A First Order Theory of Diagram Chasing

Assia Mahboubi

Nantes Université, École Centrale Nantes, CNRS, INRIA, LS2N, UMR 6004, France
Vrije Universiteit Amsterdam, The Netherlands

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

Digital Object Identifier 10.4230/LIPIcs.CSL.2024.38

Related Version
Full Version: `https://hal.science/hal-04266479`
Full Version: `https://arxiv.org/abs/2311.01790`

Supplementary Material *Software*: `https://gitlab.inria.fr/mpiquere/coq-diagram-chasing`
archived at `swh:1:dir:be0a95b35dcfefe6d05a2446be611fa81f424995`

Funding This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101001995).

Acknowledgements The authors would like to thank Kenji Maillard, Loïc Pujet and the anonymous reviewers for their useful comments and suggestions on this work.

1 Introduction

Homological algebra [12, 14] 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. Diagrams appear as early as in the introduction of Mac Lane's classic reference book [13], while Riehl's textbook devotes a section to the *art of diagram chase* [18, Section 1.6]. Lawvere and Schanuel's pedagogical introduction to category theory [11, Session 17] uses an entire session to discuss the role of graphs in diagrammatic categorical reasoning.

A diagram can be seen as a functor $F\colon J \to \mathcal{C}$, whose domain J, the indexing category, is a small category [18] sometimes also called the *shape* of the diagram [14]. 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 diagram on Figure 1 asserts that morphism e is equal to the composition of morphisms a and b, denoted by $b \circ a$, as well as to the composition $d \circ c$. Commutativity of diagrams in certain categories can be used to state more involved properties, and diagram chasing essentially consists in establishing the existence, injectivity,

32nd EACSL Annual Conference on Computer Science Logic (CSL 2024).
Editors: Aniello Murano and Alexandra Silva; Article No. 38; pp. 38:1–38:19
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

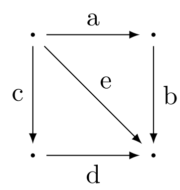


Figure 1 Square diagram with a diagonal arrow.

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 results with 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 [16, p.338]), 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 *automating* proofs of commutativity clauses. Observe for instance, that the commutativity of the square diagram of Figure 1 follows from that of the two triangle sub-diagrams: the concluding step in diagram chases usually amounts to a such a commutativity clause. Automating the mundane proofs of commutativity clauses amounts to studying a decision problem hereafter referred to as the *commerge* problem: *Given a collection of sub-diagrams of a larger diagram which commute, must the entire diagram commute?*

For this purpose, we introduce the two following formal languages.

▶ **Definition 1.** *We define a many-sorted signature* $\mathring{\Sigma}$ *with sorts the collection of finite quivers. Signature* $\mathring{\Sigma}$ *has one function symbol* $\mathrm{restr}_{m\colon Q'\to Q}$*, of arity* $Q \to Q'$*, per each quiver morphism* $m\colon Q' \to Q$ *between two quivers* Q *and* Q'*, and one predicate symbol* $\mathrm{commute}_Q$*, on sort* Q*, for each finite quiver* Q*.*

Similarly, we define a many-sorted signature Σ *with sorts the collection of finite* acyclic *quivers. Signature* Σ *has one function symbol* $\mathrm{restr}_{m\colon Q'\hookrightarrow Q}$*, of arity* $Q \to Q'$*, per each* injective *quiver morphism* $m\colon Q' \hookrightarrow Q$ *between two quivers* Q *and* Q'*, and one predicate symbol* $\mathrm{commute}_Q$*, on sort* Q*, for each finite acyclic 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 in 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 automation 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. Finally in Section 6, we formalize and study the decidability of several variants of the commerge problem.

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 $\mathrm{card}(A)$ the cardinal of a finite set A. We use the notation id for the identity map.

2.1 Quivers

In this section, we introduce some vocabulary and notations related to directed multi-graphs, also called simply *graphs* in some standard category theory textbooks [11, 13]. In this article, we depart from these texts and use instead the term *quiver*.

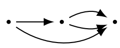
▶ **Definition 2** (General quiver, dual). *A* general quiver $\mathcal{Q}$ *is a quadruple* $(V_\mathcal{Q}, A_\mathcal{Q}, s_\mathcal{Q} \colon A_\mathcal{Q} \to V_\mathcal{Q}, t_\mathcal{Q} \colon A_\mathcal{Q} \to V_\mathcal{Q})$ *where* $V_\mathcal{Q}$ *and* $A_\mathcal{Q}$ *are two sets. The element 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* $\mathcal{Q}^\dagger := (V_\mathcal{Q}, A_\mathcal{Q}, t_\mathcal{Q}, s_\mathcal{Q})$*, which swaps the source and the target maps of* $\mathcal{Q}$*.*

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

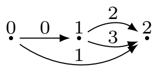
If A *is a subset of* $A_\mathcal{Q}$*, the* (spanning) restriction of $\mathcal{Q}$ to A *denoted by* $\mathcal{Q}|_A$ *is the quiver* $(V_\mathcal{Q}, A, s_\mathcal{Q}|_A, t_\mathcal{Q}|_A)$*. There is a canonical embedding* $\mathcal{Q}|_A \hookrightarrow \mathcal{Q}$*.*

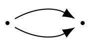
We denote by $\varnothing$ the empty quiver with no vertex and no arrow, and by $\mathring{S}$ the set of quivers Q such that V_Q and A_Q are finite subsets of $\mathbb{N}$. In this article, a *quiver* refers to an element of $\mathring{S}$. We use a non-cursive Q for elements of $\mathring{S}$, and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

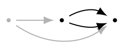
For the sake of readability, we use drawings to describe some elements of $\mathring{S}$, as for instance:



For a quiver Q denoted by such a drawing, the convention is that $V_Q = [\mathrm{card}(V_Q)]$ and $A_Q = [\mathrm{card}(A_Q)]$. From left to right, the drawn vertices correspond to $0, 1, \dots, \mathrm{card}(V_Q) - 1$. 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  into the quiver drawn above.



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

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 $0 \le k \le l$ are two integers, we denote by $\mathsf{sp}_{k,l} \colon \mathrm{PQ}_k \hookrightarrow \mathrm{PQ}_l$ the leftmost embedding of PQ_k into PQ_l, i.e., such that $(\mathsf{sp}_{k,l})_V(0) = 0$. If k and l are clear from the context, we draw $\mathsf{sp}_{k,l}$ as -∼ if $k \neq 0$ and as •∼ if $k = 0$. Moreover, we denote by $\mathsf{tp}_{k,l} \colon \mathrm{PQ}_k \hookrightarrow \mathrm{PQ}_l$ the rightmost embedding of PQ_k into PQ_l, i.e., such that $(\mathsf{tp}_{k,l})_V(k) = l$. The corresponding drawings are ∼- and ∼•. Moreover, if P is a nontrivial path-quiver, we define $\mathsf{st}_P \colon \bullet\ \bullet \hookrightarrow P$ 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 •∼•.

