned in section 3 Wilson interpreted (Euclidean) renormalisation as a coherence phenomenon between different scales. Inspired by this interpretation multi-scale analysis slices the propagator. Each slice defines a (length) scale through the behaviour (3.6). Moreover it leads to a factorization of the Gaussian measure:

$$(6.2) \quad d\mu_C(\phi) = \bigotimes_i d\mu_{C^i}(\phi^i), \quad \phi = \sum_i \phi^i$$

making then possible to perform the functional integration step by step and allowing us to construct that coherence process. When we presented multi-scale analysis we worked with graphs and scale attributions. This means that we implicitly performed the full functional integration at once. On the contrary one could well have been performed the renormalisation slice by slice. Let us assume that we integrated over all the fields down to a certain ϕ^i (included). The graphs we meet then have internal lines with scales greater than i . At external points are fields with scales (strictly) lower than i . By construction these graphs have all their internal scales higher than their external ones. They are the connected components defined in section 3. The power counting naturally factorizes into the connected components contributions. All of them are renormalisable i.e. their divergent part is local.

Now imagine we perform the functional integration form the lowest scales to the highest. The graphs we meet have their internal scales lower than their external ones: they are not local anymore. We lose the Wilson's picture. It means that the renormalisation group is directed from the highest scales down to the lowest ones: from ultraviolet to infrared.

Let us now consider the naive non-commutative ϕ^4 model (2.28). It suffers UV/IR mixing. This phenomenon has two main features. First it mixes the scales of the theory. It means that the power counting of a graph depends on the graph it is inserted in that is on what happens at scales lower than its ones. The power counting can't factorize. Second certain (non-orientable) graphs are non-local (or non-Moyal-like). With the usual inverse Laplacian as propagator, we try to define a renormalisation group directed from very small to very large distances. But what does it mean in a Moyal space? We know that small and large distances are coupled (see (2.27)) and this is what we are faced with when we try to deal with such a naive model.

Adding the x^2 term, the propagator now decays both on small and large length scales (equation (4.11)). It implements the duality between such scales. With such a propagation we define a renormalisation group from both small and large scales to some intermediate scale of order $\sqrt{\theta}$ (see (2.27) and (2.1)). The power counting factorizes into the connected component contributions. These ones are renormalisable. This suggests that to exhibit a renormalisable field theory on a non-commutative

space, we first need to find the right propagation on it.

Let us have a look at the non-commutative Gross-Neveumodel. Its Dirac operator behaves very differently than $\sqrt{-\Delta + x^2}$. Is that operator as good as the Grosse-Wulkenhaar one? I mean: does it allow to define a renormalisable field theory? Theorem 5.2 seems to answer the question. But don't forget that with only orientable interactions we avoid the problem of UV/IR mixing. Hence we have to study the non-orientable interactions [45]. Moreover even if it happens that the GN_Θ^2 propagator makes the non-orientable graphs finite (as in the Φ^4 model) there will remain some renormalisable mixing (as explained in subsection 5.3). I would then interpret this as a badly directed renormalisation group. The main question is thus: what is the precise link between the Grosse-Wulkenhaar propagator and the Moyal space? This point remains to be clarified [46].

