Galois Theory, Hopf Algebras, and Semiabelian Categories

FIELDS INSTITUTE COMMUNICATIONS

THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCES

Galois Theory, Hopf Algebras, and Semiabelian Categories

George Janelidze
Bodo Pareigis
Walter Tholen
Editors

American Mathematical Society
Providence, Rhode Island

The Fields Institute for Research in Mathematical Sciences

The Fields Institute is named in honour of the Canadian mathematician John Charles Fields (1863–1932). Fields was a visionary who received many honours for his scientific work, including election to the Royal Society of Canada in 1909 and to the Royal Society of London in 1913. Among other accomplishments in the service of the international mathematics community, Fields was responsible for establishing the world's most prestigious prize for mathematics research—the Fields Medal.

The Fields Institute for Research in Mathematical Sciences is supported by grants from the Ontario Ministry of Training, Colleges and Universities, and the Natural Sciences and Engineering Research Council of Canada. The Institute is sponsored by Carleton University, McMaster University, the University of Ottawa, the University of Toronto, the University of Waterloo, the University of Western Ontario, and York University. In addition there are several affiliated universities and corporate sponsors, from Canada and the United States.

2000 *Mathematics Subject Classification.* Primary 08Bxx, 12Hxx, 13Bxx, 14Lxx, 16Dxx, 17Bxx, 18-xx, 19Dxx, 22Axx.

Library of Congress Cataloging-in-Publication Data

Galois theory, Hopf algebras, and semiabelian categories / George Janelidze, Bodo Pareigis, Walter Tholen, editors.
p. cm. — (Fields Institute Communications; v. 43)
Includes bibliographical references.
ISBN 0-8218-3290-5 (acid-free paper)
1. Differential algebra. 2. Galois theory. 3. Hopf algebras. 4. Rings (Algebra) I. Janelidze, G. (George), 1952– II. Pareigis, Bodo. III. Tholen, W. (Walter), 1947– IV. Series.
QA247.4.G35 2004
512′.32–dc22 2004050271
CIP

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
This publication was prepared by The Fields Institute.
Visit the AMS home page at `http://www.ams.org/`

10 9 8 7 6 5 4 3 2 1 09 08 07 06 05 04

Contents

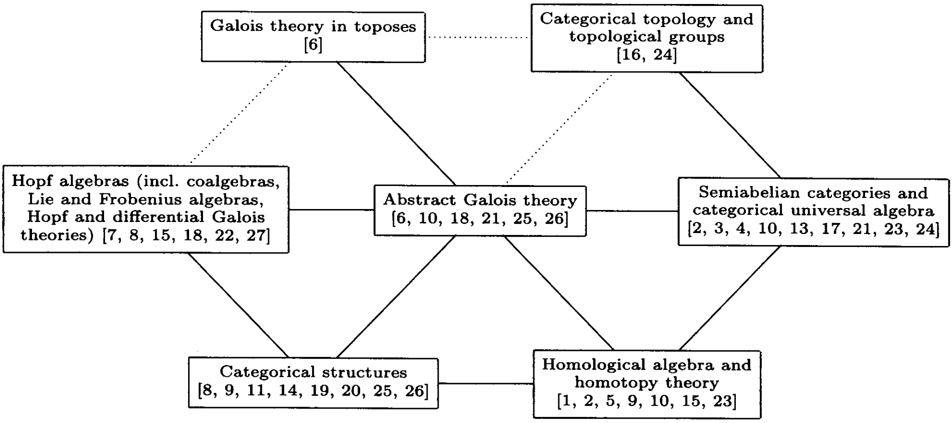
Preface

During the week of September 23–28, 2002, the editors of this volume organized a *Workshop on Categorical Structures for Descent and Galois Theory, Hopf Algebras, and Semiabelian Categories* at the Fields Institute for Research in Mathematical Sciences in Toronto. The goal of the Workshop was to bring together researchers working in the quite distinct but nevertheless interrelated and partly overlapping areas mentioned in its title. The meeting was attended by almost eighty mathematicians from various research communities and boasted twenty invited lectures and numerous contributed talks that led to an inspiring atmosphere of learning and scientific exchange.

This volume can only partially reflect the Workshop's themes but should nevertheless give the reader a good idea about the current connections among abstract Galois theories, Hopf algebras, and semiabelian categories. Here is a very brief indication of the origins of those connections. Hopf algebras arrived to the Galois theory of rings as early as the 1960s — independently of, but in fact similarly to, the way in which algebraic group schemes were introduced to the theory of étale coverings in algebraic geometry. Galois theory, in turn, was extended to elementary toposes and was then formulated in purely categorical contexts. Eventually it became general enough to even include abstractions of the theory of central extensions, to mention only one of various fairly recent developments. In fact, classically, central extensions of groups together with the homology functors $H_1(-, \mathbf{Z})$ and $H_2(-, \mathbf{Z})$ can be used to begin homological algebra, just like covering spaces together with the homotopy functors π_0 and π_1 are the starting gadgets of homotopy theory. Finally, during the past four years semiabelian categories have emerged as a very good environment in which to pursue not just basic modern algebra but in fact homological algebra of groups and other non-abelian structures categorically.

