**CENTRE DE PHYSIQUE THEORIQUE**
**CNRS - Luminy, Case 907**
**13288 Marseille Cedex 9**

# ON KREIMER'S HOPF ALGEBRA STRUCTURE OF FEYNMAN GRAPHS

**Thomas KRAJEWSKI*** and **Raimar WULKENHAAR**†‡

May 16, 1998



**Abstract**

We reinvestigate Kreimer's Hopf algebra structure of perturbative quantum field theories. In Kreimer's original work, overlapping divergences were first disentangled into a linear combination of disjoint and nested ones using the Schwinger-Dyson equation. The linear combination then was tackled by the Hopf algebra operations. We present a formulation where the coproduct itself produces the linear combination, without reference to external input.

PACS-98: 02.10.Sp Linear and multilinear algebra, 11.10.Gh Renormalization, 11.15.Bt General properties of perturbation theory

## 1 Introduction

This paper is the result of our efforts to understand the article by Dirk Kreimer on the Hopf algebra structure of perturbative quantum field theories [1]. That article was brought to our attention by Alain Connes in his talk during the Vietri conference on noncommutative geometry. Kreimer discovered that divergent Feynman graphs can be understood as elements of a Hopf algebra. The forest formula guiding the renormalization of Feynman graphs with subdivergences is reproduced by a certain interplay of product, coproduct, antipode and counit of that Hopf algebra.

We noticed that in all examples in the general part of [1], overlapping divergences – the target of the forest formula – never occurred. Thanks to a hint by Kreimer we understand now that overlapping divergences must first be disentangled into a linear combination of terms containing disjoint or nested divergences exclusively. The tool used in this procedure is the Schwinger-Dyson equation. It was shown in [2] that disentangling overlapping divergences is always possible. It is therefore no restriction that the operations of the Hopf algebra are applied only to terms containing no overlapping divergences.

*and Université de Provence and Ecole Normale Superieure de Lyon, `tkrajews@cpt.univ-mrs.fr`

†and Universität Leipzig, `raimar@cpt.univ-mrs.fr`

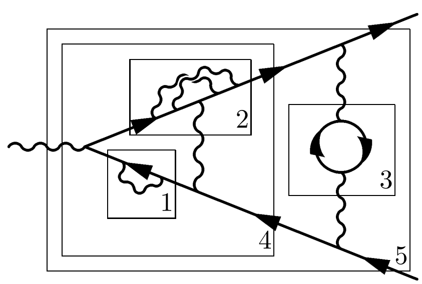
‡supported by the German Academic Exchange Service (DAAD), grant no. D/97/20386.

In this paper we present our independent approach to the problem of overlapping divergences. Our goal is to treat overlapping divergences on the same footing with disjoint and nested ones. We wish the operations of the Hopf algebra themselves to disentangle overlapping divergences, without reference to exterior input like the Schwinger-Dyson equation. We show that this aim can be achieved by endowing Kreimer's parenthesized words (PW) describing the Feynman graphs with additional information. In our formulation, a PW is a collection of all maximal forests of a Feynman graph, where identical regions in various forests are visualized. The Hopf algebra given by the set of all such extended PWs is always coassociative and has always a left counit and left antipode. It has a right counit and right antipode for certain renormalization schemes. The axiom for the left antipode yields the forest formula for any Feynman graph.

We introduce in section 2 our extended PWs and discuss in section 4 the $R$-operation of renormalization. The Hopf algebra is identified in section 5, where longer proofs are delegated to the appendix. In sections 3 and 6 we apply our methods to examples of Feynman graphs with overlapping divergences.

## 2 Feynman graphs, maximal forests and parenthesized words

Let $\Gamma$ be a Feynman graph. In the way described by Kreimer we draw boxes around superficially (UV-) divergent sectors of $\Gamma$:




$$= (((s_1)(v_2)v_4)(p_3)v_5) \qquad (1)$$

(As usual, straight lines stand for fermions and wavy lines for bosons.) A superficially divergent sector [3] is necessarily a region of $\Gamma$ which contains loops. The boxes must be drawn in such a way that no vertex of $\Gamma$ is situated on the border of the box and no line of $\Gamma$ is tangential to the border. Boxes can be deformed. During the deformation, no vertex is allowed to pass the border and at no time a line may be tangent to the border of the box. We consider boxes which differ by a deformation as identical.

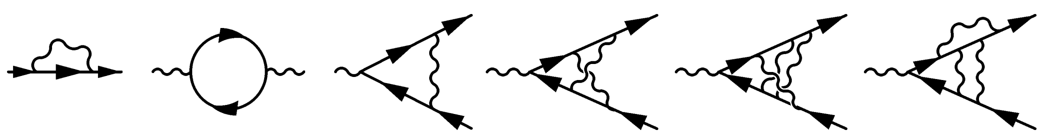
We shall work in four dimensional spacetime, but generalization is obvious. We are interested in the case typical for renormalizable gauge field theories where for each vertex $V$ of $\Gamma$ the momentum dimension of the fields meeting at $V$ is equal to the spacetime dimension (= 4). A criterium for superficial divergence of a region confined in a box is power counting. The box under consideration will contain $n_B$ bosonic and $n_F$ fermionic external legs. Ghosts are regarded as bosons here. There can only be a superficial (ultraviolet) divergence in the box if it contains at least one loop and if the power counting degree of divergence $d_{pc}$ satisfies

$$d_{pc} := 4 - n_B - \tfrac{3}{2} n_F \geq 0 \ . \qquad (2)$$

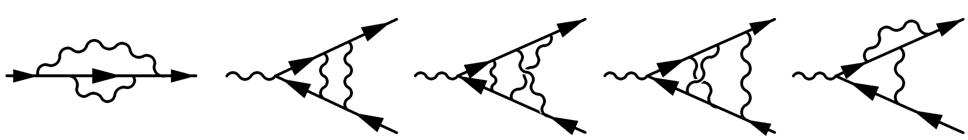
Owing to symmetries the actual degree of divergence $d$ of one graph or a sum of graphs can be lower than $d_{pc}$ calculated from (2), see ref. [3]. Examples are graphs in QED with $n_B = 3$, $n_F = 0$ (which can be omitted due to Furry's theorem) and with $n_B = 4$, $n_F = 0$ (which are superficially convergent due to gauge symmetry). Always if $d < 0$ the box must be erased. This does not mean that there cannot be divergences in the box to erase. But these non-superficial divergences must be contained in other boxes which cannot be deformed into the box we erased.

Our boxes represent the forest structure of $\Gamma$. A forest is a set of 1PI (one-particle-irreducible, i.e. the graph remains connected after cutting an arbitrary line) divergent subgraphs $\gamma \subset \Gamma$ which do not overlap. Instead, any two elements (= boxes) of a forest are either disjoint or nested. The forest structure is the collection of the maximal forests of $\Gamma$, i.e. the forests which are not contained in another forest. There are several maximal forests in general to a Feynman graph.

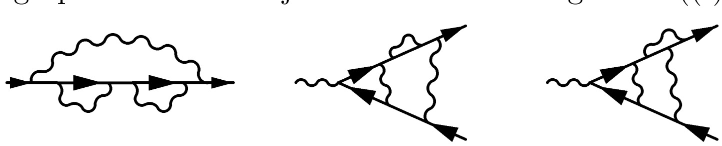
Kreimer defines [1] a recursive procedure to assign parenthesized words (PW) to the boxes of a maximal forest. The total graph $\Gamma$ stands for a certain integrand $I_\Gamma$ depending on external and internal momenta. A box is represented by a pair of opening-closing parentheses. Two nested boxes are represented by (( ) ) and two disjoint boxes by ( )( ). In an irreducible PW (iPW) the leftmost opening parenthesis matches its rightmost closing parenthesis. A primitive box contains no nested boxes and represents a graph $\gamma$ without subdivergences. Examples of primitive boxes ( ) are:


(3)

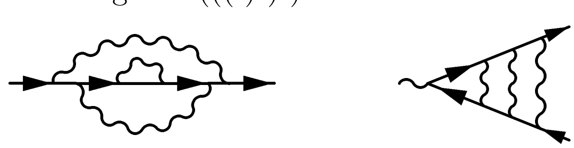
(The reader is encouraged to verify using (2) that the last three examples contain no divergent subgraphs.) We associate the integrand $I_\gamma$ defined by the vertices and propagators of $\gamma$ to such a primitive box and write the PW $(I_\gamma)$. A non-primitive box contains nested boxes. It describes a graph $\gamma$ with subdivergences $\gamma_i$, which are already characterized by PWs $X_i$. Examples for graphs with one nested subdivergence (( ) ) are:


(4)

Examples for graphs with two disjoint nested subdivergences: (( )( ) ) are:

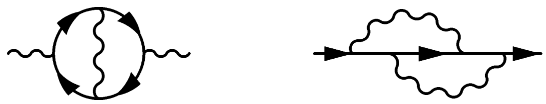

(5)

And here are two examples for graphs with a nested subdivergence which has itself a nested subsubdivergence ((( ) ) ):


(6)

If we shrink all nested boxes (=divergent subgraphs $\gamma_i$) of $\gamma$ to points, there remains a fraction $I_{\gamma/\cup\gamma_i}$ of the integrand of $\gamma$ defined by the vertices and propagators of $\gamma/\cup\gamma_i$. We write this fraction next to the right closing parenthesis and everything we have shrunk to a point (the $X_i$) between that fraction and the left opening parenthesis. The resulting PW looks like this: $\big(X_1 \ldots X_n\ I_{\gamma/\{\gamma_1\cup\cdots\cup\gamma_n\}}\big)$. Note that the order of disjoint boxes is irrelevant.

By this procedure we associate a PW to each maximal forest. As discovered by Kreimer [1], the PWs form a Hopf algebra whose antipode axiom reproduces the forest formula [4]. This assumes that overlapping divergences such as



(7)

have been disentangled into a linear combination of PWs containing disjoint and nested divergences exclusively. That procedure is an *external* one as it uses the Schwinger-Dyson equation [2]. The outcome is thus a linear combination of PWs each of them describing precisely one maximal forest.

