Bifibrational duality in non-abelian algebra and the theory of databases

by

Thomas Weighill

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics in the Faculty of Science at Stellenbosch University

Mathematics Division
Department of Mathematical Sciences
University of Stellenbosch
Private Bag X1, Matieland 7602, South Africa

Supervisor: Professor Z. Janelidze

December 2014

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 2014/09/30

Abstract

Bifibrational duality in non-abelian algebra and the theory of databases

T. Weighill

Mathematics Division
Department of Mathematical Sciences
University of Stellenbosch
Private Bag X1, Matieland 7602, South Africa

Thesis: MSc Mathematics

December 2014

In this thesis we develop a self-dual categorical approach to some topics in non-abelian algebra, which is based on replacing the framework of a category with that of a category equipped with a functor to it. We also make some first steps towards a possible link between this theory and the theory of databases in computer science. Both of these theories are based around the study of Grothendieck bifibrations and their generalisations. The main results in this thesis concern correspondences between certain structures on a category which are relevant to the study of categories of non-abelian group-like structures, and functors over that category. An investigation of these correspondences leads to a system of dual axioms on a functor, which can be considered as a solution to the proposal of Mac Lane in his 1950 paper "Duality for Groups" that a self-dual setting for formulating and proving results for groups be found. The part of the thesis concerned with the theory of databases is based on a recent approach by Johnson and Rosebrugh to views of databases and the view update problem.

Uittreksel

Bifibrasionele dualiteit in nie-abelse algebra en die teorie van databasisse

T. Weighill

Wiskunde Afdeling
Departement Wiskundige Wetenskappe
Universiteit Stellenbosch
Privaatsak X1, Matieland 7602, Suid-Afrika

Tesis: MSc Wiskunde

Desember 2014

In hierdie tesis word 'n self-duale kategoriese benadering tot verskeie onderwerpe in nie-abelse algebra ontwikkel, wat gebaseer is op die vervanging van die raamwerk van 'n kategorie met dié van 'n kategorie saam met 'n funktor tot die kategorie. Ons neem ook enkele eerste stappe in die rigting van 'n skakel tussen hierdie teorie and die teorie van databasisse in rekenaarwetenskap. Beide hierdie teorieë is gebaseer op die studie van Grothendieck bifibrasies en hul veralgemenings. Die hoof resultate in hierdie tesis het betrekking tot ooreenkomste tussen sekere strukture op 'n kategorie wat relevant tot die studie van nie-abelse groep-agtige strukture is, en funktore oor daardie kategorie. 'n Verdere ondersoek van hierdie ooreemkomste lei tot 'n sisteem van duale aksiomas op 'n funktor, wat beskou kan word as 'n oplossing tot die voorstel van Mac Lane in sy 1950 artikel "Duality for Groups" dat 'n self-duale konteks gevind word waarin resultate vir groepe geformuleer en bewys kan word. Die deel van hierdie tesis wat met die teorie van databasisse te doen het is gebaseer op 'n onlangse benadering deur Johnson en Rosebrugh tot aansigte van databasisse en die opdatering van hierdie aansigte.

Acknowledgements

I would like to thank Neil Ghani, Marino Gran, James Gray, George Janelidze, Michael Johnson and Tim Van der Linden for their interest in my work over the past two years, and Marino Gran for the invitation to visit Université catholique de Louvain. I would like to thank the scientific committees of the following conferences and workshops for the opportunity to present my work there: the International Category Theory Conference in Sydney, 2013 and Cambridge, 2014, the SAMS Annual Congress in Pietermaritzburg, 2013, the Workshop on Categorical Methods in Algebra and Topology in Coimbra, 2014, and the Workshop on Fibrations in Computing Science in Glasgow, 2014. I would like to thank Herman Engelbrecht and Gert-Jan van Rooyen for their support, and Naspers for its financial support via the MIH Media Lab. I would also like to thank all those in the Media Lab who made it an exciting place to work in. Finally, I would like to thank my supervisor Zurab Janelidze for his guidance over the last two years.

To Debbie, of whom this thesis is not worthy, but then again, neither am I.

Contents

Chapter 1

Introduction

1.1 Axiomatic contexts for group theory

In his 1950 paper *Duality for Groups* [73], Mac Lane states that

> "certain dualities arise in those theorems of group theory which deal, not with the elements of groups, but with subgroups and homomorphisms."

Thinking of subgroups as injective group homomorphisms, this observation naturally motivates the axiomatisation of the category of groups. Indeed, category theory deals with precisely those aspects of mathematics which can be described in terms of structure-preserving maps (for example, group homomorphisms).

By an axiomatisation of the category of groups, we mean the formulation of a list of axioms on a category which allows one to prove results from group theory. Such an axiomatisation would allow the formulation of general proofs which could be applied to other algebras such as rings, modules, loops or topological groups, depending of course on whether the axioms hold for these structures.

Mac Lane also mentions another advantage of doing group theory in such an axiomatic context, namely that of duality. Every category has a dual category whose objects are the same as the original category, but whose morphisms are in the opposite direction. If a result can be proved from axioms which are true of both the category and its dual, then for each result we obtain a dual result by reversing the arrows. In the paper, Mac Lane postulates that

> "a further development giving the first and second isomorphism theorems, and so on, can be made by introducing additional carefully chosen dual axioms."

Mac Lane only develops such a list of axioms for the case of abelian groups. Nonetheless, this development led to the notion of abelian category, a notion which dominated the first two decades of category theory [51] and was applied to areas such as algebraic geometry and homological algebra. Crucially, however, the category of all groups is not an abelian category. Moreover, many of the important properties shared by the category of abelian groups and the category of groups have duals which hold only for the abelian case.

The idea to axiomatise the category of all groups inspired other notions, such as protomodular categories, Mal'tsev categories and Barr-exact categories. These developments culminated in the notion of a semi-abelian category introduced by Janelidze, Márki and Tholen [51]. Subsequent developments demonstrated the power of semi-abelian categories as a context for doing group theory. The axioms defining a semi-abelian category, however, are not self-dual. In fact, if a category and its dual are both semi-abelian, then the category is necessarily abelian.

In this thesis, we make some first steps in one possible approach to Mac Lane's proposal, which is due to Z. Janelidze. The idea behind this approach is in some sense already contained in Mac Lane's paper. In the first quote, notice that Mac Lane lists subgroups and homomorphisms as separate entities, although typically we identify subgroups with monomorphisms in a category. The central idea behind the approach described in this thesis is to axiomatise not just the category of groups, but the bifibration of subgroups. By the *bifibration of subgroups*, we mean the following: let **Grp** be the category of all groups, and let **SubGrp** be the category whose objects are pairs (G, S) of groups with S a subgroup of G, and where a morphism $(G, S) \to (G', S')$ is a morphism $f : G \to G'$ in **Grp** such that the image of S under f is contained in S'. We then have a forgetful functor U (which we call the bifibration of subgroups):

$$U : \mathbf{SubGrp} \to \mathbf{Grp}$$

The functor U sends an object (G, S) to G and a morphism f to f seen as a morphism in **Grp**. Our goal is to formulate a list of axioms on a *functor* F which reflects the properties of the functor U. Thus we are indeed treating the notion of subgroup as fundamental (it is encoded by a functor, which may not be in any way associated with the structure of monomorphisms in the base category).

An immediate consequence of this idea is that the dual of an axiom for a functor $F : \mathbb{B} \to \mathbb{C}$ is that axiom stated for the *dual functor* $F^{\mathsf{op}} : \mathbb{B}^{\mathsf{op}} \to \mathbb{C}^{\mathsf{op}}$. It turns out that while the category of groups seems to lack duality, the bifibration of subgroups (and the bifibration of subobjects of semi-abelian categories in general) has many important properties which are self-dual in this sense [58].

The main idea underlying this thesis can be stated as follows: replace the axiomatic context of a category with that of a functor, and hence the notion of duality with *functorial duality*. The goal of this thesis is to explore this idea and relate it to other developments, particularly in the axiomatic treatment of groups and group-like structures. The main type of functor considered in this thesis is one which is faithful and amnestic, which we call a *form*.

There are other areas where the idea to study a certain type of functor rather than just a category is motivated by a concrete problem. In this thesis, we look at one such case, namely an approach to view updatability in databases introduced by Johnson and Rosebrugh [65].

A database can be seen as a structured way to store data. This data can be updated by inserting or deleting entries in such a way that the structure of the database is still preserved. Database states together with these updates form a category $\mathbb{D}$, which can be defined as the models for a sketch (see Chapter 6). However, in practical applications, it is often the case that a user has access to only part of the data contained in or derived from the database state. We refer to such a piece of the data as a *view*. From the perspective of the user, a view is just another database, and can be updated, and so we have another category $\mathbb{V}$ of *view states* and updates to the view. Since every update to the database immediately translates to the view in an obvious way, we have a functor:

$$V^* : \mathbb{D} \to \mathbb{V}.$$

An important question when dealing with views in practical applications is to what extent an update to a view can be propagated back to the rest of the database. This is not always possible, and the problem of finding if and how such a propagation can be done is called the *view update problem*.

The view update problem and the search for a self-dual axiomatic context for group theory are thus related superficially by the idea of replacing a category with a functor. However, it turns out that the connection is deeper than this: the two situations are linked by the notion of a fibration introduced by Grothendieck [45]. A view is defined to be completely updatable by Johnson and Rosebrugh when a restricted version of the functor V^* is a bifibration. On the other hand, being a bifibration (or having a slightly weaker property) is one of the most important axioms on a functor considered in the categorical algebra part of this thesis, since it allows the formulation of other axioms and theorems inspired by the category of groups.

1.2 Outline of the thesis

Chapter 2: Background In this chapter, we give some background necessary for the understanding and motivation of the rest of the thesis.

Chapter 3: Forms of subobjects and exact sequences In this chapter, we develop a correpondence between functors to a category and certain structures on a category given by a class of morphisms satisfying certain conditions. We characterise those functors which are, up to an equivalence, codomain functors restricted to a class of morphisms containing the identity morphisms. We also characterise those functors which are (up to an equivalence) so-called *forms of $\mathcal{N}$-short-exact-sequences* for an ideal $\mathcal{N}$ of null morphisms. We then derive a correspondence between conditions on these classes of morphisms and conditions on the correponding functors.

Chapter 4: Exact forms In this chapter, we look at the notion of *exact form* inspired by the First Isomorphism Theorem for groups, and some weaker notions. We show some links between these notions and the theory of factorisation systems [37], and also to Grandis exactness structures on a category [41]. In particular, we show that a Grandis ex4-category is the same as an Isbell bicategory [47] satisfying a certain condition.

Chapter 5: Forms of subobjects of normal categories In this chapter we give a characterisation of those categories with finite products which are normal categories [55] via self-dual conditions on the form of subobjects. We then combine this result with previous results to give characterisations of protomodular categories [13] and semi-abelian categories [51]. We end the chapter by giving an illustration of diagram chasing in the context of a form.

Chapter 6: Forms in database theory In this chapter, we given an overview of the approach to view updatability by Johnson and Rosebrugh [65], which uses the notion of a bifibration, and relate this approach to the work in the previous chapters.

1.3 Guide to the content

The terminology and results contained the Background section are taken from the literature. The proofs contained in the Background section are either the author's own or adaptation of proofs in the literature. In particular, the proofs of Proposition 8, Lemma 10 and Proposition 20 are the author's own, while the proof of Theorem 3 uses an important notion from the proof in [14] (the simplicial kernel of a relation), but is otherwise the author's own proof.

The work contained in Chapters 3, 4 and 5 is based on three joint papers with Z. Janelidze ([59],[60] and [61] respectively). Experts in the field who are familiar with these three papers may be interested in the following new results of the thesis which are not contained in these three papers:

- Chapter 3 contains a different perspective on the work in [59]. In particular, the notion of cover relation is not used. Chapter 3 also contains some generalisations of results in [59]. In particular, Theorems 4 and 5 generalise Theorem 2.4 from [59] to the case where the class of morphisms $\mathcal{M}$ is no longer required to contain only monomorphisms, and also no longer required to constitute a right factorisation system. Theorem 6 and Corollary 5 generalise Corollary 3.2 from [59]. The correspondence between the results in Chapter 3 and in [59] is also described in more detail in Chapter 3.

- In Chapter 4, Proposition 23 follows from a result (Theorem 5.7) in the third paper [61], while the observations following Lemma 31 are new.
- In Chapter 5, the results in Section 5.6 follow easily from a combination of [61] and [58] (as noted in the chapter), but they are not explicitly stated in [61]. Lemma 39 is also not contained in [61].

All the definitions in Chapter 6 are taken from the literature. The observations in Section 6.5 are (to the best of the author's knowledge) new. The results in Section 6.6 are based on remarks by Johnson and Rosebrugh [65] and Janelidze [50], but the exact statement and proofs are the author's own.

The following is a list of the main results and notions in the thesis which are solely due to the author:

- the original form of Theorem 4 in Chapter 3 (the current form incorporates a remark by Z. Janelidze);
- Theorems 5 and 6 in Chapter 3;
- the notion of an exact form in Chapter 4;
- the original form of Theorem 10 in Chapter 4 (the current form is taken from [60]);
- Propositions 23 and 24 in Chapter 4;
- condition (A4) and its role in Theorem 16 in Chapter 5;
- Proposition 26 in Chapter 5;
- the observations in Section 6.5 in Chapter 6;
- Theorem 18 in Chapter 6.

Chapter 2

Background

This chapter contains some background material which is necessary for the understanding of the subsequent chapters, and which also helps motivate the work contained in them and place this work in the context of the current literature. The first section deals with some types of functors which play an important role in the work in this thesis. In the second, we describe in more detail the motivation for the part of the thesis which is concerned with categorical algebra by examining a hypothesis due to Mac Lane. We then describe some of the work inspired by this hypothesis, namely the development of the theory of semi-abelian categories. In the third section we discuss the motivation for the work concerned with databases by discussing the view update problem, while also giving an overview of relational database theory.

We will assume familiarity with the basic notions of category, functor and natural transformation, as well as with limits and colimits, monomorphisms, epimorphisms and isomorphisms. A treatment of these topics can be found, for example, in Mac Lane's *Categories for the Working Mathematician* [75]. We will also assume a basic understanding of group theory, which can be gained, for example, from Mac Lane and Birkhoff's *Algebra* [7].

2.1 Special functors

In this section we look at some special types of functors, namely hom-functors, adjunctions and fibrations. The definitions and terminology in this section were taken from a combination of [75; 9; 10].

2.1.1 Preliminaries

For a category $\mathbb{C}$, we denote by $\mathbb{C}(C, C')$, or $\mathsf{hom}(C, C')$, the set of all morphisms from an object C to an object C' (such sets are in general called *hom-sets*). Given a functor $F : \mathbb{C} \to \mathbb{D}$, we say that F is *faithful* if for every pair of objects C, C' in $\mathbb{C}$, the induced map $\mathbb{C}(C, C') \to \mathbb{D}(F(C), F(C'))$ is injective. We say that F is *full* when this map is surjective for every pair (C, C'). We will call a functor $F : \mathbb{C} \to \mathbb{D}$ *essentially surjective* when every object in $\mathbb{D}$ is isomorphic to an object in the image of F.

For a functor $F : \mathbb{C} \to \mathbb{D}$, the *dual functor of* F is the functor $F^{\mathsf{op}} : \mathbb{C}^{\mathsf{op}} \to \mathbb{D}^{\mathsf{op}}$ whose definition on objects and arrows is the same as F when we think of objects and arrows of $\mathbb{C}^{\mathsf{op}}$ as objects and arrows (in the other direction) of $\mathbb{C}$.

Given any two categories $\mathbb{C}$, $\mathbb{D}$, we denote by $\mathbb{C}^{\mathbb{D}}$ (or sometimes $\mathsf{Fun}(\mathbb{D}, \mathbb{C})$) the category defined as follows:

- an object of $\mathbb{C}^{\mathbb{D}}$ is a functor F from $\mathbb{D}$ to $\mathbb{C}$;
- a morphism $\alpha : F \to F'$ in $\mathbb{C}^{\mathbb{D}}$ is a natural transformation from F to F'.

Given two functors $F, G : \mathbb{C} \to \mathbb{D}$, we will sometimes denote the set of all natural tranformations from F to G by $\mathsf{Nat}(F, G)$ (it is in fact nothing but the hom-set $\mathbb{C}^{\mathbb{D}}(F, G)$). There is an isomorphism of categories:

$$\mathsf{Fun}(\mathbb{D}, \mathbb{C}^{\mathbb{B}}) \cong \mathsf{Fun}(\mathbb{D} \times \mathbb{B}, \mathbb{C}) \tag{2.1.1}$$

which assigns to each functor $F : \mathbb{D} \to \mathbb{C}^{\mathbb{B}}$ the functor $F(-)(-)$, which sends an object (D, B) of $\mathbb{D} \times \mathbb{B}$ to $F(D)(B)$ and a morphism $(d : D \to D', b : B \to B')$ in $\mathbb{D} \times \mathbb{B}$ to the component of the natural transformation $F(d)$ at B. Now, $\mathbb{D} \times \mathbb{B}$ is isomorphic to $\mathbb{B} \times \mathbb{D}$, so we obtain the further isomorphism:

$$\mathsf{Fun}(\mathbb{D}, \mathbb{C}^{\mathbb{B}}) \cong \mathsf{Fun}(\mathbb{B}, \mathbb{C}^{\mathbb{D}})$$

For any functor $F : \mathbb{D} \to \mathbb{C}^{\mathbb{B}}$ and any object B, there is thus a functor $F(-)(B) : \mathbb{D} \to \mathbb{C}$, while for every morphism $b : B \to B'$ in $\mathbb{B}$ we obtain a natural transformation $F(-)(b) : F(-)(B) \to F(-)(B')$ whose component at an object D of $\mathbb{D}$ is $F(D)(b)$.

Finally, we recall some basic facts about natural transformations. Let $F : \mathbb{C} \to \mathbb{D}$ and $G, H : \mathbb{D} \to \mathbb{E}$ be functors and let $\alpha : G \Rightarrow H$ be a natural transformation. Then there is a natural transformation from GF to HF which we denote by αF, whose component at an object C in $\mathbb{C}$ is $\alpha_{F(C)}$. Now let $I : \mathbb{E} \to \mathbb{F}$ be a further functor. Then there is a natural transformation from IG to IH which we denote by $I\alpha$, whose component at an object D of $\mathbb{D}$ is $I(\alpha_D)$. Also, recall that if $\beta : H \to H'$ is another natural transformation, then we can compose α with β to form the natural transformation $\beta \circ \alpha : G \to H'$ (composition is simply done component-wise).

2.1.2 Hom-functors

We will denote the category of sets by **Set** and the category of (small) categories by **Cat**. Given a category $\mathbb{C}$, the category $\mathbb{C}^{\mathsf{op}} \times \mathbb{C}$ has as objects pairs (C, C') of objects of $\mathbb{C}$, while a morphism $(f, g) : (C, C') \to (D, D')$ is a pair (f, g) with $f : D \to C$ and $g : C' \to D'$ morphisms in $\mathbb{C}$. There is a functor

$$\mathsf{hom} : \mathbb{C}^{\mathsf{op}} \times \mathbb{C} \to \mathbf{Set},$$

called the *hom-functor*, which is defined as follows:

- for an object (C, C') of $\mathbb{C}^{\mathsf{op}} \times \mathbb{C}$, $\mathsf{hom}(C, C') = \mathbb{C}(C, C')$ is the set of all morphisms from C to C';
- for a morphism $(f, g) : (C, C') \to (D, D')$, $\mathsf{hom}(f, g)$ is the function which sends a morphism $c : C \to C'$ to the morphism $g \circ c \circ f : D \to D'$.

By the isomorphism (2.1.1) in the previous section, the hom-functor corresponds to a functor

$$Y : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}^{\mathbb{C}}.$$

This functor is called the *Yoneda functor* [75] or the *Yoneda embedding* and it has many important properties. The functor Y assigns to each object C in $\mathbb{C}$ the functor $\mathbb{C}(C, -) : \mathbb{C} \to \mathbf{Set}$, whose value at an object D is $\mathsf{hom}(C, D)$ and at a morphism f is $\mathsf{hom}(1_C, f)$. The following famous lemma is due to N. Yoneda.

Lemma 1 (Yoneda Lemma)**.** *Let $F : \mathbb{C} \to \mathbf{Set}$ be a functor and let C be an object in $\mathbb{C}$. Then there is a bijection*

$$y : \mathsf{Nat}(\mathbb{C}(C, -), F) \cong F(C)$$

which sends a natural transformation α to the image of the identity 1_C under α_C.

Proof. Let c be an element of $F(C)$. It is enough to show that there is exactly one natural transformation α such that $\alpha_C(1_C) = c$. Let f be a morphism in $\mathbb{C}(C, D)$. Then supposing that such an α does exist, $\alpha_D(f)$ is uniquely defined, as the commutative diagram below shows.

$$\begin{tikzcd}
\mathbb{C}(C, C) \arrow[r, "\alpha_C"] \arrow[d, "\mathbb{C}(C,f)"'] & F(C) \arrow[d, "F(f)"] \\
\mathbb{C}(C, D) \arrow[r, "\alpha_D"] & F(C)
\end{tikzcd}$$

Indeed, we have $\alpha_D(f) = \alpha_D \circ \mathbb{C}(C, f)(1_C) = F(f)(c)$. Now it remains to check that this always defines a natural transformation, which is easy. $\square$

The bijection y in the lemma above is in fact natural in the following two senses. Firstly, if F and G are two functors from $\mathbb{C}$ to **Set** and $\alpha : F \to G$ is a natural transformation, then the following diagram of sets commutes, where the horizontal arrows are instances of the bijection y:

$$\begin{tikzcd}
\mathsf{Nat}(\mathbb{C}(C,-),F) \arrow[r] \arrow[d, "\alpha\circ -"'] & F(C) \arrow[d, "\alpha_C"] \\
\mathsf{Nat}(\mathbb{C}(C,-),G) \arrow[r] & G(C)
\end{tikzcd}$$

Secondly, for any morphism $f : C \to D$ in $\mathbb{C}$, the following diagram of sets commutes:

$$\begin{tikzcd}
\mathsf{Nat}(\mathbb{C}(C,-),F) \arrow[r] \arrow[d, "-\circ\mathbb{C}(f,-)"'] & F(C) \arrow[d, "F(f)"] \\
\mathsf{Nat}(\mathbb{C}(D,-),F) \arrow[r] & F(D)
\end{tikzcd}$$

Note that for the first naturality to make sense, $\mathbb{C}$ needs to be a small category (see [75]).

Corollary 1. *For every category $\mathbb{C}$, the Yoneda embedding $Y : \mathbb{C}^{\mathsf{op}} \to \mathbf{Set}^{\mathbb{C}}$ is full and faithful.*

Proof. This is an easy consequence of the Yoneda Lemma. Replacing F by $\mathbb{C}(D,-)$ for some object D, we obtain a bijection:

$$y : \mathsf{Nat}(\mathbb{C}(C,-),\mathbb{C}(D,-)) \cong \mathbb{C}(D,C).$$

Looking at the proof of the Yoneda Lemma, we see that the inverse y^{-1} of this bijection is simply the action of Y on morphisms in $\mathbb{C}(D,C)$, so Y is indeed full and faithful. □

Note that there is also the dual functor $Y' : \mathbb{C} \to \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$, which is also sometimes referred to as the (covariant) Yoneda embedding. It is also clearly full and faithful. It is not hard to see that full and faithful functors reflect limits, colimits and identity morphisms, so in particular they also reflect monomorphisms, epimorphisms and isomorphisms. Thus the Yoneda embedding also reflects all these properties.

Another important property of the Yoneda embedding is that it preserves all limits which turn out to exist in $\mathbb{C}$. We first recall that limits in a functor category $\mathbb{C}^{\mathbb{B}}$ can often be computed "point-wise", as indicated by the following lemma.

Lemma 2. *Let $\mathbb{C}$ and $\mathbb{B}$ be two categories and let $F : \mathbb{D} \to \mathbb{C}^{\mathbb{B}}$ be a diagram in $\mathbb{B}$ such that for every object B, the diagram $F(-)(B) : \mathbb{D} \to \mathbb{C}$ has a limit. Then the limit of F is given by the functor $L : \mathbb{B} \to \mathbb{C}$ defined as follows:*

- *for an object B of $\mathbb{B}$, $L(B)$ is the limit of $F(-)(B) : \mathbb{D} \to \mathbb{C}$;*
- *for a morphism $b : B \to B'$, $L(b)$ is the unique arrow arising via the universal property of the limit $L(B')$ and the natural transformation $F(-)(b) \circ \alpha_B$, where α_B is the limit cone of $F(-)(B)$.*

Proof. It is easy to check that L is indeed a functor. Let D be an object of $\mathbb{D}$. For every object B of $\mathbb{B}$, we obtain a morphism $p_B : L(B) \to F(D)(B)$. The definition of L on morphisms ensures that the family $(p_B)_{B\in\mathbb{B}}$ actually constitutes a natural transformation λ_D from L to $F(D)$. Moreover, it is easy to check that by the definition of L, the family $(\lambda_D)_{D\in\mathbb{D}}$ constitutes a cone $\lambda : i_L \Rightarrow F$, where $i_L : \mathbb{D} \to \mathbb{C}^{\mathbb{B}}$ is constant functor on L. We claim that λ is a limit cone. Suppose there is a cone $i_J \Rightarrow F$ for some functor $J : \mathbb{B} \to \mathbb{C}$, where $i_J : \mathbb{D} \to \mathbb{C}^{\mathbb{B}}$ is the constant functor on J. Then for every object B, there is a cone $i_{J(B)} \Rightarrow F(-)(B)$, where $i_{J(B)} : \mathbb{D} \to \mathbb{C}$ is the constant functor on $J(B)$. This gives rise to a morphism $\alpha_B : J(B) \to L(B)$ by the property of the limit $L(B)$. If $(\alpha_B)_{B\in\mathbb{B}}$ constitutes a natural transformation α from J to L, then we are done, since then the uniqueness of α follows from the uniqueness of each α_B. It is easy to check that the naturality follows from the uniqueness part of the definition of a limit. □

Note that the above lemma can be dualized to say something about colimits in functor categories. Also, note that if $\mathbb{C}$ is complete, then so is $\mathbb{C}^{\mathbb{B}}$, with all (small) limits computed point-wise.

Lemma 3. *For a category $\mathbb{C}$ and an object C in $\mathbb{C}$, $\mathbb{C}(C,-) : \mathbb{C} \to$ **Set** preserves all limits that exist in $\mathbb{C}$.*

Combining the above two lemmas, we obtain:

Proposition 1. *For a category $\mathbb{C}$, the (covariant) Yoneda embedding $Y' : \mathbb{C} \to$ $\mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ preserves all limits that exist in $\mathbb{C}$.*

Proof. Let $D : \mathbb{D} \to \mathbb{C}$ be a diagram in $\mathbb{C}$ which has a limit. Since **Set** is complete, by Lemma 2, the limit of $Y' \circ D$ is the functor which takes an object C in $\mathbb{C}^{\mathsf{op}}$ to the limit of the diagram $\mathbb{C}(C, D(-)) : \mathbb{D} \to \mathbf{Set}$. Since the functor $\mathbb{C}(C,-)$ preserves limits by Lemma 3, this is the same as $\mathbb{C}(C, L)$, where L is the limit of D. Thus the image of the limit of D under Y' coincides on objects with the limit of $Y' \circ D$, and a straightforward calculation shows that the two functors agree on morphisms as well. Thus Y' preserves the limit of D. □

Lemma 4. *Let F and G be two functors from a category $\mathbb{C}$ to* **Set**, *and let $\alpha : F \to G$ be a natural transformation between them. Then α is a monomorphism in $\mathbf{Set}^{\mathbb{C}}$ if and only if every component of α is a monomorphism.*

Proof. Suppose α is a monomorphism, and let α_C be the component of α at an object C of $\mathbb{C}$. By the naturality part of the Yoneda Lemma, we have that the following diagram of sets commutes:

$$\begin{tikzcd}
\mathsf{Nat}(\mathbb{C}(C,-),F) \arrow[r] \arrow[d, "\alpha\circ -"] & F(C) \arrow[d, "\alpha_C"'] \\
\mathsf{Nat}(\mathbb{C}(C,-),G) \arrow[r] & G(C)
\end{tikzcd}$$

Now the map $\alpha \circ -$ is a monomorphism because α is a monomorphism, so the map α_C must also be a monomorphism, since the horizontal arrows represent bijections. The converse is easy to verify. □

The power of the Yoneda lemma is demonstrated by the following "metatheorem" taken from [11]. For an example of the application of this metatheorem, the Reader is referred to the proof of Theorem 3 in Section 1.2.6.

Metatheorem 1. *Let $\mathcal{P}$ be a statement of the form $\phi \Rightarrow \psi$, where ϕ and ψ can be expressed as conjunctions of properties in the following list:*

- *some finite diagram commutes;*
- *some morphism is a monomorphism;*
- *some morphism is an isomorphism;*
- *some finite diagram is a limit diagram;*
- *some arrow f factors (necessarily in a unique way) through a specified monomorphism m.*

Then if $\mathcal{P}$ is valid in **Set***, then it is valid in every category.*

Proof. Let $\mathbb{C}$ be any category. We have already mentioned that since $Y : \mathbb{C} \to \mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ is full and faithful, it reflects all the above properties. All functors preserve the first and third properties, while any limit preserving functor (e.g. Y') preserves the rest (recall that being a monomorphism is a limit property: a morphism m is a monomorphism if and only if its pullback along itself is an isomorphism). Thus to prove $\mathcal{P}$ it is sufficient to prove it for $\mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$. All the properties above hold in $\mathbf{Set}^{\mathbb{C}^{\mathsf{op}}}$ precisely when they hold point-wise in **Set**. Indeed, the first and third are obvious, the second is contained in Lemma 4 and the fourth is contained in Lemma 2, noting that **Set** is complete. As for last property, let F, G and M be functors from $\mathbb{C}$ to **Set**, let $\alpha : F \Rightarrow G$ be a natural transformation and let $\beta : M \Rightarrow G$ be a monomorphic natural transformation such that for every component α_C of α, $\alpha_C = \beta_C \circ e_C$ for some morphism e_C, i.e. α_C factors through β_C. It remains to show the naturality of $(e_C)_{C\in\mathbb{C}}$. Let $f : C \to D$ be a morphism in $\mathbb{C}$. Then we have:

$$\beta_D \circ M(f) \circ e_C = G(f) \circ \beta_C \circ e_C = \alpha_D \circ F(f) = \beta_D \circ e_D \circ F(f),$$

so $M(f) \circ e_C = e_D \circ F(f)$ since β_D is a monomorphism. Thus it is sufficient to prove that $\mathcal{P}$ always holds in **Set**. □

2.1.3 Adjunctions

There are a number of equivalent definitions of an adjunction. We present one particular one here, and then give a proposition (whose proof can be found in [75] for example) from which we may obtain a number of equivalent definitions.

Definition 1. *An* adjunction *from a category* $\mathbb{C}$ *to a category* $\mathbb{D}$ *is a quadruple* $(F, G, \eta, \varepsilon)$*, where*

- $F : \mathbb{C} \to \mathbb{D}$*,* $G : \mathbb{D} \to \mathbb{C}$ *are functors, and*
- $\eta : 1_{\mathbb{C}} \to GF$*,* $\varepsilon : FG \to 1_{\mathbb{D}}$ *are natural transformations,*

such that the following identities (called the triangle identities*) are satisfied:*

$$\varepsilon F \circ F\eta = 1_F$$
$$G\varepsilon \circ \eta G = 1_G.$$

We write $F \dashv G$*, and say that* G *is* right adjoint *to* F *and* F *is* left adjoint *to* G*. We call* η *and* ε *the* unit *and* counit *of the adjunction respectively.*

Proposition 2. *Let* $F : \mathbb{C} \to \mathbb{D}$ *and* $G : \mathbb{D} \to \mathbb{C}$ *be two functors. Then there is a bijection between:*

(1) adjunctions $(F, G, \eta, \varepsilon)$*;*

(2) families $(\phi_{C,D})_{C \in \mathbb{C}, D \in \mathbb{D}}$ *where for every pair of objects* C *in* $\mathbb{C}$ *and* D *in* $\mathbb{D}$*,* $\phi_{C,D}$ *is a bijection of hom-sets:*

$$\phi_{C,D} : \mathbb{C}(C, G(D)) \cong \mathbb{D}(F(C), D)$$

which is natural in C *and* D*;*

(3) natural transformations $\eta : 1_{\mathbb{C}} \to GF$ *such that every component* η_C *is universal to* G *from* C*, i.e. for any morphism* $f : C \to G(D)$*, there is a unique morphism* $d : F(C) \to D$ *in* $\mathbb{D}$ *such that* $G(d) \circ \eta_C = f$*;*

(4) natural transformations $\varepsilon : FG \to 1_{\mathbb{C}}$ *such that every component* ε_D *is universal to* D *from* F*, i.e. for any morphism* $g : F(C) \to D$*, there is a unique morphism* $c : C \to G(D)$ *in* $\mathbb{C}$ *such that* $\varepsilon_D \circ F(c) = g$*.*

When we say that the bijection of hom-sets $\mathbb{C}(C, G(D)) \cong \mathbb{D}(F(C), D)$ is *natural in* C, we mean that for any morphism $f : C \to C'$ in $\mathbb{C}$, the following diagram (of sets) commutes, while "natural in D" is defined similary.

$$\begin{tikzcd}
\mathbb{C}(C, G(D)) \arrow[r, "\phi_{C,D}"] & \mathbb{D}(F(C), D) \\
\mathbb{C}(C', G(D)) \arrow[r, "\phi_{C',D}"] \arrow[u, "\mathbb{C}(f,G(D))"'] & \mathbb{D}(F(C), D) \arrow[u, "\mathbb{D}(F(f),D)"']
\end{tikzcd}$$

We briefly mention how the bijection in Proposition 2 is established. To any adjunction in (1), we simply assign the unit and counit in (3) and (4). In (2), we define, for any pair of objects C in $\mathbb{C}$ and D in $\mathbb{D}$, a bijection $\phi_{C,D}$ as follows:

$$\phi_{C,D} : (f : C \to G(D)) \mapsto (\varepsilon_D \circ F(f) : F(C) \to D).$$

Note that if ε is the identity natural transformation, then this is just the action of F on morphisms. We now recall some important properties of adjunctions.

Proposition 3. *Any two left adjoints of a functor are naturally isomorphic. Dually, so are any two right adjoints.*

Proposition 4. *A functor $F : \mathbb{C} \to \mathbb{D}$ which has a left adjoint preserves all limits which exist in $\mathbb{C}$. Dually, if F has a right adjoint, then it preserves all colimits that exist in $\mathbb{C}$.*

It will be important for our purposes to examine the case where the categories $\mathbb{C}$ and $\mathbb{D}$ are (partially) ordered sets (also called posets). An adjunction between two such categories is nothing but a (covariant) *Galois connection*. Recall that a poset is a set X together with a reflexive, transitive and anti-symmetric relation $\leq$.

Definition 2. *A* Galois connection *from a poset $(X, \leq)$ to a poset $(Y, \leq)$ is a pair of (set-theoretic) maps $f : X \to Y$ and $g : Y \to X$ such that for any two elements $x \in X$ and $y \in Y$,*

$$x \leq g(y) \Leftrightarrow f(x) \leq y.$$

In particular, the maps f and g will be order-preserving (also called *monotone*). Applying Proposition 2 and noting that in a poset any diagram commutes (in particular, naturality squares and the component-wise versions of the triangle identities always commute), we obtain that two monotone maps $f : X \to Y$ and $g : Y \to X$ give a Galois connection if and only if for any two elements $x \in X$ and $y \in Y$, $x \leq gf(x)$ and $y \geq fg(y)$.

Finally, we deal with the notion of *equivalence* of categories.

Definition 3. *A functor $F : \mathbb{C} \to \mathbb{D}$ is an equivalence of categories if there is a functor $G : \mathbb{D} \to \mathbb{C}$ and natural isomorphisms $GF \cong 1_{\mathbb{C}}$ and $FG \cong 1_{\mathbb{D}}$.*

An *adjoint equivalence* is an adjunction $(F, G, \eta, \varepsilon)$ such that both η and ε are natural isomorphisms. Notice then that $(G, F, \varepsilon^{-1}, \eta^{-1})$ is also an adjunction, so G is both a right and left adjoint of F. The following important theorem about equivalences is taken from [75].

Theorem 1. *Let $F : \mathbb{C} \to \mathbb{D}$ be a functor. Then the following are equivalent:*

- *F is an equivalence;*
- *F is part of an adjoint equivalence $(F, G, \eta, \varepsilon)$;*
- *F is full, faithful and essentially surjective.*

2.1.4 Fibrations

We now reach the central concept of the thesis: that of a Grothendieck fibration. Fibrations were first introduced by Grothendieck in [45] as a means to develop descent theory. In this section, we give the definition of a fibration, together with a technique for constructing fibrations (and their duals, opfibrations) from functors into the category of categories **Cat**.