If $\mathcal{Q}$ is a general quiver, a morphism of the form $p \colon \mathrm{PQ}_k \to \mathcal{Q}$, for some k, is called a *path of $\mathcal{Q}$ from u to v of length k*, where $u \coloneqq p(0)$ and $v \coloneqq p(k)$. If moreover p_A is injective, it is called a *trail* [24]. A trail p such that $p_V(0) = p_V(k)$ is called a *cycle*. Two paths $p_1 \colon P_1 \to \mathcal{Q}$, $p_2 \colon P_2 \to \mathcal{Q}$ of $\mathcal{Q}$ *have the same extremities* if $p_1 \circ \mathsf{st}_{P_1} = p_2 \circ \mathsf{st}_{P_2}$. We denote by $\mathcal{BP}_{\mathcal{Q}}$, resp. $\mathcal{BT}_Q$, the set of pairs of paths, resp. of trails, of $\mathcal{Q}$ having the same extremities. Such a pair is called a *bipath*, resp. a *bitrail*.

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.

We now recall the definition of a *free category* (see for instance [13, Section II.7]).

▶ **Definition 5** (Free category). *For a general quiver $\mathcal{Q}$, the* free category over $\mathcal{Q}$*, denoted by $\langle \mathcal{Q} \rangle$ is the category with objects $\mathrm{Ob}_{\langle \mathcal{Q} \rangle} = V_{\mathcal{Q}}$ whose morphisms $\mathrm{Hom}_{\langle \mathcal{Q} \rangle}(u, v)$, 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

The purpose of this section is to introduce diagrams in a category, and to define what it means for a diagram to commute. We also introduce useful constructions later used for building new diagrams by restricting and pasting existing ones.

▶ **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 $\mathcal{C}$.*

Definition 6 is actually a special case of $\mathcal{D}$-shaped diagrams, for $\mathcal{D}$ a small category [14]. Thanks to the universal properties of free categories, it coincides with Mac Lane's Q-shaped diagrams in category $\mathcal{C}$ [13, Section II.7].

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

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

Note that term pullback here refers to pre-composition rather than to fiber products.

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

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

▶ **Definition 9** (Pushout configuration). *Consider four quivers Q, Q_1, Q_2, Q' and four embeddings $m_1 \colon Q \hookrightarrow Q_1$, $m_2 \colon Q \hookrightarrow Q_2$, $m'_1 \colon Q_1 \hookrightarrow Q'$ and $m'_2 \colon Q_2 \hookrightarrow Q'$:*

$$\begin{array}{ccc} Q & \xhookrightarrow{m_1} & Q_1 \\ \big\downarrow{\scriptstyle m_2} & & \big\downarrow{\scriptstyle m'_1} \\ Q_2 & \xhookrightarrow{m'_2} & Q'. \end{array}$$

This data is a pushout configuration *if:*

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

In this case, the triple (Q', m'_1, m'_2) is called a pushout of (Q, Q_1, Q_2, m_1, m_2).

▶ Remark 10. Any tuple (Q, Q_1, Q_2, m_1, m_2) has a pushout, and any two pushouts of the same tuple are isomorphic, see for instance Figure 2.

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

▶ **Lemma 11.** *Let $\mathcal{C}$ be some category. Consider a pushout configuration as in Definition 9. 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. ◀

2.3 Category congruences, path relations and quotient categories

We first introduce category congruences and quotients, following Mac Lane [13, Section II.8], and provide an important example thereof. However, we slightly depart from this reference by only considering the special case of quotients of categories by congruences, as it is the only one used in the present article.

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

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

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

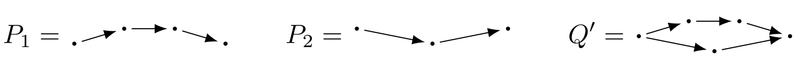


Figure 2 A pushout Q' of two path-quivers P_1 and P_2 with respect to st_{P_1} and st_{P_2}.

▶ **Proposition 13** (Quotient category). *Given such a category congruence, the data* $\mathrm{Ob}_{\mathcal{C}/r} \coloneqq \mathrm{Ob}_{\mathcal{C}}$ *and* $\mathrm{Hom}_{\mathcal{C}/r}(A, B) \coloneqq \mathrm{Hom}_{\mathcal{C}}(A, B)/r_{A,B}$ *define a category* $\mathcal{C}/r$ *called the* quotient category *of* $\mathcal{C}$ *by* r*.*

Proof. Immediate. ◀

We now introduce relations on pairs of paths with same extremities in a general quiver: the ones that are induced from a congruence on the corresponding free category are of special interest.

▶ **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* $(p, q) \in r$*. Note that* $\mathcal{BP}_{\mathcal{Q}} = \bigsqcup_{A,B \in \mathrm{Ob}_{\langle \mathcal{Q} \rangle}} \mathrm{Hom}_{\langle \mathcal{Q} \rangle}(A, B)^2$*. Such a relation* r *is called a* path relation *if it is a category congruence on* $\langle \mathcal{Q} \rangle$*. The complete path relation on* $\mathcal{Q}$*, i.e.,* $\mathcal{BP}_{\mathcal{Q}}$*, is denoted by* $\mathrm{tot}_{\mathcal{Q}}$*.*

For instance, in a small category $\mathcal{C}$ the composition axiom induces a path relation on the underlying quiver $\mathcal{Q}$. To be more precise, this relation is given by

$$\{(p_1, p_2) \in \mathcal{BP}_{\mathcal{Q}} \mid \mathrm{comp}(F \circ \Phi_{p_1}) = \mathrm{comp}(F \circ \Phi_{p_2})\}$$

where F is the canonical functor from $\langle \mathcal{Q} \rangle$ to $\mathcal{C}$. Following the usual notation of ideals, if $r_0, \ldots, r_{l-1}$ are some relations between paths with same extremities, we denote by $(r_i \mid i \in [l])$ the smallest path relation containing $r_0, \ldots, r_{l-1}$.

Let $\mathcal{Q}'$ be another general quiver and let $m \colon \mathcal{Q} \to \mathcal{Q}'$ 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 A finite characterization of commutative diagrams

Definition 8 about the commutativity of a finite diagram a priori relies on an infinite number of equations. In this section, we give an equivalent formulation based on the finite set $\mathcal{BT}_Q$.

▶ **Lemma 15.** *Let* $\mathcal{Q}$ *be a quiver. The smallest path relation containing* $\mathcal{BT}_{\mathcal{Q}} \subseteq \mathcal{BP}_{\mathcal{Q}}$ *is the total relation* $\mathrm{tot}_{\mathcal{Q}}$*.*

Proof. Denote this smallest path relation by r. It suffices to prove that any path is related with a trail. Let p be a path which is not a trail. Then p contains a nontrivial cycle, i.e., if we denote by $\gg$ the concatenation operator on paths, $p = p_1 \gg c \gg p_2$ for some paths p_1 and p_2 and some nontrivial cycle c. Since any cycle is a trail, the pair $(c, c \circ \leftharpoondown)$ belongs to $\mathcal{BT}_{\mathcal{Q}}$, where $c \circ \leftharpoondown$ is the path of length 0 based on $c_V(0)$. Hence $p = p_1 \gg c \gg p_2 \sim p_1 \gg p_2$. This last past is strictly smaller, and we conclude by infinite descent. ◀

As a direct consequence of the previous lemma, we get the following proposition.

▶ **Proposition 16.** *A diagram* D *over a finite quiver* Q *is commutative if and only if*

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

2.5 Many-sorted logic, categorical interpretation

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

Remember that signature Σ only differs from $\mathring{\Sigma}$ by restricting the allowed sorts to the acyclic quivers and the allowed quiver morphisms to embeddings.

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 $\mathrm{restr}_{m\colon Q \to Q'}(t)$, with t a term of sort Q', Q' a quiver and $m\colon Q \to Q'$ a morphism of quivers. When possible, we leave the sorts implicit and simplify the notation of symbol $\mathrm{restr}_{m\colon Q' \to Q}$ into restr_m.