Given the diversity of the backgrounds of the presenters at the Workshop, this volume cannot be expected to contain a homogeneous sequence of chapters on its themes. Rather, the reader will find a collection of beautiful but fairly independent articles on selected topics in algebra, topology, and pure category theory that should seriously contribute to the categorical unification of the subjects in question. The survey articles contained in this volume should be particularly helpful in this regard.

A rough general "map" of the topics/articles presented in this volume may be displayed as follows, with the numbers referring to the (alphabetical) list of contributions contained in the volume. Most of the papers are mentioned more than once. Solid lines represent links explicitly discussed in this volume, while dotted lines indicate other known links.



1. M. Barr, *Algebraic cohomology: the early days*
2. F. Borceux, *A survey of semi-abelian categories*
3. D. Bourn, *Commutator theory in regular Mal'cev categories*
4. D. Bourn and M. Gran, *Categorical aspects of modularity*
5. R. Brown, *Crossed complexes and homotopy groupoids as non commutative tools for higher dimensional local-to-global problems*
6. M. Bunge, *Galois groupoids and covering morphisms in topos theory*
7. S. Caenepeel, *Galois corings from the descent theory point of view*
8. B. Day and R. H. Street, *Quantum categories, star autonomy, and quantum groupoids*
9. J. W. Duskin, R. W. Kieboom, and E. M. Vitale, *Morphisms of 2-groupoids and low-dimensional cohomology of crossed modules*
10. M. Gran, *Applications of categorical Galois theory in universal algebra*
11. C. Hermida, *Fibrations for abstract multicategories*
12. J. Huebschmann, *Lie-Rinehart algebras, descent, and quantization*
13. P. T. Johnstone, *A note on the semiabelian variety of Heyting semilattices*
14. G. M. Kelly and S. Lack, *Monoidal functors generated by adjunctions, with applications to transport of structure*
15. M. Khalkhali and B. Rangipour, *On the cyclic homology of Hopf crossed products*
16. G. Lukács, *On sequentially h-complete groups*
17. J. L. MacDonald, *Embeddings of algebras*
18. A. R. Magid, *Universal covers and category theory in polynomial and differential Galois theory*
19. N. Martins-Ferreira, *Weak categories in additive 2-categories with kernels*
20. T. Palm, *Dendrotopic sets*
21. A. H. Roque, *On factorization systems and admissible Galois structures*
22. P. Schauenburg, *Hopf-Galois and bi-Galois extensions*
23. J. D. H. Smith, *Extension theory in Mal'tsev varieties*
24. L. Sousa, *On projective generators relative to coreflective classes*
25. J. J. Xarez, *The monotone-light factorization for categories via preorders*
26. J. J. Xarez, *Separable morphisms of categories via preordered sets*
27. S. Yamagami, *Frobenius algebras in tensor categories and bimodule extensions*

We express our sincere thanks to the Fields Institute for hosting and supporting the Workshop and publishing this volume. We are particularly grateful to Ms. Debbie Iscoe for her help in preparing the files. We also thank the Faculty of Arts of York University for additional financial assistance.

George Janelidze
Bodo Pareigis
Walter Tholen

Fields Institute Communications
Volume **43**, 2004

Algebraic Cohomology: The Early Days

Michael Barr
Department of Mathematics and Statistics
McGill University
805 Sherbrooke W.
Montreal, QC, H3A 2K6 Canada
barr@barrs.org

Abstract. This paper will survey the various definitions of homology theories from the first Eilenberg-Mac Lane theories for group cohomology through the Cartan-Eilenberg attempt at a uniform (co-)homology theory in algebra, cotriple cohomology theories and the various acyclic models theorems that tied them all together (as much as was possible).

1 Introduction

This paper is a mostly historical introduction to the topic of cohomology theories in algebra between 1940 and 1970 when my interests turned elsewhere. This is not to suggest that progress stopped that year, but that I did not keep up with things like crystalline cohomology, cyclic cohomology, etc., and therefore will have nothing to say about them. Much of what I report, I was directly involved in, but anything earlier than about 1962 is based either on the written record or on hearing such people as Samuel Eilenberg and Saunders Mac Lane reminisce about it. Therefore I report it as true to the best of my belief and knowledge.

I have also omitted any mention of sheaf cohomology. I had nothing to do with it and did not know the people most associated with it—Grothendieck, Godement and others. This was much more highly associated with developments in category theory which I knew little about until after the time frame I am dealing with here. The first development here was Mac Lane's paper [1950] which was the first paper to discover universal mapping properties and also attempt to define what we now call abelian categories, later given their full definition in [Buchsbaum, 1956] and [Grothendieck, 1957].

