



HAL
open science

A First Order Theory of Diagram Chasing

Assia Mahboubi, Matthieu Piquerez

► **To cite this version:**

Assia Mahboubi, Matthieu Piquerez. A First Order Theory of Diagram Chasing. 32nd EACSL Annual Conference on Computer Science Logic 2024 (CSL'24), Feb 2024, Naples (Université Federico II), Italy. hal-04266479v1

HAL Id: hal-04266479

<https://hal.science/hal-04266479v1>

Submitted on 2 Nov 2023 (v1), last revised 4 Dec 2023 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

A First Order Theory of Diagram Chasing

Assia Mahboubi 

Nantes Université, École Centrale Nantes, CNRS, INRIA, LS2N, UMR 6004, France

Matthieu Piquerez

Nantes Université, École Centrale Nantes, CNRS, INRIA, LS2N, UMR 6004, France

Abstract

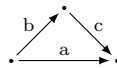
This paper discusses the formalization of proofs “by diagram chasing”, a standard technique for proving properties in abelian categories. We discuss how the essence of diagram chases can be captured by a simple many-sorted first-order theory, and we study the models and decidability of this theory. The longer-term motivation of this work is the design of a computer-aided instrument for writing reliable proofs in homological algebra, based on interactive theorem provers.

2012 ACM Subject Classification Theory of computation → Logic and verification

Keywords and phrases Diagram chasing, formal proofs, abelian categories, decidability

1 Introduction

Homological algebra [8] attaches and studies a sequence of algebraic objects, typically groups or modules, to a certain space, e.g., a ring or a topological space, in order to better understand the latter. In this field, diagram chasing is a major proof technique, which is usually carried out via a form of diagrammatic reasoning on abelian categories. A diagram can be seen as a functor $F: J \rightarrow \mathcal{C}$, whose domain J , the indexing category, is a small category [12]. Diagrams are usually represented as directed multi-graphs, also called *quivers*, whose vertices are decorated with objects of \mathcal{C} , and arrows with morphisms. Paths in such graphs thus correspond to chains of composable arrows. Diagrams allow for visualizing the existence of certain morphisms, and to study identities between certain compositions of morphisms. In particular, a diagram *commutes* when any two paths with same source and target lead to identical composite. For instance, the commutativity of the following diagram:



asserts that morphism a is equal to the composition of morphisms c and b , denoted $b \circ c$. Commutativity of diagrams in certain categories can be used to state more involved properties, and diagram chasing essentially consists in establishing the existence, injectivity, surjectivity of certain morphisms, or the exactness of some sequences, using hypotheses of the same nature. The *five lemma* or the *snake lemma* are typical examples of proofs “by diagram chasing”, also called diagram chases. On paper, diagrams help conveying in a convincing manner proofs otherwise consisting of overly pedestrian chains of equations. The tension between readability and elusiveness may however become a challenge. For instance, diagram chases may rely on non-trivial duality arguments, that is, on the fact that a property about diagrams in any abelian category remains true after reversing all the involved arrows, although the replay of a given proof *mutatis mutandis* cannot be fulfilled in general.

Motivated in part by the second author’s experience in writing intricate diagram chases (see for instance [11, p.337]), this work aims at laying the foundations of a computer-aided instrument for writing reliable proofs in homological algebra, based on interactive theorem provers. The present article discusses the design of a formal language for statements of properties amenable to proofs by diagram chasing, according to three objectives. The first is *simplicity and expressivity*: this language should be at the same time simple enough to

be implemented in a formal library, and expressive enough to encompass the desired corpus of results. Then, *duality* arguments in proofs shall follow directly from a meta-property of the language. Finally, the corresponding proof system should allow for *effective* proofs of commutativity clauses, that is, proving that the commutativity of some diagram follows from the commutativity of some other diagrams, so that these proofs can eventually be automated.

► **Definition 1.** We define the many-sorted signatures $\mathring{\Sigma}$, resp. Σ , as follows: the sorts of signature $\mathring{\Sigma}$, resp. Σ , are finite, resp. acyclic finite, quivers. The symbols of $\mathring{\Sigma}$, resp. Σ , consists of one function symbol $\text{restr}_{m:Q' \rightarrow Q}$, of arity $Q \rightarrow Q'$, per each quiver morphism, resp. each embedding between acyclic quivers, m and one predicate commute_Q on sort Q for each finite, resp. acyclic finite, quiver Q .

The thesis of the present article is that signature Σ fulfills the three above objectives. We validate this thesis by giving a first-order theory for diagrams over small and abelian small categories respectively. We state and prove a duality theorem and motivate the choice of Σ over the possibly more intuitive $\mathring{\Sigma}$ by the effectiveness objective. The rest of the article is organized as follows. We first fix some vocabulary and notations in Section 2, so as in particular to make Definition 1 precise. Then, Section 3 introduces a theory for small categories and describes its models, Section 4 discusses duality, before Section 5 provides an analogue study for abelian categories. Last, we prove in Section 6 the decidability, resp. undecidability, of commutativity clauses in Σ , resp. $\mathring{\Sigma}$, before concluding in Section 7.

2 Preliminaries

In all what follows, $\mathbb{N} := \{0, 1, \dots\}$ refers to the set of non-negative integers. If $k \in \mathbb{N}$, then $[k]$ denotes the finite collection $\{0, \dots, k-1\}$. We denote $\text{card}A$ the cardinal of a finite set A . We use the notation id for the identity map.

2.1 Quivers

► **Definition 2** (General quiver, dual). A general quiver \mathcal{Q} is a quadruple $(V_{\mathcal{Q}}, A_{\mathcal{Q}}, s_{\mathcal{Q}}: A_{\mathcal{Q}} \rightarrow V_{\mathcal{Q}}, t_{\mathcal{Q}}: A_{\mathcal{Q}} \rightarrow V_{\mathcal{Q}})$ where $V_{\mathcal{Q}}$ and $A_{\mathcal{Q}}$ are two sets. The element of $V_{\mathcal{Q}}$ are called the vertices of \mathcal{Q} and the element of $A_{\mathcal{Q}}$ are called arrows. If $a \in A_{\mathcal{Q}}$, $s_{\mathcal{Q}}(a)$ is called the source of a and $t_{\mathcal{Q}}(a)$ is called its target. The dual of a quiver \mathcal{Q} is the quiver $\mathcal{Q}^* := (V_{\mathcal{Q}}, A_{\mathcal{Q}}, t_{\mathcal{Q}}, s_{\mathcal{Q}})$, which swaps the source and the target maps of \mathcal{Q} .

► **Definition 3** (Morphism, embedding, restriction). A morphism of quiver $m: \mathcal{Q} \rightarrow \mathcal{Q}'$, is the data of two maps $m_V: V_{\mathcal{Q}} \rightarrow V_{\mathcal{Q}'}$ and $m_A: A_{\mathcal{Q}} \rightarrow A_{\mathcal{Q}'}$ such that $m_A \circ s_{\mathcal{Q}} = s_{\mathcal{Q}'} \circ m_A$ and $m_A \circ t_{\mathcal{Q}} = t_{\mathcal{Q}'} \circ m_A$. Such a morphism is called an embedding of quivers if moreover both m_V and m_A are injective. In this case we write $m: \mathcal{Q} \hookrightarrow \mathcal{Q}'$.

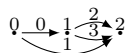
If A is a subset of $A_{\mathcal{Q}}$, the (spanning) restriction of \mathcal{Q} to A denoted $\mathcal{Q}|_A$ is the quiver $(V_{\mathcal{Q}}, A, s_{\mathcal{Q}}|_A, t_{\mathcal{Q}}|_A)$. There is a canonical embedding $\mathcal{Q}|_A \hookrightarrow \mathcal{Q}$.

We denote by \emptyset the empty quiver with no vertex and no arrow, and by \mathring{S} the set of quivers Q such that V_Q and A_Q are finite subsets of \mathbb{N} . In this article, a *quiver* refers to an element of \mathring{S} . We use a non-cursive Q for elements of \mathring{S} , and a cursive \mathcal{Q} for general quivers.

For the sake of readability, we use drawings to describe some elements of \mathring{S} , as for instance:



For a quiver Q denoted by such a drawing, the convention is that $V_Q = [\text{card}V_Q]$ and $A_Q = [\text{card}A_Q]$. From left to right, the drawn vertices correspond to $0, 1, \dots, \text{card}V_Q - 1$. Arrows are then numbered by sorting pairs (s_Q, t_Q) in increasing lexicographical order, as in:



We also use drawings to denote embeddings. The black part represents the domain of the morphism, the union of black and gray parts represents its codomain. Here is an example of an embedding of the quiver $\cdot \xrightarrow{\quad} \cdot$ into the quiver drawn above.