We denote the equality symbols by $\approx$, leaving sorts implicit. An *atom* is thus of the form $s \approx t$ with s and t two terms of the same sort, or of the form $\mathrm{commute}_Q(t)$ 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.

We shorten $\forall_Q x, \forall_Q y, P(x,y)$, resp. $\exists_Q x, \exists_Q y, P(x,y)$, into $\forall_Q x, y, P(x,y)$, resp. into $\exists_Q x, y, P(x,y)$. We write $\exists!_Q\, y, P(y)$ for formula $\big(\exists_Q\, y,\ P(y)\big) \wedge \big(\forall_Q\, y_1, y_2,\ P(y_1) \wedge P(y_2) \to y_1 \approx y_2\big)$. A formula with free variables $x_1, \ldots, x_n$ of respective sorts $Q_1, \ldots, Q_n$ is said to be of *arity* $Q_1 \times \ldots Q_n$.

We now define standard, sometimes also called Tarskian [20], semantics for first-order theories on signatures $\mathring{\Sigma}$ and Σ. We mostly follow classic presentations [9] but adapted to the context of multi-sorted signatures.

▶ **Definition 17** (Structures, models). *A* $\mathring{\Sigma}$-structure $\mathcal{M}$, *also called* interpretation of $\mathring{\Sigma}$, *is defined by:*

- *a collection of disjoint non-empty domain sets* $(\mathcal{M}_Q)_Q$*, indexed by the collection of (quiver) sorts;*
- *an interpretation of each function symbol* $\mathrm{restr}_{m\colon Q' \to Q}$*, as a function* $\mathrm{restr}^{\mathcal{M}}_{m\colon Q' \to Q}$ *with domain* $\mathcal{M}_Q$ *and codomain* $\mathcal{M}_{Q'}$*;*
- *an interpretation of each predicate symbol* $\mathrm{commute}_Q$*, as a subset* $\mathrm{commute}^{\mathcal{M}}_Q$ *of* $\mathcal{M}_Q$*.*

A given $\mathring{\Sigma}$*-structure together with a variable assignment mapping any variable of sort* Q *to an element of domain* $\mathcal{M}_Q$ *entail a truth value for any first-order formula* ϕ *on language* $\mathring{\Sigma}$*. If* ϕ *has no free variable, we write* $\mathcal{M} \models \phi$ *if* $\phi^{\mathcal{M}}$ *is true and we say that* ϕ *is valid in* $\mathcal{M}$*. A formula* ϕ *is* satisfiable *when there is a variable assignment which makes it true, and* unsatisfiable *in the opposite case. A* model of a theory $\mathcal{T}$ *on signature* $\mathring{\Sigma}$ *is the interpretation of a* $\mathring{\Sigma}$-structure *such that every formula in* $\mathcal{T}$ *is true.*

We also define analogue structures, interpretations, models for signature Σ*.*

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

▶ **Definition 18** (Categorical interpretation). *To each small category* $\mathcal{C}$*, we associate an interpretation of* Σ*, resp.* $\mathring{\Sigma}$*, 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* $m^*(D)$*.*
- $\mathrm{commute}^{\mathcal{C}}_Q$ *is the set of commutative diagrams in* $\mathcal{C}$ *over* Q*.*

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

3 A theory for diagrams in small categories

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

3.1 Axioms

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

Existence and uniqueness of the empty diagram

$$\text{EmptyEU}\colon \qquad \exists!_{\varnothing}\, x, \quad x \approx x$$

Compatibility of restrictions

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

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

Pushout

For any pushout configuration as in Definition 9, and using the same notations as this definition, we define the following formulas of arity $Q_1 \times Q_2 \times Q'$:

$$\text{Cospan}_{m_1',m_2'}(x_1, x_2, x')\colon \qquad \text{restr}_{m_1'}(x') \approx x_1 \ \wedge\ \text{restr}_{m_2'}(x') \approx x_2.$$

$$\text{PushoutEU}_{m_1,m_2,m_1',m_2'}\colon \qquad \forall_{Q_1}\, x_1, \forall_{Q_2}\, x_2, \quad \text{restr}_{m_1}(x_1) \approx \text{restr}_{m_2}(x_2) \\ \to\ \exists!_{Q'}\, x',\ \text{Cospan}_{m_1',m_2'}(x_1, x_2, x').$$

Composition

The following predicate, of arity $\bullet\!\longrightarrow\!\bullet \times \bullet\!\longrightarrow\!\bullet \times \bullet\!\longrightarrow\!\bullet$, describes composite of arrows:

$$\text{Comp}(x, y, z)\colon \qquad \exists_{\triangledown}\, w, \quad \text{restr}_{\triangledown}(w) \approx x \ \wedge\ \text{restr}_{\triangledown}(w) \approx y \\ \wedge\ \text{restr}_{\triangledown}(w) \approx z \ \wedge\ \text{commute}(w)$$

while the following one ensures the existence of compositions:

$$\text{CompE}\colon \qquad \forall_{\bullet\to\bullet}\, x, y, \quad \text{restr}_{\bullet\to\bullet}(x) \approx \text{restr}_{\bullet\to\bullet}(y)\ \to\ \exists_{\bullet\to\bullet}\, z,\ \text{Comp}(x, y, z).$$

Equality of nontrivial paths

For any two nontrivial path-quivers P_1 and P_2, it is possible to choose in a canonical way a quiver Q' together with two embeddings m_1', m_2' such that the following diagram forms a pushout configuration (as for instance on Figure 2):

$$\begin{array}{ccc} \bullet\ \bullet & \xhookrightarrow{\mathsf{st}_{P_1}} & P_1 \\ \big\downarrow \mathsf{st}_{P_2} & & \big\downarrow m_1' \\ P_2 & \xhookrightarrow{m_2'} & Q' \end{array}$$

We thus define the following predicate of arity $P_1 \times P_2$:

$$\text{EqPath}_{P_1,P_2}(x_1, x_2)\colon \qquad \text{restr}_{\bullet\ \bullet}(x_1) \approx \text{restr}_{\bullet\ \bullet}(x_2) \\ \wedge \quad \big(\forall_{Q'}\, x,\ \text{Cospan}_{m_1',m_2'}(x_1, x_2, x) \to \text{commute}(x)\big).$$

Identity

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

$$\begin{aligned}\mathrm{Id}(x,y)\colon\quad & \mathrm{restr}_{\bullet\to\cdot}(y) \approx x \quad\wedge\quad \forall_{\bullet\to\bullet}\, z, w,\\ & \big(\mathrm{Comp}(y,z,w) \to \mathrm{EqPath}(z,w)\big) \;\wedge\; \big(\mathrm{Comp}(z,y,w) \to \mathrm{EqPath}(z,w)\big)\end{aligned}$$

and the following formula ensures the existence of identity maps.

$$\mathrm{IdE}\colon\qquad \forall_{\bullet}\, x, \exists_{\bullet\to\bullet}\, y,\quad \mathrm{Id}(x,y).$$

Note that $\mathrm{Id}(x,y)$ and IdE entail that $\mathrm{restr}_{\cdot\to\bullet}(y) \approx x$.