The goal of this paper is to modify the PWs and the Hopf algebra operations in such a way that any 1PI-Feynman graph is described by *a single* PW and that all Hopf algebra operations are defined on such a PW. Our starting point is the observation that in the case of overlapping divergences there exist several maximal forests to a Feynman graph. It is clear that democracy requires to comprise all PWs associated to these maximal forests to one bigger object. We propose to build a column vector whose components are the PWs of maximal forests. The order of the components (or rows as they are long objects) of this vector is not relevant, of course. As the integrands associated to the PWs of each row are equal (up to cyclic permutations), we associate this universal integrand to our column vector.

There is one further modification necessary. Later on we are going to identify the subwords of such a vector and define the removal of subwords. Subwords represent subgraphs and the removal means shrinking the subgraphs to points. But subgraphs or subwords can occur identically in various maximal forests. If we now compare the maximal forests of a graph with removed subgraph and the maximal forests of the original graph, it is easy to see that the subgraph is removed in all maximal forest it had occurred. (An example is the step from (10) to (8) in the next section by shrinking the loop 3.) We must implement this feature in our vectors. We propose to connect by a tree of lines the closing parentheses of identical and simultaneously shrinkable boxes. If we pull out a subword of such a vector and if the subword is connected over various rows, we simply have to remove all of them.

Thus, our PWs are vectors of one-line-PWs representing the maximal forests of a Feynman graph, where the closing parentheses of simultaneously shrinkable boxes are connected. We define now the notion of a parenthesized subword (PSW) of a PW. A PSW $Y$ of $X$ is everything between a set of connected closing parentheses and its matching opening parentheses. Disconnected rows of $X$

which are accidentally between connected rows are not part of the PSW $Y$ under consideration.

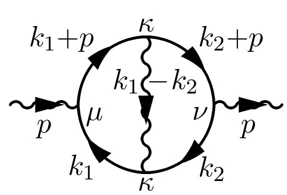
There is an algorithm which yields all PSW of a PW. Starting with the first row we run from the left through the PW until we meet a closing parenthesis. In general, it will be connected with other closing parentheses in different rows. These connected closing parentheses and their matching opening parentheses define our first PSW. We mark all these connected closing parentheses. We then go ahead and move through the first row until we arrive at the next closing parenthesis. This gives the next PSW and marks the next set of parentheses. We repeat this procedure until the rightmost closing parenthesis is reached. Then we pass to the second row and continue to search for new closing parentheses and related PSW, i.e. we ignore all parentheses marked in the previous steps. This search continues through all rows and stops at the lower right corner of our PW.

In what follows we will freely use the notions parenthesized word (PW), irreducible PW (iPW, the leftmost and rightmost parentheses match), primitive PW (no nested divergences, a special iPW) and parenthesized subword (PSW, a special iPW). We remark that a possible extension could be the inclusion of superficially convergent 1PI-graphs ($d < 0$) with subdivergences. All finite integrands fuse and stand immediately before the rightmost closing parentheses.

We will give now some examples for Feynman graphs with overlapping divergences which are represented by parenthesized words of several maximal forests. The PSW of some of these examples are discussed and further evaluated in section 6.

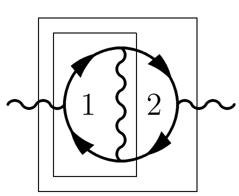
## 3 Examples for Feynman graphs with several maximal forests

In QED there is the following contribution to the photon propagator:

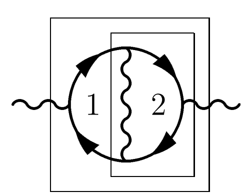



$$\begin{array}{l} ((v_1)p_2) \\ ((v_2)p_1) \end{array} \qquad (8)$$

We can draw two maximal forests of boxes. We can first draw a box around the left loop which contains the vertex correction with interior momentum $k_1$. Then we put this box into the large box which encircles both loops. Or we can first enclose the right loop by a vertex box and then put everything into *the same* large box. Graphically, the two possibilities look like this:



$= ((v_1)p_2)$ or



$= ((v_2)p_1)$ . (9)

In the first case, the innermost box is the primitive box $(v_1)$ the integrand of which is in the Feynman gauge given by

$$v_1 = e\gamma^\kappa \frac{1}{\not k_1 - \mu} e\gamma^\mu \frac{1}{\not k_1 + \not p - \mu} e\gamma_\kappa \frac{1}{(k_1-k_2)^2 - M^2} \ .$$

Here, $e$ is the electron charge, $\mu$ is the electron mass and $M$ an auxiliary photon mass to avoid IR-divergences. This vertex box is nested in the large box, so we must write $((v_1)p_2)$ as the maximal forest. The integrand $p_2$ is the interior of the large box after shrinking the small box $(v_1)$ to a point. What remains is loop 2 and the integrand is found to be

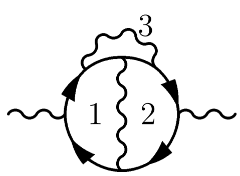
$$p_2 = \frac{1}{\not k_2 + \not p - \mu} e\gamma^\nu \frac{1}{\not k_2 - \mu} \ .$$

In the second case the loops 1 and 2 change their role and we obtain the maximal forest $((v_2)p_1)$ with

$$v_2 = e\gamma^\kappa \frac{1}{\not k_2 + \not p - \mu} e\gamma^\nu \frac{1}{\not k_2 - \mu} e\gamma_\kappa \frac{1}{(k_2 - k_1)^2 - M^2} \ , \qquad p_1 = \frac{1}{\not k_1 - \mu} e\gamma^\mu \frac{1}{\not k_1 + \not p - \mu} \ .$$

We have found two maximal forests $((v_1)p_2)$ and $((v_2)p_1)$ in this example. These two forests form the 2-line vector $\binom{((v_1)p_2)}{((v_2)p_1)}$. However, the large box occurs identically in both maximal forests. We cannot shrink it in one of them and keep it in the other. Therefore, the closing parentheses representing the large box in both rows of the vector must be connected, as we have already indicated in (8).

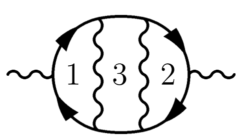
Here is a graph with two maximal forests containing a nested divergence:




$$\begin{array}{l} (((v_3)v_{13})p_2) \\ (((v_3)v_{23})p_1) \end{array} \qquad (10)$$

The vertex correction $v_3$ is nested in both vertex corrections $v_{i3}$ comprising the common loop 3 and loop $i$. The subword $(v_3)$ is identical in both maximal forests $(((v_3)v_{13})p_2)$ and $(((v_3)(v_{23})p_1)$. If we shrink it in one of them it is automatically removed in the other one. For the same reasons both maximal forests are connected at the outermost box.

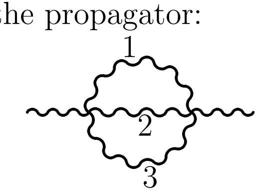
Here is now a more complicated forest structure:



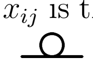

$$\begin{array}{llll} ( & (v_1) & (v_2) & p_3) \\ ((( & v_1) & & v_{13})p_2) \\ (( & & (v_2) & v_{23})p_1) \end{array} \qquad (11)$$

We have three possibilities for drawing disjoint boxes: We can take loops 1 and 2 and put them into the large box, or we can put loop 1 into the vertex box which covers loops 1 and 3 and then everything into the large box, or we can exchange the role of loops 1 and 2.

Let us also give an example from $\phi^4$-theory. There is the following second-order correction to the propagator:




$$\begin{array}{l} ((x_{23})y_1) \\ ((x_{31})y_2) \\ ((x_{12})y_3) \end{array} \qquad (12)$$

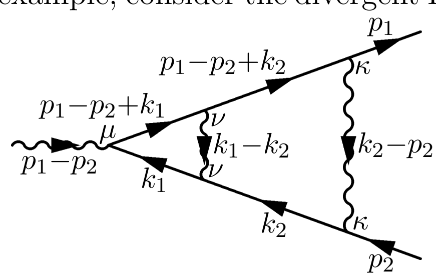
Here, $x_{ij}$ is the vertex correction  involving the lines $i, j$ and $y_k$ the tadpole graph  involving the line $k$. The three maximal forests are connected because shrinking one of them to a point forces the reduction of the other two.

## 4 Kreimer's $R$-operation [1]

To any PW $X$, Kreimer associates a second, in a certain sense equivalent copy $R[X]$. The philosophy is that $R[X]$ is a local counterterm, a point-like interaction. It is so to say a new vertex, mass or kinetic term in the Lagrangian, which itself is infinite but such that a certain combination of counterterms and divergent 1PI graphs is finite. The finite linear combination in question is given by the forest formula or – as discovered by Kreimer – by the antipode axiom of a (quasi-) Hopf algebra to construct. For renormalizability it is essential that all counterterms can be absorbed by a redefinition of physical parameters of the theory. In particular in gauge theories there are potentially more types of counterterms than physical parameters [3]. It is important then that counterterms and divergences of the sum of all graphs contributing to a certain amplitude cancel. We avoid a discussion of these subtleties by considering scalar theories or – with some care – QED.

The $R$-operation depends on the renormalization scheme, which in principle is arbitrary but fixed throughout the investigation. We shall work in the BPHZ scheme [5, 6, 4] which is the standard one in connection with the forest formula. A iPW $X$ represents one box containing a divergent Feynman graph with in general several forests of subdivergences. The box has $n_B$ bosonic and $n_F$ fermionic external legs. The superficial degree of divergence $d[X]$ of the iPW $X$ is bounded by the power counting theorem (2), $d[X] \leq 4 - n_B - \frac{3}{2} n_F$. In the BPHZ scheme the integrand $R[X]$ is the Taylor expansion until order $d[X]$ with respect to the external momenta of $X$. This implies that $R[X]$ and $X$ have the same asymptotic dependence of all internal momenta of $X$, this is Kreimer's equivalence relation $X \sim R[X]$. But this does *not* mean that $X - R[X]$ is an integrand yielding a finite integral. This is only the case if $X$ is a primitive PW without subdivergences.