Definition 4. *Let $F : \mathbb{B} \to \mathbb{C}$ be a functor. A morphism $a : X \to Y$ in $\mathbb{B}$ is* cartesian with respect to F *or* F-cartesian *if for every pair of diagrams of solid arrows below, with the left diagram in $\mathbb{B}$ and the right diagram in $\mathbb{C}$ and where the diagram on the right commutes, there is a unique morphism c, shown by the dotted arrow, such that $F(c) = h$ and the diagram on the left commutes.*

$$
\begin{tikzcd}
Z \arrow[d, dotted, "c"'] \arrow[dr, "b"] & \\
X \arrow[r, "a"'] & Y
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
F(Z) \arrow[d, "h"'] \arrow[dr, "F(b)"] & \\
F(X) \arrow[r, "F(a)"'] & F(Y)
\end{tikzcd}
$$

We recall the following, easy to prove, properties of cartesian morphisms:

Lemma 5. *Let $F : \mathbb{B} \to \mathbb{C}$ be any functor. Then:*

- *F-cartesian morphisms are closed under composition;*
- *if $a : X \to Y$ and $b : Z \to Y$ are two F-cartesian morphisms and $F(a) = F(b)$, then there is a unique isomorphism $\theta : X \to Z$ such that $F(\theta) = 1_{F(X)}$ and $b \circ \theta = a$.*

Definition 5. *A functor $F : \mathbb{B} \to \mathbb{C}$ is a* fibration *if for every object X in $\mathbb{B}$ and every morphism $f : C \to F(X)$, there exists an F-cartesian morphism $b : Y \to X$ such that $F(b) = f$. Such a morphism b will be called the* cartesian lifting of f to X.

Note that by Lemma 5, the morphism b in the definition above will be unique up to isomorphism when it exists.

Definition 6. *Let $F : \mathbb{B} \to \mathbb{C}$ and $F' : \mathbb{B}' \to \mathbb{C}$ be two fibrations. A* cartesian functor *from F to F' is a functor $H : \mathbb{B} \to \mathbb{B}'$ such that $F' \circ H = F$ and H maps F-cartesian morphisms to F'-cartesian morphisms.*

Definition 7. *For a functor $F : \mathbb{B} \to \mathbb{C}$ and an object C of $\mathbb{C}$, the* fibre of F at C *(denoted by $F^{-1}(C)$) is the subcategory of $\mathbb{B}$ whose objects are all objects X of $\mathbb{B}$ such that $F(X) = C$ and whose morphisms are those morphisms b in $\mathbb{B}$ such that $F(b) = 1_C$.*

Given a fibration $F : \mathbb{B} \to \mathbb{C}$, every morphism $f : C \to C'$ in $\mathbb{C}$ gives rise to a functor $f^* : F^{-1}(C') \to F^{-1}(C)$ which sends every object X in $F^{-1}(C')$ to the domain of a cartesian lifting of f to X. For a morphism $a : X \to X'$ in $F^{-1}(C')$, $f^*(a)$ is defined to be the unique morphism $a' : f^*(X) \to f^*(X')$ such that $F(a') = 1_C$ and which commutes with the chosen cartesian liftings of f to X and X', as given by the cartesianness property. The uniqueness of the choice of a' ensures that f^* is indeed a functor.

We now describe how to construct a fibration over a category $\mathbb{C}$ from a functor $\phi : \mathbb{C} \to \mathbf{Cat}$, where $\mathbf{Cat}$ is the category of all categories. In fact, this construction applies more generally to pseudo-functors. The notion of a pseudo-functor is weaker than that of a functor, and represents a certain kind of morphism between 2-categories. Roughly speaking, 2-categories have, in addition to objects and morphisms, 2-cells between morphisms. The category of all categories, with natural transformations as 2-cells, is a classical example of a 2-category. It is beyond the scope of this section to describe the general theory of 2-categories, and it is not needed for the rest of the thesis. However, we will define a pseudo-functor from a category $\mathbb{C}$ (i.e. a 2-category whose 2-cells are all trivial) into $\mathbf{Cat}$ in elementary terms so as to describe their relation to fibrations. For a general definition, and more on 2-categories, the Reader may consult [9].

Definition 8. *A* pseudo-functor *ϕ from a category $\mathbb{C}$ to $\mathbf{Cat}$ consists of the following data:*

- *for every object C in $\mathbb{C}$, an object $\phi(C)$ of $\mathbf{Cat}$;*
- *for every morphism $f : A \to B$, a morphism $\phi(f) : \phi(A) \to \phi(B)$;*
- *for every pair of composable morphisms $f : A \to B$, $g : B \to C$, a natural isomorphism $\gamma_{f,g} : \phi(g) \circ \phi(f) \cong \phi(g \circ f)$;*

- *for every object C of $\mathbb{C}$, a natural isomorphism $\delta_C : 1_{\phi(C)} \cong \phi(1_C)$.*

satisfying the following conditions:

- *for every triple of arrows $f : A \to B$, $g : B \to C$, $h : C \to D$,*

  $$\gamma_{gf,h} \circ (\phi(h)\gamma_{f,g}) = \gamma_{f,hg} \circ (\gamma_{g,h}\phi(f))$$

  (where $\phi(h)$ and $\phi(f)$ are functors, and hence compose with natural transformations as defined in Section 2.1.3.)

- *for every morphism $f : A \to B$ in $\mathbb{C}$,*

  $$\gamma_{1_A,f} \circ (\phi(f)\delta_A) = 1_{\phi(f)} = \gamma_{f,1_B} \circ (\delta_B\phi(f))$$

We now describe the Grothendieck construction, which constructs a fibration over a category $\mathbb{C}$ from a contravariant pseudofunctor from $\mathbb{C}$ into **Cat**. Let $\phi : \mathbb{C}^{\mathsf{op}} \to \mathbf{Cat}$ be a pseudofunctor. Let $\int \phi$ be the category defined as follows:

- an object of $\int \phi$ is a pair (A, X), where A is an object of $\mathbb{C}$ and X is an object of $\phi(A)$;
- a morphism $(A, X) \to (B, Y)$ is a pair (f, a) where $f : A \to B$ is an arrow of $\mathbb{C}$ and $a : X \to \phi(f)(Y)$ is an arrow of $\phi(A)$ (note that $\phi(f)$ is a functor from $\phi(B)$ to $\phi(A)$);
- for two morphisms $(f, a) : (A, X) \to (B, Y)$ and $(g, b) : (B, Y) \to (C, Z)$, the composite is defined as follows:

  $$(g, b) \circ (f, a) = (g \circ f, \gamma \circ \phi(f)(b) \circ a)$$

  where the second component is the composite

  $$X \xrightarrow{a} \phi(f)(Y) \xrightarrow{\phi(f)(b)} \phi(f)(\phi(g)(Z)) \xrightarrow{\gamma} \phi(g \circ f)(Z),$$

  where γ comes from the definition of the pseudofunctor ϕ.

That $\int \phi$ is a category (i.e. that composition is associative and that it has identity morphisms) follows straightforwardly from the axioms of a pseudofunctor. In particular, it is easy to check for the case when ϕ is actually a functor, i.e. when γ is always the identity. Let $F : \int \phi \to \mathbb{C}$ be the first component functor, i.e. $F(A, X) = A$ and $F(f, a) = a$.

Proposition 5. *The functor F defined above is a fibration.*

Proof. Let (B, Y) be a object $\mathbb{B}$ and $f : A \to B = F(B, Y)$ a morphism in $\mathbb{C}$. We claim that the morphism $f' = (f, 1_{\phi(f)(Y)}) : (A, \phi(f)(Y)) \to (B, Y)$ is cartesian. Let $(g, b) : (C, Z) \to (B, Y)$ be a morphism in $\mathbb{B}$, and let $h : C \to A$ be a morphism in $\mathbb{C}$ such that $f \circ h = g$. We must show the existence of a unique morphism $(h, c) : (C, Z) \to (A, \phi(f)(Y))$ in $\mathbb{B}$ such that $f' \circ (h, c) = (g, b)$. In other words, we require a morphism c from Z to $\phi(h)(\phi(f)(Y))$ such that

$$\gamma \circ \phi(h)(1_{\phi(f)(Y)}) \circ c = b,$$

where γ is the isomorphism $\phi(h)(\phi(f)(Y)) \cong \phi(g)(Y)$. Since $\phi(h)$ is a functor, this is equivalent to $\gamma \circ c = b$. Clearly $c = \gamma^{-1} \circ b$ gives the unique such morphism. □

In fact, every fibration arises (up to isomorphism) from a contravariant pseudofunctor in this way. To show this, we show how to construct a pseudofunctor ϕ from a fibration F. This requires the axiom of choice. Let $F : \mathbb{B} \to \mathbb{C}$ be a fibration. Define a pseudofunctor $\phi : \mathbb{C}^{\mathsf{op}} \to \mathbf{Cat}$ as follows (we will treat all objects and morphisms in $\mathbb{C}^{\mathsf{op}}$ as objects and morphisms of $\mathbb{C}$):

- for an object C in $\mathbb{C}$, $\phi(C)$ is the fibre of F at C;
- for a morphism $f : A \to B$ in $\mathbb{C}$, we must define a functor $\phi(f) : \phi(B) \to \phi(A)$. This functor will be nothing but (one choice of) the functor f^* defined earlier;
- for a pair of morphisms $f : A \to B$, $g : B \to C$ in $\mathbb{C}$ and an object Z in $\phi(C)$, the cartesianness condition implies a unique isomorphism $\phi(f)(\phi(g)(Z)) \cong \phi(g \circ f)(Z)$ in the fibre of A which commutes with the chosen liftings. The family of these isomorphisms indexed by Z will be the natural isomorphism $\gamma_{g,f}$.
- for an object C of $\mathbb{C}$ and an object X in $\phi(C)$, there is also a unique isomorphism $X \cong \phi(1_C)(X)$. The family of these isomorphisms indexed by X will be the natural transformation δ_C.

It is an easy application of Lemma 5 that this defines a pseudofunctor. Moreover, we will show that applying the Grothendieck construction to this pseudofunctor recovers F up to isomorphism. We will say that two fibrations F and G with the same codomain are isomorphic if there is an isomorphism I such that $G \circ I = F$ (in particular, every such isomorphism is a cartesian functor).

Proposition 6. *Let $F : \mathbb{B} \to \mathbb{C}$ be a fibration. Let ϕ be the corresponding pseudofunctor under the construction above and let $F' : \int \phi \to \mathbb{C}$ be the functor obtained from ϕ via the Grothendieck construction. Then F is isomorphic to F'.*

Proof. We define a functor $I : \mathbb{B} \to \int \phi$ as follows:

- for an object X in $\mathbb{B}$, $I(X)$ is defined to be $(F(X), X)$;
- for a morphism $b : X \to Y$, $I(b)$ is defined to be $(F(b), b') : (F(X), X) \to (F(Y), Y)$, where b' is the unique morphism from X to $b^*(Y)$ in the fibre of $F(B)$ which when composed with the cartesian lifting gives b.

The uniqueness in the definition of b' guarantees that I is a functor. We also define a functor $H : \int \phi \to \mathbb{B}$ as follows:

- for an object (A, X) in $\int \phi$, $H(A, X)$ is defined to be simply X;
- for a morphism $(f, a) : (A, X) \to (B, Y)$, $H(f, a)$ is defined to be the composite of $a : X \to f^*(Y)$ with the cartesian morphism $f' : f^*(Y) \to Y$.

We claim that this defines a functor. Let $(f, a) : (A, X) \to (B, Y)$ and $(g, b) : (B, Y) \to (C, Z)$ be two morphisms in $\int \phi$. We have the isomorphism $(g \circ f)^*(Z) \cong f^*(g^*(Z))$. By examining the definition of ϕ, we see that this isomorphism is precisely the inverse of the isomorphism γ occuring in the definition of the composite $(g, b) \circ (f, a)$, ensuring the functoriality of H. Finally, it is easy to check that $H \circ I$ and $I \circ H$ are both identity functors, and that $F' \circ I = F$, which gives the required result. □

We may ask about the other direction: what happens if we apply the Grothendieck construction to a pseudofunctor ϕ and then construct a pseudofunctor ϕ' from the resulting fibration? Since there is some choice to be made in constructing the functor ϕ', we cannot expect that the functors will be equal. However, there is always a so-called pseudo-natural isomorphism between them. In the language of 2-categories, the 2-categories $\mathsf{PsFun}(\mathbb{C}, \mathbf{Cat})$ (objects are pseudofunctors, morphisms are pseudo-natural transformations and 2-cells are so-called modifications) and $\mathsf{Fib}(\mathbb{C})$ (objects are fibrations, morphisms are cartesian functors and 2-cells are so-called cartesian natural transformations) are 2-equivalent. For a full definition and a detailed proof, the Reader is referred to Section 8.3 in [10].

Of course the Grothendieck construction applies in particular to actual functors $\mathbb{C} \to \mathbf{Cat}$. The resulting fibrations have the special property that we can choose cartesian liftings in such a way that they form a subcategory.

Definition 9. *A fibration $F : \mathbb{B} \to \mathbb{C}$ is called* split *when for every morphism $f : A \to B$ and every object Y in the fibre of B, we can choose a cartesian lifting f_Y such that the family of all such morphisms f_Y (indexing over morphisms f and objects Y) form a subcategory of $\mathbb{B}$.*

While every fibration is equivalent to a split fibration (see [10]), not every fibration is itself split. To show this, we will first describe how groups can be viewed as categories. Given a group G, we can view it as a one object category $\mathbb{G}$ whose morphisms are elements of G and where composition is given by multiplication in G. In particular, every morphism will be an isomorphism. A category where every morphism is an isomorphism is called a *groupoid*. Groups are thus in bijection with one-object groupoids. Group homomorphisms are nothing but functors between the corresponding groupoids.

Lemma 6. *Any functor $F : \mathbb{G} \to \mathbb{H}$ where F is surjective on objects and full, and where $\mathbb{G}$ and $\mathbb{H}$ are groupoids, is a fibration.*

Proof. Let $f : A \to B$ be a morphism in $\mathbb{H}$ and Y be an object in the fibre of B. Let $f' : X \to Y$ be any morphism such that $F(f') = f$ (at least one such exists by the conditions on F). We claim that f is cartesian. Let $g : C \to B$ and $h : C \to A$ be morphisms in $\mathbb{H}$ such that $f \circ h = v$ and let $b : Z \to X$ be a morphism in $\mathbb{G}$ such that $F(b) = g$. Then let $h' = f'^{-1} \circ b$. Clearly $f'h' = b$, while $F(h') = f^{-1} \circ g = h$. The morphism h' is the unique choice since f' is a monomorphism. □

Proposition 7. *Let $\mathbb{Z}$ and $\mathbb{Z}_2$ be the group of integers and the two element cyclic group respectively, both thought of as one object groupoids. The functor (i.e. group homomorphism) $F : \mathbb{Z} \to \mathbb{Z}_2$ which sends a morphism n in $\mathbb{Z}$ to $n \bmod 2$ is a fibration, but it is not split.*

Proof. The fact that F is a fibration follows from Lemma 6. Suppose F admits a splitting. Such a splitting is a choice of a single cartesian morphism in $\mathbb{Z}$ for every morphism in $\mathbb{Z}_2$ (since $\mathbb{Z}$ has only one object). These cartesian morphisms must form a subcategory, i.e. a subgroup of $\mathbb{Z}$, whose image under F must include all the morphisms of $\mathbb{Z}_2$, so the choice function must be injective. Now it remains to note that there is no subgroup of $\mathbb{Z}$ with exactly two distinct elements. □

We now mention a generalisation of the notion of fibration, due to Street [84]. One of the most important features which distinguishes Street fibrations from Grothendieck fibrations is that all equivalences are Street fibrations, but not all equivalences are Grothendieck fibrations.

Definition 10. *A functor $F : \mathbb{B} \to \mathbb{C}$ is a* Street fibration *if for every object $Y \in \mathbb{B}$ and every morphism $f : A \to F(Y)$ in $\mathbb{C}$, there is a cartesian morphism $f' : X \to Y$ and an isomorphism $h : F(X) \cong A$ such that $f \circ h = F(f')$.*

We now consider the dual notion to that of a fibration, namely the notion of an *opfibration*. A functor F is an opfibration precisely when F^{op} is a fibration. Since we will often be working with both fibrations and opfibrations, we state the definition here in full.

Definition 11. *Let $F : \mathbb{B} \to \mathbb{C}$ be a functor. A morphism $a : X \to Y$ in $\mathbb{B}$ is* cocartesian with respect to F *or* F-cocartesian *if for every pair of diagrams of solid arrows below, with the left diagram in $\mathbb{B}$ and the right diagram in $\mathbb{C}$ and where the diagram on the right commutes, there is a unique morphism c, shown by the dotted arrow, such that $F(c) = h$ and the diagram on the left commutes.*

$$
\begin{tikzcd}
 & Z \\
X \arrow[ur, "b"] \arrow[r, "a"'] & Y \arrow[u, dotted, "c"']
\end{tikzcd}
\qquad
\begin{tikzcd}
 & F(Z) \\
F(X) \arrow[ur, "F(b)"] \arrow[r, "F(a)"'] & F(Y) \arrow[u, "h"']
\end{tikzcd}
$$

Definition 12. *A functor $F : \mathbb{B} \to \mathbb{C}$ is an* opfibration *if for every object X in $\mathbb{B}$ and every morphism $f : F(X) \to B$, there exists an F-cocartesian morphism $f' : X \to Y$ such that $F(f') = f$. Such a morphism f' will be called a* cocartesian lifting of f from B.

Definition 13. *A functor which is both a fibration and an opfibration is called a* bifibration.

Given an opfibration $F : \mathbb{B} \to \mathbb{C}$, we can define for every morphism $f : A \to B$ in $\mathbb{C}$ a functor $f_* : F^{-1}(A) \to F^{-1}(B)$ which is the dual of the functor f^* corresponding to the fibration F^{op}. In particular, f_* sends an object X in $F^{-1}(A)$ to the codomain of a cocartesian lifting of f from A. The following lemma is easy to prove:

Lemma 7. *Let F be a functor and f be an F-cartesian morphism. Then if $f\alpha = f\beta$ and $F(\alpha) = F(\beta)$, then $\alpha = \beta$.*

Proposition 8. *A fibration $F : \mathbb{B} \to \mathbb{C}$ is a bifibration if and only if for every morphism f in $\mathbb{C}$, the functor f^* has a left adjoint, denoted by f_*.*

Proof. For the 'if' part, let $f : A \to B$ be a morphism in $\mathbb{C}$ and let X be an object in the fibre of A. Define f' to be the composite of the component η_X of the counit of $f_* \dashv f^*$ and the cartesian morphism $\hat{f} : f^*(f_*(X)) \to f_*(X)$. We claim that f' is F-cocartesian. Clearly $F(f') = f$, so let $u : A \to C$ and $v : B \to C$ be morphisms in $\mathbb{C}$ with $vf = u$ and let $u' : X \to Z$ be a morphism in $\mathbb{B}$ with $F(u') = u$. Consider the following diagram of solid arrows, where f''' is the unique morphism such that $F(f''') = f$ and $v^Z \circ f''' = u'$, where v^Z is the cartesian lifting of v to Z, and f'' is the cartesian lifting of f to $v^*(Z)$.

$$
\begin{tikzcd}
 & & Z \\
f^*(v^*(Z)) \arrow[r, "f''"] & v^*(Z) \arrow[ur, "v^Z"] & \\
X \arrow[u, dotted, "w"] \arrow[ur, "f'''"] \arrow[r, "f'"'] & f_*(X) \arrow[u, dotted, "v''"] \arrow[uur, dotted, "v'"'] &
\end{tikzcd}
$$

We establish some bijections (in each case see the dotted arrows):

- there is a bijection between morphisms v' such that $F(v') = v$ and $v' \circ f' = u$, and morphisms v'' in the fibre at B such that $v_Z \circ v'' \circ f' = u$ (cartesianness of v^Z);

- a morphism v'' in the fibre of B satisfies $v_Z \circ v'' \circ f' = u'$ if and only if it satisfies $v'' \circ f' = f'''$ (Lemma 7);

- there is a bijection between morphisms v'' in the fibre of B such that $v'' \circ f' = f'''$, and morphisms w in the fibre of A such that $f'' \circ w = f'''$ – for this, we use the bijection $\mathsf{hom}(f_*(X), v^*(Z)) \cong \mathsf{hom}(X, f^*(v^*(Z)))$ from the adjunction $f_* \dashv f^*$ and note that for any bijective pair (v'', w), the rectangle always commutes;

- a morphism w in the fibre of A satisfies $f'' \circ w = f'''$ if and only if it satisfies $v_Z \circ f'' \circ w = u'$ (Lemma 7).

Thus it is enough to note that since $v^Z \circ f''$ is cartesian, there is precisely one such w.

For the 'only if' part, let F be a bifibration and let $f : A \to B$ be a morphism of $\mathbb{C}$. We claim that $f_* \dashv f^*$, where f_* is defined as above. Let X be an object of the fibre of A. Then there is the morphism $f' : X \to f_*(X)$ and the cartesian morphism $f'' : f^*(f_*(X)) \to f_*(X)$, which gives a unique morphism $\eta_X : X \to f^*(f_*(X))$ such that $F(\eta_X) = 1_A$ and $f''\eta_X = f'$. Dually, for an object Y in the fibre of B, we obtain a morphism $\varepsilon_Y : f_*(f^*(Y)) \to Y$. Naturality of the family $(\eta_X)_{X \in F^{-1}(A)}$ follows from Lemma 7 and cartesianness of the morphism $f^*(f_*(X')) \to X'$. Dually, $(\varepsilon_Y)_{Y \in F^{-1}(B)}$ also gives a natural transformation. Finally, the fact that η and ε satisfy the triangle identities follows again from Lemma 7 and cartesianness and cocartesianness of the maps $f^*(Y) \to Y$ and $X \to f_*(X)$ respectively. □

We can thus conclude that every bifibration over $\mathbb{C}$ comes from a pseudo-functor from $\mathbb{C}^{\mathsf{op}}$ to **Cat** which takes every morphism in $\mathbb{C}$ to a functor which has a left adjoint. We state two more properties of fibrations without proof:

Lemma 8. *Fibrations are closed under composition.*

Lemma 9. *Fibrations are stable under pullbacks in* **Cat**.

Before we conclude this section, we should make some elementary observations on the case when F is faithful and amnestic, which will be the main context considered in the thesis. A functor $F : \mathbb{B} \to \mathbb{C}$ is *amnestic* when there are no non-identity isomorphisms in each fibre. The following proposition follows easily from the material in this section.

Proposition 9. *Let $F : \mathbb{B} \to \mathbb{C}$ be a faithful, amnestic fibration. Then:*

- *every fibre is a poset;*
- *the functors f^* are just order-preserving maps;*
- *F is split – moreover, cartesian liftings are unique;*
- *F is a bifibration if and only if every f^* is the right adjoint in a Galois connection.*

Moreover, such fibrations are in bijection with functors from $\mathbb{C}^{\mathsf{op}}$ to the category **Ord** *whose objects are posets and whose morphisms are order-preserving maps.*

Finally, we conclude this section by remarking that fibrations are ubiquitous in mathematics. They play an important role in many areas such as foundations [6], logic [48], algebraic geometry [45], topology [89], topos theory [69] and non-abelian algebra [14] (the references given are examples which readily show the importance of fibrations). We also hope to demonstrate the importance of (bi)fibrations in algebra through this thesis.

2.1.5 Comma categories

In this section, we give the definition of a comma category and relate it to the notion of fibration.

Definition 14. *Let $F : \mathbb{A} \to \mathbb{C}$ and $G : \mathbb{B} \to \mathbb{C}$ be functors. Then the* comma category *$(F \downarrow G)$ is the category defined as follows:*

- *an object of $(F \downarrow G)$ is a triple (A, f, B), with A an object of $\mathbb{A}$, B an object of $\mathbb{B}$ and $f : F(A) \to G(B)$ a morphism in $\mathbb{C}$;*
- *a morphism of $(F \downarrow G)$ from (A, f, B) to (A', f', B') is a pair (a, b) where $a : A \to A'$ is a morphism in $\mathbb{A}$ and $b : B \to B'$ is a morphism in $\mathbb{B}$ such that $G(b) \circ f = f' \circ F(a)$;*
- *composition is defined component-wise as in $\mathbb{A}$ and $\mathbb{B}$.*

Proposition 10. *Given the comma category $(F \downarrow G)$ of two functors $F : \mathbb{A} \to \mathbb{C}$ and $G : \mathbb{B} \to \mathbb{C}$, there are two obvious projection functors $U : (F \downarrow G) \to \mathbb{A}$ and $V : (F \downarrow G) \to \mathbb{B}$ and a canonical natural transformation $\alpha : F \circ U \Rightarrow G \circ V$.*

Proof. Let U be the functor $(A, f, B) \mapsto A$, $(a, b) \mapsto a$ and V be the functor $(A, f, B) \mapsto B$, $(a, b) \mapsto b$. Then α is the natural transformation whose component at (A, f, B) is simply f. □

Proposition 11. *With the notation of the previous proposition, $(F \downarrow G)$ together with U, V and α is the universal such structure, i.e. for any pair of functors $U' : \mathbb{D} \to \mathbb{A}$ and $V' : \mathbb{D} \to \mathbb{B}$ and natural transformation $\beta : F \circ U' \Rightarrow G \circ V'$, there is a unique functor $W : \mathbb{D} \to (F \downarrow G)$ such that $U \circ W = U'$, $V \circ W = V'$ and $\alpha W = \beta$.*

Proof. For an object D in $\mathbb{D}$ and a morphism $d : D \to D'$ in $\mathbb{D}$, define $W(D) = (U'(D), \beta_D, V'(D))$ and $W(d) = (U'(d), V'(d))$. It is clearly the unique functor satisfying the conditions. □

The above proposition can be taken as a definition of a *comma object* in any 2-category, since such a structure will be unique up to isomorphism. As usual, we refer to *the* comma category, with the understanding that this only specifies an object up to isomorphism. In the case of categories, there are two important cases of comma categories to consider, which each have special terminology in the literature.

- If $F = G = 1_{\mathbb{C}}$, then $(F \downarrow G)$ is the *arrow category* $\mathsf{Arr}(\mathbb{C})$. It is also isomorphic to $\mathbb{C}^{\mathbf{2}}$, where $\mathbf{2}$ is the category with two objects A and B and one non-identity morphism $A \to B$.

- If G is the inclusion of a single object C and the morphism 1_C into $\mathbb{C}$ and F is the identity on $\mathbb{C}$, then $(F \downarrow G)$ is the *slice category* $\mathbb{C}/C$. Dually, if F is the inclusion and G is the identity, then $(F \downarrow G)$ is the *coslice category* $C/\mathbb{C}$.

We now give a proposition due to R. Street [84] which links the two major topics of this section, namely adjunctions and fibrations, to comma categories. The result and the lemma below require the axiom of choice. Recall the following lemma from [9]:

Lemma 10. *A functor $F : \mathbb{B} \to \mathbb{C}$ has a right adjoint if and only if each object C in $\mathbb{C}$ admits a coreflection along F, i.e. an object $R(C)$ in $\mathbb{B}$ and a morphism $\varepsilon_C : FR(C) \to C$ which is universal to C from F.*

Proposition 12. *Let $F : \mathbb{B} \to \mathbb{C}$ be functor and let $\mathbb{A} = (1_{\mathbb{C}} \downarrow F)$. Consider the diagram below:*

$$
\begin{tikzcd}
\mathbb{B} \arrow[dr, dotted, "I"] \arrow[drr, bend left, equal] \arrow[ddr, bend right, "F"'] & & \\
& \mathbb{A} \arrow[r, "Q"'] \arrow[d, "P"'] & \mathbb{B} \arrow[d, "F"'] \\
& \mathbb{C} \arrow[r, equal] & \mathbb{C}
\end{tikzcd}
$$

where I is the canonical functor to the comma category $\mathbb{A}$. Then F is a fibration if and only if I has a right adjoint R such that $P\varepsilon = 1_P$, where ε is the counit of $F \dashv R$ (in partcular then, $I \circ R = P$).

Proof. The category $\mathbb{A}$ has objects of the form $(C, f : C \to F(B), B)$ and the functor I takes an object B to $(F(B), 1_{F(B)}, B)$ and a morphism $b : B \to B'$ to $(F(b), b)$. Let $A = (C, f, B)$ be an object of $\mathbb{A}$. To say that A has a coreflection along I is to say that there exists an object B' of $\mathbb{B}$ and morphism $(i, f') : (F(B), 1_{F(B)}, B') \to (C, f, B)$ which is universal to A from I. The universal property of (i, f') can be stated as follows: for any object B'' of $\mathbb{B}$ and any morphisms $u : F(B'') \to C$ and $v : B'' \to B$ such that $F(v) = f \circ u$, there is a unique morphism $b : B'' \to B'$ such that $i \circ F(b) = u$ and $f' \circ b = v$. If we require that i is the identity and that $F(f') = f$, then this is equivalent to saying that f' is a cartesian lifting of f to B. Suppose then that every object A has a coreflection and consider the resulting right adjoint to F. In the above notation, for i to be the identity and to have $F(f') = f$ for every A is equivalent to requiring that $P\varepsilon = 1_P$ and $F \circ R = P$ respectively. Thus for every object B in $\mathbb{B}$ and morphism $f : C \to F(B)$ in $\mathbb{C}$, there is a cocartesian lifting from f to B if and only if the functor I has a right adjoint satisfying the conditions. □

2.2 Non-abelian categorical algebra

2.2.1 Duality for groups hypothesis

In his 1950 paper *Duality for Groups* [73], Mac Lane proposes the axiomatic study of the category of groups (i.e. the category whose objects are groups and whose morphisms are group homomorphisms). He proposes the formulation of a list of dual axioms on a category which allow one to prove isomorphism theorems and other results from group theory. He then gives a list of dual axioms, suitable only for the category of abelian groups, and arrives at the notion of an *abelian bicategory*, the forerunner of the notion of an abelian category (first introduced in [20] under the name of *exact category*). Abelian categories were extensively used in the first two decades of category theory, for example in Grothendieck's famous Tohoku paper [44] (where the name of abelian category was first used) and in subsequent work on homology and algebraic geometry.

The category of all groups, however, does not satisfy Mac Lane's axioms for an abelian bicategory. For example, axiom ABC-3 in [73] requires that every subobject be "normal" (as defined in the paper), which translates as expected into the condition that every subgroup is normal, a condition which does not hold for groups in general, but does hold for all abelian groups. We may thus view Mac Lane's statement that, for the case of all (not necessarily abelian) groups,

> "a further development giving the first and second isomorphism theorems, and so on, can be made by introducing additional carefully chosen dual axioms"

as a hypothesis, which we will refer to as the *duality for groups hypothesis*. A crucial difference between the category of groups and the category of abelian groups (which we denote by **Grp** and **Ab** throughout this thesis) is that the most important properties of the latter are self-dual, while duals of similar properties of the former almost never hold. Indeed, an abelian category can be defined using self-dual axioms. Before we state these axioms, we will need the following definitions.

Definition 15. *A category $\mathbb{C}$ is called* pointed *when it has an object Z (called a* zero object*) which is both terminal and initial. Morphisms which factor through a zero object are called* zero morphisms. *We write $f = 0$, or say that f is zero, to indicate that f is a zero morphism.*

Definition 16. *Let $\mathbb{C}$ be a pointed category. For a morphism $f : X \to Y$, the* kernel *of f is a morphism $k : K \to X$ such that fk is zero and such that for any other morphism k' with fk' zero, there is a unique morphism i such that $ki = k'$. Dually, the* cokernel *of f is a morphism $c : Y \to C$ such that cf is zero and such that for any other morphism c' with $c'f$ zero, there is a unique morphism i such that $ci = c'$. Both these morphisms are unique up to isomorphism, and we denote the kernel of f by $\ker(f) : \mathsf{Ker}(f) \to X$ and the cokernel of f by $\mathsf{coker}(f) : Y \to \mathsf{Coker}(f)$.*

The axioms for an abelian category were considerably refined after Mac Lane introduced abelian bicategories. We give what is possibly the shortest list of axioms which define the modern notion of an abelian category (see [51] and the references there):

(A1) $\mathbb{C}$ has finite products and is pointed,

(A2) every morphism in $\mathbb{C}$ factorizes as a cokernel followed by a kernel.

The condition of being pointed is clearly self-dual, as is the entirety of (A2). However, (A1) only requires $\mathbb{C}$ to have finite products, and not finite coproducts. Nonetheless, it is well-known that one of the many consequences of these two simple axioms is that $\mathbb{C}$ is additive, and that therefore coproducts not only exist in $\mathbb{C}$, but coincide with products (i.e. they are biproducts). Thus the notion of an abelian category is defined by dual axioms, and so every property which holds in every abelian category will have a dual which also holds.

Just as the category of groups failed to be an abelian bicategory, so it also fails to satisfy (A1) and (A2). To see this, one only has to note that (A2) implies that every monomorphism is a kernel. Indeed, given a monomorphism f, if $f = me$ for a kernel m and cokernel e, then e is also a monomorphism and hence an isomorphism, which gives that f is a kernel. Clearly not every monomorphism in groups is a kernel – this would be to require that every subgroup is normal.

The notion of abelian category is thus not suitable for an axiomatic investigation of the totality of groups and of other, non-abelian, group-like structures (for example rings). A different axiomatic context is thus needed. In the next sections we describe one such context, that of a semi-abelian category, which provides a suitable setting for a variety of topics arising out of group theory, but whose axioms, unlike those of an abelian category, are not self-dual.

2.2.2 Factorization systems

In the category of sets, groups or indeed in any variety of universal algebras, every morphism factorises as an epimorphism followed by a monomorphism. The notion of a factorisation system, introduced by Freyd and Kelly [37], generalises this property. We will now give the definition of a factorisation system as well as that of two weaker notions, namely a prefactorisation system and a right factorisation system. We will say that a class $\mathcal{C}$ of morphisms in a category $\mathbb{C}$ is *closed under composition with isomorphisms* if $vcu \in \mathcal{C}$ whenever $c \in \mathcal{C}$ and v and u are both isomorphisms.

Definition 17. *A factorisation system [37] on a category $\mathbb{C}$ is a pair $(\mathcal{E}, \mathcal{M})$ of classes of morphisms satisfying the following conditions:*

(i) each of $\mathcal{E}$ and $\mathcal{M}$ contains the identity morphisms and are closed under composition with isomorphisms;

(ii) every morphism f in $\mathbb{C}$ can be written as $f = me$, where $m \in \mathcal{M}$ and $e \in \mathcal{E}$ (this is called an $(\mathcal{E}, \mathcal{M})$-factorisation of f);

(iii) for any commutative diagram of solid arrows below, with $m, m' \in \mathcal{M}$ and $e, e' \in \mathcal{E}$, there is a unique morphism w making the diagram commute:

$$\begin{tikzcd}
\cdot \arrow[r, "e"] \arrow[d, "u"'] & \cdot \arrow[r, "m"] \arrow[d, dotted, "w"] & \cdot \arrow[d, "v"] \\
\cdot \arrow[r, "e'"'] & \cdot \arrow[r, "m'"'] & \cdot
\end{tikzcd}$$

Note that in (iii), if v and u are isomorphisms, then so is w. Thus the $(\mathcal{E}, \mathcal{M})$-factorisation of a morphism is unique up to isomorphism. We now look at the notion of a prefactorisation system. Let $\mathbb{C}$ be a category. For two morphisms $f : X \to Y$ and $g : X' \to Y'$, we write $f \downarrow g$ if, for every commutative square of solid arrows below, there is a unique morphism i making the diagram commute.

$$\begin{tikzcd}
X \arrow[r, "v"] \arrow[d, "f"'] & Y \arrow[d, "g"] \\
X' \arrow[r, "u"'] \arrow[ur, dotted, "i"] & Y'
\end{tikzcd}$$

For a class of morphisms $\mathcal{C}$, we set

$$\mathcal{C}^{\uparrow} = \{f | \forall_{g \in \mathcal{C}} (f \downarrow g)\}, \quad \mathcal{C}^{\downarrow} = \{f | \forall_{g \in \mathcal{C}} (g \downarrow f)\}$$

Definition 18. *A prefactorisation system on a category $\mathbb{C}$ is a pair $(\mathcal{E}, \mathcal{M})$ of classes of morphisms having $\mathcal{E} = \mathcal{M}^{\uparrow}$ and $\mathcal{M} = \mathcal{E}^{\uparrow}$.*

The following proposition is easy to prove and appears in [21], where the notion of prefactorisation system was introduced. Notice that condition (iii) in the definition of a factorisation system is implied by $\mathcal{E} \subseteq \mathcal{M}^{\uparrow}$.

Proposition 13. *Factorisation systems are precisely those prefactorisation systems which satisfy condition (ii) in the definition of a factorisation system.*

We now recall a weaker notion than that of a factorisation system, which is designed to allow one to define "images" in a category. The notion first appeared under a different name in [30] (what we here call a right $\mathcal{M}$-factorisation was there called a *strong $\mathcal{M}$-image*). The following definition and name are taken from [28].

Definition 19. *Let $\mathbb{C}$ be a category and let $\mathcal{M}$ be a class of morphisms. Then a* right $\mathcal{M}$-factorisation *of a morphism f is a factorisation $f = me$, with $m \in \mathcal{M}$, such that for any commutative diagram of solid arrows below, where $n \in \mathcal{M}$, there is a unique morphism w making the diagram commute.*

$$\begin{tikzcd}
{} \arrow[d, "e"'] \arrow[dr, "u"] & \\
{} \arrow[r, dotted, "w"'] \arrow[d, "m"'] & {} \arrow[d, "n"] \\
{} \arrow[r, "v"'] & {}
\end{tikzcd}$$

We will say that a category $\mathbb{C}$ admits a right $\mathcal{M}$-factorisation system when every morphism in $\mathbb{C}$ has a right $\mathcal{M}$-factorisation. Factorisation systems are special cases of right factorisation systems, as the following proposition taken from [28] shows.