Equality for general paths

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

$$\begin{aligned}\mathrm{EqPath}_{\bullet,\bullet}(x,y)\colon\quad & x \approx y,\\ \mathrm{EqPath}_{\bullet,P}(x,y)\colon\quad & \exists_{\bullet\to\bullet}\, z,\quad \mathrm{Id}(x,z) \;\wedge\; \mathrm{EqPath}_{\bullet\to\bullet,P}(z,y),\\ \mathrm{EqPath}_{P,\bullet}(x,y)\colon\quad & \mathrm{EqPath}_{\bullet,P}(y,x).\end{aligned}$$

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 $\mathrm{EqPathRefl}_{P_1}$, $\mathrm{EqPathSym}_{P_1,P_2}$ and $\mathrm{EqPathTrans}_{P_1,P_2,P_3}$ stating that the relation EqPath is respectively reflexive, symmetric and transitive.

We also make sure to enforce the properties of a category congruence. 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 $\mathrm{PQ}_{k_i+k_i'}$.

$$\begin{aligned}&\mathrm{EqPathConcat}_{P_1,P_2,P_1',P_2'}\colon\qquad \forall x_1, x_2, x_1', x_2',\\ &\quad \mathrm{EqPath}(x_1,x_2) \;\wedge\; \mathrm{EqPath}(x_1',x_2') \;\wedge\; \mathrm{restr}_{\sim\bullet}(x_1) \approx \mathrm{restr}_{\bullet\sim}(x_1')\\ &\quad\to\quad \forall x_1'', x_2'',\quad \mathrm{Cospan}_{\sim,\sim}(x_1, x_1', x_1'') \;\wedge\; \mathrm{Cospan}_{\sim,\sim}(x_2, x_2', x_2'')\\ &\qquad\qquad\qquad\qquad\to\quad \mathrm{EqPath}(x_1'', x_2'').\end{aligned}$$

Commutativity

We relate commutativity and equality by the following formula:

$$\mathrm{ComEq}\colon\qquad \forall_{\bullet\rightrightarrows\bullet}\, x,\quad \mathrm{commute}(x) \;\to\; \mathrm{restr}_{\bullet\rightrightarrows\bullet}(x) \approx \mathrm{restr}_{\bullet\rightrightarrows\bullet}(x).$$

For any quiver Q, the following formula provides an analogue of the notion of commutativity of diagrams via the characterization given in Proposition 16,

$$\mathrm{PathCom}_Q\colon\qquad \forall_Q\, x,\quad \bigwedge_{(p_1,p_2)\in\mathcal{BT}_Q} \mathrm{EqPath}(\mathrm{restr}_{p_1}(x), \mathrm{restr}_{p_2}(x)) \quad\leftrightarrow\quad \mathrm{commute}(x).$$

where we recall that $\mathcal{BT}_Q$ denotes the (finite) set of pairs of trails of Q having the same extremities.

▶ **Definition 19.** *Theory $\mathring{\mathcal{T}}_{\mathrm{cat}}$, over signature $\mathring{\Sigma}$, consists of the following formulas:*

- EmptyEU, CompE, IdE, ComEq,
- $\mathrm{RestrComp}_{m,m'}$ *for any pair of maps m and m' such that the codomain of m is the domain of m',*
- $\mathrm{PushoutEU}_{m_1,m_2,m_1',m_2'}$ *for a pushout configuration as in Definition 9,*
- $\mathrm{EqPathRefl}_{P_1}, \mathrm{EqPathSym}_{P_1,P_2}, \mathrm{EqPathTrans}_{P_1,P_2,P_3}, \mathrm{EqPathConcat}_{P_1,P_2,P_3,P_4}$ *for any quadruple of path-quivers P_1, P_2, P_3 and P_4,*
- $\mathrm{PathCom}_Q$ *for any quiver Q.*

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

3.2 Models

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

▶ **Theorem 20.** *Every categorical interpretation of Σ, resp. $\mathring{\Sigma}$, is a model of $\mathcal{T}_{\mathrm{cat}}$, resp. $\mathring{\mathcal{T}}_{\mathrm{cat}}$. Moreover, any model $\mathcal{M}$ of $\mathcal{T}_{\mathrm{cat}}$, resp. $\mathring{\mathcal{T}}_{\mathrm{cat}}$, has an isomorphic categorical interpretation.*

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

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

- the formulas PushoutEU follows from Lemma 11;
- IdE and CompE come from the existence of the identity map and the existence of the composition respectively;
- for two diagrams D_1 and D_2 respectively over path-quivers P_1 and P_2, the formula $\mathrm{EqPath}_{P_1,P_2}(D_1, D_2)$ corresponds to the relation $\mathrm{comp}(D_1) = \mathrm{comp}(D_2)$, which is a path relation;
- ComEq and $\mathrm{PathCom}_Q$ follow from Proposition 16.

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, resp. a, the corresponding embedding $v\colon \bullet \hookrightarrow Q$, resp. $a\colon \bullet \longrightarrow \bullet \hookrightarrow Q$. Here is a crucial lemma for the proof of the theorem.

▶ **Lemma 21** (General pushout). *Let $\mathcal{M}$ be a model of $\mathcal{T}_{\mathrm{cat}}$. Let Q be an acyclic quiver, $\beta_V\colon V_Q \to \mathcal{M}_{\bullet}$ and $\beta_A\colon A_Q \to \mathcal{M}_{\bullet\to\bullet}$ be two maps. Then the following statements are equivalent.*

1. *There exists an element β of $\mathcal{M}_Q$ such that, for any $v \in V_Q$ and any $a \in A_Q$, $\mathrm{restr}_v^{\mathcal{M}}(\beta) = \beta_V(v)$ and $\mathrm{restr}_a^{\mathcal{M}}(\beta) = \beta_A(a)$.*
2. *For any $a \in A_Q$, $\mathrm{restr}^{\mathcal{M}}_{\bullet\to\cdot}(\beta_A(a)) = \beta_V(s_Q(a))$ and $\mathrm{restr}^{\mathcal{M}}_{\cdot\to\bullet}(\beta_A(a)) = \beta_V(t_Q(a))$.*

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. In the case of an empty Q, 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 V_{Q_1}. We have the following pushout configuration:

$$\begin{array}{ccc} \varnothing & \xhookrightarrow{m_1} & Q_1 \\ \big\downarrow{\scriptstyle m_2} & & \big\downarrow{\scriptstyle m_1'} \\ \bullet & \xhookrightarrow[m_2' = v_0]{} & Q \end{array}$$

Point 2 holds trivially. Let us prove the existence and uniqueness of an element 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, $\mathrm{restr}^{\mathcal{M}}_{m_1}(\beta_1) = \mathrm{restr}^{\mathcal{M}}_{m_2}(\beta_2)$. Hence we can apply $\mathrm{PushoutEU}_{m_1,m_2,m_1',m_2'}$ to get a unique element $\beta \in \mathcal{M}_Q$ such that $\mathrm{Cospan}^{\mathcal{M}}_{m_1',m_2'}(\beta_1, \beta_2, \beta)$. From RestrComp and the induction hypothesis, it follows that for $\beta' \in \mathcal{M}_Q$, $\mathrm{Cospan}^{\mathcal{M}}_{m_1',m_2'}(\beta_1, \beta_2, \beta')$ is equivalent to $\mathrm{restr}^{\mathcal{M}}_v(\beta') = \beta_V(v)$ 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 $m_1 \colon \bullet\,\bullet \hookrightarrow Q_1$ mapping the first point to $s_Q(a_0)$ and the second point to $t_Q(a_0)$. Let $m_2 := \mathsf{st}_{\bullet\to\bullet} = \bullet \longrightarrow \bullet$. Once again, we get a pushout configuration:

$$\begin{array}{ccc} \bullet\;\bullet & \xhookrightarrow{m_1} & Q_1 \\ \big\downarrow{\scriptstyle m_2} & & \big\downarrow{\scriptstyle m_1'} \\ \bullet\longrightarrow\bullet & \xhookrightarrow[m_2' := a_0]{} & Q \end{array}$$