References

- [1] Alain Connes, M. R. Douglas, and A. Schwarz. “Noncommutative geometry and matrix theory: Compactification on tori”. *JHEP*, **02**:003, 1998. [arXiv:hep-th/9711162](https://arxiv.org/abs/hep-th/9711162).
- [2] Nathan Seiberg and Edward Witten. “String theory and noncommutative geometry”. *JHEP*, **09**:032, 1999. [arXiv:hep-th/9908142](https://arxiv.org/abs/hep-th/9908142).
- [3] Razvan Gurau, Jacques Magnen, Vincent Rivasseau, and Fabien Vignes-Tourneret. “Renormalization of non-commutative Φ_4^4 field theory in x space”. *Commun. Math. Phys.*, **267** (2): 515–542, 2006a. [arXiv:hep-th/0512271](https://arxiv.org/abs/hep-th/0512271).
- [4] Harald Grosse and Raimar Wulkenhaar. “Power-counting theorem for non-local matrix models and renormalisation”. *Commun. Math. Phys.*, **254** (1):91–127, 2005a. [arXiv:hep-th/0305066](https://arxiv.org/abs/hep-th/0305066).
- [5] Harald Grosse and Raimar Wulkenhaar. “Renormalisation of ϕ^4 -theory on noncommutative \mathbb{R}^4 in the matrix base”. *Commun. Math. Phys.*, **256** (2):305–374, 2005b. [arXiv:hep-th/0401128](https://arxiv.org/abs/hep-th/0401128).
- [6] Vincent Rivasseau, Fabien Vignes-Tourneret, and Raimar Wulkenhaar. “Renormalization of noncommutative ϕ^4 -theory by multi-scale analysis”. *Commun. Math. Phys.*, **262**:565–594, 2006. [arXiv:hep-th/0501036](https://arxiv.org/abs/hep-th/0501036).
- [7] T. Filk. “Divergencies in a field theory on quantum space”. *Phys. Lett.*, **B376**:53–58, 1996.
- [8] Iouri Chepelev and Radu Roiban. “Convergence theorem for non-commutative Feynman graphs and renormalization”. *JHEP*, **03**:001, 2001. [arXiv:hep-th/0008090](https://arxiv.org/abs/hep-th/0008090).
- [9] Iouri Chepelev and Radu Roiban. “Renormalization of quantum field theories on noncommutative \mathbb{R}^d . i: Scalars”. *JHEP*, **05**:037, 2000. [arXiv:hep-th/9911098](https://arxiv.org/abs/hep-th/9911098).
- [10] E. Langmann, R. J. Szabo, and K. Zarembo. “Exact solution of quantum field theory on noncommutative phase spaces”. *JHEP*, **01**:017, 2004. [arXiv:hep-th/0308043](https://arxiv.org/abs/hep-th/0308043).
- [11] Michael R. Douglas and Nikita A. Nekrasov. “Noncommutative field theory”. *Rev. Mod. Phys.*, **73**:977–1029, 2001. [arXiv:hep-th/0106048](https://arxiv.org/abs/hep-th/0106048).
- [12] Dirk Kreimer. “On the hopf algebra structure of perturbative quantum field theories”. *Adv. Theor. Math. Phys.*, **2**:303–334, 1998. [arXiv:q-alg/9707029](https://arxiv.org/abs/q-alg/9707029).
- [13] Alain Connes and Dirk Kreimer. “Renormalization in Quantum Field Theory and the Riemann-Hilbert Problem I: The Hopf Algebra Structure of Graphs and the Main Theorem”. *Commun. Math. Phys.*, **210** (1):249–273, March 2000. URL <http://www.springerlink.com/content/j3hwa3qnen9uec9h/?p=824382c0bd2c4071a0e52468fc400417&pi=4>.
- [14] Alain Connes and Dirk Kreimer. “Renormalization in Quantum Field Theory and the Riemann–Hilbert Problem II: The β -Function, Diffeomorphisms and the Renormalization Group”. *Commun. Math. Phys.*, **216** (1):215–241, January 2001. URL <http://www.springerlink.com/content/5qe9upa1j9e57rfw/?p=824382c0bd2c4071a0e52468fc400417&pi=3>.
- [15] Alain Connes and Matilde Marcolli. “From physics to number theory via noncommutative geometry. II: Renormalization, the Riemann-Hilbert correspondence, and motivic Galois theory”. 2004. [arXiv:hep-th/0411114](https://arxiv.org/abs/hep-th/0411114).