2000 *Mathematics Subject Classification.* Primary 01A60, 18-03; Secondary 18G10, 18G30, 18G35.

Key words and phrases. Homology and cohomology, triples, acyclic models.

This research has been supported by the NSERC of Canada.

1.1 Acknowledgment. I would like to thank the referee who took an inexcusably careless draft and read it with care and many—far too many than should have been necessary—valuable suggestions for improvements. Any remaining errors and obscurities are, of course, mine.

2 Eilenberg-Mac Lane cohomology of groups

2.1 The background. Eilenberg escaped from Poland in 1939 and spent the academic year 1939-40 at the University of Michigan. Mac Lane, then a junior fellow at Harvard, was invited to speak at Michigan. He was attempting to compute what I will call the Baer group of a group π with coefficients in a π-module A. This group, which I will denote $B(\pi, A)$, can be described as follows. Consider an exact sequence

$$0 \longrightarrow A \longrightarrow \Pi \longrightarrow \pi \longrightarrow 1$$

(Of course, that terminology did not exist in those days; Mac Lane—and Baer—would have said that A was a commutative normal subgroup of Π and $\pi = \Pi/A$.) Since A is normal subgroup of Π, it is a Π-module by conjugation. Since A is abelian, the action of A on itself is trivial, so that the Π-action induces a π-action on A, which may, but need not be the action we began with. Then $B(\pi, A)$ is the class of all such exact sequences that induce the π-action we started with. Say that the sequence above is equivalent to

$$0 \longrightarrow A \longrightarrow \Pi' \longrightarrow \pi \longrightarrow 1$$

if there is a homomorphism (necessarily an isomorphism) $f : \Pi \longrightarrow \Pi'$ such that

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r] \arrow[d, "\mathrm{id}"] & \Pi \arrow[r] \arrow[d, "f"] & \pi \arrow[r] \arrow[d, "\mathrm{id}"] & 0 \\
0 \arrow[r] & A \arrow[r] & \Pi' \arrow[r] & \pi \arrow[r] & 0
\end{tikzcd}$$

commutes. Then the equivalence classes of such sequences are a set. Moreover there is a way of adding two such equivalence classes that makes $B(\pi, A)$ into an abelian group. If $0 \longrightarrow A \longrightarrow \Pi \longrightarrow \pi \longrightarrow 0$ and $0 \longrightarrow A \longrightarrow \Pi' \longrightarrow \pi \longrightarrow 0$ are two such sequences, we first form the pullback $P = \Pi \times_\pi \Pi'$, let S be the subgroup of all elements of P of the form $(a, -a)$ for $a \in A$. It is immediate that S is a normal subgroup of P and then we can let $\Pi'' = P/S$. It is easy to see that the two maps $A \longrightarrow \Pi \longrightarrow P$ and $A \longrightarrow \Pi' \longrightarrow P$ are rendered equal by the projection $P \longrightarrow P/S$ so there is a canonical map $A \longrightarrow \Pi''$. Also the map $P \longrightarrow \pi$ vanishes on S and hence induces $\Pi'' \longrightarrow \pi$. A simple computation shows that $0 \longrightarrow A \longrightarrow \Pi'' \longrightarrow \pi \longrightarrow 0$ is exact. That defines the sum of the two sequences. The negative of the sequence $0 \longrightarrow A \longrightarrow \Pi \longrightarrow \pi \longrightarrow 0$ is simply the sequence in which the inclusion of $A \longrightarrow \Pi$ is negated. Since not every element has order 2 this also provides an example showing that equivalence of sequences is not simply isomorphism of the middle term.

Note that in a sequence the kernel A is not merely an abelian group, but also a π-module. This is because Π acts on the normal subgroup A by conjugation and A, being abelian, acts trivially on itself so that action extends to an action of π. The Baer group is the set of equivalence classes of extensions that induce the given module structure on A as a π-module.

Mac Lane's talk included the computations he was carrying out to compute the Baer group. It would have been along these lines: an extension determines and is determined by a function $\pi \times \pi \longrightarrow A$ satisfying a certain condition and two such functions determined equivalent extensions if there was a function $\pi \longrightarrow A$ that generated the difference of the two original functions. Details will be given later.

Apparently during Mac Lane's talk, Eilenberg noted that some of the calculations that Mac Lane was carrying out were the same as those he was doing in connection with certain cohomology groups. As they remarked in one of their first papers [Eilenberg and Mac Lane, 1942b], "This paper originated from an accidental observation that the groups obtained by Steenrod [1940] were identical with some groups that occur in the purely algebraic theory of *extension of groups*." The computations by Steenrod refer to some first cohomology classes (they are described as homology classes of infinite cycles, but I assume that they were really based on duals of chain groups since the dual of an infinite sum is an infinite product, whose elements can look like infinite sums of chains, especially if you do not have a clear idea of contravariant functors) of the complements of solenoids embedded in a 3-sphere. This is likely the second cohomology group of the solenoid itself. See [Eilenberg and Mac Lane, 1941] for more on infinite cycles.

