Towards an Algebraic Theory of Boolean Circuits

Yves Lafont[*]

February 12, 2003

Abstract

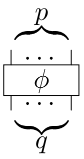
Boolean circuits are used to represent programs on finite data. *Reversible Boolean circuits* and *quantum Boolean circuits* have been introduced to modelize some physical aspects of computation. Those notions are essential in complexity theory, but we claim that a deep mathematical theory is needed to make progress in this area. For that purpose, the recent developments of *knot theory* is a major source of inspiration.

Following the ideas of Burroni, we consider logical gates as *generators* for some algebraic structure with two compositions, and we are interested in the *relations* satisfied by those generators. For that purpose, we introduce *canonical forms* and *rewriting systems.* Up to now, we have mainly studied the *basic case* and the *linear case*, but we hope that our methods can be used to get presentations by generators and relations for the (reversible) *classical case* and for the (unitary) *quantum case.*

Keywords: boolean circuit; reversible gate; monoidal category; presentation by generators and relations; canonical form; rewriting; symmetric group; alternating group; linear group; orthogonal group.

1 Introduction

We use *diagrams* to represent certain kinds of maps. If p and q are natural numbers, $\phi : p \to q$ stands for a diagram with p *inputs* and q *outputs.* It is pictured as follows:

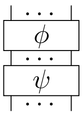


Typically, such a diagram represents:

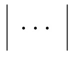
- a map from $\{1, \dots, p\}$ to $\{1, \dots, q\}$ (*basic case*);
- a map from X^p to X^q, where X is a given set (*classical case*);
- a K-linear map from K^p to K^q, where K is a given field (*linear case*);
- a K-linear map from $\bigotimes^p V$ to $\bigotimes^q V$, where V is a given vector space over a field K, and $\bigotimes^n V$ stands for the n-ary tensor product $V \otimes \cdots \otimes V$ (*quantum case*).

The basic case corresponds to *control flow diagrams* and the classical case to *data flow diagrams.*

Diagrams may be composed in two different ways. For any $\phi : p \to q$ and $\psi : q \to r$, we have a diagram $\psi \circ \phi : p \to r$, which corresponds to the usual composition of maps, and which is pictured as follows:

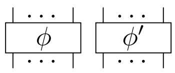


This *vertical* (or *sequential*) composition is associative, and we have an *identity diagram* $\mathrm{id}_p : p \to p$ for each p, such that $\phi \circ \mathrm{id}_p = \phi = \mathrm{id}_q \circ \phi$ for any $\phi : p \to q$. This id_p is pictured as follows:



[*] Université de la Méditerranée & Institut de Mathématiques de Luminy, UPR 9016 du CNRS, 163 avenue de Luminy, case 930, 13288 Marseille Cedex 9, France. E-mail: lafont@iml.univ-mrs.fr

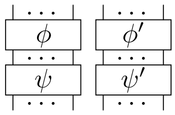
For any $\phi : p \to q$ and $\phi' : p' \to q'$, we have a diagram $\phi \,|\, \phi' : p + p' \to q + q'$ which is pictured as follows:



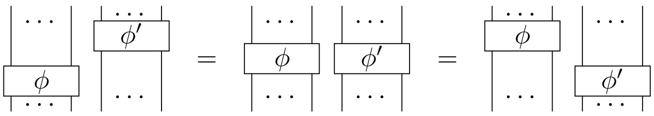
If ϕ represents f and ϕ' represents f', the interpretation g of $\phi \,|\, \phi'$ depends on the considered case:

- in the basic case, g is the *disjoint union* (or *coproduct*) $f \uplus f'$ defined by $g(i) = f(i)$ for $i = 1, \ldots, p$ and $g(p + i) = q + f'(i)$ for $i = 1, \ldots, p'$;
- in the classical case, g is the *cartesian product* $f \times f'$ defined by $g(x_1, \ldots, x_{p+p'}) = (y_1, \ldots, y_{q+q'})$ where $(y_1, \ldots, y_q) = f(x_1, \ldots, x_p)$ and $(y_{q+1}, \ldots, y_{q+q'}) = f'(x_{p+1}, \ldots, x_{p+p'})$;
- in the linear case, g is the *direct sum* $f \oplus f'$ defined by $g(u \oplus u') = f(u) \oplus f'(u')$ for $u \in K^p$ and $u' \in K^{p'}$. Note that g coincides with the cartesian product $f \times f'$;
- in the quantum case, g is the *tensor product* $f \otimes f'$ defined by $g(u \otimes u') = f(u) \otimes f'(u')$ for $u \in \bigotimes^p V$ and $u' \in \bigotimes^{p'} V$.

This *horizontal* (or *parallel*) composition is associative, and the *void diagram* $\mathrm{id}_0 : 0 \to 0$ is such that $\phi \,|\, \mathrm{id}_0 = \phi = \mathrm{id}_0 \,|\, \phi$ for any $\phi : p \to q$. Furthermore, we have $\mathrm{id}_p \,|\, \mathrm{id}_{p'} = \mathrm{id}_{p+p'}$, and the two compositions are compatible in the following sense: for any $\phi : p \to q$, $\psi : q \to r$, $\phi' : p' \to q'$, and $\psi' : q' \to r'$, we have $(\psi \circ \phi) \,|\, (\psi' \circ \phi') = (\psi \,|\, \psi') \circ (\phi \,|\, \phi')$. This diagram is pictured as follows:

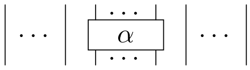


In particular, for any $\phi : p \to q$ and $\phi' : p' \to q'$, we get $(\phi \,|\, \mathrm{id}_{q'}) \circ (\mathrm{id}_p \,|\, \phi') = \phi \,|\, \phi' = (\mathrm{id}_q \,|\, \phi') \circ (\phi \,|\, \mathrm{id}_{p'})$. This corresponds to the following picture:

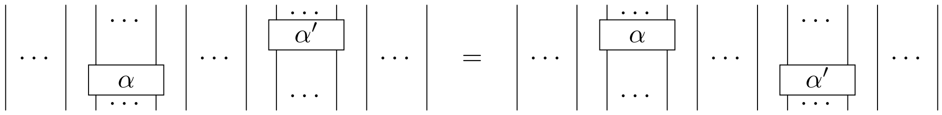


All this can be summarized as follows: the diagrams are the *morphisms* of a (strict) *monoidal category* whose *objects* are natural numbers (with addition). See [Mac71] for the notion of monoidal category. Moreover, this monoidal category is *freely generated* by a given list of atomic diagrams called *cells*. In other words, all diagrams are built from identities and cells using vertical and horizontal composition, and an equality between two diagrams holds only if it follows from the above properties.

An *elementary diagram* is a diagram ξ of the form $\mathrm{id}_i \,|\, \alpha \,|\, \mathrm{id}_j$ where α is a cell:



It is easy to see that any diagram ϕ is a vertical composition of elementary diagrams $\xi_1 \circ \cdots \circ \xi_n$, but this decomposition is not unique. In fact, two decompositions define the same diagram if and only if they are equivalent modulo the following *commutation rule*:



In particular, all decompositions of ϕ have the same length. This common length is called the *size* of ϕ: it is the total number of cells in ϕ.

Diagrams may be interpreted in any monoidal category. We have already seen four examples:

- sets with disjoint union (basic case) or with cartesian product (classical case);
- vector spaces with direct sum (linear case) or with tensor product (quantum case).

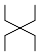
We may also consider *monoidal subcategories* obtained by restricting the class of objects or the class of morphisms. For instance, we may limit our study to *finite sets* or to *finite dimensional spaces*, to *bijective*, *surjective*, or *injective maps*, and whenever it makes sense, to *monotone*, *orthogonal*, or *unitary maps.* In this paper, we give *presentations by generators and relations* for some of those monoidal categories.

2 The basic case

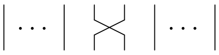
In this section, we consider the monoidal category $\mathfrak{F}$ of finite sets $\{1, \ldots, n\}$ with disjoint union, and some of its monoidal subcategories. Essentially, we follow [Laf95a], except for the case of even permutations which was not handled in that paper.

2.1 Permutations

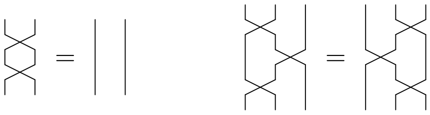
Let $\mathfrak{S}$ be the monoidal subcategory of $\mathfrak{F}$ whose morphisms are permutations. We represent the unique transposition of $\{1, 2\}$ by a cell $\tau : 2 \to 2$, which is pictured as follows:



Note that the transposition τ_i of $\{1, \ldots, n\}$ which exchanges i with $i+1$ is represented by the elementary diagram $\mathrm{id}_{i-1} \mid \tau \mid \mathrm{id}_{n-i-1}$:



Obviously, τ satisfies the relations $\tau \circ \tau = \mathrm{id}_2$ and $(\tau \mid \mathrm{id}_1) \circ (\mathrm{id}_1 \mid \tau) \circ (\tau \mid \mathrm{id}_1) = (\mathrm{id}_1 \mid \tau) \circ (\tau \mid \mathrm{id}_1) \circ (\mathrm{id}_1 \mid \tau)$, which are pictured as follows:



Of course, those equalities hold between the interpretations, not between the diagrams.

Theorem 1 *The generator τ and the above two relations form a presentation of $\mathfrak{S}$.*

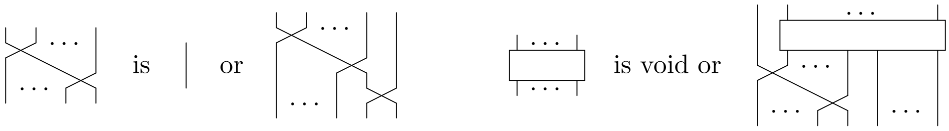
This means that τ generates $\mathfrak{S}$, and if two diagrams represent the same permutation, they are equivalent modulo the above relations.

If we restrict this presentation of the monoidal category $\mathfrak{S}$ to permutations of $\{1, \ldots, n\}$, we get the usual presentation of the *symmetric group* $\mathfrak{S}_n$ by the generators $\tau_1, \ldots, \tau_{n-1}$ and the following relations:

$$\tau_i^2 = 1, \qquad \tau_i \tau_{i+1} \tau_i = \tau_{i+1} \tau_i \tau_{i+1}, \qquad \tau_i \tau_j = \tau_j \tau_i \text{ for } j > i+1.$$

However, our approach has several advantages: we present all $\mathfrak{S}_n$ at the same time; we need one generator instead of $n-1$, and two relations instead of $n(n-1)/2$; the relations $\tau_i \tau_j = \tau_j \tau_i$ for $j > i+1$ are made implicit by the commutation rule. Furthermore, we shall see that a monoidal category such as $\mathfrak{F}$, which is not a groupoid, can also be presented in this way, even though it cannot be decomposed into monoids.

To show theorem 1, we introduce the notions of *stairs* and *canonical forms*, which are defined by the following *grammar*:

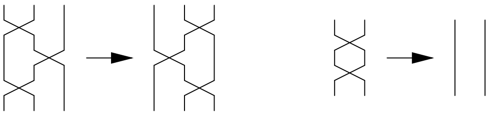


Lemma 1 *Any permutation f of $\{1, \ldots, n\}$ is represented by a unique canonical form.*

This is proved by induction on n:

- If $n = 0$, then f is the identity on the empty set, which is represented by the void diagram.
- If $n \geq 1$, we get a permutation of $\{1, \ldots, n-1\}$ by removing 1 from the domain of f and $f(1)$ from its codomain, and by renumbering both of them. In the canonical form of f, the size of the stairs is $f(1) - 1$, and the remaining part is given by the induction hypothesis.

Lemma 2 *Any diagram $\phi : n \to n$ reduces to a canonical form $\widehat{\phi}$ by the following rules:*



This is proved by double induction on n and the size m of ϕ:

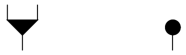
- If $m = 0$, then $\phi = \mathrm{id}_n$, which is a canonical form. This is easily checked by induction on n.
- If $m \geq 1$, then $\phi = \xi \circ \psi$ where ξ is an elementary diagram and ψ is a diagram of size $m - 1$ which reduces to a canonical form $\widehat{\psi}$ by induction hypothesis. Therefore, ϕ reduces to $\xi \circ \widehat{\psi}$ and there are four possible cases: see figure 1. In the first case, we use the first rule and the induction hypothesis for $n - 1$; in the second case, we use the second rule and we get a canonical form; in the third case, we get a canonical form; in the fourth case, we use the induction hypothesis for $n - 1$. Note that in the first and in the last case, we also need the commutation rule, which is implicit.

Now we can prove theorem 1: τ generates $\mathfrak{S}$ by lemma 1, and if two diagrams ϕ and ψ represent the same permutation, then ϕ reduces to $\widehat{\phi}$ and ψ to $\widehat{\psi}$ by lemma 2. Since all those diagrams represent the same permutation, we have $\widehat{\phi} = \widehat{\psi}$ by lemma 1, so that ϕ and ψ are equivalent modulo the relations.

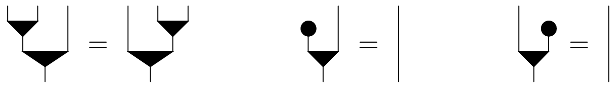
In fact, our rules form a *terminating rewrite system*: see appendix A. Since they do not increase the size of diagrams, a canonical form is always a diagram of minimal size for the map it represents.

2.2 Monotone maps

Let $\mathfrak{M}$ be the monoidal subcategory of $\mathfrak{F}$ whose morphisms are monotone maps. We introduce two generators $\mu : 2 \to 1$ and $\eta : 0 \to 1$, which are interpreted in the obvious way and pictured as follows:

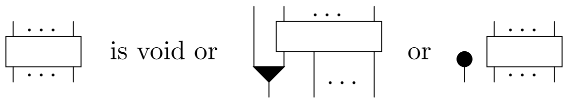


Obviously, μ and η satisfy the relations $\mu \circ (\mu \,|\, \mathrm{id}_1) = \mu \circ (\mathrm{id}_1 \,|\, \mu)$, $\mu \circ (\eta \,|\, \mathrm{id}_1) = \mathrm{id}_1$, and $\mu \circ (\mathrm{id}_1 \,|\, \eta) = \mathrm{id}_1$, which are pictured as follows:



Theorem 2 *The generators μ, η and the above three relations form a presentation of $\mathfrak{M}$.*

To show this, we introduce the following notion of canonical form:

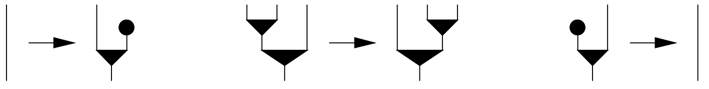


Lemma 3 *Any monotone map $f : \{1, \ldots, p\} \to \{1, \ldots, q\}$ is represented by a unique canonical form.*

This is proved by induction on $n = p + q$:

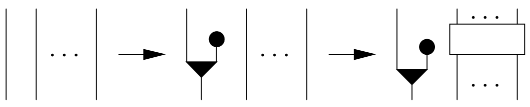
- If $n = 0$, then f is the identity on the empty set, which is represented by the void diagram.
- If $p \geq 1$ and $f(1) = 1$, we get a monotone map by removing 1 from the domain of f, and by renumbering it. The canonical form is of the second type, and the remaining part is given by the induction hypothesis.
- Otherwise, $q \geq 1$ and 1 is not in the image of f: we get a monotone map by removing 1 from the codomain of f, and by renumbering it. The canonical form is of the third type, and the remaining part is given by the induction hypothesis.

Lemma 4 *Any diagram $\phi : p \to q$ reduces to a canonical form $\widehat{\phi}$ by the following rules:*



This is proved by double induction on $n = p + q$ and the size m of ϕ:

- If $m = 0$, then $\phi = \mathrm{id}_p$, which reduces to a canonical form. This is proved by induction on p, using the first rule:



- If $m \geq 1$, then $\phi = \xi \circ \psi$ where ξ is an elementary diagram and ψ is a diagram of size $m-1$ which reduces to a canonical form $\widehat{\psi}$ by induction hypothesis. Therefore, ϕ reduces to $\xi \circ \widehat{\psi}$ and there are seven possible cases: see figure 2. In the first case, we use the second rule and the induction hypothesis for $n-1$; in the second case, we use the third rule and we get a canonical form; in the fifth case, we get a canonical form; in the other four cases, we use the induction hypothesis for $n-1$.

Theorem 2 follows from lemma 3 and lemma 4.

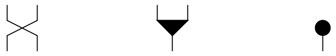
Note that identity diagrams are not in canonical form, so that the above rewrite system is not terminating. This can be fixed by adopting the alternative canonical form of figure 3, which consists in decomposing any monotone map $f : \{1, \dots, p\} \to \{1, \dots, q\}$ into a monotone surjection $g : \{1, \dots, p\} \to \{1, \dots, n\}$ followed by a monotone injection $h : \{1, \dots, n\} \to \{1, \dots, q\}$, and the terminating rewrite system of figure 4. By the way, we get:

- a presentation of the monoidal subcategory $\mathfrak{M}^{\mathrm{surj}}$ of $\mathfrak{M}$ whose morphisms are *monotone surjections* by one generator μ and the relation $\mu \circ (\mu \mid \mathrm{id}_1) = \mu \circ (\mathrm{id}_1 \mid \mu)$;

- a presentation of the monoidal subcategory $\mathfrak{M}^{\mathrm{inj}}$ of $\mathfrak{M}$ whose morphisms are *monotone injections* by one generator η and no relation.

2.3 Maps

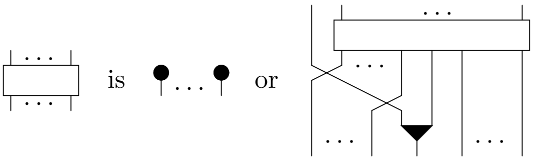
Any map $f : \{1, \dots, p\} \to \{1, \dots, q\}$ can be decomposed into a permutation $g : \{1, \dots, p\} \to \{1, \dots, p\}$ followed by a monotone map $h : \{1, \dots, p\} \to \{1, \dots, q\}$. This decomposition is not unique, but it shows that the monoidal category $\mathfrak{F}$ is generated by τ, μ, and η:



These generators satisfy the relations of figure 5, which consist of: the two relations for $\mathfrak{S}$; the first two relations for $\mathfrak{M}$; three extra relations $\tau \circ (\mu \mid \mathrm{id}_1) = (\mathrm{id}_1 \mid \mu) \circ (\tau \mid \mathrm{id}_1) \circ (\mathrm{id}_1 \mid \tau)$, $\tau \circ (\eta \mid \mathrm{id}_1) = \mathrm{id}_1 \mid \eta$, and $\mu \circ \tau = \mu$. Note that the third relation for $\mathfrak{M}$ is derivable: see figure 7.

Theorem 3 *The generators τ, μ, η and the seven relations of figure 5 form a presentation of $\mathfrak{F}$.*

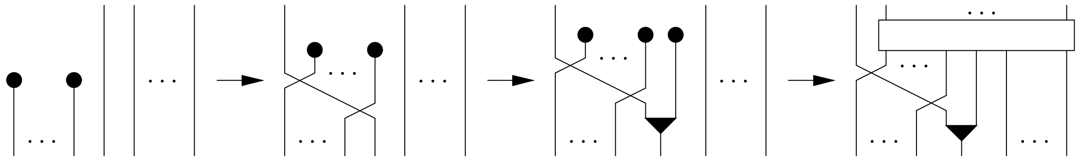
To show this, we introduce the following notion of canonical form:



It is easy to see that any map is represented by a unique canonical form. Now we introduce the rules of figure 6, which are derivable from the above relations: see figure 7. To show that any identity diagram reduces to a canonical form, we need a slightly more general result:

Lemma 5 *The diagram $\eta \mid \cdots \mid \eta \mid \mathrm{id}_p$ reduces to a canonical form.*

This is proved by induction on p, using the two reversible rules:



Theorem 3 follows from the fact that any diagram reduces to a canonical form by the rules of figure 6. In fact, it is also possible to show that our presentation is minimal: see [Mas97a].

Again, the rewrite system of figure 6 is not terminating, but we can adopt the alternative canonical form of figure 8, which consists in decomposing any map $f : \{1, \dots, p\} \to \{1, \dots, q\}$ into a surjection $g : \{1, \dots, p\} \to \{1, \dots, n\}$ followed by a monotone injection $h : \{1, \dots, n\} \to \{1, \dots, q\}$, and the terminating rewrite system of figure 9: see appendix A. By the way, we get:

- a presentation of the monoidal subcategory $\mathfrak{F}^{\mathrm{surj}}$ of $\mathfrak{F}$ whose morphisms are *surjections* by τ, μ, and the five relations of figure 5 which do not involve η;

- a presentation of the monoidal subcategory $\mathfrak{F}^{\mathrm{inj}}$ of $\mathfrak{F}$ whose morphisms are *injections* by τ, η, and the three relations of figure 5 which do not involve μ.