► **Definition 4** (Path-quiver). *The path-quiver of length k , denoted PQ_k , is the quiver with $k + 1$ vertices and k arrows $([k + 1], [k], \text{id}, (i \mapsto i + 1))$.*

A path-quiver can be drawn as:



with at least one vertex. Such a path-quiver is called *nontrivial* if it has at least two vertices.

If $0 \leq k \leq l$ are two integers, we denote by $\text{sp}_{k,l}: \text{PQ}_k \hookrightarrow \text{PQ}_l$ the leftmost embedding of PQ_k into PQ_l , i.e., such that $(\text{sp}_{k,l})_V(0) = 0$. If k and l are clear from the context, we draw $\text{sp}_{k,l}$ as \sim if $k \neq 0$ and as \curvearrowright if $k = 0$. Moreover, we denote by $\text{tp}_{k,l}: \text{PQ}_k \hookrightarrow \text{PQ}_l$ the rightmost embedding of PQ_k into PQ_l , i.e., such that $(\text{tp}_{k,l})_V(k) = l$. The corresponding drawings are \curvearrowleft and \dashrightarrow . Moreover, if P is a nontrivial path-quiver, we define $\text{st}_P: \cdot \cdot \hookrightarrow P$ to be the embedding mapping the first vertex on the leftmost vertex of P and the second vertex on the rightmost vertex of P . We denote this embedding \dashrightarrow .

If \mathcal{Q} is a general quiver, a morphism of the form $p: \text{PQ}_k \rightarrow \mathcal{Q}$, for some k , is called a *path of \mathcal{Q} from u to v of length k* , where $u := p(0)$ and $v := p(k)$. Two paths $p_1: P_1 \hookrightarrow \mathcal{Q}$, $p_2: P_2 \hookrightarrow \mathcal{Q}$ of \mathcal{Q} have the same extremities if $p_1 \circ \text{st}_{P_1} = p_2 \circ \text{st}_{P_2}$. We denote by $\mathcal{BP}_{\mathcal{Q}}$ the set of pair of paths of \mathcal{Q} having the same extremities. Let \mathcal{Q}' be another general quiver and $m: \mathcal{Q} \rightarrow \mathcal{Q}'$ be a morphism. Then we define $m_*(p) := m \circ p$.

A general quiver is *acyclic* if any path of this quiver is an embedding. The set of acyclic quivers in \mathring{S} is denoted by S .

► **Definition 5** (Free category). *For a general quiver \mathcal{Q} , the free category over \mathcal{Q} , denoted $\langle \mathcal{Q} \rangle$ is the category with objects $\text{Ob}_{\langle \mathcal{Q} \rangle} = V_{\mathcal{Q}}$ whose morphisms $\text{Hom}_{\langle \mathcal{Q} \rangle}(u, v)$, for two vertices u and v are the paths from u to v . The identity map from u to u is the empty path, and the composition is defined as the concatenation of paths.*

Note that a morphism m of quivers induces a functor between the corresponding categories that we denote by Φ_m . In the other direction, any small category \mathcal{C} has an *underlying quiver*.

2.2 Diagrams

We can now introduce diagrams in a category, and a few useful specific examples thereof.

► **Definition 6** (Diagram). *For any category \mathcal{C} and any quiver Q , a diagram in \mathcal{C} over Q is a functor from $\langle Q \rangle$ to \mathcal{C} .*

Let P be a path-quiver from vertex u to vertex v . To a diagram $D: \langle P \rangle \rightarrow \mathcal{C}$ over P one can associate the corresponding composition of morphisms in the category \mathcal{C} , which is an element of $\text{Hom}_{\mathcal{C}}(D(u), D(v))$. We denote this element $\text{comp}(D)$. By convention, when the path-quiver P is trivial, $\text{comp}(D)$ is the identity map $\text{id}_{D(u)}$.

► **Definition 7** (Pullback). Let Q, Q' be two quivers and let D be a diagram over Q . Let $m: Q' \rightarrow Q$ be a morphism of quivers. We define the pullback of D by m , denoted $m^*(D)$, as the diagram $D \circ \Phi_m$ over Q' .

► **Definition 8** (Commutative diagram). For any category \mathcal{C} and any quiver Q , a diagram D over Q is commutative if $\text{comp}(p_1^*(D))$ and $\text{comp}(p_2^*(D))$ coincide for any two paths p_1 and p_2 in Q with same extremities, that is:

$$\forall (p_1, p_2) \in \mathcal{BP}_Q, \quad \text{comp}(p_1^*(D)) = \text{comp}(p_2^*(D)). \quad (1)$$

The next lemma allows to reduce the number of distinct diagrams involved in a formula.

► **Lemma 9.** Let \mathcal{C} be some category. Consider two diagrams D_1 and D_2 in \mathcal{C} over Q_1 and Q_2 , respectively. If the pullback of D_1 by m_1 coincides with the pullback of D_2 by m_2 , i.e.,

$$m_1^*(D_1) = m_2^*(D_2)$$

then, there exists a unique diagram D' over Q' such that D_1 (resp. D_2) is the pullback of D' by m'_1 (resp. m'_2), i.e.,

$$D_1 = m_1'^*(D') \quad \text{and} \quad D_2 = m_2'^*(D')$$

Proof. Immediate. ◀

► **Definition 10** (Pushout). Consider four quivers Q, Q_1, Q_2, Q' and four maps $m_1: Q \rightarrow Q_1$, $m_2: Q \rightarrow Q_2$, $m'_1: Q_1 \rightarrow Q'$ and $m'_2: Q_2 \rightarrow Q'$:

$$\begin{array}{ccc} Q & \xrightarrow{m_1} & Q_1 \\ \downarrow m_2 & & \downarrow m'_1 \\ Q_2 & \xrightarrow{m_2} & Q' \end{array}$$

This data is a pushout configuration if:

- $m'_1 \circ m_1 = m'_2 \circ m_2$,
- $Q' = \text{Im}(m'_1) \cup \text{Im}(m'_2)$, (i.e., $V_{Q'} = \text{Im}(m'_{1,V}) \cup \text{Im}(m'_{2,V})$ and $A_{Q'} = \text{Im}(m'_{1,A}) \cup \text{Im}(m'_{2,A})$),
- $\text{Im}(m'_1 \circ m_1) = \text{Im}(m'_1) \cap \text{Im}(m'_2)$.

Consider a pushout configuration as in Definition 10. The triple (Q', m'_1, m'_2) is called a pushout of (Q, Q_1, Q_2, m_1, m_2) . Such a pushout always exists, and any two such pushouts are isomorphic. Moreover, if m_1 and m_2 are embeddings, then so are m'_1 and m'_2 .

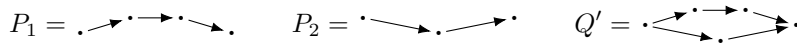
► **Lemma 11.** Let \mathcal{C} be some category. Consider two diagrams D_1 and D_2 in \mathcal{C} over Q_1 and Q_2 , respectively. If the pullback of D_1 by m_1 coincides with the pullback of D_2 by m_2 , i.e.,

$$m_1^*(D_1) = m_2^*(D_2)$$

then there exists a unique diagram D' over Q' such that D_1 (resp. D_2) is the pullback of D' by m'_1 (resp. m'_2), i.e.,

$$D_1 = m_1'^*(D') \quad \text{and} \quad D_2 = m_2'^*(D')$$

Proof. Immediate. ◀



■ **Figure 1** A pushout Q' of two path-quivers P_1 and P_2 with respect to st_{P_1} and st_{P_2} .

2.3 Category relations, path relations and quotient categories

We first name relations on the morphisms of a category that are compatible with composition:

► **Definition 12** (Category relation). *A category relation r on \mathcal{C} is by definition the data of an equivalence relation $r_{A,B} \subseteq \text{Hom}(A,B)^2$ for any pair of objects A and B such that, for any objects A, B, C and any morphisms $f, g \in \text{Hom}(A, B)$ and $f', g' \in \text{Hom}(B, C)$,*

$$f \sim g \text{ and } f' \sim g' \implies f' \circ f \sim g' \circ g,$$

where we write $h \sim h'$ if h and h' are in relation, i.e., $(h, h') \in r_{A,B}$. Such a relation is said complete if $r_{A,B} = \text{Hom}(A, B)^2$ for any pair of objects A and B .

► **Lemma 13** (Quotient category). *Given such a category relation, we define the quotient category \mathcal{C}/r given by $\text{Ob}_{\mathcal{C}/r} := \text{Ob}_{\mathcal{C}}$ and $\text{Hom}_{\mathcal{C}/r}(A, B) = \text{Hom}_{\mathcal{C}}(A, B)/r_{A,B}$ is indeed a category for the induced composition.*

Proof. Immediate. ◀

We now name the relations between the paths of a general quiver induced by the corresponding identities of morphism composites, in the associated free category.