To give an example, consider the divergent Feynman graph with subdivergence




$$= ((v_1)v_2)\;, \qquad (13)$$

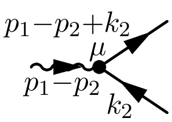
$$v_1 = e\gamma^\nu \frac{1}{\not{k}_1 - \mu} e\gamma^\mu \frac{1}{\not{k}_1 + (\not{p}_1 - \not{p}_2 + \not{k}_2) - \not{k}_2 - \mu} e\gamma_\nu \frac{1}{(k_1 - k_2)^2 - M^2}\;,$$

$$v_2 = \frac{1}{\not{p}_1 - \not{p}_2 + \not{k}_2 - \mu} e\gamma^\kappa \frac{1}{(k_2 - p_2)^2 - M^2} e\gamma_\kappa \frac{1}{\not{k}_2 - \mu}\;.$$

We have written $v_1$ in a form where its external momenta $p_1 - p_2 + k_2$ and $k_2$ are explicit. The two subwords of $((v_1)v_2)$ are clearly $(v_1)$ and $((v_1)v_2)$. Let us compute $R[(v_1)]$. It has 2 fermionic and 1 bosonic external legs, hence $d[(v_1)] \leq 0$, and actually $d[(v_1)] = 0$. In the BPHZ scheme we take the Taylor expansion of $(v_1)$ in its external momenta $p_1 - p_2 + k_2$ and $k_2$ until order 0. This gives

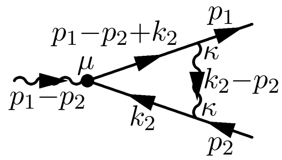
$$R[(v_1)] = (v_1)|_{p_1 - p_2 + k_2 = k_2 = 0} = e\gamma^\nu \frac{1}{\not{k}_1 - \mu} e\gamma^\mu \frac{1}{\not{k}_1 - \mu} e\gamma_\nu \frac{1}{k_1^2 - M^2}$$

$$=$$




(14a)

$$(R[(v_1)]v_2) = e\gamma^\nu \tfrac{1}{\not k_1-\mu} e\gamma^\mu \tfrac{1}{\not k_1-\mu} e\gamma_\nu \tfrac{1}{k_1^2-M^2} \tfrac{1}{\not p_1-\not p_2+\not k_2-\mu} e\gamma^\kappa \tfrac{1}{(k_2-p_2)^2-M^2} e\gamma_\kappa \tfrac{1}{\not k_2-\mu}$$

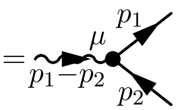
$=$ 

(14b)

The asymptotic behavior of $(v_1)$ and $R[(v_1)]$ with respect to $k_1$ is identical, this is symbolized by the equivalence relation $(v_1) \sim R[(v_1)]$. We also see that $R[(v_1)]$ defines a local counterterm. The integral $\int d^4k_1 \ \mathrm{tr}\{(v_1) - R[(v_1)]\}$ is finite.

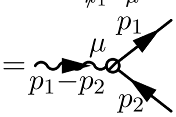
We can now apply the $R$-operation to the PWs $(R[(v_1)]v_2)$ and $((v_1)v_2)$, which both have 2 fermionic and 1 bosonic external legs and $d[((v_1)v_2)] = d[(R[(v_1)]v_2)] = 0$. We have to take the Taylor expansion of these PWs in their external momenta $p_1$ and $p_2$ until order 0, which gives

$$R[(R[(v_1)]v_2)] = e\gamma^\nu \tfrac{1}{\not k_1-\mu} e\gamma^\mu \tfrac{1}{\not k_1-\mu} e\gamma_\nu \tfrac{1}{k_1^2-M^2} \tfrac{1}{\not k_2-\mu} e\gamma^\kappa \tfrac{1}{k_2^2-M^2} e\gamma_\kappa \tfrac{1}{\not k_2-\mu}$$

$=$ 

(15a)

$$R[((v_1)v_2)] = e\gamma^\nu \tfrac{1}{\not k_1-\mu} e\gamma^\mu \tfrac{1}{\not k_1-\mu} e\gamma_\nu \tfrac{1}{(k_1-k_2)^2-M^2} \tfrac{1}{\not k_2-\mu} e\gamma^\kappa \tfrac{1}{k_2^2-M^2} e\gamma_\kappa \tfrac{1}{\not k_2-\mu}$$

$=$ 

(15b)

Observe that $R[(R[(v_1)]v_2)]$ and $(R[(v_1)]v_2)$ have the same asymptotic behavior with respect to $k_1$ and $k_2$, the same is true for the pair $((v_1)v_2)$ and $R[((v_1)v_2)]$. Both $R[(R[(v_1)]v_2)]$ and $R[((v_1)v_2)]$ define local counterterms, but both integrals $\int d^4k_2 d^4k_1 \ \mathrm{tr}\{((v_1)v_2) - R[((v_1)v_2)]\}$ and $\int d^4k_2 d^4k_1 \ \mathrm{tr}\{((v_1)v_2) - R[(R[(v_1)]v_2)]\}$ are *infinite*. To obtain a finite expression one has to include $(R[(v_1)]v_2)$ in a way given by the forest formula.

We must say a few words how equivalence is defined quantitatively. Renormalization schemes depend on some regularization parameter $\epsilon$. Infinities correspond to pole terms in $\epsilon$. In terms of $\epsilon$, Kreimer gives the following definition of equivalence:

$$X \sim Y \qquad \text{iff} \qquad \lim_{\hbar\to 0, \epsilon\to 0} \{X - Y\} = 0 \ . \tag{16}$$

The equivalence $R[X] \sim X$ has to be regarded in this sense. It is important to understand that $R[X] \sim X$ does not imply $R[X]Y \sim XY$. The reason is that if $Y$ has pole terms in $\epsilon$ then in the product $(R[X] - X)Y$ also terms of order $\epsilon$ in $R[X] - Y$ become essential. It turns out that the full set of properties of a Hopf algebra can only be guaranteed if equivalence works for products, in a certain sense. The precise condition to the the renormalization map $R$ is

$$R\big[\prod_i R[X_i] \ \prod_j Y_j\big] = \prod_i R[X_i] \ \prod_j R[Y_j] \ . \tag{17}$$

We indicate by $X \approx Y$ that under the condition (17) we have $X \sim Y$, but that in general equivalence is not guaranteed.

In the BPHZ scheme there is no regularization parameter $\epsilon$, so we cannot use the definition (16). Nevertheless, $R$ is defined for any Feynman graph, and we say that $X \sim Y$ iff $Y = X$ or $Y = R[X]$. The condition (17) makes sense, and we have $R^2 = R$ by construction. We remark that superficially convergent graphs with subdivergences (if included, see the remark at the end of section 2) are annihilated by $R$. This is clear in the BPHZ scheme, because a Taylor expansion until order $d < 0$ makes no sense. In what follows we work on a general level without specifying the renormalization scheme and its $R$-operation.

## 5 The Hopf algebra

Following the work of Kreimer [1] we will now equip the PWs with the structure of a (left-) Hopf algebra. This goes in four steps. First, we would like to consider the set $\mathcal{A}$ of all PWs (which include from now on its $R$-equivalents) as a vector space. Stricty speaking, the product definition below will force us to introduce a refinement of PWs, and $\mathcal{A}$ is the set of such refined PWs. We enlarge formally this set $\mathcal{A}$ by all rational linear combinations of PWs. This makes $\mathcal{A}$ to a formal vector space over the field $\mathbf{Q}$ of rational numbers, $\mathbf{Q}$ just for simplicity.

The second step makes $\mathcal{A}$ to an algebra by defining a product $m$. This is an operation which assigns to a sum of pairs of elements of $\mathcal{A}$ a new one. Actually only $\mathbf{Q}$-equivalence classes of pairs are essential as $m(qX, Y) = m(X, qY) = q\, m(X, Y)$, for $X, Y \in \mathcal{A}$ and $q \in \mathbf{Q}$. Thus, $m$ operates on the tensor product. Suppose we want to multiply two iPWs $(X)$ and $(Y)$ to a new PW $Z = m[(X) \otimes (Y)]$. There are clearly many possibilities. We could write the PW $Z = (X)(Y)$ which corresponds to two disjoint divergences. Or we could insert $(X)$ into $(Y)$ and write $Z = ((X)Y)$. This corresponds to a subdivergence $X$ nested in the divergence $Y$. Or we could exchange the role of $X$ and $Y$. Kreimer always takes the disjoint product $(X)(Y)$. We believe that this is not justified. Let us look again at the example (14b), where we have computed the diagram $(R[(v_1)]v_2)$. On our way to recover the forest formula we will be forced to identify this term with $m[R[(v_1)] \otimes (v_2)]$. The integrand is by chance equal to $R[(v_1)](v_2)$. But how can we distinguish it from two disjoint vertices $R[(v_1)]$ and $(v_2)$? In writing $(R[(v_1)]v_2)$ we say unambiguously that the counterterm $R[(v_1)]$ is inserted into the vertex correction $(v_2)$. This is important because with that interpretation, formula (17) is much less restrictive. In particular, we do not need (17) for 1PI-graphs.