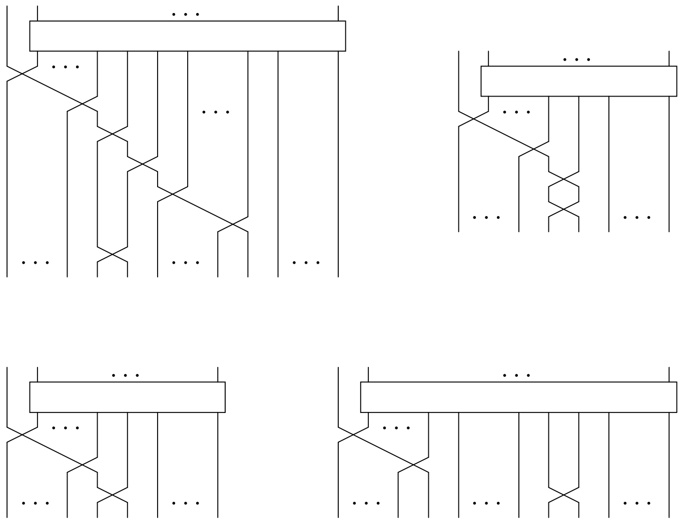


Figure 1: The four cases of lemma 2

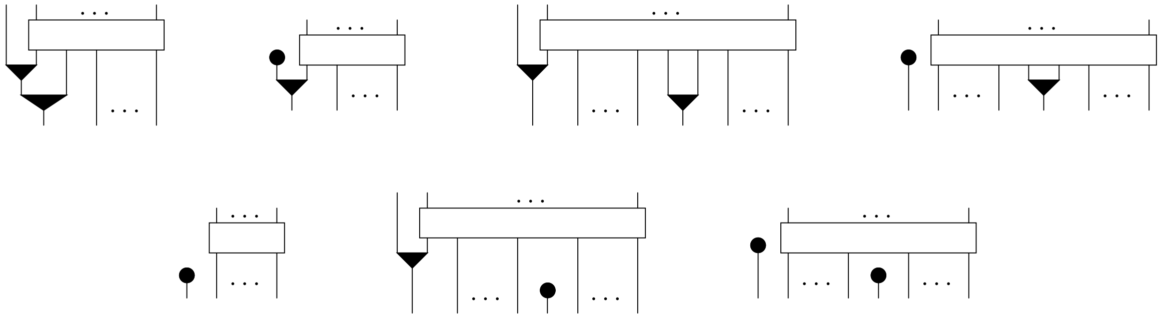


Figure 2: The seven cases of lemma 4

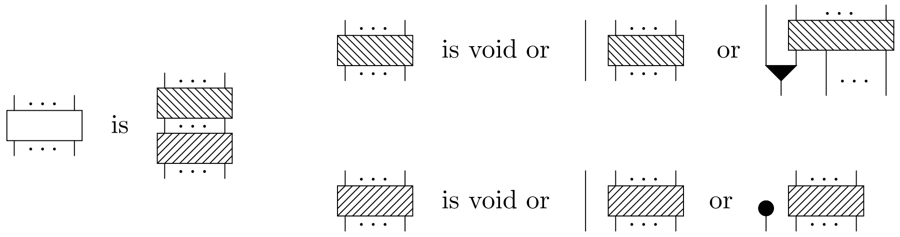


Figure 3: Alternative canonical form for $\mathfrak{M}$

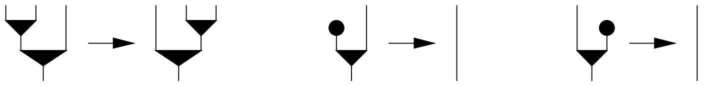


Figure 4: Alternative rules for $\mathfrak{M}$

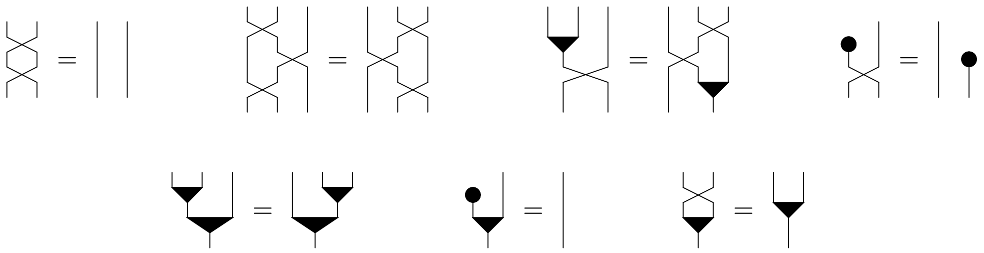


Figure 5: Relations for $\mathfrak{F}$

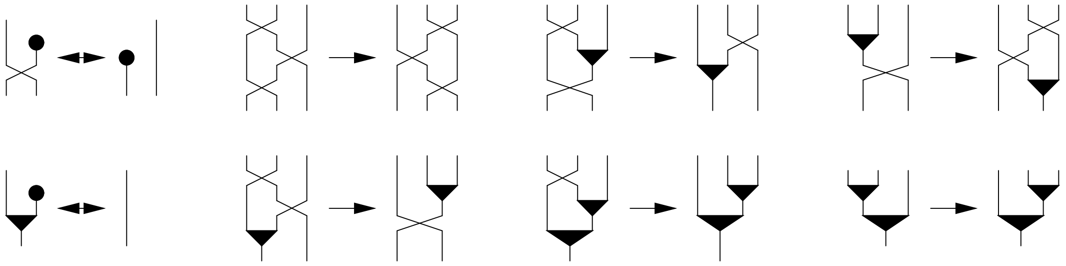


Figure 6: Rules for $\mathfrak{F}$

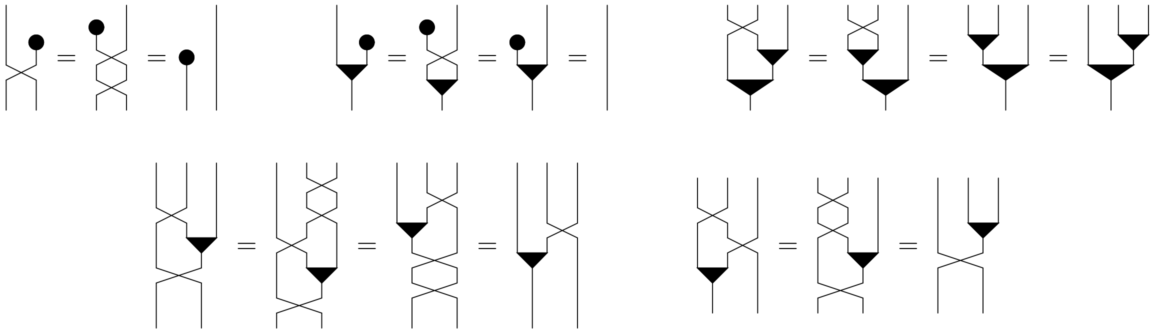


Figure 7: Deriving rules for $\mathfrak{F}$

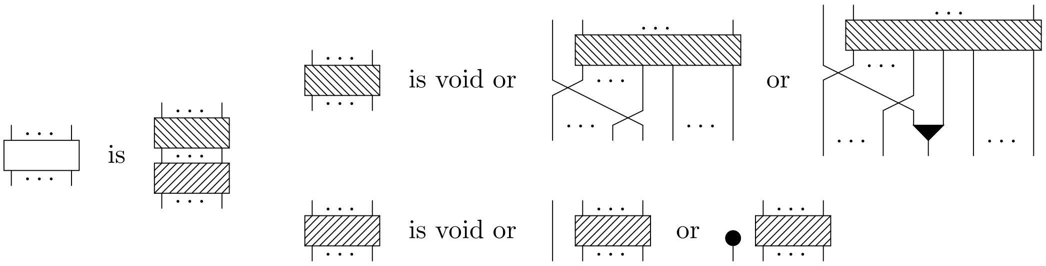


Figure 8: Alternative canonical form for $\mathfrak{F}$

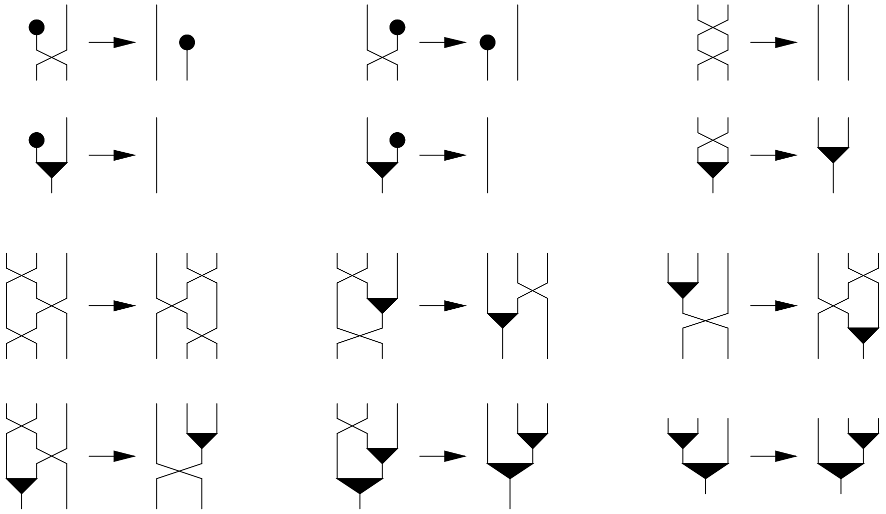
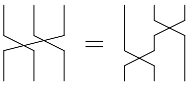


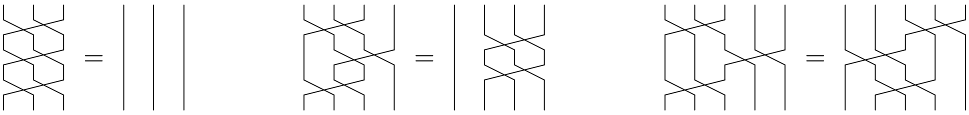
Figure 9: Alternative rules for $\mathfrak{F}$

2.4 Even permutations

Finally, we consider the monoidal subcategory $\mathfrak{A}$ of $\mathfrak{S}$ whose morphisms are *even permutations*. An even permutation is a product of an even number of transpositions. The simplest one is a cyclic permutation of order 3. So we introduce a generator $\omega : 3 \to 3$, which is pictured and defined as follows:

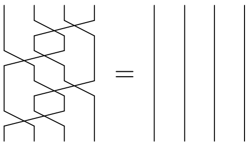


Obviously, ω satisfies the following relations:



Theorem 4 *The generator ω and the above three relations form a presentation of $\mathfrak{A}$.*

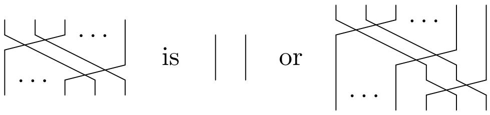
Note that the second relation can be replaced by the following one:



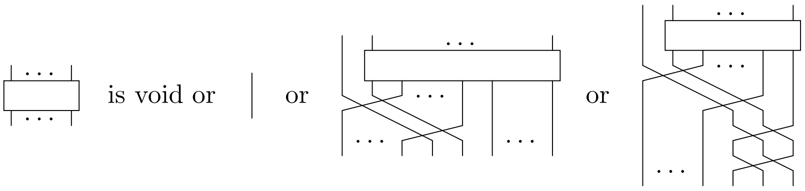
If we restrict this presentation of the monoidal category $\mathfrak{A}$ to even permutations of $\{1, \dots, n\}$, we get a presentation of the *alternating group* $\mathfrak{A}_n$ by the generators $\omega_1, \dots, \omega_{n-2}$ and the following relations:

$$\omega_i^3 = 1, \qquad (\omega_i\omega_{i+1})^2 = 1, \qquad \omega_i\omega_{i+2}\omega_i = \omega_{i+1}\omega_i\omega_{i+2}, \qquad \omega_i\omega_j = \omega_j\omega_i \text{ for } j > i + 2.$$

To show theorem 4, we introduce the following notion of *stairs*:



Canonical forms are defined as follows:



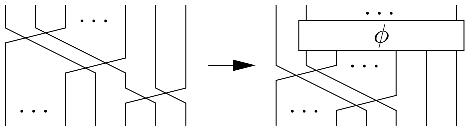
Lemma 6 *Any even permutation f of $\{1, \dots, n\}$ is represented by a unique canonical form.*

This is proved by induction on n:

- If $n \leq 2$, then f is an identity which is represented by the canonical form id_n.
- If $n > 2$ and $f(1) < n$, let $p = f(1)$ and $h = g^{-1} \circ f$ where g is the even permutation of $\{1, \dots, n\}$ defined by $g(1) = p$, $g(2) = p + 1$, $g(i) = i - 2$ for $3 \leq i \leq p + 1$, and $g(i) = i$ for $i > p + 1$. Clearly, $h(1) = 1$ so that h is of the form $1 \uplus f'$ where f' is an even permutation of $\{1, \dots, n - 1\}$. So we get a canonical form of the third type: the size of the stairs is $p - 1$ and the remaining part is given by the induction hypothesis.
- Similarly, if $n > 2$ and $f(1) = n$, we get a canonical form of the fourth type.

We introduce the rules of figure 10, which are derivable from the above relations: see figure 11.

Lemma 7 *Using the first rule of figure 10, we get the following reduction (for some diagram ϕ):*



This is proved by induction on the size of the stairs. Using this, one proves:

Lemma 8 *Any diagram $\phi : n \to n$ reduces to a canonical form $\widehat{\phi}$ by the rules of figure 10.*

Theorem 4 follows from lemma 6 and lemma 8.

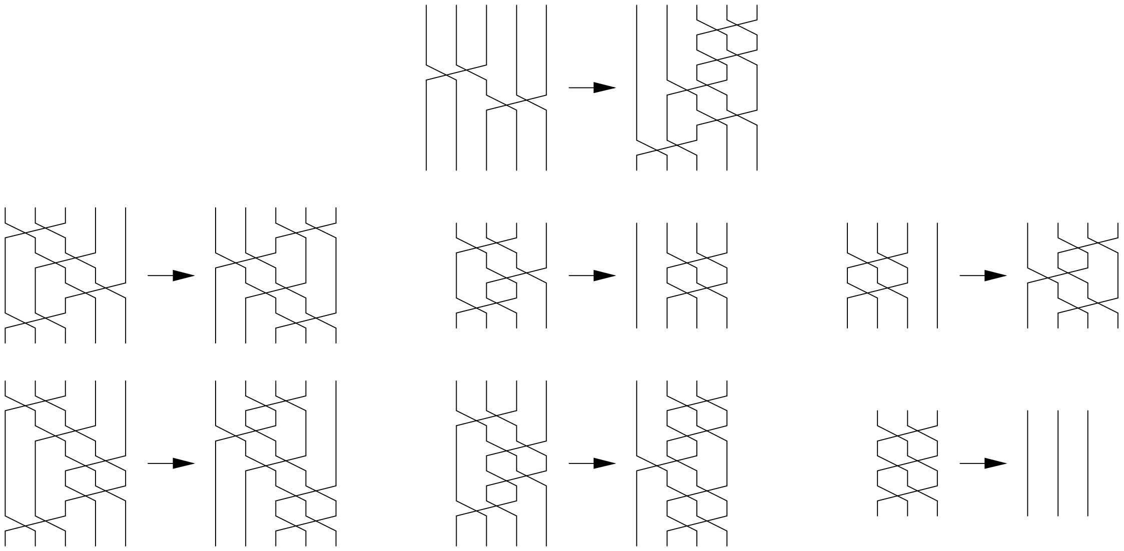


Figure 10: Rules for $\mathfrak{A}$

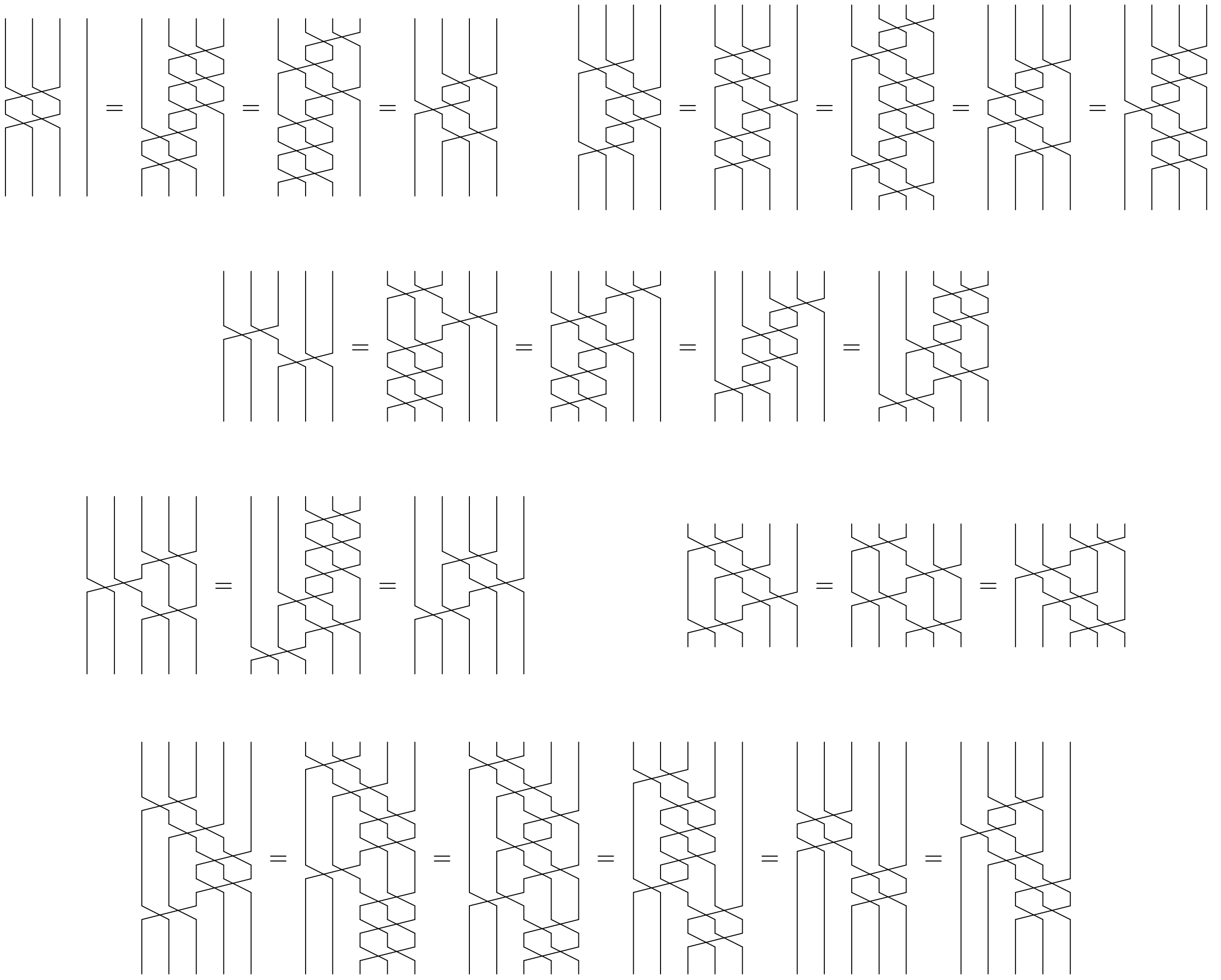


Figure 11: Deriving rules for $\mathfrak{A}$

3 The linear case

We consider the monoidal category $\mathbf{L}(K)$ of finite dimensional vector spaces K^n over a field K, with direct sum. A linear map $f : K^p \to K^q$ is represented by its *matrix*, with q *rows* and p *columns*. Vertical composition corresponds to the product of matrices, and horizontal composition to the direct sum:

$$A \oplus B = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix} \text{ where } \mathbf{0} \text{ stands for a block of } 0.$$

We write $e_1, \dots, e_n$ for the canonical basis of K^n. There are two major *symmetries* of $\mathbf{L}(K)$:

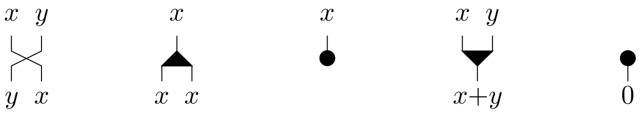
- *transposition* (or *duality*), which corresponds to a symmetry with respect to the diagonal of the matrix. Note that transposition reverts the order of vertical composition;
- *conjugacy*, which corresponds to a central symmetry of the matrix. We call it so because the conjugate of a linear permutation $f : K^n \to K^n$ is $g \circ f \circ g^{-1} = g \circ f \circ g$ where $g : K^n \to K^n$ is the linear involution defined by $g(e_i) = e_{n+1-i}$, or equivalently, by $g(x_1, \dots, x_n) = (x_n, \dots, x_1)$. Note that conjugacy reverts the order of horizontal composition.

For any map $f : \{1, \dots, p\} \to \{1, \dots, q\}$ in $\mathfrak{F}$, we define two dual linear maps:

- $f_* : K^p \to K^q$ such that $f_*(e_i) = e_{f(i)}$;
- $f^* : K^q \to K^p$ such that $f^*(x_1, \dots, x_q) = (x_{f(1)}, \dots, x_{f(p)})$.

So we get two natural *embeddings* of $\mathfrak{F}$ into $\mathbf{L}(K)$: a *covariant* one and a *contravariant* one.