► **Definition 14.** *A relation between paths with same extremities in \mathcal{Q} is by definition a subset of $\mathcal{BP}_{\mathcal{Q}}$. If $r \subseteq \mathcal{BP}_{\mathcal{Q}}$ is such a relation then, for $(p, q) \in \mathcal{BP}_{\mathcal{Q}}$, we write $p \sim q$ if $(p, q) \in r$. The complete path relation on \mathcal{Q} , i.e., $\mathcal{BP}_{\mathcal{Q}}$, is denoted $\text{tot}_{\mathcal{Q}}$. Note that $\mathcal{BP}_{\mathcal{Q}} = \bigsqcup_{A, B \in \text{Ob}(\mathcal{Q})} \text{Hom}_{\langle \mathcal{Q} \rangle}(A, B)^2$. Such a relation r is called a path relation if it is a category relation on $\langle \mathcal{Q} \rangle$.*

For instance, the equality of compositions in a small category \mathcal{C} induces a path relation on the underlying quiver \mathcal{Q} . If r_1, \dots, r_l are some relations between paths with same extremities, we denote by (r_1, \dots, r_l) the smallest path relation containing r_1, \dots, r_l .

Let \mathcal{Q}' be another general quiver and let $m: \mathcal{Q} \rightarrow \mathcal{Q}'$ be a morphism. If $r \subseteq \mathcal{BP}_{\mathcal{Q}}$, we denote by $m_*(r)$ the relation induced by the image by m of r in $\mathcal{BP}_{\mathcal{Q}'}$.

2.4 Many-sorted logic, categorical interpretation

We first recall a few basic definitions mostly pertaining to many-sorted logic, applied to the signatures introduced by Definition 1, and we set the corresponding notations.

Let us first fix a countable set X , so that for each quiver Q in \mathring{S} (resp. in S), elements of the set $X_Q := X \times \{Q\}$ are the *variables of sort Q* . A *term* of sort Q either is a variable of sort Q or has the form $\text{restr}_m: Q' \rightarrow Q(t)$, with t a term of sort Q . When possible, we leave the sorts implicit and simplify the notation of symbol $\text{restr}_m: Q' \rightarrow Q(t)$ into restr_m .

We denote the equality symbols by \approx . An *atom* is thus of the form $s \approx t$ with s and t two terms of the same sort, or of the form $\text{commute}_Q(t)$ with t a term of sort Q . We consider first-order many-sorted formulas and write the sort of quantifiers as a subscript, i.e., $\exists_Q x_Q, \phi$ and $\forall_Q x_Q, \phi$ where $Q \in \mathring{S}, x_Q \in X_Q$ and ϕ is a formula. In what follows, we however drop sort subscripts when they are clear from the context.

We write $\exists! y, P(y)$ for formula $(\exists y, P(y)) \wedge (\forall y_1, y_2, P(y_1) \wedge P(y_2) \rightarrow y_1 \approx y_2)$. A formula with free variables x_1, \dots, x_n of respective sorts Q_1, \dots, Q_n is said to be of *arity* $Q_1 \times \dots \times Q_n$.

► **Definition 15** (Models). *Let $Y \subseteq X \times \mathring{S}$ be a set of variables. An interpretation (also called a model) \mathcal{M} of \mathring{S} over Y is a map such that*

- each sort Q is mapped to a domain set, denoted \mathcal{M}_Q ,
- each variable $x \in Y$ of sort Q is mapped to an element $x^{\mathcal{M}} \in \mathcal{M}_Q$,
- each function $\text{restr}_m: Q' \rightarrow Q$ is mapped to function $\text{restr}_m^{\mathcal{M}}: \mathcal{M}_{Q'} \rightarrow \mathcal{M}_Q$,
- each predicate commute_Q is mapped to a subset $\text{commute}_Q^{\mathcal{M}}$ of \mathcal{M}_Q .

An interpretation is extended in the usual way to all formulas whose free variables are in Y . An interpretation \mathcal{M} of $\dot{\Sigma}$ is an interpretation over the empty set.

We define interpretations (models) of Σ in a similar fashion.

The evaluation of a formula ϕ under an interpretation \mathcal{M} is either true or false depending on the truthiness of the interpretation of the formula. If ϕ has no free variable, we write $\mathcal{M} \models \phi$ if $\phi^{\mathcal{M}}$ is true.

A theory is a set of formulas for a certain signature. Remember that signature Σ only differs from $\dot{\Sigma}$ by restricting the allowed sorts (to the acyclic quivers). If \dot{T} is a theory for $\dot{\Sigma}$, we thus define the restriction of \dot{T} to Σ , denoted $\dot{T}|_{\Sigma}$, as the subset of formulas of \dot{T} which are well-formed with respect to Σ . A model of a theory is a model such that the interpretation of every formula of the theory is true.

The prototypical models of the signatures introduced in Definition 1 are actually diagrams over a certain category.

► **Definition 16** (Categorical interpretation). *To each small category \mathcal{C} , we associate an interpretation of Σ , resp. $\dot{\Sigma}$, that we also denote by \mathcal{C} , as follows.*

- To each sort Q we associate the set \mathcal{C}_Q of diagrams in \mathcal{C} over Q .
- restr_m is interpreted as the function mapping a diagram D to the diagram $m^*(D)$.
- $\text{commute}_Q^{\mathcal{C}}$ is the set of commutative diagrams in \mathcal{C} over Q .

We call such an interpretation a categorical interpretation of Σ , resp. $\dot{\Sigma}$.

3 A theory for diagrams over small categories

This section introduces a theory whose models of can be seen as categorical interpretations.

3.1 Axioms

We now introduce the different axioms of the theory. A formula F with free variables x_1, \dots, x_n is written $F(x_1, \dots, x_n)$ so as to clarify the sorts of each variable in the arity of F .

Existence and uniqueness of the empty diagram

$$\text{EmptyEU:} \quad \exists!_{\emptyset} x, \quad x \approx x$$

Compatibility of restrictions

For any quivers Q, Q', Q'' and morphisms $m: Q \rightarrow Q'$ and $m': Q' \rightarrow Q''$, we define:

$$\text{RestrComp}_{m, m'}: \quad \forall_{Q''} x'', \quad \text{restr}_m(\text{restr}_{m'}(x'')) \approx \text{restr}_{m' \circ m}(x'').$$

Pushout

For any pushout configuration as in Definition 10, and using the same notations as this definition, we define the following formulas of arity $Q_1 \times Q_2 \times Q'$:

$$\text{Pushout}_{m'_1, m'_2}(x_1, x_2, x'): \quad \text{restr}_{m'_1}(x') \approx x_1 \wedge \text{restr}_{m'_2}(x') \approx x_2.$$

$$\text{PushoutE}_{m_1, m_2, m'_1, m'_2} : \quad \forall x_1, x_2, \text{restr}_{m_1}(x_1) \approx \text{restr}_{m_2}(x_2) \\ \rightarrow \exists! x', \text{Pushout}_{m'_1, m'_2}(x_1, x_2, x').$$

Composition

The following formula, of arity $\cdot \longrightarrow \cdot \times \cdot \longrightarrow \cdot \times \cdot \longrightarrow \cdot$, describes composite of arrows:

$$\text{Comp}(x, y, z) : \quad \exists \cdot \xrightarrow{\cdot} \cdot w, \text{restr}_{\cdot \xrightarrow{\cdot} \cdot}(w) \approx x \wedge \text{restr}_{\cdot \xrightarrow{\cdot} \cdot}(w) \approx y \\ \wedge \text{restr}_{\cdot \xrightarrow{\cdot} \cdot}(w) \approx z \wedge \text{commute}(w)$$

while the following one ensures the existence of compositions:

$$\text{CompE} : \quad \forall x, y, \exists z, \text{Comp}(x, y, z).$$

Equality of nontrivial paths

For any two nontrivial path-quivers P_1 and P_2 , and Q' , m'_1 , m'_2 such that the following diagram forms a pushout configuration (as for instance on Figure 1)

$$\begin{array}{ccc} \cdot \cdot & \xrightarrow{\text{st}_{P_1}} & P_1 \\ \downarrow \text{st}_{P_2} & & \downarrow m'_1 \\ P_2 & \xrightarrow{m'_2} & Q' \end{array}$$

we define the following formula of arity $P_1 \times P_2$:

$$\text{EqPath}_{P_1, P_2}(x_1, x_2) : \quad \text{restr}_{\cdot \cdot}(x_1) \approx \text{restr}_{\cdot \cdot}(x_2) \\ \wedge (\forall x, \text{Pushout}_{m'_1, m'_2}(x_1, x_2, x) \rightarrow \text{commute}(x)).$$