Later in the same paper, in a footnote, they remark, "Group extensions are discussed by Baer [1934], Hall [1938], Turing [1938], Zassenhaus [1937] and elsewhere. Much of the discussion in the literature treats the case in which G but not H is assumed to be abelian and in which G is not necessarily in the centre of H." What the latter sentence means is that they were looking at extensions of the form

$$0 \longrightarrow G \longrightarrow H \longrightarrow H/G \longrightarrow 0$$

in which G and H are abelian, while the literature was mainly discussing the case that G was an abelian normal subgroup of H and the conjugation action of H (and therefore H/G) on G was non-trivial.

When I gave my talk, I said (and believed, relying on what I thought Eilenberg had said to me 30 years ago, from what he remembered from 30 years before that) that what Eilenberg was working on was calculating the cohomology groups of $K(\pi, 1)$, a space that had fundamental group π and no other homotopy. This was described as the main motivation in the later joint work [1947a, b]. But this goes back only to a paper of Hopf's from 1942 in which he showed that if X is a suitable space, then the fundamental group $\pi_1(X)$ determines the cokernel of the map from $\pi_2(X) \longrightarrow H_2(X)$. Thus in the particular case of a $K(\pi, 1)$, whose π_2 is 0, Hopf's result says that $\pi_1(X)$ determines $H_2(X)$ (as well, of course, as $H_1(X)$, which is the commutator quotient of π_1). It is not clear quite when it was realized that π determines all the homology and cohomology groups of a $K(\pi, 1)$, but it must have been before 1947. That paper came after Hochschild's cohomology theory for commutative algebras, which had, it would appear, no connection with topology. I am indebted to Johannes Hübschmann for pointing out some of the pre-history that I was previously unaware of.

2.2 The Eilenberg-Mac Lane groups. Here is a brief description of the Eilenberg-Mac Lane theory. For a non-negative integer n, an n-cochain on π with coefficients in A is a function $f : \pi^n \longrightarrow A$. When $n = 0$, this is simply an element of A. Denote the set of such functions by $C^n(\pi, A)$. Define a function $\delta_n : C^n(\pi, A)$

$\longrightarrow C^{n+1}(\pi, A)$ by

$$(\delta_n f)(x_1, \ldots, x_{n+1}) = x_1 f(x_2, \ldots, x_{n+1})$$

$$+\sum_{i=1}^{n}(-1)^i f(x_1, \ldots, x_i x_{i+1}, \ldots, x_{n+1}) + (-1)^{n+1} f(x_1, \ldots, x_n)$$

The first term uses the action (which could be the trivial, or identity, action) of π on A. It is a simple exercise to show that $\delta_n \circ \delta_{n-1} = 0$ so that $\mathrm{im}(\delta_{n-1}) \subseteq \ker(\delta_n)$ so that we can define

$$H^n(\pi, A) = \ker(\delta_n)/\mathrm{im}(\delta_{n-1})$$

When $n = 0$, the cohomology is simply the kernel of δ_0. The elements of $\ker(\delta_n)$ are called n-cocycles and those of $\mathrm{im}(\delta_{n-1})$ are called n-coboundaries. Cochains that differ by a coboundary are said to be cohomologous.

2.3 Interpretation of $H^2(\pi, A)$. We will see how the group $H^2(\pi, A)$ is the same as $B(\pi, A)$. Given an extension

$$0 \longrightarrow A \xrightarrow{j} \Pi \xrightarrow{p} \pi \longrightarrow 1$$

let s be a set-theoretic section of p. This means that s is a function, not necessarily a homomorphism, $\pi \longrightarrow \Pi$ such that $p \circ s = \mathrm{id}$. We can and will suppose that $s(1) = 1$. If we write $j(a) = (a, 1)$ for $a \in A$, $s(x) = (0, x)$ for $x \in \pi$, and (a, x) for $(a, 1)(0, x)$ then the fact that every element of Π can be written in the form $j(a)s(x)$ for $a \in A$ and $x \in \Pi$ means that we can identify the underlying set of Π with $A \times \pi$. Since j is a homomorphism, we see that $(a, 0)(b, 0) = (a + b, 0)$. Although s is not a homomorphism, p is and that forces the second coordinate of $(0, x)(0, y)$ to be xy so that we can write $(0, x)(0, y) = (f(x, y), xy)$ with $f \in C^2(\pi, A)$. Note that $(0, x)(a, 1) = (xa, 1)(0, x)$ since the action of π on A is just conjugation. It follows that