We start with the field $K = \mathbb{Z}_2 = \{0, 1\}$ of integers modulo 2. It is indeed a first step towards an algebraic theory of *Boolean circuits*. We introduce five generators $\tau : 2 \to 2$, $\delta : 1 \to 2$, $\varepsilon : 1 \to 0$, $\mu : 2 \to 1$, and $\eta : 0 \to 1$, which are pictured and interpreted as follows:



This means that the matrices for τ, δ, and μ are the following:

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \end{pmatrix}$$

Note that each generator is its own conjugate and the transpose of another generator: see figure 12. These generators satisfy the relations of figure 13, which consist of: the two relations for $\mathfrak{S}$, which are self-dual; the five other relations for $\mathfrak{F}$, corresponding to the covariant embedding of $\mathfrak{F}$ into $\mathbf{L}(\mathbb{Z}_2)$; the five dual relations, corresponding to the contravariant embedding of $\mathfrak{F}$ into $\mathbf{L}(\mathbb{Z}_2)$; the four relations $\delta \circ \mu = (\mu \,|\, \mu) \circ (\mathrm{id}_1 \,|\, \tau \,|\, \mathrm{id}_1) \circ (\delta \,|\, \delta)$, $\delta \circ \eta = \eta \,|\, \eta$, $\varepsilon \circ \mu = \varepsilon \,|\, \varepsilon$, and $\varepsilon \circ \eta = \mathrm{id}_0$; one extra relation $\mu \circ \delta = \eta \circ \varepsilon$, which is specific to $\mathbb{Z}_2$. Note that four of these relations correspond to the following *axioms*:

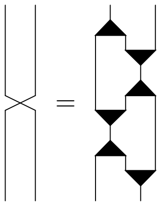
$$(x + y) + z = x + (y + z), \qquad 0 + x = x, \qquad x + y = y + x, \qquad x + x = 0.$$

Those axioms define the *theory of vector spaces over* $\mathbb{Z}_2$. Therefore, any map $f : \mathbb{Z}_2^n \to \mathbb{Z}_2$ is represented by a *term* with n *variables* $x_1, \dots, x_n$, and two such terms represent the same map if and only if they are equivalent modulo those axioms. Following [Bur93] and [Laf95a], we get:

Theorem 5 *The generators τ, δ, ε, μ, η and the relations of figure 13 form a presentation of* $\mathbf{L}(\mathbb{Z}_2)$.

We shall give an independent proof of this theorem in section 3.2.

Note that τ is a *superfluous generator*, since it is definable in terms of δ and μ:



However, removing τ would seriously complicate the presentation.

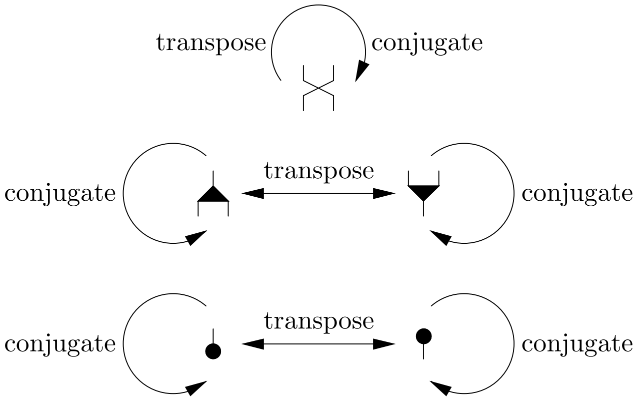


Figure 12: Symmetries of the generators for $\mathbf{L}(\mathbb{Z}_2)$

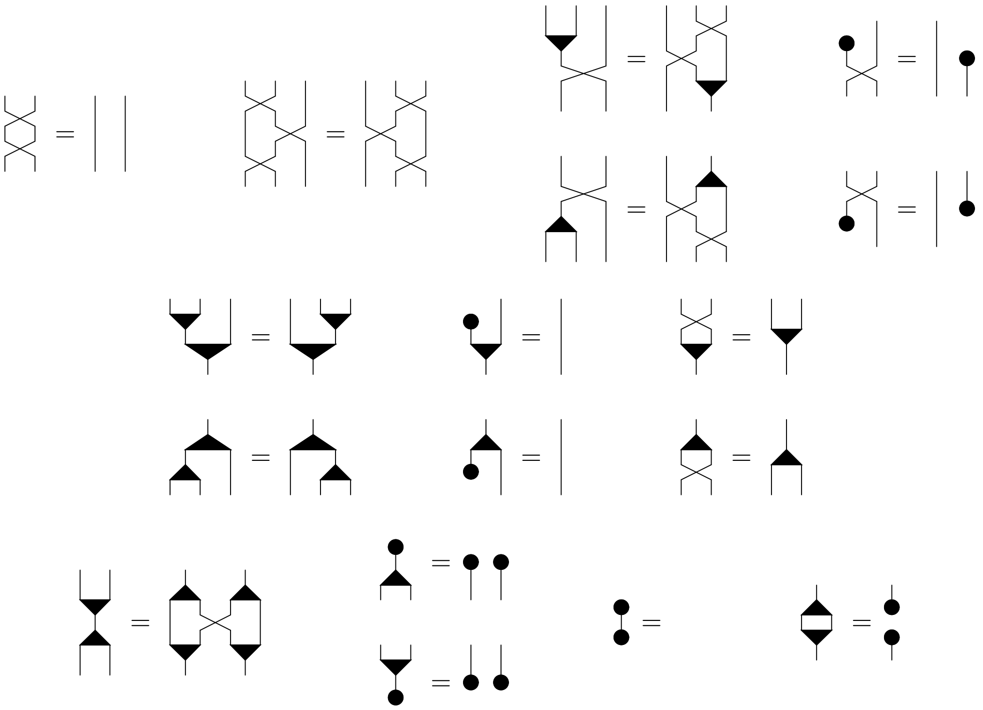


Figure 13: Relations for $\mathbf{L}(\mathbb{Z}_2)$

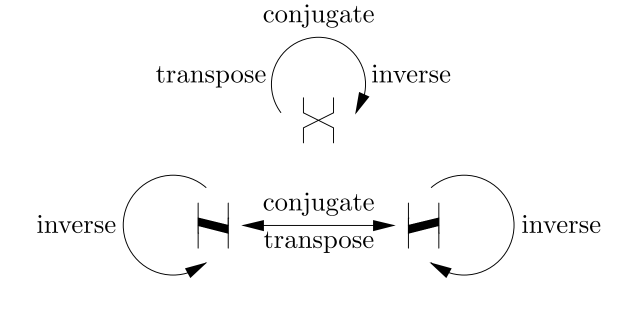


Figure 14: Symmetries of the generators for $\mathbf{GL}(\mathbb{Z}_2)$

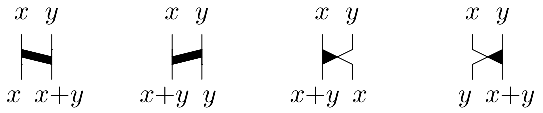
3.1 Linear permutations

Let $\mathbf{GL}(K)$ be the monoidal subcategory of $\mathbf{L}(K)$ whose morphisms are *linear permutations*. In this case, there is an extra symmetry, *inversion*, which reverts the order of vertical composition.

Note that any $\alpha : 2 \to 2$ defines *elementary operations* on matrices:

- Multiplying $\psi : p \to q$ by $\mathrm{id}_{i-1} \mid \alpha \mid \mathrm{id}_{q-i-1}$ on the *left* corresponds to an elementary operation on *rows* i and $i+1$ of the matrix. In that case, we say that we *apply* α *to rows* i *and* $i+1$.
- Multiplying $\psi : p \to q$ by $\mathrm{id}_{j-1} \mid \alpha \mid \mathrm{id}_{p-j-1}$ on the *right* corresponds to an elementary operation on *columns* j and $j+1$ of the matrix. In that case, we say that we *apply* α *to columns* j *and* $j+1$.

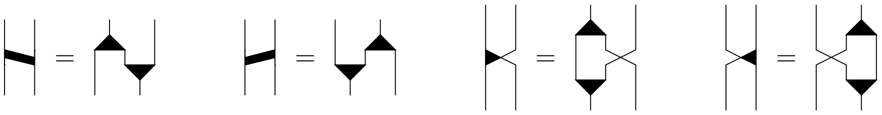
Again, we consider the case of $\mathbb{Z}_2$. Apart from id_2 and τ, there are four linear permutations of $\mathbb{Z}_2^2$:



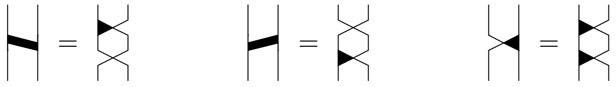
The corresponding matrices are the following:

$$\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix} \quad \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} \quad \begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix} \quad \begin{pmatrix}0 & 1\\ 1 & 1\end{pmatrix}$$

The symmetries are given in figure 14. Of course, each of them is definable in terms of τ, δ, and μ:

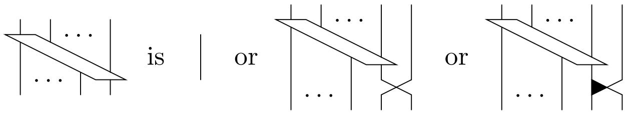


But δ and μ are not permutations. In fact, $\mathbf{GL}(\mathbb{Z}_2)$ is generated by τ and any of the four above generators, for instance the third one, that we call κ. The other ones are indeed superfluous:

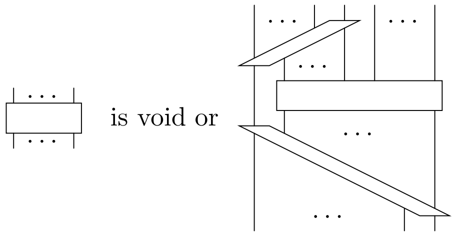


Theorem 6 *The generators τ, κ, and the six relations of figure 15 form a presentation of* $\mathbf{GL}(\mathbb{Z}_2)$.

To show this, we need an extended notion of *stairs*:



Of course, there is a dual notion of *antistairs*. We consider a first notion of canonical form:



It is obtained by the following algorithm, which applies to any invertible matrix A of order n:

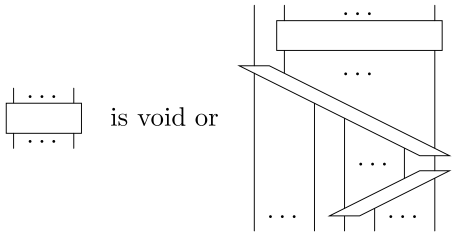
1. Consider the last row of A, and let j be the last index for which the coefficient is 1.
2. While $j > 1$, apply τ or κ^{-1} to columns $j-1$ and j, so that this index becomes $j-1$.
3. By construction, the last row is now $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$. So, if we consider the first column, n is the last index i for which the coefficient is 1.
4. While $i > 1$, apply τ or κ^{-1} to rows $i-1$ and i, so that this index becomes $i-1$, and row $i-1$ becomes $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$.

To sum up, we have made the following transformations:

$$\begin{pmatrix} * & \cdots & * & * & * & \cdots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ * & \cdots & * & * & * & \cdots & * \\ * & \cdots & * & 1 & 0 & \cdots & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \\ 1 & 0 & \cdots & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}$$

In particular, we get a matrix of the form $1 \oplus A'$, where A' is an invertible matrix of order $n-1$. The antistairs are given by step 2, the stairs by step 4, and the rest of the canonical form is obtained by applying the algorithm to the matrix A'. See figure 16 for an example.

As a by-product, we get a simple algorithm for inverting matrices: see appendix B. It is easy to see that this canonical form is unique, but it is not suitable for our purpose, because the stairs and the antistairs are far away. For that reason, we shall use an alternative notion of canonical form:



Lemma 9 *Any linear permutation $f : K^n \to K^n$ is represented by a unique canonical form.*

This is proved by induction on n, using the following algorithm, which applies to any invertible matrix A of order n:

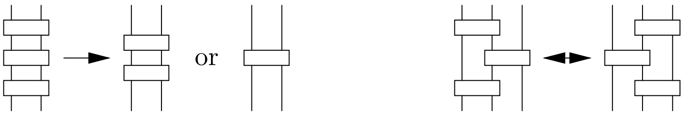
1. The matrix obtained by forgetting the first column of A is of rank $n-1$. Since 1 is the unique invertible element of $\mathbb{Z}_2$, there is a unique non-trivial linear relation $a_1u_1 + \cdots + a_nu_n = 0$, where $u_1, \ldots, u_n$ are the rows of this truncated matrix. Let i be the first index such that $a_i = 1$.
2. While $i < n$, apply τ or κ^{-1} to rows i and $i+1$ of A, so that this index becomes $i+1$.
3. By construction, the last row is now $\begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$.
4. Proceed as in the previous algorithm.

To sum up, we have made the following transformations:

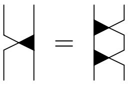
$$\left(\begin{matrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \\ * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{matrix}\right) \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \\ * \\ \vdots \\ * \end{matrix} \longrightarrow \left(\begin{matrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \\ 1 & 0 & \cdots & 0 \end{matrix}\right) \begin{matrix} 0 \\ \vdots \\ 0 \\ 1 \end{matrix} \longrightarrow \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix}$$

The extra column on the right contains $a_1, \ldots, a_n$. Again, we get a matrix of the form $1 \oplus A'$. This A' is obtained by removing the first column and one row of the original matrix. See figure 17 for an example.

In figure 18, we introduce two groups of rules for the following transformations:



Those fourteen rules are derivable from the six relations of figure 15. Indeed, five rules are already in the presentation, five more are derived in figure 19, and we get three more by duality. It remains to show that the last rule of figure 18 is derivable from the presentation. This is done in figure 20, using the superfluous generator κ^2, which is both the inverse and the conjugate of κ:



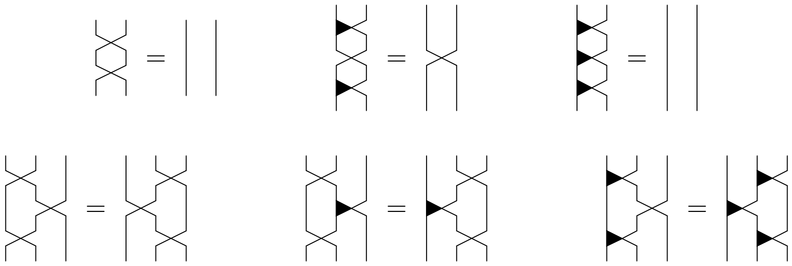


Figure 15: Relations for $\mathbf{GL}(\mathbb{Z}_2)$

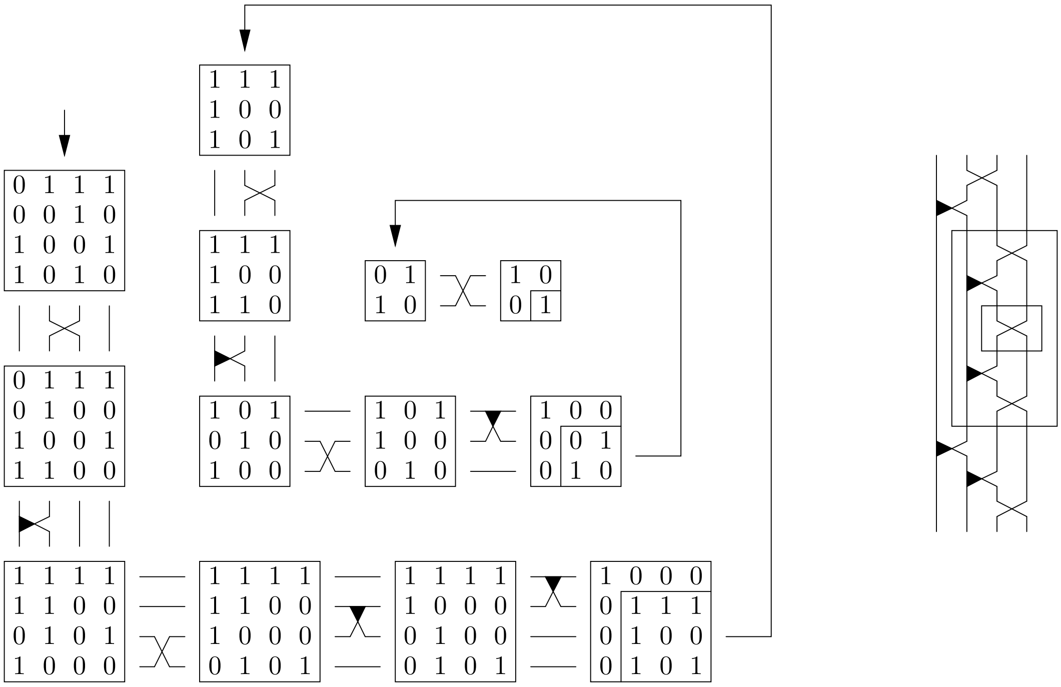


Figure 16: Computing the first canonical form of a matrix in $\mathbf{GL}(\mathbb{Z}_2)$

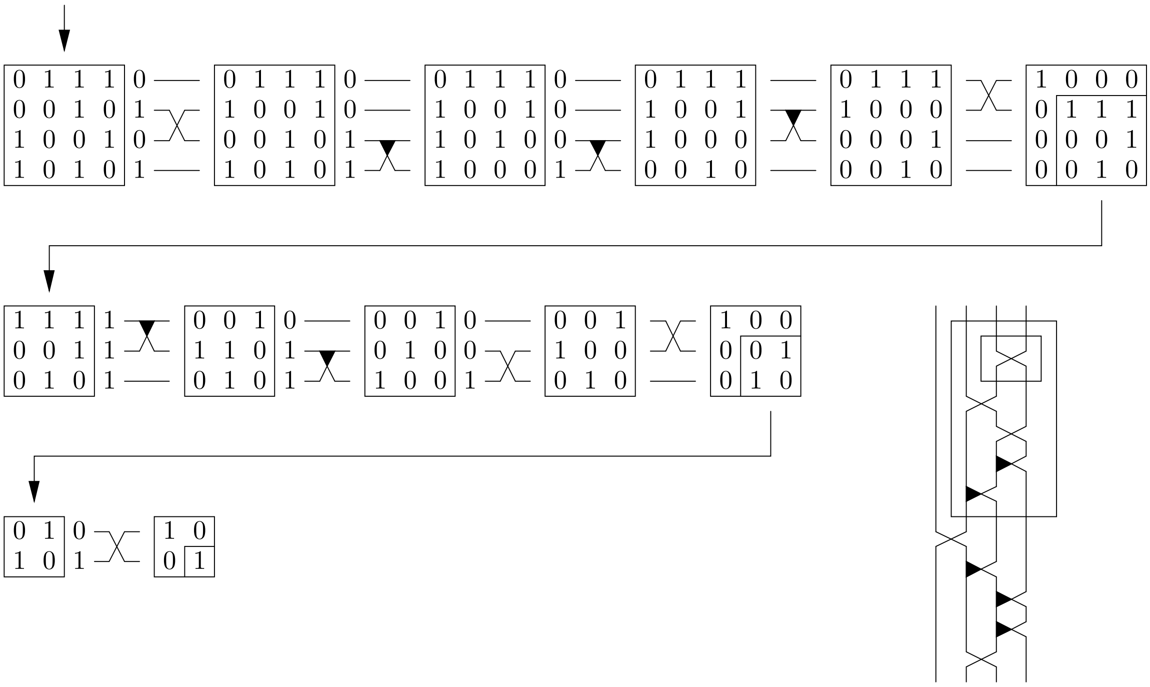


Figure 17: Computing the second canonical form of a matrix in $\mathbf{GL}(\mathbb{Z}_2)$