Proposition 14. *Let $\mathbb{C}$ be a category and let $\mathcal{M}$ be a class of morphisms. Then $\mathcal{M}$ is part of a factorisation system $(\mathcal{E}, \mathcal{M})$ if and only if $\mathcal{M}$ is closed under composition with isomorphisms, every morphism admits a right $\mathcal{M}$ factorisation and $\mathcal{M}$ is closed under composition. Moreover, the class $\mathcal{E}$ is uniquely determined by $\mathcal{M}$ when this is the case.*

Example 1. *An example of a class $\mathcal{M}$ for which every morphism has a right $\mathcal{M}$-factorisation, but which is not part of a factorisation system, is the class of all normal monomorphisms in* **Grp***: every morphism factors through the normal closure of its image, but since normal monomorphisms are not closed under composition, $\mathcal{M}$ cannot be part of a factorisation system by Proposition 14. It* *should be noted that normal epimorphisms* are *closed under composition in* **Grp**, *another example of a property whose dual does not hold in the category of groups.*

Definition 20. *Let $(\mathcal{E}, \mathcal{M})$ be a factorisation system on a category $\mathbb{C}$. We call $(\mathcal{E}, \mathcal{M})$* stable *when the class $\mathcal{E}$ is pullback stable, i.e. when for any morphism $f : X \to Y$ and any morphism $e : E \to Y$ with $e \in \mathcal{E}$, the pullback of e along f is also in $\mathcal{E}$.*

The class $\mathcal{M}$ in a factorisation system $(\mathcal{E}, \mathcal{M})$ is always pullback-stable, so in particular, a stable factorisation system allows us to compute the factorisation of the pullback of a morphism f along a morphism g "piece-by-piece". More precisely, if $f = me$ is the factorisation of a morphism f, then in the diagram below, where both squares and hence also the outer rectangle are pullbacks, $f' = m'e'$ is the factorisation of f', where f' is the pullback of f along g.

$$\begin{tikzcd} {} \arrow[r, "e"] & {} \arrow[r, "m"] & {} \\ {} \arrow[u] \arrow[r, "e'"'] & {} \arrow[u] \arrow[r, "m'"'] & {} \arrow[u, "g"'] \end{tikzcd}$$

Example 2. *For a category $\mathbb{C}$, let $\mathcal{C}$ be the class of all morphisms and let $\mathcal{I}$ be the class of all isomorphisms. Then there are always two factorisation systems $(\mathcal{C}, \mathcal{I})$ and $(\mathcal{I}, \mathcal{C})$ on $\mathbb{C}$, both of which are stable.*

Example 3. *The category* **Top** *of topological spaces admits two factorisations systems:*

- *$(\mathcal{E}, \mathcal{M})$ where $\mathcal{E} =$ all quotient maps and $\mathcal{M} =$ all injective maps;*
- *$(\mathcal{E}, \mathcal{M})$ where $\mathcal{E} =$ all surjective maps and $\mathcal{M} =$ all embeddings of spaces.*

Only the second of these is stable, however.

2.2.3 Regular and normal categories

Regular categories were introduced in [3]. One way to view the axioms of a regular category is as the minimal list of axioms required to establish a *calculus of relations*, i.e. the ability to compose relations in a meaningful way.

Definition 21. *A category $\mathbb{C}$ is* regular *if it satisfies the following conditions:*

(S1a) *$\mathbb{C}$ has finite limits,*

(S1b) *every morphism in $\mathbb{C}$ can be factorized as a regular epimorphism (i.e. a coequalizer) followed by a monomorphism.*

(S1c) regular epimorphisms are stable under pullback

Lemma 11. *Let e be a regular epimorphism and m be a monomorphism in a category $\mathbb{C}$. Then $e \downarrow m$.*

Let RegEpi be the class of all regular epimorphisms, and let Mono be the class of all monomorphisms. Lemma 11 thus states that $\mathsf{RegEpi} \subseteq \mathsf{Mono}^{\uparrow}$. Combining this observation with (S1b), we can equivalently state the definition of a regular category as a finitely complete category which admits a stable $(\mathsf{RegEpi}, \mathsf{Mono})$-factorisation system.

Example 4. *Any variety of universal algebras is regular (the required factorisation is simply the image factorisation).*

Example 5. *While the category of topological spaces does satisfy conditions (S1a) and (S1b), it was already mentioned that it does not satisfy condition (S1c) (regular epimorphisms are precisely the quotient maps). Its dual (i.e.* $\mathbf{Top}^{\mathsf{op}}$*), is regular, however.*

Definition 22. *A category $\mathbb{C}$ is* normal *[55] if it is pointed, regular and every regular epimorphism is a normal epimorphism (i.e. a cokernel).*

An equivalent definition for a normal category is a pointed, finitely complete category which admits a stable $(\mathsf{NormEpi}, \mathsf{Mono})$-factorisation system (where $\mathsf{NormEpi}$ is the class of all normal epimorphisms). An example of a category which is pointed and regular but is not normal is the category of pointed sets. The notion of a normal category generalises a classical result of group theory known as the First Isomorphism Theorem.

Theorem 2 (First Isomorphism Theorem)**.** *Any homomorphism of groups $f : G \to H$ factors as $f = m \circ \theta \circ e$, where $e : G \to G/\mathsf{Ker}(f)$ is a quotient map, θ is an isomorphism, and $m : \mathsf{Im}(f) \to H$ is the inclusion of the image of f into H.*

Cokernels in **Grp** are precisely the quotient maps, which coincide with the surjective maps and hence also with the regular epimorphisms. As a direct consequence of the First Isomorphism Theorem, we thus have:

Proposition 15. **Grp** *is a normal category.*

2.2.4 Internal relations and Barr-exact categories

A relation from a set X to a set Y can be viewed as a subset of the product $X \times Y$. This notion can be generalised to give the notion of an *internal relation* in a category $\mathbb{C}$. Recall that a pair of morphisms $(f : X \to Y, g : X \to Z)$ in a category $\mathbb{C}$ are called *jointly monic* if $(fh = fh' \wedge gh = gh') \Rightarrow h = h'$ for all pairs of morphisms $h, h' : W \to X$. If $\mathbb{C}$ admits products, then this is equivalent to the morphism $(f, g) : X \to Y \times Z$ being a monomorphism.

Definition 23. *Let* $\mathbb{C}$ *be a category. An* internal relation *from an object* X *to an object* Y *is triple* $(R, r_1 : R \to X, r_2 : R \to Y)$ *such that* r_1 *and* r_2 *are jointly monic.*

Let (R, r_1, r_2) be a relation from an object X to an object Y. Then for every object S in $\mathbb{C}$, the functor $\mathsf{hom}(S, -)$ sends (R, r_1, r_2) to the diagram:

$$\mathsf{hom}(S, X) \xleftarrow{r_1 \circ -} \mathsf{hom}(S, R) \xrightarrow{r_2 \circ -} \mathsf{hom}(S, Y)$$

We have already noted that the functor $\mathsf{hom}(S, -)$ preserves all limits that exist in $\mathbb{C}$. Being a relation is a property that can be expressed in terms of limits which always exist. Indeed, a pair of morphisms (r_1, r_2) is jointly monic if and only if the following is a limit diagram (i.e. the object R together with the identity morphisms form the limit cone over the rest of the diagram):

$$\begin{tikzcd}
R \arrow[r, equal] \arrow[d, equal] & R \arrow[d, "r_1"] \arrow[dr, bend left, "r_2"] & \\
R \arrow[r, "r_1"'] \arrow[rr, bend right, "r_2"'] & X & Y
\end{tikzcd}$$

Thus $\mathsf{hom}(S, -)$ sends a relation in $\mathbb{C}$ to a relation in **Set**. Now let (R, r_1, r_2) be a relation from X to X in a category $\mathbb{C}$. We say that (R, r_1, r_2) is *reflexive/symmetric/transitive* when for every object S, the relation

$$\mathsf{hom}(S, X) \xleftarrow{r_1 \circ -} \mathsf{hom}(S, R) \xrightarrow{r_2 \circ -} \mathsf{hom}(S, X)$$

is reflexive/symmetric/transitive. An *internal equivalence relation* in $\mathbb{C}$ is an internal relation which is reflexive, symmetric and transitive.

In the presence of finite limits, reflexivity, symmetry and transitivity can be expressed without reference to the hom-functor. For a finitely complete category $\mathbb{C}$, a relation (R, r_1, r_2) from X to X is reflexive if and only if there is a common splitting of (r_1, r_2), i.e. a morphism δ such that $r_1 \circ \delta = r_2 \circ \delta = 1_X$. The relation (R, r_1, r_2) is symmetric if and only if there exists a morphism $\sigma : R \to R$ such that $r_1 = r_2 \circ \sigma$ and $r_2 = r_1 \circ \sigma$. Let the pullback of r_2 along r_1 be given by the following diagram:

$$\begin{tikzcd}
R \times_X R \arrow[r, "p_2"] \arrow[d, "p_1"'] & R \arrow[d, "r_1"] \\
R \arrow[r, "r_2"'] & X
\end{tikzcd}$$

Then (R, r_1, r_2) is transitive if and only if there is a morphism $\tau : R \times_X R \to R$ such that $r_1 \circ p_1 = r_1 \circ \tau$ and $r_2 \circ p_2 = r_2 \circ \tau$.

We have already mentioned that regular categories allow one to compose relations. Let $\mathbb{C}$ be a regular category, let (R, r_1, r_2) be a relation from an object X to an object Y and let (S, s_1, s_2) be a relation from Y to an object Z. Let (T', t'_1, t'_2) be the pullback of r_2 along s_1 as shown in the following diagram:

$$\begin{tikzcd}
X & & Y & & Z \\
& R \arrow[ul, "r_1"] \arrow[ur, "r_2"'] & & R \arrow[ul, "s_1"'] \arrow[ur, "s_2"'] & \\
& & T' \arrow[ul, "t'_1"] \arrow[ur, "t'_2"'] & &
\end{tikzcd}$$

The triple $(T, r_1 \circ t'_1, s_2 \circ t'_2)$ may no longer be a relation. However, we can factorise the morphism $(r_1 \circ t'_1, s_2 \circ t'_2) : T' \to X \times Z$ as a regular epi followed by a monomorphism $(t_1, t_2) : T \to X \times Z$, which will be defined to be the composite of R and S (denoted by RS).

Note that it is not true in general that the composite of two equivalence relations is again an equivalence relation. It is however easy to check that the composite of two reflexive relations is again a reflexive relation.

Definition 24. *For a morphism f in a category $\mathbb{C}$, the* kernel pair *of f (denoted by $(\mathsf{Eq}(f), \pi_1, \pi_2)$) is the pullback of f along itself. The relation* induced by a morphism *f is its kernel pair.*

Proposition 16. *For any morphism f in a category $\mathbb{C}$, the kernel pair*

$$(\mathsf{Eq}(f), \pi_1, \pi_2)$$

is an internal equivalence relation (when it exists).

We sometimes refer to just the two morphisms π_1, π_2 as a kernel pair. It is easy to check that every regular epimorphism is in particular the coequalizer of its kernel pair viewed in this way. Conversely, a kernel pair is always the kernel pair of its coequalizer. Regular categories always have coequalizers of kernel pairs: indeed, if π_1, π_2 are the kernel pair of a morphism f, then the regular epi e occuring in the factorisation of f is the coequalizer of π_1 and π_2.

Given an object C in a category $\mathbb{C}$ which has coequalizers of kernel pairs, we can consider the poset K_C whose objects are kernel pairs with codomain C and where for two kernel pairs (π_1, π_2) and (π'_1, π'_2), $(\pi_1, \pi_2) \leq (\pi'_1, \pi'_2)$ if there is a morphism r such that $\pi'_1 \circ r = \pi_1$ and $\pi'_2 \circ r = \pi_2$ (such a morphism will always be unique). We can also consider the poset E_C whose objects are regular epis with domain C and where $e \leq e'$ if e' factors through e. It is easy to show using the above remarks that there is an isomorphism between K_C

and E_C which sends a kernel pair to its coequalizer and whose inverse sends a regular epi to its kernel pair.

Definition 25. *An internal equivalence relation is called* effective *when it is the kernel pair of some morphism. A category $\mathbb{C}$ is called* Barr-exact *[3] if it is regular and every internal equivalence relation is effective.*

Examples of Barr-exact categories include all varieties of universal algebras. In particular, **Grp** is Barr-exact.

2.2.5 Protomodular categories

Let $\mathbb{C}$ be a category and let B be an object in $\mathbb{C}$. Then the *category of points of B in $\mathbb{C}$* (denoted by $\mathsf{Pt}_{\mathbb{C}}(B)$) is defined as follows (see for example [51]):

- objects of $\mathbb{C}$ are triples (E, p, s) with morphisms $p : E \to B$, $s : B \to E$ such that $p \circ s = 1_B$;
- a morphism $f : (E, p, s) \to (E', p', s')$ is a morphism f in $\mathbb{C}$ such that $p'f = p$ and $fs = s'$.

Given any morphism $f : X \to Y$ in $\mathbb{C}$, there is a pullback functor $f^* : \mathsf{Pt}_{\mathbb{C}}(Y) \to \mathsf{Pt}_{\mathbb{C}}(X)$, whose definition we now recall (note that it will be defined only up to natural isomorphism). For an object (E, p, s) of $\mathsf{Pt}_{\mathbb{C}}(Y)$, let the pullback of p' along f be given by the following diagram:

$$\begin{tikzcd} E' \arrow[r, "f'"] \arrow[d, "p'"'] & E \arrow[d, "p"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

Then the pullback functor sends (E, p, s) to the object (E', p', s'), where s' is the unique morphism such that $p' \circ s' = 1_X$ and $f' \circ s' = s \circ f$ obtained from the property of the pullback. For a morphism $g : (E_1, p_1, s_1) \to (E_2, p_2, s_2)$, let $f^*((E_1, p_1, s_1)) = (E_1', p_1', s_1')$ and $f^*((E_2, p_2, s_2)) = (E_2', p_2', s_2')$, with f_1' and f_2' the pullbacks of f along p_1 and p_2 respectively. Then we have $p_2 \circ (g \circ f_1') = f \circ p_1'$, so there is a morphism $g' : E_1' \to E_2'$ such that $p_2' \circ g' = p_1'$. It remains to show that $g' \circ s_1' = s_2'$ for g' to be a morphism in $\mathsf{Pt}_{\mathbb{C}}(X)$. We have:

$$f_2' \circ (g' \circ s_1') = g \circ f_1' \circ s_1' = g \circ s_1 \circ f = s_2 \circ f = f_2' \circ s_2'$$

and

$$p_2' \circ (g' \circ s_1') = p_1' \circ s_1' = 1_X = p_2' \circ s_2'$$

so by the uniqueness part of the property of the pullback, we have that $g' \circ s_1' = s_2'$. The functor f^* sends the morphism g to g'. It is easy to check that this defines a functor.

Definition 26. *A category $\mathbb{C}$ is called* (Bourn)-protomodular *[13] if it has pullbacks and if for every morphism f in $\mathbb{C}$, the pullback functor f^* reflects isomorphisms (i.e. $f^*(\theta)$ is an isomorphism if and only if θ is).*

Lemma 12. *A pointed category $\mathbb{C}$ with zero object 0 is protomodular if and only if for every zero morphism $z_Y : 0 \to Y$, the pullback functor z_Y^* reflects isomorphisms.*

Proof. Suppose z_Y^* reflects isomorphisms for every zero morphism z_Y. Let $f : X \to Y$ be a morphism in $\mathbb{C}$. Then $f \circ z_X = z_Y$, so it is easy to check that $z_X^* \circ f^* = z_Y^*$ for some choice of z_Y^* (as remarked earlier, it is defined only up to natural isomorphism). Thus since z_Y^* reflects isomorphisms and z_X^* preserves isomorphisms (any functor does), it is easy to see that f must reflect isomorphisms as well. The converse is obvious. □

For a zero morphism $z_Y : 0 \to Y$, z_Y^* is nothing but the functor which associates to each morphism $p : E \to Y$ its kernel object $\mathsf{Ker}(p)$. Thus a pointed category $\mathbb{C}$ is protomodular if and only if for every object X in $\mathbb{C}$, the kernel functor $\mathsf{ker}_X : \mathsf{Pt}_{\mathbb{C}}(X) \to \mathbb{C}(\cong \mathsf{Pt}_{\mathbb{C}}(0))$ reflects isomorphisms. This is equivalent to the *Split Short Five Lemma* holding in $\mathbb{C}$, i.e. that in any commutative diagram of the form of the one below, with $k = \mathsf{ker}(p)$ and $l = \mathsf{ker}(q)$ and p and q split epimorphisms, w is an isomorphism if u and v are.

$$\begin{tikzcd} L \arrow[r, "l"] \arrow[d, "u"'] & F \arrow[r, "q"] \arrow[d, "w"] & C \arrow[d, "v"] \\ K \arrow[r, "k"'] & E \arrow[r, "p"'] & B \end{tikzcd} \tag{2.2.1}$$

In a normal category $\mathbb{C}$, split epimorphisms are always normal epimorphisms, so for a normal category to be protomodular it is sufficient to have that the *Short Five Lemma* (which appears under the name of *ABC extension property* in [73]) holds, i.e. the above property but where p and q are only required to be cokernels (they will then in particular be the cokernels of k and l respectively).

Proposition 17. Grp *is protomodular.*

Proof. It is enough to prove the Short Five Lemma, and the proof is classical. Consider a diagram of the form of 2.2.1, where $k = \mathsf{ker}(p)$, $l = \mathsf{ker}(q)$, p and q are the cokernels of k and l respectively, and u and v are both isomorphisms. To prove that w is a isomorphism, it is sufficient to prove that its kernel is trivial and that it is surjective. Suppose $w(f) = 0$ for some $f \in F$. Then $vq(f) = pw(f) = 0$, so since v is an isomorphism, $q(f) = 0$. Thus f is in the kernel of q, so there is an element $x \in L$ such that $l(x) = f$. Then $ku(x) = wl(x) = 0$, but u and k are injective, so $x = 0$ and thus $f = l(x) = 0$ as required. Now let e be an element of E. Let f be an element in the inverse image of $v^{-1}p(e)$ under q. Consider the element $e' = w(f) \cdot e^{-1}$. Clearly $p(e') = 0$, so there must be an element y of K such that $k(y) = e'$. Consider then the element $lu^{-1}(y^{-1}) \cdot f$. We have:

$$w(lu^{-1}(y^{-1}) \cdot f) = k(y^{-1}) \cdot w(f) = e'^{-1} \cdot w(f) = e$$

so w is surjective as required. □

The Short Five Lemma is one of the results that can be obtained from the axioms for a abelian bicategory introduced by Mac Lane. In fact, since those axioms are self-dual, it is only necessary to prove that w has trivial kernel, and the other half follows by duality. This type of argument cannot be used straightforwardly in case of groups, because of the lack of duality there. At the end of Chapter 5, we discuss an alternative proof of this lemma using the new context developed in this thesis.

Protomodularity is essentially a condition on split epimorphisms. It is thus useful to look more closely at split epimorphisms in **Grp**. They turn out to be precisely the projections from so-called *semi-direct products*. This property holds in the more general setting of semi-abelian categories (with semi-direct products suitably defined) [17], and leads to the study of *internal object actions* [12]. The following definition can be found in [7].

Definition 27. *Let G and H be groups and let θ be a group homomorphism from H to the automorphism group* $\mathsf{Aut}(G)$ *of G. Then the* semi-direct product of G and H relative to θ, *denoted by $G \rtimes_\theta H$, is defined as follows:*

- *the underlying set of $G \rtimes_\theta H$ is the same as $G \times H$,*
- *for two elements (g, h) and (g', h'), using multiplicative notation,*

$$(g, h) \cdot (g', h') = (g \cdot (\theta(h)(g')), h'h)$$

.

Notice that for every semi-direct product we have the following two group homomorphisms: firstly, the projection $\pi_2 : G \rtimes_\theta H \to H$ which sends (g, h) to h and secondly, the injection $\delta : H \to G \rtimes_\theta H$ which sends h to $(1, h)$.

Proposition 18. *Let $p : G \to H$ and $s : H \to G$ be group homomorphisms such that $p \circ s = 1_G$, and let N be the kernel of p (seen as a subgroup of G for simplicity). Then there is an isomorphism ϕ from the semi-direct product $N \rtimes_\theta H$ to G, where θ sends an element $h \in H$ to the automorphism $n \mapsto s(h) \cdot n \cdot s(h)^{-1}$ of N. Moreover, $p \circ \phi = \pi_1$, where $\pi_2 : N \rtimes_\theta H \to H$ is the projection $(n, h) \mapsto h$ and $s = \phi \circ \delta$, where $\delta : H \to N \rtimes_\theta H$ is the injection $h \mapsto (1, h)$.*

Proof. We define a map $\phi : N \rtimes_\theta H \to G$ as follows: $\phi(n,h) = n \cdot (s(h))$. We claim that this is a group homomorphism. Let (n,h) and (n',h') be two elements of $N \rtimes_\theta H$. Then:

$$\phi((n,h)\cdot(n',h')) = \phi(ns(h)n's(h)^{-1}, hh') = ns(h)n's(h') = \phi(n,h)\cdot\phi(n',h').$$

Thus ϕ is a group homomorphism, $p(\phi(n,h)) = p(n \cdot s(h)) = 1 \cdot ps(h) = h = \pi_1(n,h)$ and $s = \phi \circ \delta$ is obvious. It remains to show that ϕ is an isomorphism. For this it is enough to show that it is surjective and has kernel the zero group. Suppose that $\phi(n,h) = 1$ (where 1 is the unit of G). Then $n \cdot s(h) = 1$, so $p(n \cdot s(h)) = h = 1$, hence $n = 1$, which shows that the kernel of ϕ is trivial. Now let g be an element of G. Then $n' = g \cdot (sp(g))^{-1}$ is in the kernel of p, i.e. in N. Thus $g = n' \cdot sp(g) = \phi(n', p(g))$. □

2.2.6 Mal'tsev categories

By now we have all the ingredients necessary to describe semi-abelian categories, but we should make some remarks on a related notion, namely the notion of a Mal'tsev category. Mal'tsev categories were introduced by Carboni, Kelly and Pedicchio [22], although since then the requirement that the category be Barr-exact has been dropped. This form of the definition is given in [23].

Definition 28. *A category $\mathbb{C}$ is Mal'tsev if it has finite limits and satisfies the following condition:*

(M) every reflexive internal relation is an equivalence relation.

The roots of this notion go back to a classical theorem of A. Mal'tsev [76], which states that the composition of congruences on any object in a variety $\mathbb{X}$ is commutative if and only if the theory of $\mathbb{X}$ contains a ternary term μ satisfying the term equations

$$\mu(x,x,y) = x = \mu(y,x,x).$$

Such varieties were later called *Mal'tsev varieties* [83]. As expected, those varieties of universal algebras which form Mal'tsev categories are precisely the Mal'tsev varieties (see [36] and [87]). Thus we can already conclude that **Grp** is a Mal'tsev category: simply set $\mu(x,y,z) = xy^{-1}z$ (using multiplicative notation).

There are a number of conditions on a category which in various settings are equivalent to condition (M) – see [19] for a survey of these, as well as two other characterisations of Mal'tsev categories using so-called *approximate Mal'tsev operations.*

Observe that in a regular Mal'tsev category, the composite of two equivalence relations is again an equivalence relation. Define an order on relations from an object X to itself in the obvious way, namely that $(R, r_1, r_2) \leq$

(R', r'_1, r'_2) (we will sometimes just say that R is *contained* in R') if there is a morphism $u : R \to R'$ such that $r_1 \circ u = r'_1$ and $r_2 \circ u = r_2$ (such a morphism will necessarily be unique). Then it is easy to show that the composite RS of two equivalence relations on an object X is the join (i.e. least upper bound) of R and S in the partially ordered set of equivalence relations.

We end this section by proving the following theorem taken from [14]. The proof given in [14] goes via another characterisation of Mal'tsev categories; here we present a direct proof which relies heavily on the Yoneda Lemma (see Section 2.1.2).

Theorem 3. *Every finitely complete protomodular category $\mathbb{C}$ (so, in particular, any pointed protomodular category) is a Mal'tsev category.*

Before we prove this proposition, we need some further facts about relations in categories. For a reflexive relation R from a set X to itself (which we can view as an internal relation (R, r_1, r_2) in **Set**), it is easy to see that R is an equivalence relation if and only if the following holds:

$$xRy \wedge xRz \Rightarrow yRz.$$

Otherwise stated, let R_0 be the set of triples (x, y, z) with xRy and xRz and K_2 be the set of triples (x, y, z) with xRy, xRz and yRz. Then R is an equivalence relation if the obvious inclusion $j : K_2 \to R_0$ is an isomorphism. Both K_2 and R_0 can be expressed as a limit of a diagram (K_2 is called the simplicial kernel of R – see for example [14]). Indeed, R_0 is just the pullback of r_1 along itself, while K_2 is the limit of the diagram of solid arrows below (where the dotted arrows show the limit cone):

$$\begin{tikzcd}
 & & X & & \\
 & X & & X & \\
R \arrow[ur, "r_1"'] \arrow[uurr, bend left, "r_2"] & & R \arrow[ul, "r_1"'] \arrow[ur, "r_2"] & & R \arrow[ul, "r_2"] \arrow[uull, bend right, "r_1"'] \\
 & & K_2 \arrow[ull, dotted, "k_1"] \arrow[u, dotted, "k_2"] \arrow[urr, dotted, "k_3"'] & &
\end{tikzcd}$$

Thus for a reflexive relation in any finitely complete category $\mathbb{C}$, we can define K_2 and R_0. Since $r_1 k_1 = r_1 k_2$, we have a morphism $j : K_2 \to R_0$ such that $p_1 j = k_1$ and $p_2 j = k_2$, which for sets is precisely the inclusion we earlier required to be an isomorphism.

Thus we have that the following are equivalent, where the last equivalence follows from the fact that the Yoneda embedding reflects isomorphisms (which are the same as component-wise isomorphisms):

- R is an equivalence relation;
- for every object Y, $\mathsf{hom}(Y, -)$ takes R to an equivalence relation;
- for every object Y, $\mathsf{hom}(Y, j)$ is an isomorphism;
- j is an isomorphism.

Proof of Proposition 3. Let R be a reflexive relation in $\mathbb{C}$ and consider the situation described above. We are required to show that the morphism j is an isomorphism. We have the following cone:

$$
\begin{tikzcd}
 & & X & & \\
 & X & & X & \\
R \arrow[uurr, bend left, "r_2"] \arrow[ur, "r_1"'] \arrow[drr, equal] & & R \arrow[ul, "r_1"] \arrow[ur, "r_2"'] \arrow[d, equal] & & R \arrow[uull, bend right, "r_1"'] \arrow[ul, "r_2"] \\
 & & R \arrow[urr, "sr_2"'] & &
\end{tikzcd}
$$

so there is a morphism $\sigma_1 : R \to K_2$ such that $k_1\sigma_1 = k_2\sigma_2 = 1_R$ and $k_3\sigma_1 = sr_2$. Let (R_0, p_1, p_2) be the pullback of r_1 along itself. Then p_1 and p_2 have a common splitting δ. The two morphisms $j\sigma_1 : R \to R_0$ and $\delta : R \to R_0$ are in fact equal by the uniqueness part of the property of the pullback. Hence j is a morphism in $\mathsf{Pt}_{\mathbb{C}}(R)$ from (K_2, k_1, σ_1) to (R_0, p_1, δ). Consider now the pullback functor $s^* : \mathsf{Pt}_{\mathbb{C}}(R) \to \mathsf{Pt}_{\mathbb{C}}(X)$. Let

$$s^*(j) = j' : (K_2', k_1', \sigma_1') \to (R_0', p_1', \delta').$$

We claim that j' is an isomorphism. Since j' is defined using only limits and composites, we can apply Metatheorem 1. Thus, consider the same situation, but where $\mathbb{C}$ is **Set**. Note that **Set** is obviously not a Mal'cev category itself, so j will not in general be an isomorphism. However, we will now show that j' always is.

We have already given a description of j, so j' is simply the pullback of j along the joint splitting of the relation. The object R_0' is then the set of all triples (x, x, y) with xRx and xRy, while K_2' is the set of all triples (x, x, y) with xRx, xRy and (now redundantly) xRy. Clearly the inclusion j' is an isomorphism. Returning to the case of an arbitrary category $\mathbb{C}$ which is finitely complete and protomodular, we obtain that j' is an isomorphism also in the case of $\mathbb{C}$. Now we apply the protomodularity of $\mathbb{C}$. Since $\mathbb{C}$ is protomodular, s^* reflects isomorphisms, and thus j must be an isomorphism. $\square$

2.2.7 Semi-abelian categories

Semi-abelian categories were defined in [51], and represent a culmination of a number of developments towards an axiomatic context corresponding to the category of groups.

Definition 29. *A category* $\mathbb{C}$ *is* semi-abelian *if it satisfies the following conditions:*

(S1) $\mathbb{C}$ *is (Barr-)exact,*

(S2) $\mathbb{C}$ *is (Bourn-)protomodular, and*

(S3) $\mathbb{C}$ *is pointed and has binary coproducts.*

It was already remarked that all varieties of universal algebras satisfy condition (S1) in the definition of a semi-abelian category. Moreover, every variety of universal algebras admits coproducts (for example, it is well-known that the coproduct of two groups G and H is given by the so-called *free product* of G and H). Thus those varieties of universal algebras which form semi-abelian categories are precisely the pointed protomodular ones. Recalling that **Grp** is indeed pointed, Proposition 17 gives:

Proposition 19. Grp *is semi-abelian.*

We should also mention that a characterisation of those varieties of universal algebras which form semi-abelian categories is given in [18]. The characterisation makes it clear that groups are indeed an example of a semi-abelian category. However, the direct proof of the Short Five Lemma is illuminating for other reasons, as well as not requiring any background other than classical group theory, hence we included it in the section on protomodularity.

We now prove that every semi-abelian is normal. This is of course to be expected, since normal categories capture the First Isomorphism Theorem for groups, one of the most important results which distinguish groups from other algebraic structures. A proof is given already in [51], though not using the term “normal”. Here we give a slightly more detailed proof which highlights how conditions about normal epimorphisms can be translated into conditions about split epimorphisms by considering kernel pairs, motivating in some sense the condition of protomodularity.

Lemma 13. *Every semi-abelian category is finitely cocomplete.*

Proof. Let $\mathbb{C}$ be a semi-abelian category. Since it is pointed, it is enough to have pushouts. Moreover, following [51], in the presence of coproducts, it is enough to have pushouts of split epimorphisms along split epimorphisms, since the following two diagrams have isomorphic colimits:

$$\begin{tikzcd} B & \\ C \arrow[u, "g"] \arrow[r, "f"'] & A \end{tikzcd} \qquad \begin{tikzcd} A+B & \\ A+C+B \arrow[u, "1_A+\begin{pmatrix} g \\ 1_B \end{pmatrix}"] \arrow[r, "\begin{pmatrix} 1_A \\ f \end{pmatrix}+1_B"'] & A+B \end{tikzcd}$$

Since split epimorphisms are always regular epimorphisms, it is sufficient to have pushouts of regular epis along regular epis. Let $f : X \to Y$ and $g : X \to Z$ be two regular epis. Let (R, r_1, r_2) and (S, s_1, s_2) (R and S for short) be the kernel pairs of f and g respectively. Let (T, t_1, t_2) be the composite of R and S. We have coequalizers of kernel pairs, and since $\mathbb{C}$ is Barr-exact and Mal'tsev (Proposition 3), (T, t_1, t_2) is the kernel pair of some morphism. Let h be its coequalizer. Then $h = uf = vg$ for some unique morphisms u and v since R and S are both contained in T. We claim that the square below is a pushout:

$$\begin{tikzcd} {} \arrow[r, "u"] & {} \\ {} \arrow[u, "f"] \arrow[r, "g"'] & {} \arrow[u, "v"'] \end{tikzcd}$$

Let p and q be morphisms such that $pf = qg$. Then the equivalence relation (K, k_1, k_2) induced by $pf = qg$ contains both R and S. Since T is the join of R and S, K must contain T as well, and thus $pf = gq$ factors through $h = uf = vg$. The rest follows easily from the fact that f and g are both epi. □

Proposition 20. *Every semi-abelian category is normal.*

Proof. Let $\mathbb{C}$ be a semi-abelian category. It is enough to prove that every regular epimorphism is normal. Let f be a regular epimorphism. Then f decomposes as $f = \mathsf{coker}(\mathsf{ker}(f)) \circ m = e \circ m$, with e a normal epimorphism (here we use the fact that $\mathbb{C}$ is cocomplete to produce the cokernel). Consider the kernel pairs $(\mathsf{Eq}(f), \pi_1, \pi_2)$, $(\mathsf{Eq}(e), \pi_1', \pi_2')$ of f and e respectively. Since $f\pi_1' = f\pi_2'$, there is a morphism $\phi : \mathsf{Eq}(e) \to \mathsf{Eq}(f)$ such that $\pi_1' = \pi_1 \circ \phi$ and $\pi_2' = \pi_2 \circ \phi$. Both e and f are regular epimorphisms, so they are in particular coequalizers of their kernel pairs. Thus for e and f to be isomorphic, it is sufficient to show that ϕ is an isomorphism. We have the diagram below, where both inner squares are pullbacks:

$$\begin{tikzcd}
{} \arrow[r, "0"] & {} \arrow[r, "f"] & {} \\
{} \arrow[u] \arrow[r, "\mathsf{ker}(\pi_1)"'] & {} \arrow[u, "\pi_1"] \arrow[r, "\pi_2"'] & {} \arrow[u, "f"']
\end{tikzcd}$$

The outer rectangle is also a pullback, so the kernel object of π_1 is isomorphic to the kernel object of f. Similarly, the kernel object of π_1' is isomorphic to the kernel object of e, which in turn is isomorphic to the kernel object of f. Thus, since both π_1 and π_1' are split epis (a kernel pair is always a reflexive relation), ϕ is an isomorphism by the Split Short Five Lemma, i.e. by protomodularity. $\square$

That semi-abelian categories provide a good generalisation of the category of groups should be attested to by two facts: firstly, that the category of groups is an example of a semi-abelian category, and secondly, that the context of a semi-abelian category is suitable for the treatment of important topics arising from group theory. The former fact has been demonstrated already. As to the latter, semi-abelian categories and some weaker versions thereof have indeed been used in treating a variety of such topics, including torsion theories [16], commutator theory [34], homology [33] and cohomology [40].

2.3 Databases and views

In this section, we give a brief overview of databases and views of databases. In particular, we give some background on the *relational database model*, since this will be the model considered as a guiding example in Chapter 6. The relational model, originally introduced in [26], is currently the dominant technology for storing data. More recently, non-relational trends such as the noSQL movement (see for example [72]) have gained popularity. Non-relational database management systems such as Apache Cassandra [81] have been used by large companies such as Twitter, Netflix and eBay. Interestingly, a recent paper by Meijer and Bierman suggests that relational and noSQL database models should in fact be viewed as dual to one another [79].

2.3.1 Components of a relational database

The following terminology is adapted from [27]. Note that the structure of a relational database described here is the conceptual structure. In other words, it is the structure of the database as perceived by the user, and does not necessarily bear any resemblance to the way the data is physically stored.

2.3.1.1 Domains

Individual data values in the relational model are considered to be atomic – i.e. they cannot be further decomposed from the perspective of the model. A set of all possible data values of a certain type is called a *domain*. Examples might include the set of all integers in a certain range, or of all strings of a certain length containing characters from a specified list.

2.3.1.2 Relations

Let D_1, ..., D_n be sequence of domains. Then a *relation* on D_1, ..., D_n consists of

- a *heading* consisting of a fixed set of *attributes* A_1, ..., A_n
- a *body* consisting of a relation $R \subseteq D_1 \times \ldots \times D_n$.

A typical visualisation of a relation is as a table, whose column headings are the attributes A_1, ..., A_n and whose rows are the elements of R (each such row is called a *tuple*). Note that relations are considered to be unordered, so that rearranging the domains and attributes in the same way give the same relation (even though it will only be isomorphic as a set).

Sometimes the term "relation" refers to a fixed heading with a body that is time-dependent, i.e. varying as data is inserted and deleted over time. Part of the purpose of the work of Johnson and Rosebrugh as described in the last chapter of this thesis is to clarify the ambiguity between abstract *database schemas* and database states which may vary over time. Roughly speaking, a database schema is a collection of relations whose bodies are all empty, together with some specified keys (see the next section), which acts as a template for a database.

2.3.1.3 Keys

Suppose a relation on domains D_1, ..., D_n is given by a heading A_1, ..., A_n and a relation R. Then a *candidate key* is a set $\{A_{x_1}, A_{x_2}, \ldots, A_{x_m}\} \subset \{A_1, \ldots, A_n\}$ which is minimal with the property that the projections π_{x_1}, ..., π_{x_m} are jointly monic for every possible state of the relation R. In other words, it is a minimal set of attributes which uniquely determines any tuple in the relation. In particular, every relation has at least one candidate key, since the set of morphisms $\{A_1, \ldots, A_n\}$ is jointly monic, and there are only finitely many subsets of this set.