$$(a, x)(b, y) = (a, 1)(0, x)(b, 1)(0, y) = (a, 1)(xb, 1)(0, x)(0, y)$$

$$= (a + xb, 1)(f(x, y), xy) = (a + xb, 1)(f(x, y), 1)(0, xy)$$

$$= (a + xb + f(x, y), 1)(0, xy) = (a + xb + f(x, y), xy)$$

Next I claim that the associative law of group multiplication forces f to be a 2-cocyle. In fact, we have for a, b, $c \in A$ and x, y, $z \in \pi$,

$$((a, x)(b, y))(c, z) = (a + xb + f(x, y), xy)(c, z)$$

$$= (a + xb + f(x, y) + xyc + f(xy, z), xyz)$$

while

$$(a, x)((b, y)(c, z)) = (a, x)(b + yc + f(y, z), yz)$$

$$= (a + xb + xyc + xf(y, z) + f(x, yz), xyz)$$

and comparing them we see that

$$f(x, y) + f(xy, z) = xf(y, z) + f(x, yz)$$

which is equivalent to $\delta_2 f = 0$.

The same computation shows that if we begin with a 2-cocycle f and define multiplication on $A \times \pi$ by the formula

$$(a,x)(b,y) = (a + xb + f(x,y), xy)$$

we have an associative multiplication. To see it is a group first suppose that $f(x,1) = 0$ for all $x \in \pi$. It is then an easy computation to show that $(0,1)$ is a right identity and then that $(-x^{-1}a - x^{-1}f(x,x^{-1}), x^{-1})$ is a right inverse for (a,x). We will show below that the group extensions constructed using cohomologous cocycles are equivalent. We use that fact here by showing that any cocycle is cohomologous to one for which $f(x,1) = 0$. From

$$\delta f(x,1,1) = xf(1,1) - f(x,1) - f(x,1) + f(x,1) = 0$$

we see that $f(x,1) = xf(1,1)$. Then for $g(x) = f(x,1)$,

$$(f - \delta g)(x,1) = f(x,1) - xg(1) + g(x) - g(x) = f(x,1) - xf(1,1) = 0$$

Finally, we claim that equivalent extensions correspond to cohomologous cocycles. For suppose that f and f' are 2-cocycles and g is a 1-cochain such that $f - f' = \delta g$. Suppose we denote the two possible multiplications by $*$ and $*'$ so that

$$(a,x) * (b,y) = (a + xb + f(x,y), xy)$$

$$(a,x) *' (b,y) = (a + xb + f'(x,y), xy)$$

Define $\alpha : A \times \pi \longrightarrow A \times \pi$ by $\alpha(a,x) = (a + g(x), x)$. Then

$$\begin{aligned}
\alpha((a,x) * (b,y)) &= \alpha(a + xb + f(x,y), xy) \\
&= (a + xb + f(x,y) + g(xy), xy) \\
&= (a + xb + f'(x,y) + xg(y) + g(x), xy) \\
&= (a + g(x), x) *' (b + g(y), y) \\
&= \alpha(a,x) *' \alpha(b,y)
\end{aligned}$$

which shows that α is a homomorphism, while it is obviously invertible. Clearly the isomorphism commutes with the inclusion of A and the projection on π.

Conversely, suppose f and f' are cocycles that give equivalent extensions. Again, we will denote the two multiplications by $*$ and $*'$. Since we are trying to show that f and f' are cohomologous, we can replace them by cohomologous cocycles that satisfy $f(x,1) = f'(x,1) = 0$ for all $x \in \pi$. The fact that the extensions are actually equivalent (not merely isomorphic) implies that the isomorphism α has the property that elements of A are fixed and that the second coordinate of $\alpha(a,x)$ is x. In particular, we can write $\alpha(0,x) = (g(x), x)$. Then

$$\begin{aligned}
\alpha(a,x) &= \alpha((0,x) * (x^{-1}a, 1)) = \alpha(0,x) *' \alpha(x^{-1}a, 1) \\
&= (g(x), x) *' (x^{-1}a, 1) = (a + g(x), x)
\end{aligned}$$

Then from

$$\begin{aligned}(a+xb+f(x,y)+g(xy),xy) &= \alpha(a+xb+f(x,y),xy)\\ &= \alpha((a,x)*(b,y))\\ &= \alpha(a,x)*'\alpha(b,y)\\ &= (a+g(x),x)*'(b+g(y),y)\\ &= (a+g(x)+xb+xg(y)+f'(xy),xy)\end{aligned}$$

we conclude that $f-f'=\delta g$.

2.4 Interpretations in other low dimensions. The definition of $H^0(\pi,A)$ as the kernel of δ_0 makes it obvious that it is simply

$$\{a\in A\mid xa=a \text{ for all } x\in\pi\}$$

otherwise known as the group of fixed elements of A and denoted A^π.

The kernel of δ_1 consists of those maps $d:\pi\longrightarrow A$ such that $d(xy)=xd(y)+d(x)$. In the case that π acts as the identity on A, this is exactly a homomorphism of π to the abelian group A. For this reason, such a function is sometimes called a crossed homomorphism. We prefer to call it a derivation for compatibility with other examples. The image of δ_0 consists of those derivations of the form $d(x)=xa-a$ for some element $a\in A$. These are called the inner derivations and $H^1(\pi,A)$ is simply the quotient group of derivations modulo inner derivations.