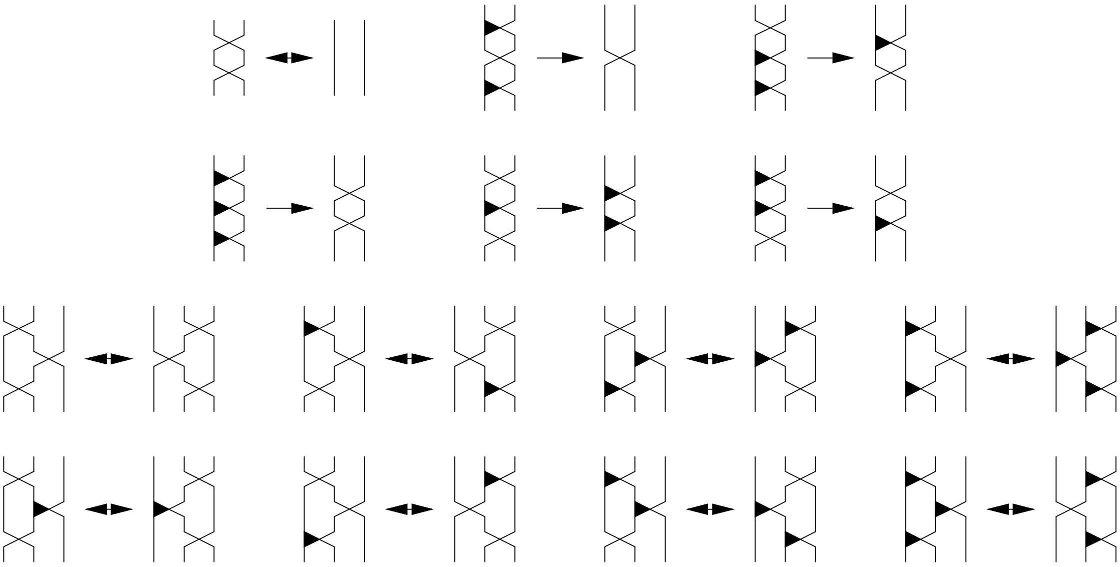


Figure 18: Rules for $\mathbf{GL}(\mathbb{Z}_2)$

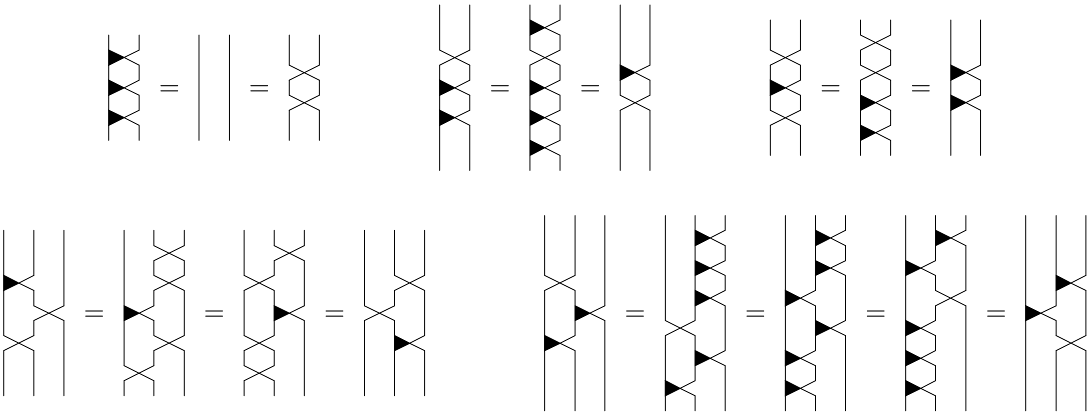


Figure 19: Deriving rules for $\mathbf{GL}(\mathbb{Z}_2)$

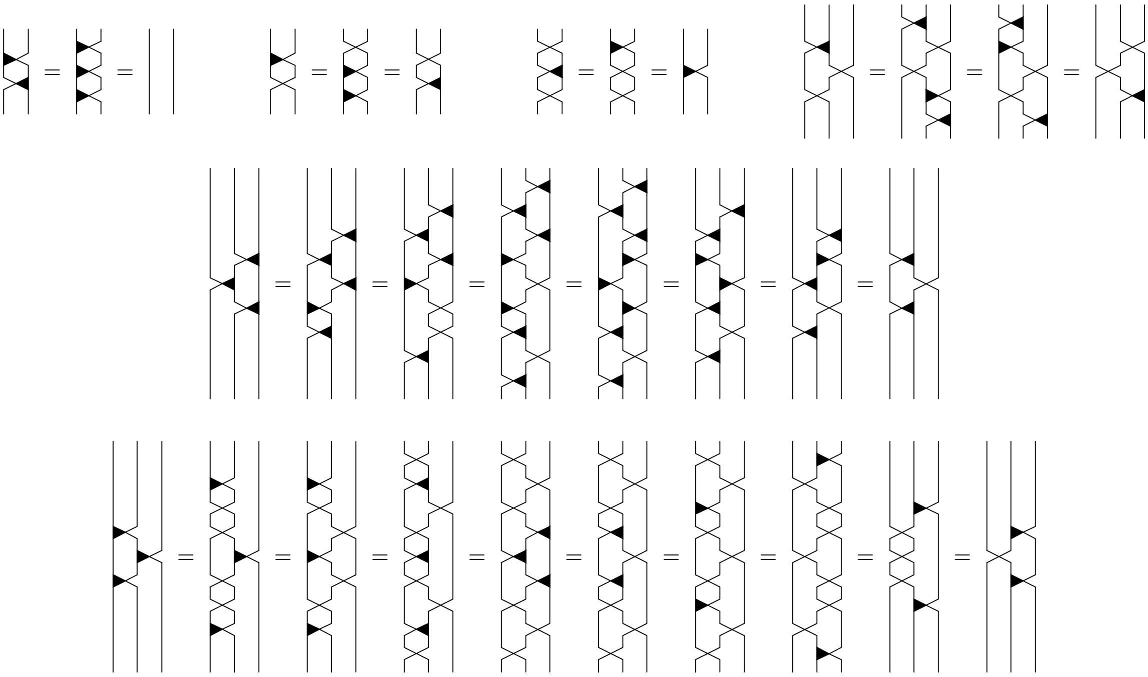
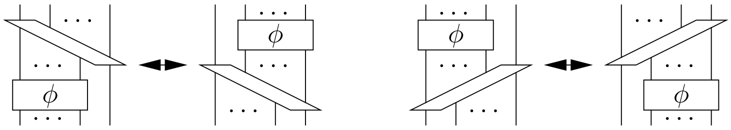


Figure 20: Deriving the last rule for $\mathbf{GL}(\mathbb{Z}_2)$

Lemma 10 *For any diagram ϕ, we have the following commutations:*

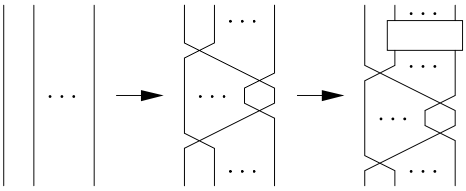


This is easily proved by induction on the size m of ϕ, using the second group of rules. Of course, the stairs (respectively the antistairs) may be changed by this commutation.

Lemma 11 *Any diagram $\phi : n \to n$ reduces to a canonical form $\widehat{\phi}$ by the rules of figure 18.*

This is proved by double induction on n and the size m of ϕ:

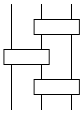
- If $m = 0$, then $\phi = \mathrm{id}_n$, which reduces to a canonical form. This is proved by induction on n, using the first rule:



- If $m \geq 1$, then $\phi = \xi \circ \psi$ where ξ is an elementary diagram and ψ is a diagram of size $m-1$ which reduces to a canonical form $\widehat{\psi}$ by induction hypothesis. Therefore, ϕ reduces to $\xi \circ \widehat{\psi}$ and there are four possible cases (figure 21): in the first case, we use the second group of rules and the induction hypothesis for $n-1$; in the second case, we get a canonical form; in the third case, we use the crucial transformation of figure 22, which itself uses lemma 10 and the first group of rules, and the induction hypothesis for $n-1$; in the fourth case, we use the second group of rules twice and the induction hypothesis for $n-1$.

Theorem 6 follows from lemma 9 and lemma 11.

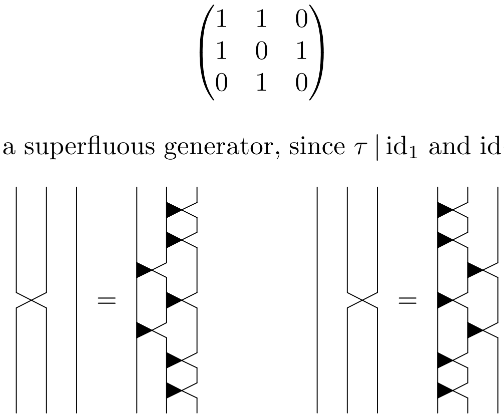
Note that the rewrite system of figure 18 is not terminating. It is easy to give an alternative system such that identity diagrams are in canonical form, but the existence of a terminating rewrite system for $\mathbf{GL}(\mathbb{Z}_2)$ is an open question. Indeed, it is essential that the rules of the second group are used in both directions: one for stairs and one for antistairs. Furthermore, the fact that we need both stairs and antistairs does not rely on our choice of generators. Indeed, even if we take all linear permutations of $\mathbb{Z}_2^2$ as generators, there are linear permutations of $\mathbb{Z}_2^3$ which cannot be decomposed as follows:



In other words, there are invertible matrices of order 3 which are not of the form $(1 \oplus A)(B \oplus 1)(1 \oplus C)$. Here is an example:

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Finally, note that τ is almost a superfluous generator, since $\tau \mid \mathrm{id}_1$ and $\mathrm{id}_1 \mid \tau$ are definable in terms of κ:



This means that τ is only necessary in dimension 2, but again, removing it would seriously complicate the presentation.

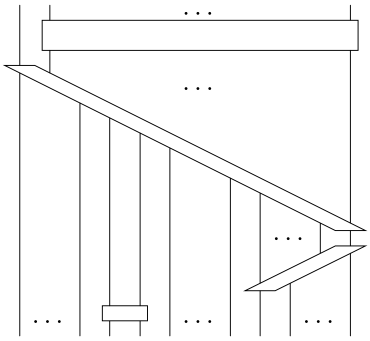


Figure 21: The four cases of lemma 11

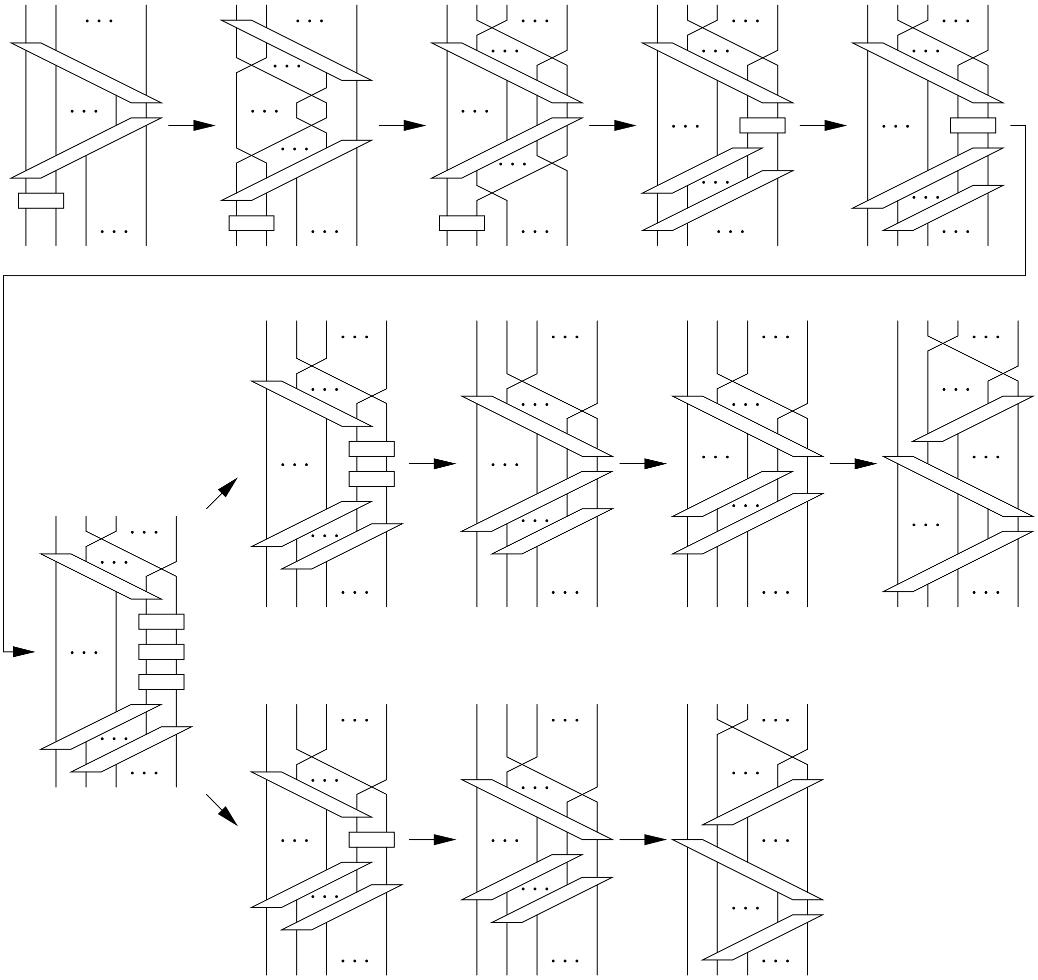
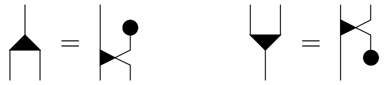


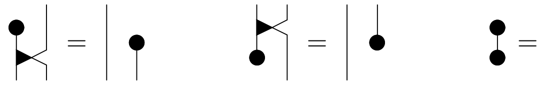
Figure 22: Crucial transformation for the third case of lemma 11

3.2 Linear maps

In order to extend our presentation of $\mathbf{GL}(\mathbb{Z}_2)$ to $\mathbf{L}(\mathbb{Z}_2)$, the generators ε and η are clearly needed, but δ and μ are superfluous, since they are definable in terms of κ, ε, and η:

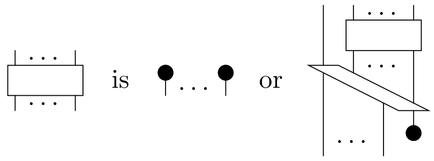


Obviously, κ, ε, and η satisfy the following relations:



Theorem 7 *The generators τ, κ, ε, η, and the six relations of figure 15 together with the above three ones form a presentation of* $\mathbf{L}(\mathbb{Z}_2)$*.*

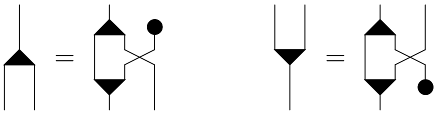
To show this, we use the following notion of canonical form:



It is a straightforward transcription of the matrix: the stairs correspond to the first column of the matrix, and within stairs, τ stands for 0 and κ for 1. See figure 23 for an example.

We introduce the rules of figure 24, which are derivable from our presentation: see figure 25. To show that any identity diagram reduces to a canonical form, an analogue of lemma 5 is needed: see figure 26. Theorem 7 follows from the fact that any diagram reduces to a canonical form by the rules of figure 24.

We get theorem 5 as a corollary. Indeed, we have defined a translation F of diagrams built with τ, κ, ε, η into diagrams built with τ, δ, ε, μ, η, and a translation G in the opposite direction. By applying F to the nine relations of theorem 7, we get relations which are derivable from those of figure 13. Furthermore, the following two relations are also derivable from those of figure 13:



This implies that for any diagram ϕ built with τ, δ, ε, μ, η, we have $\phi \equiv F(G(\phi))$ modulo the relations of figure 13. Therefore, if ϕ and ψ represent the same linear map, we have $G(\phi) \equiv G(\psi)$ modulo the nine relations of theorem 7, so that $\phi \equiv F(G(\phi)) \equiv F(G(\psi)) \equiv \psi$ modulo the relations of figure 13.

Again, the rewrite system of figure 24 is not terminating, but we can adopt the alternative canonical form of figure 27, which uses the two superfluous generators δ and μ, and the rewrite system of figure 28. We conjecture that this rewrite system is terminating: see appendix A. The upper part of this canonical form is obtained by the following algorithm, which applies to any matrix A without zero row:

1. If the first column is zero, remove it. Otherwise, let i be the last index for which the coefficient of the first column is 1.

2. If row i is not of the form $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$, replace this 1 by 0 and insert the row $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$ between rows $i-1$ and i.

3. While $i > 1$, apply τ or κ^{-1} to rows $i-1$ and i, or in case row $i-1$ is of the form $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$, remove it, so that in all cases, this index becomes $i-1$ and row $i-1$ becomes $\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}$.

To sum up, we have made the following transformations:

$$\begin{pmatrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \\ 1 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} \longrightarrow \begin{pmatrix} * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \\ 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & * & \cdots & * \\ & & & \\ \vdots & \vdots & & \vdots \\ & & & \\ 0 & * & \cdots & * \end{pmatrix}$$

$$\boxed{\begin{matrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{matrix}}$$