To cut a long story short, we are motivated to modify Kreimer's product definition. But how does the multiplication operator $m$ know whether it must insert a subword or take a disjoint product? One idea is to add more information to the tensor product. For example, if we have a tensor product of $(X_1)(X_2)(X_3)(X_4)$ with $(Y_1)(Y_2)(Y_3)$ and we want $m$ to insert $(X_2)$ and $(X_3)$ into $(Y_2)$ and $(X_4)$ into $(Y_1)$, we indicate this by the following assignment of horizontal brackets:

$$m\Big[(X_1)(X_2)(X_3)(X_4) \otimes (Y_1)(Y_2)(Y_3)\Big] = (X_1)\,((X_4)Y_1)\,((X_3)(X_2)Y_2)\,(Y_3)\ .$$

This multiplication rule also works across $R$-operations. For example, if we have a tensor product of $R[(X_1)]R[(X_2)](X_3)$ with $R[(Y_1)]((Y_2)Y_3)$ and we want $m$ to insert $R[(X_2)]$ and $(X_3)$ into $R(Y_1)]$ and $R[(X_1)]$ into the subword $(Y_2)$ of $((Y_2)Y_3)$, we indicate this by

$$m\Big[R[(X_1)]R[(X_2)](X_3)\otimes R[(Y_1)]((Y_2)Y_3)\Big]=R[((X_3)R[(X_2)]Y_1)]\,((R[(X_1)]Y_2)Y_3)\,.$$

This is easy. However, such brackets across tensor products seem to be a severe violation of algebraic principles. Actually, these horizontal brackets are a convenient visualization of very complicated index structures. By construction there is always a countable number of divergent 1PI-Feynman graphs, hence a countable number of iPWs. Thus, we can assign a number to the rightmost closing parenthesis of each iPW. We can now label the subwords of the iPW $X_n$ according to their position and their depth, for instance

$$X_n = ((((\ .\ )_{n,221}\ .\ )_{n,22}(\ .\ )_{n,21}\ .\ )_{n,2}(\ .\ )_{n,1}\ .\ )_n\ .$$

If we now pull out a subword (that operation will be rigorously defined below), say $((\ .\ )_{n,221}\ .\ )_{n,22}$, then the label structure tells us that the reminder (also characterized below) $X_n' = (((\ .\ )_{n,21}\ .\ )_{n,2}(\ .\ )_{n,1}\ .\ )_n$ is a PW where a subword is missing, just because the word with the same parenthesis arrangements as $X_n'$ would carry another label than ${}_n$, and that the missing subword carries the label ${}_{n,22}$ at its closing parenthesis. Now, there are five possibilities for the relative position of any two PWs $X_k$ and $Y_l$ occurring in the tensor product $X_k \otimes Y_l$:

1. They intersect, which means that there exists a parenthesis with the same label in $X_k$ and $Y_l$, such as $(\ .\ )_{n,221}$ and $((\ .\ )_{n,221}\ .\ )_{22}$ above.
2. They are disjoint, such as $(\ .\ )_{n,21}$ and $(\ .\ )_{n,1}$ above.
3. The closing parenthesis of $X_k$ is missing in $Y_l$ and $\mathrm{depth}(k) - \mathrm{depth}(l) = 1$, such as $(\ .\ )_{n,221}$ and $(\ .\ )_{n,22}$ above.
4. The closing parenthesis of $X_k$ is missing in $Y_l$ and $\mathrm{depth}(k) - \mathrm{depth}(l) > 1$, such as $(\ .\ )_{n,221}$ and $((\ .\ )_{n,21}\ .\ )_{n,2}$ above.
5. The closing parenthesis of $Y_l$ is missing in $X_k$.

We define $m[X_k \otimes Y_l] = 0$ in cases 1,4,5. The multiplication $m$ builds the disjoint product in case 2 and inserts the immediately missing subword $X_k$ into the correct position in $Y_l$ in case 3.

All this complicated index structure is encoded in the brackets across tensor products. The brackets are much easier to memorize, but they are completely equivalent to true tensor products and large towers of indices.

We shall define the multiplication as an operator

$$m : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$$

acting as above explained. This multiplication is noncommutative in general. If there is a horizontal bracket we always insert the left factor into the right factor, never the right into the left. If there is no bracket we build the disjoint product which is commutative. The multiplication $m$ is always associative,

$$m \circ (\mathrm{id} \otimes m) = m \circ (m \otimes \mathrm{id})\ .$$

This is clear for disjoint products, but also for brackets: If we have three factors related by brackets, it is obviously the same to insert first the left into the middle and then everything into the right factor, or to insert first the middle into the right and finally the left into the same place in the middle factor which is now considered as a subword of the right factor. We further define a formal unit $e$ by

$$m[e \otimes X] = m[X \otimes e] = X \qquad \forall\, X \in \mathcal{A} \ .$$

The unit $e$ is *not* considered as a PSW. It is convenient to consider $e$ as produced by an operation

$$E : \mathbf{Q} \to \mathcal{A} \ , \qquad E(q) = qe \ .$$

The third step is to make $\mathcal{A}$ to a coalgebra. The operations of a coalgebra are the duals of the algebra operations. Dual means turning the arrows. For instance, the dual of the above unit $E$, the counit $\bar{e}$, will be a formal operation given by

$$\bar{e} : \mathcal{A} \to \mathbf{Q} \ , \qquad \bar{e}[qe] := q \ , \qquad \bar{e}[X] := 0 \ \ \forall\, X \neq e \ , \quad X \in \mathcal{A} \ .$$

Now comes a physically significant ingredient of our coalgebra, the coproduct $\Delta$. A product was the assignment of one element to sums of pairs of other elements, the pairs being connected by horizontal brackets. Hence, a coproduct will be the splitting of one element into sums of connected pairs of other elements, in symbols

$$\Delta : \mathcal{A} \to \overbrace{\mathcal{A} \otimes \mathcal{A}} \ .$$

The philosophy is that $\Delta$ provides the splitting of a 1PI-graph $\Gamma$ into a formal sum of tensor products of all possible divergent subgraphs $\gamma_i$ (left factor) by the fraction $\Gamma/\gamma_i$ obtained by reducing $\gamma_i$ to a point (right factor). The left factors are, moreover, treated by the $R$-operation and a horizontal bracket connects them to the places in $\Gamma/\gamma_i$ where they had been before.

Let us formalize this idea. The graph $\Gamma$ is represented by a PW $X$ describing its forest structure. Let $X$ be an iPW and $Y$ be a PSW of $X$ in the sense of section 2. We are going to define the fraction $X/Y$. If $Y = X$ we define $X/X = e$ (no brackets between $X$ and $e = X/X$). Otherwise we label the rows of $X$. Each row of $Y$ is a substring of one determined row of $X$. We give to the $Y$-rows the labels of the $X$-rows they are contained in. These labels could be ambiguous but we fix one choice for all subwords of $X$. We delete all rows of $X$ which have no counterpart in $Y$. Let $X'$ be the result of this cutting procedure. Now, there is a $1:1$-correspondence between the rows of $X'$ and $Y$. It remains to remove in each $X'$-row its related $Y$-row and to put one end of a string at this position, the other end is attached to the $Y$-row. The result is an $n$-line iPW $X/Y$ connected by $n$ strings (or brackets) to the $n$-line PSW $Y$.

Let us now compare $X/Y$ with a second PSW $Z$ of $X$. Assume that both have some rows of common labels. We remove all but the common rows of $Z$ and $X/Y$ and moreover all of the remaining related rows, where the $Z$-row is not a subword of its corresponding $X/Y$-row. (This happens if $Z$ is a subword of $Y$.) The results are two $n'$-line PWs $Z'$ and $(X/Y)'$ where each row of $Z'$ is a subword of its corresponding $(X/Y)'$-row. Again we remove in each $(X/Y)'$-row its related $Z'$-row and put one end of a string at this position, the other end is attached to the $Z'$-row. The outcome is the $n'$-line iPW $(X/Y)/Z$ connected by $n'$ strings with the $n'$-line PW $Z'$ and by the same number $n'$ of strings with the $n'$-line PW $Y'$ obtained from $Y$ by deleting its rows whose labels do not meet the labels of the rows of $Z'$. If after deleting the rows nothing remains from $X/Y$ and $Z$ then we put $(X/Y)/Z = 0$. This procedure can be repeated for various PSWs $Y_{i_1}, \ldots, Y_{i_k}$ of $X$ and gives $n_k$-line PWs

$$X/(Y_{i_1}Y_{i_2}\cdots Y_{i_k}) \equiv ((\ldots((X/Y_{i_1})/Y_{i_2})/\ldots)/Y_{i_k})$$

connected by each $n_k$ strings to the $n_k$-line PWs $Y'_{i_1}, \ldots, Y'_{i_k}$. If we have a disjoint product $X = \prod_i X_i$ and $Y$ is a PSW of $X$, then $Y$ is a PSW of precisely one $X_j$ or one $X_j$ itself. In that case we define $X/Y := \{X_j/Y\} \prod_{i\neq j} X_i$.

Now, the coproduct of a PW $X$ containing the PSWs $Y_1, \ldots Y_n$ is defined by

$$\begin{aligned} &\Delta[e] := e \otimes e\,, \\ &\Delta[X] := e \otimes X + \sum_{i_1<i_2<\cdots<i_k} \Big\{ R[Y_{i_k}] \cdots R[Y_{i_2}] R[Y_{i_1}] \otimes X/(Y_{i_1}Y_{i_2}\cdots Y_{i_k}) \Big\}\,, \end{aligned} \tag{18}$$

where the sum runs over all ordered subsets $\{i_1, \ldots, i_k\} \subset \{1, 2, \ldots, n\}$. The order of the factors and products is not important in this definition, but we must avoid taking identical terms several times. We have omitted the primes which indicate that only the common rows of $Y_{i_1}, \ldots Y_{i_k}$ are taken.

Our algebra $\mathcal{A}$ also contains elements of the type $R[X]$, where $X$ is a PW. Kreimer gives two possible definitions for $\Delta \circ R$,

$$\Delta[R[X]] = \Delta[X]\ , \tag{19a}$$

$$\Delta[R[X]] = (\mathrm{id} \otimes R) \circ \Delta[X]\ . \tag{19b}$$

He chooses to work with (19a). We prefer (19b), because $R[X]$ is always a local counterterm $\bullet$. The philosophy is that $\Delta$ splits a graph into subgraphs and reminders. Hence, both of them should be local counterterms in this example, $\Delta[\bullet] = \sum \bullet \otimes \bullet$, and the natural definition is (19b) or

$$\Delta[R[X]] := e \otimes R[X] + \sum_{i_1<i_2<\cdots<i_k} \Big\{ R[Y_{i_k}] \cdots R[Y_{i_2}] R[Y_{i_1}] \otimes R[X/(Y_{i_1}Y_{i_2}\cdots Y_{i_k})]' \Big\}\,. \tag{20}$$

Here, the prime means that $R[X/X]$ has to be replaced by $e$ instead of $R[e]$. This can be easily interpreted in terms of PSWs. The PSWs $Y_i$ of $R[X]$ are identical with the PSWs of $X$, except for the total PW $R[X]$. The fraction $R[X]/Y_i$ obtained by removing $Y_i$ in $R[X]$ clearly coincides with $R[X/Y_i]$, except for $R[X]/R[X] = e$.