Identity

The following formula, of arity $\cdot \times \cdot \longrightarrow \cdot$, defines the identity map:

$$\text{Id}(x, y) : \quad \text{restr}_{\cdot \cdot}(y) \approx x \wedge \forall z, w, \\ (\text{Comp}(y, z, w) \rightarrow \text{EqPath}(z, w)) \wedge (\text{Comp}(z, y, w) \rightarrow \text{EqPath}(z, w))$$

and the following formula ensures the existence of identity maps.

$$\text{IdE} : \quad \forall x, \exists y, \text{Id}(x, y).$$

Equality for general paths

For any nontrivial path-quiver P , we define the following formulas:

$$\text{EqPath}_{\cdot \cdot}(x, y) : \quad x \approx y, \\ \text{EqPath}_{\cdot, P}(x, y) : \quad \exists z, \text{Id}(x, z) \wedge \text{EqPath}_{\cdot \cdot, P}(z, y), \\ \text{EqPath}_{P, \cdot}(x, y) : \quad \text{EqPath}_{\cdot, P}(y, x).$$

Hence, we can see EqPath as a relation with arity any pair of path-quivers. We ensure this relation to be an equivalence relation by defining for any path-quivers P_1, P_2 and P_3 the three formulas EqPathRef_{P_1} , $\text{EqPathSym}_{P_1, P_2}$ and $\text{EqPathTrans}_{P_1, P_2, P_3}$ stating that the relation EqPath is respectively reflexive, symmetric and transitive.

We also make sure to enforce the properties of a category relation. For this purpose, for any four path-quivers $P_1, P'_1, P_2,$ and $P'_2,$ of respective length k_1, k'_1, k_2 and $k'_2,$ we define the following formula where bound variables x_1, x_2, x'_1, x'_2 respectively have sort P_1, P_2, P'_1 and P'_2 and the sort of $x''_i, i \in \{1, 2\},$ is $\text{PQ}_{k_i+k'_i}.$

$$\begin{aligned} \text{EqPathConcat}_{P_1, P_2, P'_1, P'_2} : \quad & \forall x_1, x_2, x'_1, x'_2, \\ & \text{EqPath}(x_1, x_2) \wedge \text{EqPath}(x'_1, x'_2) \wedge \text{restr}_{\sim}(x_1) \approx \text{restr}_{\sim}(x'_1) \\ \rightarrow \quad & \forall x''_1, x''_2, \quad \text{Pushout}_{\sim, \sim}(x_1, x'_1, x''_1) \wedge \text{Pushout}_{\sim, \sim}(x_2, x'_2, x''_2) \\ & \rightarrow \quad \text{EqPath}(x''_1, x''_2). \end{aligned}$$

Commutativity

We relate commutativity and equality by the following formula:

$$\text{ComEq} : \quad \forall \cdot \curvearrowright \cdot x, \quad \text{commute}(x) \rightarrow \text{restr}_{\curvearrowright}(\cdot)(x) \approx \text{restr}_{\curvearrowright}(\cdot)(x).$$

For any quiver $Q,$ the following formula provides an analogue of the notion of commutativity of diagrams given in Definition 8, Equation (1):

$$\text{PathCom}_Q : \quad \forall_Q x, \quad \bigwedge_{(p_1, p_2) \in \mathcal{BP}_Q} \text{EqPath}(\text{restr}_{p_1}(x), \text{restr}_{p_2}(x)) \leftrightarrow \text{commute}(x).$$

where we recall that \mathcal{BP}_Q denotes the set of pair of paths of Q having the same extremities.

► **Definition 17.** *Theory $\mathring{\mathcal{T}}_{\text{cat}},$ over signature $\mathring{\Sigma},$ consists of the following formulas:*

- EmptyEU, CompE, IdE, ComEq,
- $\text{RestrComp}_{m, m'}$ for any pair of maps m and m' as in Section 3.1,
- $\text{PushoutE}_{m_1, m_2, m'_1, m'_2}$ for a pushout configuration as in Definition 10,
- $\text{EqPathRefl}_{P_1}, \text{EqPathSym}_{P_1, P_2}, \text{EqPathTrans}_{P_1, P_2, P_3}, \text{EqPathConcat}_{P_1, P_2, P_3, P_4}$ for any quadruple of path-quivers P_1, P_2, P_3 and $P_4,$
- PathCom_Q for any quiver $Q.$

Theory \mathcal{T}_{cat} is defined as the restriction of $\mathring{\mathcal{T}}_{\text{cat}}$ to $\Sigma.$

3.2 Models

Models of $\mathcal{T}_{\text{cat}},$ resp. $\mathring{\mathcal{T}}_{\text{cat}}$ are in fact exactly what we have called categorical interpretations.

► **Theorem 18.** *Every categorical interpretation of $\Sigma,$ resp. $\mathring{\Sigma},$ is a model of $\mathcal{T}_{\text{cat}},$ resp. $\mathring{\mathcal{T}}_{\text{cat}}.$ Moreover, any model \mathcal{M} of $\mathcal{T}_{\text{cat}},$ resp. $\mathring{\mathcal{T}}_{\text{cat}},$ has an isomorphic categorical interpretation.*

Proof. We only prove the theorem for $\mathcal{T}_{\text{cat}},$ as the proof for $\mathring{\mathcal{T}}_{\text{cat}}$ is similar.

If \mathcal{C} is a small category, a routine check shows that the associated model \mathcal{C} of Σ verifies the theory $\mathcal{T}_{\text{cat}}.$ For instance,

- the formulas PushoutE follows from the remark in Definition 10,
- IdE and CompE come from the existence of the identity map and the existence of the composition respectively,
- for two path-quivers P_1, P_2 and two diagrams D_1 and D_2 over them, $\text{EqPath}_{P_1, P_2}(D_1, D_2)$ is the relation $\text{comp}(D_1) = \text{comp}(D_2),$ which is a path relation by Lemma 13,
- ComEq and PathCom $_Q$ follow from Equation (1).

Let us prove the other direction. Let Q be an acyclic quiver. By an abuse of the notations, if $v \in V_Q$, resp. $a \in A_Q$, we also denote by v the corresponding embedding $v: \cdot \hookrightarrow Q$, resp. $a: \cdot \rightarrow \cdot \hookrightarrow Q$. Here is a crucial lemma for the proof of the theorem.

► **Lemma 19** (General pushout). *Let \mathcal{M} be a model of \mathcal{T}_{cat} . Let Q be an acyclic quiver, $\beta_V: V_Q \rightarrow \mathcal{M}$, and $\beta_A: A_Q \rightarrow \mathcal{M}_{\cdot, \cdot}$ be two maps. Then the following statements are equivalent.*

1. *There exists an element β of \mathcal{M}_Q such that $\text{restr}_v^{\mathcal{M}}(\beta) = \beta_V(v)$ and $\text{restr}_a^{\mathcal{M}}(\beta) = \beta_A(a)$ for any $v \in V_Q$ and any $a \in A_Q$,*
 2. *For any $a \in A_Q$, $\text{restr}_{\cdot, \cdot}^{\mathcal{M}}(\beta_A(a)) = \beta_V(s_Q(a))$ and $\text{restr}_{\cdot, \cdot}^{\mathcal{M}}(\beta_A(a)) = \beta_V(t_Q(a))$.*
- Moreover, when both statements hold, the element β is unique.

Proof. The fact that the first point induces the second one follows directly from `RestrComp`. Hence we focus on the other direction and on the uniqueness.

We proceed by induction on the structure of Q . If $Q = \emptyset$ is empty, the second point holds trivially. The first point and the uniqueness follows from `EmptyEU`.

Assume first that the lemma holds for some quiver Q_1 with no arrow. Let Q be the quiver Q_1 with an extra vertex v_0 . Let β_V and β_A be two maps as in the statement of the lemma, and $\beta_{1,V}$ the restriction of β_V to $Q_{1,V}$. We have the following pushout configuration:

$$\begin{array}{ccc} \emptyset & \xrightarrow{m_1} & Q_1 \\ \downarrow m_2 & & \downarrow m'_1 \\ \cdot & \xrightarrow{m'_2} & Q \\ & m'_2 = v_0 & \end{array}$$

Point 2 holds trivially. Let us prove the existence and the uniqueness of β satisfying the property of Point 1. By induction, we get a unique $\beta_1 \in \mathcal{M}_{Q_1}$ compatible with $\beta_{1,V}$ and β_A . Set $\beta_2 := \beta_V(v_0)$. By `EmptyEU`, $\text{restr}_{m_1}^{\mathcal{M}}(\beta_1) = \text{restr}_{m_2}^{\mathcal{M}}(\beta_2)$. Hence we can apply `PushoutE` $_{m_1, m_2, m'_1, m'_2}$ to get a unique element $\beta \in \mathcal{M}_Q$ such that `Pushout` $^{\mathcal{M}}(\beta_1, \beta_2, \beta)$. Using `RestrComp` and the induction hypothesis, we can check that for $\beta' \in \mathcal{M}_Q$, there is an equivalence between `Pushout` $^{\mathcal{M}}(\beta_1, \beta_2, \beta')$ and $\text{restr}_v^{\mathcal{M}}(\beta') = \beta_V(v)$ for any $v \in V_Q$. This concludes the induction.