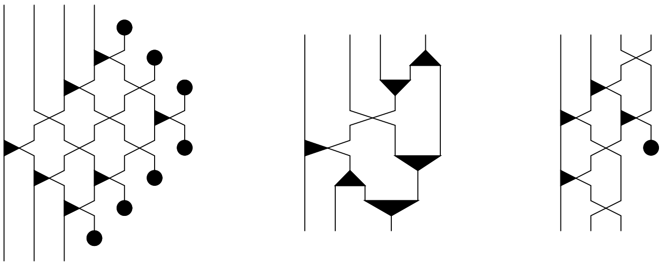


Figure 23: The canonical forms of a matrix in $\mathbf{L}(\mathbb{Z}_2)$

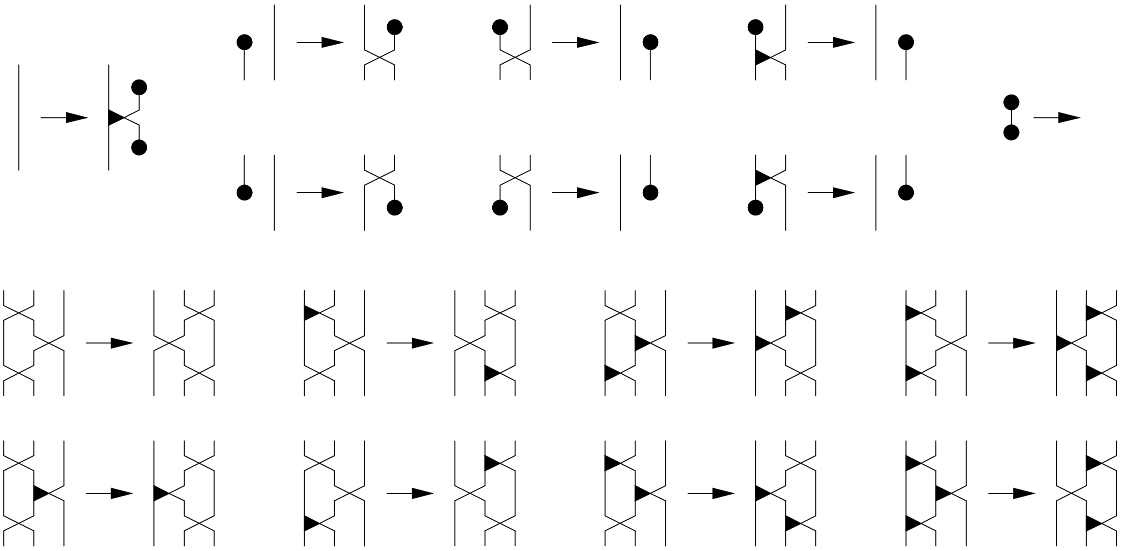


Figure 24: Rules for $\mathbf{L}(\mathbb{Z}_2)$

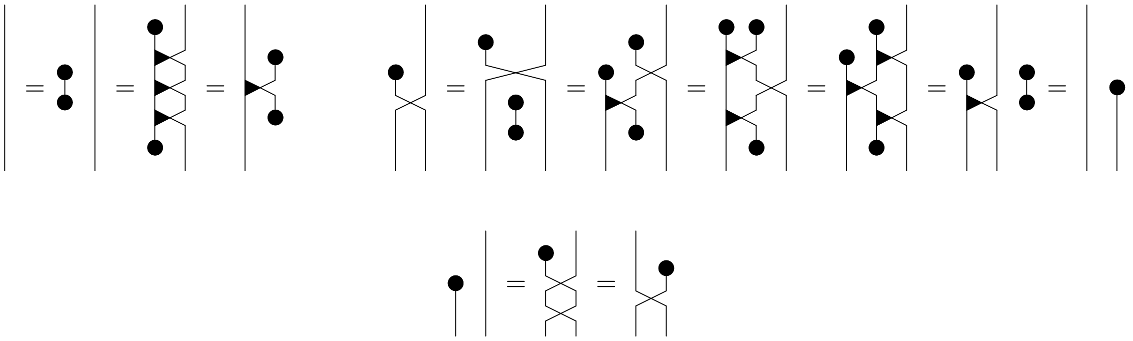


Figure 25: Deriving rules for $\mathbf{L}(\mathbb{Z}_2)$

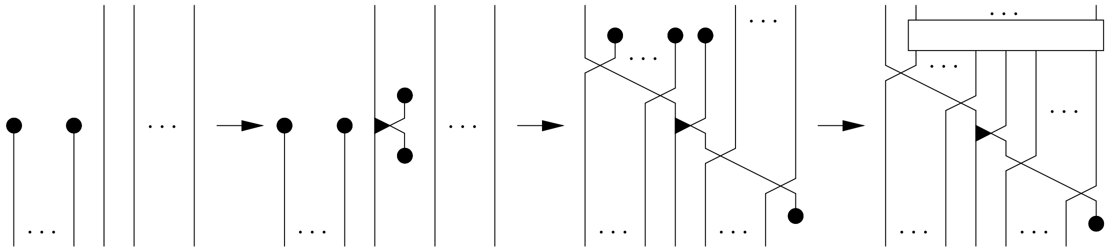


Figure 26: Expansion of identities for $\mathbf{L}(\mathbb{Z}_2)$

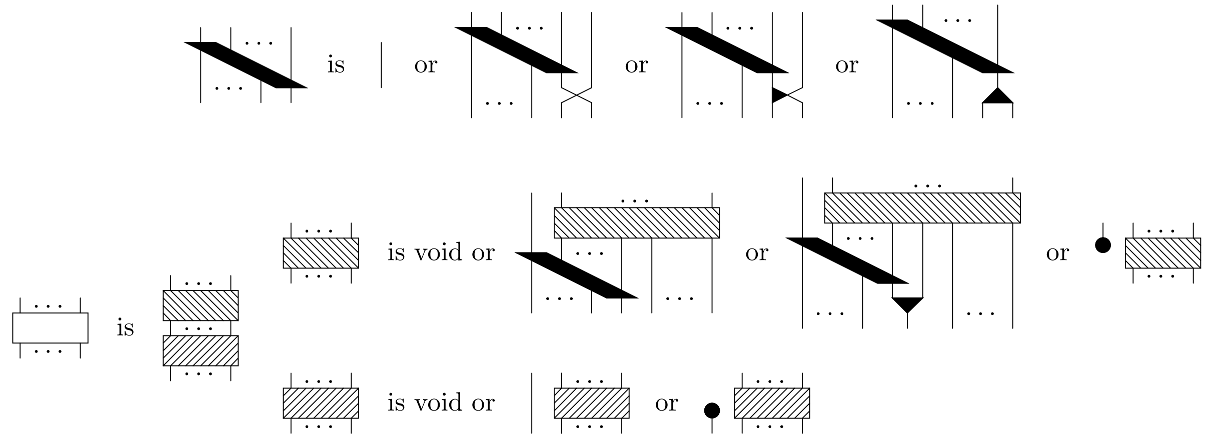


Figure 27: Alternative canonical form for $\mathbf{L}(\mathbb{Z}_2)$

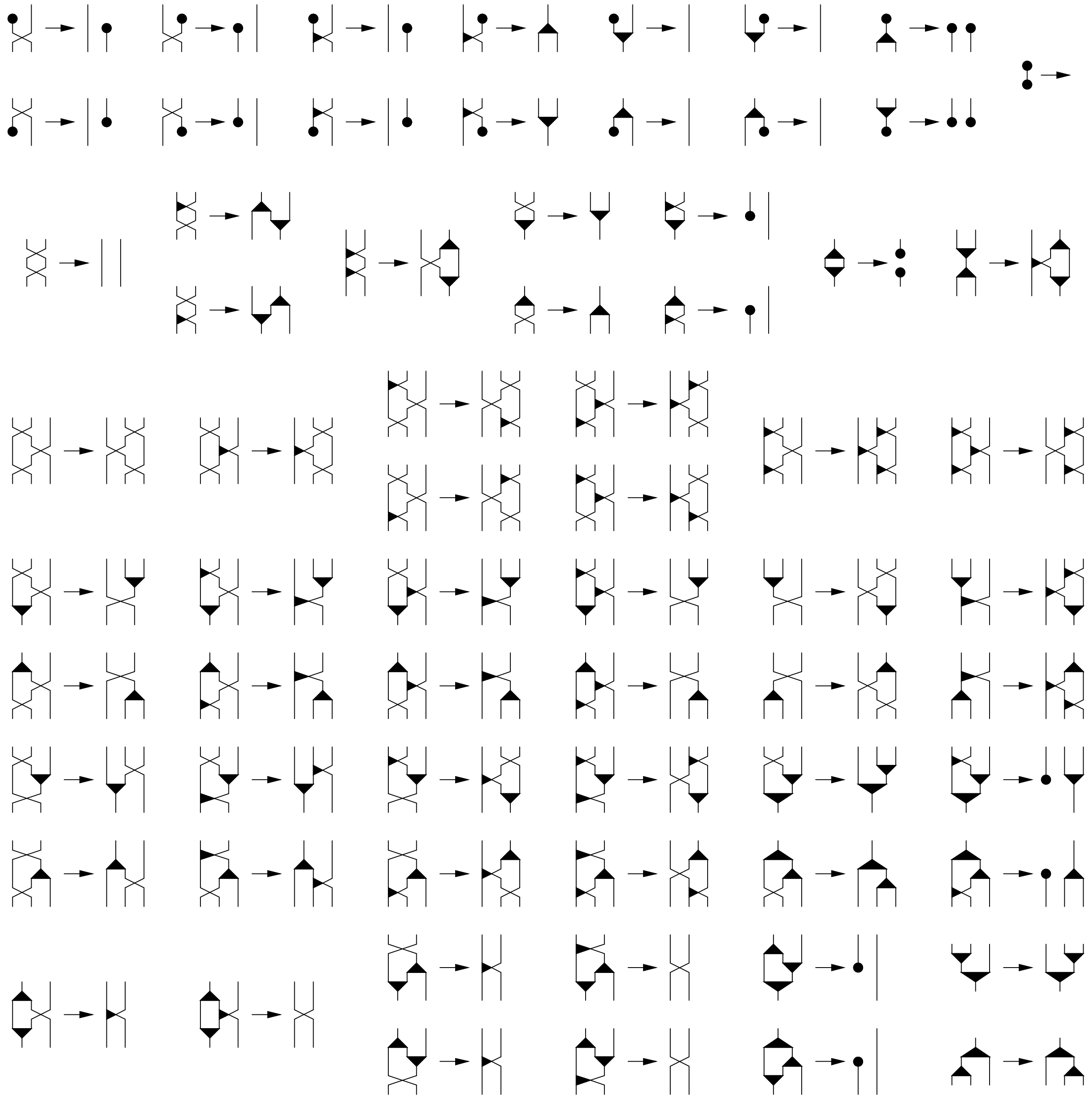
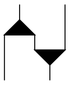


Figure 28: Alternative rules for $\mathbf{L}(\mathbb{Z}_2)$

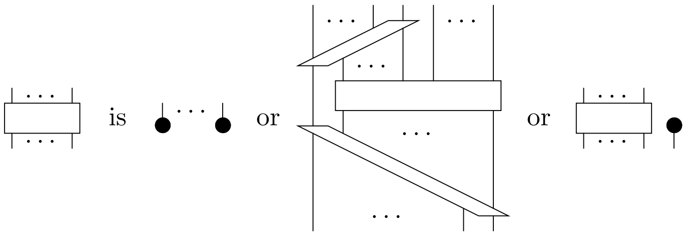
In particular, we get a matrix of the form $1 \oplus A'$, and by construction, A' has no zero row. The μ cell (if any) is given by step 2, the (generalized) stairs by step 3, and the rest of the canonical form is obtained by applying the algorithm to the matrix A'. See figure 23 for an example.

Note that the rewrite system of figure 28 extends the one of figure 9. Furthermore, it is self-dual, so that the canonical form of the dual is always the dual of the canonical form. This may not be clear from the grammar of figure 27, but remember that we have an implicit commutation rule.

Note also that the canonical form of a linear permutation may contain the generators δ and μ, which are not permutations, as in the following example:



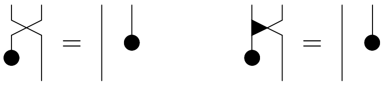
Instead, we can generalize the first canonical form of linear permutations as follows:



This canonical form is obtained by the same algorithm as in the case of linear permutations, except that the last row may be zero, and one may end with a degenerate matrix. See figure 23 for an example. This algorithm can also be used to compute the rank of a matrix. Note the following points:

- The canonical form of a linear permutation contains only τ and κ.
- The canonical form of a linear surjection contains only τ, κ, and ε.
- The canonical form of a linear injection contains only τ, κ, and η.

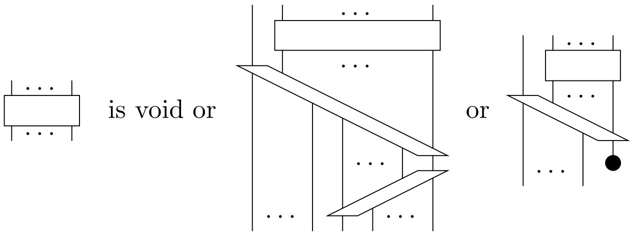
In particular, the monoidal subcategory $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$ of $\mathbf{L}(\mathbb{Z}_2)$ whose morphisms are *linear surjections* is generated by τ, κ, and ε. We have already seen that these generators satisfy the following relations:



Theorem 8 *The generators τ, κ, ε, and the six relations of figure 15 together with the above two ones form a presentation of* $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$.

By duality, we get a presentation of the monoidal subcategory $\mathbf{L}^{\mathrm{inj}}(\mathbb{Z}_2)$ of $\mathbf{L}(\mathbb{Z}_2)$ whose morphisms are *linear injections* by τ, κ, η, and the six relations of figure 15 together with the dual of the above ones.

To show theorem 8, we consider the following canonical form of linear surjections, which generalizes the second canonical form of linear permutations:

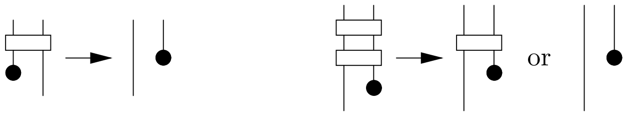


It is obtained by the following algorithm, which applies to any matrix A with q independent rows:

1. The matrix obtained by forgetting the first column of A is of rank $q-1$ or q.
2. If the rank is q, remove this first column. Otherwise, there is a unique non-trivial linear relation $a_1u_1 + \cdots + a_nu_n = 0$, where $u_1, \ldots, u_n$ are the rows of this truncated matrix.
3. Proceed as in the case of linear permutations.

We get a canonical form of the second type if the rank is $q-1$, or of the third type if the rank is q, and in that case, the stairs corresponds to the first column of the matrix. See figure 29 for an example.

In figure 30, we introduce rules for the following transformations:



Those rules are derivable from the six relations of figure 15 together with the above two ones: see figure 31.

Theorem 8 follows from the fact that any diagram reduces to a canonical form by the rules of figure 18 together with those of figure 30. To show this, we proceed as in the case of linear permutations, using the crucial transformation of figure 32.

The canonical form of linear surjections can be generalized to linear maps: see figure 33. The bottom part of this canonical form is obtained by the following algorithm, which takes any system $u_1, \ldots, u_q$ and returns a free subsystem of the same rank:

1. Apply the algorithm to $u_2, \ldots, u_q$: it returns a free subsystem $v_1, \ldots, v_n$.
2. If u_1 is independent of $v_1, \ldots, v_n$, return $u_1, v_1, \ldots, v_n$. Otherwise, return $v_1, \ldots, v_n$.

This algorithm is applied to the *rows* of the matrix. One gets a canonical form of the second type if u_1 is independent of $v_1, \ldots, v_n$, or of the third type if $u_1 = a_1v_1 + \cdots + a_nv_n$, and in that case, the antistairs are given by the coefficients $a_1, \ldots, a_n$. In both cases, the rest of the canonical form is given by step 1. Finally, the algorithm returns a matrix with independent rows, corresponding to a linear surjection, from which one gets the upper part of the canonical form. See figure 35 for an example.

The canonical form of figure 33 corresponds to the unique decomposition of any linear map $f : \mathbb{Z}_2^p \to \mathbb{Z}_2^q$ into a linear surjection $g : \mathbb{Z}_2^p \to \mathbb{Z}_2^n$ followed by a *monotone linear injection* $h : \mathbb{Z}_2^n \to \mathbb{Z}_2^q$, where n is the rank of f. Here, we assume that $\mathbb{Z}_2^n$ is equipped with the following *antilexicographical ordering*:

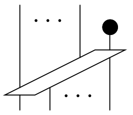
$$(0, x_2, \ldots, x_n) < (1, x_2, \ldots, x_n), \qquad (x_1, \ldots, x_n) < (y_1, \ldots, y_n) \text{ whenever } (x_2, \ldots, x_n) < (y_2, \ldots, y_n).$$

It is indeed easy to see that the bottom part of a canonical form corresponds to a monotone linear injection. The converse is proved by induction on q, using the following lemma:

Lemma 12 *Let A be the matrix of a monotone linear injection $f : \mathbb{Z}_2^p \to \mathbb{Z}_2^q$ and let B be the matrix obtained by removing the first row of A. Then B is the matrix of a monotone linear injection or A is of the form $1 \oplus C$ where C is the matrix of a monotone linear injection.*

By definition of the ordering, B is indeed the matrix of a monotone linear map $g : \mathbb{Z}_2^p \to \mathbb{Z}_2^{q-1}$. If g is not injective, then $g(u) = 0$ for some $u \neq 0$ in $\mathbb{Z}_2^p$. Since f is injective, $f(u) = e_1$, which is the smallest element after 0 in $\mathbb{Z}_2^q$. Since f is a monotone injection, $u = e_1$. Now, if $j > 1$, then $e_j < e_1 + e_j$ and $f(e_j) < f(e_1 + e_j) = f(e_1) + f(e_j) = e_1 + f(e_j)$, so that the first component of $f(e_j)$ must be 0. Therefore, A is of the form $1 \oplus C$, and it is easy to check that C is the matrix of a monotone linear injection.

Monotone linear injections form a monoidal subcategory of $\mathbf{L}^{\mathrm{inj}}(\mathbb{Z}_2)$ which is not finitely generated. In fact, it is generated by all maps of the form $f(x_1, \ldots, x_n) = (a_1x_1 + \cdots + a_nx_n, x_1, \ldots, x_n)$, which are represented by diagrams of the following form:



Note also that if we apply the above decomposition to a linear injection, we get a linear permutation followed by a monotone linear injection. This means that monotone linear injections provide a canonical choice of basis for subspaces of $\mathbb{Z}_2^n$.

There is a dual notion of *comonotone linear surjection.* Any linear map can be uniquely decomposed into a comonotone linear surjection followed by a linear injection. Again, if we apply this decomposition to a linear surjection, we get a comonotone linear surjection followed by a linear permutation. In fact, the comonotone linear surjections coincide with the *conjugate* of linear surjections whose matrices are in *row-reduced echelon form.* This row-reduced echelon form is just what we get when we apply the Gauss algorithm to solve a linear system.

Finally, any linear map can be uniquely decomposed into a comonotone linear surjection followed by a linear permutation followed by a monotone linear injection. This decomposition corresponds to the canonical form of figure 34. See figure 35 for an example.

$$\boxed{\begin{matrix} 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{matrix}}$$

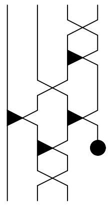


Figure 29: The canonical form of a matrix in $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$

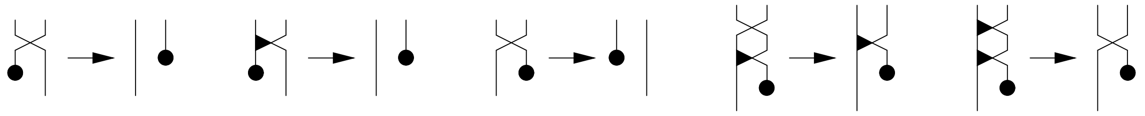


Figure 30: Extra rules for $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$

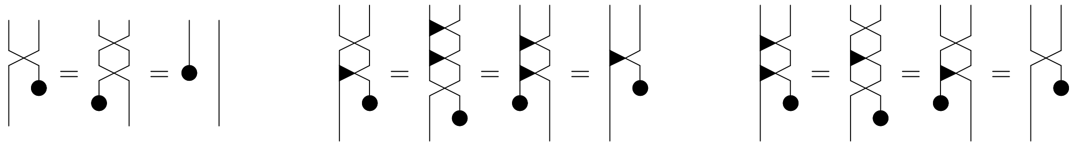


Figure 31: Deriving rules for $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$

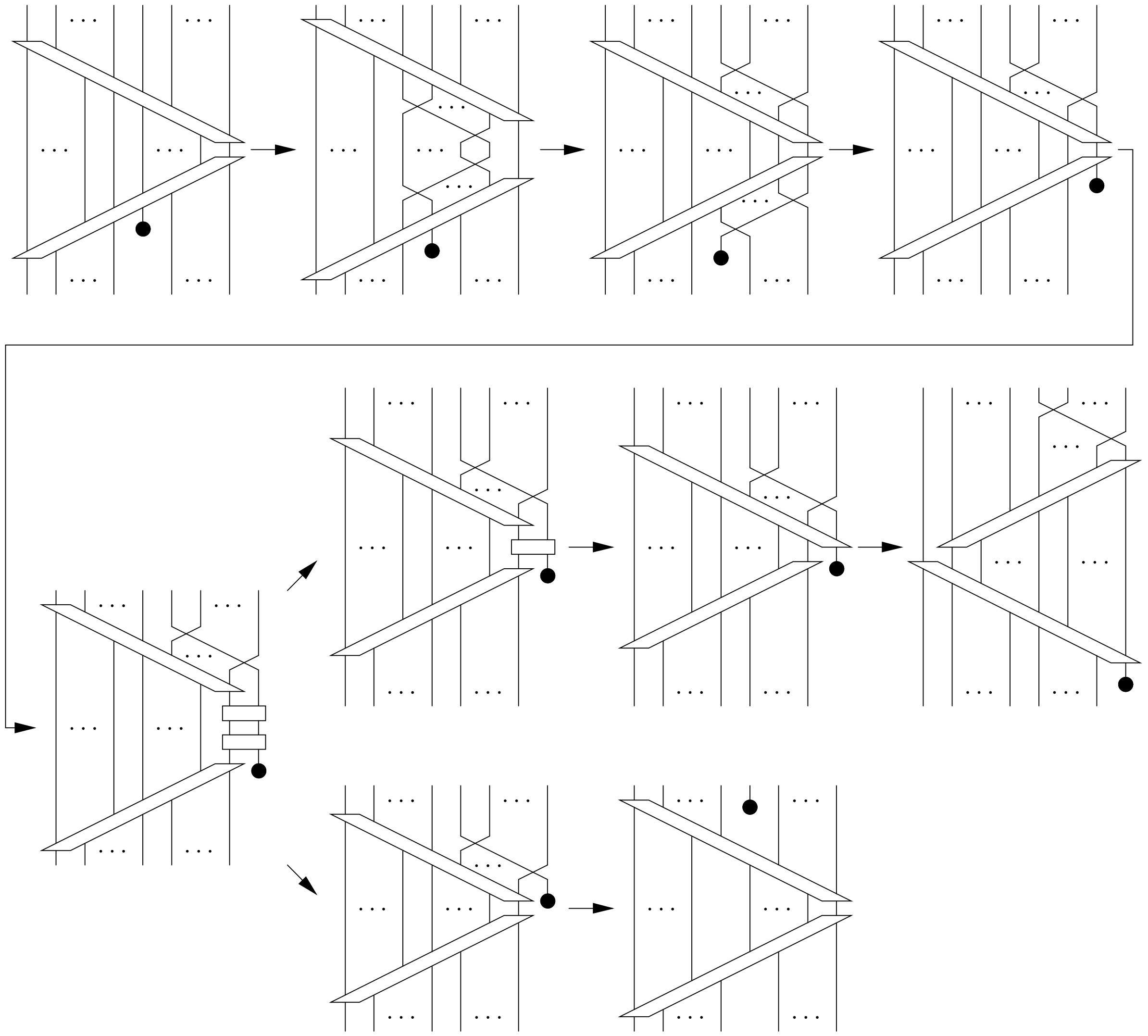


Figure 32: Crucial transformation for $\mathbf{L}^{\mathrm{surj}}(\mathbb{Z}_2)$

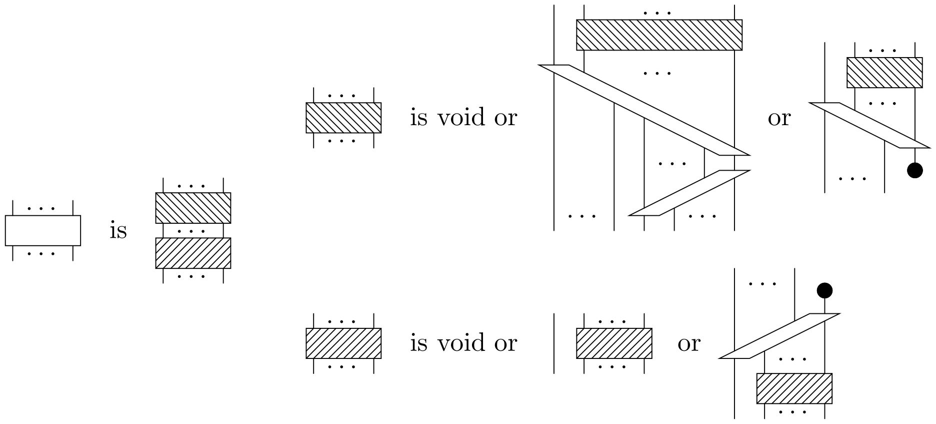


Figure 33: Another canonical form for $\mathbf{L}(\mathbb{Z}_2)$

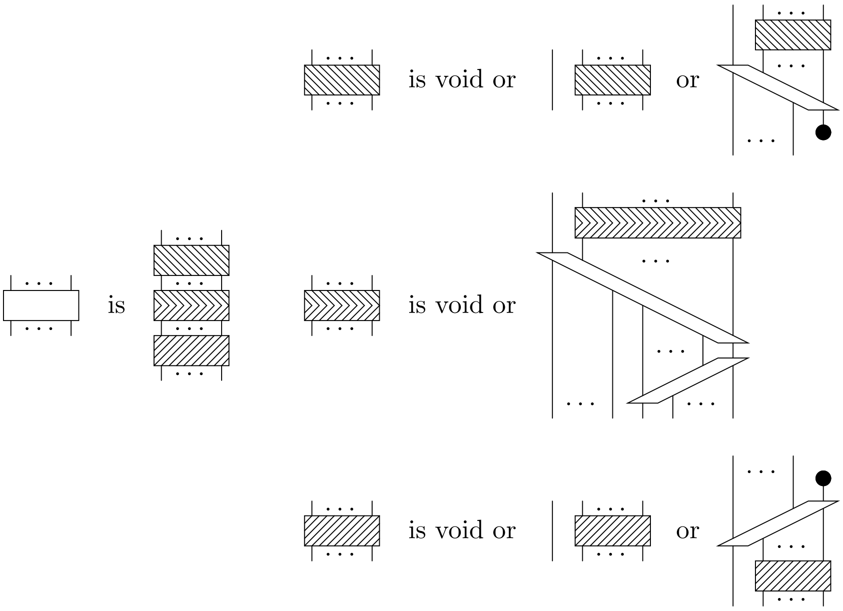
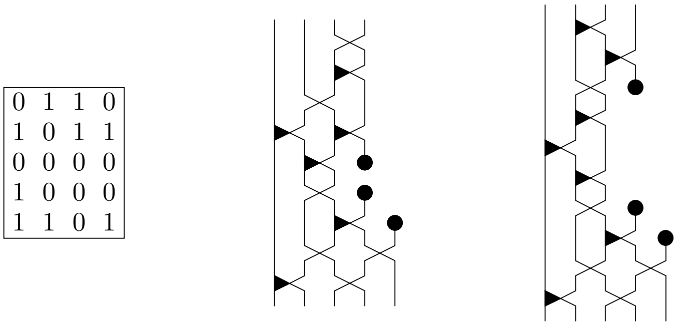


Figure 34: Yet another canonical form for $\mathbf{L}(\mathbb{Z}_2)$

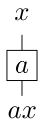


$$\boxed{\begin{array}{cccc} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{array}}$$

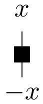
Figure 35: Other canonical forms of a matrix in $\mathbf{L}(\mathbb{Z}_2)$

3.3 Arbitrary field

The results of the previous sections can be generalized to an arbitrary field K. First we introduce a generator $\mathrm{H}_a : 1 \to 1$ for each $a \in K$. It is pictured and interpreted as follows:



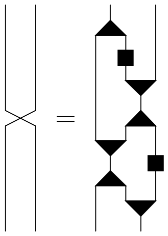
In particular, we write σ for H_{-1}, which is pictured and interpreted as follows:



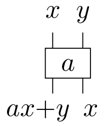
The presentation of $\mathbf{L}(K)$ is the one of figure 13, where the last relation is replaced by the nine relations of figure 36. The last six relations correspond to the following axioms for vector spaces:

$$a(x+y) = ax + ay, \qquad a0 = 0, \qquad a(bx) = (ab)x, \qquad 1x = x, \qquad (a+b)x = ax + bx, \qquad 0x = 0.$$

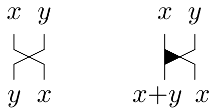
Note that H_0 and H_1 are superfluous generators. Furthermore, in the case of a finite field, it is well known that the multiplicative group of invertible elements is cyclic, so that the H_a for $a \neq 0$ are definable in terms of a single one. Note also that τ is always definable in terms of δ, μ, and σ:



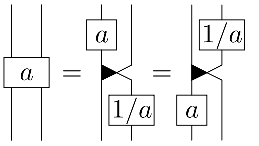
For $\mathbf{GL}(K)$, we keep the H_a for $a \neq 0$, and we introduce a new generator $\mathrm{K}_a : 2 \to 2$ for each $a \in K$, which is pictured and interpreted as follows:



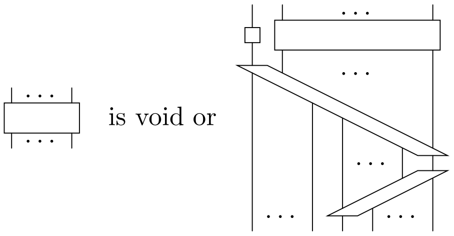
In particular, we write τ for K_0 and κ for K_1 as in the case of $\mathbb{Z}_2$:



Note that the K_a for $a \neq 0$ are definable in terms of κ and the H_a:



Stairs and *antistairs* are built with the K_a, and the second canonical form is generalized as follows:

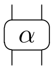


Any diagram built with the H_a for $a \neq 0$ and the K_a reduces to such a canonical form by the rules of figure 37. From this, it is possible to get a presentation of $\mathbf{GL}(K)$ with τ, κ, and the H_a for $a \neq 0$ as generators. Again, a single H_a is needed in the case of a finite field.

