Diagram rewriting for orthogonal matrices: a study of critical peaks *

Yves Lafont & Pierre Rannou

Institut de Mathématiques de Luminy, UMR 6206 du CNRS
Université de la Méditerranée (Aix-Marseille 2)

February 11, 2008

Abstract

Orthogonal diagrams represent decompositions of orthogonal matrices, corresponding to isometries of $\mathbb{R}^n$, into elementary ones: 1-dimensional symmetries and 2-dimensional rotations. A convergent rewrite system for this structure was introduced by the first author.

One of the rules, which is similar to the Yang-Baxter equation, involves a map $h : [0, \pi[^3 \to [0, \pi[^3$. To study the algebraic properties of h, we use the confluence of critical peaks in our rewrite system, and we introduce *parametric diagrams* describing the calculation of angles of rotations generated by rewriting. In particular, h satisfies the *tetrahedron equation* (also called Zamolodchikov equation).

1 Introduction

Diagrams are widely used for computation in various fields of mathematics and physics, like category theory, knot theory, proof theory, quantum electrodynamics and relativity. Formally, a diagram is an element of the free 2-*monoid* (or strict monoidal category) generated by some 2-*computad* (or 2-*polygraph*). See [Pow91, Bur93].

Typical examples are *boolean circuits* and *quantum circuits*, which are respectively interpreted in the 2-monoid of finite sets with Cartesian product and in the 2-monoid of finite dimensional (complex) vector spaces with tensor product. See [Laf03, Ran07]. Here we study the case of *orthogonal diagrams*, which are interpreted in the 2-monoid of finite dimensional (real) vector spaces with direct sum.

The starting point is the classical decomposition of rotations in $\mathbb{R}^3$ (*Euler angles*), which can be generalized to a decomposition of isometries in $\mathbb{R}^n$. From this, we get a convergent rewrite system. In that case, critical peaks are not used to prove confluence, since it holds by construction, but to derive some algebraic properties of the rules.

Here are some motivations for such a study: First, similar rules appear in the theory of quantum circuits, for which no complete presentation is known. Moreover, we must study enough examples in order to get a general theory of confluence for diagram rewriting. Finally, there are interesting connections between rewriting and homology, which should extend to diagram rewriting. See [Laf07] for a survey.

*This work is partially supported by ANR project *Invariants algébriques des systèmes informatiques*.

2 Rotations of $\mathbb{R}^3$

The matrix of an isometry is an orthogonal matrix. Such a matrix has determinant ± 1. If the determinant is 1, it corresponds to a rotation. In particular, the following matrices correspond to rotations of respective axes Ox and Oz in $\mathbb{R}^3$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix}, \quad \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Theorem 1 *(Euler angles)*

Any rotation of $\mathbb{R}^3$ can be decomposed into three rotations, possibly preceded by an axial symmetry, where:

- *the axes of the three rotations are, in order, Ox, Oz, and Ox;*
- *the angles of the three rotations are in $[0, \pi[$;*
- *the axis of the symmetry is Ox, Oy, or Oz.*

Moreover, this decomposition is unique if the second angle is $\neq 0$. Otherwise, we get a single rotation of axis Ox, possibly preceded by an axial symmetry.

We get the following decomposition for a rotation matrix:

$$R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{pmatrix} \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} u & 0 & 0 \\ 0 & v & 0 \\ 0 & 0 & w \end{pmatrix}$$

where $\alpha, \beta, \gamma \in [0, \pi[$, $u, v, w = \pm 1$, $uvw = \det R = 1$, and $\gamma = 0$ whenever $\beta = 0$.

This is called a *(left) canonical decomposition*. It is *standard* if the angles of the three rotations are $\neq 0$ and if there is no axial symmetry. In that case, α, β, γ are called the *Euler angles* of the rotation.

By exchanging Ox and Oz, we get the notions of *right canonical decomposition* and of *right standard decomposition*.

Lemma 1 *The canonical decomposition of a rotation matrix is standard if and only if its lower left coefficient and its upper right coefficient are positive.*

Proof: The decomposition is standard if and only if $\alpha, \beta, \gamma \neq 0$ and $u = v = w = 1$. Moreover, the lower left coefficient and the upper right coefficient are:

$$\begin{aligned} a &= u \sin\gamma \sin\beta, \\ b &= w \sin\beta \sin\alpha. \end{aligned}$$

Since $\alpha, \beta, \gamma \in [0, \pi[$, we have $a, b > 0$ if and only if $\alpha, \beta, \gamma \neq 0$ and $u = v = w = 1$.