Assume more generally that the lemma holds for some acyclic quiver Q_1 . Let Q be an acyclic quiver obtained from Q_1 by adding one arrow a_0 . Let β_V and β_A be two maps as before and $\beta_{1,A}$ the restriction of β_A to A_{Q_1} . Let $m_1: \cdot \hookrightarrow Q_1$ mapping the first point to $s_Q(a_0)$ and the second point to $t_Q(a_0)$. Let $m_2 := \text{st}_{\cdot, \cdot} = \cdot \rightarrow \cdot$. Once again, we get a pushout configuration:

$$\begin{array}{ccc} \cdot \cdot & \xrightarrow{m_1} & Q_1 \\ \downarrow m_2 & & \downarrow m'_1 \\ \cdot \rightarrow \cdot & \xrightarrow{m'_2} & Q \\ & m'_2 = a_0 & \end{array}$$

Assume that Point 2 holds. By induction, we get a unique element $\beta_1 \in \mathcal{M}_{Q_1}$ compatible with β_V and $\beta_{1,A}$. Set $\beta_2 := \beta_A(a_0)$. We have already proven the lemma for the quiver $\cdot \cdot$. Hence we deduce that

$$\text{restr}_{m_1}^{\mathcal{M}}(\beta_1) = \text{restr}_{m_2}^{\mathcal{M}}(\beta_2).$$

We can apply `PushoutE` as before to get Point 1 as well as the uniqueness part. This concludes the proof of the lemma. ◀

We now continue the proof of Theorem 18. Let \mathcal{M} be a model of \mathcal{T}_{cat} . We define the general quiver \mathcal{Q} associated to \mathcal{M} by

$$V_{\mathcal{Q}} := \mathcal{M}_{\cdot}, \quad A_{\mathcal{Q}} := \mathcal{M}_{\cdot, \dots}, \quad s_{\mathcal{Q}} := \text{restr}_{\cdot, \dots}^{\mathcal{M}} \quad \text{and} \quad t_{\mathcal{Q}} := \text{restr}_{\cdot, \dots}^{\mathcal{M}}.$$

Set $\tilde{\mathcal{C}} := \langle \mathcal{Q} \rangle$. Thanks to Lemma 19, to each acyclic quiver Q and to each element $\beta \in \mathcal{M}_Q$, we can associate a unique diagram $\tilde{\Psi}(\beta)$ in $\tilde{\mathcal{C}}$ verifying:

- for any $v \in V_Q$, $\tilde{\Psi}(\beta)(v) = \text{restr}_v^{\mathcal{M}}(\beta)$.
- for any $a \in A_Q$, $\tilde{\Psi}(\beta)(a)$ is the path of length one with arrow $\text{restr}_a^{\mathcal{M}}(\beta)$.

The image of $\tilde{\Psi}$ is exactly the set of diagrams whose morphisms are paths of lengths one.

There is another important map. Let A and B be two objects of $\tilde{\mathcal{C}}$, and let $p \in \text{Hom}(A, B)$. Recall that p is just a path from A to B in Q . Let k be the length of p . By Lemma 19, there exists a unique element $\Theta(p) \in \mathcal{M}_{\text{PQ}_k}$ such that

- for each $v \in V_{\text{PQ}_k}$, $\text{restr}_v^{\mathcal{M}}(\Theta(p)) = p(v)$,
- for each $a \in A_{\text{PQ}_k}$, $\text{restr}_a^{\mathcal{M}}(\Theta(p)) = p(a)$.

Relation $\text{EqPath}^{\mathcal{M}}$ thus induces a relation r on morphisms of $\tilde{\mathcal{C}}$. Moreover, EqPathRef , EqPathSym , EqPathTrans and EqPathConcat , together with Lemma 19, make r a category relation. We can hence define the category $\mathcal{C} := \tilde{\mathcal{C}}/r$. Now $\tilde{\Psi}$ induces a map $\Psi: \mathcal{M}_Q \rightarrow \mathcal{C}_Q$ for any quiver Q , and we claim that Ψ induces a model isomorphism between \mathcal{M} and \mathcal{C} .

Let Q be any acyclic quiver. We first prove that Ψ is injective. Let $\beta, \gamma \in \mathcal{M}_Q$ such that $\Psi(\beta) = \Psi(\gamma)$. For any vertex $v \in V_Q$, $\tilde{\Psi}(\beta)(v) = \tilde{\Psi}(\gamma)(v)$, i.e., $\text{restr}_v^{\mathcal{M}}(\beta) = \text{restr}_v^{\mathcal{M}}(\gamma)$. Let a be an arrow of Q . Then we have the relation $\tilde{\Psi}(\beta)(a) \sim \tilde{\Psi}(\gamma)(a)$. By definition of $\text{EqPath}_{\dots, \dots, \dots}$, we validate the premise of ComEq , and thus the equality $\tilde{\Psi}(\beta)(a) = \tilde{\Psi}(\gamma)(a)$, i.e., $\text{restr}_a^{\mathcal{M}}(\beta) = \text{restr}_a^{\mathcal{M}}(\gamma)$. By the uniqueness part of Lemma 19, we get $\beta = \gamma$.

We now consider the surjectivity of Ψ . It suffices to prove that any morphism $p \in \text{Hom}_{\tilde{\mathcal{C}}}(A, B)$ is in relation via r to a path of length one. Indeed, in such a case, for any diagram \tilde{D} in $\tilde{\mathcal{C}}$ over Q , one can find another diagram \tilde{D}' over Q whose morphisms are path of size one and such that any morphism of \tilde{D} is in relation with the corresponding morphism of \tilde{D}' . Hence the induced diagrams in \mathcal{C} are equal. Moreover, \tilde{D}' is in the image of $\tilde{\Psi}$, and we would get the surjectivity.

Let P be a path-quiver and let $\beta \in \mathcal{M}_P$. We have to find an element $\gamma \in \mathcal{M}_{\dots}$ such that $\text{EqPath}_{P, \dots}^{\mathcal{M}}(\beta, \gamma)$. If P has length one, this is trivial. If P has length zero, then by IdE , there exists γ such that $\text{Id}^{\mathcal{M}}(\beta, \gamma)$. Moreover, by the definition $\text{EqPath}_{\dots, \dots}$ and using the reflexivity of EqPath , we get $\text{EqPath}_{\dots, \dots}^{\mathcal{M}}(\beta, \gamma)$. If P has length two, then by CompE , we can find an element γ such that

$$\text{Comp}(\text{restr}_{\dots, \dots}^{\mathcal{M}}(\beta), \text{restr}_{\dots, \dots}^{\mathcal{M}}(\beta), \gamma).$$

The commutativity of the triangle induces $\text{EqPath}_{\dots, \dots, \dots}^{\mathcal{M}}(\beta, \gamma)$.

For P with length $k > 2$, we work by induction. Using EqPathTrans , it suffices to find γ over PQ_{k-1} such that $\text{EqPath}_{P, \text{PQ}_{k-1}}^{\mathcal{M}}(\beta, \gamma)$. To do so, we see P as the pushout of PQ_2 and PQ_{k-2} along $m'_1 = \sim: \text{PQ}_2 \rightarrow P$ and $m'_2 = \sim: \text{PQ}_{k-2} \rightarrow P$. By the case $k = 2$, we can find $\gamma_1 \in \mathcal{M}_{\dots}$ such that $\text{EqPath}(\text{restr}_{m'_1}^{\mathcal{M}}(\beta), \gamma_1)$. We set $\gamma_2 = \text{restr}_{m'_2}^{\mathcal{M}}(\beta)$. In particular we have $\text{EqPath}(\gamma_2, \gamma_2)$. By EqPathConcat , we get $\text{EqPath}(\beta, \gamma)$ where $\gamma \in \mathcal{M}_{\text{PQ}_{k-1}}$ is such that $\text{Pushout}_{m'_1, m'_2}(\gamma_1, \gamma_2, \gamma)$, and the result follows.

We have shown that Ψ induces a bijection between the corresponding domains. In order to conclude the proof, it remains to prove that Ψ commutes with restr and with commute . The commutativity with restr follows from the definition of Ψ and the formulas RestrComp .