3.4 Isometries

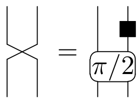
Finally, we consider the monoidal subcategory $\mathbf{O}$ of $\mathbf{GL}(\mathbb{R})$ whose morphisms are *isometries*. An isometry is a linear map $f : \mathbb{R}^n \to \mathbb{R}^n$ which preserves the Euclidean metrics of $\mathbb{R}^n$. The matrix of an isometry is called an *orthogonal matrix*: its columns form an *orthonormal basis*, that is a system $u_1, \ldots, u_n$ such that each u_i has length 1, and u_i is orthogonal to u_j for $j \neq i$. Note that if all coefficients of u_1 are zero, but the first one, then the matrix is of the form $\pm 1 \oplus A$ where A is an orthogonal matrix of order $n-1$.

Apart from the identity, σ is the only isometry of $\mathbb{R}$. The isometries of $\mathbb{R}^2$ are rotations or symmetries. So we introduce a generator $\mathrm{R}_\alpha : 2 \to 2$ for each $\alpha \in \mathbb{R}$, which is pictured as follows:

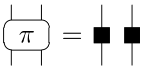


It stands for the rotation of angle α, whose matrix is $\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$.

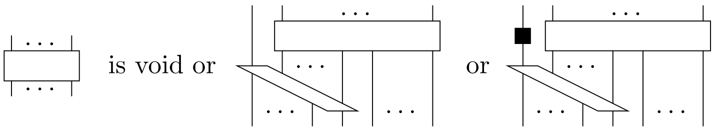
Symmetries are definable in terms of the R_α and σ. For instance, τ can be decomposed as follows:



Note also that R_π is definable in terms of σ:



Therefore, we shall only use R_α for $\alpha \in]0, \pi[$. *Stairs* are built with those R_α, and there is a simple notion of canonical form which is similar to the canonical form of permutations in the basic case:



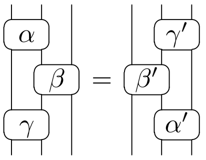
It is obtained by the following algorithm, which applies to any orthogonal matrix A of order n:

1. Consider the first column of A, and let i be the last index for which the coefficient is $\neq 0$.
2. While $i > 1$, apply R_α^{-1} to rows $i-1$ and i, so that this index becomes $i-1$. For that purpose, choose α such that $\cot\alpha = a_{i-1}/a_i$, where a_j stands for the first coefficient of row j.

By the previous remark, we get a matrix of the form $\pm 1 \oplus A'$ where A' is an orthogonal matrix of order $n-1$. The type of the canonical form is given by the sign ± 1, the stairs by step 2, and the rest of the canonical form is obtained by applying the algorithm to the matrix A'.

In dimension 3, we get the canonical form of figure 38. The parameters $\alpha, \beta, \gamma \in]0, \pi[$ which appear in the bottom part of this canonical form are called *Euler angles*. In the first case, where three angles are actually needed, we say that the decomposition is *generic*.

Lemma 13 *If* $\alpha, \beta, \gamma \in]0, \pi[$*, there are* $\alpha', \beta', \gamma' \in]0, \pi[$ *such that* $(\mathrm{R}_\gamma \mid \mathrm{id}_1) \circ (\mathrm{id}_1 \mid \mathrm{R}_\beta) \circ (\mathrm{R}_\alpha \mid \mathrm{id}_1)$ *is of the form* $(\mathrm{id}_1 \mid \mathrm{R}_{\alpha'}) \circ (\mathrm{R}_{\beta'} \mid \mathrm{id}_1) \circ (\mathrm{id}_1 \mid \mathrm{R}_{\gamma'})$.



The parameters α', β', γ' are given by the above algorithm. In this case, it happens that no σ appears in the canonical form, and the decomposition is generic. Obviously, the generators satisfy the extra relations of figure 39. Using the same argument as for theorem 1, one proves:

Theorem 9 *The generators* σ*,* R_α *for* $\alpha \in]0, \pi[$*, and the relations of lemma 13 together with those of figure 39 form a presentation of* $\mathbf{O}$.

Similarly, it is possible to get a presentation of the monoidal subcategory $\mathbf{SO}$ of $\mathbf{O}$ whose morphisms are *rotations* with the R_α for $\alpha \in]0, 2\pi[$ as generators. All this can be easily generalized to get a presentation of the monoidal subcategory $\mathbf{U}$ of $\mathbf{GL}(\mathbb{C})$ whose morphisms are *unitary maps* and similarly for the monoidal subcategory $\mathbf{SU}$ of $\mathbf{U}$ whose morphisms are *special unitary maps.* Using this, one gets presentations for the groups $\mathbf{O}_n$, $\mathbf{SO}_n$, $\mathbf{U}_n$, and $\mathbf{SU}_n$.

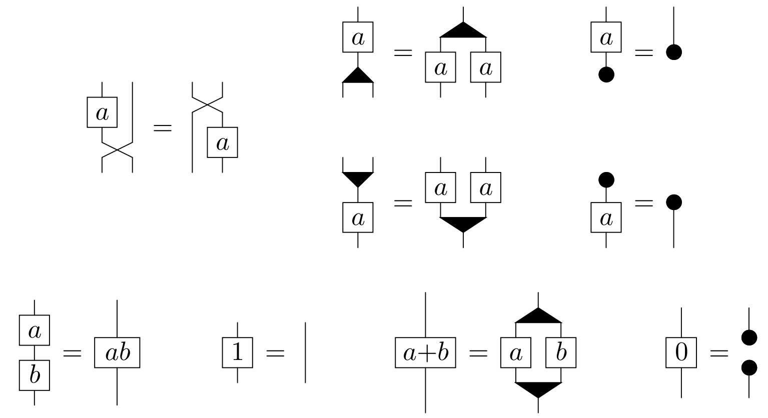


Figure 36: Extra relations for $\mathbf{L}(K)$

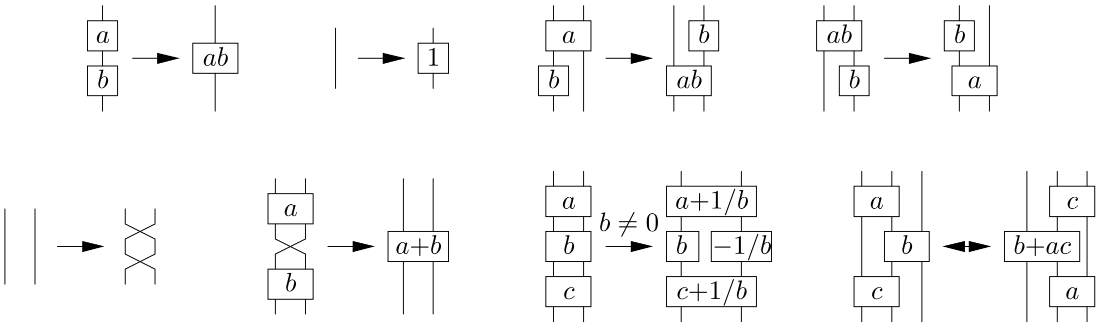


Figure 37: Rules for $\mathbf{GL}(K)$

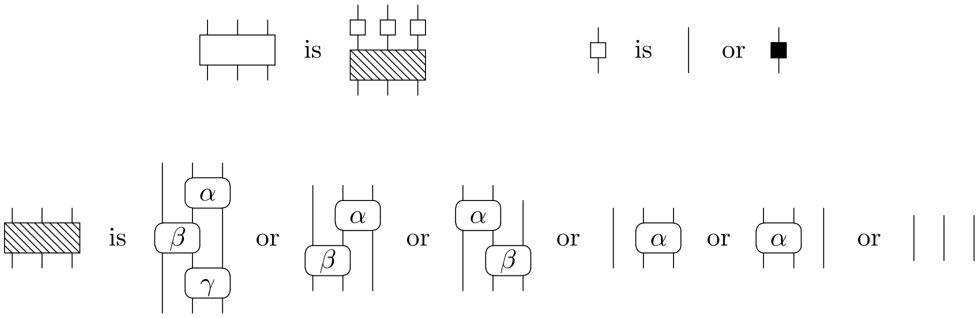


Figure 38: Canonical form for $\mathbf{O}_3$

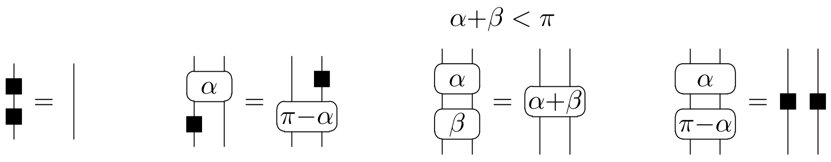
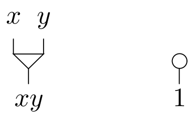


Figure 39: Extra relations for $\mathbf{O}$

4 The classical case

We consider the monoidal category $\mathfrak{F}[k]$ of finite sets $\mathbb{Z}_k^n$ with cartesian product. Again we start with the *Boolean case* $k = 2$. Since $\mathfrak{F}[2]$ is an extension of $\mathbf{L}(\mathbb{Z}_2)$, we have already τ, δ, ε, μ, and η as generators. We introduce two new ones $\mu' : 2 \to 1$ and $\eta' : 0 \to 1$, which are pictured and interpreted as follows:



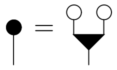
It is well known that any map from $\mathbb{Z}_2^n$ to $\mathbb{Z}_2$ is a polynomial, and any two polynomials define the same map if and only if they are equal modulo the axiom $x^2 = x$. Following [Bur93] and [Laf95a], we get:

Theorem 10 *The generators τ, δ, ε, μ, η, μ', η', and the relations of figure 13 together with those of figure 40 form a presentation of $\mathfrak{F}[2]$.*

Diagrams built with those generators can be seen as *Boolean circuits*: The generators stand respectively for *exchange*, *duplication*, *erasing*, *xor gate*, *false*, *and gate*, and *true*. The last six relations of figure 40 correspond to the following axioms for Boolean algebras:

$$(xy)z = x(yz), \qquad 1x = x, \qquad xy = yx, \qquad (x+y)z = xz + yz, \qquad 0x = 0, \qquad xx = x.$$

Note that η is a superfluous generator:



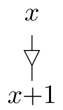
Furthermore, four of the six relations involving η can be removed modulo this definition.

Theorem 10 can also be proved directly by using a suitable notion of canonical form. Moreover, we get:

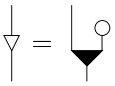
- a presentation of the monoidal subcategory $\mathbf{Z}(\mathbb{Z}_2)$ of $\mathfrak{F}[2]$ whose morphisms are *zero-preserving maps*, by removing η' and all relations involving η' from the presentation of $\mathfrak{F}[2]$;
- a presentation of the monoidal subcategory $\mathbf{A}(\mathbb{Z}_2)$ of $\mathfrak{F}[2]$ whose morphisms are *affine maps*, by removing μ' and all relations involving μ' from the presentation of $\mathfrak{F}[2]$.

4.1 Affine permutations

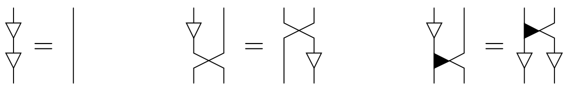
Now, we consider the monoidal subcategory $\mathbf{GA}(\mathbb{Z}_2)$ of $\mathfrak{F}[2]$ whose morphisms are *affine permutations*. We introduce a generator $\nu : 1 \to 1$ for *negation*, which is pictured and interpreted as follows:



Of course, ν is definable in terms of μ and η':



It satisfies the following relations:



Theorem 11 *The generators τ, κ, ν, and the relations of figure 15 together with the above three ones form a presentation of $\mathbf{GA}(\mathbb{Z}_2)$.*

To show this, it suffices to notice that any affine permutation $f : \mathbb{Z}_2^n \to \mathbb{Z}_2^n$ is of the form $h \circ g$ where g is a linear permutation and h is the *translation* defined by $h(x_1, \ldots, x_n) = (x_1 + a_1, \ldots, x_n + a_n)$, with $(a_1, \ldots, a_n) = f(0, \ldots, 0)$. Therefore, we have the canonical form of figure 41 for affine permutations. Any diagram reduces to such a canonical form by the rules of figure 18 and those of figure 42. The latter are derivable: three of them are already in the presentation and the other two are derived in figure 43.

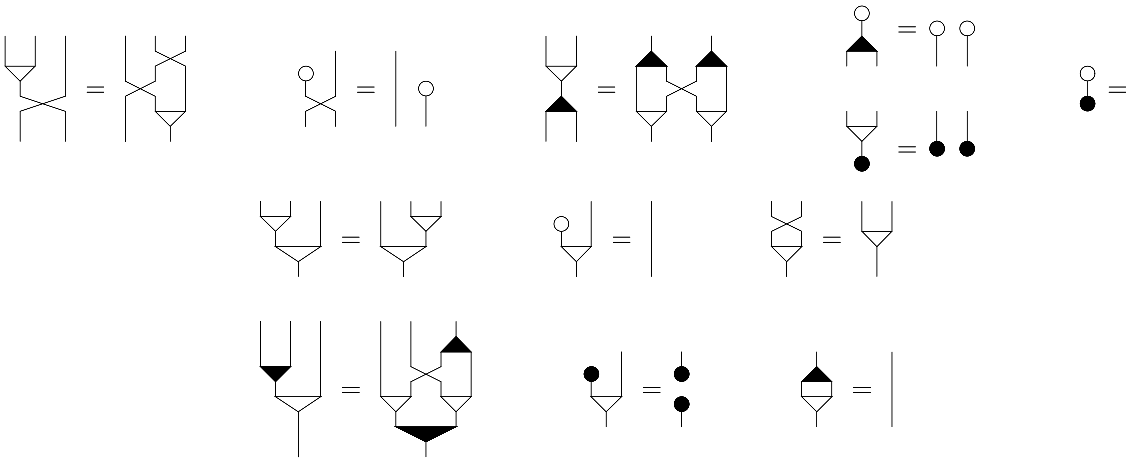


Figure 40: Extra relations for $\mathfrak{F}[2]$

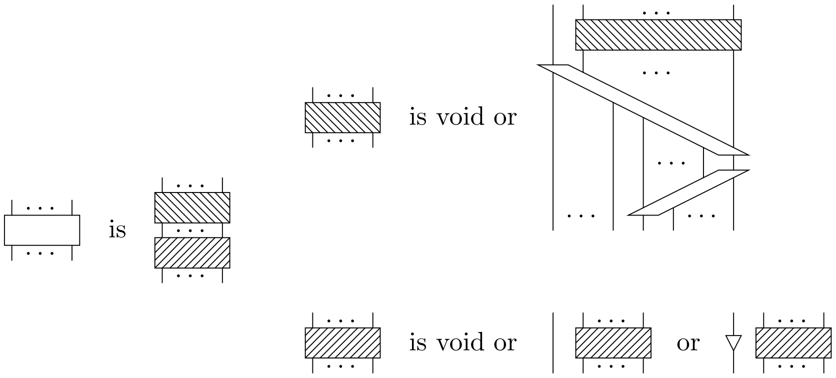


Figure 41: Canonical form for $\mathbf{GA}(\mathbb{Z}_2)$

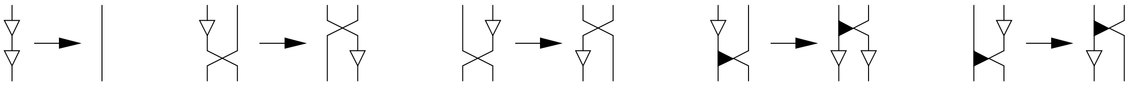


Figure 42: Extra rules for $\mathbf{GA}(\mathbb{Z}_2)$

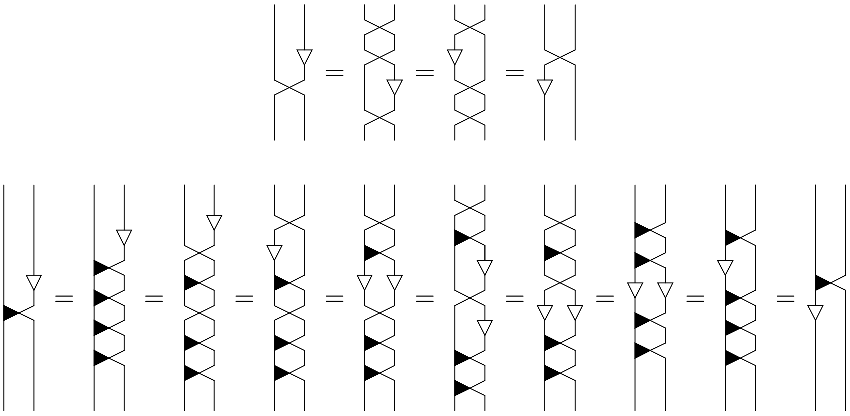
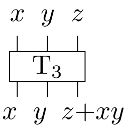


Figure 43: Deriving rules for $\mathbf{GA}(\mathbb{Z}_2)$

4.2 Classical permutations

It is easy to see that, for $n \leq 2$, any permutation of $\mathbb{Z}_2^n$ is affine, but it is not the case for $n = 3$. So we introduce the *3-bit Toffoli gate* $\mathrm{T}_3 : 3 \to 3$, which is pictured and interpreted as follows:



This T_3 is an involution: It corresponds to the transposition of $\mathbb{Z}_2^3$ which exchanges $(1,1,0)$ with $(1,1,1)$. More generally, we introduce the *n-bit Toffoli gate* $\mathrm{T}_n : n \to n$ for each $n \geq 1$, corresponding to the transposition of $\mathbb{Z}_2^n$ which exchanges $(1,\ldots,1,0)$ with $(1,\ldots,1,1)$. For instance, T_1 is the negation ν and T_2 is the linear involution $\tau \circ \kappa$.

It happens that the monoidal subcategory $\mathfrak{S}[2]$ of $\mathfrak{F}[2]$ whose morphisms are permutations is *not finitely generated*. This follows from the following remark:

Lemma 14 *If f is a finite permutation, then $f \times \mathrm{id}_{\mathbb{Z}_2}$ and $\mathrm{id}_{\mathbb{Z}_2} \times f$ are even permutations.*

Indeed, for any decomposition of f into n transpositions, we get a decomposition of $f \times \mathrm{id}_{\mathbb{Z}_2}$ into $2n$ transpositions, and similarly for $\mathrm{id}_{\mathbb{Z}_2} \times f$.