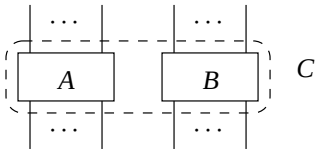
Corollary 1 *The left canonical decomposition of a rotation matrix is standard if and only if its right canonical decomposition is standard.*

3 Orthogonal diagrams

We introduce orthogonal diagrams. A diagram on n wires is interpreted as an isometry of $\mathbb{R}^n = \mathbb{R} \oplus \cdots \oplus \mathbb{R}$, or equivalently, as an orthogonal $n \times n$ matrix. The gates represent elementary isometries in low dimension.

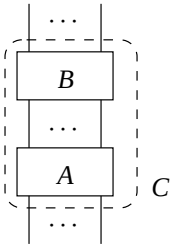
Compositions of diagrams are interpreted as follows:

- Let A and B be diagrams respectively with n and m wires, interpreted by orthogonal matrices M_A and M_B. Their parallel composition is the following diagram:



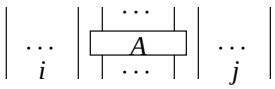
  It is interpreted by the matrix $M_C = M_A \oplus M_B = \left(\begin{array}{c|c} M_A & 0 \\ \hline 0 & M_B \end{array}\right)$.

- If $n = m$, the sequential composition of A and B is the following diagram:

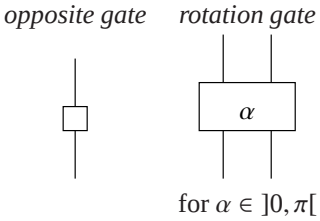


  It is interpreted by the matrix $M_C = M_A M_B$.

Remark: The identity on $\mathbb{R}$ is represented by a wire. Hence, the matrix $Id_i \oplus M_A \oplus Id_j$ is represented by the following diagram:



There are two kinds of gates:

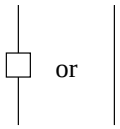


The first one is interpreted by the scalar -1 and the second one by the matrix

$$R_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

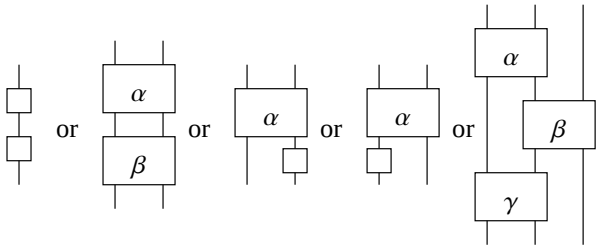
Definition 1 *Canonical diagrams are defined by induction on the number of wires:*

- *A canonical diagram on 1 wire is:*



- *If $n > 0$, the general form of a canonical diagram on n wires is given in figure 1, where C_1 and C_{n-1} are canonical diagrams respectively on 1 and $n-1$ wires, and $0 \leq k \leq n-1$.*

Remark: An orthogonal diagram is canonical if and only if it contains no sub-diagram of the following form:



Theorem 1 can be generalized as follows:

Theorem 2 *Any isometry of $\mathbb{R}^n$ can be represented by a unique canonical diagram.*

Consider the rewriting rules of figure 2. The last one transforms a right standard form into a left standard form with the same interpretation: The angles α', β', γ' are given by corollary 1, which asserts the existence and the uniqueness of this left standard form. There are some complicated formulas for those angles, but we shall not use them.

Remark: If a diagram D reduces to D', then D and D' have the same interpretation.

For the last rule, this holds by construction. The other cases are obvious. For instance, the second, the fourth and the fifth rule correspond to the following identities:

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos(\pi-\alpha) & -\sin(\pi-\alpha) \\ \sin(\pi-\alpha) & \cos(\pi-\alpha) \end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} \cos(\pi-\alpha) & -\sin(\pi-\alpha) \\ \sin(\pi-\alpha) & \cos(\pi-\alpha) \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{pmatrix}$$

Lemma 2 *The system is noetherian.*

Proof: For each diagram on n wires we define a vector $(p_1, \ldots, p_{n-1}, q)$ as follows:

- p_i is the number of occurrences of binary gates having their left input on wire i;
- $q = \Sigma_{A\in\Xi}(f(A)+1)$, where Ξ is the set of occurrences of unary gates and $f(A)$ is the number of occurrences of binaries gate above A.