The compatibility with the predicate commute can be reduced to the compatibility of EqPath via the formula PathCom. Let P_1 and P_2 be two path-quivers and $\beta_1 \in \mathcal{M}_{P_1}$ and $\beta_2 \in \mathcal{M}_{P_2}$. These elements correspond to paths p_1 and p_2 in \mathcal{Q} . Using the definitions of the different elements, we get the following chain of equivalences:

$$\begin{aligned} \text{EqPath}^c(\Psi(\beta_1), \Psi(\beta_2)) &\Leftrightarrow \text{comp}_{\mathcal{C}}(\Psi(\beta_1)) = \text{comp}_{\mathcal{C}}(\Psi(\beta_2)) \\ &\Leftrightarrow \text{comp}_{\tilde{\mathcal{C}}}(\tilde{\Psi}(\beta_1)) \sim \text{comp}_{\tilde{\mathcal{C}}}(\tilde{\Psi}(\beta_2)) \Leftrightarrow p_1 \sim p_2 \Leftrightarrow \text{EqPath}(\beta_1, \beta_2). \end{aligned}$$

This concludes the proof of the theorem. \blacktriangleleft

4 Duality

Signatures of Definition 1 are tailored to enforce a built-in, therefore easy to prove, duality principle, which we now make precise. Recall from Definition 2 that duality is an involution on quivers, and also on acyclic quivers. We define the dual of a formula over $\dot{\Sigma}$ as follows:

- if $m: Q \rightarrow Q'$ is a morphism, then $m^*: Q^* \rightarrow Q'^*$ is defined by $m_V^* = m_V$ and $m_A^* = m_A$,
- if $x = (x, Q)$ is a variable in $X \times \dot{S}$ then $x^* := (x, Q^*)$,
- $(\text{restr}_m(x))^* := \text{restr}_{m^*}(x^*)$ and $(\text{commute}_Q(x))^* := \text{commute}_{Q^*}(x^*)$,
- $(x \approx y)^* := x^* \approx y^*$,
- $(\forall_Q x, \phi)^* := \forall_{Q^*} x^*, \phi^*$ and $(\exists_Q x, \phi)^* := \exists_{Q^*} x^*, \phi^*$,
- $(\phi \wedge \psi)^* := \phi^* \wedge \psi^*$, etc.

For Y a set of variables, Y^* denotes $\{x^* \mid x \in Y\}$. For any theory \mathcal{T} , \mathcal{T}^* denotes $\{\phi^* \mid \phi \in \mathcal{T}\}$.

For \mathcal{M} an interpretation of $\dot{\Sigma}$ over a set of variables Y , we define its dual model \mathcal{M}^* as:

- $\mathcal{M}_Q^* := \mathcal{M}_{Q^*}$,
- for $x \in Y^*$, $x^{\mathcal{M}^*} := (x^*)^{\mathcal{M}}$.
- $\text{restr}_m^{\mathcal{M}^*} := \text{restr}_{m^*}^{\mathcal{M}}$ and $\text{commute}_Q^{\mathcal{M}^*} := \text{commute}_{Q^*}^{\mathcal{M}}$,

The duality involution also restricts to formulas, theories and models over Σ .

► **Example 20.** If \mathcal{C} is a small category, then the dual interpretation \mathcal{C}^* is isomorphic to the model of the dual category, both with respect to $\dot{\Sigma}$ and to Σ .

► **Theorem 21 (Duality theorem).** *Let ϕ be a formula with free variables included in $Y \subseteq X \times S$, resp. in $Y \subseteq X \times \dot{S}$, and let \mathcal{M} be a model of Σ , resp. of $\dot{\Sigma}$. Then*

$$\mathcal{M} \models \phi \iff \mathcal{M}^* \models \phi^*.$$

Proof. More generally, a formula is provable if and only if its dual is provable. \blacktriangleleft

► **Remark 22.** The duality principle has some useful direct consequences:

- If ϕ is a valid, resp. satisfiable, resp. unsatisfiable, formula, so is ϕ^* .
- Let \mathcal{T} be theory such that any model of \mathcal{T} verifies \mathcal{T}^* . If ϕ is a valid, resp. satisfiable, resp. unsatisfiable, formula among models of \mathcal{T} , so is ϕ^* .
- We have the following reciprocal. Let \mathcal{T} be a theory such that any model \mathcal{M} of \mathcal{T} verifies that $\mathcal{M}^* \models \mathcal{T}$, then every model of \mathcal{T} verifies \mathcal{T}^* .

The following fact follows directly from this last point, Theorem 18 and Example 20.

► **Proposition 23.** *Models of \mathcal{T}_{cat} verify $\mathcal{T}_{\text{cat}}^*$, and models of $\dot{\mathcal{T}}_{\text{cat}}$ verify $\dot{\mathcal{T}}_{\text{cat}}^*$.*

5 A theory for diagrams over abelian categories

We now introduce a theory whose models are diagrams in small abelian categories. We rely on the set of axioms given by Freyd in [5]. This reference is particularly well-suited for our purpose. Indeed, the author does not impose the homomorphisms between any two objects of an abelian category to form a group, but this fact rather follows from the axioms.

Let us first introduce common notions of category theory in our logic. Here is a formula of arity $\cdot \longrightarrow \cdot$ which corresponds to monicity of a map in a category.

$$\begin{aligned} \text{Mono}(x) : \quad & \forall \begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array} y, \quad \text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(y) \approx x \\ & \wedge \text{commute}(\text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(y)) \wedge \text{commute}(\text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(y)) \\ & \rightarrow \text{commute}(\text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(y)). \end{aligned}$$

The dual formula is called Epi.

Let Q be a quiver. We define the *cone of Q* as the quiver

$$\text{cone}(Q) := (V_Q \sqcup \{v_0\}, A_Q \sqcup \{a_v \mid v \in V_Q\}, s_{\text{cone}(Q)}, t_{\text{cone}(Q)}),$$

where $s_{\text{cone}(Q)}$ and $t_{\text{cone}(Q)}$ are extensions of s_Q and t_Q by $s_Q(a_v) = v_0$ and $t_Q(a_v) = v$. We define by $i_Q: Q \hookrightarrow \text{cone}(Q)$ the corresponding embedding. If $m: Q \rightarrow Q'$ is a morphism of quivers, we get a canonical morphism $\text{cone}(m): \text{cone}(Q) \rightarrow \text{cone}(Q')$.

Abusing the notations, if a is an arrow of Q , we also denote by $a: \cdot \longrightarrow \cdot \hookrightarrow Q$ the corresponding morphism. We then introduce the usual notion of cones of diagrams by the following formula of arity $Q \times \text{cone}(Q)$.

$$\text{Cone}_Q(x, y) : \quad \text{restr}_{i_Q}(y) \approx x \quad \wedge \quad \bigwedge_{a \in A_Q} \text{commute}(\text{restr}_{\text{cone}(a)}(y)).$$

Here is the notion of limit.

$$\begin{aligned} \text{Limit}_Q(x, y) : \quad & \text{Cone}_Q(x, y) \quad \wedge \\ & \forall z, \quad \text{Cone}_Q(x, z) \rightarrow \exists! w, \quad \text{Cone}(y, w) \wedge \text{restr}_{\text{cone}(i_Q)}(w) \approx z. \end{aligned}$$

We also introduce the dual notion $\text{Colimit}_Q := \text{Limit}_Q^*$.

The introduction of monos, epis, limits and colimits allows to state the axioms of abelian category given by Freyd [5]. First, we define zero objects and kernels as follows.

$$\begin{aligned} \text{Zero}(x) : \quad & \forall y, \quad \text{Limit}_{\emptyset}(y, x) \wedge \text{Colimit}_{\emptyset}(y, x), \\ \text{Ker}(x, y) : \quad & \exists \begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array} z, \quad \text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(z) \approx x \quad \wedge \quad \text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(z) \approx y \\ & \wedge \quad \text{Zero}(\text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(z)) \quad \wedge \quad \text{Limit}(\text{restr}_{\begin{array}{c} \xrightarrow{\quad} \\ \xrightarrow{\quad} \\ \xrightarrow{\quad} \end{array}}(z), z). \end{aligned}$$

We also define $\text{Coker} := \text{Ker}^*$.

We define the category \mathcal{T}_{ab} as the extension of \mathcal{T}_{cat} by the following formulas.

$$\begin{aligned} \text{ZeroE} : \quad & \exists x, \quad \text{Zero}(x), \\ \text{ProductE} : \quad & \forall x, \exists y, \quad \text{Limit}_{\cdot, \cdot}(x, y), \\ \text{CoproductE} : \quad & \text{ProductE}^*, \\ \text{KerE} : \quad & \forall x, \exists y, \quad \text{Ker}(x, y), \\ \text{CokerE} : \quad & \text{KerE}^*, \\ \text{MonoNormal} : \quad & \forall x, \quad \text{Mono}(x) \rightarrow \exists y, \quad \text{Ker}(y, x), \\ \text{EpiNormal} : \quad & \text{MonoNormal}^*. \end{aligned}$$

The following theorem states that \mathcal{T}_{ab} is a theory for diagrams over abelian categories.