- [16] E. Leichtnam. “Scaling group flow and Lefschetz trace formula for laminated spaces with p -adic transversal”. ArXiv Mathematics e-prints, March 2006. [arXiv:math/0603576](https://arxiv.org/abs/math/0603576).
- [17] Abdelmalek Abdesselam. “The Jacobian conjecture as a problem of perturbative quantum field theory”. *Annales Henri Poincaré*, **4**:199–215, 2003. [arXiv:math/0208173](https://arxiv.org/abs/math/0208173).
- [18] Razvan Gurau, Vincent Rivasseau, and Fabien Vignes-Tourneret. “Propagators for non-commutative field theories”. *Ann. H. Poincaré*, **7** (7-8):1601–1628, 2006b. [arXiv:hep-th/0512071](https://arxiv.org/abs/hep-th/0512071), doi:10.1007/s00023-006-0296-2.
- [19] Fabien Vignes-Tourneret. “Renormalization of the orientable non-commutative Gross-Neveu model”. *Ann. H. Poincaré*, **8** (3):427–474, June 2007. [arXiv:math-ph/0606069](https://arxiv.org/abs/math-ph/0606069).
- [20] J. M. Gracia-Bondía and J. C. Várilly. “Algebras of distributions suitable for phase space quantum mechanics. I”. *J. Math. Phys.*, **29**:869–879, 1988.
- [21] Shiraz Minwalla, Mark Van Raamsdonk, and Nathan Seiberg. “Noncommutative perturbative dynamics”. *JHEP*, **02**:020, 2000. [arXiv:hep-th/9912072](https://arxiv.org/abs/hep-th/9912072).
- [22] Victor Gayral. “Heat-kernel approach to UV/IR mixing on isospectral deformation manifolds”. *Annales Henri Poincaré*, **6**:991–1023, 2005. [arXiv:hep-th/0412233](https://arxiv.org/abs/hep-th/0412233).
- [23] K. G. Wilson and John B. Kogut. “The renormalization group and the epsilon expansion”. *Phys. Rept.*, **12** (2):75–200, 1974.
- [24] Barry Simon. *Functionnal Integration and Quantum Physics*, volume 86 of *Pure and applied mathematics*. Academic Press, New York, 1979.
- [25] Harald Grosse and Raimar Wulkenhaar. “Renormalisation of ϕ^4 -theory on noncommutative \mathbb{R}^2 in the matrix base”. *JHEP*, **12**:019, 2003. [arXiv:hep-th/0307017](https://arxiv.org/abs/hep-th/0307017).
- [26] H. Grosse and H. Steinacker. “Renormalization of the noncommutative ϕ^3 model through the kontsevich model”. *Nucl. Phys. B*, **746**:202–226, 2006a. [arXiv:hep-th/0512203](https://arxiv.org/abs/hep-th/0512203).
- [27] Harald Grosse and Harold Steinacker. “A nontrivial solvable noncommutative ϕ^3 model in 4 dimensions”. *JHEP*, **0608**:008, 2006b. [arXiv:hep-th/0603052](https://arxiv.org/abs/hep-th/0603052).
- [28] E. Langmann, R. J. Szabo, and K. Zarembo. “Exact solution of noncommutative field theory in background magnetic fields”. *Phys. Lett.*, **B569**:95–101, 2003. [arXiv:hep-th/0303082](https://arxiv.org/abs/hep-th/0303082).