There are of course some consistency conditions to fulfill before we can call $\mathcal{A}$ a coalgebra. One of these conditions to $\Delta$ is coassociativity, which is derived from associativity by turning the arrows: If we split one element into sum of pairs, it must be the same to split the left or the right factor further. In symbols, coassociativity means

$$(\mathrm{id}\otimes\Delta)\circ\Delta[X] = (\Delta\otimes\mathrm{id})\circ\Delta[X]\ , \qquad \forall\, X\in\mathcal{A}\ . \tag{21}$$

We give the proof in proposition 1 in the appendix. For the choice (19a), coassociativity was only satisfied under the additional condition (17). We suppose that Kreimer had rejected (19b) because of problems with the antipode defined below. These problems are due to his product definition and disappear (partly) with our modification. We recall that there are also physical reasons in favour of (19b) and for the horizontal brackets across tensor products. However, (19b) does not remove all problems. The 'counit' $\bar{e}$ is only a left counit and becomes a true counit under the condition (17), see proposition 2. This means that we have

$$(\bar{e}\otimes\mathrm{id})\circ\Delta[X] = X\ , \qquad \forall\, X\in\mathcal{A}\,, \tag{22a}$$

which is good, but only the weak property

$$(\mathrm{id}\otimes\bar{e})\circ\Delta[X] \approx X \tag{22b}$$

in general. We see that without condition (17), $\mathcal{A}$ is only some sort of left-coalgebra. Moreover, the 'antipode' $S$ defined below turns out to be only a left antipode in general.

So far we have equipped $\mathcal{A}$ with the structures of an algebra and a coalgebra. Both merge to a bialgebra if $\Delta$ is an algebra homomorphism,

$$\Delta\circ m[X\otimes Y] = \tilde{m}[\Delta[X]\otimes\Delta[Y]]\ , \qquad \forall\, X,Y\in\mathcal{A}\ . \tag{23}$$

We must check that it is possible to define $\tilde{m} : (\mathcal{A}\otimes\mathcal{A})\otimes(\mathcal{A}\otimes\mathcal{A}) \to \mathcal{A}\otimes\mathcal{A}$ appropriately. This definition depends on the brackets (which we have omitted) joining the various $\mathcal{A}$-factors. If $X$ and $Y$ in (23) are disjoint, we also take the disjoint product

$$\tilde{m}\Big[(\prod_i R[X_i]\otimes X')\otimes(\prod_j R[Y_j]\otimes Y')\Big] := (\prod_j R[Y_j]\ \prod_i R[X_i])\otimes(X'Y')\ . \tag{24a}$$

It is evident now that (23) is fulfilled for the disjoint product, because the subwords of $XY$ are the subwords $X_i$ of $X$ and $Y_i$ of $Y$ together. Extension of the disjoint $\tilde{m}$ to more than two factors is obvious. For the tensor product related by brackets we define

$$\tilde{m}\Big[(\prod_i R[X_i]\otimes X')\otimes(\prod_j R[Y_j]\otimes Y')\Big] := 0\ , \tag{24b}$$

$$\tilde{m}\Big[(\prod_i R[X_i]\otimes X')\otimes(\prod_j R[Y_j]\otimes Y')\Big] := (\prod_j R[Y_j]\ \prod_i R[X_i])\otimes m[(X'\otimes Y')]\ , \tag{24c}$$

$$\tilde{m}\Big[(R[X]\otimes e)\otimes(\prod_j R[Y_j]\otimes Y')\Big] := (\prod_j R[Y_j]\ R[X])\otimes Y'\ , \quad Y'\neq e\ , \tag{24d}$$

$$\tilde{m}\Big[(R[X]\otimes e)\otimes(R[Y]\otimes e)\Big] := m[X\otimes R[Y]]\otimes e\ . \tag{24e}$$

This is easy to understand. In (24b) we would have the PSW $X'$ on the right of its reminder $Y_j$, which is impossible. Thus, the generic case is (24c) where $X$ is a PSW of $Y'$. This implies that $X$ and all its subwords $X_i$ are disjoint from all $Y_j$. A special case is $X' = e$. As long as $Y' \neq e$, the generic formula (24c) remains correct and reads (24d). For $Y' = e$, however, the product on lhs of (24e) must give the insertion of $X$ and not $R[X]$ into $R[Y]$. These definitions remain unchanged for $R[X']$ and $R[Y']$ instead of $X'$ and $Y'$. The product of several disjoint factors $X^k$ related by horizontal brackets to $Y$ is defined via associativity, which for two factors $X^k$ reads $\tilde{m} \circ (\tilde{m} \otimes \mathrm{id} \otimes \mathrm{id}) = \tilde{m} \circ (\mathrm{id} \otimes \mathrm{id} \otimes \tilde{m})$. The proof that (23) is fulfilled in the case where $X$ is related to $Y$ by a bracket is straightforward.

The last step extends the bialgebra to a Hopf algebra. On a Hopf algebra there exists the additional structure of an antipode $S : \mathcal{A} \to \mathcal{A}$, which is the dual of the inverse in an algebra. Our algebra does not have an inverse (except for $e^{-1} = e$), nevertheless it has an antipode, which will provide the link to the forest formula:

$$\begin{aligned}
S[e] &= e \;, && && (25a)\\
S[XY] &= S[Y]S[X] \;, && \forall X, Y \in \mathcal{A} \;, && (25b)\\
S[X] &= -X - m \circ (\mathrm{id} \otimes S) \circ P_2 \circ \Delta[X] \;, && \forall \text{ iPW } X \in \mathcal{A} \;, && (25c)\\
S[R[X]] &= -R[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \circ \Delta[X]] \;, && \forall \text{ iPW } X \in \mathcal{A} \;, && (25d)
\end{aligned}$$

where $P_2 = (\mathrm{id} - E \circ \bar{e}) \otimes (\mathrm{id} - E \circ \bar{e})$. The antipode is by (25) recursively defined, because in $P_2 \circ \Delta[X]$ only smaller words than $X$ survive, and for primitive words $(x)$ we simply have $S[(x)] = -(x)$ and $S[R[(x)]] = -R[(x)]$. We prove in proposition 3 that $S$ is only a left antipode and becomes a true antipode for renormalization schemes fulfilling (17),

$$\begin{aligned}
m \circ (S \otimes \mathrm{id}) \circ \Delta[X] &\sim E \circ \bar{e}[X] && (26a)\\
m \circ (\mathrm{id} \otimes S) \circ \Delta[X] &\approx E \circ \bar{e}[X] \;. && (26b)
\end{aligned}$$

Formula (26a) relies deeply on the fact that for $X$ being an iPW the equation

$$\begin{aligned}
m \circ (S \otimes \mathrm{id}) \circ \Delta[X] &= (\mathrm{id} - R)\Big[X + \sum_T \Big\{ m\Big[ \prod_{i \in T} (-R[\bar{X}_i\overline{]) \otimes X/(\prod_{i \in T}}X_i) \Big] \Big\} \Big] \\
&= (\mathrm{id} - R)[\bar{X}] \;, \qquad\qquad (27a)\\
R[\bar{X}_i] &:= -S[R[X_i]] \;,
\end{aligned}$$

reproduces Bogoliubov's recurrence formula of renormalization [7]. In this equation, $X_i \neq X$, $i = 1, \dots, n$, are the proper PSWs of $X$ and $T$ the set of all ordered subsets of $\{1, \dots, n\}$. Denoting by $X_{ij} \neq X_i$, $j = 1, \dots n_i$, the proper PSW of $X_i$, we can write

$$\begin{aligned}
R[\bar{X}_i] \equiv -S[R[X_i]] &= R[X_i + m \circ (S \otimes \mathrm{id}) \circ P_2 \circ \Delta[X_i]] \\
&= R\Big[X_i + \sum_{T_i} \Big\{ m\Big[ \prod_{j \in T_i} S[R[X_{ij}\overline{]] \otimes X_i/(\prod_{j \in T_i}}X_{ij}) \Big] \Big\} \Big] \;. \qquad (27b)
\end{aligned}$$

Thus, $\bar{X}_i$ has the same structure as $\bar{X}$, and we obtain indeed a recurrence formula. The integrand $\bar{X}$ associated to an integrand $X$ is pre-finite, which means that all subdivergences are compensated. The remaining superficial divergence is compensated by $\mathrm{id}-R$. In any recurrence step, the pre-finite $\bar{X}$ is given by $X$ plus the sum over all different products of pre-finite integrands of disjoint subdivergences, where each such product is multiplied by the fraction obtained by shrinking the subdivergences of $X$ to points. It is important to include all combinations of disjoint subdivergences, which are encoded in the set of maximal forests.

This means that in describing a Feynman graph $\Gamma$ with subdivergences by a parenthesized word $X$, we must somehow include in $X$ all maximal forests of $\Gamma$. That is why we have written the maximal forests as lines of $X$. The maximal forests are defined by the relative position of the subdivergences. Each time we meet an overlap of subdivergences we have a branching of forests. Having defined the forests we must say how to detect the disjoint subdivergences. Forests contain by definition no overlapping divergences, so the only problem is to avoid nested divergences. This is now achieved by taking from $\Gamma$ any subdivergence and shrinking it to a point, from the rest we take again any $\Gamma$-subdivergence and shrink it, and so on. By variation of the choices we find all products of disjoint subdivergences. It is extremely important that if a subdivergence occurs in two or more forests, we must shrink it in all of them simultaneously. If not, we could remove in the next step the same subdivergence again, or a subsubdivergence. All that must be avoided, and we did it by connecting the closing parentheses of simultaneously shrinkable boxes. Finally, as the disjoint product is commutative, we must get rid of the multiplicities. We achieved this by assigning numbers to the subdivergences and by considering only the ordered products.