Now, assume that we have cells $\alpha_1 : p_1 \to p_1, \ldots, \alpha_k : p_k \to p_k$ in $\mathfrak{S}[2]$ and let $m = \max(p_1, \ldots, p_k)$. Then any diagram $\phi : n \to n$ built with those generators represents an even permutation of $\mathbb{Z}_2^n$ whenever $n > m$. To show this, it suffices to consider the case of an elementary diagram, and the lemma applies in that case. In particular, the odd permutation T_{m+1} is not definable in terms of $\alpha_1, \ldots, \alpha_k$. Therefore, we consider the monoidal subcategory $\mathfrak{A}[2]$ of $\mathfrak{S}[2]$ whose morphisms are *even permutations.*

Theorem 12 *$\mathfrak{A}[2]$ is contained in the monoidal subcategory of $\mathfrak{S}[2]$ generated by τ, T_1, T_2, and T_3.*

These generators represent transpositions, which are not in $\mathfrak{A}[2]$, but we have the following corollaries:

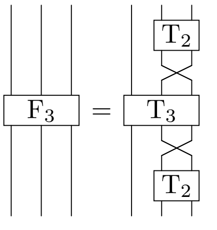
Corollary 1 *$\mathfrak{A}[2]$ is finitely generated.*

Indeed, there are finitely many even permutations of $\mathbb{Z}_2^n$ with $n \leq 3$, and by the theorem, any even permutation of $\mathbb{Z}_2^n$ with $n \geq 4$ is definable in terms of $\tau \mid \mathrm{id}_1$, $\mathrm{id}_1 \mid \tau$, $\mathrm{T}_i \mid \mathrm{id}_1$, and $\mathrm{id}_1 \mid \mathrm{T}_i$ for $i = 1, 2, 3$. In fact, three generators suffice: see [Mus97].

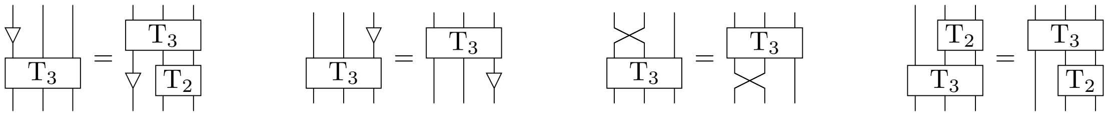
Corollary 2 *$\mathfrak{S}[2]$ is generated by τ and the T_n for $n \geq 1$.*

For any permutation f of $\mathbb{Z}_2^n$, it suffices apply the theorem to f if it is even, or to $\mathrm{T}_n \circ f$ if it is odd.

Theorem 12 is proved in [Mus97], using the *Fredkin gate* $\mathrm{F}_3 : 3 \to 3$ instead of T_3: It corresponds to the transposition of $\mathbb{Z}_2^3$ which exchanges $(1,0,1)$ with $(1,1,0)$. This F_3 is definable in terms of τ, T_2, and T_3:



The proof is based on the fact that any even permutation can be decomposed into 3-cycles. It does not use any notion of canonical form for $\mathfrak{S}[2]$ or $\mathfrak{A}[2]$. Therefore, it is not very helpful for getting presentations of those monoidal categories. Of course, some obvious commutations are satisfied by the generators:



But it is clear that other relations are needed.

Finally, note that $\mathfrak{S}[k]$ is not finitely generated whenever k is an even number, since we have an analogue of lemma 14. On the other hand, Peter Selinger showed that $\mathfrak{S}[k]$ is finitely generated by unary and binary gates whenever k is an odd number (private communication).

Acknowledgements

This paper is dedicated to Albert Burroni who inspired this theory. The author wish also to thank Serge Burckel, Julien Cassaigne, and Peter Selinger for fruitful interactions.

References

[Bur93] **A. Burroni**, *Higher Dimensional Word Problem*, Theoretical Computer Science 115, 1993, pp. 43–62.

[LP91] **Y. Lafont, A. Prouté**, *Church-Rosser property and homology of monoids*, Mathematical Structures in Computer Science 1 (3), Cambridge University Press, 1991, pp. 297–326.

[Laf92] **Y. Lafont**, *Penrose diagrams and 2-dimensional rewriting*, Applications of Categories in Computer Science (ed. M.P. Fourman, P.T. Johnstone & A.M. Pitts), LMSLNS 177, Cambridge University Press, 1992, pp. 191–201.

[Laf95a] **Y. Lafont**, *Equational reasoning with 2-dimensional diagrams*, Term Rewriting, Lecture Notes in Computer Science 909, Springer-Verlag, 1995, pp. 170-195.

[Laf95b] **Y. Lafont**, *A new finiteness condition for monoids presented by complete rewriting systems (after Craig C. Squier)*, Journal of Pure and Applied Algebra 98, North-Holland, 1995, pp. 229-244.

[Law63] **F.L. Lawvere**, *Functorial Semantics of Algebraic Theories*, Proc. Nat. Acad. Sci. USA, 1963.

[Mac71] **S. Mac Lane**, *Categories for the Working Mathematician*, GTM 5, Springer-Verlag, 1971.

[Mas97a] **A. Massol**, *Minimality of the system of seven equations for the category of finite sets*, Theoretical Computer Science, 176, 1997, pp. 347–353.

[Mas97b] **A. Massol**, *Calcul symbolique avec des diagrammes de Penrose*, thèse de doctorat, Université d'Aix-Marseille II, 1997.

[Mus97] **J. Musset**, *Générateurs et relations pour les circuits booléens réversibles*, rapport de stage, Institut de Mathématiques de Luminy, preprint 97-32.[1]

[Squ87] **C. C. Squier**, *Word problems and a homological finiteness condition for monoids*, Journal of Pure and Applied Algebra, 49, 1987, pp. 201–217.

[SOK94] **C. C. Squier, F. Otto, Y. Kobayashi**, *A finiteness condition for rewriting systems*, Theoretical Computer Science, 131, 1994, pp. 271–294.

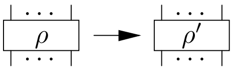
[1] available on `ftp://iml.univ-mrs.fr/pub/lafont/musset.ps.gz`

A Rewriting

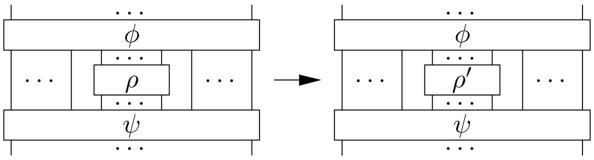
The theory of rewriting is well established in the case of *words* and in the case of *terms*. Following [Laf92], we explain how it can be generalized to *diagrams*. Detailed proofs will not be given here.

A.1 Rewrite rules

A *rewrite system* is given by a set of cells and a set of *rules*. Each rule of the system is of the form $\rho \to \rho'$ where $\rho, \rho' : p \to q$ are two diagrams built with the cells of the system:



If such a rule is given, and if $\phi : m \to i + p + j$ and $\psi : i + q + j \to n$ are any diagrams, we write $\psi \circ (\mathrm{id}_i \mid \rho \mid \mathrm{id}_j) \circ \phi \to \psi \circ (\mathrm{id}_i \mid \rho' \mid \mathrm{id}_j) \circ \phi$. This is called an *elementary reduction*:



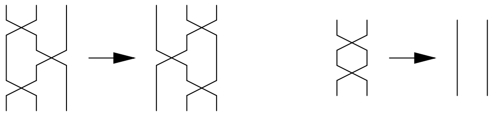
We write $\to^*$ for the *iterated reduction*, which is the reflexive transitive closure of $\to$. Similarly, we write $\leftrightarrow^*$ for the reflexive transitive closure of $\leftrightarrow$, which is itself the symmetric closure of $\to$. Note that $\phi \leftrightarrow^* \psi$ when ϕ and ψ are equivalent modulo the rules (considered as relations).

We say that a diagram ϕ is *reduced* if there is no ϕ' such that $\phi \to \phi'$. We say that ψ is a *reduced form* of ϕ if $\phi \to^* \psi$ and ψ is reduced. Note that the commutation rule does not count as a reduction. For the question of finding canonical decompositions, independently of the rewrite rules, see [Mas97b].

A.2 Termination

A rewrite system is *terminating* (or *noetherian*) if there is no infinite reduction $\phi_0 \to \phi_1 \to \phi_2 \to \phi_3 \to \cdots$ In other words, any reduction strategy terminates and leads to a reduced form.

Consider for instance the rewrite system of lemma 2 for $\mathfrak{S}$:



This system is terminating because the first rule moves one cell to the right and the second rule removes two cells. To make this argument precise, we define the natural number $\|\phi\|$ for any diagram ϕ as follows:

- If $\xi = \mathrm{id}_i \mid \sigma \mid \mathrm{id}_j$ is an elementary diagram, then $\|\xi\| = j + 1$.
- If $\phi = \xi_1 \circ \cdots \circ \xi_n$ where $\xi_1, \ldots, \xi_n$ are elementary diagrams, then $\|\phi\| = \|\xi_1\| + \cdots + \|\xi_n\|$.

This definition does not depend on the decomposition of ϕ because the number of inputs of the generator σ is the same as its number of outputs. Since we have $\|\phi\| > \|\phi'\|$ whenever $\phi \to \phi'$, the length of any reduction starting from a diagram ϕ is bounded by $\|\phi\|$.

The rewrite system of figure 9 for $\mathfrak{F}$ is also terminating, but the previous argument cannot be extended. Instead, we interpret any diagram $\phi : p \to q$ as a strictly monotonic map $[\phi] : \mathbb{N}^{*p} \to \mathbb{N}^{*q}$, where $\mathbb{N}^*$ is the set of strictly positive integers, and $\mathbb{N}^{*n}$ is equipped with the following partial order (*product order*):

$$(x_1, \ldots, x_n) \leq (y_1, \ldots, y_n) \text{ whenever } x_1 \leq y_1,\ \ldots,\ x_n \leq y_n.$$

It suffices to give the interpretation of each generator: see figure 44. For any strictly monotone maps $f, g : \mathbb{N}^{*p} \to \mathbb{N}^{*q}$, we write $f < g$ if $f(x_1, \ldots, x_p) < g(x_1, \ldots, x_p)$ for all $(x_1, \ldots, x_p) \in \mathbb{N}^{*p}$. This relation is compatible with horizontal and vertical composition, and we have $[\rho] > [\rho']$ for each rule $\rho \to \rho'$: see figure 45. Therefore, $[\phi] > [\phi']$ whenever $\phi \to \phi'$. Finally, the length of any reduction starting from a diagram ϕ is bounded by $n_1 + \cdots + n_q$ where $(n_1, \ldots, n_q)$ is $[\phi](1, \ldots, 1)$.

We conjecture that the rewrite system of figure 28 for $\mathbf{L}(\mathbb{Z}_2)$ is also terminating, but another method is needed to show this, since there is no way of interpreting the cell $\varepsilon : 1 \to 0$ as a strictly monotone map.

A.3 Confluence

Termination ensures the existence of a reduced form for any diagram, but uniqueness is also needed to decide whether two diagrams are equivalent. The following result is standard in rewriting theory:

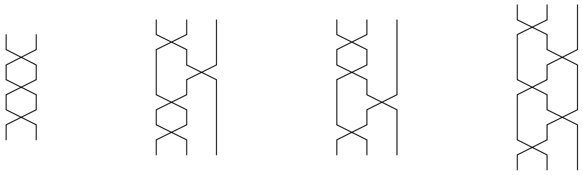
Lemma 15 *For a terminating rewrite system, the following properties are equivalent :*

1. *If $\phi \rightarrow^* \psi$ and $\phi \rightarrow^* \psi'$ where ψ and ψ' are reduced, then $\psi = \psi'$ (uniqueness of the reduced form);*
2. *If $\phi \leftrightarrow^* \phi'$, then there exists ψ such that $\phi \rightarrow^* \psi$ and $\phi' \rightarrow^* \psi$ (Church-Rosser property);*
3. *If $\phi \rightarrow^* \phi'$ and $\phi \rightarrow^* \phi''$, then there exists ψ such that $\phi' \rightarrow^* \psi$ and $\phi'' \rightarrow^* \psi$ (confluence);*
4. *If $\phi \rightarrow \phi'$ and $\phi \rightarrow \phi''$, then there exists ψ such that $\phi' \rightarrow^* \psi$ and $\phi'' \rightarrow^* \psi$ (local confluence).*

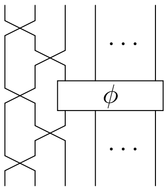
$$\begin{tikzcd}
& \phi \arrow[dl, "*"'] \arrow[dr, "*"] & \\
\psi & = & \psi'
\end{tikzcd}
\qquad
\begin{tikzcd}
\phi \arrow[rr, leftrightarrow, "*"] \arrow[dr, dashed, "*"'] & & \phi' \arrow[dl, dashed, "*"] \\
& \psi &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \phi \arrow[dl, "*"'] \arrow[dr, "*"] & \\
\phi' \arrow[dr, dashed, "*"'] & & \phi'' \arrow[dl, dashed, "*"] \\
& \psi &
\end{tikzcd}
\qquad
\begin{tikzcd}
& \phi \arrow[dl] \arrow[dr] & \\
\phi' \arrow[dr, dashed, "*"'] & & \phi'' \arrow[dl, dashed, "*"] \\
& \psi &
\end{tikzcd}$$

Indeed, it is easy to show that each item implies the next one, and the last one implies the first one by noetherian induction. Such a rewrite system is called *canonical.* In the case of words or terms, it suffices to check local confluence for a finite number of conflicts between rules called *critical peaks.*

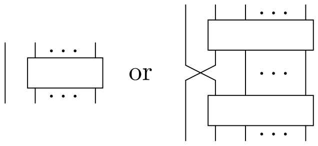
In the rewrite system of lemma 2 for $\mathfrak{S}$, there are four obvious conflicts:



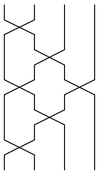
Each of the above diagrams contains two instances of the left member of a rule, and those instances have a common cell. However, this list is not complete, and indeed, there was a gap in [Laf92]. Because of the commutation rule, the general form of the last conflict is the following:



We call this a *global conflict.* Fortunately, its confluence can be proved by noetherian induction on the diagram ϕ: see figure 46. Hence, it suffices to consider the case where ϕ is reduced. By induction on the size, it is easy to see that such a diagram is of one of the following two types:



Finally, it suffices to consider one extra conflict, corresponding to the case where ϕ is of the second type:



To sum up, there are five *critical peaks* and all of them are confluent: see figure 47. Therefore, our system is canonical. In fact, the reduced forms are the canonical forms of section 2.1, so that confluence also follows from lemmas 1 and 2.