- [29] Edwin Langmann. “Interacting fermions on noncommutative spaces: Exactly solvable quantum field theories in $2n+1$ dimensions”. *Nucl. Phys.*, **B654**:404–426, 2003. [arXiv:hep-th/0205287](https://arxiv.org/abs/hep-th/0205287).
- [30] Razvan Gurau and Vincent Rivasseau. “Parametric representation of non-commutative field theory”. *Commun. Math. Phys.*, **272**:811, 2007. [arXiv:math-ph/0606030](https://arxiv.org/abs/math-ph/0606030).
- [31] Vincent Rivasseau and Fabien Vignes-Tourneret. “Renormalisation of non-commutative field theories”. In *Renormalisation et théories de Galois*. Luminy, March 2006. [arXiv:hep-th/0702068](https://arxiv.org/abs/hep-th/0702068).
- [32] Adrian Tanasa and Fabien Vignes-Tourneret. “Hopf Algebra of Non-commutative Field Theory”. *J. Noncommut. Geom.*, **2** (1):125–139, 2008. [arXiv:0707.4143](https://arxiv.org/abs/0707.4143).
- [33] H. O. Girotti, M. Gomes, A. Y. Petrov, V. O. Rivelles, and A. J. da Silva. “The three-dimensional noncommutative nonlinear sigma model in superspace”. *Physics Letters B*, **521**:119–126, November 2001. [arXiv:hep-th/0109222](https://arxiv.org/abs/hep-th/0109222).
- [34] H. O. Girotti, M. Gomes, V. O. Rivelles, and A. J. da Silva. “A consistent noncommutative field theory: the Wess-Zumino model”. *Nuclear Physics B*, **587**:299–310, October 2000. [arXiv:hep-th/0005272](https://arxiv.org/abs/hep-th/0005272).
- [35] Harald Grosse and Harold Steinacker. “Exact renormalization of a noncommutative ϕ^3 model in 6 dimensions”. *Adv. Theor. Math. Phys.*, 2008. [arXiv:hep-th/0607235](https://arxiv.org/abs/hep-th/0607235).
- [36] P. K. Mitter and P. H. Weisz. “Asymptotic scale invariance in a massive Thirring model with $U(n)$ symmetry”. *Phys. Rev.*, **D8**:4410–4429, 1973.
- [37] David J. Gross and Andre Neveu. “Dynamical symmetry breaking in asymptotically free field theories”. *Phys. Rev.*, **D10**:3235, 1974.
- [38] Christoph Kopper, J. Magnen, and V. Rivasseau. “Mass generation in the large N Gross-Neveu model”. *Commun. Math. Phys.*, **169**:121–180, 1995.
- [39] K. Gawedzki and A. Kupiainen. “Gross-Neveu model through convergent perturbation expansions”. *Commun. Math. Phys.*, **102** (1):1–30, 1986.
- [40] Joel Feldman, Jacques Magnen, Vincent Rivasseau, and R. Sénéor. “A Renormalizable Field Theory: The Massive Gross-Neveu Model in two Dimensions”. *Commun. Math. Phys.*, **103** (1):67–103, 1986.
- [41] Vincent Rivasseau and Adrian Tanasa. “Parametric representation of “critical” noncommutative QFT models”. *Commun. Math. Phys.*, **279**:355, 2008. [arXiv:hep-th/0701034](https://arxiv.org/abs/hep-th/0701034).