Then, one checks that each rule makes this vector decrease for the lexicographic order. Since this order is well founded, we are done.

Figure 1: General form of a canonical diagram

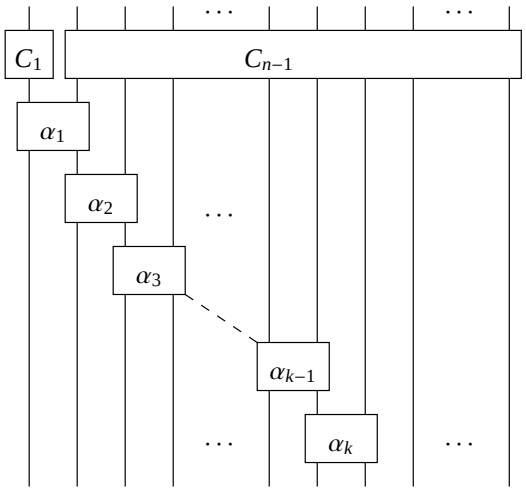
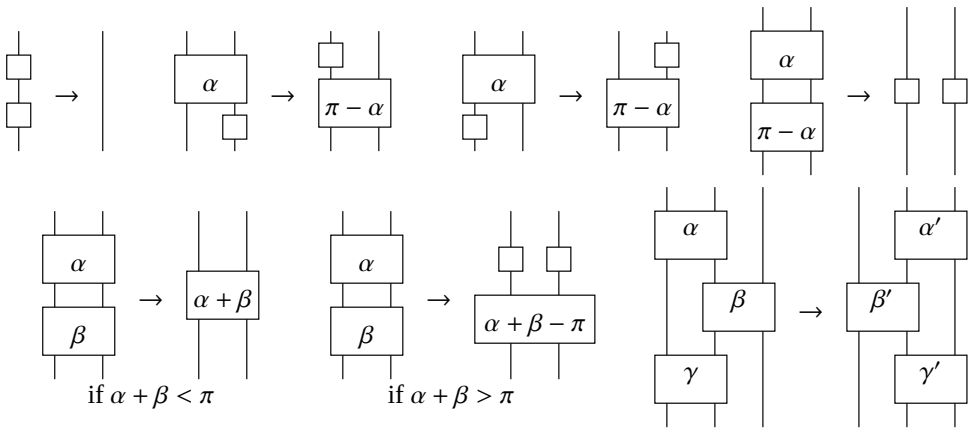


Figure 2: Rules for orthogonal diagrams



Lemma 3 *Every orthogonal diagram reduces to a unique canonical diagram.*

Proof: By double induction on the number of wires and the number of gates.

Consider a diagram $D_{n,m}$ with n wires and m gates and let A be its last gate. In other words, $D_{n,m}$ consists of some diagram $D_{n,m-1}$ followed by A.

By induction hypothesis, $D_{n,m-1}$ reduces to a canonical diagram $D'_{n,m'-1}$. Hence, $D_{n,m}$ reduces to $D'_{n,m'}$ which consists of $D'_{n,m'-1}$ followed by A. It remains to reduce $D'_{n,m'}$. There are several configurations, depending on the type and the position of A:

- if A is an opposite gate, there are three cases: see figure 3;
- if A is a rotation, there are four cases: see figure 4.

After reduction, we obtain a new diagram, where some unary gate may appear just below C_1 and some (unary or binary) gate may appear just below C_{n-1}. The first one can always be eliminated using the first rule, and the second one can be eliminated by applying the induction hypothesis for $n-1$ wires.

Uniqueness follows from theorem 2 and the remark.