Note that the condition of being a candidate key depends on the allowable states for the relation R, i.e. its values over time. This will either be enforced by whatever software is used to manage the database, or it will be an external rule that must be adhered to. For example, an ID number might uniquely identify any person, but only if duplicate ID numbers are prohibited. For a given relation, one candidate key is designated as the *primary key*, and the rest will be called *alternate keys*. Primary keys allow the user of a database to access any single tuple uniquely, and thus are an important part of any relational database system.

A *foreign key* is an attribute (or combination of atttributes) in a relation R_1 whose values are required to match those of the primary key in a relation R_2. In particular, the foreign key and corresponding primary key should be defined on the same set of domains. Foreign keys are what allow relationships between different relations in the database.

2.3.1.4 Integrity rules

The two integrity rules below were introduced in Codd's original paper [26]. They will not be expressed mathematically, since they rely on the notion of *null value*, a value used by a database management system to indicate incomplete data. For a mathematical approach to null values, see for example [64].

- *Entity integrity*: No attribute which is part of a primary key of a relation is allowed to accept null values.

- *Referential integrity*: If FK is a foreign key in a relation R_1 matching a primary key PK in a relation R_2, then every value of FK must be either (a) equal to the value of PK in some tuple of R_2 or (b) wholly null (i.e. every component of the tuple is null).

2.3.2 Queries and Views

A query language allows a user to extract, manipulate and combine information held in the database. Currently, SQL is by far the most widely-used query language. It was first defined by Chamberlin and others at the IBM Research Laboratory [25]. It includes *data definition* functions which are used to manipulate the database schema by creating relations (called "tables" in SQL) or deleting them, as well as functions for manipulating data stored in a relational database (inserting or deleting tuples for example).

We will not describe the syntax of SQL in detail here, since it is not necessary for the work in Chapter 6. We will show just one example of a query, along the lines of the example given in Chapter 6. Suppose a database contains a table (i.e. a relation) called `suppliers` containing all the suppliers for a certain company, with attributes 'Name' and 'Location'. The following query (a SELECT query) will extract every row of that table where the 'Location' attribute is 'Cape Town', and return it to the user:

```
SELECT *
FROM suppliers
WHERE suppliers.Location = 'Cape Town';
```

Note that the string 'Cape Town' must be included in the domain of the 'Location' attribute for this to be defined.

A *view*, as defined in SQL, is a named, derived table. Consider again the above example of suppliers for a company. One possible view would be the table consisting of all those suppliers based in Cape Town. To obtain this view, the user will execute the query

```
CREATE VIEW capetown_suppliers AS
SELECT *
FROM suppliers
WHERE suppliers.Location = 'Cape Town';
```

We will not explain the syntax in detail here, but comparing this query with the one above we see that the view, like any view in SQL, is obtained from a query (in this case, a SELECT query on the table suppliers). This is what is meant by a "derived" table. Note that while it was the case here, it is not necessary that a view contain every attribute of a table.

Views can also be seen as virtual tables. Upon executing the query above, the user may manipulate (or attempt to manipulate) the view as if it were a table in the database, using the SQL language. The central question of Chapter 6 is to what extent this is possible.

2.3.3 View updates

The *view update problem* is the problem of determining when and how changes to a view can be propagated to the underlying database. For example, consider a view of the `suppliers` table which only contains the names and not the locations of the suppliers. This view is returned by the following query:

```
CREATE VIEW capetown_suppliers AS
SELECT name
FROM suppliers;
```

If the user inserts a new supplier 'Supplier X' into this view (thinking of it as a table), and the change must be propagated back to the underlying database, then a choice must be made for the 'Location' attribute for this supplier. Possible solutions include a default value, or the `null` value, but in general these are not always allowed by the database system. In other words, not all views are updatable. More detailed examples of updatable and non-updatable views can be found in Date's book [27], Johnson and Rosebrugh's paper [65] and Chapter 6 of this thesis.

In practice, the view update problem is dealt with by database management systems in an ad hoc way: certain views are defined to be updatable, with no general criteria for updatability. It is precisely the lack of a rigorous, general treatment of the view update problem that motivated the work in [65] and described in Chapter 6 of this thesis.

Views of databases abound in practical applications. They are used, for example, when a user does not have permission to, or does not desire to, access all of a company's data (for example, a cashier in a store, or an external auditor). Views and view updates are especially important in distributed systems and crowd-sourcing websites, where users must be able to make changes to certain parts of the available data without jeopardising the rest of the database. View updatability is also important in the context of linked databases which are required to interoperate (see [62]).

A survey of mathematical approaches to view updatability is given in [65], together with links to the work of the authors. For example, an important alternative approach to the view update problem is given by the notion of *lens* [8], which turns out to be related to the work in [65] via the more general notion of *c-lens* [68], a notion which turns out to be nothing but a Grothendieck opfibration – the central notion for this thesis.

Chapter 3

Forms of subobjects and exact sequences

3.1 Introduction

The theme of this thesis can be described as follows: replace the axiomatic context of a category $\mathbb{C}$, with that of a functor $F : \mathbb{B} \to \mathbb{C}$. A natural consequence of this idea is that the notion of duality at the level of a category $\mathbb{C}$ is replaced by the notion of functorial duality: the dual of the statement about F is that statement stated for $F^{\mathsf{op}} : \mathbb{B}^{\mathsf{op}} \to \mathbb{C}^{\mathsf{op}}$. In this chapter, we describe how to translate certain structures on a category $\mathbb{C}$ into functors $F : \mathbb{B} \to \mathbb{C}$. This is done by way of "classification" theorems: theorems which classify, using axioms on a functor F, different types of structures on categories that are typically encountered in categorical algebra.

In this chapter, we look primarily at two such structures on a category $\mathbb{C}$. The first is a class of morphisms $\mathcal{M}$ in $\mathbb{C}$ which allows one to work with "direct images". By this we mean that $\mathcal{M}$ satisfies the following axioms due to Ehrbar and Wyler [30]:

(M_1) $\mathcal{M}$ is closed under composition with isomorphisms;

(M_2) every morphism in $\mathbb{C}$ admits a *right $\mathcal{M}$-factorisation*, i.e. a factorisation $f = me$, with $m \in \mathcal{M}$, such that for any commutative diagram of solid arrows below, where $n \in \mathcal{M}$, there is a unique morphism w making the diagram commute.

$$\begin{tikzcd}
\cdot \arrow[d, "e"'] \arrow[dr, "u"] & \\
\cdot \arrow[r, dotted, "w"'] \arrow[d, "m"'] & \cdot \arrow[d, "n"] \\
\cdot \arrow[r, "v"'] & \cdot
\end{tikzcd}$$

The second structure we consider is that of a class $\mathcal{N}$ of morphisms in $\mathbb{C}$ which we think of as playing the role played by zero morphisms in a pointed category. This idea goes back to Ehresmann [31], Lavendhomme [71], as well as Kelly [70]. Such a structure together with certain axioms was used by Grandis in his "categorical foundations of homological algebra" [41]. The fundamental requirement on $\mathcal{N}$ is that it is closed under right and left composition with arbitrary morphisms (such a class is called an *ideal of null morphisms*), a property which zero morphisms in a pointed category always have, for example. One may then introduce the notions of kernel, cokernel and exactness of a sequence relative to $\mathcal{N}$. This allows for the development of homological algebra in the context of a category equipped with an ideal – see [43] for the complete theory, as well as examples and applications.

These two structures are related in that they both represent a way to generalise results and constructions from abelian categories to non-abelian ones: the class $\mathcal{M}$ replaces the notion of the (cokernel, kernel)-factorisation system, while the ideal $\mathcal{N}$ replaces the notion of pointedness (and hence the classical notions of cokernel and kernel). What is more surprising is that the work in this chapter shows that they are mathematically linked: the axioms which characterise functors arising from classes $\mathcal{M}$, together with their (functorial) duals, are precisely those that characterise functors arising from ideals $\mathcal{N}$. This is made more precise in the subsequent sections.

In this chapter, we first look at the general case of codomain functors, of which the notion of *form of $\mathcal{M}$-subobjects* arising from a class $\mathcal{M}$ is a special case. We then restrict to the case of faithful, amnestic functors (called *forms*), and derive a correspondence between classes $\mathcal{M}$ of monomorphisms satisfying (M_1) and (M_2) and functors satisfying certain conditions. We then deal with functors which are so-called *forms of $\mathcal{N}$-short-exact sequences* for an ideal $\mathcal{N}$ of null morphisms, and relate the axioms which characterise these functors to those corresponding to forms of $\mathcal{M}$-subobjects.

Some of the results in this chapter can be obtained via another route, one which goes via the work of Z. Janelidze on *cover relations* [54]. This approach is contained in a joint paper with Z. Janelidze [59].

3.2 Codomain functors

Let $\mathbb{C}$ be a category and $\mathbb{C}^{\mathbf{2}}$ be the category of arrows of $\mathbb{C}$. A class $\mathcal{C}$ of morphisms in $\mathbb{C}$ can be viewed as a full subcategory of $\mathbb{C}^{\mathbf{2}}$, which we denote simply by $\mathcal{C}$, following a suggestion of Tholen. Thus the category denoted by $\mathcal{C}$ has as objects morphisms $c : A \to C$ in the class $\mathcal{C}$ and as morphisms commutative squares of the form

$$\begin{tikzcd} C \arrow[r, "f"] & D \\ A \arrow[u, "c"] \arrow[r, "g"'] & B \arrow[u, "c'"'] \end{tikzcd} \tag{3.2.1}$$

where c and c' are in the class $\mathcal{C}$. Following the notation used in [30], we write such a square as $(f, g) : c \to c'$. The codomain functor $\mathsf{Cod}_{\mathcal{C}} : \mathcal{C} \to \mathbb{C}$ takes such a square to $f : C \to D$ and the domain functor $\mathsf{Dom}_{\mathcal{C}} : \mathcal{C} \to \mathbb{C}$ takes it to $g : A \to B$. Many conditions on $\mathcal{C}$ can be translated into conditions on $\mathsf{Cod}_{\mathcal{C}}$, as the following two lemmas show.

Lemma 14. *Suppose $\mathcal{C}$ contains all the identity morphisms. Then the square (3.2.1) is cartesian with respect to $\mathsf{Cod}_{\mathcal{C}}$ if and only if it is a pullback.*

Lemma 15. *Suppose $\mathcal{C}$ contains all the identity morphisms. Then the square (3.2.1) is cocartesian with respect to $\mathsf{Cod}_{\mathcal{C}}$ if and only if $c' \circ g$ is a right $\mathcal{C}$-factorisation of $f \circ c$.*

In particular, the notion of pullback becomes dual to the notion of right $\mathcal{C}$-factorisation when we think in terms of the codomain functor $\mathsf{Cod}_{\mathcal{C}}$. This is not surprising, since these two notions typically correspond to the notions of inverse image and direct image, respectively, in a category [30].

Lemma 16. *Let $\mathcal{C}$ be a class of morphisms in a category $\mathbb{C}$ containing the identity morphisms. Let $\mathsf{Id}_{\mathcal{C}} : \mathbb{C} \to \mathcal{C}$ be the functor which takes every object C in $\mathbb{C}$ to the identity morphism on C and every morphism $f : C \to D$ to $(f, f) : 1_C \to 1_D$. Then*

(1) there are two adjunctions $(\mathsf{Cod}_{\mathcal{C}}, \mathsf{Id}_{\mathcal{C}}, \gamma, 1)$ and $(\mathsf{Id}_{\mathcal{C}}, \mathsf{Dom}_{\mathcal{C}}, 1, \delta)$, where 1 denotes the identity natural transformation on $1_{\mathbb{C}}$ and where for a morphism $c : A \to C$ in $\mathcal{C}$, the components of the natural transformations γ and δ are given by $\gamma_c = (1_C, c) : c \to 1_C$ and $\delta_c = (c, 1_C) : 1_A \to c$;

(2) the class $\mathcal{C}$ is the image of all the components of δ under $\mathsf{Cod}_{\mathcal{C}}$, and dually, the image of all the components of γ under $\mathsf{Dom}_{\mathcal{C}}$;

(3) every component of δ is cocartesian with respect to $\mathsf{Cod}_{\mathcal{C}}$, and dually, every component of γ is cartesian with respect to $\mathsf{Dom}_{\mathcal{C}}$.

Proof. (1) is easy to verify and (2) is obvious. By Lemma 15, the first part of (3) is equivalent to the statement that every morphism $c : A \to C$ in $\mathcal{C}$ has a right $\mathcal{C}$-factorisation $c \circ 1_A$, which is easy to check. The second part of (3) can be proved dually. □

By an *equivalence* from a functor $F : \mathbb{B} \to \mathbb{C}$ to a functor $F' : \mathbb{B}' \to \mathbb{C}$, we mean an adjoint equivalence (E, D, ϕ, ψ) from $\mathbb{B}$ to $\mathbb{B}'$ such that $F'E = F$, $FD = F'$, $F\phi = 1_F$ and $F'\psi = 1_{F'}$. We will call two functors $F : \mathbb{B} \to \mathbb{C}$ and $F' : \mathbb{B}' \to \mathbb{C}$ *equivalent* when there is an equivalence from F to F'. The following lemma is easy to verify.

Lemma 17. *Let (E, D, ϕ, ψ) be an equivalence from a functor F to a functor F'. Then E and D preserve cartesian and cocartesian morphisms.*

Theorem 4. *For any functor $F : \mathbb{B} \to \mathbb{C}$, the following are equivalent:*

(1) there exists a class $\mathcal{C}$ of morphisms in $\mathbb{C}$ containing the identity morphisms such that $\mathsf{Cod}_{\mathcal{C}}$ and F are equivalent;

(2) there exist adjunctions $(F, T, \eta, 1)$ and $(T, L, \alpha, \varepsilon)$ with T a full functor such that every component ε_B of ε is cocartesian with respect to F (where 1 is the identity natural transformation on $1_{\mathbb{C}}$);

(3) there exists a full functor T such that $FT = 1_{\mathbb{C}}$ and an adjunction $(T, L, \alpha, \varepsilon)$ such that every component ε_B of ε is cocartesian with respect to F.

Proof. (1) $\Rightarrow$ (2) Suppose (E, D, ϕ, ψ) is the adjoint equivalence between $\mathbb{B}$ and $\mathcal{C}$ which witnesses the equivalence of F and $\mathsf{Cod}_{\mathcal{C}}$. Set $T = D \circ \mathsf{Id}_{\mathcal{C}}$ and $L = \mathsf{Dom}_{\mathcal{C}} \circ E$. Then since E is a left and right adjoint of D, we obtain the required adjunctions. In particular, the fact that $F \dashv T$ has counit the identity follows from $\mathsf{Cod}_{\mathcal{C}}\psi = 1_{\mathsf{Cod}_{\mathcal{C}}}$. The functor T is clearly full since it is the composite of the full functors $\mathsf{Id}_{\mathcal{C}}$ and D. In the notation of Lemma 16, the counit of $T \dashv L$ will be $\varepsilon = D\delta E \circ \phi$. Each component ε_B is E-cocartesian (as indeed any morphism in $\mathbb{B}$ is), and $E(\varepsilon_B)$ is $\mathsf{Cod}_{\mathcal{C}}$-cocartesian (it is isomorphic to $\delta_{E(B)}$, which is $\mathsf{Cod}_{\mathcal{C}}$-cocartesian by Lemma 16). It is easy to check that this makes ε_B F-cocartesian as required.

(2) $\Rightarrow$ (3) is obvious.

(3) $\Rightarrow$ (1): We have the natural transformation $F\varepsilon : L \Rightarrow F$. Since $\mathbb{C}^{\mathbf{2}}$ is the comma category $(\mathbb{C} \downarrow \mathbb{C})$, this gives rise to a functor $E : \mathbb{B} \to \mathbb{C}^{\mathbf{2}}$. Explicitly, E takes an object B to the arrow $F(\epsilon_B)$, and a morphism $b : B \to B'$ in $\mathbb{B}$ to the square:

$$\begin{tikzcd} F(B) \arrow[r, "F(b)"] & F(B') \\ L(B) \arrow[u, "F(\varepsilon_B)"] \arrow[r, "L(b)"'] & L(B') \arrow[u, "F(\varepsilon_{B'})"'] \end{tikzcd}$$

We will show that E is full and faithful. Consider a commutative square of the following form:

$$\begin{tikzcd} F(B) \arrow[r, "f"] & F(B') \\ L(B) \arrow[u, "F(\varepsilon_B)"] \arrow[r, "g"'] & L(B') \arrow[u, "F(\varepsilon_{B'})"'] \end{tikzcd} \tag{3.2.2}$$

For a morphism $b : B \to B'$, $g = L(b)$ if and only if the diagram below commutes:

$$\begin{tikzcd} B \arrow[r, "b"] & B' \\ TL(B) \arrow[u, "\varepsilon_B"] \arrow[r, "T(g)"'] & TL(B') \arrow[u, "\varepsilon_{B'}"'] \end{tikzcd} \tag{3.2.3}$$

Thus (3.2.2) is the image of b under E if and only if $F(b) = f$ and (3.2.3) commutes. By cocartesianness of ε_B, there is a unique such morphism b.

The image of E is nothing but a class of morphisms $\mathcal{C}'$. Let $\mathcal{C}$ be the closure of $\mathcal{C}'$ under right composition with isomorphisms. If we restrict E to the functor $E' : \mathbb{B} \to \mathcal{C}$, the functor E' will be an equivalence, since $\mathcal{C}$ is contained in the essential image of E. The fullness of T ensures that for every object C in $\mathbb{C}$, $\varepsilon_{T(C)}$ is an isomorphism. Hence so is $F(\varepsilon_{T(C)}) : LT(C) \to C$, and thus $\mathcal{C}$ will contain the identity on C (simply right compose with its inverse). Thus $\mathcal{C}$ contains all the identity morphisms. Assuming the axiom of choice, every equivalence is part of an adjoint equivalence, and it is easy to check that we can construct the adjunction (E', D, ϕ, ψ) such that it satisfies all the conditions required to make it an equivalence between F and $\mathsf{Cod}_{\mathcal{C}}$. $\square$

By *equivalence classes of functors*, we always mean equivalence classes of functors under the equivalence relation "F is equivalent to G". For convenience, we will denote condition (2) in the above theorem by (L):

(L) there exist adjunctions $(F, T, \eta, 1)$ and $(T, L, \alpha, \varepsilon)$ with T a full functor such that every component ε_B of ε is cocartesian with respect to F.

Theorem 5. *Let $\mathbb{C}$ be a category. Then there is a bijection between:*

(1) classes of morphisms which contain the identity morphisms and are closed under right composition with isomorphisms, and

(2) equivalence classes of functors $F : \mathbb{B} \to \mathbb{C}$ satisfying (L).

Proof. To each class $\mathcal{C}$ of morphisms we assign the equivalence class of the functor $\mathsf{Cod}_{\mathcal{C}}$ (all functors in this class satisfy (L) by Theorem 4), while to each equivalence class represented by a functor F we assign the closure under right composition with isomorphisms of the family $(F(\varepsilon_B))_{B \in \mathbb{B}}$, which we will denote by F^* (this is the class $\mathcal{C}$ from the proof of Theorem 4). We need to show that this is well-defined. Firstly, we, we need to show that F^* is invariant under the choice of adjoints T and L. A straightforward calculation shows that for any two chains of adjunctions $F \dashv T \dashv L$ and $F \dashv T' \dashv L'$, there is a natural isomorphism θ such that $\varepsilon' \circ \theta = \varepsilon$, where ε' and ε are the counits of $T' \dashv L'$ and $T \dashv L$ respectively. This gives the required result. Secondly, we need to show that the action $F \mapsto F^*$ is invariant under equivalence. Suppose F and

F' are two equivalent functors satisfying condition (L). By symmetry, it is enough to show that $F^* \subseteq F'^*$. Suppose (E, D, ϕ, ψ) witnesses the equivalence of F and F', and F' is part of a chain of adjoints $F' \dashv T' \dashv L'$, where ε' is the counit of $T' \dashv L'$. By the previous remark, we can calculate F^* using the chain $F \dashv D \circ T' \dashv L' \circ E$. The counit of this last adjunction is $D\varepsilon' E \circ \phi$. Consider a component $D\varepsilon'_{E(B)} \circ \phi_B$. Its image under F is $F'\varepsilon'_{E(B)}$, which is in F'^*, so we have $F^* \subseteq F'^*$ as required. Now all that remains is to show that this assignment gives a bijection. By Lemma 16 we have that $\mathsf{Cod}^*_{\mathcal{C}} = \mathcal{C}$, and the fact that Cod_{F^*} is equivalent to F follows from the proof of Theorem 4. □

It is now possible to combine Theorem 5 with other well-known results to obtain a correspondence between conditions on a class of morphisms and conditions on a functor F. For example, let $\mathcal{M}$ be a class of morphisms containing the identity morphisms. Then it follows from Lemma 15 that the following are equivalent:

- (M_1) and (M_2) (see Introduction) hold;
- $\mathsf{Cod}_{\mathcal{M}}$ is an opfibration.

It is easy to check that the property of being an opfibration is preserved under equivalence (it is not true in general that Grothendieck opfibrations are closed under composition with equivalences of categories, but the additional requirements in the definition of an equivalence of functors ensures this result). Also note that any class satisfying (M_1) and (M_2) contains the identity morphisms, and that under (M_2), condition (M_1) is equivalent to the weaker condition:

(M'_1) $\mathcal{M}$ is closed under right composition with isomorphisms.

Thus we immediately obtain:

Corollary 2. *Let $\mathbb{C}$ be a category. Then there is a bijection between:*

(1) classes of morphisms which satisfy (M_1) *and* (M_2)*, and*

(2) equivalence classes of opfibrations $F : \mathbb{B} \to \mathbb{C}$ satisfying (L)*.*

Again when $\mathcal{M}$ contains all the identity morphisms, the following are equivalent by Lemma 14:

- all pullbacks along morphisms in $\mathcal{M}$ exist and $\mathcal{M}$ is pullback-stable;
- $\mathsf{Cod}_{\mathcal{M}}$ is a fibration.

We thus obtain another bijection between:

- classes $\mathcal{M}$ of morphisms satisfying (M_1') such that all pullbacks along morphisms in $\mathcal{M}$ exist and $\mathcal{M}$ is pullback-stable, and
- equivalence classes of fibrations $F : \mathbb{B} \to \mathbb{C}$ satisfying (L).

3.3 Forms of $\mathcal{M}$-subobjects

We will call a functor F a *form* when it is faithful and amnestic. In particular, the fibres of such a functor will be (partially) ordered sets. Note that in the original definition in [58], the term "form" was reserved for faithful amnestic functors which are also bifibrations. In this section we borrow terminology freely from [59].

Definition 30. *A form $F : \mathbb{B} \to \mathbb{C}$ is called* locally bounded above *if every fibre $F^{-1}(C)$ has a terminal object (i.e. a top element), which we denote by $T(C)$.*

Definition 31. *Given a form $F : \mathbb{B} \to \mathbb{C}$ which is locally bounded above, a* left universalizer *[58] of an object B in $\mathbb{B}$, denoted by $\mathsf{lun}(B) : \mathsf{Lun}(B) \to F(B)$, is a terminal morphism in $\mathbb{C}$ among those morphisms $f : A \to F(B)$ in $\mathbb{C}$ with the property that there exists a morphism f' from $T(A)$ to B with $F(f') = f$.*

In other words, for an object B with a left universaliser $\mathsf{lun}(B) : \mathsf{Lun}(B) \to F(B)$, there is a (unique) morphism $\varepsilon_B : T(\mathsf{Lun}(B)) \to F(B)$ with $F(\varepsilon_B) = \mathsf{lun}(B)$ and moreover, for any other morphism $f : T(C) \to B$, where C is an object of $\mathbb{C}$, there is a morphism i such that $F(\varepsilon_B) \circ i = F(f)$. Consider the following condition on a form F:

(LE) for every object C and morphism $f : C \to C'$ in $\mathbb{C}$, if the fibre of C has a terminal object $T(C)$, then there is a (unique) cocartesian lifting of f from $T(C)$, whose codomain we denote by $f \cdot 1$.

The object $f \cdot 1$ was called the *left norm of f* in [58]. When we define something to be $f \cdot 1$, we are in particular implying that it (and the terminal object in fibre at the domain of f) exists. For a form F satisfying (LE), the notion of left universaliser can be suggestively rephrased:

Lemma 18. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form which is locally bounded above and which satisfies* (LE). *Then the assignment $C \mapsto T(C)$ extends to a functor in a unique way such that $F \circ T = 1_{\mathbb{C}}$. An object B has a left universaliser if and only if it has a coreflection along T.*

Proof. For a morphism $f : C \to C'$, define $T(f)$ to be the composite of the cocartesian lifting of f from $T(C)$ and the unique morphism in the fibre at C' from $f \cdot 1$ to $T(C')$. Since F is faithful, this is the unique way to define T, and the uniqueness ensures that T is a functor (moreover, T will be full and faithful). The second part is now easy to verify. □

The following proposition follows easily from the lemma above:

Proposition 21. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form satisfying* (LE). *Then*

(1) there is an adjunction $(F, T, \eta, 1)$ if and only if F is locally bounded above (where 1 is the identity natural transformation on $1_{\mathbb{C}}$), and

(2) when this is the case, then there is an adjunction $(T, L, \alpha, \varepsilon)$ if and only if every object B in $\mathbb{B}$ has a left universalizer.

Note that when both (1) and (2) in Proposition 21 hold, condition (LE) will ensure that F is an opfibration, since every object is the image of a cocartesian morphism with domain a terminal object. The following lemma gives a useful fact about cocartesian morphisms for faithful functors:

Lemma 19. *Let $F : \mathbb{B} \to \mathbb{C}$ be a faithful functor and let f and g be a composable pair of morphisms in $\mathbb{B}$. Then g is cocartesian if $g \circ f$ is.*

Definition 32. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form which is locally bounded above. We say that an object B is* conormal *if it is equal to $f \cdot 1$ for some morphism f in $\mathbb{C}$, i.e. it is the codomain of a cocartesian morphism with domain a terminal object. We say that a form $F : \mathbb{B} \to \mathbb{C}$ which is locally bounded above is* conormal *if every object B in $\mathbb{B}$ is conormal.*

Lemma 20. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form which is locally bounded above, and let B be an object of $\mathbb{B}$ which has a left universaliser $\mathsf{lun}(B)$. Then B is conormal if and only if $\mathsf{lun}(B) \cdot 1 = B$.*

Proof. This is a straightforward application of Lemma 19. □

Two forms $F : \mathbb{C} \to \mathbb{C}$ and $F' : \mathbb{B}' \to \mathbb{C}$ are equivalent as functors if and only if they are *isomorphic*, i.e. there exists an isomorphism I such that $F'I = F$. Equivalence classes of forms will thus be referred to as *isomorphism classes*. Notice that every faithful functor is equivalent to a form – we simply identify isomorphic objects in each fibre. Moreover, cartesian and cocartesian liftings are preserved by this process by Lemma 17.

When $\mathcal{C}$ is a class of morphisms containing the identity morphisms, it is easy to check that $\mathsf{Cod}_{\mathcal{C}}$ is faithful if and only if every morphism in $\mathcal{C}$ is a monomorphism. For a class $\mathcal{M}$ of monomorphisms, we will call the form which is equivalent to $\mathsf{Cod}_{\mathcal{M}}$ the *form of $\mathcal{M}$-subobjects*, and denote it by

$$\overline{\mathsf{Cod}_{\mathcal{M}}} : \overline{\mathcal{M}} \to \mathbb{C}.$$

We say that a form $F : \mathbb{B} \to \mathbb{C}$ which is locally bounded above *admits left universalisers* when every object B in $\mathbb{B}$ has a left universaliser. Combining the results in this section with Theorems 4 and 5 (and noting that a right inverse T of F will be full whenever F is faithful), we obtain the following two corollaries:

Corollary 3. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form. Then there exists a class $\mathcal{M}$ of monomorphisms containing the identity morphisms such that the form of $\mathcal{M}$-subobjects is isomorphic to F if and only if F is locally bounded above, conormal and admits left universalisers.*

Corollary 4. *Let $\mathbb{C}$ be a category. Then there is a bijection between:*

(1) classes of monomorphisms which satisfy (M_1) *and* (M_2)*, and*

(2) isomorphism classes of faithful amnestic opfibrations $F : \mathbb{B} \to \mathbb{C}$ which are locally bounded above, admit left universalisers and are conormal.

Under this bijection, each class $\mathcal{M}$ is assigned to the form of $\mathcal{M}$-subobjects, while to each form F we assign the class of all left universalisers for F.

This second corollary is Theorem 2.4 from [59]. The approach to this result in [59], however, goes via a more general correspondence between faithful amnestic opfibrations which are locally bounded above and conormal (but which do not necessarily admit left universalisers) and reflexive and transitive cover relations [58].

Recall that a category $\mathbb{C}$ is *finitely $\mathcal{M}$-complete* in the sense of Dikranjan and Tholen [28] when it satisfies (M_1) and (M_2) and all pullbacks along morphisms in $\mathcal{M}$ exist (the class $\mathcal{M}$ will then be pullback-stable). Thus the bijection in Corollary 4 restricts to a bijection between

- classes $\mathcal{M}$ of monomorphisms such that $\mathbb{C}$ is finitely $\mathcal{M}$-complete, and
- isomorphisms classes of faithful amnestic bifibrations (called *biforms*) over $\mathbb{C}$ which are locally bounded above, conormal and admit left universalizers;

It is also well known that a class $\mathcal{M}$ satisfying (M_1) and (M_2) is part of a (uniquely determined) factorisation system $(\mathcal{E}, \mathcal{M})$ in the sense of Freyd and Kelly [37] if and only if $\mathcal{M}$ is closed under composition. Thus we obtain a bijection between

- classes $\mathcal{M}$ of monomorphisms which are part of a factorization system $(\mathcal{E}, \mathcal{M})$, and
- isomorphisms classes of left forms over $\mathbb{C}$ which are locally bounded above, conormal, admit left universalizers, and for which the class of left universalizers is closed under composition.

As a direct corollary of Proposition 1.4.20 in [58] and the bijection between cover relations and forms established in [59], we obtain that this last condition, namely that the class of left universalisers is closed under composition, is equivalent to the following one:

- for every left universaliser $m : M \to Y$, the push-forward functor $m^* : F^{-1}(M) \to F^{-1}(Y)$ is full.

We should remark that the results obtained in these last two sections are closely related to the work in [46], where only the case when $\mathcal{M}$ is a class of morphisms which is part of a factorisation system $(\mathcal{E}, \mathcal{M})$ is considered.

3.4 Forms of $\mathcal{N}$-exact sequences

A class $\mathcal{N}$ of morphisms in a category $\mathbb{C}$ is called an *ideal of null morphisms* [31; 70; 71] if it satisfies the following condition:

(N) for any triple of morphisms $f : A \to B$, $g : B \to C$ and $h : C \to D$, $g \in \mathcal{N} \Rightarrow hgf \in \mathcal{N}$.

Definition 33. *Let $\mathbb{C}$ be a category and $\mathcal{N}$ be an ideal of null morphisms. Then an $\mathcal{N}$-kernel of a morphism $f : A \to B$ is a morphism $k : K \to A$ such that $fk \in \mathcal{N}$ and such that for any other morphism $k' : K' \to A$ such that $fk' \in \mathcal{N}$, there is a unique morphism i such that $ki = k'$. An $\mathcal{N}$-cokernel of a morphism g is defined dually.*

A pair (f, g) of composable morphisms (i.e. for which $g \circ f$ is defined) will be called $\mathcal{N}$-short-exact when f is the $\mathcal{N}$-kernel of g and g is the $\mathcal{N}$-cokernel of f. Let $\mathcal{P}$ be a class of composable pairs of morphisms in a category $\mathbb{C}$. We consider some conditions on $\mathcal{P}$ inspired by the example of $\mathcal{N}$-short-exact sequences for an ideal $\mathcal{N}$:

(P_1) for any two pairs (f, g) and (f', g') in $\mathcal{P}$ and any morphism h as shown in the diagram below, there is a bijection between morphisms i such that $hf = f'i$ and morphisms j such that $jg = g'h$:

$$\begin{tikzcd} \cdot \arrow[r, "f"] \arrow[d, dotted, "i"'] & \cdot \arrow[r, "g"] \arrow[d, "h"] & \cdot \arrow[d, dotted, "j"] \\ \cdot \arrow[r, "f'"'] & \cdot \arrow[r, "g'"'] & \cdot \end{tikzcd}$$

(P_2) for every object C in $\mathbb{C}$, there are composable pairs $(1_C, c)$ and $(k, 1_C)$ in $\mathcal{P}$, with c an epi and k a mono.

Lemma 21. *Conditions* (P_1) *and* (P_2) *together imply the following condition:*

(P_3) *for every pair (f, g) in $\mathcal{P}$, f is mono and g is epi.*

Proof. Suppose (P_1) and (P_2) hold and let (f, g) be a pair in $\mathcal{P}$. Let α and β be two morphisms such that $f\alpha = f\beta$. Consider the diagram:

$$\begin{tikzcd} {} \arrow[r, equal] \arrow[d, dotted, "\alpha"'] & {} \arrow[r, "c"] \arrow[d, "f\circ\alpha"] & {} \arrow[d, dotted, "j"] \\ {} \arrow[r, "f"'] & {} \arrow[r, "g"'] & {} \end{tikzcd}$$

where $(1, c)$ is in $\mathcal{P}$ and c is epi. Then there is at most one morphism j such that $jc = gh$, which means there is at most one morphism i such that $fi = f\alpha = f\beta$, which gives $\alpha = \beta$. The fact that g is epi follows dually. $\square$

Given any class $\mathcal{P}$ of composable pairs of morphisms, we can consider $\mathcal{P}$ as a full subcategory (which we also denote by $\mathcal{P}$) of the category $\mathbb{C}^{\mathbf{2}} \times_{\mathbb{C}} \mathbb{C}^{\mathbf{2}}$ of composable pairs of morphisms in $\mathbb{C}$. We can then consider the restricted "midpoint" functor $\mathsf{Mid}_{\mathcal{P}} : \mathcal{P} \to \mathbb{C}$ which takes a composable pair $(f : A \to B, g : B \to C)$ to B. When (P_3) holds, this functor is faithful, and so we can produce a form

$$\overline{\mathsf{Mid}_{\mathcal{P}}} : \overline{\mathcal{P}} \to \mathbb{C}$$

by identifying isomorphic objects in each fibre. We call this form the *form of $\mathcal{P}$-sequences*. We now give definitions for the duals of some of the notions which appeared in the previous chapter.

Definition 34. *A form $F : \mathbb{B} \to \mathbb{C}$ is called* locally bounded below *if every fibre $F^{-1}(C)$ has an initial object (i.e. a bottom element), which we denote by $I(C)$. A form F which is locally bounded below and above will be called* locally bounded.

Definition 35. *Given a form $F : \mathbb{B} \to \mathbb{C}$ which is locally bounded below, a* right universalizer *[58] of an object B in $\mathbb{B}$, denoted by $\mathsf{run}(B) : F(B) \to \mathsf{Run}(B)$, is an initial morphism in $\mathbb{C}$ among those morphisms $f : F(B) \to C$ in $\mathbb{C}$ with the property that there exists a morphism f' from B to $I(C)$ with $F(f') = f$.*

For a form which is locally bounded below, we denote the domain of the cartesian lifting of a morphism $f : A \to B$ to $I(B)$ by $0 \cdot f$ (when it exists).

Definition 36. *Let $F : \mathbb{B} \to \mathbb{C}$ be a form which is locally bounded below. We say that an object B is* normal *if it is equal to $0 \cdot f$ for some morphism f in $\mathbb{C}$. We say that a form $F : \mathbb{B} \to \mathbb{C}$ which is locally bounded below is* normal *if every object B in $\mathbb{B}$ is normal.*

We also have a notion which is dual to the notion of form of $\mathcal{M}$-subobjects. Given a class $\mathcal{E}$ of morphisms, we can consider the domain functor $\mathsf{Dom}_{\mathcal{E}} : \mathcal{E} \to \mathbb{C}$. If $\mathcal{E}$ is a class of epis, this functor is equivalent to a form

$$\overline{\mathsf{Dom}_{\mathcal{E}}} : \overline{\mathcal{E}} \to \mathbb{C},$$

which we call the *form of $\mathcal{E}$-quotients*.