In conclusion, our modified definition of a parenthesized word that keeps track of different maximal forests and connects simultaneously shrinkable boxes is the correct language for Bogoliubov's recurrence formula [7]. This formula has an explicit solution, Zimmermann's forest formula [4]. Both are reproduced by coproduct and antipode of a (left-) Hopf algebra via $m \circ (S \otimes \mathrm{id}) \circ \Delta$.

We remark that the crucial formula (26a) is actually a stronger equivalence $\simeq$. Due to the forest formula (27), the difference between left and right hand sides is *finite* in any renormalization scheme. This suggests to give up the goal of building a true Hopf algebra whose axioms are fulfilled modulo the weak equivalence $\approx$ and building instead a left-Hopf algebra $(\mathcal{A}, m, E, \Delta, \bar{e}, S)$ whose axioms are

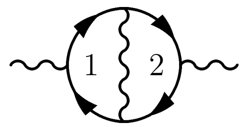
$$
\begin{aligned}
m \circ (m \otimes \mathrm{id}) &= m \circ (\mathrm{id} \otimes m) \;, & (\Delta \otimes \mathrm{id}) \circ \Delta &= (\mathrm{id} \otimes \Delta) \circ \Delta \;, \\
(\bar{e} \otimes \mathrm{id}) \circ \Delta &= \mathrm{id} \;, & m \circ (S \otimes \mathrm{id}) \circ \Delta &\simeq E \circ \bar{e} \;.
\end{aligned}
$$

That could be an alternative as it is unclear so far what the Hopf algebra structure of perturbative quantum field theories is good for. There are some relations [9] to the recent article by Connes and Moscovici [8], but further work to clarify this point is necessary.

## 6 Two examples for the coproduct and the forest formula

We compute here the coproducts and forest formulas for two striking examples of section 3. By PSW we shall always mean proper PSW, we write the trivial PWs explicitly. The PSWs of

$$X = \begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big] \tag{8}$$




are obviously

$$Y_1 = (v_1)\,, \qquad Y_2 = (v_2)\,. \tag{$8_\mathrm{s}$}$$

Let us compute $X/Y_1$. The only row of $Y_1$ can only be related to the upper row of $X$ so that $X' = ((v_1)p_2)$. To obtain $X/Y_1$ we must remove $Y_1$ from $X'$, the result is $X/Y_1 = (p_2)$. There remains the horizontal bracket attached to $Y_1$ which points into $(p_2)$. Analogously one finds

$$X/Y_1 = (p_2)\,, \qquad X/Y_2 = (p_1)\,, \qquad X/(Y_1Y_2) = 0\,, \tag{$8_\mathrm{r}$}$$

because $Y_2 = (v_2)$ is not a PSW of $X/Y_1 = (p_2)$. Therefore, the coproduct reads

$$\Delta[X] = e \otimes \begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big] + R\left[\begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big]\right] \otimes e + R[(v_1)] \otimes (p_2) + R[(v_2)] \otimes (p_1)\,. \tag{$8_\Delta$}$$

Let us now apply the operator $m \circ (S \otimes \mathrm{id})$. To avoid unnecessary calculation we use the general result (27a),

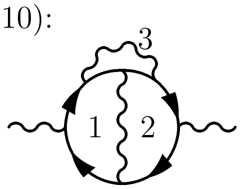
$$m \circ (S \otimes \mathrm{id}) \circ \Delta[X] = (\mathrm{id} - R)[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \circ \Delta[X]]\ .$$

The projection $P_2$ removes all terms containing the unit $e$ so that in our case we have $P_2 \circ \Delta[X]] = R[(v_1)] \otimes (p_2) + R[(v_2)] \otimes (p_1)$. This gives

$$\begin{aligned}
&m \circ (S \otimes \mathrm{id}) \circ \Delta[X] \\
&\quad = (\mathrm{id} - R)\left[\begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big] + m\Big[S[R[(v_1)]] \otimes (p_2)\Big] + m\Big[S[R[(v_2)]] \otimes (p_1)\Big]\right] \\
&\quad = (\mathrm{id} - R)\left[\begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big] - m\Big[R[(v_1)] \otimes (p_2)\Big] - m\Big[R[(v_2)] \otimes (p_1)\Big]\right] \\
&\quad = (\mathrm{id} - R)\left[\begin{matrix}((v_1)p_2)\\((v_2)p_1)\end{matrix}\Big] - (R[(v_1)]p_2) - (R[(v_2)]p_1)\right]\ .
\end{aligned} \tag{$8_\mathrm{f}$}$$

Let us do the same steps for example (10):

$$X = \begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] , \qquad (10)$$




$$Y_1 = \begin{matrix}(v_3)\\(v_3)\end{matrix}\Big] , \qquad Y_2 = ((v_3)v_{13}) , \qquad Y_3 = ((v_3)v_{23}) , \qquad (10_{\mathrm{s}})$$

$$X/Y_1 = \begin{matrix}((v_{13})p_2)\\((v_{23})p_1)\end{matrix}\Big] , \quad X/Y_2 = (p_2) , \qquad X/Y_3 = (p_1) , \qquad (10_{\mathrm{r}})$$

$$X/(Y_1Y_2) = X/(Y_1Y_3) = X/(Y_2Y_3) = X/(Y_1Y_2Y_3) = 0 ,$$

$$\Delta[X] = e \otimes \begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] + R\left[\begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big]\right] \otimes e + R\left[\begin{matrix}(v_3)\\(v_3)\end{matrix}\Big]\right] \otimes \begin{matrix}((v_{13})p_2)\\((v_{23})p_1)\end{matrix}\Big]$$
$$+ R[((v_3)v_{13})] \otimes (p_2) + R[((v_3)v_{23})] \otimes (p_1) , \qquad (10_\Delta)$$

$$m \circ (S \otimes \mathrm{id}) \otimes \Delta[X]$$
$$= (\mathrm{id} - R)\Bigg[\begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] + m\Big[S\Big[R\Big[\begin{matrix}(v_3)\\(v_3)\end{matrix}\Big]\Big]\Big] \otimes \begin{matrix}((v_{13})p_2)\\((v_{23})p_1)\end{matrix}\Big]\Big]$$
$$+ m\big[S[R[((v_3)v_{13})]] \otimes (p_2)\big] + m\big[S[R[((v_3)v_{23})]] \otimes (p_1)\big]\Bigg]$$
$$= (\mathrm{id} - R)\Bigg[\begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] - m\Big[R\Big[\begin{matrix}(v_3)\\(v_3)\end{matrix}\Big]\Big] \otimes \begin{matrix}((v_{13})p_2)\\((v_{23})p_1)\end{matrix}\Big]\Big]$$
$$- m\Big[\big\{R[((v_3)v_{13})] + R[m \circ (S \otimes \mathrm{id}) \circ P_2\Delta[((v_3)v_{13})]]\big\} \otimes (p_2)\Big]$$
$$- m\Big[\big\{R[((v_3)v_{23})] + R[m \circ (S \otimes \mathrm{id}) \circ P_2\Delta[((v_3)v_{23})]]\big\} \otimes (p_1)\Big]\Bigg]$$
$$= (\mathrm{id} - R)\Bigg[\begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] - \begin{matrix}((R[(v_3)]v_{13})p_2)\\((R[(v_3)]v_{23})p_1)\end{matrix}\Big]$$
$$- (R[((v_3)v_{13})]p_2) - m\Big[R\big[m\big[S[R[(v_3)]] \otimes (v_{13})\big]\big] \otimes (p_2)\Big]$$
$$- (R[((v_3)v_{23})]p_1) - m\Big[R\big[m\big[S[R[(v_3)]] \otimes (v_{23})\big]\big] \otimes (p_1)\Big]\Bigg]$$
$$= (\mathrm{id} - R)\Bigg[\begin{matrix}(((v_3)v_{13})p_2)\\(((v_3)v_{23})p_1)\end{matrix}\Big] - \begin{matrix}((R[(v_3)]v_{13})p_2)\\((R[(v_3)]v_{23})p_1)\end{matrix}\Big] - (R[((v_3)v_{13})]p_2)$$
$$+ (R[(R[(v_3)]v_{13})]p_2) - (R[((v_3)v_{23})]p_1) + (R[(R[(v_3)]v_{23})]p_1)\Bigg] . \qquad (10_{\mathrm{f}})$$

Example (11) is similar to (10) and left as an exercise to the reader. Example (12) is the obvious generalization of (8) to three maximal forests.

## Appendix: Verification of the Hopf algebra properties

**Proposition 1** *The coproduct* $\Delta$ *is coassociative,* $(\Delta \otimes \mathrm{id}) \circ \Delta = (\mathrm{id} \otimes \Delta) \circ \Delta$.