There is also an interpretation of H^3 that gave information (more limited than H^2) for extensions with non-abelian kernels. We give no proofs here, but content ourselves with a brief description.

Suppose that

$$1\longrightarrow G\longrightarrow \Pi\longrightarrow \pi\longrightarrow 1 \qquad (*)$$

is an exact sequence of groups. This means that G is a normal subgroup of Π and the quotient is π. Since G is normal, Π acts on G by conjugation. This gives a map $\Pi\longrightarrow \operatorname{Aut}(G)$ the group of automorphisms of G. Unless G is commutative, this does not vanish on G and hence does not give a natural map $\pi\longrightarrow\operatorname{Aut}(G)$. However, if $\operatorname{In}(G)$ denotes the (normal) subgroup of $\operatorname{Aut}(G)$ consisting of the inner automorphisms, then the composite map $\Pi\longrightarrow\operatorname{Aut}(G)\longrightarrow\operatorname{Aut}(G)/\operatorname{In}(G)$ does vanish on G and hence induces a natural map $\theta:\pi\longrightarrow\operatorname{Aut}(G)/\operatorname{In}(G)$. We will say that θ is induced by $(*)$.

Now let $\mathscr{Z}(G)$ denote the centre of G. One may check that the centre of a normal subgroup is also a normal subgroup so that $\mathscr{Z}(G)$ is also a Π-module, but one that G acts trivially on, so that $\mathscr{Z}(G)$ is a π-module. Now it turns out that any homomorphism $\theta:\pi\longrightarrow\operatorname{Aut}(G)/\operatorname{In}(G)$ induces, after a number of choices, a cocycle in $C^3(\pi,\mathscr{Z}(G))$ whose cohomology class we denote $[\theta]$.

A homomorphism $\theta:\pi\longrightarrow\operatorname{Aut}(G)/\operatorname{In}(G)$ induces a π-module structure on $\mathscr{Z}(G)$. Simply choose, for each $x\in\pi$, an element $\overline{\theta}(x)\in\operatorname{Aut}(G)$ whose class mod $\operatorname{In}(G)$ is $\theta(x)$ and define $xz=\overline{\theta}(x)(z)$ for $z\in\mathscr{Z}(G)$. This is well-defined since inner automorphisms are trivial on the centre.

The main result is contained in the following.

Theorem 2.1 *Suppose G is a group whose centre $\mathscr{Z}(G)$ is a π-module and that $\theta : \pi \longrightarrow \mathrm{Aut}(G)/\mathrm{In}(G)$ is a homomorphism that induces the given action of π on $\mathscr{Z}(G)$. Then*

1. *the cohomology class $[\theta]$ in $H^3(\pi, \mathscr{Z}(G))$ does not depend on the arbitrary choices made;*
2. *the cohomology class $[\theta] = 0$ if and only if θ comes from an extension of the form $(*)$;*
3. *the equivalence class of extensions $(*)$ that give rise to a given cohomology class $[\theta]$ are in 1-1 correspondence with the elements of $H^2(\pi, \mathscr{Z}(G))$; and*
4. *given a π-module A, every element of $H^3(\pi, A)$ has the form $[\theta]$ for some group G and some homomorphism $\theta : \pi \longrightarrow \mathrm{Aut}(G)/\mathrm{In}(G)$ such that $A \cong \mathscr{Z}(G)$ as π-modules, the latter with the π-action induced by θ.*

The 1-1 correspondence in the third point above is actually mediated by a principal homogeneous action of $H^2(\pi, \mathscr{Z}(G))$ on extensions $(*)$ ([Barr, 1969]). The class $[\theta]$ is called the **obstruction** of $[\theta]$ (to arising from an extension) and the last clause says that every element of $H^3(\pi, A)$ is the obstruction to some homomorphism's coming from an extension.

The most striking application of that theory was that if $\mathscr{Z}(G) = 1$, then the equivalence classes of extensions

$$1 \longrightarrow G \longrightarrow \Pi \longrightarrow \pi \longrightarrow 1$$

is in 1-1 correspondence with the homomorphisms

$$\theta : \pi \longrightarrow \mathrm{Aut}(G)/G$$

3 Hochschild cohomology of associative algebras

Gerhard Hochschild defined a cohomology theory for associative algebras in [1945, 1946]. Formally, his definitions look almost identical to those of Eilenberg and Mac Lane.