Theorem 6. *The following are equivalent for any form $F : \mathbb{B} \to \mathbb{C}$:*

(1) there exists a class $\mathcal{P}$ of pairs of composable morphisms satisfying (P_1) *and* (P_2) *such that F is isomorphic to the form of $\mathcal{P}$-sequences;*

(2) F satisfies (L) *and its (functorial) dual;*

Proof. (1) $\Rightarrow$ (2): For a class $\mathcal{P}$ of composable pairs of morphisms, let $\mathcal{M}$ (resp. $\mathcal{E}$) be the class of all morphisms which appear as a first (resp. second) component in some element of $\mathcal{P}$, and let $\pi_1 : \mathcal{P} \to \mathcal{M}$ and $\pi_2 : \mathcal{P} \to \mathcal{E}$ be the first and second projection functors respectively. Conditions (P_1) and (P_2) ensure that both π_1 and π_2 are equivalences, and that both $\mathcal{E}$ and $\mathcal{M}$ contain the identity morphisms. Moreover, π_1 and π_2 form part of equivalences between the functor $\mathsf{Mid}_{\mathcal{P}}$ and $\mathsf{Cod}_{\mathcal{M}}$ and $\mathsf{Dom}_{\mathcal{E}}$ respectively. Thus the form of $\mathcal{P}$-sequences is isomorphic to $\overline{\mathsf{Cod}_{\mathcal{M}}}$ and $\overline{\mathsf{Dom}_{\mathcal{E}}}$, and hence satisfies (L) and its dual by Corollary 3 and the dual result.

(2) $\Rightarrow$ (1): By Corollary 3 (and the dual result), we have isomorphisms $I : \mathbb{B} \to \overline{\mathcal{M}}$ and $J : \mathbb{B} \to \overline{\mathcal{E}}$ for some classes $\mathcal{M}$ (of monos) and $\mathcal{E}$ (of epis) such that $M \circ I = F = E \circ J$, where M and E are the forms of $\mathcal{M}$-subobjects and $\mathcal{E}$-quotients respectively. This gives rise to a functor $H : \mathbb{B} \to \overline{\mathcal{M}} \times_{\mathbb{C}} \overline{\mathcal{E}}$, where $\overline{\mathcal{M}} \times_{\mathbb{C}} \overline{\mathcal{E}}$ is the pullback of M along E. While H may not be an isomorphism, it is easy to check that it is full and faithful (the fact that F itself is faithful plays an important role here). Thus restricting H to its image, it becomes an isomorphism. The image of H is nothing but a class of composable pairs of (isomorphism classes of) morphisms. In other words, H is isomorphic to a form of $\mathcal{P}$-sequences for some $\mathcal{P}$. Since I and J are isomorphisms, $\pi_1 : \overline{\mathcal{M}} \times_{\mathbb{C}} \overline{\mathcal{E}} \to \overline{\mathcal{M}}$ and $\pi_2 : \overline{\mathcal{M}} \times_{\mathbb{C}} \overline{\mathcal{E}} \to \overline{\mathcal{E}}$ will be faithful, so $\mathcal{P}$ satisfies condition (P_1). Condition (P_2) is satisfied since $\mathcal{M}$ and $\mathcal{E}$ both contain the identities and are classes of monos and epis respectively.

□

Consider the following condition on a class $\mathcal{P}$ of pairs of composable morphisms:

(P_0) if $(f, g) \in \mathcal{P}$, then $(fu, vg) \in \mathcal{P}$ for any isomorphisms u and v such that the composites fu and vg are defined.

A class $\mathcal{P}$ of morphisms satisfying (P_0), (P_1) and (P_2) is completely determined by the class $\mathcal{M}$ of those morphisms which appear as a first component in $\mathcal{P}$, since every morphism m in $\mathcal{M}$ is part of exactly one pair (m, e), up to left composition of e by an isomorphism. This allows us to restrict the bijection given in Theorem 5 to obtain:

Corollary 5. *Let $\mathbb{C}$ be a category. Then there is a bijection between:*

- *classes of pairs of composable morphisms $\mathcal{P}$ satisfying* (P_0)*,* (P_1) *and* (P_2)*.*

- *isomorphism classes of forms F which satisfy* (L) *and its (functorial) dual.*

Under this bijection, we assign to $\mathcal{P}$ the form of $\mathcal{P}$-sequences, while to a form F we assign the class of all pairs (m, e) such that m is a left universaliser and e is a right universaliser of the same object B.

Any class $\mathcal{P}$ satisfying (P_1) and (P_2) determines an ideal $\mathcal{N}$ of null morphisms relative to which identity morphisms have cokernels and kernels. Indeed, define the ideal as follows: a morphism $f : A \to B$ is in $\mathcal{N}$ if $f = mc$ for some morphisms m and c with $(1_A, c) \in \mathcal{P}$. Clearly, by condition (P_1), a morphism f satisfies this condition if and only if it satisfies the dual one, i.e. $f = ke$, for some morphisms k and e with $(k, 1_B) \in \mathcal{P}$. Using these two equivalent definitions, one easily sees that the class $\mathcal{N}$ is an ideal of null morphisms. The kernel and cokernel of an identity morphism 1_C are simply the morphisms k and c occuring in (P_2) respectively.

We can also describe this construction in terms of the corresponding form of $\mathcal{P}$-sequences P. A morphism f is defined to be in $\mathcal{N}$ precisely when it fits inside a diagram of the following form:

$$\begin{tikzcd} {} \arrow[r, equal] \arrow[d, "e"'] & {} \arrow[r, "c"] \arrow[d, "f"] & {} \arrow[d, "m"] \\ {} \arrow[r, "d"'] & {} \arrow[r, equal] & {} \end{tikzcd}$$

where the top and bottom pairs are in $\mathcal{P}$. In fact, the top (resp. bottom) pair is a representative of the terminal (resp. initial) objects in the fibre of P at the domain (resp. codomain) of f. Thus a morphism $f : A \to B$ is in $\mathcal{N}$ if and only if it is the image of a morphism $f' : T(A) \to I(B)$ from a terminal object to an initial object. For an arbitrary form F, we call such a morphism f an *F-null morphism.*

Lemma 22. *Let $\mathcal{P}$ be a class of composable pairs of morphisms satisfying* (P_1) *and* (P_2) *and let $\mathcal{N}$ be the corresponding ideal of null morphisms. Then any pair (f, g) in $\mathcal{P}$ constitutes a short exact sequence relative to $\mathcal{N}$.*

Proof. By duality, it is enough to show that f is the kernel of g. Suppose gh is in $\mathcal{N}$ for a morphism h. Then we have the following diagram of solid arrows:

$$\begin{tikzcd} {} \arrow[r, equal] \arrow[d, dotted, "i"'] & {} \arrow[r, "c"] \arrow[d, "h"] & {} \arrow[d] \\ {} \arrow[r, "f"'] & {} \arrow[r, "g"'] & {} \end{tikzcd}$$

where gh factors through c and $(1, c)$ is in $\mathcal{P}$ by definition. But then, by (P_1), we have a morphism i shown by the dotted arrow which makes the diagram commute, and moreover, it is the unique such since c is epi. $\square$

Given a ideal of null morphisms $\mathcal{N}$, we can consider the class $\mathsf{Ex}(\mathcal{N})$ of all short exact sequences relative to $\mathcal{N}$. We will call the form of $\mathsf{Ex}(\mathcal{N})$-sequences the *form of $\mathcal{N}$-short-exact sequences.*

Theorem 7. *For a given category $\mathbb{C}$, there is a bijection between:*

(1) ideals of null morphisms $\mathcal{N}$ in $\mathbb{C}$ admitting all cokernels and kernels;

(2) isomorphism classes of binormal biforms $F : \mathbb{B} \to \mathbb{C}$ which are locally bounded above and below, and admit all left and right universalisers;

Under this bijection, each ideal $\mathcal{N}$ is assigned to the form of $\mathcal{N}$-short-exact-sequences, while to each form F we assign the class of all F-null morphisms.

Proof. By Corollary 5, forms satisfying the conditions in (2) are in bijection with classes $\mathcal{P}$ satisfying conditions (P_0)-(P_2) and whose form of $\mathcal{P}$-sequences is a bifibration. We claim that such classes are in bijection with ideals $\mathcal{N}$ admitting all kernels and cokernels. One direction of the bijection will be given by assigning to an ideal $\mathcal{N}$ the class of $\mathcal{N}$-short-exact sequences, while the other direction will be the construction of an ideal from a class $\mathcal{P}$ given earlier. Given an ideal $\mathcal{N}$, $\mathcal{N}$ is clearly equal to the ideal corresponding to the class of $\mathcal{N}$-short-exact sequences (a morphism is in $\mathcal{N}$ if and only if it factors through a cokernel of an identity). Now let $\mathcal{P}$ be a class of composable pairs of morphisms satisfying the conditions and consider the biform of $\mathcal{P}$-sequences. It is easy to show that if f is a morphism in $\mathbb{C}$, then the morphism shown by the diagram below is cartesian if and only if m is the $\mathcal{N}$-kernel of f.

$$\begin{tikzcd} \cdot \arrow[r, "m"] \arrow[d] & \cdot \arrow[r, "e"] \arrow[d, "f"] & \cdot \arrow[d] \\ \cdot \arrow[r, "k"'] & \cdot \arrow[r, equal] & \cdot \end{tikzcd}$$

Thus the corresponding ideal $\mathcal{N}$ admits all kernels and cokernels. We have by Lemma 22 that $\mathcal{P}$ is contained in $\mathsf{Ex}(\mathcal{N})$. On the other hand, given any $\mathcal{N}$-short-exact sequence $(k : W \to X, e : X \to Y)$, the first component of the cartesian lifting of e to the pair $(k', 1_Y) \in \mathcal{P}$ must be an $\mathcal{N}$-kernel of e by the above remark. Thus (k, e) must be in $\mathcal{P}$ by condition (P_0). This shows that every $\mathcal{N}$-short-exact sequence is in $\mathcal{P}$, so $\mathcal{P} = \mathsf{Ex}(\mathcal{N})$. Composing this bijection with the one given earlier in the proof gives the required result. $\square$

A consequence of Corollay 4 and it dual result is that there is a bijection between isomorphism classes of forms satisfying the conditions in (2) in Theorem 7 above, and pairs $(\mathcal{E}, \mathcal{M})$, where $\mathcal{M}$ is a class of monomorphisms satisfying (M_1) and (M_2), and $\mathcal{E}$ is a class of epimorphisms satisfying the dual conditions, such that the form of $\mathcal{M}$-subobjects is isomorphic to the form of $\mathcal{E}$-quotients. Composing with the bijection in Theorem 7, we obtain that an ideal $\mathcal{N}$ admitting all kernels and cokernels is the same as such a pair $(\mathcal{E}, \mathcal{M})$, where by "is the same as" we mean that there is a bijection between these two types of structures – namely, for an ideal $\mathcal{N}$ the corresponding pair $(\mathcal{E}, \mathcal{M})$ consists of the class $\mathcal{E}$ of $\mathcal{N}$-cokernels and the class $\mathcal{M}$ of $\mathcal{N}$-kernels.

Theorem 7 is also contained in [59]. The approach to the result there goes via a correspondence between ideals $\mathcal{N}$ for which every morphism $f : B \to C$ in $\mathbb{C}$ is part of an $\mathcal{N}$-exact (not necessarily short exact) sequence

$$A \xrightarrow{g} B \xrightarrow{f} C \xrightarrow{h} D$$

(with the notion of $\mathcal{N}$-exact suitably defined) and binormal biforms which are locally bounded (but do not necessarily admit left and right universalisers).

A pair $(\mathbb{C}, \mathcal{N})$, where $\mathbb{C}$ is a category and $\mathcal{N}$ is an ideal of null morphisms in $\mathbb{C}$, is an ex2-category in the sense of Grandis [41] if $\mathbb{C}$ admits all cokernels and kernels relative to $\mathcal{N}$ and moreover these cokernels and kernels are closed under composition. Combining Theorem 7 above with the remark at the end of the last section, we see that for a given category $\mathbb{C}$ there is a bijection between:

- ideals of null morphisms $\mathcal{N}$ such that $(\mathbb{C}, \mathcal{N})$ constitutes an ex2-category, and

- isomorphism classes of binormal biforms F to $\mathbb{C}$ which are locally bounded, admit all left and right universalisers and where for every left/right universaliser, the push-forward/pull-back functor induced by it is full.

It is worth noting that throughout this section, we have only dealt with faithful functors, i.e. forms, whereas in the previous sections we were able to sometimes obtain bijections for more general functors. Since the generalisation of pointed categories to categories equipped with an ideal has led to developments in non-abelian homological algebra, it would be interesting to see if a notion of homological algebra could be developed with the form of $\mathcal{N}$-exact sequences replaced by a non-faithful functor satisfying similar axioms.

3.5 Concluding remarks

Let F be a biform, and consider the following weaker versions of condition (L) and its dual:

(L') F is locally bounded above and admits left universalisers.

(R') F is locally bounded below and admits right universalisers.

A biform F satisfying both of these conditions is precisely one which fits into the following picture:

$$
\begin{tikzcd}
\mathbb{B} \arrow[d, bend right=60, "R"'] \arrow[d, bend right=25, leftarrow, "I"] \arrow[d, "F"] \arrow[d, bend left=25, leftarrow, "T"'] \arrow[d, bend left=60, "L"] \\
\mathbb{C}
\end{tikzcd}
$$

where the counit of $F \dashv T$ and the unit of $I \dashv F$ are the identity (and as a result, the unit of $T \dashv L$ and the counit of $R \dashv I$ are isomorphisms). This structure, which is simplicial in character, is similar to the *pointed combinatorial exactness structures* considered in [52]. In the case of pointed combinatorial structures, however, the categories $\mathbb{C}$ and $\mathbb{B}$ are replaced with sets.

The most important case of a form which we will consider in this thesis is that of the *form of subobjects* in a category $\mathbb{C}$, i.e. the form of $\mathcal{M}$-subobjects with $\mathcal{M}$ the class of all monomorphisms in $\mathbb{C}$. The guiding example of such a form is the form of subgroups over the category of groups. In this case, the functors in the simplicial picture above represent well-known notions in group theory. Indeed, T and I take a group G to G and the zero group respectively (seen as subgroups of G), while L and R take a subgroup S of G to S and the quotient group G/S respectively (in case S is not normal, this is the quotient of G by the normal closure of S). The idea (due to Z. Janelidze) behind the work in [59; 60; 61], as well as in this thesis, is to use the context of such a simplicial structure, together with some additional axioms, to establish isomorphism theorems and other results for groups in a self-dual way. This simplicial stucture was mentioned in a talk by Z. Janelidze [57], but the description there contained a mistake, which the work presented in this chapter fixes. Forms of subobjects of group-like structure will be further studied in the subsequent chapters.

Chapter 4

Exact forms

Adapted from: Z. Janelidze and T. Weighill, Duality in non-abelian algebra II. From Isbell bicategories to Grandis exact categories (submitted; preprint at http://math.sun.ac.za/cms/).

4.1 Introduction

In the previous chapter, we defined the notion of a form and the related notions of left and right universalisers, amongst others. The idea behind the development of the theory of forms is to define a self-dual axiomatic context in which isomorphism theorems and diagram lemmas for non-abelian algebras such as groups can be established. The next step should thus be to organise a suitable list of self-dual axioms which capture important aspects of the category of groups. In this chapter we look at one such axiom (axiom (E) below), and the corresponding notion of *exact form*.

The First Isomorphism Theorem for groups states that every group homomorphism $f : G \to H$ factorises as follows:

$$\begin{tikzcd} G \arrow[rrr, "f"] \arrow[dr, "e"'] & & & H \\ & G/\mathsf{ker}(f) \arrow[r, "\theta"'] & \mathsf{Im}(f) \arrow[ur, "m"'] & \end{tikzcd}$$

where m is the inclusion of the image of f into H, θ is an isomorphism, and e is the evident quotient map. This can be expressed categorically as the following statement about the category **Grp** of all groups: every morphism f factorises as a normal epi e followed by a monomorphism m. The dual of this statement does not hold in **Grp** (it does, however, hold in the category **Ab** of abelian groups).

Recalling the notation and terminology of the previous chapter, we can state the First Isomorphism Theorem for groups in a self-dual way. Let $\mathbb{C}$ be the category of groups and let $F : \mathbb{B} \to \mathbb{C}$ be the bifibration of subgroups (i.e. the form of subobjects over **Grp**). Then it turns out that the First Isomorphism Theorem can be expressed as follows:

(E) every morphism f in $\mathbb{C}$ in factorises as $f = \mathsf{lun}(f \cdot 1) \circ \mathsf{run}(0 \cdot f)$.

In this chapter, we examine this statement, which we consider as a condition on a form F, as well as some weaker conditions. These conditions turn out to be closely related to the notion of prefactorisation and factorisation systems [37]. They also turn out to classify, via the correspondence established in the previous chapter, various exactness axioms considered by Grandis in [41; 42; 43].

4.2 Preliminaries

We begin by introducing some new notation suitable for the type of forms we will be dealing with in this chapter. In general, these forms will always be locally bounded, but will not admit all left and right universalisers, and will not be conormal or normal. Nonetheless, many important results carry over from the previous chapter (see for example Lemma 27 later in this section).

Let $F : \mathbb{B} \to \mathbb{C}$ be a form. For an object X of $\mathbb{C}$, we write 1^X (and sometimes simply 1) for the upper bound of $F^{-1}(X)$, when it exists, and 0^X (or simply 0) for the lower bound, again when it exists. When a morphism $f : X \to Y$ has a cocartesian lifting at an object $A \in F^{-1}(X)$, the codomain of this cocartesian lifting will be denoted by fA (and sometimes by $f \cdot A$). Dually, when f has a cartesian lifting at $B \in F^{-1}(Y)$, the domain of this lifting will be denoted by Bf (and sometimes by $B \cdot f$). When we say that Bf (or fA) *is defined* we are making a claim/assumption that the object which it is supposed to represent exists, i.e. the (co)cartesian lifting of f at B (at A) exists. In particular, for a locally bounded form, $f \cdot 1$ and $0 \cdot f$ are as defined in the previous chapter. The following lemmas are consequences of standard properties of (co)cartesian liftings of morphisms:

Lemma 23. *Consider two morphisms $f : X \to Y$ and $g : Y \to Z$, and an object $A \in F^{-1}(X)$. If $f \cdot A$ is defined, then $g \cdot (f \cdot A)$ is defined if and only if $(g \circ f) \cdot A$ is defined, and $g \cdot (f \cdot A) = (g \circ f) \cdot A$ when they are defined. Dually, for any object $B \in F^{-1}(Z)$, if $B \cdot g$ is defined then $(B \cdot g) \cdot f$ is defined if and only if $B \cdot (g \circ f)$ is defined, and $(B \cdot g) \cdot f = B \cdot (g \circ f)$ when they are defined.*

Lemma 24. *Consider a morphism $f : X \to Y$ and two objects $A_1, A_2 \in F^{-1}(X)$. If both $f \cdot A_1$ and $f \cdot A_2$ are defined, then $A_1 \leqslant A_2$ implies $f \cdot A_1 \leqslant f \cdot A_2$. Dually, for any two objects $B_1, B_2 \in F^{-1}(Y)$, if both $B_1 \cdot f$ and $B_2 \cdot f$ are defined, then $B_1 \leqslant B_2$ implies $B_1 \cdot f \leqslant B_2 \cdot f$.*

Henceforth in this section we work in a category $\mathbb{C}$ equipped with a form F which is *bounded* in the following sense:

Definition 37. *A form F over a category $\mathbb{C}$ is said to be* bounded *when it is locally bounded and for any morphism $f : X \to Y$ in $\mathbb{C}$, both $f \cdot 1$ and $0 \cdot f$ are defined.*

In other words, a bounded form is a locally bounded form satisfying condition (LE) and its dual from the previous chapter. Note that the notion of a bounded form is *self-dual*: a form $F : \mathbb{B} \to \mathbb{C}$ is bounded if and only if the *dual form* $F^{\mathsf{op}} : \mathbb{B}^{\mathsf{op}} \to \mathbb{C}^{\mathsf{op}}$ is bounded. Also note that in a bounded form, left and right universalisers can be defined equivalently as follows: a left universalizer of an object $B \in F^{-1}(Y)$ is a morphism $f : X \to Y$ which is terminal with the property that $f \cdot 1^X \leqslant B$, while a right universalizer of an object $A \in F^{-1}(X)$ is a morphism $f : X \to Y$ which is initial with the property that $A \leqslant 0^Y \cdot f$. The following lemma is easy to prove.

Lemma 25. *Any left universalizer f of an object $B \in F^{-1}(Y)$ is necessarily a monomorphism, and f' is another left universalizer of the same object B if and only if $f' = fi$ for some isomorphism i. Dually, a right universalizer of an object $A \in F^{-1}(X)$ is an epimorphism, and if f is a right universalizer of A then f' is also a right universalizer of A if and only if $f' = jf$ for a isomorphism j.*

The proof of the above lemma uses the fact that if $f : X \to Y$ is an isomorphism then $f \cdot 1^X = 1^Y$ (and dually, $0^Y \cdot f = 0^X$). In general, a morphism $f : X \to Y$ such that $f \cdot 1^X = 1^Y$ will be called a *thick morphism*. Dually, when $0^Y \cdot f = 0^X$ we say that f is *thin*.

A class $\mathcal{C}$ of morphisms is said to be a *left class* if it has the following properties:

- $\mathcal{C}$ contains all identity morphisms.
- $\mathcal{C}$ is closed under composition.
- If $fg \in \mathcal{C}$ then $g \in \mathcal{C}$.

Dually, $\mathcal{C}$ is said to be a *right class* if the first two conditions above hold, and $fg \in \mathcal{C}$ always implies $f \in \mathcal{C}$. Notice that the class of split monomorphisms is the smallest left class, and the class of split epimorphisms is the smallest right class. Note also that a left/right class always contains all isomorphisms. The class of all mono(/epi)morphisms is another example of a left(/right) class.

Lemma 26. *The class of all thin morphisms is a left class, and dually, the class of all thick morphisms is a right class.*

Recall that an object $A \in F^{-1}(X)$ is said to be *normal* when $A = 0^Y \cdot f$ for some morphism $f : X \to Y$ in $\mathbb{C}$, and that, dually, an object $B \in F^{-1}(Y)$ is said to be *conormal* when $B = f \cdot 1^X$ for some morphism $f : X \to Y$.

Lemma 27. *Consider a left universalizer $f : X \to Y$ of an object $B \in F^{-1}(Y)$. Then:*

- *f is also a left universalizer of $f \cdot 1^X$.*
- *B is conormal if and only if $f \cdot 1^X = B$.*
- *f is an isomorphism if and only if it is thick.*

Dually, if f is a right universalizer of an object $A \in F^{-1}(X)$, then f is also a right universalizer of $0^Y \cdot f$ and A is normal if and only if $0^Y \cdot f = A$, and finally, f is an isomorphism if and only if it is thin.

A morphism $f : X \to Y$ is said to be an *embedding* if for any two conormal objects $A_1, A_2 \in F^{-1}(X)$ we have

$$f \cdot A_1 \leqslant f \cdot A_2 \quad \Rightarrow \quad A_1 \leqslant A_2.$$

Dually, f is a *coembedding* if for any two normal objects $B_1, B_2 \in F^{-1}(Y)$ we have

$$A_1 \cdot f \leqslant A_2 \cdot f \quad \Rightarrow \quad A_1 \leqslant A_2.$$

Lemma 28. *The class of all (co)embeddings is a left (right) class.*

4.3 The general theory

We recall from [37] (see also [21], and the Background section in this thesis) some notation and some very basic notions and results from the theory of factorization systems. A morphism e is said to be *orthogonal* to a morphism m, written as $e \downarrow m$, if any commutative square of solid arrows

$$\begin{tikzcd} \bullet \arrow[r, "e"] \arrow[d, "v"'] & \bullet \arrow[d, "v"] \arrow[dl, dotted, "d"'] \\ \bullet \arrow[r, "m"'] & \bullet \end{tikzcd} \tag{4.3.1}$$

admits a unique diagonal fill-in d which makes the two triangles inside the square commute. For two classes $\mathcal{E}$ and $\mathcal{M}$ of morphisms, we write $\mathcal{E} \downarrow \mathcal{M}$ when $e \downarrow m$ for all $e \in \mathcal{E}$ and $m \in \mathcal{M}$. Then, for a class $\mathcal{C}$ of morphisms,

$$\mathcal{C}^{\downarrow} = \{m \mid \mathcal{C} \downarrow \{m\}\}, \quad \mathcal{C}^{\uparrow} = \{e \mid \{e\} \downarrow \mathcal{C}\}.$$

A *prefactorization system* is a pair $(\mathcal{E}, \mathcal{M})$ such that $\mathcal{E} = \mathcal{M}^{\uparrow}$ and $\mathcal{E}^{\downarrow} = \mathcal{M}$. A *factorization system* is a prefactorization system $(\mathcal{E}, \mathcal{M})$ such that any morphism f decomposes as $f = me$ where $e \in \mathcal{E}$ and $m \in \mathcal{M}$. Already in a prefactorization system both classes $\mathcal{E}$ and $\mathcal{M}$ contain isomorphisms and are closed under composition. A factorization system can be equivalently defined as a pair $(\mathcal{E}, \mathcal{M})$ such that $\mathcal{E} \downarrow \mathcal{M}$ and in addition:

- both $\mathcal{E}$ and $\mathcal{M}$ contain identity morphisms are are closed under composition with isomorphisms;
- any morphism f decomposes as $f = me$ where $e \in \mathcal{E}$ and $m \in \mathcal{M}$.

Lemma 29. *In a category equipped with a bounded form, we have $e \downarrow m$ in any of the following cases:*

- *when e is thick and m is a left universalizer;*
- *when e is a right universalizer and m is thin.*

Proof. Suppose e is thick and m is a left universalizer. Consider a commutative square (4.3.1) of solid arrows. Since m is a monomorphism (Lemma 25), it suffices to show that $md = u$ for some morphism d. Since m is a left universalizer of $m \cdot 1$ (Lemma 27), the existence of such d will follow from the inequality $u \cdot 1 \leqslant m \cdot 1$. Since e is thick, we have: $u \cdot 1 = u \cdot (e \cdot 1) = (u \circ e) \cdot 1 = (m \circ v) \cdot 1 = m \cdot (v \cdot 1)$ (Lemma 23). At the same time, $m \cdot (v \cdot 1) \leqslant m \cdot 1$ (Lemma 24). Orthogonality of e and m in the case when e is a right universalizer and m is thin follows by duality. $\square$

Throughout the rest of this chapter, for a bounded form F, we write $\mathcal{R}_F$ to denote the class of right universalizers and $\mathcal{L}_F$ denote the class of left universalizers.

Definition 38. *A bounded form F over a category $\mathbb{C}$ is said to be*

- *a* pre-exact form *when every conormal object has a left universalizer and every normal object has a right universalizer;*
- *an* orthogonal form *when it is a pre-exact form and $\mathcal{R}_F \downarrow \mathcal{L}_F$;*
- *a* closed orthogonal form *when it is an orthogonal form and in addition $\mathcal{R}_F = \mathcal{R}_F^{\downarrow\uparrow}$ and $\mathcal{L}_F = \mathcal{L}_F^{\uparrow\downarrow}$;*
- *an* exact form *when it is a pre-exact form and $(\mathcal{R}_F, \mathcal{L}_F)$ is a prefactorization system.*

Proposition 22. *For any pre-exact form F we have: $\mathcal{L}_F^{\uparrow}$ is the class of thick morphisms and $\mathcal{R}_F^{\downarrow}$ is the class of thin morphisms.*

Proof. Already by Lemma 29, $\mathcal{L}_F^{\uparrow}$ contains all thick morphisms. If a morphism $f : X \to Y$ belongs to $\mathcal{L}_F^{\uparrow}$, then we obtain a commutative diagram

$$\begin{tikzcd}
X \arrow[r, "f"] \arrow[d] & Y \arrow[d, "1_Y"] \arrow[dl, dotted, "d"'] \\
\mathsf{Lun}(f1) \arrow[r, "\mathsf{lun}(f1)"'] & Y
\end{tikzcd}$$

This gives that $\mathsf{lun}(f1)$ is a split epimorphism. As $\mathsf{lun}(f1)$ is also a monomorphism (Lemma 25), we obtain that it is an isomorphism. Then $f1 = 1$ (Lemma 27), showing that f is thick. Dually, $\mathcal{R}_F^{\downarrow}$ is the class of thin morphisms. □

Note that as a consequence of the above proposition, we get that for a pre-exact form F the following conditions are equivalent:

(i) F is an orthogonal form.

(ii) Every left universalizer is thin.

(iii) Every right universalizer is thick.

We now characterize orthogonality of forms via existence of special factorizations of morphisms:

Theorem 8. *For any bounded form F over a category $\mathbb{C}$ the following conditions are equivalent:*

(i) F is an orthogonal form.

(ii) Each morphism $f : X \to Y$ in $\mathbb{C}$ admits a factorization $f = m\theta e$ where m is a left universalizer of $f1$ and e is a right universalizer of $0f$.

Proof. (i) ⇒ (ii): Suppose (i) holds. Consider any morphism $f : X \to Y$ in $\mathbb{C}$. Since $f1$ is conormal, it has a left universalizer. Dually, $0f$ has a right universalizer. We then obtain a commutative diagram, where the morphisms u and v arise from the universal properties of the given left and right universalizers, respectively:

$$\begin{tikzcd} X \arrow[r, "\mathsf{run}(0f)"] \arrow[d, "u"'] \arrow[dr, "f"] & \mathsf{Run}(0f) \arrow[d, "v"] \\ \mathsf{Lun}(f1) \arrow[r, "\mathsf{lun}(f1)"'] & Y \end{tikzcd}$$

Now, orthogonality produces a diagonal fill-in:

$$\begin{tikzcd} X \arrow[r, "\mathsf{run}(0f)"] \arrow[d, "u"'] & \mathsf{Run}(0f) \arrow[d, "v"] \arrow[dl, dotted, "\theta"'] \\ \mathsf{Lun}(f1) \arrow[r, "\mathsf{lun}(f1)"'] & Y \end{tikzcd}$$

The zigzag in the above diagram is the desired factorization of f.

(ii) ⇒ (i): Suppose that (ii) holds. For a conormal object $B = f \cdot 1^X$, the left universalizer of B is the morphism m in the factorization of f given by (ii), which shows that every conormal object has a left universalizer. Dually, every normal object has a right universalizer. So F is a pre-exact form. With

Proposition 22 in mind, to show that F is an orthogonal form, it suffices to show that any left universalizer is thin. If f is a left universalizer, then f is a left universalizer of $f1$ (Lemma 27), which forces the composite θe in the factorization $f = m\theta e$ given by (ii) to be an isomorphism (Lemma 25). This implies that e is a split monomorphism. Since e is a right universalizer, it is also an epimorphism (Lemma 25). Hence e is an isomorphism. Now, e is a right universalizer of $0f$, and so $0f = 0$ (Lemma 27), showing that f is thin. $\square$

Closed orthogonal and exact forms can be also characterized via the presence of suitable factorizations of morphisms, as the two theorems below show. In fact, as we will see, much more can be said in these two cases. We state both theorems before presenting their proofs:

Theorem 9. *For any bounded form F over a category $\mathbb{C}$ the following conditions are equivalent:*

(i) F is a closed orthogonal form.

(ii) Each morphism $f : X \to Y$ in $\mathbb{C}$ admits a factorization $f = m\theta e$ where m is a thin left universalizer, e is a thick right universalizer, and θ is both thick and thin.

(iii) Each morphism $f : X \to Y$ in $\mathbb{C}$ admits a factorization $f = m\theta e$ where m is a left universalizer of $f1$, e is a right universalizer of $0f$, and θ is both thick and thin.

(iv) F is an orthogonal form with both $\mathcal{R}_F$ and $\mathcal{L}_F$ closed under composition.

(v) Every conormal object has a left universalizer which is a thin embedding, and dually, every normal object has right universalizer which is a thick coembedding.

(vi) F is an orthogonal form and the pairs $(\mathcal{R}_F, \mathcal{R}_F^{\downarrow})$ and $(\mathcal{L}_F^{\uparrow}, \mathcal{L}_F)$ are factorization systems.

Theorem 10. *For any bounded form F over a category $\mathbb{C}$ the following conditions are equivalent:*

(i) F is an exact form.

(ii) F is a closed orthogonal form and any morphism in $\mathbb{C}$ that is both thick and thin is an isomorphism.

(iii) Each morphism $f : X \to Y$ in $\mathbb{C}$ admits a factorization $f = me$ where m is a thin left universalizer, and e is a thick right universalizer.

(iv) Each morphism $f : X \to Y$ in $\mathbb{C}$ admits a factorization $f = me$ where m is a left universalizer of $f1$, and e is a right universalizer of $0f$.

(v) F is a pre-exact form and the pair $(\mathcal{R}_F, \mathcal{L}_F)$ is a factorization system.

Proof of Theorem 9. (i)⇒(ii): Suppose F is a closed orthogonal form. By Theorem 8, we have a factorization

$$\begin{tikzcd}
X \arrow[rrr, "f"] \arrow[dr, "e=\mathsf{run}(0f)"'] & & & Y \\
& \mathsf{Run}(0f) \arrow[r, "\theta"'] & \mathsf{Lun}(f1) \arrow[ur, "\mathsf{lun}(f1)=m"'] &
\end{tikzcd}$$

Since the form is orthogonal, e is thick and m is thin (Proposition 22). We would like to show that θ is both thick and thin. Since $\theta 1$ is conormal, it has a left universalizer. Since $\mathcal{L}_F = \mathcal{L}_F^{\uparrow\downarrow}$, a composite of two left universalizers is a left universalizer. In particular, the composite $m \circ \mathsf{lun}(\theta 1)$ is a left universalizer. Then, it is a left universalizer of $(m \circ \mathsf{lun}(\theta 1)) \cdot 1$ (Lemma 27). Since $\mathsf{lun}(\theta 1) \cdot 1 = \theta 1$ (Lemma 27), we have:

$$(m \circ \mathsf{lun}(\theta 1)) \cdot 1 = m \cdot (\mathsf{lun}(\theta 1) \cdot 1) = m \cdot \theta 1 = (m \circ \theta) \cdot 1$$

(Lemma 23). Now, since e is thick, we further have:

$$(m \circ \theta) \cdot 1 = (m \circ \theta) \cdot (e1) = (m \circ \theta \circ e) \cdot 1 = f1$$

(Lemma 23). Thus, $m \circ \mathsf{lun}(\theta 1)$ is a left universalizer of $f1$. But so is m, and hence $\mathsf{lun}(\theta 1)$ is an isomorphism (Lemma 25). Then, $\theta 1 = 1$ (Lemma 27). This shows that θ is thick. By a dual argument, θ is also thin.

(ii)⇒(iii): Suppose (ii) holds. Consider a factorization

$$\begin{tikzcd}
X \arrow[rrr, "f"] \arrow[dr, "e"'] & & & Y \\
& \bullet \arrow[r, "\theta"'] & \bullet \arrow[ur, "m"'] &
\end{tikzcd}$$

where e is a thick right universalizer, m is a thin left universalizer, and θ is both thick and thin. Since both e and θ are thick, so is their composite θe (Lemma 26). Then $m1 = m \cdot (\theta e \cdot 1) = (m\theta e) \cdot 1 = f1$ (Lemma 23). So m is a left universalizer of $m1 = f1$ (Lemma 27). Dually, e is a right universalizer of $0f$.

(iii)⇒(iv): Suppose (iii) holds. Then F is an orthogonal form by Theorem 8, and so by Proposition 22, every left universalizer is thin and every right universalizer is thick. We show that $\mathcal{L}_F$ is closed under composition. Consider a composite $m_1 m_2$ of two left universalizers, and a commutative diagram of solid arrows

$$\begin{tikzcd}
\bullet \arrow[rr, "m_1"] \arrow[dr, "e=\mathsf{run}(0m_2m_1)"'] & & \bullet \arrow[rr, "m_2"] & & \bullet \\
& \mathsf{Run}(0m_2m_1) \arrow[rr, "\theta"'] & & \mathsf{Lun}(m_2m_1 1) \arrow[ur, "\mathsf{lun}(m_2m_1 1)=m"'] \arrow[ul, dotted, "d_1"] \arrow[ulll, dotted, "d_2"] &
\end{tikzcd}$$

obtained by (iii). Since e is thick and θ is thick, so is their composite θe (Lemma 26). The orthogonality $\theta e \downarrow m_2$ (Lemma 29) produces a morphism d_1 such that $m_2 d_1 = m$ and $d_1 \theta e = m_1$. Next, the orthogonality $\theta e \downarrow m_1$ produces a morphism d_2 such that $d_2\theta e$ is an identity morphism, and $m_1 d_2 = d_1$. Altogether, since the m's are monomorphisms (Lemma 25), we obtain that θe is an isomorphism (with inverse d_2). Then, since $m_2 m_1 = m\theta e$ and m is a left universalizer, $m_2 m_1$ is also a left universalizer (Lemma 25). This shows that $\mathcal{L}_F$ is closed under composition. Dually, $\mathcal{R}_F$ is closed under composition.