- [42] Harald Grosse and Raimar Wulkenhaar. “The beta-function in duality-covariant noncommutative ϕ^4 -theory”. *Eur. Phys. J.*, **C35**:277–282, 2004. [arXiv:hep-th/0402093](https://arxiv.org/abs/hep-th/0402093).
- [43] Margherita Disertori, Razvan Gurau, Jacques Magnen, and Vincent Rivasseau. “Vanishing of Beta Function of Non Commutative Φ_4^4 Theory to all orders”. *Phys. Lett. B*, **649** (1):95–102, May 2007. [arXiv:hep-th/0612251](https://arxiv.org/abs/hep-th/0612251).
- [44] Ahmed Lakhouda, Fabien Vignes-Tourneret, and Jean-Christophe Wallet. “One-loop Beta Functions for the Orientable Non-commutative Gross-Neveu Model”. *Eur. Phys. J. C*, **52**:735–742, 2007. [arXiv:hep-th/0701170](https://arxiv.org/abs/hep-th/0701170).
- [45] Fabien Vignes-Tourneret. “Non-commutative field theory under magnetic field”. Work in progress, 2008.
- [46] Fabien Vignes-Tourneret. Work in progress.

This page intentionally left blank

Titles in This Series

- 539 **Kurusch Ebrahimi-Fard, Matilde Marcolli, and Walter D. van Suijlekom, Editors**, Combinatorics and physics, 2011
- 538 **José Ignacio Cogolludo-Agustín and Eriko Hironaka, Editors**, Topology of algebraic varieties and singularities, 2011
- 537 **César Polcino Milies, Editor**, Groups, algebras and applications, 2011
- 536 **Kazem Mahdavi, Debbie Koslover, and Leonard L. Brown III, Editors**, Cross disciplinary advances in quantum computing, 2011
- 535 **Maxim Braverman, Leonid Friedlander, Thomas Kappeler, Peter Kuchment, Peter Topalov, and Jonathan Weitsman, Editors**, Spectral theory and geometric analysis, 2011
- 534 **Pere Ara, Fernando Lledó, and Francesc Perera, Editors**, Aspects of operator algebras and applications, 2011
- 533 **L. Babinkostova, A. E. Caicedo, S. Geschke, and M. Scheepers, Editors**, Set theory and its applications, 2011
- 532 **Sergiy Kolyada, Yuri Manin, Martin Möller, Pieter Moree, and Thomas Ward, Editors**, Dynamical numbers: Interplay between dynamical systems and number theory, 2010
- 531 **Richard A. Brualdi, Samad Hedayat, Hadi Kharaghani, Gholamreza B. Khosrovshahi, and Shahriar Shahriari, Editors**, Combinatorics and graphs, 2010
- 530 **Vitaly Bergelson, Andreas Blass, Mauro Di Nasso, and Renling Jin, Editors**, Ultrafilters across Mathematics, 2010
- 529 **Robert Sims and Daniel Ueltschi, Editors**, Entropy and the Quantum, 2010
- 528 **Alberto Farina and Enrico Valdinoci, Editors**, Symmetry for Elliptic PDEs, 2010
- 527 **Ricardo Castaño-Bernard, Yan Soibelman, and Ilia Zharkov, Editors**, Mirror symmetry and tropical geometry, 2010
- 526 **Helge Holden and Kenneth H. Karlsen, Editors**, Nonlinear partial differential equations and hyperbolic wave phenomena, 2010
- 525 **Manuel D. Contreras and Santiago Díaz-Madrigal, Editors**, Five lectures in complex analysis, 2010
- 524 **Mark L. Lewis, Gabriel Navarro, Donald S. Passman, and Thomas R. Wolf, Editors**, Character theory of finite groups, 2010
- 523 **Aiden A. Bruen and David L. Wehlau, Editors**, Error-correcting codes, finite geometries and cryptography, 2010
- 522 **Oscar García-Prada, Peter E. Newstead, Luis Álvarez-Cónsul, Indranil Biswas, Steven B. Bradlow, and Tomás L. Gómez, Editors**, Vector bundles and complex geometry, 2010
- 521 **David Kohel and Robert Rolland, Editors**, Arithmetic, geometry, cryptography and coding theory 2009, 2010
- 520 **Manuel E. Lladser, Robert S. Maier, Marni Mishna, and Andrew Rechnitzer, Editors**, Algorithmic probability and combinatorics, 2010
- 519 **Yves Félix, Gregory Lupton, and Samuel B. Smith, Editors**, Homotopy theory of function spaces and related topics, 2010
- 518 **Gary McGuire, Gary L. Mullen, Daniel Panario, and Igor E. Shparlinski, Editors**, Finite fields: Theory and applications, 2010
- 517 **Tewodros Amdeberhan, Luis A. Medina, and Victor H. Moll, Editors**, Gems in experimental mathematics, 2010
- 516 **Marlos A.G. Viana and Henry P. Wynn, Editors**, Algebraic methods in statistics and probability II, 2010
- 515 **Santiago Carrillo Menéndez and José Luis Fernández Pérez, Editors**, Mathematics in finance, 2010
- 514 **Arie Leizarowitz, Boris S. Mordukhovich, Itai Shafrir, and Alexander J. Zaslavski, Editors**, Nonlinear analysis and optimization II, 2010