*Proof.* Let $X$ be an iPW which is not $R[X']$. Let $X_i \neq X$, $i = 1, \dots, n$, be the proper PSW of $X$. Let $T$ be the set of all ordered subsets of $\{1, 2, \dots, n\}$. We write the contribution of the trivial PSW $X$ of $X$ explicitly:

$$\Delta[X] = e \otimes X + R[X] \otimes e + \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes X/(\prod_{i \in T} X}_i) \Big\} \; .$$

This gives

$$\begin{aligned}
&(\mathrm{id} \otimes \Delta) \circ \Delta[X] \qquad\qquad \text{(A.1)}\\
&= e \otimes \Big\{ e \otimes X + R[X] \otimes e + \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes X/(\prod_{i \in T} X}_i) \Big\} \Big\} + R[X] \otimes e \otimes e \\
&+ \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes e \otimes X/(\prod_{i \in T} X}_i) \Big\} + \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes R[X/(\prod_{i \in T} X}_i)] \otimes e \Big\} \\
&+ \sum_T \Big\{ \prod_{i \in T} R[X_i] \otimes \sum_{T'} \Big\{ \prod_{j \in T'} R\big[\{X/(\prod_{i \in T} X_i)\}_j\big] \otimes \{X/(\prod_{i \in T} X_i)\}/\big(\prod_{j \in T'} \{X/(\prod_{i \in T} X_i)\}_j\big) \Big\} \Big\},
\end{aligned}$$

where $\{X/(\prod_{i \in T} X_i)\}_j$ are the proper PSW of $X/(\prod_{i \in T} X_i)$, $j = 1, \dots, n' < n$, and $T'$ is the set of all ordered subsets of $\{1, \dots, n'\}$. The following terms can be rearranged:

$$\begin{aligned}
&e \otimes e \otimes X \\
&+ \Big\{ e \otimes R[X] \otimes e + R[X] \otimes e \otimes e + \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes R[X/(\prod_{i \in T} X}_i)] \Big\} \otimes e \Big\} \\
&= (\Delta \otimes \mathrm{id})(e \otimes X + R[X] \otimes e) \qquad\qquad \text{(A.2)}
\end{aligned}$$

so that there remain

$$\sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes e \otimes X/(\prod_{i \in T} X}_i) \Big\} + e \otimes \sum_T \Big\{ \prod_{i \in T} R[X_i\overline{] \otimes X/(\prod_{i \in T} X}_i) \Big\} \quad \text{and} \qquad \text{(A.3)}$$

$$\sum_{T,T'} \Big\{ \prod_{i \in T} R[X_i] \otimes \Big\{ \prod_{j \in T'} R\big[\{X/(\prod_{i \in T} X_i)\}_j\big] \otimes \{X/(\prod_{i \in T} X_i)\}/\big(\prod_{j \in T'} \{X/(\prod_{i \in T} X_i)\}_j\big) \Big\} \Big\}. \qquad \text{(A.4)}$$

We investigate $\{X/(\prod_{i \in T} X_i)\}_j$. Either this is a PSW of $X$ or not. If not there must exist a PSW $X_m$ of $X$ and some PSWs $X_k$ with $k \in T^m \subset T$ such that $\{X/(\prod_{i \in T} X_i)\}_j = X_m/(\prod_{k \in T^m} X_k)$. This means that $T' = T_1 \oplus T_2$ (both $T_1, T_2$ can be empty but not the sum) and

$$\prod_{j \in T'} R\big[\{X/(\prod_{i \in T} X_i)\}_j\big] = \prod_{l \in T_1} R[X_l] \prod_{m \in T_2} R[X_m/(\prod_{k_m \in T^m} X_{k_m})] \; .$$

Let us assume that $T_2$ contains at least two elements $m_1, m_2$ and perform the factorization

$$\{X/(\prod_{i \in T} X_i)\}/\big(\{X_{m_1}/(\prod_{k_1 \in T^{m_1}} X_{k_1})\}\{X_{m_2}/(\prod_{k_2 \in T^{m_2}} X_{k_2})\}\big) \; . \qquad \text{(A.5)}$$

Recall that $T^{m_1}\subset T$ and $T^{m_2}\subset T$ and assume that $X_n \in T^{m_1}\cap T^{m_2}$. The fraction (A.5) will only be non-zero if $X_{m_1}/(\prod_{k_1\in T^{m_1}} X_{k_1})$ and $X_{m_2}/(\prod_{k_2\in T^{m_2}} X_{k_2})$ occur together and disjoint in at least one row of $X/(\prod_{i\in T} X_i)$. These rows correspond to those rows of $X$ each of which contain all $X_i$, $i\in T$, too. But each $X_i$ occurs precisely once in any row, so does the $X_n$ in question, hence it will either occur in $T^{m_1}$ or in $T^{m_2}$, but never in both. Therefore, we have a direct sum decomposition $T = T_3 \oplus \bigoplus_{m\in T_2} T^m$ and (A.4) takes the form

$$\begin{aligned}
(\mathrm{A.4}) = & \sum_{\{T_1,T_2,T_3,\bigcup_{m\in T_2}T^m\}} \Big\{ \prod_{i\in T_3} R[X_i] \prod_{m\in T_2} \{ \prod_{k_m\in T^m} R[X_{k_m}] \}\otimes \\
& \otimes \prod_{l\in T_1} R[X_l] \prod_{m\in T_2} R\big[X_m/(\prod_{k_m\in T^m} X_{k_m})\big] \otimes X/(\prod_{m\in T_2} X_m \prod_{l\in T_1} X_l \prod_{i\in T_3} X_i)\Big\} \\
= & \sum_T \Big\{ \{ \sum_{T_3\subset T} \prod_{i\in T_3} R[X_i] \otimes \sum_{T_1\subset T/T_3} \prod_{l\in T_1} R[X_l] \Big\}\times \\
& \times \{ \prod_{m\in T_2 = T/(T_1\oplus T_3)} \sum_{T^m} \Big\{ \prod_{k_m\in T^m} R[X_{k_m}] \otimes R[X_m/(\prod_{k_m\in T^m} X_{k_m})]\Big\}\} \otimes X/(\prod_{j\in T} X_j)\Big\}.
\end{aligned}$$

Note that $T_1, T_2, T_3$ can be empty, in that case the missing product over $R[X_j]$ has to be replaced by $e$. If $T_2$ is empty then the sum over $T_1 = T/T_3$ has to be omitted. Observe that neither $T_1 \oplus T_2$ nor $T_3 \oplus T_2$ can be empty, but these two terms $T_2 = \emptyset$ and either $T_1 = \emptyset$ or $T_3 = \emptyset$ are precisely those of (A.3). All together can be rewritten as

$$\begin{aligned}
(\mathrm{A.3}) + (\mathrm{A.4}) = & \sum_T \Big\{ \prod_{j\in T} \{ e\otimes R[X_j] + R[X_j]\otimes e + \\
& + \sum_{T^j} \Big\{ \prod_{k_j\in T^j} R[X_{k_j}] \otimes R[X_j/(\prod_{k_j\in T^j} X_{k_j})]\Big\}\} \otimes X/(\prod_{j\in T} X_j)\Big\} \\
= & (\Delta\otimes \mathrm{id})\{ \sum_T \prod_{j\in T} R[X_j] \otimes X/(\prod_{j\in T} X_j)\}\,, \qquad (\mathrm{A.6})
\end{aligned}$$

and we conclude

$$(\mathrm{A.2}) + (\mathrm{A.3}) + (\mathrm{A.4}) = (\Delta\otimes\mathrm{id})\circ\Delta[X] = (\mathrm{id}\otimes\Delta)\circ\Delta[X]\ . \qquad (\mathrm{A.7})$$

To finish the proof of coassociativity of $\Delta$ we must write down

$$\begin{aligned}
(\mathrm{id}\otimes\Delta)\circ\Delta[R[X]] &= (\mathrm{id}\otimes\Delta)\circ(\mathrm{id}\otimes R)\circ\Delta[X] \\
&= (\mathrm{id}\otimes\mathrm{id}\otimes R)\circ(\mathrm{id}\otimes\Delta)\circ\Delta[X] \\
&= (\mathrm{id}\otimes\mathrm{id}\otimes R)\circ(\Delta\otimes\mathrm{id})\circ\Delta[X] \\
&= (\Delta\otimes\mathrm{id})\circ\Delta[R[X]]\ , \\
(\mathrm{id}\otimes\Delta)\circ\Delta[XY] &= \hat{m}\big[\{(\mathrm{id}\otimes\Delta)\circ\Delta[X]\}\otimes\{(\mathrm{id}\otimes\Delta)\circ\Delta[Y]\}\big] \\
&= \hat{m}\big[\{(\Delta\otimes\mathrm{id})\circ\Delta[X]\}\otimes\{(\Delta\otimes\mathrm{id})\circ\Delta[Y]\}\big] \\
&= (\Delta\otimes\mathrm{id})\circ\Delta[XY]\ .
\end{aligned}$$

We have defined $\hat{m}[\{X'\otimes X''\otimes X'''\}\otimes\{Y'\otimes Y''\otimes Y'''\}] := X'Y'\otimes X''Y''\otimes X'''Y'''$, if a set of brackets connects $\{X', X'', X'''\}$ and a disjoint set of brackets connects $\{Y', Y'', Y'''\}$. □

**Proposition 2** *The 'counit' $\bar{e}$ is only a left counit, it fulfills $(\bar{e} \otimes \mathrm{id}) \circ \Delta = \mathrm{id}$, but in general only the weak relation $(\mathrm{id} \otimes \bar{e}) \circ \Delta \approx \mathrm{id}$.*

*Proof.* An element of $\mathcal{A}$ is a formal linear combinations of disjoint products $X = \prod_i X_i \; \prod_j R[Y_j]$, where $X_i, Y_i$ are iPWs. The case $X = e$ is trivial. We have

$$\Delta[X] = \prod_i R[X_i] \; \prod_j R[Y_j] \otimes e + e \otimes \prod_i X_i \; \prod_j R[Y_j] + \sum Z \otimes Z' \;, \qquad X \neq e \;,$$

where $Z, Z'$ stand for terms which do not contain the unit $e$ and which are annihilated by $\bar{e}$. Hence,

$$\begin{aligned}(\bar{e} \otimes \mathrm{id}) \circ \Delta[X] &= \prod_i X_i \; \prod_j R[Y_j] = X \;, \\ (\mathrm{id} \otimes \bar{e}) \circ \Delta[X] &= \prod_i R[X_i] \; \prod_j R[Y_j] \approx R[X] \sim X \;.\end{aligned}$$

In the last line we need (17) to obtain equivalence with $X$. □

**Proposition 3** *The 'antipode' $S$ is only a left antipode, it fulfills $m \circ (S \otimes \mathrm{id}) \circ \Delta \sim E \circ \bar{e}$ but in general only the weak property $m \circ (\mathrm{id} \otimes S) \circ \Delta \approx E \circ \bar{e}$.*