(iv)⇒(v): Suppose that (iv) holds. Consider a left universalizer $f : X \to Y$ and two conormal objects $A_1, A_2 \in F^{-1}(X)$. Then $A_1 = \mathsf{lun}(A_1)\cdot 1$ and $A_2 = \mathsf{lun}(A_2)\cdot 1$ (Lemma 27). Suppose $f\cdot A_1 \leqslant f\cdot A_2$. By (iv), the composite $f\mathsf{lun}(A_2)$ is a left universalizer, and hence it is a left universalizer of $(f\mathsf{lun}(A_2))\cdot 1$ (Lemma 27). Now, $(f\mathsf{lun}(A_2))\cdot 1 = f\cdot(\mathsf{lun}(A_2)\cdot 1) = f\cdot A_2$ (Lemma 23), and similarly, $(f\mathsf{lun}(A_1))\cdot 1 = f\cdot A_1$. So, $f\cdot A_1 \leqslant f\cdot A_2$ implies $f\mathsf{lun}(A_1) = f\mathsf{lun}(A_2)u$ for some morphism u. Since f is a monomorphism (Lemma 25), we get $\mathsf{lun}(A_1) = \mathsf{lun}(A_2)u$. This in turn gives $A_1 = \mathsf{lun}(A_1)\cdot 1 = (\mathsf{lun}(A_2)u)\cdot 1 = \mathsf{lun}(A_2)\cdot(u\cdot 1)$ (Lemma 23). Finally, $\mathsf{lun}(A_2)\cdot(u\cdot 1) \leqslant \mathsf{lun}(A_2)\cdot 1 = A_2$ (Lemma 24) and so $A_1 \leqslant A_2$. This shows that any left universalizer is an embedding. Dually, any right universalizer is a coembedding. Since by (iv) the form is orthogonal, to obtain (v) it remains to apply Proposition 22.

(v)⇒(vi): Suppose (v) holds. Then the form is pre-exact and so by Proposition 22, $\mathcal{L}_F^{\uparrow}$ is the class of thick morphisms and $\mathcal{R}_F^{\downarrow}$ is the class of thin morphisms. (v) implies that every left universalizer is thin, and so $\mathcal{R}_F^{\downarrow} \supseteq \mathcal{L}_F$ which is the same as $\mathcal{R}_F \downarrow \mathcal{L}_F$. The form F is therefore orthogonal. Consider a morphism $f : X \to Y$. By Theorem 8, it decomposes as $f = \mathsf{lun}(f1)\circ\theta\circ\mathsf{run}(0f)$. We claim that $\theta\circ\mathsf{run}(0f)$ is thick. Indeed, on the one hand we have $f1 = \mathsf{lun}(f1)\cdot 1$ (Lemma 27), and on the other hand, $f1 = (\mathsf{lun}(f1)\circ\theta\circ\mathsf{run}(0f))\cdot 1 = \mathsf{lun}(f1)\cdot((\theta\circ\mathsf{run}(0f))\cdot 1)$ (Lemma 23). Since $\mathsf{lun}(f1)$ is an embedding, it follows that $1 \leqslant (\theta\circ\mathsf{run}(0f))\cdot 1$, and hence $1 = (\theta\circ\mathsf{run}(0f))\cdot 1$, showing that $\theta\circ\mathsf{run}(0f)$ is thick. So any morphism f decomposes as a thick morphism followed by a left universalizer, which together with Lemma 29 and the fact that $\mathcal{L}_F^{\uparrow}$ and $\mathcal{L}_F$ contain identity morphisms and are closed under composition with isomorphisms (Lemmas 25 and 26), show that $(\mathcal{L}_F^{\uparrow}, \mathcal{L}_F)$ is a factorization system. Dually, $(\mathcal{R}_F, \mathcal{R}_F^{\downarrow})$ is a factorization system.

(vi)⇒(i): When $(\mathcal{L}_F^{\uparrow}, \mathcal{L}_F)$ and $(\mathcal{R}_F, \mathcal{R}_F^{\downarrow})$ are factorization systems we have $\mathcal{L}_F = \mathcal{L}_F^{\uparrow\downarrow}$ and $\mathcal{R}_F = \mathcal{R}_F^{\downarrow\uparrow}$. □

Proof of Theorem 10. (i)⇒(ii): Any exact form is a closed orthogonal form. When F is pre-exact, $\mathcal{R}_F^{\downarrow}\cap\mathcal{L}_F^{\uparrow}$ is the class of morphisms which are both thick and thin (Proposition 22). When $(\mathcal{R}_F, \mathcal{L}_F)$ is a prefactorization system, $\mathcal{R}_F^{\downarrow}\cap\mathcal{L}_F^{\uparrow} = \mathcal{L}_F\cap\mathcal{R}_F$ is the class of isomorphisms.

(ii)⇒(iii) follows from Theorem 9.

(iii)⇒(iv) by the same argument as the one used to prove (ii)⇒(iii) in Theorem 9, with θ there to be taken as an identity morphism.

(iv)⇒(v): Suppose (iv) holds. By Theorem 9, F is a closed orthogonal form and so it pre-exact and we have $\mathcal{L}_F = \mathcal{L}_F^{\uparrow\downarrow}$ and $\mathcal{R}_F \subseteq \mathcal{L}_F^{\uparrow}$. The existence of factorizations described in (iv) leave to show that F is an exact form, and for this it suffices to show that $\mathcal{L}_F^{\uparrow} \subseteq \mathcal{R}_F$. Recall from Proposition 22 that $\mathcal{L}_F^{\uparrow}$ is the class of thick morphisms. So we must show that any thick morphism is a right universalizer. If a morphism $f : X \to Y$ is thick, then since $f1 = 1$ the morphism m in the factorization $f = me$ given by (iv) is an isomorphism (Lemma 27). Since in the same factorization e is a right universalizer, we get that $f = me$ is a right universalizer (Lemma 25).

(v)⇒(i) since any factorization system is a prefactorization system. □

4.4 The binormal case

Recall from the previous chapter that a form is *binormal* when in each fibre every object is both normal and conormal. In such a form, to say that every (co)normal object has a right (left) universalizer, is the same as to say that the form admits right (left) universalisers. Also, a binormal form is bounded if and only if it is locally bounded and is a bifibration. For a general form F, we denote the class of all F-null morphisms (as defined in the previous section) by F^*.

Lemma 30. *In a category $\mathbb{C}$ equipped with a binormal pre-exact form F, a morphism $f : X \to Y$ is thin if and only if it is F^*-mono in the sense of [43], i.e. f has a kernel which belongs to F^*. Dually, f is thick if and only if f is F^*-epi in the sense of [43], i.e. f has a cokernel which belongs to F^*.*

Proof. Suppose f is thin. Then its kernel k is a left universalizer of $0f = 0$. This implies that $k1 = 0$, i.e. $k \in F^*$. Conversely, if $k \in F^*$ then $0 = k1 = 0f$ and so f is thin. □

Recall that a *semiexact category* or *ex1-category* in the sense of M. Grandis [41; 43] is a pair $(\mathbb{C}, \mathcal{N})$ where $\mathbb{C}$ is a category, and $\mathcal{N}$ is an ideal of null morphisms in $\mathbb{C}$ admitting kernels and cokernels, such that $\mathcal{N}$ is a *closed ideal*, i.e. any morphism from $\mathcal{N}$ factors through an isomorphism which belongs to $\mathcal{N}$. By Theorem 1.5.4 in [43], this latter additional requirement is equivalent to every kernel being $\mathcal{N}$-mono, and is also equivalent to every cokernel being $\mathcal{N}$-epi. So in view of the above lemma, in our language this is stating that a binormal pre-exact form defines a closed ideal if and only if every left universalizer is thin, and if and only if every right universalizer is thick. Thanks to Proposition 22, we know that for a pre-exact form the last two conditions are equivalent to orthogonality, and so we obtain:

Theorem 11. *A binormal bounded form F over a category $\mathbb{C}$ is an orthogonal form if and only if the pair $(\mathbb{C}, F^*)$ is a Grandis semiexact category (which is the same as a Grandis ex1-category).*

After this theorem, the implication (i)⇒(ii) of Theorem 8 in the binormal case becomes the basic known result that in a semiexact category any morphism has a *normal factorization* — see Section 1.5.5 in [43].

An *ex2-category* in the sense of M. Grandis [41; 43] is an ex1-category in which the class of kernels and the class of cokernels are both closed under composition. So, from Theorem 9 we reobtain the following result from the previous chapter:

Theorem 12. *A binormal bounded form F over a category $\mathbb{C}$ is a closed orthogonal form if and only if the pair $(\mathbb{C}, F^*)$ is a Grandis ex2-category.*

The equivalences (iii)⇔(iv)⇔(v) of Theorem 9 are established in Section 2.1.3 of [43], in the context of semiexact categories.

An *exact category* or an *ex4-category* in the sense of M. Grandis [41; 43] is a semiexact category in which any morphism f factories as $f = me$ where m is a kernel of a cokernel of f and e is a cokernel of a kernel of f. In the language of the underlying binormal form, this is the same as to say that m is a left universalizer of $f1$ and e is a right universalizer of $0f$. So, Theorem 10 gives:

Theorem 13. *A binormal bounded form F over a category $\mathbb{C}$ is an exact form if and only if the pair $(\mathbb{C}, F^*)$ is a Grandis exact category (which is the same as a Grandis ex4-category).*

In the context of semiexact categories, the equivalence (iv)⇔(ii) of Theorem 10 is made explicit in Section 2.2.6 of [43]. Thus we have the following hierarchy of bijections (where the first row is obtained from Corollary 4.2 in [59]):

$$
\begin{array}{ccc}
\begin{array}{c}\text{ideals of null morphisms}\\ \text{admitting enough exact sequences}\end{array} & \approx & \begin{array}{c}\text{isomorphism classes of}\\ \text{binormal bounded forms}\end{array}\\
\cup & & \cup\\
\begin{array}{c}\text{ideals of null morphisms}\\ \text{admitting kernels and cokernels}\end{array} & \approx & \begin{array}{c}\text{isomorphism classes of}\\ \text{binormal pre-exact forms}\end{array}\\
\cup & & \cup\\
\begin{array}{c}\text{ex1-category structures}\\ \text{(= semiexact category structures)}\end{array} & \approx & \begin{array}{c}\text{isomorphism classes of}\\ \text{binormal orthogonal forms}\end{array}\\
\cup & & \cup\\
\text{ex2-category structures} & \approx & \begin{array}{c}\text{isomorphism classes of}\\ \text{binormal closed orthogonal forms}\end{array}\\
\cup & & \cup\\
\begin{array}{c}\text{ex4-category structures}\\ \text{(= exact category structures)}\end{array} & \approx & \begin{array}{c}\text{isomorphism classes of}\\ \text{binormal exact forms}\end{array}
\end{array}
\tag{4.4.1}
$$

4.5 Grandis exact categories via Isbell bicategories

As we saw in Theorem 10, any exact form gives rise to a proper factorization system, i.e. a factorization system $(\mathcal{E}, \mathcal{M})$ where $\mathcal{E}$ is a class of epimorphisms and $\mathcal{M}$ is a class of monomorphisms [37]. We now ask the question: which proper factorizations systems arise from binormal exact forms? We obtain the answer by combining Theorem 7 from the previous chapter (and the subsequent remark) and Theorem 10 above.

Theorem 14. *For any category $\mathbb{C}$ there is a bijection*

$$
\begin{array}{ccc}
\begin{array}{c}\textit{proper factorization systems}\\ (\mathcal{E},\mathcal{M})\ \textit{such that the form}\\ \textit{of } \mathcal{E}\textit{-quotients is isomorphic}\\ \textit{to the form of } \mathcal{M}\textit{-subobjects}\end{array} & \approx & \begin{array}{c}\textit{isomorphism classes of}\\ \textit{binormal exact forms } F \textit{ over } \mathbb{C}\end{array}
\end{array}
$$

given by assigning to an isomorphism class of a binormal exact form F the factorization system $(\mathcal{E}, \mathcal{M})$ where $\mathcal{E}$ is the class of right universalizers for F and $\mathcal{M}$ is the class of left universalizers for F.

Recall that an Isbell bicategory [47] is a category equipped with a class $\mathcal{E}$ of morphisms called *projections* and a class $\mathcal{M}$ of morphisms called *injections*, such that the pair $(\mathcal{E}, \mathcal{M})$ is a proper factorization system on $\mathbb{C}$. The above theorem together with Theorem 13 lead to the following conclusion:

Corollary 6. *A Grandis exact category is the same as an Isbell bicategory in which the form of $\mathcal{E}$-quotients is isomorphic to the form of $\mathcal{M}$-subobjects,* *where $\mathcal{E}$ is the class of projections and $\mathcal{M}$ is the class of injections of an Isbell bicategory (which become the classes of cokernels and kernels, respectively, of the Grandis exact category).*

After this, one may ask: which proper factorization systems arise from exact forms that are not necessarily binormal? The answer is *all*, as we now show:

Proposition 23. *For any proper factorization system $(\mathcal{E}, \mathcal{M})$ on a category $\mathbb{C}$ there exists an exact form F over $\mathbb{C}$ such that $\mathcal{E}$ is the class of right universalizers and $\mathcal{M}$ is the class of left universalizers for F.*

Proof. First, we construct the domain $\mathbb{B}$ of F. Objects of $\mathbb{B}$ are triples (X, E, M) where

- X is an object of $\mathbb{C}$,
- E is a class of morphisms of $\mathbb{C}$ having the following property: either $E = \varnothing$ or there exists a morphism $e \in \mathcal{E}$ (which we will call a *generator* of E) such that the domain of e is X, and E is the class $E = \langle e]$ of all composites ue where u is any morphism in $\mathbb{C}$,
- M is a class of morphisms of $\mathbb{C}$ having the following property: either $M = \varnothing$ or there exists a morphism $m \in \mathcal{M}$ (which we will call a *generator* of M) such that the codomain of m is X, and M is the class $M = [m\rangle$ of all composites mv where v is any morphism in $\mathbb{C}$,
- Exactly one of E and M is the empty set.

A morphism $f : (X, E, M) \to (X', E', M')$ in $\mathbb{B}$ is a morphism $f : X \to X'$ in $\mathbb{C}$ such that the following conditions hold:

- for any morphism $m \in M$ we have $fm \in M'$,
- for any morphism $e \in E'$ we have $ef \in E$,
- if $M = \varnothing$ and $E' = \varnothing$ then $f \in M'$ or $f \in E$.

It is a routine to verify that composition and identity morphisms can be defined in $\mathbb{B}$ via composition and identity morphisms in $\mathbb{C}$. Then, mapping a morphism $f : (X, E, M) \to (X', E', M')$ to the morphism $f : X \to X'$ defines a faithful functor $F : \mathbb{B} \to \mathbb{C}$. In fact, F is even amnestic, and hence it is a form. In each fibre $F^{-1}(X)$, the top element 1^X is given by the triple $(X, \varnothing, [1_X\rangle)$, while the bottom element is given by the triple $(X, \langle 1_X], \varnothing)$. So, F is locally bounded. It is not difficult to show that F is in fact a bounded form where for each morphism $f : X \to Y$ we have $f \cdot 1^X = (Y, \varnothing, [m\rangle)$ and and $0^Y \cdot f = (X, \langle e], \varnothing)$ where m and e constitute an $(\mathcal{E}, \mathcal{M})$-factorization $f = me$ of f. Then, it is again not difficult to show that m is a left universalizer of $f \cdot 1^X$ and e is a right universalizer of $0^Y \cdot f$, and hence F is an exact form. We also get at once that every morphism in $\mathcal{M}$ is a left universalizer and every morphism in $\mathcal{E}$ is a right universalizer. Moreover, we get that $\mathcal{M}$ is the class of left universalizers of conormal objects and $\mathcal{E}$ is the class of right universalizers of normal objects. Via Lemma 27 this implies that $\mathcal{M}$ is the class of all left universalizers and $\mathcal{E}$ is the class of all right universalizers. $\square$

4.6 Exactness up to a class of morphisms and Grandis ex3-categories

In analogy with the terminology used in [41; 43], we call a morphism f in a category equipped with a bounded form *exact up to* Θ, where Θ is a class of morphisms in the category, if f admits a factorization $f = m\theta e$ where m is a left universalizer of $f1$, e is a right universalizer of $0f$, and $\theta \in \Theta$.

By Theorems 8, 9 and 10, a bounded form over a category is

- orthogonal, if and only if any morphism in the category is exact up to the class of all morphisms;
- closed orthogonal, if and only if any morphism in the category is exact up to the class of morphisms which are both thick and thin;
- exact, if and only if any morphism in the category is exact up to the class of identity morphisms, or equivalently, up to the class of isomorphisms.

In view of the fourth bijection in the table (4.4.1), the *ex3-categories* or *homological categories* in the sense of M. Grandis [41; 43] are precisely those categories equipped with a binormal closed orthogonal form in which any morphism f which decomposes as $f = em$ where m is a left universalizer and e is a right universalizer such that $0e \leqslant m1$, is exact up to the class of identity morphisms.

4.7 Exact forms of subobjects

For a category $\mathbb{C}$, recall that the form of subobjects is the form obtained from the codomain functor $\mathsf{Cod}_{\mathcal{M}}$, where $\mathcal{M}$ is the class of all monomorphisms. In particular, this form is always conormal, locally bounded above and admits left universalisers. We introduce some notation specific to forms of subobjects. Let $F : \mathbb{B} \to \mathbb{C}$ be the form of subobjects over a category $\mathbb{C}$. Then:

- objects of $\mathbb{B}$ will be pairs $(X, [m])$ where X is an object in $\mathbb{C}$ and $[m]$ is an equivalence class of monomorphisms m having codomain X, for the equivalence relation under which m_1 is equivalent to m_2 when $m_1 = m_2 i$ for some isomorphism i;

- a morphism $f : (X, [m]) \to (Y, [n])$ in $\mathbb{B}$ will be a morphism $f : X \to Y$ in $\mathbb{C}$ such that $fm = nj$ for some morphism j;
- $F(X, [m]) = X$ and $F(f) = f$.

Lemma 31. *Let $\mathbb{C}$ be a pointed category with form of subobjects $F : \mathbb{B} \to \mathbb{C}$. Then*

(1) the bottom element in the fibre at an object C of $\mathbb{C}$ is the (isomorphism class of) the zero morphism to C;

(2) the zero morphisms in $\mathbb{C}$ coincide with the F-null morphisms;

(3) for a morphism f in $\mathbb{C}$, $0 \cdot f$ exists if and only if f has a kernel k, in which case $0 \cdot f = [k]$;

(4) the cokernels in $\mathbb{C}$ coincide with the right universalisers for F when $\mathbb{C}$ has kernels of cokernels.

Proof. (1) and (2) are easy to prove. Recall that cartesian morphisms for a form of subobjects are always given by pullbacks, which gives (3). The right universaliser of an object $[m]$ is clearly the cokernel of m, while any cokernel c is the cokernel of its kernel, and hence the right universaliser of $[k]$, where k is its kernel. This gives (4). □

For convenience, we introduce the following condition on a category $\mathbb{C}$:

(M) every morphism in $\mathbb{C}$ admits a right $\mathcal{M}$-factorisaton, where $\mathcal{M}$ is the class of all monomorphisms.

We thus have the following characterisation of forms of subobjects satisfying the conditions in this chapter:

- The form of subobjects F of a pointed category $\mathbb{C}$ is bounded if and only if $\mathbb{C}$ satisfies (M) and admits all kernels.
- When this is the case, F will be pre-exact if and only if $\mathbb{C}$ admits all cokernels of kernels (i.e. F admits right universalisers of normal objects).
- Since monomorphisms always have trivial kernel, F will be orthogonal as soon as it is pre-exact (since left universalisers will then be thin).
- To additionally require that F is closed orthogonal it is enough to require cokernels to be closed under composition.

Finally, by Theorem 10, we have the following proposition, which shows that the exactness axiom on a form of subobjects captures the first isomorphism theorem.

Proposition 24. *Let $\mathbb{C}$ be a pointed category with form of subobjects F. Then F is exact if and only if $\mathbb{C}$ admits all kernels and a (cokernel, mono)-factorisation system.*

4.8 Some examples

We begin by recalling some examples from [43]. The forms of normal subobjects in the following pointed categories are binormal and orthogonal, but not closed orthogonal:

- the category of groups, where normal subobjects are given by normal subgroups;
- the category of rings without unit, where normal subobjects are given by ideals.

The forms of normal subobjects in the following pointed categories are binormal closed orthogonal, with the corresponding ex2-category being homological, but not exact:

- the category of lattices and Galois connections, where normal subobjects are given by principal down-closed sets;
- the category of commutative monoids, where normal subobjects are given by submonoids H that satisfy $[a + h \in H \ \wedge \ h \in H] \Rightarrow [a \in H]$;
- the category of topological vector spaces (over a given topological field), where normal subobjects are given by closed linear subspaces;
- the category of Banach spaces (over the field of reals or complex numbers) and bounded linear mappings, where normal subobjects are given by closed linear subspaces;
- the category of pointed sets, where normal subobjects are given by subsets containing the base point.

In any abelian category, the form of normal subobjects (which are the same as subobjects) is binormal exact. The form of subobjects in any normal category in the sense of [55] is exact, but in general is not binormal; in fact, to ask binormality would be equivalent to ask that the normal category is abelian. A principal example of a non-abelian normal category is the category of groups (where subobjects are given by subgroups). In these cases, any object in a fibre is still conormal. A natural example of an exact form where this is no longer the case is the form of additive subgroups of unitary rings, over the category of unitary rings. In this form, for a unitary ring R, the fibre $F^{-1}(R)$ consists of additive subgroups of R, and a left/right action by a ring homomorphism is given by taking the inverse/direct image of an additive subgroup along the homomorphism. Then, conormal objects are subrings, while normal objects are ideals. But in fact, non-binormal examples of exact forms abound, as witnessed by Proposition 23.

Chapter 5

Forms of subobjects of normal categories

Adapted from: Z. Janelidze and T. Weighill, Duality in non-abelian algebra III. Normal categories and 0-regular varieties (submitted; preprint at http://math.sun.ac.za/cms/).

5.1 Introduction

In this chapter, we relate the theory of forms developed so far to other axiomatic developments aimed at the study of group-like structures. We do so by characterising different types of categories which play a role in categorical algebra via conditions on their forms of subobjects. Our aim is to obtain a list of "classification theorems" of the following form: a category $\mathbb{C}$ is of a certain type if and only if its form of subobjects satisfies certain conditions. Moreover, these conditions should be self-dual.

The main result of this chapter is a characterisation of *normal categories* [55], where the notion of exact form introduced in the previous chapter plays an important role. Indeed, normal categories can be seen as pointed regular categories in which the First Isomorphism Theorem (the theorem which inspired the notion of exact form) holds.

We then combine the characterisation of normal categories with the results in [58] (which we recall) to obtain characterisations of pointed regular [3] protomodular [13] categories (also called homological categories [11]) and pointed Barr-exact protomodular categories (which are the same as semi-abelian categories [51] when the existence of coproducts is required). In doing so, we organise a list of axioms on a form which can be used as a self-dual axiomatic context for group theory. Some results which can be established in this context were given in a talk by Z. Janelidze [57]. They include, for example, the Zassenhaus Lemma and the Jordan-Hölder Theorem.

At the end of the chapter we give a proof of the Short Five Lemma as an example of diagram-chasing in the context of forms satisfying self-dual conditions, and relate this proof to the the classical group-theoretic proof given in the Background section, as well as to proofs by Mac Lane and Grandis in different contexts.

Some further characterisation of special types of normal categories in terms of so-called *Wyler joins* for a form, as well as some additional results, can be found in a joint paper with Z. Janelidze [61].

5.2 Normal categories

In this paper we will be looking at self-dual axioms on the *bifbration of subobjects* of a category, i.e. a form of subobjects which is at the same time a bifibration. From the results in Chapter 3 we have:

Lemma 32. *The form of subobjects over a category* $\mathbb{C}$ *is a bifibration if and only if every morphism in* $\mathbb{C}$ *admits a right* $\mathcal{M}$*-factorisation (for* $\mathcal{M}$ *the class of all monomorphisms) and all pullbacks along monomorphisms.*

Recall that a *normal category* [55] is a category $\mathbb{C}$ having the following properties:

(N1) $\mathbb{C}$ has a zero object and pullbacks (and hence all finite limits).

(N2) Any morphism f in $\mathbb{C}$ factorizes as $f = me$ where m is a monomorphism and e is a normal epimorphism.

(N3) The class of normal epimorphisms in $\mathbb{C}$ is stable under pullbacks (along arbitrary morphisms).

Equivalently, a normal category is the same as a regular category [3] which is pointed and where every regular epimorphism is a normal epimorphism. Although the present name for this concept was used for the first time in [55], the concept itself is a very old one, which has roots both in early investigations in categorical algebra and in universal algebra (see [53; 51] and the references there). In particular, algebraic normal categories were first studied in [35], where they were called *varieties with ideals*. Various aspects of normal categories have been extensively studied also in recent literature [38; 39; 56; 77].

If we remove axiom (N3), then we get what we call in this chapter an *unstably normal category*. To emphasise the difference, we may refer to normal categories as *stably normal categories*. In this section we show that in any unstably normal category, the stability axiom (N3) can be decomposed into the following stability axioms:

(N4) The class of normal epimorphisms in $\mathbb{C}$ is stable under pullbacks along arbitrary monomorphisms.

(N5) The class of *normal morphisms* in $\mathbb{C}$, i.e. those morphisms which factor as a normal epimorphism followed by a normal monomorphism, is stable under pullbacks (along arbitrary morphisms).

(N6) The pullback of any normal epimorphism in $\mathbb{C}$ along a split epimorphism is a normal morphism in the above sense.

Theorem 15. *For an unstably normal category, the axiom (N3) is equivalent to the conjunction of (N4) and (N5), as well as to the conjunction of (N4) and (N6).*

Proof. (N3) implies (N4) trivially. We can also get (N5) from (N3) since the class of normal monomorphisms is stable under pullbacks (along arbitrary morphisms) in any pointed category.

Note that (N5) implies (N6) simply because any normal epimorphism is a normal morphism.

It remains to show that (N4) and (N6) imply (N3). Any morphism $f : X \to Y$ in $\mathbb{C}$ can be decomposed as

$$\begin{tikzcd} & X \times Y \arrow[dr, "\pi_2"] & \\ X \arrow[ur, "(1_X,f)"] \arrow[rr, "f"'] & & Y \end{tikzcd}$$

where $(1_X, f)$ is a split monomorphism, and the product projection π_2 is a split epimorphism (the morphism $(0, 1_Y) : Y \to X \times Y$, where 0 is the zero morphism $Y \to X$, is a right inverse of π_2). So it suffices to show that the class of normal epimorphisms is stable under pullbacks along split monomorphisms and along split epimorphisms. We already have the first of these by (N4). To show the second, we will use the fact that because of (N2), any split epimorphism in $\mathbb{C}$ is a normal epimorphism. Let

$$\begin{tikzcd} X_1 \times_Y X_2 \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & X_2 \arrow[d, "f_2"] \\ X_1 \arrow[r, "f_1"'] & Y \end{tikzcd}$$

be a pullback where f_1 is a normal epimorphism and f_2 is a split epimorphism, with a right inverse $g : Y \to X_2$. Then π_1 is also a split epimorphism (with a right inverse given by the pair $(1_{X_1}, gf_1)$). Let $k : K \to X_2$ be a kernel of f_2. Then there is a morphism $u : K \to X_1 \times_Y X_2$ such that $\pi_2 u = k$ (it is given by the pair (0,k)). By (N6), the morphism π_2 is normal and so it decomposes as $\pi_2 = me$ where e is a normal epimorphism and m is a kernel of some morphism $h : Y \to Z$. Since k factors through π_2, and f_2 is a normal epimorphism, we obtain a factorization

$$\begin{tikzcd} X_2 \arrow[r, "h"] \arrow[d, "f_2"'] & Z \\ Y \arrow[ur, "v"'] & \end{tikzcd}$$

Next, we claim that v is a null morphism, and hence so is h. Since f_1 and π_1 are both epimorphisms, it suffices to show that $vf_1\pi_1$ is a null morphism, which is certainly true as $vf_1\pi_1 = vf_2\pi_2 = h\pi_2$. Since h is a null morphism, its kernel m must be an isomorphism. This makes π_1 into a normal epimorphism, since it factorises as $\pi_1 = me$ where e is such. $\square$

5.3 Axioms (A1–4)

In the case when $F : \mathbb{B} \to \mathbb{C}$ is a bounded biform, we can give a third equivalent definition for left and right univeralisers. Given an object $W \in F^{-1}(Y)$, a left universaliser of an object $W \in F^{-1}(Y)$ is a morphism $f : X \to Y$ such that $1^X = Wf$ and for any morphism $f' : X' \to Y$ such that $1^{X'} = Wf$, there exists a unique morphism $j : X' \to X$ such that $f' = fj$. Given an object $V \in F^{-1}(X)$, a right universalizer of an object $W \in F^{-1}(Y)$ is a morphism $f : X \to Y$ such that $fV = 0^Y$ and for any morphism $f' : X \to Y'$ such that $f'V = 0^{Y'}$, there exists a unique morphism $j : Y \to Y'$ such that $f' = jf$. This is in fact the original definition in [58], where only the context of a biform is considered.

We will see in the subsequent section that for the form of subobjects in an unstably normal category, normal morphisms in following sense are the same as in the sense of Section 5.2: a *normal morphism* (in a category equipped with a bounded biform) is a morphism $f : X \to Y$ such that $f1^X$ is normal, and hence binormal since $f1^X$ is automatically conormal. Dually, a *conormal morphism* is a morphism $f : X \to Y$ such that $0^Y f$ is conormal (binormal).

For a biform $F : \mathbb{B} \to \mathbb{C}$ and a morphism $f : X \to Y$ in $\mathbb{C}$, we call the push-forward functor $f^* : F^{-1}(X) \to F^{-1}(Y)$ (i.e. the functor whose action on objects is $A \mapsto f \cdot A$) the *left action* of f. Dually, the functor $B \mapsto B \cdot f$ will be called the *right action* of f. Together, the left and right action make up the change of base adjunction induced by f. A morphism $f : X \to Y$ is said to be an *injection* or a *surjection* when the left action of f is an injection or a surjection, respectively. Note that the left action is injective if and only if the right action is surjective, and when this is the case the left action is a right inverse of the right action — all for a given morphism f. Dually, the left action is surjective if and only if the right action is injective (and then the left action is the left inverse of the right action).

We are now ready to state the self-dual axioms on a bounded biform F, which will be used to characterise normal categories:

(A1) F is exact, i.e. any morphism $f : X \to Y$ in the ground category $\mathbb{C}$ decomposes as $f = me$ where e is a right universalizer of $0^Y f$ and m is a left universalizer of $f1^X$.

(A2) For any object X in the ground category $\mathbb{C}$, we have: 1^X is normal and 0^X is conormal.

(A3) All left universalizers are injections, and dually, all right universalizers are surjections.

(A4) The class of normal morphisms is stable under those pullbacks which exist in the category, and dually, the class of conormal morphisms is stable under those pushouts which exist in the category.

5.4 The main results

In this section we prove the following:

Theorem 16. *The form of subobjects of a category $\mathbb{C}$ having finite products is a bounded biform and satisfies the axioms (A1-4) if and only if $\mathbb{C}$ is a normal category.*

We will prove this theorem as follows: first, we will show that the axioms (A1) and (A2) together are equivalent to $\mathbb{C}$ being an unstably normal category. Then, we will show that in an unstably normal category (A3) is equivalent to the stability axiom (N4), and (A4) to (N5). The rest will follow from Theorem 15.

We begin by considering a category $\mathbb{C}$ with a terminal object equipped with a form $F : \mathbb{B} \to \mathbb{C}$ satisfying (A1) and (A2). It turns out that this already implies that $\mathbb{C}$ is pointed, and that the zero morphisms in $\mathbb{C}$ are nothing but the F-null morphisms. Recall the following result from the previous chapter:

Lemma 33. *If axiom (A1) holds, then a morphism $f : X \to Y$ is a right universalizer if and only if $f1^X = 1^Y$. Dually, under axiom (A1), f is a left universalizer if and only if $0^Y f = 0^X$. In particular, if $f1^X = 1^Y$ (i.e. f is thick) and $0^Y f = 0^X$ (i.e. f is thin), then f is an isomorphism.*

Note also that for any biform $F : \mathbb{B} \to \mathbb{C}$ and any morphism f in $\mathbb{C}$, $f \cdot 0 = 0$ and $1 \cdot f = f$ since left/right adjoints preserve colimits/limits.

Lemma 34. *Let $\mathbb{C}$ be a category which has a terminal object and let $F : \mathbb{B} \to \mathbb{C}$ be a bounded biform satisfying (A1) and (A2). Then, an object X in $\mathbb{C}$ is a terminal object in $\mathbb{C}$ if and only if $1^X = 0^X$, i.e. if and only if $|F^{-1}(X)| = 1$.*

Proof. Suppose X is a terminal object. Consider the right universalizer $e : X \to Y$ of 1^X, and the unique morphism $u : Y \to X$. Note that since by axiom (A2) the object 1^X is normal, it must have a right universaliser. Then $ue1^X = u0^Y = 0^X$. Since X is a terminal object, $ue = 1_X$, and so $1^X = 0^X$. Conversely, suppose $1^X = 0^X$. Consider the unique morphism $f : X \to T$ from X to a terminal object T. Since $1^X = 0^X$ and $1^T = 0^T$, f is thick and thin and hence an isomorphism. □

Recall that an *extremal monomorphism* is a monomorphism f such that every time we have a factorization $f = me$ where e is an epimorphism, it follows that e is an isomorphism. Then, it follows trivially from axiom (A1) that any extremal monomorphism is a left universalizer. In particular, this means that any split monomorphism, and more generally, any equalizer is a left universalizer. Dually, any *extremal epimorphism* (which is defined dually to an extremal monomorphism), and hence in particular any split epimorphism and any coequalizer is a right universalizer.

Lemma 35. *Let $\mathbb{C}$ be a category which has a terminal object and let $F : \mathbb{B} \to \mathbb{C}$ be a bounded biform satisfying (A1) and (A2). Then $\mathbb{C}$ is a pointed category. Moreover, a morphism $n : X \to Y$ in $\mathbb{C}$ is a null morphism if and only if $n1^X = 0^Y$, or equivalently, $1^X = 0^Y n$.*

Proof. For the first part, we need to show that a terminal object X in $\mathbb{C}$ is at the same time an initial object. Consider an object Y in $\mathbb{C}$. We will show that there is a unique morphism from some terminal object X to Y. Consider a left universalizer $f : X \to Y$ of 0^Y. Then $1^X = 0^Y f = 0^X$ and by Lemma 34, X is a terminal object. Suppose $f' : X \to Y$ is also a morphism from X to Y. Then f' is clearly thin, so it is a left universaliser of $f'1^X = 0^Y$, so it is isomorphic to f. Since X is terminal, it must be equal to f. This shows that $f : X \to Y$ is the unique morphism from the terminal object X to Y.

We now prove the second part of the lemma. Suppose $n : X \to Y$ is a zero morphism. Then it factors through a zero object Z. Since $1^Z = 0^Z$, we obtain that $n1^X = 0^Y$. Conversely, suppose $n1^X = 0^Y$. Then $1^X = 0^Y n$ and so n factors through the left universalizer of 0^Y. As we already showed above, the domain of this left universalizer is a zero object, and so n is a zero morphism. □

Remark 1. *Note that a binormal exact form satisfies (A1) and (A2). Thus we obtain the following corollary of the above lemmas by combining then with the results of the previous chapter: if $(\mathbb{C}, \mathcal{N})$ is a Grandis ex4-category [41] and $\mathbb{C}$ has a terminal object, then $\mathbb{C}$ is pointed and $\mathcal{N}$ is the class of zero morphisms. In particular, $\mathbb{C}$ will be exact in the sense of Puppe-Mitchell [82; 80]. If in addition $\mathbb{C}$ admits binary products, it will be abelian.*

Now let $\mathbb{C}$ be a category having finite products, and let F be its form of subobjects, which we assume to be a bounded biform. We have already remarked (Lemma 32) that since F is a bifibration, all pullbacks along monomorphisms exist.

Lemma 36. *If axiom (A1) holds then $\mathbb{C}$ has all finite limits.*

Proof. Since $\mathbb{C}$ already has all finite products, it is sufficient to show that $\mathbb{C}$ has pullbacks. In fact, it is sufficient to show that $\mathbb{C}$ has pullbacks of split monomorphisms along split monomorphisms. Indeed, any pullback

$$\begin{tikzcd} X_1 \times_Y X_2 \arrow[r, "\pi_2"] \arrow[d, "\pi_1"'] & X_2 \arrow[d, "f_2"] \\ X_1 \arrow[r, "f_1"'] & Y \end{tikzcd}$$

can be computed via the pullback

$$\begin{tikzcd} X_1 \times_Y X_2 \arrow[rr, "(\pi_1,\pi_2)"] \arrow[d, "(\pi_1,\pi_2)"'] & & X_1 \times X_2 \arrow[d, "1_{X_1}\times(f_2,1_{X_2})"] \\ X_1 \times X_2 \arrow[r, "(1_{X_1},f_1)\times 1_{X_2}"'] & (X_1 \times Y) \times X_2 \approx & X_1 \times (Y \times X_2) \end{tikzcd}$$

of split monomorphisms $(1_{X_1}, f_1) \times 1_{X_2}$ and $1_{X_1} \times (f_2, 1_{X_2})$. Split monomorphisms are in particular monomorphisms, and pullbacks of monomorphisms exist by Lemma 32. □

Since $\mathbb{C}$ has all kernels and is pointed, we have the following by Lemma 31 from the previous chapter:

Lemma 37. *Suppose axioms (A1) and (A2) hold. Consider a morphism $f : X \to Y$ in $\mathbb{C}$. A right universalizer of $f1^X$ is the same as a cokernel of f. In particular, this implies that the class of right universalizers coincides with the class of normal epimorphisms.*

Thus far, by using only axioms (A1) and (A2), we have obtained that $\mathbb{C}$ is a finitely complete pointed category in which any morphism decomposes as a normal epimorphism followed by a monomorphism. Moreover, we also know that monomorphisms are the same as left universalisers, and normal epimorphisms are the same as right universalisers. We also have the converse: if $\mathbb{C}$ is a finitely complete pointed category in which any morphism decomposes as a normal epimorphism followed by a monomorphism, then the axioms (A1) and (A2) will hold. Indeed, by Proposition 24 in the previous chapter, it is enough to show that every top element has a right universaliser, i.e. that identity morphisms have cokernels, which is obviously true.