The setting of this theory is that of an associative algebra Λ over a (commutative) field K. It was later observed that the definitions can be given when K is any commutative ring, but the resultant extension theory is limited to those extensions that split as K-modules. Let A be a two-sided Λ-module. Define $C^n(\Lambda, A)$ to be the set of all n-linear functions $\Lambda^n \longrightarrow A$. Define $\delta_n : C^n(\Lambda, A) \longrightarrow C^{n+1}(\Lambda, A)$ by the formula

$$\begin{aligned}(\delta_n f)(x_1, \ldots, x_{n+1}) = x_1 f(x_2, \ldots, x_{n+1}) \\ + \sum_{i=1}^{n} (-1)^i f(x_1, \ldots, x_i x_{i+1}, \ldots, x_{n+1}) \\ + (-1)^{n+1} f(x_1, \ldots, x_n) x_{n+1}\end{aligned}$$

This differs from the formula for group cohomology only in that the last term is multiplied on the right by x_{n+1}. One can make the formulas formally the same by making any left π-module into a two-sided π-module with the identity action on the right. Alternatively, you can use two-sided modules, using the same coboundary formula as for the Hochschild cohomology. This does not really give a different theory, since you can make a two-sided π-module into a left π-module by the formula $x * a = xax^{-1}$ without changing the cohomology.

Much the same interpretation of the low dimensional cohomology holds for the Hochschild cohomology as for groups. Derivations are defined slightly differently: $d(xy) = xd(y) + d(x)y$, that is the Leibniz formula for differentials. Commutative normal subgroups are replaced by ideals of square 0.

The limitation to extensions that split as modules was avoided by having two kinds of cocycles (in degree 2; n in degree n), one to express the failure of additive splitting and one for the multiplication. This was first done by Mac Lane [1958] and then generalized to algebras in [Shukla 1961].

4 Chevalley-Eilenberg cohomology of Lie algebras

The third cohomology theory that was created during the half decade after the war was the Chevalley-Eilenberg cohomology of Lie algebras [1948]. The formulas are a little different, although the conclusions are much the same. If $\mathfrak{g}$ is a Lie algebra over the field K, then a $\mathfrak{g}$-module A is abelian group with an action of $\mathfrak{g}$ that satisfies $[x, y]a = x(ya) - y(xa)$. It can be considered a two-sided module by defining $ax = -xa$, a fact we will make use of when discussing the Cartan-Eilenberg cohomology. A derivation $d : \mathfrak{g} \longrightarrow A$ satisfies $d[xy] = xd(y) - yd(x)$. An n-cochain is still an n-linear map $f : \mathfrak{g}^n \longrightarrow M$, but it is required to alternate. That is $f(x_1, \ldots, x_n) = 0$ as soon as two arguments are equal. The coboundary formula is also quite different. In fact the coboundary formulas used for groups and associative algebras would not map alternating functions to alternating functions. The definition found by Chevalley and Eilenberg is given by the following formula, in which the hat, ^, is used to denote omitted arguments.

$$\delta f(x_0, \ldots, x_n) = \sum_{i=1}^{n} (-1)^i x_i f(x_0, \ldots, \hat{x}_i, \ldots, x_n)$$
$$+ \sum_{0 \le i < j \le n} (-1)^{i+j} f([x_i, x_j], x_0, \ldots, \hat{x}_i, \ldots, \hat{x}_j, \ldots, x_n)$$

The formula arose naturally as the infinitesimal version of the de Rham coboundary formula on Lie groups.

The same kinds of interpretations in low dimensions hold as in the other two cases.

4.1 Comments on these definitions. Two of these three definitions arose from topology in which homology and cohomology were, by the 1940s, coming to be well understood. The cohomology of π is the cohomology of the space $K(\pi, 1)$ and the cohomology of a real or complex Lie algebra is that of the corresponding Lie group. As far as I know, there was no topological motivation behind Hochschild's theory. Of course, the interpretations in dimensions ≤ 3 presumably gave some confidence that the basic theory was correct. Nonetheless, the definitions, viewed as purely algebraic formulas, were inexplicable and *ad hoc*. Among the questions one might raise was the obvious, "What is a module?" To this question, at least, we will give Beck's surprising and entirely convincing answer.

5 Cartan-Eilenberg cohomology

When, as a graduate student in 1959, I took a course from David Harrison on homological algebra, the definitions of cohomology of groups and associative algebras given above were the definitions I learned. (If Lie algebras were mentioned, I do not recall it.) The course was mainly concerned with Ext, Tor and the like, leading to a proof of the Auslander-Buchsbaum theorem. I imagine the book of Cartan-Eilenberg [1956] was mentioned, but I do not think I ever looked at it. The purchase date inscribed in own copy is April, 1962, just after I had finished typing my thesis. My thesis was on commutative algebra cohomology, which I will discuss later, and the methods of that book do not work in that case for reasons I will explain.

When I got to Columbia as a new instructor, the Cartan-Eilenberg book was the bible of homological algebra and I gradually learned its methods. Basically, they had found a uniform method for defining cohomology theories that included the three theories I have described. Only later did it become apparent their methods applied only to the three cohomology theories it was based on and not on any other.