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 $\bullet\,\bullet$. Hence we deduce that

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

We can apply PushoutEU 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 20. Let $\mathcal{M}$ be a model of $\mathcal{T}_{\mathrm{cat}}$. We define the general quiver $\mathcal{Q}$ associated to $\mathcal{M}$ by

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

Set $\widetilde{\mathcal{C}} := \langle \mathcal{Q} \rangle$. Thanks to Lemma 21, to each acyclic quiver Q and to each element $\beta \in \mathcal{M}_Q$, we can associate a unique diagram $\widetilde{\Psi}(\beta)$ in $\widetilde{C}$ verifying:

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

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

Let us now introduce another important map. Let A and B be two objects of $\widetilde{\mathcal{C}}$, and let $p \in \mathrm{Hom}_{\widetilde{\mathcal{C}}}(A, B)$. Recall that p is just a path from A to B in $\mathcal{Q}$. Let k be the length of p. By Lemma 21, there exists a unique element $\Theta(p) \in \mathcal{M}_{\mathrm{PQ}_k}$ such that

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

Relation $\mathrm{EqPath}^{\mathcal{M}}$ thus induces a relation r on morphisms of $\widetilde{\mathcal{C}}$. Moreover, EqPathRefl, EqPathSym, EqPathTrans and EqPathConcat, together with Lemma 21, make r a category congruence. We can hence define the category $\mathcal{C} := \widetilde{\mathcal{C}}/r$. Now $\widetilde{\Psi}$ induces a map $\Psi\colon \mathcal{M}_Q \to \mathcal{C}_Q$ 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 $v \in V_Q$, $\widetilde{\Psi}(\beta)(v) = \widetilde{\Psi}(\gamma)(v)$, i.e., $\mathrm{restr}_v^{\mathcal{M}}(\beta) = \mathrm{restr}_v^{\mathcal{M}}(\gamma)$. Let a be an arrow of Q. Then we have the relation $\widetilde{\Psi}(\beta)(a) \sim \widetilde{\Psi}(\gamma)(a)$. By definition of $\mathrm{EqPath}_{\bullet\to\bullet,\bullet\to\bullet}$, we validate the premise of ComEq, and thus the equality $\widetilde{\Psi}(\beta)(a) = \widetilde{\Psi}(\gamma)(a)$, i.e., $\mathrm{restr}_a^{\mathcal{M}}(\beta) = \mathrm{restr}_a^{\mathcal{M}}(\gamma)$. By the uniqueness part of Lemma 21, we get $\beta = \gamma$.

We now consider the surjectivity of Ψ. It suffices to prove that any morphism $p \in \mathrm{Hom}_{\widetilde{\mathcal{C}}}(A, B)$, for any $A, B \in \mathrm{Ob}_{\widetilde{\mathcal{C}}}$, is in relation via r to a path of length one. Indeed, in such a case, for any diagram $\widetilde{D}$ in $\widetilde{\mathcal{C}}$ over Q, one can find another diagram $\widetilde{D}'$ over Q whose morphisms are path of size one and such that any morphism of $\widetilde{D}$ is in relation with the corresponding morphism of $\widetilde{D}'$. Hence the induced diagrams in $\mathcal{C}$ are equal. Moreover, $\widetilde{D}'$ is in the image of $\widetilde{\Psi}$, 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 $\gamma \in \mathcal{M}_{\bullet\to\bullet}$ such that $\mathrm{EqPath}^{\mathcal{M}}_{P,\bullet\to\bullet}(\beta, \gamma)$. If P has length one, this is trivial. If P has length zero, then by IdE, there exists γ such that $\mathrm{Id}^{\mathcal{M}}(\beta, \gamma)$. Moreover, by the definition $\mathrm{EqPath}_{\bullet,\bullet\to\bullet}$ and using the reflexivity of EqPath, we get $\mathrm{EqPath}^{\mathcal{M}}_{\bullet,\bullet\to\bullet}(\beta, \gamma)$. If P has length two, then by CompE, we can find an element γ such that

$$\mathrm{Comp}(\mathrm{restr}^{\mathcal{M}}_{\bullet\to\bullet\to\bullet}(\beta), \mathrm{restr}^{\mathcal{M}}_{\bullet\to\bullet\to\bullet}(\beta), \gamma).$$

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

For P with length $k > 2$, we proceed by induction on the length of P. Using EqPathTrans, it suffices to find γ over PQ_{k-1} such that $\mathrm{EqPath}^{\mathcal{M}}_{P,\mathrm{PQ}_{k-1}}(\beta, \gamma)$. To do so, we see P as the pushout of PQ_2 and PQ_{k-2} along $m_1' = \smile\colon \mathrm{PQ}_2 \to P$ and $m_2' = \smile\colon \mathrm{PQ}_{k-2} \to P$. By the case $k = 2$, we can find $\gamma_1 \in \mathcal{M}_{\bullet\to\bullet}$ such that $\mathrm{EqPath}(\mathrm{restr}^{\mathcal{M}}_{m_1'}(\beta), \gamma_1)$. We set $\gamma_2 = \mathrm{restr}^{\mathcal{M}}_{m_2'}(\beta)$. In particular we have $\mathrm{EqPath}(\gamma_2, \gamma_2)$. By EqPathConcat, we get $\mathrm{EqPath}(\beta, \gamma)$ where $\gamma \in \mathcal{M}_{\mathrm{PQ}_{k-1}}$ is such that $\mathrm{Cospan}_{m_1', m_2'}(\gamma_1, \gamma_2, \gamma)$, 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 function restr and with predicate 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 and the characterization of commutativity given in Proposition 16. 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:

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

This concludes the proof of the theorem. ◀

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, as well as on acyclic quivers. We define the dual of a formula over $\mathring{\Sigma}$ as follows:

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

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

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

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

The duality involution also induces involutions respectively on formulas, theories and models over Σ.

▶ **Example 22.** If $\mathcal{C}$ is a small category, then the dual interpretation $\mathcal{C}^\dagger$ is isomorphic to the model of the dual category, both with respect to $\mathring{\Sigma}$ and to Σ.

▶ **Theorem 23** (Duality theorem)**.** *Let ϕ be a formula with free variables included in $Y \subseteq X \times S$, resp. in $Y \subseteq X \times \mathring{S}$, and let $\mathcal{M}$ be a model of Σ, resp. of $\mathring{\Sigma}$. Then*

$$\mathcal{M} \models \phi \quad \iff \quad \mathcal{M}^\dagger \models \phi^\dagger.$$

Proof. The proof is direct. ◀

▶ Remark 24. The duality principle has some useful direct consequences:

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

The following fact follows directly from this last point, Theorem 20 and Example 22.

▶ **Proposition 25.** *Models of $\mathcal{T}_{\mathrm{cat}}$ verify $\mathcal{T}_{\mathrm{cat}}^\dagger$, and models of $\mathring{\mathcal{T}}_{\mathrm{cat}}$ verify $\mathring{\mathcal{T}}_{\mathrm{cat}}^\dagger$.*

5 A theory for diagrams in 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 [7]. 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 an abelian group, but this fact rather follows from the axioms.