► **Theorem 24.** *The categorical interpretation induced by any small abelian category is a model of \mathcal{T}_{ab} . Conversely, any model of \mathcal{T}_{ab} is isomorphic to the categorical interpretation associated to some small abelian category.*

Proof. This follows from Theorem 18 and from [5, Chapter 2]. ◀

► **Proposition 25.** *The theory \mathcal{T}_{ab} implies its dual $\mathcal{T}_{\text{ab}}^*$.*

Proof. The theory \mathcal{T}_{cat} implies its dual by Proposition 23. Moreover, ZeroE clearly implies its dual. Finally, for the other axioms we added, we also added their dual. ◀

6 Decidability of the commerge problem

In this section, we use the notations of Section 2.3. Let Q be a quiver, $k \in \mathbb{N}$ and, for each $i \in [k]$, let Q_i be a quiver and $m_i: Q_i \rightarrow Q$ be a morphism. We define the following formula:

$$\text{Commerge}_{m_0, \dots, m_{k-1}}: \quad \forall_Q x, \quad \bigwedge_{i=0}^{k-1} \text{commute}(\text{restr}_{m_i}(x)) \quad \rightarrow \quad \text{commute}(x).$$

► **Definition 26.** *Notations as above, the acyclic, resp. cyclic, commerge problem for morphisms, resp. embeddings, m_0, \dots, m_{k-1} and for a theory \mathcal{T} is the problem of deciding the validity of $\text{Commerge}_{m_0, \dots, m_{k-1}}$ among models of Σ , resp. $\mathring{\Sigma}$, verifying the theory \mathcal{T} .*

We recall that a *thin category* is a category with at most one morphism between any pair of objects. Let $\text{tot}_{Q_i} = \mathcal{BP}_{Q_i}$ be the complete path relation on Q_i . Set $r_i := m_{i*}(\text{tot}_{Q_i})$ for $i \in [k]$. Recall that $(r_i)_{i \in [k]}$ is the smallest path relation containing the r_i for all $i \in [k]$.

► **Lemma 27.** *Notation as above, the formula $\text{Commerge}_{m_0, \dots, m_{k-1}}$ is valid among model of \mathcal{T}_{cat} , resp. $\mathring{\mathcal{T}}_{\text{cat}}$, if and only if $\langle Q \rangle / (r_i)_{i \in [k]}$ is a thin category.*

Proof. Set $\mathcal{C} := \langle Q \rangle / (r_i)_{i \in [k]}$. It is a model of \mathcal{T}_{cat} , resp. $\mathring{\mathcal{T}}_{\text{cat}}$. Moreover, the canonical diagram $D: \langle Q \rangle \rightarrow \mathcal{C}$ verifies the premise of $\text{Commerge}_{m_0, \dots, m_{k-1}}$. If \mathcal{C} is not thin, then there are two paths p and q in $\langle Q \rangle$ with the same extremities which are not in relation. Then $\text{comp}(p^*(D))$ is the class of p in the quotient, which is different of the class of q , that is of $\text{comp}(q^*(D))$. Hence D is not commutative.

For the other direction, by Theorem 18, it suffices to study diagrams in small categories. It is easy to check that any diagram D' over Q in a category \mathcal{C}' which verifies the condition of $\text{Commerge}_{m_0, \dots, m_{k-1}}$ factors through D , i.e., $D' = \Psi \circ D$ for some functor $\Psi: \mathcal{C} \rightarrow \mathcal{C}'$. If \mathcal{C} is thin, then for any two paths p and q with same extremities in Q ,

$$\text{comp}(p^*(D')) = \Psi(\text{comp}(p^*(D))) = \Psi(\text{comp}(q^*(D))) = \text{comp}(q^*(D')).$$

Hence $\text{Commerge}_{m_0, \dots, m_{k-1}}$ is valid. ◀

► **Theorem 28.** *The acyclic commerge problem for \mathcal{T}_{cat} is decidable for any tuple of embeddings.*

Proof. By Lemma 27, it suffices to decide if $\langle Q \rangle / (r_i)_{i \in [k]}$. Since $\langle Q \rangle$ and the $\langle Q_i \rangle$, $i \in [k]$ are finite, we can compute relation $(m_{i*}(\text{tot}_{Q_i}) \mid i \in [k])$ and check it is complete. ◀

► **Proposition 29.** *There exists a tuple of morphisms for which the cyclic commerge problem for $\mathring{\mathcal{T}}_{\text{cat}}$ is undecidable.*

Proof. We proceed by reduction to an undecidability result, due to Markov [9]. For B an arbitrary finite set and $\langle B \rangle$ the associated free monoid, let M be the finitely presentable monoid $\langle B \rangle / R$, for R a finite subset of $\langle B \rangle$. The triviality of M is undecidable.

Let B, R and M as above. Let $Q = (\{v\}, B \sqcup \{e\}, s_Q, t_Q)$ be a quiver with one vertex and loops labeled by elements of B plus one loop e . To each element $\rho \in R$ corresponds a path $p_\rho: \text{PQ}_{k_\rho} \rightarrow Q$, for some $k_\rho \in \mathbb{N}$. Let Q_ρ be a pushout of the morphisms $\dashv: \cdot \cdot \hookrightarrow \text{PQ}_{k_\rho}$ and $\cdot \dashrightarrow \cdot$. Let m_ρ be the extension to Q_ρ of p_ρ obtained by mapping the new arrow onto e . Also set $m_e: Q \rightarrow Q$ which maps the loop on e . We claim that the cyclic comerge problem for the m_ρ and m_e is undecidable. Indeed, $\langle Q \rangle / (m_e^*(\text{tot } Q), (m_\rho^*(\text{tot } Q_\rho))_{\rho \in R})$ is the category associated to the monoid $M = \langle B \rangle / R$. Hence this category is thin if and only if the monoid is trivial. Markov Theorem [9] and Lemma 27 conclude the proof. \blacktriangleleft

We strengthen the previous proposition to the case of embeddings.

► Theorem 30. *There exists a tuple of embeddings for which the cyclic comerge problem for \mathcal{T}_{cat} is undecidable.*

Proof. Let B, R, M as in the proof of Proposition 29. Let Q be the quiver $(\{v\}, B, s_Q, t_Q)$. If $k \geq 2$, we define the quiver \mathring{Q}^k as

$$\mathring{Q}^k := \left(\{v_i \mid i \in [k]\}, \{b_{i,j} \mid b \in B, 0 \leq i < j < k\} \sqcup \{e_{i,j} \mid i, j \in [k]\}, s_{\mathring{Q}^k}, t_{\mathring{Q}^k} \right) \text{ where}$$

$$s_{\mathring{Q}^k}(b_{i,j}) = v_i, \quad s_{\mathring{Q}^k}(e_{i,j}) = v_i, \quad t_{\mathring{Q}^k}(b_{i,j}) = v_j, \quad t_{\mathring{Q}^k}(e_{i,j}) = v_j.$$

We have a projection $\pi: \langle \mathring{Q}^k \rangle \rightarrow \langle Q \rangle$, which maps $b_{i,j}$ on b and $e_{i,j}$ on id_v , and a section $\iota: \langle Q \rangle \rightarrow \langle \mathring{Q}^k \rangle$ defined by mapping v onto v_0 and b onto $b_{0,k-1} \circ e_{k-1,0}$, where, as usual, we denote in a same way an arrow and the corresponding path of length one.

For A any subset of $A_{\mathring{Q}^k}$, let $m_A: \mathring{Q}^k|_A \hookrightarrow \mathring{Q}^k$ be the canonical embedding, and let $r_A := m_{A^*}(\text{tot } \mathring{Q}^k|_A)$. Set $r' := ((r_A)_{A \in \mathcal{A}})$ for $\mathcal{A} \subset 2^{A_{\mathring{Q}^k}}$ defined as the set containing

- $A_e := \{e_{i,j} \mid i, j \in [k]\}$,
- for $i < j$ and $b \in B$,

$$A_{b_{i,j}} := \left\{ \underbrace{e_{0,i}}_{\text{if } i \neq 0}, b_{i,j}, \underbrace{e_{j,k-1}}_{\text{if } j \neq k-1}, b_{0,k-1} \right\}.$$

We claim that π and ι induce an equivalence of category between $\langle Q \rangle$ and $\langle \mathring{Q}^k \rangle / r'$. From the definition of A_e , for any $i, j, l \in [k]$, we have $e_{i,j} \circ e_{j,l} \sim e_{i,l}$ and $e_{i,i} \sim \text{id}_i$. Now the definition of $A_{b_{i,j}}$, for $b \in B$ and $0 \leq i < j < k$, induces that $e_{0,i} \circ b_{i,j} \circ e_{j,k-1} \sim b_{0,k-1}$. These relations generate all r' , and they become equalities by applying the projection. Hence $\pi_*: \langle \mathring{Q}^k \rangle / r' \rightarrow \langle Q \rangle$ is well-defined. Clearly $\pi \circ \iota$ is identity. Concerning the other direction, for $b \in B$ and $0 \leq i < j < k$, we have

$$\iota \circ \pi(b_{i,j}) = b_{0,k-1} \circ e_{k-1,0} \sim e_{0,i} \circ b_{i,j} \circ e_{j,k-1} \circ e_{k-1,0} \sim e_{0,i} \circ b_{i,j} \circ e_{j,0}.$$

Hence we get a natural transformation η between the identity functor and $\iota \circ \pi_*$ by setting $\eta_i := e_{i,0} \in \text{Hom}_{\langle \mathring{Q}^k \rangle / r'}(i, \iota \circ \pi(i) = 0)$. Since $e_{i,0}$ is an isomorphism, we conclude that there is an equivalence of category between $\langle Q \rangle$ and $\langle \mathring{Q}^k \rangle / r'$.