We have thus proved the following:

Theorem 17. *The form of subobjects of a category $\mathbb{C}$ having finite products is a bounded biform and satisfies the axioms (A1) and (A2), if and only if $\mathbb{C}$ is an unstably normal category (i.e. it satisfies axioms (N1) and (N2)).*

Theorem 16 now follows from the following proposition and Theorem 15:

Proposition 25. *For an unstably normal category $\mathbb{C}$ with its form of subobjects, (A3) is equivalent to (N4), and (A4) is equivalent to (N5).*

Proof. First, we note that injectivity of left universalisers always holds. Indeed, for a form of subobjects this is equivalent to monomorphisms being closed under composition, which is always true (see Chapter 3).

Surjectivity of a morphism $f : X \to Y$ states that for every subobject $[n]$ of Y, we have: $f([n]f) = [n]$. Recall that here $[n]f$ is given by pulling back n along f:

$$\begin{tikzcd} \bullet \arrow[r, "g"] \arrow[d, "m"'] & \bullet \arrow[d, "n"] \\ X \arrow[r, "f"'] & Y \end{tikzcd}$$

To get $f([n]f)$, we need to decompose the composite fm into a normal epimorphism followed by a monomorphism. Say $fm = n'e$ is such a decomposition. Then, $[n'] = f([n]f)$ and examining the commutative square

$$\begin{tikzcd} \bullet \arrow[r, "g"] \arrow[d, "e"'] & \bullet \arrow[d, "n"] \\ \bullet \arrow[r, "n'"'] & Y \end{tikzcd}$$

we see that the claim $[n'] = [n]$ is equivalent to g being a normal epimorphism. So, f is surjective if and only if for any monomorphism n with codomain Y, the pullback of f along n is a normal epimorphism. Since normal epimorphisms are the same as right universalisers, this proves that (A3) is equivalent to (N4).

The equivalence of (A4) and (N5) is easy to verify. □

As a final remark, we note that the form of subobjects of an unstably normal category can be characterised as the unique conormal bounded biform over that category satisfying (A1) and (A2):

Proposition 26. *Let $\mathbb{C}$ be a category with finite products and $F : \mathbb{B} \to \mathbb{C}$ be a bounded conormal biform satisfying (A1) and (A2). Then F is the form of subobjects over $\mathbb{C}$.*

Proof. Clearly F admits left universalisers and is locally bounded above. Thus it is the form of $\mathcal{L}$-subobjects, where $\mathcal{L}$ is the class of all left universalisers (see Chapter 3). It is thus enough to prove that every monomorphism $m : X \to Y$ is a left universaliser, which by (A1) is equivalent to every monomorphism being thin, i.e. that $0 \cdot m = 0$. By Lemma 35, $\mathbb{C}$ is pointed, and null morphisms coincide with F-null morphisms. Consider then the left universaliser $m' : M' \to X$ of $0 \cdot m$ and the zero morphism z from M' to X. Then both $m \circ m'$ and $m \circ z$ are F-null, so they are zero morphisms and thus they coincide. Since m is mono, we have $m' = z$, which gives that m is thin as required. □

Corollary 7. *A category $\mathbb{C}$ is a normal category if and only if there exists a bounded conormal biform over $\mathbb{C}$ satisfying (A1–4).*

The notion of form of subobjects is much more familiar than that of a bounded conormal form, however, hence the phrasing of Theorem 16. Also, the self-dual nature of the result is better revealed when one considers only the form of subobjects.

5.5 Some examples

In this section, we first recall a characterisation of those varieties which form normal categories, from which we may obtain a wide variety of examples of normal categories (and thus forms satisfying (A1–4)). We then give some examples which illustrate the independence of some subsets of the axioms (A1–4), although the question of the independence of each individual axiom is still open.

5.5.1 Normal varieties

A classical result in universal algebra obtained in [35] states that a pointed variety of universal algebras is a normal category if and only if its algebraic theory contains a constant 0 and a sequence $d_1, \dots, d_n$ of binary terms satisfying

$$\bigwedge_{i=1}^{n} d_i(x, y) = 0 \quad \Leftrightarrow \quad x = y.$$

Recall that pointedness of a variety is equivalent to it having a unique constant 0. Among normal varieties are thus the varieties of groups, abelian groups, rings without unit, modules, loops, and many other group-like structures.

5.5.2 Groups with equivalence relations

In [61], the question of whether for an unstably normal category the axioms (N4) and (N5) (as well as (N4) and (N6)) are independent of each other was left open; each example of an unstably normal category that the authors of [61] have examined either satisfies both (N4) and (N6) (and hence is a stably normal category by virtue of the above theorem), or satisfies neither (N4) nor (N6). For instance, such is the category $\mathbf{Ab}_{\mathsf{Eq}}$ where

- objects are pairs (A, E) where A is an abelian group and E is an equivalence relation on the underlying set of A,
- a morphism $f : (A, E) \to (A', E')$ is a group homomorphism $f : A \to A'$ such that $a_1 E a_2 \Rightarrow f(a_1) E' f(a_2)$ for all $a_1, a_2 \in A$,

- composition of morphisms is defined as the usual composition of group homomorphisms.

The fact that $\mathbf{Ab}_{\mathsf{Eq}}$ is an unstably normal category is an easy consequence of the fact that $\mathbf{Ab}$ is a normal category. To see that $\mathbf{Ab}_{\mathsf{Eq}}$ does not satisfy (N4), consider the pullback

$$\begin{tikzcd}
& 1 \arrow[r, mapsto] & 1 & \\
1 \arrow[d, mapsto] & (\mathbb{Z}_4, E_3) \arrow[r] \arrow[d] & (\mathbb{Z}_2, E_2) \arrow[d] & 1 \arrow[d, mapsto] \\
2 & (\mathbb{Z}_8, E_1) \arrow[r] & (\mathbb{Z}_4, E_0) & 2 \\
& 1 \arrow[r, mapsto] & 1 &
\end{tikzcd}$$

where

- E_0 is the smallest equivalence relation on $\mathbb{Z}_4$ under which $0 \sim 1 \sim 2$,
- E_1 is the smallest equivalence relation on $\mathbb{Z}_8$ under which $0 \sim 1$ and $5 \sim 6$,
- E_2 is the largest equivalence relation on $\mathbb{Z}_2$,
- and E_3 is smallest equivalence relation on $\mathbb{Z}_4$.

The bottom horizontal morphism in the above pullback is a normal epimorphism, the right vertical morphism is a monomorphism, while the top horizontal morphism is not a normal epimorphism since the equivalence relation generated by the image of E_3 is the smallest equivalence relation on $\mathbb{Z}_2$, which is different from E_2. To see that $\mathbf{Ab}_{\mathsf{Eq}}$ does not satisfy (N6), consider the pullback

$$\begin{tikzcd}
& & (1,0) \arrow[r, mapsto] & (1,0) & & \\
& & (0,1) \arrow[r, mapsto] & (2,1) & & \\
(1,0) \arrow[d, mapsto] & (0,1) \arrow[d, mapsto] & (\mathbb{Z}_8 \times \mathbb{Z}_2, E_3') \arrow[r] \arrow[d] & (\mathbb{Z}_4 \times \mathbb{Z}_2, E_2') \arrow[d] & (1,0) \arrow[d, mapsto] & (0,1) \arrow[d, mapsto] \\
1 & 4 & (\mathbb{Z}_8, E_1) \arrow[r] & (\mathbb{Z}_4, E_0) & 1 & 2 \\
& & 1 \arrow[r, mapsto] & 1 & &
\end{tikzcd}$$

where

- E_0 and E_1 are the same as above.
- E'_2 is the smallest equivalence relation on $\mathbb{Z}_4 \times \mathbb{Z}_2$ under which $(0,0) \sim (1,0) \sim (2,0)$ and $(0,1) \sim (2,1)$,
- E'_3 is the smallest equivalence relation on $\mathbb{Z}_8 \times \mathbb{Z}_2$ under which $(0,0) \sim (1,0)$ and $(5,0) \sim (6,0)$.

As before, the bottom horizontal morphism in the above pullback is a normal epimorphism. The right vertical morphism is a split epimorphism. Now, the top horizontal morphism is not a normal morphism. To see why, first note that from the fact that the underlying map is surjective it follows that this morphism is a normal morphism if and only if it is a normal epimorphism. This morphism cannot be a normal epimorphism since for the smallest equivalence relation containing the image of E'_3 under this morphism, we do not have $(0,1) \sim (2,1)$ and hence this equivalence relation does not coincide with E'_2.

For the same reasons, other similarly constructed categories such as the category $\mathbf{Grp}_{\mathsf{Eq}}$ of groups with equivalence relations, are examples of unstably normal categories which satisfy neither (N4) nor (N6) (and hence do not satisfy (N5) either). Note that any unstably normal quasi-variety of universal algebras will necessarily be stably normal, since any quasi-variety is a regular category.

5.5.3 Bounded meet semi-lattices, and groupoid structures

Let $\mathbb{C}$ be a regular category. Then its form of subobjects is a biform. It is not bounded because, in general, an object X might not have a least subobject. Indeed, it is not difficult to find an example of a regular category where objects do not have least subobjects: any meet semi-lattice with an upper bound but not a lower bound is one (regular epimorphisms in this category are the same as identity morphisms, while any morphism is a monomorphism). Requiring the biform of subobjects of a regular category to be bounded is equivalent to requiring that any object in the category has a least subobject. In particular, this could be a bounded meet semi-lattice. In this example, the class of right universalisers coincides with the class of identity morphisms, while normal morphisms are morphisms whose domain is the initial object — the lower bound of the meet semi-lattice. Is it then not difficult to see that in this example (A1), (A3) and (A4) hold, while (A2) fails. Axioms (A1), (A3) and (A4) hold more generally for all sequentiable categories in the sense of Bourn [15], as the study of these categories in [15] reveals. For the form of subobjects of a sequentiable category, right universalisers coincide with cokernels in the sense of [15], which, by Proposition 2 in [15], are the same as regular epimorphisms. Then, axiom (A2) for a sequentiable category becomes simply pointedness. As explained in [15], the category of groupoid structures over a fixed set of objects gives an example of a sequentiable category. It is pointed if and only if the fixed set is either empty or a singleton — in the first case we get a singleton category, and in the second case we get the category of groups. So, to get a non-pointed example, it suffices to choose a set having at least two elements.

5.5.4 Pointed sets

As remarked above, the biform of subobjects in a regular category $\mathbb{C}$ is bounded if and only if every object in $\mathbb{C}$ has a least subobject. Let $\mathbb{C}$ be such a regular category. It is not difficult to see that all right universalisers for the form of subobjects in $\mathbb{C}$ are regular epimorphisms. The converse need not be the case: it is not the case, for instance, in the category $\mathbf{Set}_*$ of pointed sets, where regular epimorphisms are precisely the surjective maps of pointed sets which are injective beyond their kernels. The property that every regular epimorphism is a right universaliser is equivalent to validity of (A1), for the form of subobjects of any category $\mathbb{C}$ as above. At the same time, the fact that right universalisers are regular epimorphisms ensures that axiom (A3) holds. In fact, axiom (A4) will also hold. This follows from the fact that normal morphisms are precisely those whose regular image is a pullback of a least subobject — since in a regular category regular images commute with pullbacks, pullback stability of the class of normal morphisms reduces to the trivial fact that a pullback of a pullback of a least subobject is a pullback of a least subobject. Note that we need not worry about the conormal part of axiom (A4) since in the form of subobjects all objects in the fibres are conormal. So, for a category $\mathbb{C}$ as above, the form of subobjects satisfies both (A3) and (A4). The category of sets gives an example of such $\mathbb{C}$, with neither (A1) nor (A2) being satisfied. The category of pointed sets still does not satisfy (A1), as already remarked above, but it satisfies (A2) as a simple consequence of the fact that in this category every subobject is normal.

5.6 Protomodular and semi-abelian categories

Given a biform $F : \mathbb{B} \to \mathbb{C}$, a morphism f in $\mathbb{C}$ is called *left/right cartesian* [58] if the image of the left/right action of f is down/up-closed. The form F is called (left/right) cartesian when every morphism in $\mathbb{C}$ is (left/right) cartesian. In Corollary 1 in [58], it was shown that for a regular category $\mathbb{C}$, the following are equivalent:

- the form of subobjects F over $\mathbb{C}$ is cartesian;
- $\mathbb{C}$ is protomodular.

The following lemma is taken from [58]:

Lemma 38. *Let $F : \mathbb{B} \to \mathbb{C}$ be a locally bounded biform and $f : X \to Y$ be a morphism in $\mathbb{C}$. Then f is left cartesian if and only if for any object A in $F^{-1}(Y)$, $f \cdot (A \cdot f) = A \wedge f1$. Dually, f is right cartesian if and only if for any object B in $F^{-1}(X)$, $(f \cdot B) \cdot f = B \vee 0f$.*

Corollary 8. *For a category $\mathbb{C}$ having finite products, the following are equivalent:*

(1) $\mathbb{C}$ is pointed, regular and protomodular (i.e. it is a homological category);

(2) the form of subobjects over F is a bounded biform, satisfies the axioms (A1) and (A2), and is cartesian.

Proof. (1) $\Rightarrow$ (2): It is well-known that every homological category is normal (see for example [15]), so the result follows easily from Corollary 1 in [58].

(1) $\Rightarrow$ (3): Clearly cartesianness together with (A1) gives (A3), so we have that $\mathbb{C}$ satisfies (N1), (N2) and (N4). By Theorem 15, it remains to show that it satisfies (N6). By the remarks in the previous sections, normal epimorphisms (and in particular split epimorphisms) are precisely the right universalisers for F. Thus let $e : E \to Y$ and $f : X \to Y$ be two right universalisers, with f a split epi, and let $g : A \to X$ be the pullback of e along f as shown in the diagram below:

$$\begin{tikzcd} A \arrow[r, "g"] \arrow[d, "h"'] & X \arrow[d, "f"] \\ E \arrow[r, "e"'] & Y \end{tikzcd}$$

We are required to show that $g1$ is normal. In fact, we will show that $g1 = 1$, which together with (A2) gives the result. By cartesianness, if $g1 \geq 0f$, then $g1 = g1 \vee 0f = (fg1) \cdot f = (eh1) \cdot f = (e1) \cdot f = 1$, since h will also be a split epi, and hence a thick morphism under (A1). Translating the requirement $g1 \geq 0f$ back into the language of monomorphisms, we see that it is sufficient for the kernel of f to factor through g. This follows easily from the property of the pullback. □

A biform $F : \mathbb{B} \to \mathbb{C}$ is called *left (resp. right) stable* [58] when conormality (resp. normality) of objects is preserved under right (resp. left) action of left (resp. right) universalisers. A form which is both right and left stable is called simply *stable*. In Theorem 2 in [58] it was shown that for a pointed regular category with binary coproducts, the following are equivalent:

- the form of subobjects F over $\mathbb{C}$ is cartesian and stable;
- $\mathbb{C}$ is semi-abelian.

We obtain the following corollary:

Corollary 9. *For a category $\mathbb{C}$ having finite products and coproducts, the following are equivalent:*

(1) $\mathbb{C}$ is semi-abelian;

(2) the form of subobjects over F is a bounded biform, satisfies the axioms (A1) and (A2), is cartesian and is stable.

Thus the axioms (A1) and (A2) on a bounded biform, together with cartesianness and stability, provide a suitable context for studying group-like structures. Indeed, one only has to further require that the form is conormal to completely characterise forms of suobjects of semi-abelian categories (Lemma 7). However, when proving results in the context of a form satisfying these conditions, we do not use the fact that the form is conormal, since this is precisely the non-self-dual part of the properties of such a form.

We also have the following characterisation of abelian categories, which is easy to verify from the results in this chapter (for example from the remark following Lemma 35).

Proposition 27. *For a category $\mathbb{C}$ having finite products, the following are equivalent:*

- *$\mathbb{C}$ is abelian;*
- *the form of subobjects over F is a bounded binormal biform satisfying (A1).*

Note that any exact binormal form is cartesian; this was already noted (in the language of ideals of null morphisms) in [41].

5.7 Formal diagram chasing

We end this chapter with an illustration of diagram chasing in the context of a form satisfying self-dual conditions, as first introduced in [58]. Since we proved the Short Five Lemma for groups in the Background section, we prove a version of this lemma for forms here. In fact we prove a slightly more general lemma, based on the formulation in [74]. The proof is based on the "subobject chasing" argument introduced by Mac Lane in [74] and further developed by Grandis in [43].

Lemma 39. *Let $F : \mathbb{B} \to \mathbb{C}$ be a bounded biform which satisfies (A2) and which is cartesian. Then given any commutative diagram of the following form in $\mathbb{C}$:*

$$\begin{tikzcd} X \arrow[r, "l"] \arrow[d, "u"'] & Y \arrow[r, "r"] \arrow[d, "w"] & Z \arrow[d, "v"] \\ X' \arrow[r, "l'"'] & Y' \arrow[r, "r'"'] & Z' \end{tikzcd}$$

where r and r' are thick, l and l' are thin, $l \cdot 1 = 0 \cdot r$ and $l' \cdot 1 = 0 \cdot r'$, we have that w is thin/thick if both v and u are thin/thick. In particular, w is an isomorphism if both v and u are.

Proof. Suppose u and v are thin. Since v is thin,

$$(r' \circ w) \cdot (0 \cdot w) = 0 \Rightarrow (v \circ r) \cdot (0 \cdot w) = 0 \Rightarrow r \cdot (0 \cdot w) = 0$$

since left actions preserve bottom elements. Thus $0 \cdot w \leq 0 \cdot r = l \cdot 1$. Then by cartesianness,

$$0 \cdot w = l \cdot (0 \cdot (w \circ l)) = 0 \cdot (u \circ l') = 0.$$

which gives that w is thin. The fact that w is thick when u and v are thick follows dually, and the fact that w is an isomorphism when it is thick and thin follows from condition (A1). □

Notice that we are able to prove one half of the lemma and have the other half by dual arguments, since all the axioms used are self-dual. This is in contrast to the group-theoretic proof given in the Background section, where injectivity and surjectivity had to be proved separately using different methods.

The proof given here is based on the proof of the Five Lemma in [41], where it is also the case that one half is proved and the other deduced by duality. The proof in [41] is in the context of a "modular semi-exact category" of which any abelian category is an example, but of which the category of groups is not. Modular semi-exact categories correspond (under the bijection in Theorem 7) to cartesian binormal orthogonal forms. Thus, the difference between the proof given here and the one in [41] is that the arguments used here will also apply to the case of homological categories, since we never required the form to be binormal. Indeed, it is not hard to see, using the results of this chapter, that if F is the form of subobjects of a homological category, this lemma translates into the classical Short Five Lemma for pointed categories.

Chapter 6

Bifibrations in database theory

6.1 Introduction

In this chapter, we explore some connections between the work in the previous three chapters and the theory of databases. In particular, we look at an approach to the view update problem introduced by Johnson and Rosebrugh [65], where the notion of bifibration plays an important role. The goal of this chapter is to give an overview of the bifibrational model of view updatability, with the hope of providing a starting point for further applications of the theory of bifibrations, in particular that part which plays a role in categorical algebra, to the theory of databases. Some elementary remarks and elaborations on the work of Johnson and Rosebrugh are also made in this chapter.

We first look at the notion of a sketch, and the *sketch data model* developed by Johnson and Rosebrugh. We then give Johnson and Rosebrugh's definition of view updatability in this context, before looking at two examples which illustrate a connection between the functor from database states to view states and the functors encountered in the previous chapters. The first example is that of a codomain bifibration. In particular, the codomain and subobject bifibrations over the category of finite groups are examples of bifibrations arising from views. The second example is that of a domain functor, which in general does not represent an updatable view, for reasons which we discuss in detail. Finally, we look at a simplified version of the view update problem, which leads us to a proposition giving necessary and sufficient conditions for a reflection whose domain is finitely complete to be a fibration. This proposition uses the notion of semi-left-exact reflection in the sense of [24] (see also [21]).

6.2 Sketches and models

The notion of a sketch is due to Ehresmann [32]. The following definitions are taken from [5]:

Definition 39. *A* sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$ *consists of a graph* $\mathcal{G}$*, a set* $\mathcal{D}$ *of diagrams in* $\mathcal{G}$ *, a set* $\mathcal{L}$ *of finite cones in* $\mathcal{G}$ *and a set* $\mathcal{C}$ *of finite cocones in* $\mathcal{G}$*. A morphism of sketches from a sketch* $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$ *to a sketch* $\mathcal{S}' = (\mathcal{G}', \mathcal{D}', \mathcal{L}', \mathcal{C}')$ *is a graph homomorphism* $F : \mathcal{G} \to \mathcal{G}'$ *which sends diagrams in* $\mathcal{D}$*,* $\mathcal{L}$ *and* $\mathcal{C}$ *to diagrams in* $\mathcal{D}'$*,* $\mathcal{L}'$ *and* $\mathcal{C}'$ *respectively.*

By a *cone* in $\mathcal{G}$, we mean a graph homomorphism $D : \mathcal{G}' \to \mathcal{G}$, a distinguished node d in $\mathcal{G}$ and for each object g in $\mathcal{G}'$ a distinguished arrow $l_g : d \to D(g)$. A *cocone* is defined dually.

Definition 40. *A model of* $\mathcal{S}$ *in a category* $\mathbb{C}$ *is a graph homomorphism from* $\mathcal{G}$ *to the underlying graph of* $\mathbb{C}$ *which sends diagrams in* $\mathcal{D}$*,* $\mathcal{L}$ *and* $\mathcal{C}$ *to commutative diagrams, limit cones and colimit cones respectively.*

Every category $\mathbb{C}$ has an *underlying sketch*, whose graph is the underlying graph of $\mathcal{G}$ and where the diagrams in $\mathcal{D}$, $\mathcal{L}$ and $\mathcal{C}$ are precisely the commutative, limit and colimit diagrams in $\mathbb{C}$ respectively. Thus a model of a sketch $\mathcal{S}$ in $\mathbb{C}$ can be equivalently defined as a morphism of sketches from $\mathcal{S}$ to the underlying sketch of $\mathbb{C}$.

Every sketch $\mathcal{S}$ determines (up to equivalence) a category $\mathsf{Th}(\mathcal{S})$ which is finitely complete and cocomplete, which we call the *theory of* $\mathcal{S}$, and a model M_0 of $\mathcal{S}$ in $\mathsf{Th}(\mathcal{S})$ which has the universal property that for any model M of $\mathcal{S}$ in a category $\mathbb{C}$, there is a (unique up to natural isomorphism) finite limit- and colimit-preserving functor $F : \mathsf{Th}(\mathcal{S}) \to \mathbb{C}$ such that $F \circ M_0 = M$ [4].

By a (limit-class, colimit-class)-sketch, we mean a sketch where $\mathcal{L}$ and $\mathcal{C}$ are required to only contain cones and cocones from the specified limit class and colimit class respectively. The theory of such a sketch is then only closed under the limits and colimits in these classes, and has the corresponding universal property. For example, a *finite limit sketch* $\mathcal{S}$ is a sketch for which $\mathcal{C}$ is empty, and whose theory is thus only finitely complete. Models of such a sketch in a category $\mathbb{C}$ are in bijection with functors from $\mathsf{Th}(\mathcal{S})$ to $\mathbb{C}$ which preserve finite limits.

Definition 41. *Let* $M, M' : \mathsf{Th}(\mathcal{S}) \to \mathbb{C}$ *be two models of a sketch* $\mathcal{S}$ *in a category* $\mathbb{C}$ *(thought of as functors from the theory of* $\mathcal{S}$*). Then a* homomorphism of models *from* M *to* M' *is a natural transformation from* M *to* M'*.*

We could of course define homomorphisms of models without referring to theories – we need only adapt the notion of natural transformation to apply to a homomorphism of graphs between a graph $\mathcal{G}$ and the underlying graph of a category $\mathbb{C}$ in the obvious way.

Given a sketch $\mathcal{S}$ and a category $\mathbb{C}$, we thus have the category of models $\mathsf{Mod}(\mathcal{S}, \mathbb{C})$ whose objects are models of $\mathcal{S}$ in $\mathbb{C}$ and whose morphisms are model homomorphisms. Every category of universal algebras, for example. is (up to equivalence) a category of models for a sketch (see [5]). In practice, the sketches for even simple algebras can have lengthy definitions (see [5] for a sketch corresponding to semi-groups). It will be enough for us to observe that there is a sketch $\mathcal{S}$ whose category of models in **Set** is equivalent to **Grp** (depending on the definition of a group, we can even say that it *is* **Grp**).

We now show that for any sketch $\mathcal{S}$, there is a sketch $\mathcal{S}_2$ (which we will call the *arrow sketch* of $\mathcal{S}$) whose category of models in a category $\mathbb{C}$ is isomorphic to the category of arrows of $\mathsf{Mod}(\mathcal{S}, \mathbb{C})$. It is constructed as follows. Consider a sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$, and let $\mathcal{S}_2 = (\mathcal{G}_2, \mathcal{D}_2, \mathcal{L}_2, \mathcal{C}_2)$ be the sketch obtained as follows:

- the graph $\mathcal{G}_2$ is the disjoint union of $\mathcal{G}$ with itself, together with a family of arrows $\alpha_g : (g, 0) \to (g, 1)$, one for each object g of $\mathcal{G}$;

- the set $\mathcal{D}_2$ contains the images of all the diagrams in $\mathcal{D}$ under the inclusions $I_0, I_1 : \mathcal{G} \to \mathcal{G}_2$, as well as, for every arrow $f : g \to h$ in $\mathcal{G}$, the following "naturality" diagram:

$$\begin{tikzcd} (g,0) \arrow[r, "(f{,}0)"] \arrow[d, "\alpha_g"'] & (h,0) \arrow[d, "\alpha_h"] \\ (g,1) \arrow[r, "(f{,}1)"'] & (h,1) \end{tikzcd}$$

- the sets $\mathcal{L}_2$ and $\mathcal{C}_2$ contain the images of all the diagrams in $\mathcal{L}$ and $\mathcal{C}$ respectively under the inclusions $I_0, I_1 : \mathcal{G} \to \mathcal{G}'$.

It is easy to check that a model M for $\mathcal{S}_2$ in a category $\mathbb{C}$ is nothing but a pair of models M_1 and M_2 of $\mathcal{S}$ together with a model homomorphism $f : M_1 \to M_2$. Indeed, the inclusion of the naturality squares in $\mathcal{D}_2$ ensures that $(M(\alpha_g))_{g \in \mathcal{G}}$ gives a model homomorphism. Thus we have an isomorphism:

$$\mathsf{Mod}(\mathcal{S}_2) \cong \mathsf{Mod}(\mathcal{S})^{\mathbf{2}}$$

Given any sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$ and an arrow $f : g \to h$ in $\mathcal{G}$, we can recover those models M of $\mathcal{S}$ such that $M(f)$ is a monomorphism as the models for a new sketch $\mathcal{S}'$. We obtain $\mathcal{S}'$ by firstly adding an additional arrow $1_g : g \to g$ to $\mathcal{G}$, and then add the following diagram to $\mathcal{L}$:

$$\begin{tikzcd} g \arrow[r, "f"] & h \\ g \arrow[u, "1_g"] \arrow[r, "1_g"'] & g \arrow[u, "f"'] \end{tikzcd}$$

Any model M must send such a diagram to a pullback, and since $M(1_g) = M(1_g)$ (note, however, that it need not be equal to $1_{M(g)}$), we have that f must be monic. Conversely, if $M(f)$ is a monomorphism for a model M, then sending 1_g to the identity on $M(g)$ clearly defines a model of $\mathcal{S}'$. We say that an arrow f in $\mathcal{G}$ is a *specified monomorphism* to mean that every model sends f to a monomorphism [67]. In particular, we can specify each α_g in $\mathcal{G}_2$ above to be a monomorphism, and arrive at a sketch (which we will denote by $\mathcal{S}_m$), whose models are pairs of models M_1 and M_2 of $\mathcal{S}$ together with a model homomorphism whose every component is a monomorphism (in particular, it will be a monomorphism in $\mathsf{Mod}(\mathcal{S}, \mathbb{C})$).

Finally, we give a construction which will play an important role in the subsequent sections. Suppose we have two sketches $\mathcal{S}$ and $\mathcal{S}'$ and a model of $\mathcal{S}$ in $\mathsf{Th}(\mathcal{S}')$. Then for any category $\mathbb{C}$ there is a functor:

$$F^* : \mathsf{Mod}(\mathcal{S}', \mathbb{C}) \to \mathsf{Mod}(\mathcal{S}, \mathbb{C})$$

which takes a model M to $M' \circ F$, where M' is the functor from $\mathsf{Th}(\mathcal{S}')$ to $\mathbb{C}$ corresponding to M.

6.3 The sketch data model

The main idea of the sketch data model is that database schemas should correpond to sketches, while database states should correspond to models. The sketch data model has been the used in a number of successful major information system consultancies by Johnson and Rosebrugh, demonstrating its power in real-world applications (see [65] and the references there).

Definition 42. *An* EA (entity-attribute) sketch *[65] is a (finite limit, finite discrete cocone)-sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$ such that $\mathcal{L}$ contains a specified cone in $\mathcal{L}$ with empty base (denoted by* 1*). Arrows with domain 1 are called* elements*. Nodes which are vertices of cocones all of whose injections are elements are called* attributes*. Nodes which are neither attributes, nor 1, are called* entities*.*

Relational databases, as described in the Background section, are easy to represent as EA sketches. Given a relational database, we construct an EA sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$ as follows. For every domain D we add an attribute (in the sense of the above definition) to $\mathcal{S}$. For every relation R with attributes $(A_1, \ldots, A_n)$ and corresponding domains $(D_1, \ldots, D_n)$, we add a node R and a family of arrows $(r_i : R \to D_i)_{i \in \{1,2,\ldots,n\}}$, together with a limit diagram specifying that this family of arrows be monic. If the family $\{A_{x_1}, A_{x_2}, \ldots, A_{x_m}\}$ is a primary key, we add a limit cone specifying that the family of arrows $\{r_{x_1}, r_{x_2}, \ldots, r_{x_m}\}$ be jointly monic. Finally, if a relation R has an attribute A which is a foreign key referencing an attribute A' in a relation R', and which takes values in a domain D, then we add an arrow $a : R \to R'$ to $\mathcal{G}$ and the following diagram to $\mathcal{D}$:

$$\begin{tikzcd} R \arrow[r, "a"] \arrow[dr, "d"'] & R' \arrow[d, "d'"] \\ & D \end{tikzcd}$$

where d and d' are the arrows from R and R' corresponding to the attributes A and A' respectively. It is easy to generalise this to the case when the foreign key consists of multiple columns.

Given a sketch $\mathcal{S}$ corresponding to a database scheme, the theory of $\mathcal{S}$ has the interesting property that it contains many of the important queries implemented by a typical database system. For example, suppose a database scheme consists of three tables `tableA`, `tableB` and `tableC`. Suppose `tableA` and `tableB` both contain a foreign key column `columnX` which references the primary key in `tableC`. The corresponding EA sketch $\mathcal{S}$ for this schema has an underlying graph containing the following diagram:

$$\begin{tikzcd} 1 & & \mathtt{tableC} & \\ & \mathtt{tableA} \arrow[ur, "f"] & & \mathtt{tableB} \arrow[ul, "g"'] \end{tikzcd}$$

A typical query on this database would be (in SQL):

```
SELECT * FROM tableA INNER JOIN tableC
ON tableA.columnX = tableB.columnX;
```

This query returns a new table, each row of which is a pair of rows, one row from `tableA` and one row from `tableB`, such that the entries in `columnX` of each row are the same. Suppose M is a database state for $\mathcal{S}$. Then this query simply forms the pullback of $M(f)$ along $M(g)$. Thus the query is represented by the pullback of f along g in $\mathsf{Th}(\mathcal{S})$.

Another important query already mentioned is a SELECT query. The result of such a query can also be represented by an object of the theory. Suppose that the column `columnX` contains a number, and we want to select all those rows in `tableC` for which that number is 2. In SQL, we would execute the following query:

```
SELECT * FROM tableC WHERE columnX=2;
```

In order to represent this query as an object in the theory of a sketch, an "integers" attribute N must be added to the sketch. We should also have an arrow from `tableC` to N which we think of as interpreting the value of `columnX` (call it i). Now, suppose we have a database state D of this sketch where N is taken to the set $\{1, 2, \ldots, n\}$ and each coproduct injection $i_m : 1 \to N$ is taken to the morphism from the singleton set which picks out the number m.

In this case, we can compute the result of the select query above as the image under D of the limit of the following diagram (which is an object of $\mathsf{Th}(\mathcal{S})$):

$$\begin{tikzcd} & \texttt{tableC} \arrow[d, "i"'] \\ 1 \arrow[r, "i_2"'] & N \end{tikzcd}$$

Note that the result of this query is always relative to the choice of $D(N)$, i.e. the chosen representation of the number attribute.

Remark 2. *While it is not noted explicitly in [65], it is fairly clear that a sketch need only have one attribute. Every value in every row in a database, regardless of whether we view it as a number or a string, can be (and practically speaking already is) encoded uniquely as a natural number (since the possibilities for such fields are always finite). We can therefore define one entity which is the coproduct of sufficiently many copies of* 1 *as the sole attribute in the sketch.*

Not only can EA sketches capture many of the aspects of traditional relational databases and their query languages, but the sketch data model further allows the use of additional constraints which, while natural to consider, aren't typically expressible in a language like SQL. For example, the class $\mathcal{D}$ of commuting diagrams can enforce certain real-world constraints on a database (see [63]).

6.4 Updates and views

Now that we have defined an EA sketch, we can define database states and updates. The following definition is taken from [65].

Definition 43. *Given an EA sketch* $\mathcal{S}$*, a* database state *for* $\mathcal{S}$ *is a model of* $\mathcal{S}$ *in* $\mathbf{set}_0$*, the category of all finite sets. An* insert update *(resp.* delete update*) for a database state* D *is a monomorphism* $D \to D'$ *(resp.* $D' \to D$*) in the category* $\mathsf{Mod}(\mathcal{S})$.

As mentioned in the Background section, a *view* of a database allows a user to manipulate part of the data contained in or derived from the whole database. In SQL, for example, the `create view` query returns the result of a query on the database as a virtual table, which the user can manipulate [88]. The authors of [65] argue that views of databases should be allowed to include multiple tables, as well as tables derived from the larger database via queries. With this in mind, the following definition is given in [65]:

Definition 44. *A view of an EA sketch $\mathcal{S}$ is an EA sketch $\mathcal{V}$ together with a model V of $\mathcal{V}$ in $\mathsf{Th}(\mathcal{S})$ (or equivalently, a sketch morphism V from $\mathcal{V}$ to the underlying sketch of $\mathsf{Th}(\mathcal{S})$).*

As demonstrated in a previous section, any view V gives rise to a functor (which we call the *substitution functor*):

$$V^* : \mathsf{Mod}(\mathcal{S}, \mathbf{set}_0) \to \mathsf{Mod}(\mathcal{V}, \mathbf{set}_0)$$

from database states for $\mathcal{S}$ to database states for $\mathcal{V}$ (which we call *view states*).

The *view update problem* can be phrased in this language as follows: given an update to a view state, is there a "best" update to the database state which extends this update? The authors of [65] define such a "best" insert/delete as follows:

Definition 45. *Let $V : \mathcal{V} \to \mathsf{Th}(\mathcal{S})$ be a view and let D be a database state for $\mathcal{S}$. Let $i : T \to T'$ be an insert update with $T = V^*(D)$. Then i is a* propagatable insert *(relative to D) if there exists an update $i' : D \to D'$ in $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$ such that $V^*(i') = i$ and for any insert updates $j : T' \to T''$, $k : T \to T''$ and $k' : D \to D''$ with $V^*(k') = k = j \circ i$, there exists a unique insert update j' such that $j' \circ i' = k'$ and $V^*(j') = j$.*

In other words, given a database state D and the corresponding view state T, an update to T is propagatable whenever it can be extended to the rest of the database D in a "universal" way. Propagatable deletes are defined dually:

Definition 46. *Let $V : \mathcal{V} \to \mathsf{Th}(\mathcal{S})$ be a view and let D be a database state for $\mathcal{S}$. Let $i : T' \to T$ be an delete update with $T = V^*(D)$. Then i is a* propagatable delete *(relative to D) if there exists an update $i' : D' \to D$ in $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$ such that $V^*(i') = i$ and for any insert updates $j : T'' \to T'$, $k : T'' \to T$ and $k' : D'' \to D$ with $V^*(k') = k = i \circ j$, there exists a unique insert update j' such that $i' \circ j' = k'$ and $V^*(j') = j$.*

If every insert/delete update on a view state T is propagatable, then we call the view state *insert/delete updatable*. An *insert/delete updatable view* is one for which every view state is insert/delete updatable.