Let us first formulate common notions of category theory in the languages introduced in Section 2.5. Here is a predicate of arity $\cdot\longrightarrow\cdot$ which corresponds to monicity of a map in a category.

$$\begin{aligned}\mathrm{Mono}(x): \quad & \forall_{\cdot\rightrightarrows\cdot\to\cdot}\, y, \quad \mathrm{restr}_{\cdot\rightrightarrows\cdot\to\cdot}(y) \approx x \\ & \wedge\ \mathrm{commute}(\mathrm{restr}_{\cdot\rightrightarrows\cdot\to\cdot}(y)) \ \wedge\ \mathrm{commute}(\mathrm{restr}_{\cdot\rightrightarrows\cdot\to\cdot}(y)) \\ & \to\ \mathrm{commute}(\mathrm{restr}_{\cdot\rightrightarrows\cdot\to\cdot}(y)).\end{aligned}$$

The dual predicate to $\mathrm{Mono}(x)$ is named $\mathrm{Epi}(x)$.

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

$$\mathrm{cone}(Q) \coloneqq (V_Q \sqcup \{v_0\},\ A_Q \sqcup \{a_v \mid v \in V_Q\}, s_{\mathrm{cone}(Q)}, t_{\mathrm{cone}(Q)}),$$

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

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

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

We now formulate the notion of limit.

$$\begin{aligned} &\mathrm{Limit}_Q(x, y)\colon \qquad \mathrm{Cone}_Q(x, y) \quad \wedge \\ &\qquad \forall_{\mathrm{cone}(Q)}\, z, \quad \mathrm{Cone}_Q(x, z) \ \to\ \exists!_{\mathrm{cone}(\mathrm{cone}(Q))}\, w,\ \mathrm{Cone}(y, w) \wedge \mathrm{restr}_{\mathrm{cone}(i_Q)}(w) \approx z. \end{aligned}$$

We also introduce the dual $\mathrm{Colimit}_Q(x, y) \coloneqq \mathrm{Limit}^{\dagger}_{Q^{\dagger}}(x, y)$.

The introduction of monos, epis, limits and colimits allows to state Freyd's axioms of abelian categories [7]. First, we define zero objects and kernels of respective arity $\bullet$ and $\bullet \longrightarrow \bullet \times \bullet \longrightarrow \bullet$ as follows:

$$\begin{aligned} \mathrm{Zero}(x)\colon &\quad \forall_{\varnothing}\, y, \quad \mathrm{Limit}_{\varnothing}(y, x) \ \wedge\ \mathrm{Colimit}_{\varnothing}(y, x), \\ \mathrm{Ker}(x, y)\colon &\quad \exists_{\bullet \rightrightarrows \bullet \rightrightarrows \bullet}\, z, \qquad \mathrm{restr}_{\bullet \to \bullet}(z) \approx x \quad \wedge \quad \mathrm{restr}_{\bullet \to \bullet}(z) \approx y \\ &\qquad \wedge \quad \mathrm{Zero}(\mathrm{restr}_{\bullet}(z)) \quad \wedge \quad \mathrm{Limit}(\mathrm{restr}_{\bullet \to \bullet}(z), z). \end{aligned}$$

We also define $\mathrm{Coker}(x, y) \coloneqq \mathrm{Ker}^{\dagger}(x, y)$.

We define the category $\mathcal{T}_{\mathrm{ab}}$, resp. $\mathring{\mathcal{T}}_{\mathrm{ab}}$ as the extension of $\mathcal{T}_{\mathrm{cat}}$, resp. $\mathring{\mathcal{T}}_{\mathrm{cat}}$, by the following formulas.

$$\begin{aligned} \mathrm{ZeroE}\colon &\quad \exists_{\bullet}\, x, \quad \mathrm{Zero}(x), \\ \mathrm{ProductE}\colon &\quad \forall_{\bullet\,\bullet}\, x, \exists_{\mathrm{cone}(\bullet\,\bullet)}\, y, \quad \mathrm{Limit}_{\bullet\,\bullet}(x, y), \\ \mathrm{CoproductE}\colon &\quad \mathrm{ProductE}^{\dagger}, \\ \mathrm{KerE}\colon &\quad \forall_{\bullet \to \bullet}\, x, \exists_{\bullet \to \bullet}\, y, \quad \mathrm{Ker}(x, y), \\ \mathrm{CokerE}\colon &\quad \mathrm{KerE}^{\dagger}, \\ \mathrm{MonoNormal}\colon &\quad \forall_{\bullet \to \bullet}\, x, \quad \mathrm{Mono}(x) \ \to\ \exists_{\bullet \to \bullet}\, y,\ \mathrm{Ker}(y, x), \\ \mathrm{EpiNormal}\colon &\quad \mathrm{MonoNormal}^{\dagger}. \end{aligned}$$

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

▶ **Theorem 26.** *The categorical interpretation induced by any small abelian category is a model of $\mathcal{T}_{\mathrm{ab}}$. Conversely, any model of $\mathcal{T}_{\mathrm{ab}}$ is isomorphic to the categorical interpretation associated to some small abelian category.*

Proof. This follows from Theorem 20 and from [7, Chapter 2]. ◀

▶ **Proposition 27.** *The theory* $\mathcal{T}_{\mathrm{ab}}$ *implies its dual* $\mathcal{T}_{\mathrm{ab}}^{\dagger}$.

Proof. The theory $\mathcal{T}_{\mathrm{cat}}$ implies its dual by Proposition 25. Moreover, ZeroE clearly implies its dual. Finally, the other axioms have their dual in the theory, by definition thereof. ◀

▶ Remark 28. The theorem and the proposition also hold for theory $\mathring{\mathcal{T}}_{\mathrm{ab}}$.

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 $m_i \colon Q_i \to Q$ be a morphism. We define the following formula:

$$\mathrm{Commerge}_{m_0,\dots,m_{k-1}}\colon \qquad \forall_Q\, x, \quad \bigwedge_{i=0}^{k-1} \mathrm{commute}(\mathrm{restr}_{m_i}(x)) \quad \to \quad \mathrm{commute}(x).$$

▶ **Definition 29.** *Notations as above, the* acyclic, resp. cyclic, commerge problem for morphisms, or embeddings, $m_0, \dots, m_{k-1}$ and for a theory $\mathcal{T}$ *on language* Σ, *resp.* $\mathring{\Sigma}$, *is the problem of deciding the validity of* $\mathrm{Commerge}_{m_0,\dots,m_{k-1}}$ *in models of theory* $\mathcal{T}$.

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

▶ **Lemma 30.** *Notations as above, the formula* $\mathrm{Commerge}_{m_0,\dots,m_{k-1}}$ *is valid for models of* $\mathcal{T}_{\mathrm{cat}}$*, resp.* $\mathring{\mathcal{T}}_{\mathrm{cat}}$*, if and only if* $\langle Q\rangle/(r_i \mid i \in [k])$ *is a thin category.*

Proof. Set $\mathcal{C} \coloneqq \langle Q\rangle/(r_i \mid i \in [k])$. It is a model of $\mathcal{T}_{\mathrm{cat}}$, resp. $\mathring{\mathcal{T}}_{\mathrm{cat}}$. Moreover, the canonical diagram $D \colon \langle Q\rangle \to \mathcal{C}$ verifies the premise of $\mathrm{Commerge}_{m_0,\dots,m_{k-1}}$. 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 $\mathrm{comp}(p^*(D))$ is the class of p in the quotient, which is different of the class of q, that is of $\mathrm{comp}(q^*(D))$. Hence D is not commutative.