To sum up, we get the following result:

Theorem 3 *The system is convergent. In other words, it is noetherian and confluent.*

4 Critical peaks

Let P be a set. Assume we have a partition $P^2 = \Delta^0 \cup \Delta^- \cup \Delta^+$ and the following maps:

- $f : P \to P$ such that $f(\alpha) = \beta$ if and only if $(\alpha, \beta) \in \Delta^0$;
- $g_- : \Delta^- \to P$ and $g_+ : \Delta^+ \to P$;
- $h : P^3 \to P^3$ given by $h(\alpha, \beta, \gamma) = (h_1(\alpha, \beta, \gamma), h_2(\alpha, \beta, \gamma), h_3(\alpha, \beta, \gamma))$.

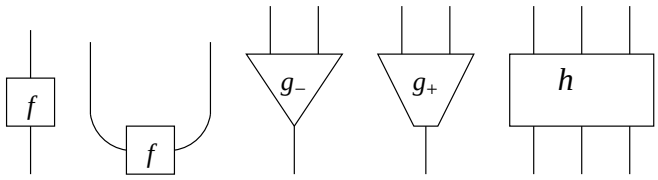
We write $\alpha \mid \beta$ if $(\alpha, \beta) \in \Delta^0$, $\alpha \smile \beta$ if $(\alpha, \beta) \in \Delta^-$, and $\alpha \frown \beta$ if $(\alpha, \beta) \in \Delta^+$.

We are mainly interested in the case where $P =]0, \pi[$ and:

- $\alpha \mid \beta$ if $\alpha + \beta = \pi$, $\alpha \smile \beta$ if $\alpha + \beta < \pi$, and $\alpha \frown \beta$ if $\alpha + \beta > \pi$;
- $f(\alpha) = \pi - \alpha$, $g_-(\alpha, \beta) = \alpha + \beta$, and $g_+(\alpha, \beta) = \alpha + \beta - \pi$;
- h corresponds to the last rule of figure 2.

We consider *generalized orthogonal diagrams* with parameters $\alpha \in P$, and the rewrite system $\mathcal{H}$ of figure 5.

In order to represent calculations on parameters, we also need the following gates for *parametric diagrams*:



Each gate is interpreted by the corresponding map, which may be partial: In particular, the second one is interpreted by the predicate $\alpha \mid \beta$. Since each one has a distinct shape, we shall omit labels in parametric diagrams.