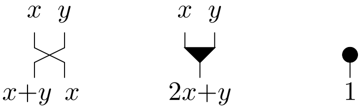


Figure 44: Interpretation of the generators

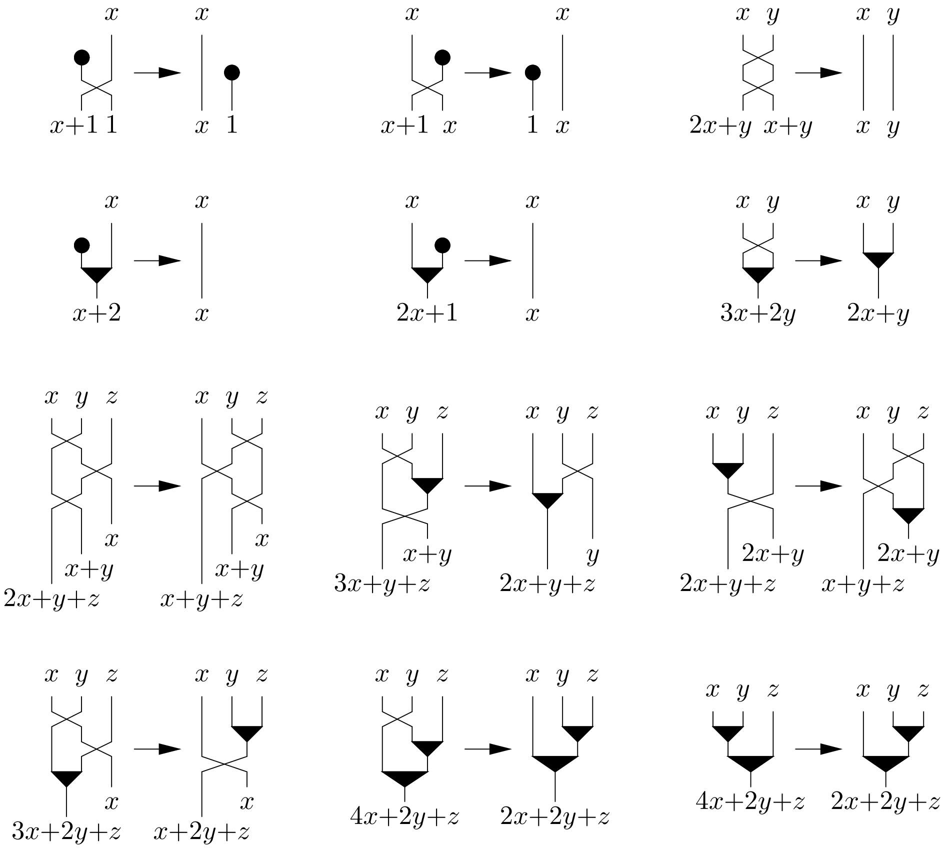


Figure 45: Termination for 𝔉

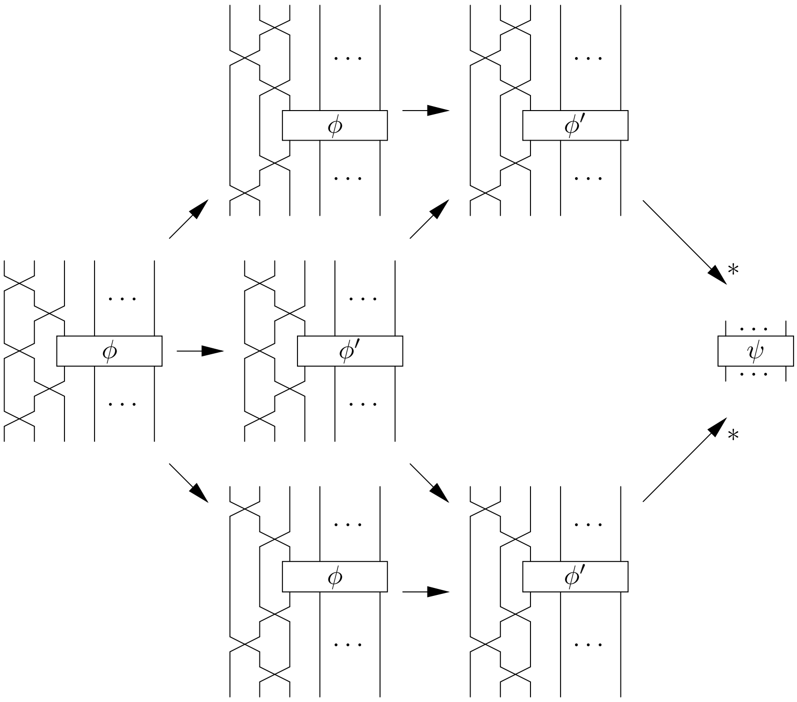


Figure 46: Proving the confluence of a global conflict by noetherian induction

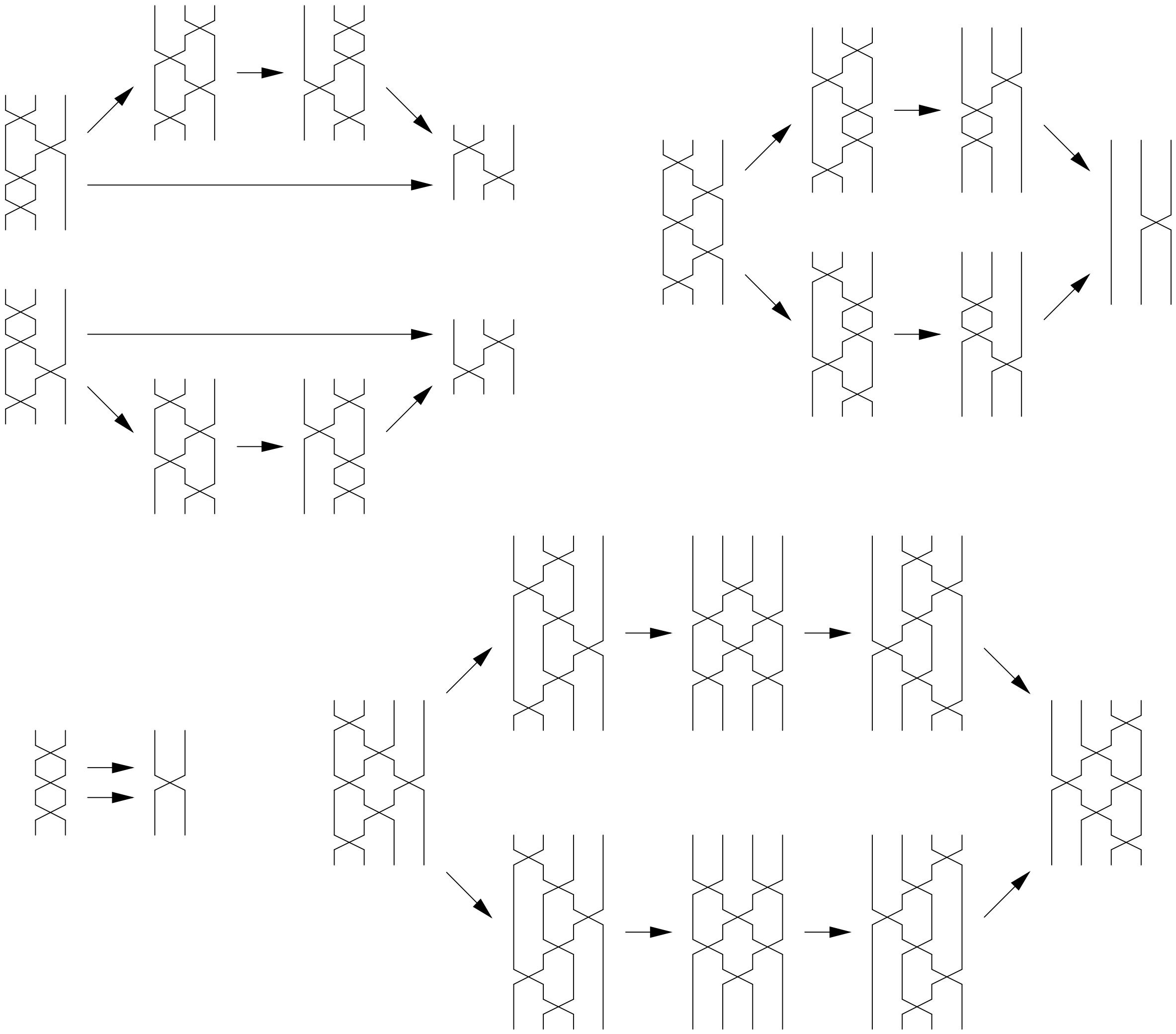


Figure 47: Confluence of the five critical peaks for $\mathfrak{S}$

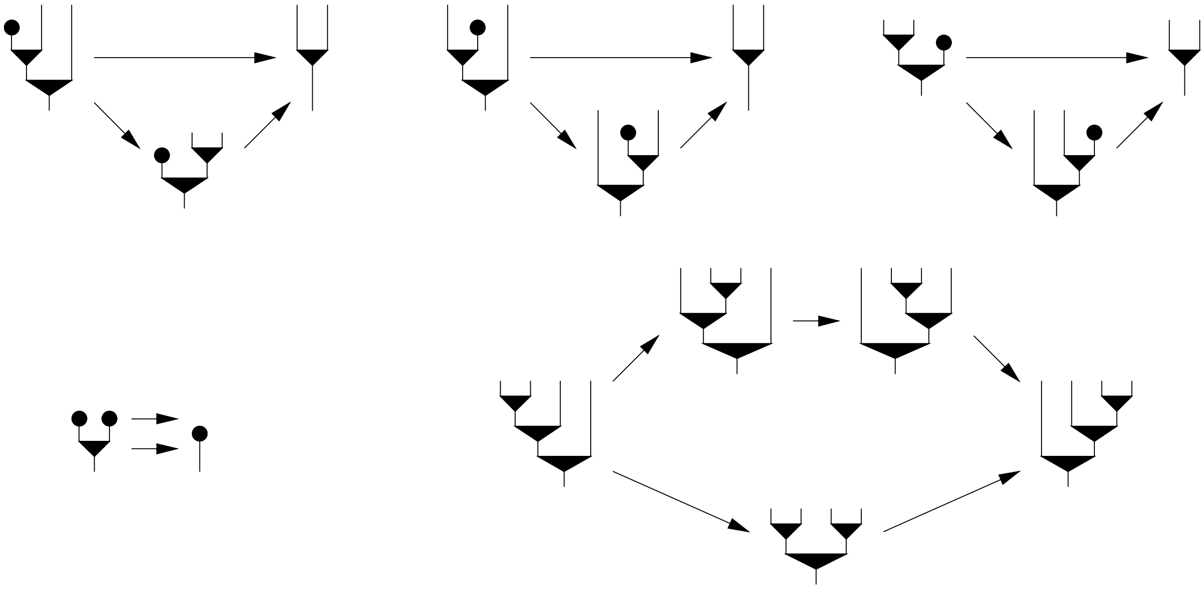
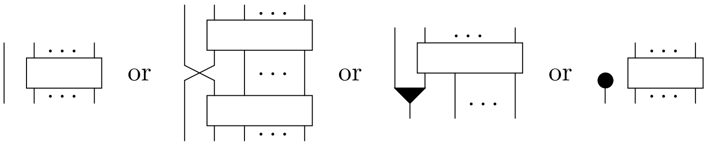


Figure 48: Confluence of the five critical peaks for $\mathfrak{M}$

The rewrite system of figure 4 for $\mathfrak{M}$ is also canonical. In this case, there is no global conflict and the five critical peaks are confluent: see figure 48. In fact, this system can be identified with the well-know term rewrite system for the theory of monoids:

$$(xy)z \to x(yz), \qquad 1x \to x, \qquad x1 \to x.$$

In the rewrite system of figure 9 for $\mathfrak{F}$, there are six global conflicts: see figure 49. Again, it suffices to consider the case where ϕ is reduced, and such a diagram is of one of the following four types:



Furthermore, if ϕ is of the fourth type, we get a *reducible conflict*. This means that a third rule applies, and the confluence of this conflict can be deduced from the confluence of a smaller conflict: see figure 50. Finally, the global conflicts lead to $6 \times 3 = 18$ critical peaks, and all together, there are 68 critical peaks: see figure 51. The confluence can be checked case by case, but it also follows from section 2.3.

In our examples, the study of critical peaks is not the only way to prove confluence, but the notion is interesting anyway:

- In [Squ87] and [SOK94], critical peaks are used to compute *homological invariants* of a monoid. See also [LP91] and [Laf95b] for an introduction. This should be extended to monoidal categories.
- In [Mac71], the critical peaks of figure 48 are used to prove the *coherence theorem* for (non strict) *monoidal categories*. Similarly, the canonical system for $\mathfrak{F}$ can be used to prove the coherence theorem for *symmetric monoidal categories*.

A.4 Terms versus diagrams

Any first-order equational theory with n function symbols and r equations corresponds to a presentation by $n+3$ generators and $r+3n+7$ relations. This presentation consists of:

- one generator $\alpha : m \to 1$ for each function symbol of arity m, and the 3 generators $\tau : 2 \to 2$, $\delta : 1 \to 2$, and $\varepsilon : 1 \to 0$;
- one relation for each equation, 3 relations for each function symbol (see figure 52), and the dual of the 7 relations for $\mathfrak{F}$.

For instance, the theory of Boolean algebras with 4 function symbols and 10 equations corresponds to the presentation of theorem 10 for $\mathfrak{F}[2]$ with 7 generators and 29 relations. This idea comes from [Bur93]. It is based on Lawvere's interpretation of algebraic theories by means of *Cartesian categories*: see [Law63].

There is a similar correspondence for rewrite systems: Any *left linear* canonical term rewrite system with n function symbols, r rules, and p critical peaks corresponds to a canonical diagram rewrite system with $n+3$ generators, $r+4n+12$ rules, and $p+4r+n^2+14n+68$ critical peaks. In addition to the r original rules, we have 4 extra rules for each function symbol (see figure 53) and the dual of the 12 rules for $\mathfrak{F}$.

Left linearity is a strong restriction: It means that variables occur once in the left member of each rule. For instance, $x(y+z) \to xy+xz$ is left linear, but not $x+x \to 0$. Here is the crucial observation: the left member of such a rule corresponds to a diagram (in fact a tree) with no τ, δ, or ε, and there is no critical peak between the original rules and those coming from the rules for $\mathfrak{F}$. Furthermore, one can check that all new global conflicts are reducible. Therefore, it suffices to consider the 5 types of critical peaks between the 3 types of rules (see figure 54).

This proliferation of rules and critical peaks is the price to pay for a decomposition of equational reasoning into more elementary steps. However, we may find canonical systems which are not of the above form:

- If our conjecture about the termination of the rewrite system of figure 28 holds, then we get a canonical system for $\mathbf{L}(\mathbb{Z}_2)$, whereas there is no canonical term rewrite system for the theory of vector spaces over $\mathbb{Z}_2$, simply because the commutativity $x+y=y+x$ cannot be oriented. This problem is usually handled by introducing *rewriting modulo associativity and commutativity*, but we claim that our approach is less ad hoc.
- There are useful algebraic structures, such as *braids*, *tangles*, and *Hopf algebras*, which are naturally expressed with diagrams, but not with terms. Unfortunately, we have no interesting example of canonical system of this kind. For instance, the existence of a finite canonical rewrite system for the monoidal category of braids is an open question.

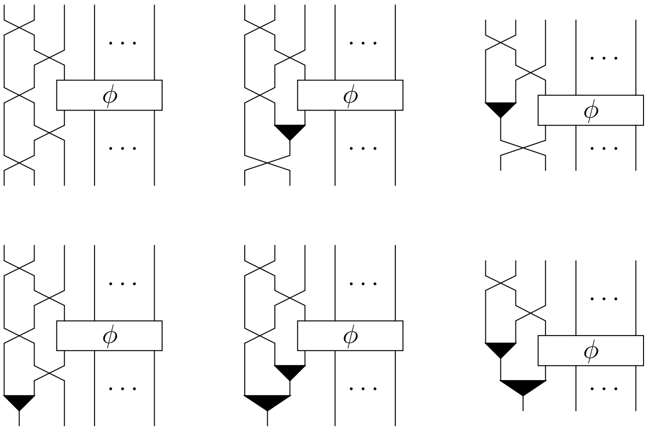


Figure 49: The six global conflicts for $\mathfrak{F}$

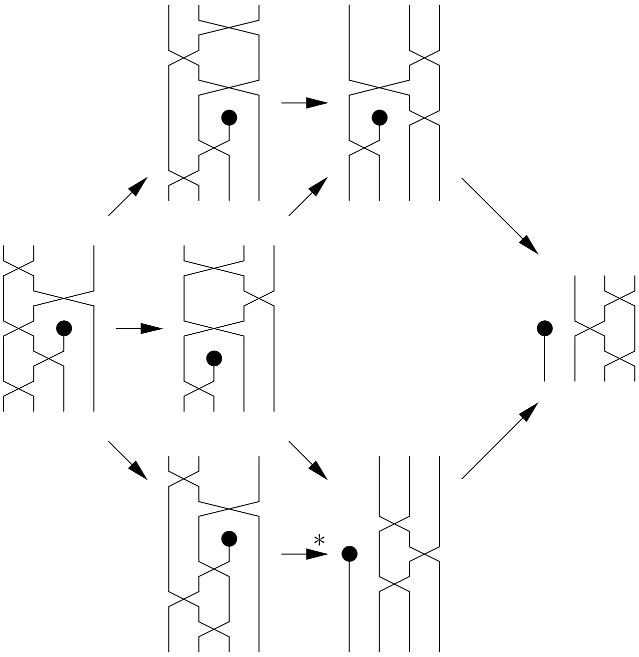


Figure 50: Confluence of a reducible conflict

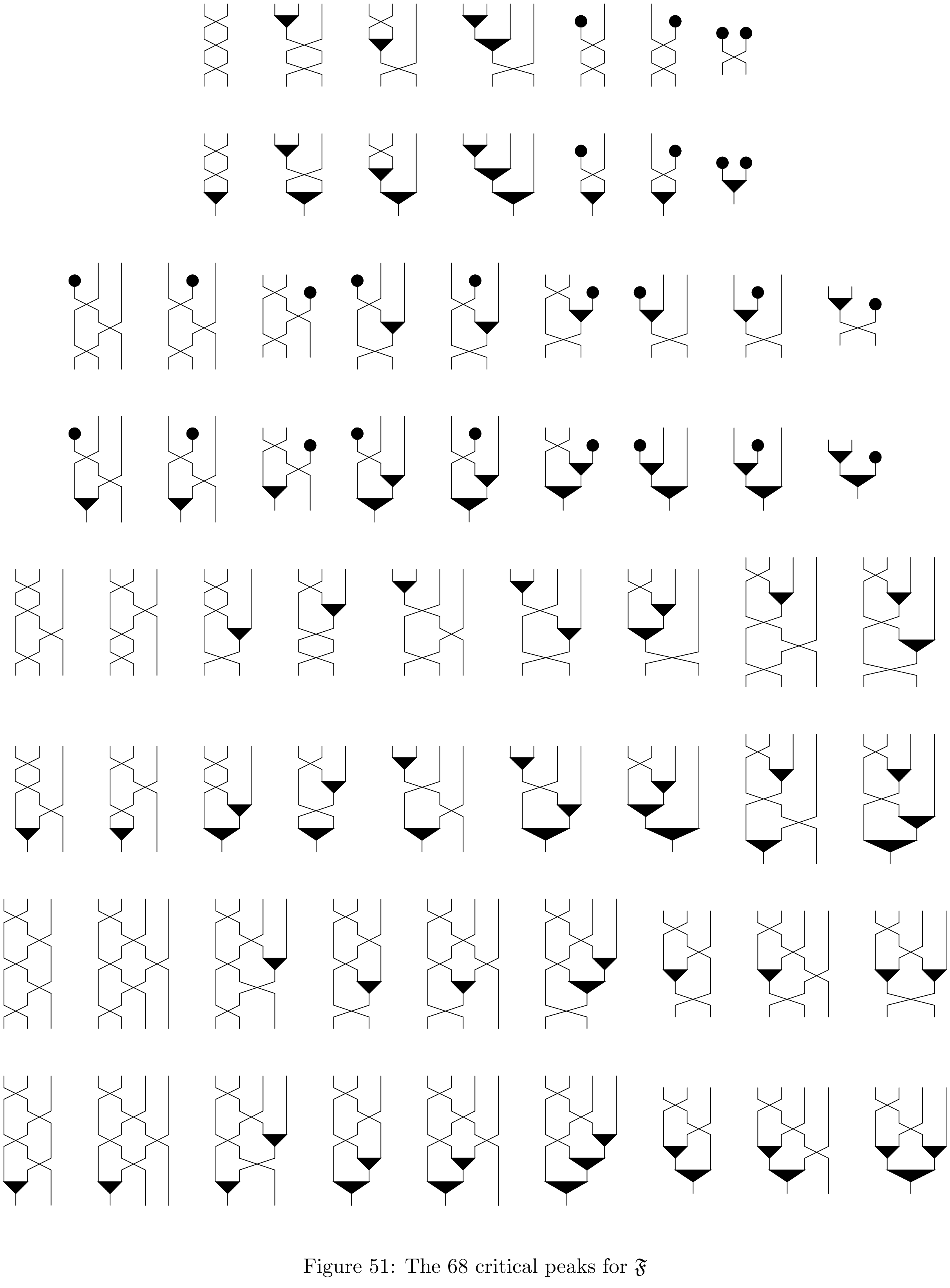


Figure 51: The 68 critical peaks for $\mathfrak{F}$

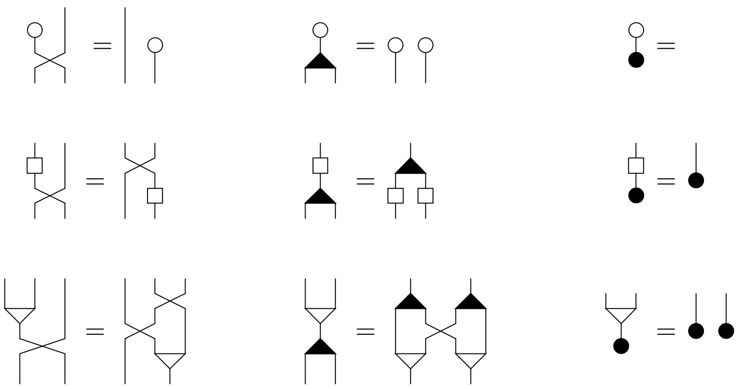


Figure 52: Relations for function symbols of arity 0, 1, or 2

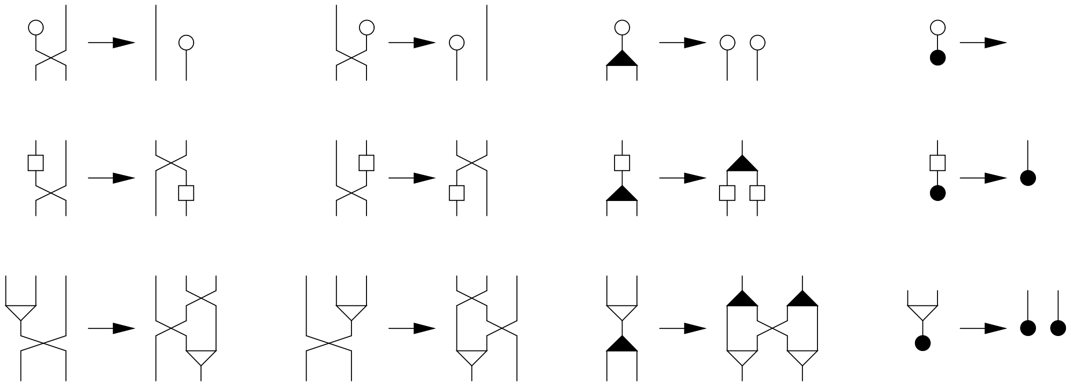


Figure 53: Rules for function symbols of arity 0, 1, or 2

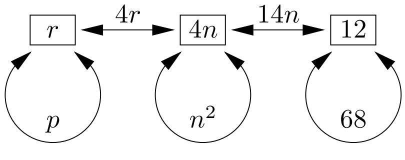
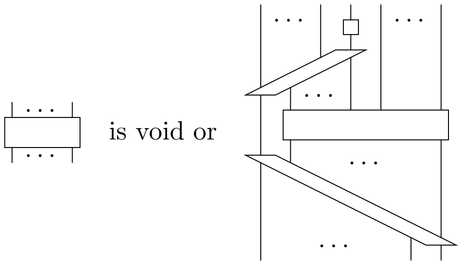


Figure 54: The 5 types of critical peaks between the 3 types of rules

B Inverting matrices

Knowing a decomposition of an invertible matrix into elementary ones, it is easy to compute its inverse. Therefore, any notion of canonical form for $\mathbf{GL}(K)$ should give an inversion algorithm for matrices.

Consider the following canonical form for $\mathbf{GL}(K)$, which generalizes the first canonical form for $\mathbf{GL}(\mathbb{Z}_2)$:



It suggests a variant of the Gauss algorithm, which applies to any invertible matrix A of order n over K:

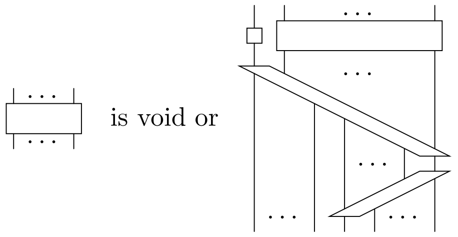
1. Create a counter k with initial value 1, and two matrices P and Q with initial value $\mathbf{I}$ (the identity matrix of order n).
2. Consider the last row of A, and let j be the last index for which the coefficient a is not 0.
3. Divide column j by a so that this coefficient becomes 1, and apply the same operation to P.
4. While $j > k$, apply an elementary operation to columns $j-1$ and j, so that this index becomes $j-1$, and apply the same operation to P.
5. Now, if we consider column p, n is the last index i for which the coefficient is not 0 (in fact, it is 1).
6. While $i > k$, apply an elementary operation to rows $i-1$ and i, so that this index becomes $i-1$, and apply the same operation to Q.
7. If $k < n$, increment k and go back to step 2. Otherwise, return the product PQ.

Note that the final value of A is $\mathbf{I}$. To see that this algorithm computes the inverse, we write $A_0, \dots, A_n$ for the successive values of A, and similarly for P and Q. Since $P_0 = Q_0 = \mathbf{I}$, it is easy to see that $A_i = Q_i A_0 P_i$ for each i. In particular, $A_n = Q_n A_0 P_n = \mathbf{I}$, so that $A_0^{-1} = P_n Q_n$.

In the case where the last coefficient of the last row is not 0, it amounts to applying the following formula:

$$\begin{pmatrix} A & u \\ v & b \end{pmatrix}^{-1} = \begin{pmatrix} C^{-1} & -\frac{1}{b}C^{-1}u \\ -\frac{1}{b}vC^{-1} & \frac{1}{b}+\frac{1}{b^2}vC^{-1}u \end{pmatrix} \text{ where } C = A - \tfrac{1}{b}uv.$$

Consider now the second canonical form for $\mathbf{GL}(K)$, which generalizes the first canonical form for $\mathbf{GL}(\mathbb{Z}_2)$:



Here the antistairs are obtained by solving a linear system, which means that we must first compute the inverse of some submatrix. We shall not give the details here, but in the case where the matrix obtained by forgetting the first row and the first column is invertible, it amounts to applying the following formula:

$$\begin{pmatrix} a & u \\ v & B \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{c} & -\frac{1}{c}uB^{-1} \\ -\frac{1}{c}B^{-1}v & B^{-1}+\frac{1}{c}B^{-1}vuB^{-1} \end{pmatrix} \text{ where } c = a - uB^{-1}v.$$

Note that the above formulas are two instances of a more general one:

$$\begin{pmatrix} A & U \\ V & B \end{pmatrix}^{-1} = \begin{pmatrix} C^{-1} & -C^{-1}UB^{-1} \\ -B^{-1}VC^{-1} & B^{-1}+B^{-1}VC^{-1}UB^{-1} \end{pmatrix} \text{ where } C = A - UB^{-1}V.$$