*Proof.* We first apply both operators to $X$ and $R[X]$, where $X \neq e$ is a iPW:

$$\begin{aligned} m \circ (S \otimes \mathrm{id}) \circ \Delta[X] &= m \circ (S \otimes \mathrm{id})[e \otimes X + R[X] \otimes e + P_2 \Delta[X]] \\ &= X + S[R[X]] + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X] \\ &= X - R\big[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X]\big] + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X] \\ &= (\mathrm{id} - R)\big[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X]\big] \\ &\sim 0 = E \circ \bar{e}[X] \;, \end{aligned}$$

$$\begin{aligned} m \circ (S \otimes \mathrm{id}) \circ \Delta[R[X]] &= m \circ (S \otimes \mathrm{id})[e \otimes R[X] + R[X] \otimes e + P_2 \Delta[R[X]]] \\ &= R[X] + S[R[X]] + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[R[X]] \\ &= R[X] - R\big[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X]\big] + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[R[X]] \\ &= 0 = E \circ \bar{e}[R[X]] \;, \end{aligned}$$

$$\begin{aligned} m \circ (\mathrm{id} \otimes S) \circ \Delta[X] &= m \circ (\mathrm{id} \otimes S)[e \otimes X + R[X] \otimes e + P_2 \Delta[X]] \\ &= S[X] + R[X] + m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[X] \\ &= -(X + m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[X]) + R[X] + m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[X] \\ &= -(\mathrm{id} - R)[X] \sim 0 = E \circ \bar{e}[X] \;, \end{aligned}$$

where we have used

$$\sum_T m\big[\prod_{i \in T} R[X_i] \otimes R[X/(\prod_{i \in T} X_i)]\big] = R\big[\sum_T m\big[\prod_{i \in T} R[X_i] \otimes X/(\prod_{i \in T} X_i)\big]\big] \;, \qquad \text{(A.8)}$$

for $X_i \neq X$ being the proper PSW of $X$. The remaining case is more complicated:

$$\begin{aligned} m \circ (\mathrm{id} \otimes S) \circ \Delta[R[X]] &= S[R[X]] + R[X] + m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[R[X]] \\ &= -R\big[X + m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X]\big] + R[X] + m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[R[X]] \\ &= m \circ (\mathrm{id} \otimes S) \circ P_2 \Delta[R[X]] - R\big[m \circ (S \otimes \mathrm{id}) \circ P_2 \Delta[X]\big] \;. \qquad \text{(A.9)} \end{aligned}$$

We transform the first term, using (A.8) and the definition of $S$ acting on $R[\,.\,]$:

$$\begin{aligned} & m\circ(\mathrm{id}\otimes S)\circ P_2\Delta[R[X]] \qquad\qquad (A.10)\\ &\quad = -m[P_2\Delta[R[X]]] - m\circ(\mathrm{id}\otimes\{R\circ m\circ(S\otimes\mathrm{id})\circ P_2\Delta\})\circ P_2\Delta[X]\\ &\quad = -R\Big[m[P_2\Delta[X]] + m\circ(\mathrm{id}\otimes m)\circ(\mathrm{id}\otimes S\otimes\mathrm{id})\circ(\mathrm{id}\otimes P_2\Delta)\circ P_2\Delta[X]\Big]\;. \end{aligned}$$

Now observe that due to coassociativity of $\Delta$ we have

$$\begin{aligned} (\mathrm{id}\otimes P_2\Delta)\circ P_2\Delta[X] &= P_3\circ(\mathrm{id}\otimes\Delta)\circ\Delta[X] = P_3\circ(\Delta\otimes\mathrm{id})\circ\Delta[X]\\ &= (P_2\otimes\mathrm{id})\circ(\Delta\otimes\mathrm{id})\circ P_2\Delta[X]\;, \end{aligned}$$

with $P_3 = (\mathrm{id}{-}E\circ\bar{e})\otimes(\mathrm{id}{-}E\circ\bar{e})\otimes(\mathrm{id}{-}E\circ\bar{e})$. Note that $\Delta$ is multiplicative, not $(P_2\Delta)$. Using also associativity of $m$ we can write

$$\begin{aligned} & -R\big[m\circ(\mathrm{id}\otimes m)\circ(\mathrm{id}\otimes S\otimes\mathrm{id})\circ(\mathrm{id}\otimes P_2\Delta)\circ P_2\Delta[X]\big]\\ &\qquad = -R\big[m\circ(m\otimes\mathrm{id})\circ(\mathrm{id}\otimes S\otimes\mathrm{id})\circ(P_2\otimes\mathrm{id})\circ(\Delta\otimes\mathrm{id})\circ P_2\Delta[X]\big]\;. \end{aligned}$$

We have computed $(\Delta\otimes\mathrm{id})\circ P_2\Delta[X]$ in (A.6). By inspection of that formula we find that $(P_2\otimes\mathrm{id})(\Delta\otimes\mathrm{id})\circ P_2\Delta[X]$ equals $(\Delta\otimes\mathrm{id})\circ P_2\Delta[X]-$ (A.3), which gives

$$\begin{aligned} & -R\big[m\circ(\mathrm{id}\otimes m)\circ(\mathrm{id}\otimes S\otimes\mathrm{id})\circ(\mathrm{id}\otimes P_2\Delta)\circ P_2\Delta[X]\big]\\ &\quad = -R\Big[\sum_T m\Big[\prod_{j\in T}\{m\circ(\mathrm{id}\otimes S)\circ\Delta[R[X_j]]\}\otimes X/\prod_{j\in T}X_j\Big]\Big]\\ &\quad + R\Big[\sum_T m\Big[\prod_{j\in T}S[[R[X_j]]]\otimes X/\prod_{j\in T}X_j\Big]\Big] + R\Big[\sum_T m\Big[\prod_{j\in T}[R[X_j]]\otimes X/\prod_{j\in T}X_j\Big]\Big]\;. \end{aligned}$$

The last term cancels $-R[m[P_2\Delta[X]]]$ in (A.10) and the middle term cancels $-R[m\circ(S\otimes\mathrm{id})\circ P_2\Delta[X]]$ in (A.9). We end up with the same problem as before, to calculate $m\circ(\mathrm{id}\otimes S)\circ\Delta[R[X_i]]$, however, these $X_i$ are *smaller* than the original $X$. This leads to an iteration which stops if $X_i$ is primitive, and for primitive $X_i$ we have

$$\begin{aligned} m\circ(\mathrm{id}\otimes S)\circ\Delta[R[X_i]] &= m\circ(\mathrm{id}\otimes S)\circ(e\otimes R[X_i]+R[X_i]\otimes e)\\ &= S[R[X_i]]+R[X_i] = 0\;. \end{aligned}$$

The conclusion is $m\circ(\mathrm{id}\otimes S)\circ\Delta[R[X]] = 0 = E\circ\bar{e}[R[X]]$ for all iPW $X$.

It remains to apply $m\circ(\mathrm{id}\otimes S)\circ\Delta$ and $m\circ(S\otimes\mathrm{id})\circ\Delta$ to disjoint products $\prod_i X_i\prod_j R[Y_j]$. Here we have the multiplicativity of $\Delta$ (23) and $S$ (25b) at disposal. The result is zero if at least one $R[Y_j]$ is present. Otherwise we have

$$\begin{aligned} m\circ(\mathrm{id}\otimes S)\circ\Delta[\prod_i X_i] &= \prod_i\{R[X_i]-X_i\}\\ &\sim R\big[\prod_i\{R[X_i]-X_i\}\big]\approx 0 = E\circ\bar{e}[\prod_i X_i]\;, \qquad (A.11)\\ m\circ(S\otimes id)\circ\Delta[\prod_i X_i] &= \prod_i\big\{(\mathrm{id}-R)\big[X_i+m\circ(S\otimes\mathrm{id})\circ P_2\Delta[X_i]\big]\big\}\\ &\sim 0 = E\circ\bar{e}[\prod_i X_i]\;. \qquad (A.12) \end{aligned}$$

In (A.11) note that $R[X_i]-X_i$ is *divergent* in general and that the multiplication of $R[X_j]-X_j$ by a divergent term is not equivalent to zero. Equivalence follows for renormalization schemes satisfying (17). In (A.12) however, $(\mathrm{id}-R)\big[X_i + m\circ(S\otimes\mathrm{id})\circ P_2\Delta[X_i]\big]$ is *convergent as it reproduces the forest formula*, see (27). Now, multiplication of $(\mathrm{id}-R)\big[X_i+m\circ(S\otimes\mathrm{id})\circ P_2\Delta[X_i]\big]$ by a convergent term is equivalent to zero. It is even strongly equivalent ($\simeq$) to zero which means that the integral is *finite*. The fact that $m\circ(S\otimes id)\circ\Delta$ gives the forest formula is essential for $S$ being a left antipode in any renormalization scheme. □


## Acknowledgments

We are grateful to Dirk Kreimer for explaining us the way he treats overlapping divergences. We would like to thank Bruno Iochum, Serge Lazzarini and Thomas Schücker for discussions.


## References


[1] D. Kreimer, *On the Hopf algebra structure of perturbative quantum field theories*, Adv. Theor. Math. Phys. **2** (1998), and `q-alg/9707029`.

[2] D. Kreimer, *Renormalization and Knot Theory*, J. Knot Th. Ram. **6** (1997) 479–581, and `q-alg/9607022`.

[3] C. Itzykson and J. B. Zuber, *Quantum Field Theory*, McGraw-Hill (1980)

[4] W. Zimmermann, *Convergence of Bogoliubov's method of renormalization in momentum space*, Commun. Math. Phys. **15** (1969) 208–234.

[5] N. N. Bogoliubov and O. S. Parasiuk, *On the multiplication of the causal function in the quantum theory of fields*, Acta Math. **97** (1957) 227–266.

[6] K. Hepp, *Proof of the Bogoliubov-Parasiuk theorem on renormalization*, Commun. Math. Phys. **2** (1966) 301–326.

[7] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, J. Wiley (1980).

[8] A. Connes and H. Moscovici, *Hopf algebras, cyclic cohomology and the transverse index theorem*, Vietri conference on noncommutative geometry, March 1998.

[9] A. Connes and D. Kreimer, in preparation