Figure 3: Configurations for a unary gate

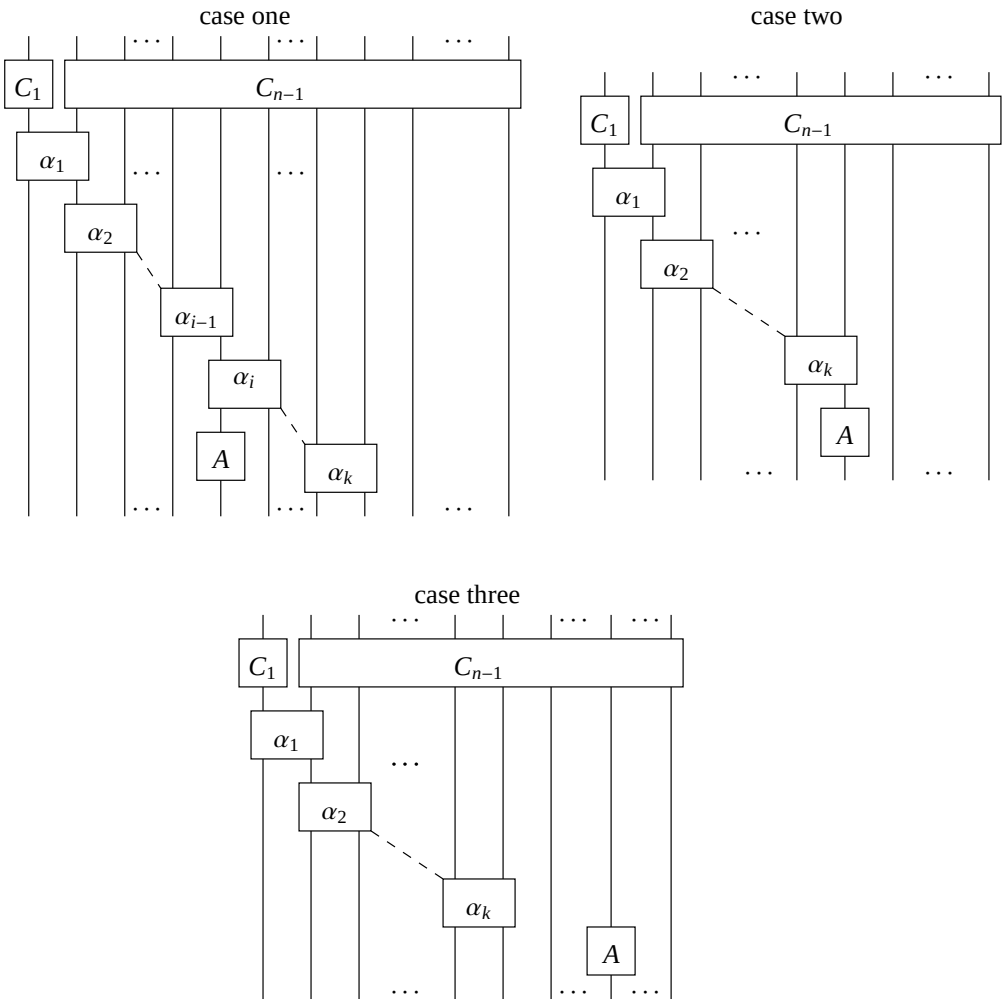
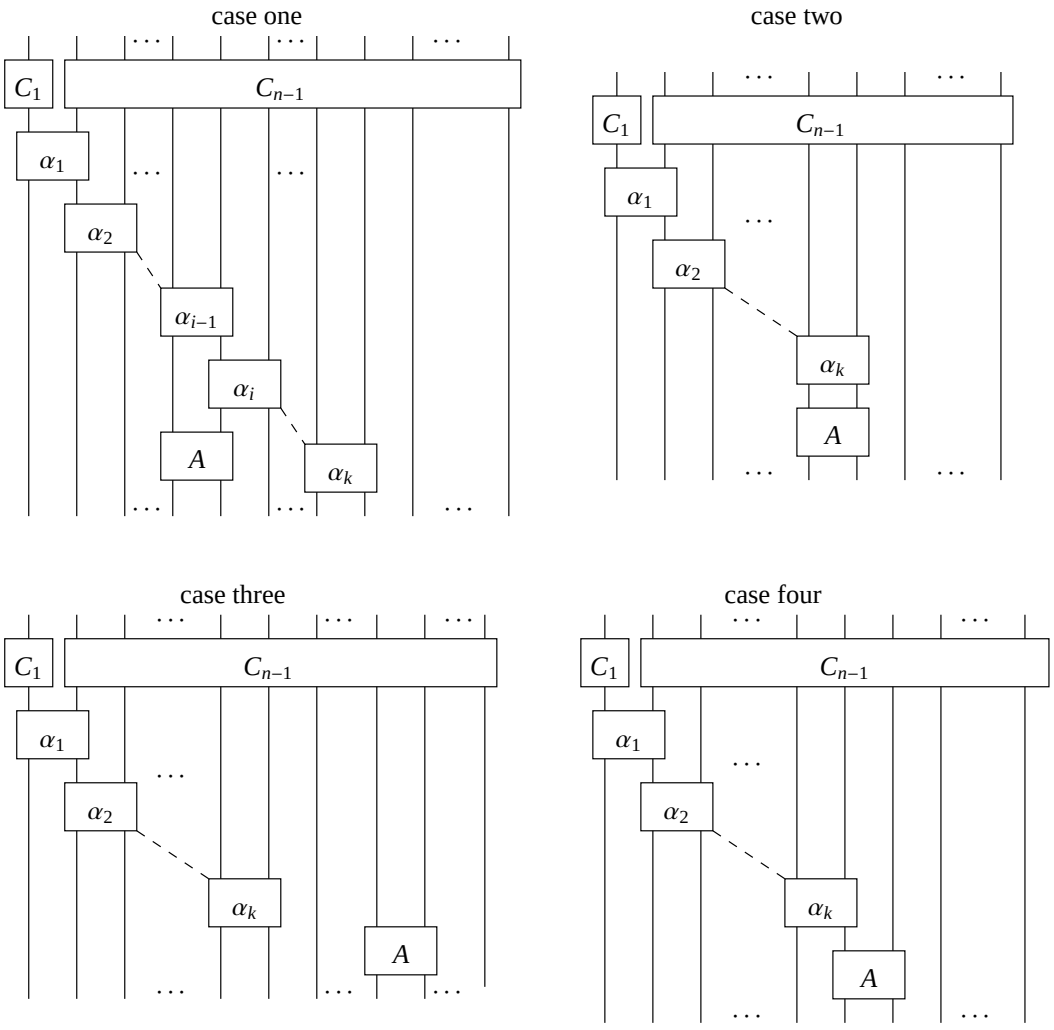