Recall that M is the monoid $\langle B \rangle / R$. Assume that k is greater than the longest word in R . To any word $\rho = b^1 b^2 \dots b^l$ in R corresponds a subset

$$A_\rho := \{e_{0,l}, b_{0,1}^1, b_{1,2}^2, \dots, b_{l-1,l}^l\} \subseteq A_{\mathring{Q}^k}.$$

Let $\mathcal{A}' := \mathcal{A} \cup \{A_\rho \mid \rho \in R\}$. We claim that $\text{Commerge}_{(m_A)_{A \in \mathcal{A}'}}$ is undecidable. Indeed, $\mathcal{Q}^k / ((r_A)_{A \in \mathcal{A}'})$ is equivalent as a category to $Q / ((\pi_*(r_{A_\rho}))_{\rho \in R})$ which is the category of the monoid M . Once again, we conclude using Lemma 27 and Markov Theorem. \blacktriangleleft

► **Theorem 31.** *The theory \mathcal{T}_{cat} contains undecidable formulas.*

Proof. Let M , B , R and Q as in the proof of Theorem 30. Let k be the size of the longest word in R plus one. We define the quiver Q^k as

$$Q^k := \left(\{v_i \mid i \in [k]\}, \{b_{i,j} \mid b \in B, 0 \leq i < j < k\} \sqcup \{e_{i,j} \mid 0 \leq i < j < k\}, s_{Q^k}, t_{Q^k} \right)$$

where

$$s_{Q^k}(b_{i,j}) = v_i, \quad s_{Q^k}(e_{i,j}) = v_i, \quad t_{Q^k}(b_{i,j}) = v_j, \quad t_{Q^k}(e_{i,j}) = v_j.$$

Note that Q^k is acyclic. Let

$$\mathcal{A}'' := \{A_{b,i,j} \mid b \in B, 0 \leq i < j < k\} \cup \{A_\rho \mid \rho \in R\},$$

where $A_{b,i,j}$ and A_ρ are defined as in the proof of Theorem 30. For $A \in \mathcal{A}''$, let $m_A: Q^k|_A \hookrightarrow Q^k$ be the corresponding embedding. Consider the formula

$$\begin{aligned} \text{CommergeWithId}_M: \quad \forall_{Q^k} x, \quad & \bigwedge_{0 \leq i < j < k} \text{Id}(\text{restr}_{v_i}(x), \text{restr}_{e_{i,j}}(x)) \\ & \wedge \bigwedge_{A \in \mathcal{A}''} \text{commute}(\text{restr}_A(x)) \quad \rightarrow \quad \text{commute}(x). \end{aligned}$$

The projection $\pi: Q^k \rightarrow Q$ induces a pullback π^* between diagrams on Q and that on Q^k . It is easy to check that π^* induces a bijection between the diagrams on Q , verifying the condition of the commerge problem described in the proof of Proposition 29, and the diagrams on Q^k verifying the condition of CommergeWithId_M . Moreover the bijection preserves commutativity. Hence we conclude the proof as for Proposition 29. \blacktriangleleft

7 Conclusion

We have shown that the many-sorted signature Σ is expressive enough to formulate a theory \mathcal{T}_{cat} and its extension \mathcal{T}_{ab} , whose models are exactly and respectively diagrams in small categories, and that in small abelian categories. Restricting sorts to acyclic quivers makes the commerge problem for \mathcal{T}_{cat} decidable. A companion file [1] to this submission illustrates how to implement a deep embedding of formulas of Σ using the Coq proof assistant [13]; its content should be easy to transpose to other proof systems. Theorem `duality_theorem_with_theory` shall bring a duality principle, i.e., that a formula of the language is valid if and only if its dual is valid, to any formalized definition of abelian categories. Theorem 28 results in a complete decision procedure for commutativity clauses. The optimizations that make it work on concrete examples however go beyond the scope of the present article.

Similar concerns have motivated the implementation of the accomplished Globular proof assistant [2], for higher-dimensional category theory. The closest related work we are aware of yet seem unpublished at the time of writing. Lafont's categorical diagram editor [7], based on the Unimath library [14] and Barras and Chabassier's graphical interface for diagrammatic proofs [3] both provide a graphical interface for generating Coq proof scripts and visualizing Coq goals as diagrams. No specific automation is however provided. Himmel [6] describes a

formalization of abelian categories in Lean [4], including proofs of the five lemma and of the snake lemma, and proof (semi-)automation tied to this specific formalization. Duality arguments are not addressed. Monbru [10] also discusses automation issues in diagram chases, and provides heuristics for generating them automatically, albeit expressed in a pseudo-language.

References

- 1 http://matthieu.piquerez.fr/partage/FANL_duality.v.
- 2 Krzysztof Bar, Aleks Kissinger, and Jamie Vicary. Globular: an online proof assistant for higher-dimensional rewriting. *Log. Methods Comput. Sci.*, 14(1), 2018. doi:10.23638/LMCS-14(1:8)2018.
- 3 Luc Chabassier and Bruno Barras. A graphical interface for diagrammatic proofs in proof assistants. Contributed talks in the 29th International Conference on Types for Proofs and Programs (TYPES 2023), 2023. <https://types2023.webs.upv.es/TYPES2023.pdf>.
- 4 Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. The lean theorem prover (system description). In Amy P. Felty and Aart Middeldorp, editors, *Automated Deduction - CADE-25 - 25th International Conference on Automated Deduction, Berlin, Germany, August 1-7, 2015, Proceedings*, volume 9195 of *Lecture Notes in Computer Science*, pages 378–388. Springer, 2015. doi:10.1007/978-3-319-21401-6_26.
- 5 Peter Freyd. *Abelian categories. An introduction to the theory of functors*. Harper’s Series in Modern Mathematics. Harper & Row Publishers, New York, 1964.
- 6 Markus Himmel. Diagram chasing in interactive theorem proving. Bachelorarbeit. Karlsruhe Institut für Technologie, 2020. <https://pp.ipd.kit.edu/uploads/publikationen/himmel20bachelorarbeit.pdf>.
- 7 Ambroise Lafont. A categorical diagram editor to help formalising commutation proofs. <https://amblafont.github.io/graph-editor/index.html>.
- 8 Saunders Mac Lane. *Homology*. Class. Math. Berlin: Springer-Verlag, reprint of the 3rd corr. print. 1975 edition, 1995.
- 9 Andrey A. Markov. The impossibility of algorithms for the recognition of certain properties of associative systems. *Dokl. Akad. Nauk SSSR, n. Ser.*, 77:953–956, 1951.
- 10 Yannis Monbru. Towards automatic diagram chasing. M1 report. École Normale Supérieure Paris-Saclay, 2022. https://github.com/ymonbru/Diagram-chasing/blob/main/MONBRU_Yannis_Rapport.pdf.
- 11 Matthieu Piquerez. *Tropical Hodge theory and applications*. PhD thesis, Institut Polytechnique de Paris, November 2021. URL: <https://theses.hal.science/tel-03499730#>.
- 12 Emily Riehl. *Category Theory in Context*. Dover Publications, 2017. <https://math.jhu.edu/~eriehl/context.pdf>.
- 13 The Coq Development Team. The coq proof assistant, June 2023. doi:10.5281/zenodo.8161141.
- 14 Vladimir Voevodsky. Univalent semantics of constructive type theories. In Jean-Pierre Jouannaud and Zhong Shao, editors, *Certified Programs and Proofs - First International Conference, CPP 2011, Kenting, Taiwan, December 7-9, 2011. Proceedings*, volume 7086 of *Lecture Notes in Computer Science*, page 70. Springer, 2011. doi:10.1007/978-3-642-25379-9_7.