For the other direction, by Theorem 20, 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 $\mathrm{Commerge}_{m_0,\dots,m_{k-1}}$ factors through D, i.e., $D' = F \circ D$ for some functor $F \colon \mathcal{C} \to \mathcal{C}'$. If $\mathcal{C}$ is thin, then for any two paths p and q with same extremities in Q, we have:

$$\mathrm{comp}(p^*(D')) = F(\mathrm{comp}(p^*(D))) = F(\mathrm{comp}(q^*(D))) = \mathrm{comp}(q^*(D')).$$

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

▶ **Theorem 31.** *The acyclic commerge problem for embeddings and* $\mathcal{T}_{\mathrm{cat}}$ *is decidable.*

Proof. By Lemma 30, it suffices to decide if $\langle Q\rangle/(r_i \mid i \in [k])$ is a thin category. Since the set of paths in $\langle Q\rangle$ is finite, we can compute the relation $({m_i}_*(\mathrm{tot}_{Q_i}) \mid i \in [k])$ extensively and check whether it is complete. ◀

▶ **Proposition 32.** *The cyclic commerge problem for morphisms and* $\mathring{\mathcal{T}}_{\mathrm{cat}}$ *is undecidable.*

Proof. We proceed by reduction to an undecidability result, independently due to Adyan and Rabin [2, 17]. For B an arbitrary finite set and $\langle B\rangle$ the associated free monoid, let M be the finitely presentable monoid $\langle B \mid r = 1, r \in R\rangle$, for R a nonempty finite subset of $\langle B\rangle \setminus \{1\}$. In particular, any finitely presentable group is of this form. Hence, by the Adyan-Rabin theorem, determining the triviality of M from B and R 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 $p_\rho \colon \mathrm{PQ}_{k_\rho} \to Q$, for some $k_\rho \in \mathbb{N}$. Let Q_ρ be a pushout of the morphisms $\bullet\!\rightsquigarrow\!\bullet \colon \bullet\ \bullet \hookrightarrow \mathrm{PQ}_{k_\rho}$ and $\bullet \longrightarrow \bullet$. Let m_ρ be the extension to Q_ρ of p_ρ obtained by mapping the new arrow onto e. Also set $m_e \colon \circlearrowleft \to Q$ which maps the loop on e. Now, $\langle Q\rangle/\big(m_e^*(\mathrm{tot}_{\circlearrowleft}), (m_\rho^*(\mathrm{tot}_{Q_\rho}) \mid \rho \in R)\big)$ is the category associated to the monoid M. Hence this category is thin if and only if the monoid is trivial. The Adyan-Rabin theorem and Lemma 30 conclude the proof. ◀

We strengthen the previous proposition to the case of embeddings.

▶ **Theorem 33.** *The cyclic commerge problem for embeddings and $\mathring{\mathcal{T}}_{\mathrm{cat}}$ is undecidable.*

Proof. Let B, R, M as in the proof of Proposition 32. Let Q be the quiver $(\{v\}, B, s_Q, t_Q)$ (note that we removed the arc e). If $n \geq 2$, we define the quiver $\mathring{Q}^n$ as

$$\mathring{Q}^n \coloneqq \Big(\{v_i \mid i \in [n]\}, \{b_{i,j} \mid b \in B, 0 \leq i < j < n\} \sqcup \{e_{i,j} \mid i,j \in [n]\}, s_{\mathring{Q}^n}, t_{\mathring{Q}^n}\Big) \text{ where}$$
$$s_{\mathring{Q}^n}(b_{i,j}) = v_i, \quad s_{\mathring{Q}^n}(e_{i,j}) = v_i, \quad t_{\mathring{Q}^n}(b_{i,j}) = v_j, \quad t_{\mathring{Q}^n}(e_{i,j}) = v_j.$$

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

For A any subset of $A_{\mathring{Q}^n}$, let $m_A \colon \mathring{Q}^n|_A \hookrightarrow \mathring{Q}^n$ be the canonical embedding, and let $r_A \coloneqq m_{A*}(\mathrm{tot}_{\mathring{Q}^n|_A})$. Set $r' \coloneqq (r_A \mid A \in \mathcal{A})$ for $\mathcal{A} \subset \mathcal{P}(A_{\mathring{Q}^n})$ defined as the set containing

- $A_e \coloneqq \{e_{i,j} \mid i,j \in [n]\}$,
- for $i < j$ and $b \in B$,

$$A_{b,i,j} \coloneqq \{\underbrace{e_{0,i}}_{\text{if } i \neq 0}, b_{i,j}, \underbrace{e_{j,n-1}}_{\text{if } j \neq n-1}, b_{0,n-1}\}.$$

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

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

Hence we get a natural transformation η between the identity functor and $\iota \circ \pi_*$ by setting $\eta_i \coloneqq e_{i,0} \in \mathrm{Hom}_{\langle \mathring{Q}^n \rangle / r'}(i, \iota \circ \pi(i) = v_0)$. Since $e_{i,0}$ is an isomorphism, we conclude that there is an equivalence of category between $\langle Q \rangle$ and $\langle \mathring{Q}^n \rangle / r'$.

Recall that M is the finitely presentable monoid $\langle B \mid r = 1, r \in R \rangle$. Assume that n is greater than the longest word in R. To any word $\rho = b^0 b^1 \cdots b^{l-1}$ in R corresponds a subset

$$A_\rho \coloneqq \{e_{0,l}, b^0_{0,1}, b^1_{1,2}, \ldots, b^{l-1}_{l-1,l}\} \subseteq A_{\mathring{Q}^n}.$$

Let $\mathcal{A}' \coloneqq \mathcal{A} \cup \{A_\rho \mid \rho \in R\}$. Then, $\langle \mathring{Q}^n \rangle / (r_A \mid A \in \mathcal{A}')$ is equivalent as a category to $\langle Q \rangle / (\pi_*(r_{A_\rho}) \mid \rho \in R)$ which is the category of the monoid M. Hence the validity of $\mathrm{Commerge}_{(m_A)_{A \in \mathcal{A}'}}$ is equivalent to the triviality of M. Once again, we conclude using Lemma 30 and the Adyan-Rabin theorem. ◀