Figure 4: Configurations for a rotation



Lemma 4 *The system $\mathcal{H}$ is noetherian.*

Proof: By the same argument as for lemma 2.

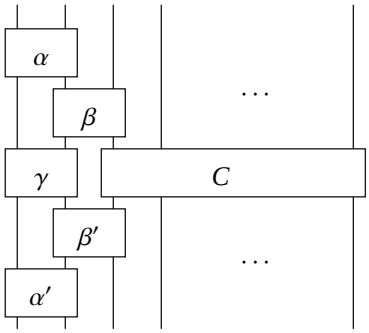
Theorem 4 *The following statements are equivalent:*

1. *$\mathcal{H}$ is confluent;*
2. *the* critical peaks *of figure 6 are confluent;*
3. *the four maps satisfy the identities of figure 7.*

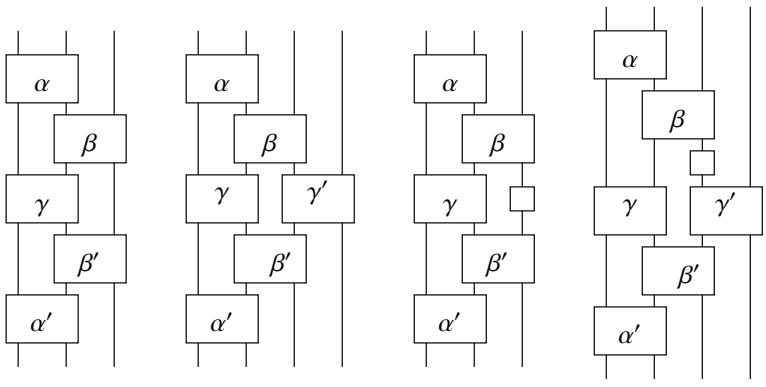
Proof: Obviously, we have $1 \Rightarrow 2$.

Conversely, assume D reduces in one step to D' and to D''. If the rules apply to disjoint subdiagrams of D, then D' and D'' reduce to a common diagram in one step and we are done. Otherwise, we have a conflict and there are two cases:

- if one of the rules of the conflict is not ternary, the conflict appears in figure 6;
- if both rules are ternary, we get a *global conflict* of the following form:



Here, C stands for an arbitrary diagram, but in fact, it suffices to consider the cases where C is a normal form: See appendix A of [Laf03]. Therefore, we get four cases:



The first two cases appear in figure 6, whereas the two other ones can be decomposed into simpler conflicts. Hence, we get $2 \Rightarrow 1$.

Now, assume that each critical peak of figure 6 is confluent. This means that we have two reductions leading to a common diagram. Each one yields a parametric diagram