TITLES IN THIS SERIES

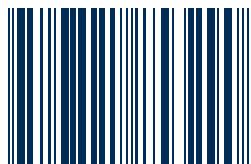
- 513 **Arie Leizarowitz, Boris S. Mordukhovich, Itai Shafrir, and Alexander J. Zaslavski, Editors**, Nonlinear analysis and optimization I, 2010
- 512 **Albert Fathi, Yong-Geun Oh, and Claude Viterbo, Editors**, Symplectic topology and measure preserving dynamical systems, 2010
- 511 **Luise-Charlotte Kappe, Arturo Magidin, and Robert Fitzgerald Morse, Editors**, Computational group theory and the theory of groups, II, 2010
- 510 **Mario Bonk, Jane Gilman, Howard Masur, Yair Minsky, and Michael Wolf, Editors**, In the Tradition of Ahlfors-Bers, V, 2010
- 509 **Primitivo B. Acosta-Humánez and Francisco Marcellán, Editors**, Differential algebra, complex analysis and orthogonal polynomials, 2010
- 508 **Martin Berz and Khodr Shamseddine, Editors**, Advances in p -Adic and non-archimedean analysis, 2010
- 507 **Jorge Arvesú, Francisco Marcellán, and Andrei Martínez-Finkelshtein, Editors**, Recent trends in orthogonal polynomials and approximation theory, 2010
- 506 **Yun Gao, Naihuan Jing, Michael Lau, and Kailash C. Misra, Editors**, Quantum affine algebras, extended affine Lie algebras, and their applications, 2010
- 505 **Patricio Cifuentes, José García-Cuerva, Gustavo Garrigós, Eugenio Hernández, José María Martell, Javier Parcet, Alberto Ruiz, Fernández Soria, José Luis Torrea, and Ana Vargas, Editors**, Harmonic analysis and partial differential equations, 2010
- 504 **Christian Ausoni, Kathryn Hess, and Jérôme Scherer, Editors**, Alpine perspectives on algebraic topology, 2009
- 503 **Marcel de Jeu, Sergei Silvestrov, Christian Skau, and Jun Tomiyama, Editors**, Operator structures and dynamical systems, 2009
- 502 **Viviana Ene and Ezra Miller, Editors**, Combinatorial Aspects of Commutative Algebra, 2009
- 501 **Karel Dekimpe, Paul Igodt, and Alain Valette, Editors**, Discrete groups and geometric structures, 2009
- 500 **Philippe Briet, François Germinet, and Georgi Raikov, Editors**, Spectral and scattering theory for quantum magnetic systems, 2009
- 499 **Antonio Giambruno, César Polcino Milies, and Sudarshan K. Sehgal, Editors**, Groups, rings and group rings, 2009
- 498 **Nicolau C. Saldanha, Lawrence Conlon, Rémi Langevin, Takashi Tsuboi, and Paweł Walczak, Editors**, Foliations, geometry and topology, 2009
- 497 **Maarten Bergvelt, Gaywalee Yamskulna, and Wenhua Zhao, Editors**, Vertex operator algebras and related areas, 2009
- 496 **Daniel J. Bates, GianMario Besana, Sandra Di Rocco, and Charles W. Wampler, Editors**, Interactions of classical and numerical algebraic geometry, 2009
- 495 **G. L. Litvinov and S. N. Sergeev, Editors**, Tropical and idempotent mathematics, 2009
- 494 **Habib Ammari and Hyeonbae Kang, Editors**, Imaging microstructures: Mathematical and computational challenges, 2009
- 493 **Ricardo Baeza, Wai Kiu Chan, Detlev W. Hoffmann, and Rainer Schulze-Pillot, Editors**, Quadratic Forms—Algebra, Arithmetic, and Geometry, 2009
- 492 **Fernando Giráldez and Miguel A. Herrero, Editors**, Mathematics, Developmental Biology and Tumour Growth, 2009
- 491 **Carolyn S. Gordon, Juan Tirao, Jorge A. Vargas, and Joseph A. Wolf, Editors**, New developments in Lie theory and geometry, 2009
- 490 **Donald Babbitt, Vyjayanthi Chari, and Rita Fioresi, Editors**, Symmetry in mathematics and physics, 2009

This book is based on the mini-workshop Renormalization, held in December 2006, and the conference Combinatorics and Physics, held in March 2007. Both meetings took place at the Max-Planck-Institut für Mathematik in Bonn, Germany.

Research papers in the volume provide an overview of applications of combinatorics to various problems, such as applications to Hopf algebras, techniques to renormalization problems in quantum field theory, as well as combinatorial problems appearing in the context of the numerical integration of dynamical systems, in noncommutative geometry and in quantum gravity.

In addition, it contains several introductory notes on renormalization Hopf algebras, Wilsonian renormalization and motives.

ISBN 978-0-8218-5329-0



9 780821 853290

CONM/539

AMS on the Web
www.ams.org