The definitions resemble the definitions for cocartesian and cartesian morphisms, but with the added complication of only allowing monomorphisms. For a sketch $\mathcal{S}$, we denote by $\mathsf{Mon}(\mathcal{S})$ the subcategory of $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$ consisting of all the monomorphisms – if $\mathcal{S}$ is an EA sketch, then this is nothing but the category of database states and insert updates for $\mathcal{S}$. We thus have the following proposition from [65]:

Proposition 28. *Let $\mathcal{S}$ and $\mathcal{V}$ be sketches and let $V : \mathcal{V} \to \mathsf{Th}(\mathcal{S})$ be a view such that V^* preserves monomorphisms. Then the restricted functor $V^*_m : \mathsf{Mon}(\mathcal{S}) \to \mathsf{Mon}(\mathcal{V})$ is a fibration/opfibration if and only if V is delete/insert updatable.*

Note that whenever $\mathcal{S}$ is *keyed* in the sense of [65], then every morphism of database states for $\mathcal{S}$ is a monomorphism (see [67]). Also, the functor V^* will always preserve monomorphisms when V is an inclusion into $\mathcal{S}$ itself.

Suppose we have a view V for which $V_m^* : \mathsf{Mon}(\mathcal{S}) \to \mathsf{Mon}(\mathcal{V})$ is a bifibration. Given a morphism $f : A \to B$ in $\mathsf{Mon}(\mathcal{V})$, consider the change of base adjunction $f^* \dashv f_*$ induced by f. Recall that an important concept in the previous chapters was when the unit/counit if this adjunction is trivial (for forms this is equivalent to f_* being injective/surjective). In terms of the database schema, this would mean that the insert/delete update represented by f is *reversible*, i.e. performing the update on the view and then undoing it returns the database to the original state.

Developing connections of this kind between bifibrations in algebra and in database theory could help illuminate aspects of both. It may be interesting to look at cases where some of the conditions introduced in the previous chapters hold for a view bifibration, and see if these cases have any real-world meaning. This is left for future work.

6.5 Updatable view schemas

In [66] and [65], various examples of updatable and non-updatable views are given (although the definition of updatability differs subtly between these two papers – see the section on update definitions below). Here we look at only two examples of views. The first is motivated from the codomain functors in the previous chapters, while the second is similar, but simpler, and illustrates the lack of duality in the definition of updatability arising from the monomorphic requirement in the definition. This requirement is discussed further in the final subsection.

6.5.1 Codomain and subobject views

Groups can be defined as models in **Set** for a finite limit sketch $\mathcal{S}$. Finite groups are models of this sketch in $\mathbf{set}_0$, so finite groups are database states for an EA sketch $\mathcal{S}$. Models (or database states) for the arrow sketch $\mathcal{S}_2$ (as defined earlier), are then group homomorphisms. The second injection $I_1 : \mathcal{S} \to \mathcal{S}_2$ gives rise to a substitution functor

$$I_1^* : \mathsf{Mod}(\mathcal{S}_2, \mathbf{set}_0) \to \mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$$

which is precisely the codomain functor Cod over the category of finite groups. Is I_1, seen as a view, updatable? Clearly I_1^* is a bifibration, but we need to consider the restriction to monomorphisms. We will denote the category of finite groups by **grp** and the subcategory consisting of all monomorphisms by $\mathbf{grp}_M$.

Lemma 40. *Let $\mathbb{C}$ be a category, and let $\mathbb{C}^{\mathbf{2}}$ be its category of arrows. If $\mathbb{C}$ has an initial object, then a morphism $(f, g) : a \to b$ in $\mathbb{C}^{\mathbf{2}}$ is a monomorphism if and only if both f and g are monomorphisms in $\mathbb{C}$.*

Proof. If $\mathbb{C}$ has an initial object, then it is easy to see that $\mathsf{Cod} : \mathbb{C}^{\mathbf{2}} \to \mathbb{C}$ has a left adjoint. We have already noted that the domain functor Dom has a left adjoint, so both Cod and Dom preserve limits, and hence also monomorphisms, which gives the required result. $\square$

It makes sense then to consider the subcategory $\mathbf{grp}^{\mathbf{2}}_M$ of $\mathbf{grp}^{\mathbf{2}}$ consisting of all the monomorphisms in $\mathbf{grp}^{\mathbf{2}}$, and the restricted functor

$$I^*_m : \mathbf{grp}^{\mathbf{2}}_M \to \mathbf{grp}_M$$

The cocartesian liftings and cartesian liftings for the codomain functor restrict to this context. Indeed, recall that given morphisms $f : A \to B$, $g : C \to D$ and a monomorphism $m : B \to D$, the Cod-cocartesian lifting of m from f is $(m, 1_A) : f \to m \circ f$, while the cartesian lifting of m to g is given by the pullback of m along g, and it is straightforward to check that these morphisms are also cocartesian and cartesian with respect to the restricted functor I^*_m. Thus the view I_1 is both delete and insert updatable.

This argument holds more generally for any finite limit sketch $\mathcal{S}$, not just for the sketch for groups. Recall that a category is the category of models in **Set** for some finite limit sketch if and only if it is a locally finitely presentable category (see [1]). In particular, these categories are always complete and cocomplete, and so have pullbacks and initial objects. Thus if we include models in **Set** as database states, every codomain functor over a locally finitely presentable category is a substitution functor arising from a view. Moreover, this view is updatable.

Practically, this result (that the view I_1 is always updatable) is not surprising. Roughly speaking, if a view of a database contains a collection of tables such that the rest of the database resembles a copy of the view with references into the view state, then any change to the view can easily be propagated to the rest of the database. In particular, inserting into the view does not change the rest of the database at all (see the construction of the cocartesian lifting above), while any delete updates must "cascade" into the rest of the database – this is the notion represented by taking the pullback.

The subobject functor over **grp** also arises from an updatable view. Let $\mathcal{S}$ be the sketch for groups, and recall from Section 6.2 that we can specify each α_g in $\mathcal{S}_2$ to be a monomorphism to obtain the sketch $\mathcal{S}_m$. The category of models for $\mathcal{S}_m$ is then the category of model homomorphisms $f : M \to M'$, each of whose components is a monomorphism. Clearly for the sketch of groups this is equivalent to f being a monomorphism in $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$. More generally, we have the following result:

Lemma 41. *Let $\mathcal{S}$ be a finite limit sketch whose theory has finite hom-sets. Then a homomorphism of models $m : M \to M'$ is a monomorphism in* $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$ *if and only if every component of m is a monomorphism.*

Proof. The category $\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0)$ can be equivalently expressed as the category of finite limit preserving functors from $\mathsf{Th}(\mathcal{S})$ to $\mathbf{set}_0$. In particular, this category contains all the representable functors from $\mathsf{Th}(\mathcal{S})$, so the result follows from the Yoneda lemma (see the proof in Background for the case of arbitrary presheaves) □

Thus the category $\mathsf{Mod}(\mathcal{S}_m, \mathbf{set}_0)$ is isomorphic to the full subcategory of the arrow category $\mathbf{grp^2}$ with objects all the monomorphisms. Now consider the view $I_1 : \mathcal{S} \to \mathcal{S}_m$. The substitution functor arising from this view is the subobject functor over $\mathbf{grp}$, and it is easy to see that the view I_1 is updatable. Indeed, the functor I_1^* restricted to monomorphisms is simply the codomain functor over $\mathbf{grp}_M$.

6.5.2 Domain views

We cannot immediately conclude that the view I_0 (i.e. the first injection $\mathcal{S} \to \mathcal{S}_2$) has all its view states updatable, since pushouts do not necessarily exist in $\mathsf{Mon}(\mathcal{S})$. This is precisely the non-dual part of the definition of updatability: that we only deal with monomorphisms. The following discussion is meant to illuminate this issue.

Suppose we have the following simple example of an EA sketch $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \mathcal{L}, \mathcal{C})$, namely that $\mathcal{G}$ has only one node with no edges, and $\mathcal{D}$, $\mathcal{L}$ and $\mathcal{C}$ are all empty (apart from the distinguished node 1). As a database, this is just a single entity, and states for $\mathcal{S}$ are just finite sets. Consider the corresponding arrow sketch $\mathcal{S}_2$. Models of $\mathcal{S}_2$ are then functions between finite sets. Now, consider the view $I_0 : \mathcal{S} \to \mathcal{S}_2$. The functor I_0^* is the domain functor over $\mathbf{set}_0$, which we know to be a bifibration.

Let D be a state for $\mathcal{S}$, i.e. a pair of sets D_1 and D_2 and a function $f : D_1 \to D_2$, and let $m : D_1 \to T$ be a monomorphism, i.e. an insert update on the view state D_1. The cocartesian lifting of m from f with respect to Cod is given by the pushout of f along m, shown in the diagram below:

$$\begin{tikzcd} D_2 \arrow[r, "m'"] & T' \\ D_1 \arrow[u, "f"] \arrow[r, "m"'] & T \arrow[u, "g"'] \end{tikzcd} \tag{6.5.1}$$

The morphism m' will be a monomorphism, but this is not necessarily a cocartesian lifting with respect to the restricted functor

$$I_{0m}^* : \mathsf{Mon}(\mathcal{S}_2) \to \mathsf{Mon}(\mathcal{S})$$

In fact, there is no cocartesian lifting of m from f in general. To see this, let $D_1 = \{1\}$, $D_2 = \{1,2\}$ and $f = m$ be the inclusion $D_1 \to D_2$. There are the following commutative diagrams:

$$\begin{tikzcd} \{1,2\} \arrow[r, equal] & \{1,2\} \\ \{1\} \arrow[u, "f"] \arrow[r, "m"'] & \{1,2\} \arrow[u, equal] \end{tikzcd} \qquad \begin{tikzcd} \{1,2\} \arrow[r, equal] & \{1,2\} \\ \{1\} \arrow[u, "f"] \arrow[r, "m"'] & \{1,2\} \arrow[u, "h"'] \end{tikzcd}$$

where h is the map $1 \mapsto 1$, $2 \mapsto 1$. Suppose $(n,m) : f \to g$ is a cocartesian lifting of m from f, where g is a morphism from $\{1,2\}$ to X, as shown in the diagram below:

$$\begin{tikzcd} \{1,2\} \arrow[r, "g'"] & X \\ \{1\} \arrow[u, "f"] \arrow[r, "m"'] & \{1,2\} \arrow[u, "g"'] \end{tikzcd}$$

Looking at the left-hand diagram, we conclude that g must be a split monomorphism. But looking at the right-hand diagram, we conclude that there must be a monomorphism k from X to $\{1,2\}$ such that $kg = h$, which is not possible.

6.5.3 Definitions of update

Consider the meaning of the pushout diagram (6.5.1) in the previous section in the language of databases. Recalling the construction of pushouts in $\mathbf{set}_0$, the object T' above is the disjoint union of D_2 and the complement of the image of m. In the language of databases, the update m adds to the rows of entity D_1. Each new row must reference a row of D_2, but since no universal choice can be made from D_2 itself, D_2 is expanded to freely include references for the new rows. However, another possible update to the database consistent with the view update m consists in manually choosing the values referenced by the new rows. In order that the propagated update be universal, it will be necessary to reassign the freely added values later on, but this is a modification of data, not an insert update.

The example in the previous section is in some sense trivial, but it illuminates some aspects of the definition of an update. Adding rows freely might not be something that a database management system implements, in which case we might want to conclude the view I_0 is not insert updatable. However, allowing modification of data to be included in the definition of an update allows for a universal propagation of the update which would also make sense in the context of databases, namely the use of “placeholder” values (i.e. freely added values). Thus depending on the context, it may or may not be sensible to consider non-monomorphic updates.

We should also note that an earlier definition of propagatable view update by the same authors in [66] seems to omit one of the requirements in Definition 45: it requires the lifting i' to be a monomorphism, but does not require the morphism j' to be a monomorphism. This is significant: Example 16 in [66] is for example not true under the definition given here (which is taken from the later paper [65]), as shown by the example in the previous section.

6.6 Bifibrations between categories of presheaves

By a linear sketch, we mean a sketch which is required to have both $\mathcal{L}$ and $\mathcal{C}$ empty (in other words a (none, none)-sketch in the terminology of Section 6.2). Let $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \varnothing, \varnothing)$ be a (linear) sketch and $V : \mathcal{V} \to \mathcal{S}$ be a full sketch inclusion. Further, assume that $\mathcal{G}$ is finite and acyclic, so that the theory of $\mathcal{S}$ will also be finite. In this section we show that the functor $V^* : \mathsf{Mod}(\mathcal{S}, \mathbf{set}_0) \to \mathsf{Mod}(\mathcal{V}, \mathbf{set}_0)$ is always a bifibration. Note that this does not mean that the view V is updatable, even with the addition of the required cone 1, since updatability requires us to look at the functor V^* restricted to monomorphisms. However, the result provides some insight into the view updatability problem for simple sketches, especially if the definition of update is relaxed to include natural transformations which are not monomorphisms. Note that a closely related result was already mentioned without proof in [66].

Since $\mathcal{S}$ is a linear sketch, we have the isomorphism:

$$\mathsf{Mod}(\mathcal{S}, \mathbf{set}_0) \cong \mathbf{set_0}^{\mathsf{Th}(\mathcal{S})}$$

so, writing $\mathbb{S} = \mathsf{Th}(\mathcal{S})$ and $\mathbb{V} = \mathsf{Th}(\mathcal{V})$, and denoting the inclusion $\mathbb{V} \to \mathbb{S}$ by V', we are required to prove that the functor

$$V'^* : \mathbf{set_0}^{\mathbb{S}} \to \mathbf{set_0}^{\mathbb{V}}$$

is a bifibration. We will prove this via a more general result regarding reflections.

Lemma 42. *Let $F : \mathbb{B} \to \mathbb{C}$ be a functor with a right adjoint T such that the counit of $F \circ T$ is the identity. Let $f, g : X \to T(B)$ be two morphisms in $\mathbb{B}$ such that $F(f) = F(g)$. Then $f = g$.*

Proof. We have the bijection $\mathsf{hom}(X, T(B)) \to \mathsf{hom}(F(X), B)$ whose action on morphisms is that of F, since the counit is trivial. $\square$

The following proposition is based on a remark by Janelidze [50].

Proposition 29. *Let $F : \mathbb{B} \to \mathbb{C}$ be a functor with a right adjoint T such that the counit of $F \dashv T$ is the identity. Then the following are equivalent:*

(1) F is a Street fibration (see Background section);

(2) $\mathbb{B}$ admits all pullbacks of the following form:

$$\begin{tikzcd} P \arrow[r, "q"] \arrow[d, "p"'] & B \arrow[d, "\eta_B"] \\ T(X) \arrow[r, "f'"'] & TF(B) \end{tikzcd}$$

where η is the unit of $F \dashv T$, and F preserves such pullbacks (i.e. F is a semi-left-exact reflection in the sense of [24]).

Proof. Let $F \dashv T$ be such an adjunction, and consider a commutative square in $\mathbb{B}$ of the following form, where $F(p)$ is an isomorphism:

$$\begin{tikzcd} P \arrow[r, "f''"] \arrow[d, "p"'] & B \arrow[d, "\eta_B"] \\ T(X) \arrow[r, "f'"'] & TF(B) \end{tikzcd}$$

We claim that this square is a pullback if and only if f'' is cartesian. It follows from Lemma 42 that given any morphism $u : Z \to B$, there exists a morphism $v : Z \to T(X)$ such that $f' \circ v = \eta_B \circ u$ if and only if there exists a morphism $g : F(Z) \to X$ such that $F(u) = F(f') \circ g$ (the morphism g will then be $F(v)$). Since p is an isomorphism, this is equivalent to the existence of a morphism $g' : F(Z) \to F(P)$ such that $F(f'') \circ g' = F(u)$. For such a morphism v and the corresponding morphisms g and g', a morphism $w : Z \to P$ satisfies $f'' \circ w = u$ and $p \circ w = v$ if and only if satisfies $f'' \circ w = u$ and $F(w) = g'$. It is now easy to see that the square above is a pullback if and only if f'' is cartesian.

(1) $\Rightarrow$ (2): Consider an object B in $\mathbb{B}$ and a morphism $f' : T(X) \to TF(B)$. The object B is in the fibre of $TF(B)$, so let $f'' : P \to B$ be the cartesian lifting of $F(f')$ to B (i.e. such that $F(f') \circ \theta = F(f'')$ for some isomorphism $\theta : F(P) \to X$). By adjunction, there is a morphism $p : P \to T(X)$ such that $F(p) = \theta$. By the above result, p will then be the pullback of η_B along f'. Moreover, since $F(\eta_B) = 1_{F(B)}$ and $F(p)$ is an isomorphism, F preserves this pullback.

(2) $\Rightarrow$ (1): Let $f : A \to F(Y)$ be a morphism in $\mathbb{C}$. Consider the pullback:

$$\begin{tikzcd} P \arrow[r, "f'"] \arrow[d, "p"'] & Y \arrow[d, "\eta_Y"] \\ T(A) \arrow[r, "T(f)"'] & TF(Y) \end{tikzcd}$$

Since F is semi-left-exact, the morphism $F(p)$ must be the pullback of $F(\eta_Y) = 1_{F(Y)}$ along $FT(f) = f$, i.e. an isomorphism such that $F(f') = f \circ F(p)$. Thus, f' is cartesian by the earlier result. $\square$

Note that semi-left-exact reflections were called *admissible* reflections by Janelidze in his Galois theory [49] (for the connection between the two notions and more on semi-left-exact reflections, see [21]).

Corollary 10. *Suppose $F : \mathbb{B} \to \mathbb{C}$ has a right adjoint T and a left adjoint I such that the counit of $F \dashv T$ and the unit of $I \dashv F$ are both trivial. Then if $\mathbb{B}$ admits pullbacks and pushouts, then F is a Street bifibration.*

Proof. Since F has a left and right adjoint, it preserves all limits and colimits, so the result follows from the previous result and its dual. □

Remark 3. *Recall that a functor F having left and right adjoints $I \dashv F \dashv T$ such that the counit of $F \dashv T$ and the unit of $I \dashv F$ are both trivial was one of the most important contexts considered in the previous chapters. By the above result, the requirement that F be a Street bifibration is equivalent to the existence of certain pullbacks and pushouts. Note, however, that the existence of the adjoints I and T is stronger than F being locally bounded (see Chapter 3).*

Returning to the functor V'^*, we notice that its domain is finitely complete and cocomplete (limits and colimits are computed point-wise). Moreover, since pullbacks are computed pointwise in $\mathbf{set}_0^{\mathbb{V}}$ and $\mathbf{set}_0^{\mathbb{S}}$, it is easy to see that we can always choose the pullback in (2) in Proposition 29 such that $F(p)$ is in fact an identity and not only an isomorphism. Thus, V'^* will be a Grothendieck bifibration as soon as it has left and right adjoints for which the corresponding unit and counit are trivial. Adjoints to V'^* always exist, and they are given by so-called Kan extensions (see for example [75]).

Definition 47. *Let $\mathbb{A}$, $\mathbb{B}$ and $\mathbb{C}$ be categories, and let $F : \mathbb{A} \to \mathbb{B}$ be a functor. Consider the functor $F^* : \mathbb{C}^{\mathbb{B}} \to \mathbb{C}^{\mathbb{A}}$. If it has a left adjoint $F_!$, then $F_!$ is called the* left Kan extension *operation along F. For a functor $H : \mathbb{A} \to \mathbb{C}$, we call $F_!(H)$ the* left Kan extension of H along F. *If F has a right adjoint F_*, then F_* is called the* right Kan extension *operation along F. For a functor $H : \mathbb{A} \to \mathbb{C}$, we call $F_*(H)$ the* right Kan extension of H along F.

Recall that whenever $\mathbb{C}$ is complete and cocomplete, the left and right Kan extension of any functor $H : \mathbb{A} \to \mathbb{C}$ along any functor $F : \mathbb{A} \to \mathbb{B}$ exists (see for example [75]), so that the functors $F_!$ and F_* in the above definition always exist. In fact, when $\mathbb{A}$ and $\mathbb{B}$ are finite, it is enough for $\mathbb{C}$ to have finite limits and colimits. Also from [75], we recall that when F is full and faithful, the unit and counit of the adjunctions $F_! \dashv F^*$ and $F^* \dashv F_*$ respectively are trivial. Together with Proposition 29, we thus have:

Theorem 18. *Let $\mathcal{S} = (\mathcal{G}, \mathcal{D}, \varnothing, \varnothing)$ be a linear sketch where $\mathcal{G}$ is finite and acyclic, and let $V : \mathcal{V} \to \mathcal{S}$ be a full inclusion. Then the substitution functor $V^* : \mathsf{Mod}(\mathcal{S}, \mathbf{set}_0) \to \mathsf{Mod}(\mathcal{V}, \mathbf{set}_0)$ is a bifibration.*

Chapter 7

Conclusion

7.1 Concluding remarks

At the end of the first chapter, we arrived at the following structure: a faithful amnestic functor F which fits into the following picture:

$$
\begin{tikzcd}
\mathbb{B} \arrow[d, "F \;\dashv"] \arrow[d, bend right=80, "R \;\dashv"'] \arrow[d, bend left=80, "L"] \\
\mathbb{C} \arrow[u, bend left=40, "I \;\dashv"] \arrow[u, bend right=40, "T \;\dashv"']
\end{tikzcd}
$$

where the counit of $F \dashv T$ and the unit of $I \dashv F$ are the identity. We noted that many important concepts in group theory such as quotient groups, subgroup inclusions and zero subgroups are already represented in this picture when we think of F as the bifibration of subgroups.

In order to represent the notions of kernel and image, we must require that F be a bifibration. The last proposition in the previous chapter gives an equivalent way to state this: F is a bifibration if and only if certain pullbacks and pushout exist in $\mathbb{B}$. Moreover, these pushouts and pullbacks are familiar, and play a role in categorical Galois theory and the theory of factorisation systems. Proposition 29 should in some sense have appeared earlier in the thesis. However, the author arrived to it while thinking about the work in Chapter 6, so it is included there.

7.2 Future work

7.2.1 Group theory

The most obvious next step is to use the context developed here as a means for doing group theory in a self-dual way. Already we have seen that the First Isomorphism Theorem can be stated in a self-dual way in the context of such a structure, and that other self-dual axioms capture the notions of normal, protomodular and semi-abelian categories. In each case, the only non-self-dual part is the requirement is that the form is a form of subobjects, which is often equivalent to the form being conormal (see Chapter 4).

The technique for proving isomorphism theorems and other results in such a context should be to assume that objects are conormal only so far as will ensure that the statement of the theorem remains self-dual. Hopefully this way of formulating and proving results for groups leads to new conceptual links between results, as well as more efficient proofs. There is still much work to be done in this direction, however. The first proofs in this line were given in a recent talk by Z. Janelidze in Sydney [57].

7.2.2 Other topics in categorical algebra

It would be interesting to see which other areas of categorical algebra could be extended to the context of a form. For example, the very first steps towards generalising homological algebra have already been made. A basic treatment of homological algebra in this context is work in progress by Z. Janelidze and Van der Linden, and was presented in detail in [85]. Z. Janelidze has also remarked on a number of occasions that it may be interesting to develop a theory of closure operators in the context of forms. This became the starting point for the PhD research in progress of his student A. Abdalla. Some basic results about torsion theories for a form have been formulated by the author and discussed in a recent talk [86].

7.2.3 Views of databases

The question of extending the definition of update to a database to include arbitrary modifications and not just inserts and deletes was already mentioned. It would also be good to continue to apply the mathematical theory of view updates to real-world problems, and to investigate real-world problems in order to inspire new theory. In particular, the idea of applying this theory to large networks of database systems is particularly interesting.

Links between the conditions on a bifibration which arise in categorical algebra, and aspects of bifibrations arising from views should also be investigated. As mentioned in the previous chapter, it would be interesting to see if these conditions have real-world meanings in the context of databases, and can inform the process of defining and implementing view updatability.

Bibliography

[1] J. Adámek and J. Rosicky, Locally presentable and accessible categories, London Mathematical Society Lecture Note Series 189, Cambridge University Press, 1994.

[2] F. Bancilhon and N. Spyratos, Update semantics of relational views, ACM Transactions on Database Systems (TODS) 6(4), 1981, 557–575.

[3] M. Barr, P. A. Grillet and D. H. van Osdol, Exact categories and categories of sheaves, Lecture Notes in Mathematics 236, Springer-Verlag, 1971.

[4] M. Barr and C. Wells, Toposes, triples and theories, New York: Springer-Verlag, 1985.

[5] M. Barr and C. Wells, Category theory for computing science, New York: Prentice Hall, 1990.

[6] J. Bénabou, Fibered categories and the foundations of naive category theory, Journal of Symbolic Logic 50.01, 1985, 10–37.

[7] G. Birkhoff and S. Mac Lane, Algebra, AMS Chelsea Publishing, 1967.

[8] A. Bohannon, J. Vaughan and B. Pierce, Relational Lenses: A language for updatable views, Proceedings of Principles of Database Systems (PODS) 2006, 338–347.

[9] F. Borceux, Handbook of Categorical Algebra 1: Basic Category Theory, Encyclopedia of Mathematics and its Applications Vol. 50, Cambridge University Press, 1994.

[10] F. Borceux, Handbook of Categorical Algebra 2: Categories and Structures, Encyclopedia of Mathematics and its Applications Vol. 51, Cambridge University Press, 1994.

[11] F. Borceux and D. Bourn, Mal'cev, protomodular, homological and semi-abelian categories, Springer, 2004.

[12] F. Borceux, G. Janelidze and G. M. Kelly, Internal object actions, Comment. Math. Univ. Carolinae 46.2, 2005, 235–255.

[13] D. Bourn, Normalization equivalence, kernel equivalence and affine categories, Lecture Notes in Mathematics 1488, Springer-Verlag, 1991, 43–62.

[14] D. Bourn, Mal'cev categories and fibrations of pointed objects, Appl. Categorical Structures 4, 1996, 307–327.

[15] D. Bourn, 3×3 lemma and Protomodularity, Journal of Algebra 236, 2001, 778-795.

[16] D. Bourn and M. Gran, Torsion theories in homological categories, Journal of Algebra 305.1, 2006, 18–47.

[17] D. Bourn and G. Janelidze, Protomodularity, descent, and semidirect products, Theory and Applications of Categories 4.2, 1998, 37–46.

[18] D. Bourn and G. Janelidze, Characterisation of protomodular varieties of universal algebras, Theory and Applications of Categories 11.6, 2003, 143–147.

[19] D. Bourn and Z. Janelidze, Approximate Mal'tsev operations, Theory and Applications of Categories 21, 2008, 152–171.

[20] D. A. Buchsbaum, Exact categories and duality, Transactions of the American Mathematical Society, 1955, 1–34.

[21] A. Carboni, G. Janelidze, G. M. Kelly, and R. Paré, On localization and stabilization for factorization systems, Applied Categorical Structures 5, 1997, 1–58.

[22] A. Carboni, J. Lambek, and M. C. Pedicchio, Diagram chasing in Mal'cev categories, Journal of Pure and Applied Algebra 69, 1990, 271–284.

[23] A. Carboni, M. C. Pedicchio, N. Pirovano, Internal graphs and internal groupoids in Mal'cev categories, Category Theory 1991, Montreal, CMS Conf. Proc. Vol. 13, 1991, 97–109

[24] C. Cassidy, M. Hebert, and G. M. Kelly, Reflective subcategories, localizations, and factorization systems, Journal of Australian Mathematical Society (Series A), 1985, 287-329.

[25] D. D. Chamberlin and R. F. Boyce, SEQUEL: A structured English query language, Proceedings of the 1974 ACM SIGFIDET (now SIGMOD) workshop on Data description, access and control, ACM, 1974.

[26] E. F. Codd, A relational model of data for large shared data banks, Communications of the ACM 13.6, 1970, 377–387.

[27] C. J. Date, Introduction to Database Systems, Addison-Wesley, 2004.

[28] D. Dikranjan and W. Tholen, Categorical structure of closure operators, Mathematics and its Applications 346, Kluwer Academic Publishers, 1995.

[29] Z. Diskin, Y. Xiong and K. Czarnecki, From state- to delta-based bidirectional model transformations: The symmetric case, Model Driven Engineering Languages and Systems, Springer Berlin Heidelberg, 2011, 304–318.

[30] H. Ehrbar and O. Wyler, On subobjects and images in categories, Technical Report, Department of Mathematical Sciences, Carnegie Mellon University, 1968.

[31] C. Ehresmann, Sur une notion générale de cohomologie, C. R. Acad. Sci. Paris 259, 1964, 2050–2053.

[32] C. Ehresmann, Sur les structures algébriques, C. R. Acad. Sci. Paris Sér. A-B 264, 1967.

[33] T. Everaert and T. Van der Linden, Baer invariants in semi-abelian categories II: Homology, Theory and Applications of Categories 12.4, 2004, 195-224.

[34] T. Everaert and T. Van der Linden, Relative commutator theory in semi-abelian categories, Journal of Pure and Applied Algebra 216.8, 2012, 1791–1806.

[35] K. Fichtner, Varieties of universal algebras with ideals, Mat. Sbornik, N. S. 75 (117), 1968, 445-453 (English translation: Math. USSR Sbornik 4, 1968, 411-418).

[36] G. D. Findlay, Reflexive homomorphic relations, Canadian Mathematical Bulletin 3, 1960, 131–132.

[37] P. J. Freyd and G. M. Kelly, Categories of continuous functors I, Journal of Pure and Applied Algebra 2, 1972, 169-191.

[38] M. Gran and T. Everaert, Monotone-light factorisation systems and torsion theories, Bulletin des Sciences Mathématiques 137, 2013, 996-1006.

[39] M. Gran and Z. Janelidze, Star-regularity and regular completions, Journal of Pure and Applied Algebra 218, 2014, 1771-1782.

[40] M. Gran and T. Van der Linden, On the second cohomology group in semi-abelian categories, Journal of Pure and Applied Algebra 212.3, 2008, 636–651.

[41] M. Grandis, On the categorical foundations of homological and homotopical algebra, Cah. Top. Géom. Diff. Catég. 33, 1992, 135–175.

[42] M. Grandis, Homological Algebra: The interplay of homology with distributive lattices and orthodox semigroups, World Scientific Publishing Co., Singapore, 2012.

[43] M. Grandis, Homological Algebra In Strongly Non-Abelian Settings, World Scientific Publishing Co., Singapore, 2013.

[44] A. Grothendieck, Sur quelques points d'algebre homologique, Tohoku Math. J. 9, 1957, 119–221.

[45] A. Grothendieck, Technique de descente et théorèmes d'existence en géométrie algébrique. I. Généralités. Descente par morphismes fidèlement plats, Séminaire N. Bourbaki 5, Exposé 190, 1959, 299–327.

[46] J. Hughes and B. Jacobs, Factorization systems and fibrations: Toward a fibred Birkhoff variety theorem, Electronic Notes in Theoretical Computer Science 69, 2003, 156–182.

[47] J. R. Isbell, Algebras of uniformly continuous functions, Annals of Mathematics (2) 68, 1958, 96–125.

[48] B. Jacobs, Categorical logic and type theory, Studies in logic and the foundations of mathematics Vol. 141, Elsevier, 1999.

[49] G. Janelidze, The fundamental theorem of Galois theory, Math. USSR-Sb. 64, 1989, 359–374.

[50] G. Janelidze, An email sent to the category theory mailing list, 4 August 2014 `http://www.mta.ca/ cat-dist/`.

[51] G. Janelidze, L. Márki and W. Tholen, Semi-abelian categories, Journal of Pure and Applied Algebra 168, 2002, 367–386.

[52] G. Janelidze and L. Márki, Kurosh–Amitsur radicals via a weakened Galois connection, Communications in Algebra, 31, 2003, 241–258.

[53] G. Janelidze, L. Márki and A. Ursini, Ideals and clots in universal algebra and in semi-abelian categories, Journal of Algebra 307, 2007, 191-208.

[54] Z. Janelidze, Cover relations on categories, Appl. Categor. Struct. 17, 2009, 351-371.

[55] Z. Janelidze, The pointed subobject functor, 3×3 lemmas, and subtractivity of spans, Theory and Applications of Categories 23, 2010, 221-242.

[56] Z. Janelidze, An axiomatic survey of diagram lemmas for non-abelian group-like structures, Journal of Algebra 370, 2012, 387-401.

[57] Z. Janelidze, Bifibrational duality for groups and in general categories, Category Theory 2013, Sydney (Australia), July 2013.

[58] Z. Janelidze, On the form of subobjects in semi-abelian and regular protomodular categories, (special volume dedicated to the sixtieth birthday of George Janelidze), Online First, December 2013.

[59] Z. Janelidze and T. Weighill, Duality in non-abelian algebra I. From cover relations to Grandis ex2-categories, Theory and Applications of Categories 29.11, 2014, 315–331.

[60] Z. Janelidze and T. Weighill, Duality in non-abelian algebra II. From Isbell bicategories to Grandis exact categories. Preprint at http://math.sun.ac.za/cms/.

[61] Z. Janelidze and T. Weighill, Duality in non-abelian algebra III. Normal categories and 0-regular varieties. Preprint at http://math.sun.ac.za/cms/.

[62] M. Johnson, Enterprise Software with Half-Duplex Interoperations, In Doumeingts, Mueller, Morel and Vallespir (eds), Enterprise Interoperability: New Challenges and Approaches, Springer-Verlag, 2007, 521–530.

[63] M. Johnson and C. N. G. Dampney, On the value of commutative diagrams in information modelling, Algebraic Methodology and Software Technology (AMAST'93), Springer, London, 1994.

[64] M. Johnson and S. Kasangian, A Relational Model of Incomplete Data without `NULL`s, CRPIT (Conferences in Research and Practice in Information Technology), Vol 109, 2010, 89–94.

[65] M. Johnson and R. Rosebrugh, Fibrations and universal view updatability, Theoretical Computer Science 388.1, 2007, 109–129.

[66] M. Johnson, R. Rosebrugh and C. N. G. Dampney, View updates in a semantic data modelling paradigm, Database Conference, 2001. ADC 2001. Proceedings. 12th Australasian. IEEE, 2001.

[67] M. Johnson, R. Rosebrugh and R. J. Wood, Entity-relationship-attribute designs and sketches, Theory and Applications of Categories 10.3, 2002, 94–112.

[68] M. Johnson, R. Rosebrugh and R. J. Wood, Lenses, Fibrations, and Universal Translations, Mathematical Structures in Computer Science, 22, Cambridge University Press, 2012, 25–42

[69] P. T Johnstone, Sketches of an Elephant: A Topos Theory Compendium, Oxford University Press, 2003.

[70] G. M. Kelly, Complete functors in homology. I. Chain maps and endomorphisms, Proc. Cambridge Philos. Soc. 60, 1964, 721-735.

[71] R. Lavendhomme, La notion d'idéal dans la théorie des catégories, Ann. Soc. Sci. Bruxelles Sér. 1 79, 1965, 5-25.

[72] N. Leavitt, Will NoSQL databases live up to their promise?, Computer 43.2, 2010, 12–14.

[73] S. Mac Lane, Duality for groups, Bulletin of the American Mathematical Society 56, 1950, 485-516.

[74] S. Mac Lane, Homology, Berlin-Göttingen-Heidelberg, 1963.

[75] S. Mac Lane, Categories for the working mathematician, Second Edition, Graduate Texts in Mathematics 5, Springer-Verlag, New York, 1998.

[76] A. I. Mal'tsev, On the general theory of algebraic systems, Mat. Sbornik, N.S. 35 (77), 1954, 3–20 (in Russian); English translation: American Mathematical Society Translations (2) 27, 1963, 125–341.

[77] N. Martins-Ferreira, A. Montoli and M. Sobral, Semidirect products and split short five lemma in normal categories, Applied Categorical Structures, 2013.

[78] N. R. McConnell and T. Stokes, Radicals of 0-regular algebras, Acta Math. Hungar. 113, 2006, 19-37.

[79] E. Meijer and G. Bierman, A co-relational model of data for large shared data banks, Communications of the ACM 54.4, 2011, 49–58.

[80] B. Mitchell, Theory of categories, Academic Press, 1965.

[81] Planet Cassandra, `http://planetcassandra.org/` (accessed Sept 2014).

[82] D. Puppe, Korrespondenzen in abelschen Kategorien, Mathematische Annalen 148.1, 1962, 1–30.

[83] J. D. H. Smith, Mal'cev varieties, Lecture Notes in Mathematics 554, Springer, 1976.

[84] R. Street, Fibrations in bicategories, Cahiers de Topologie et Géométrie Différentielle Catégoriques, 21.2, 1980, 111–160.

[85] T. Van der Linden, Lecture on ordered sets in homological algebra, 1st Workshop on Mathematical Structures, December 2013, `http://math.sun.ac.za/cms/`.

[86] T. Weighill, On torsion theories relative to a class of morphisms, Category Theory Seminar, Louvain-la-Neuve, June 2014.

[87] H. Werner, A Mal'cev condition for admissible relations, Algebra Universalis 3, 1973, 263.

[88] M. Widenius and D. Axmark, MySQL reference manual: documentation from the source, O'Reilly Media Inc., 2002.

[89] O. Wyler, Top categories and categorical topology, General Topology and its Applications 1.1, 1971, 17–28.