Figure 5: Rules for generalized orthogonal diagrams

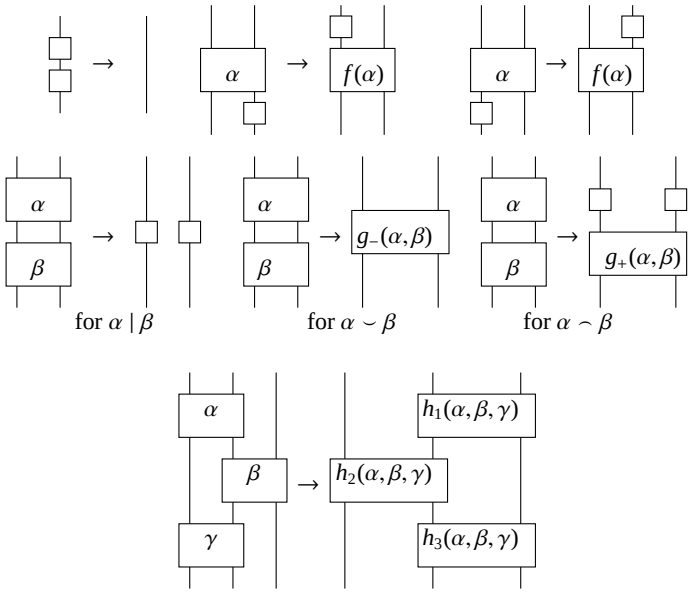


Figure 6: Critical peaks

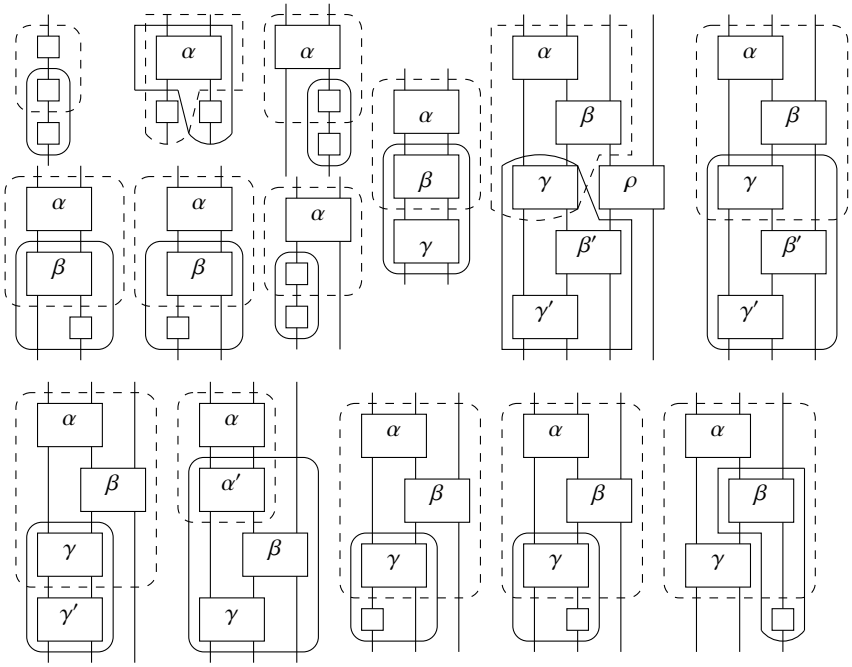
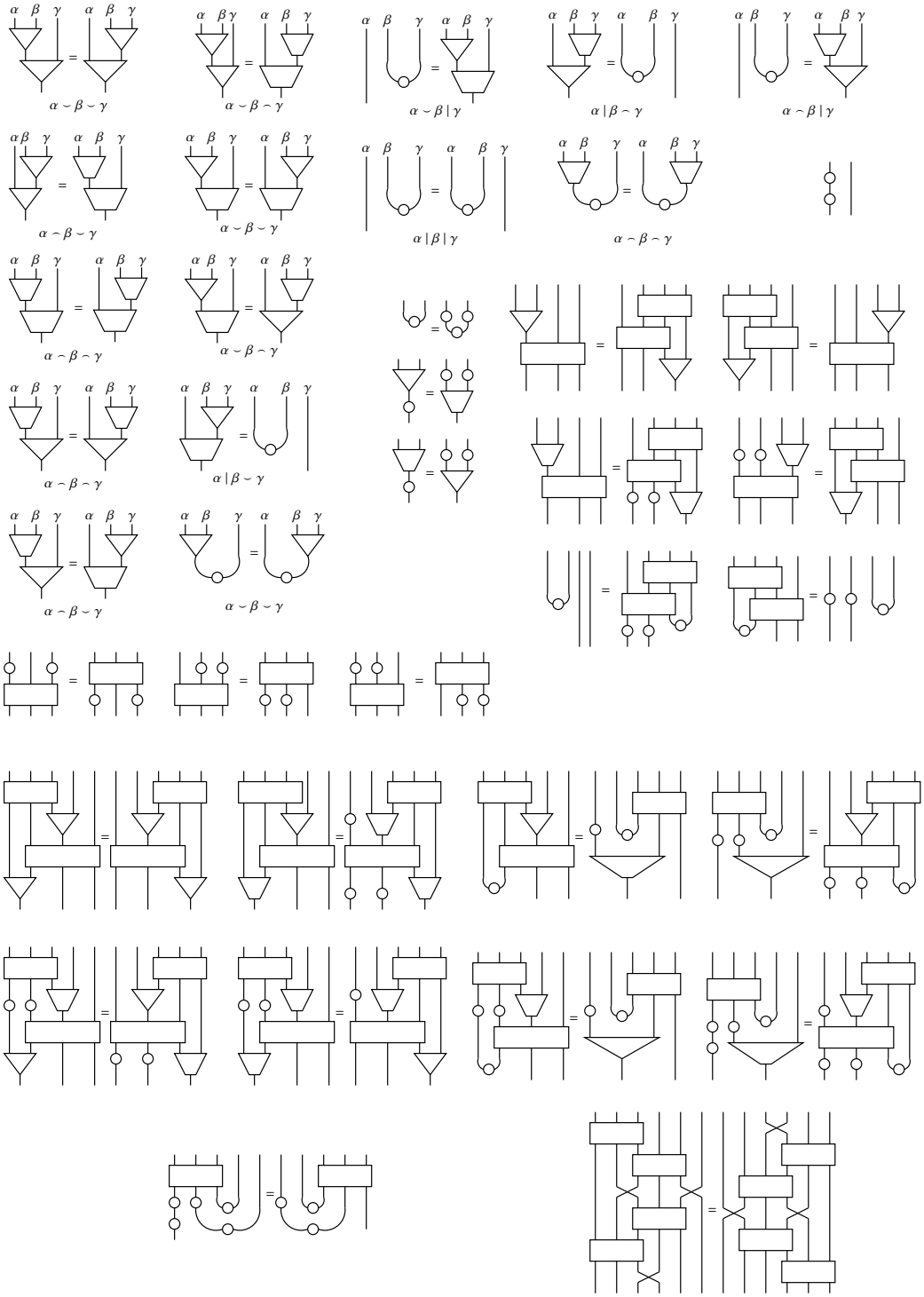