▶ **Theorem 34.** *The theory $\mathcal{T}_{\mathrm{cat}}$ is undecidable.*

Proof. Let M, B, R be as in the proof of Theorem 33. Consider the acyclic quiver $Q :=$ $(\{v_0, v_1\}, B, s_Q, t_Q)$ where $s_Q(b) = v_0$ and $t_Q(b) = v_1$ for any $b \in B$. To a word $\rho =$ $b^0 b^1 \cdots b^{l-1}$ with $b^0, \ldots, b^{l-1} \in B$, for some $l \in \mathbb{N}$, we associate the predicate of arity Q

$$\mathrm{EqId}_\rho(x)\colon \qquad \exists_{\mathrm{PQ}_l}\, y, \qquad \bigwedge_{i\in[l]} \mathrm{restr}_{a_i}(y) \approx \mathrm{restr}_{b^i}(x) \quad \wedge \quad \mathrm{EqPath}_{\bullet,\mathrm{PQ}_l}(\mathrm{restr}_{\bullet\!\sim}(y), y)$$

where, for $i \in [l]$, $a_i\colon \bullet\!\longrightarrow\!\bullet \to \mathrm{PQ}_l$ maps the arc of $\bullet\!\longrightarrow\!\bullet$ onto the i-th arc of PQ_l and, as usual, for $b \in B$, $b\colon \bullet\!\longrightarrow\!\bullet \to Q$ maps the arc of $\bullet\!\longrightarrow\!\bullet$ onto the arc $b \in A_Q$. If D is a diagram over Q in some category $\mathcal{C}$, then $\mathrm{EqId}_\rho(D)$ is equivalent to the fact that $D(v_0) = D(v_1)$ and $D(b^{l-1}) \circ \cdots \circ D(b^0) = \mathrm{id}_{D(v_0)}$. In particular, a diagram $D\colon \langle Q\rangle \to \mathcal{C}$ verifies $\bigwedge_{\rho\in R} \mathrm{EqId}_\rho(D)$ if and only if it factorizes through the category associated to the monoid M. Hence, the triviality of the monoid M is equivalent to the validity of the following formula in $\mathcal{T}_{\mathrm{cat}}$:

$$\forall_Q\, x, \quad \bigwedge_{\rho\in R} \mathrm{EqId}_\rho(x) \;\rightarrow\; \mathrm{commute}(x).$$

The Adyan-Rabin theorem concludes the proof. ◀

7 Conclusion

We have shown that the many-sorted signature Σ is expressive enough to formulate a theory $\mathcal{T}_{\mathrm{cat}}$ whose models are exactly diagrams in small categories, as well as an extension $\mathcal{T}_{\mathrm{ab}}$ of $\mathcal{T}_{\mathrm{cat}}$ whose models are exactly diagrams in small *abelian* categories. Restricting sorts to acyclic quivers and morphisms to embeddings makes the commerge problem for $\mathcal{T}_{\mathrm{cat}}$ decidable, that is, one can decide when the commutativity of a given diagram follows from that of a given collection of sub-diagrams. Generalizing this study to monoidal categories [19] or more generally to higher category theory does not seem immediate. However, type-theoretic approaches have been successfully applied to devise a syntactic description of weak ω-categories [6] and of opetopes [22], and both these works laid the foundations of prototype proof assistants.

The signatures and theories presented in this article are shaped by their subsequent usage as *interfaces* in a computer-aided tool for diagram chases. This tool eventually produces formal proofs of the corresponding theorems for a given mathematical structure. Interfaces should indeed be convenient enough to fullfil for concrete applications, e.g., abelian groups. This motivation explains some seemingly odd choices, including the use of a commutation predicate instead of the arguably more natural equivalence relation on paths.

A companion file [1] to this submission illustrates how to implement a deep embedding of formulas of Σ using the Coq proof assistant [21]; its content should be easy to transpose to other proof systems. Theorem `duality_theorem_with_theory` expresses a general duality principle. It can be specialized, e.g., to any formalized definition of abelian categories, and the resulting instance of the theorem ensures that a formula of the language is valid for abelian categories if and only if its dual is valid. This companion file however does not provide any specific such formal definition of abelian categories. Theorem 31 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 about diagrammatic reasoning have motivated the implementation of the accomplished Globular proof assistant [3], for higher-dimensional category theory. Based on higher-dimensional rewriting, it implements various algorithms for constructing and comparing diagrams in higher categories. It is geared towards visualization rather than formal verification. The closest related work we are aware of seem unpublished at the time of writing. Lafont's categorical diagram editor [10], based on the Unimath library [23] and Barras and Chabassier's graphical interface for diagrammatic proofs [4] both provide a graphical interface for generating Coq proof scripts and visualizing Coq goals as diagrams. Himmel [8] describes a formalization of abelian categories in Lean [5], 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 [15] 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 Sergei Ivanovich Adyan. Algorithmic undecidability of problems of recognition of certain properties of groups. *Dokl. Akad. Nauk SSSR*, 103:533–535, 1955.

3 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.

4 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. URL: https://types2023.webs.upv.es/TYPES2023.pdf.

5 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.

6 Eric Finster and Samuel Mimram. A type-theoretical definition of weak ω-categories. In *32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, Reykjavik, Iceland, June 20-23, 2017*, pages 1–12. IEEE Computer Society, 2017. doi:10.1109/LICS.2017.8005124.

7 Peter Freyd. *Abelian categories. An introduction to the theory of functors.* Harper's Series in Modern Mathematics. Harper & Row Publishers, New York, 1964.

8 Markus Himmel. Diagram chasing in interactive theorem proving. Bachelorarbeit. Karlsruher Institut für Technologie, 2020. URL: https://pp.ipd.kit.edu/uploads/publikationen/himmel20bachelorarbeit.pdf.

9 Wilfrid Hodges. *A shorter model theory.* Cambridge: Cambridge University Press, 1997.

10 Ambroise Lafont. A categorical diagram editor to help formalising commutation proofs. https://amblafont.github.io/graph-editor/index.html.

11 F. William Lawvere and Stephen H. Schanuel. *Conceptual mathematics. A first introduction to categories.* Cambridge: Cambridge University Press, 2nd ed. edition, 2009.

12 Saunders Mac Lane. *Homology.* Class. Math. Berlin: Springer-Verlag, reprint of the 3rd corr. print. 1975 edition, 1995.

13 Saunders Mac Lane. *Categories for the working mathematician*, volume 5 of *Grad. Texts Math.* New York, NY: Springer, 2nd ed edition, 1998.

14 J. Peter May. *A concise course in algebraic topology.* Chicago, IL: University of Chicago Press, 1999.

15 Yannis Monbru. Towards automatic diagram chasing. M1 report. École Normale Supérieure Paris-Saclay, 2022. URL: https://github.com/ymonbru/Diagram-chasing/blob/main/MONBRU_Yannis_Rapport.pdf.

16 Matthieu Piquerez. *Tropical Hodge theory and applications.* PhD thesis, Institut Polytechnique de Paris, November 2021. URL: https://theses.hal.science/tel-03499730#.

17 Michael O. Rabin. Recursive unsolvability of group theoretic problems. *Ann. Math. (2)*, 67:172–194, 1958. `doi:10.2307/1969933`.

18 Emily Riehl. *Category Theory in Context*. Dover Publications, 2017. URL: `https://math.jhu.edu/~eriehl/context.pdf`.

19 Peter Selinger. A survey of graphical languages for monoidal categories. In Bob Coecke, editor, *New Structures for Physics*, volume 813 of *Lecture Notes in Physics*, pages 289–355. Springer, 2011. `doi:10.1007/978-3-642-12821-9_4`.

20 Alfred Tarski. The semantic conception of truth and the foundations of semantics. Philos. Phenomenol. Res. 4, 341-376 (1944)., 1944.

21 The Coq Development Team. The coq proof assistant, June 2023. `doi:10.5281/zenodo.8161141`.

22 Cédric Ho Thanh, Pierre-Louis Curien, and Samuel Mimram. A sequent calculus for opetopes. In *34th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2019, Vancouver, BC, Canada, June 24-27, 2019*, pages 1–12. IEEE, 2019. `doi:10.1109/LICS.2019.8785667`.

23 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`.

24 Douglas B. West. *Introduction to graph theory*. New Delhi: Prentice-Hall of India, 2nd ed. edition, 2005.