Figure 7: Parametric identities



(which is empty in few cases) representing calculation on parameters, and both calculations must give the same result. Hence, we get an identity between two parametric diagrams. Three examples are given in figure 8: In each case, calculations have been drawn over reductions to show how the corresponding parametric diagrams are built.

In fact, some critical peaks of figure 6 yield several identities because there are several cases to consider according to the conditions satisfied by the parameters. By a careful analysis, we obtain the list of identities of figure 7. So we get $2 \Leftrightarrow 3$ and we are done.

Note that half of the identities of figure 7 do not involve the ternary gate corresponding to the map h. All those identities are trivially satisfied in the case of orthogonal diagrams: For instance, the first one corresponds to the associativity of $+$. We are mainly interested in the other identities, which express properties of our map h. The last one is known as the *tetrahedron equation* (or *Zamolodchikov equation*).

5 Conclusion

We used critical peaks to study the properties of h, for which we obtained a list of 19 identities. However, there are many redundancies in this list. In a future work, we shall explain how this list can be reduced to a shorter one expressed in terms of *undirected parametric diagrams*.

It is also important to notice that diagram rewriting is more complicated than word or term rewriting. See [Gui06] for a general theory of termination. A general theory of confluence including the notion of global conflict should also be developed.

References

[Bur93] Albert Burroni. Higher-dimensional word problems with applications to equational logic. *Theoretical Computer Science*, 115:43–62, 1993.

[Gui06] Yves Guiraud. Termination orders for 3-dimensional rewriting. *Journal of Pure and Applied Algebra*, 207(2):341–371, 2006.

[Laf03] Yves Lafont. Towards an algebraic theory of boolean circuits. *Journal of Pure and Applied Algebra*, 184:257–310, 2003.

[Laf07] Yves Lafont. Algebra and geometry of rewriting. *Applied Categorical Structures*, 15:415–437, 2007.

[Pow91] John Power. An n-categorical pasting theorem. In *Category Theory, Proc. Int. Conf., Como/Italy 1990*, number 1488 in Lect. Notes Math., pages 326–358, 1991.

[Ran07] Pierre Rannou. Théorie algébrique des circuits quantiques, circuits orthogonaux, et circuits paramétriques. Master thesis, 2007.

Figure 8: Confluence of 3 critical peaks