The way that Cartan and Eilenberg had proceeded was based on the observation that in all three cases, for each object X of the category of interest, there was an enveloping algebra X^e that had the property that the coefficient modules for the cohomology were exactly the X^e-modules. And in all three cases, there was some module—call it X_J—for which the cohomology groups could be described as $\mathrm{Ext}_{X^e}(X_J, M)$.

5.1 Comparison of Cartan-Eilenberg with earlier theories. Here is how you see that the Cartan-Eilenberg (CE) cohomology of a group is the same as the Eilenberg-Mac Lane (EM) cohomology. To compute the EM cohomology of a group π with coefficients in a π-module A, you let $C^n(\pi, A) = \mathrm{Hom}_{\mathsf{Set}}(\pi^n, A)$. By using various adjunctions, we see that

$$\begin{aligned} C^n(\pi, A) = \mathrm{Hom}_{\mathsf{Set}}(\pi^n, A) &\cong \mathrm{Hom}_{\mathsf{Ab}}(\mathbf{Z}(\pi^n), A) \\ &\cong \mathrm{Hom}_{\mathsf{Ab}}(\mathbf{Z}(\pi)^{\otimes n}, A) \cong \mathrm{Hom}_{\mathbf{Z}(\pi)}(\mathbf{Z}(\pi)^{\otimes(n+1)}, A) \end{aligned}$$

Here $M^{\otimes n}$ stands for the nth tensor power of a module M. It is easy to compute that the boundary operator $\delta : C^n(\pi, A) \longrightarrow C^{n+1}(\pi, A)$ is induced by the map $\partial : \mathbf{Z}(\pi)^{\otimes(n+2)} \longrightarrow \mathbf{Z}(\pi)^{\otimes(n+1)}$ defined by

$$\begin{aligned} \partial(x_0 \otimes \cdots \otimes x_{n+1}) = &\sum_{i=0}^{n} (-1)^i x_0 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_{n+1} \\ &+ (-1)^{n+1} x_0 \otimes \cdots \otimes x_n \end{aligned}$$

I claim that if you set $C_n(\pi) = \mathbf{Z}(\pi)^{\otimes(n+1)}$ with the boundary operator ∂, the resultant complex is a projective resolution of $\mathbf{Z}$ made into a π-module by having each element of π act as the identity. Let us note that the action of π on $\mathbf{Z}(G)^{\otimes(n+1)}$ is on the first coordinate, in accordance with the adjunction isomorphism above. Thus $\mathbf{Z}(\pi)^{\otimes(n+1)} = \mathbf{Z}(\pi) \otimes \mathbf{Z}(\pi)^{\otimes n}$ and $\mathbf{Z}(\pi)^{\otimes n}$ is just the free abelian group generated by π^n and hence $\mathbf{Z}(\pi)^{\otimes(n+1)}$ is the free π-module generated by π^n. The augmentation $\mathbf{Z}(\pi) \longrightarrow \mathbf{Z}$ is the linear map that takes the elements of π to the integer 1. In order to show that $\mathrm{Ext}_\pi(\mathbf{Z}, M)$ is the Eilenberg-Mac Lane cohomology of π with coefficients in M, it is sufficient to show that the augmented chain complex

$$\cdots \longrightarrow \mathbf{Z}(\pi)^{\otimes(n+1)} \longrightarrow \mathbf{Z}(\pi)^{\otimes n} \longrightarrow \cdots \longrightarrow \mathbf{Z}(\pi) \longrightarrow \mathbf{Z} \longrightarrow 0$$

is acyclic (that is, is an exact sequence), which implies that the unaugmented complex is a projective resolution of $\mathbf{Z}$ as a $\mathbf{Z}(\pi)$-module. We will actually show that the augmented complex is contractible as a complex of $\mathbf{Z}$-modules, although not as a complex of $\mathbf{Z}(\pi)$-modules. Define $s_{-1} : \mathbf{Z} \longrightarrow \mathbf{Z}(\pi)$ to be the linear map that sends 1 to $1 \in \pi$ (that is, the identity of π). For $n > 0$, define $s_{n-1} : \mathbf{Z}(\pi)^{\otimes n} \longrightarrow \mathbf{Z}(\pi)^{\otimes(n+1)}$ to be the linear map that takes $x_1 \otimes \cdots \otimes x_n$ to $1 \otimes x_1 \otimes \cdots \otimes x_n$. Then $\partial_0 \circ s_{-1} = \mathrm{id}$ clearly. For $n = 1$, we have

$$(\partial_1 \circ s_0 + s_{-1}\partial_0)(x) = \partial_1(1 \otimes x) + s_{-1}(1) = x - 1 + 1 = x$$

For $n > 1$,

$$\begin{aligned}\partial_n \circ s_{n-1}(x_1 \otimes \cdots \otimes x_n) &= \partial_n(1 \otimes x_1 \otimes \cdots \otimes x_n)\\ &= x_1 \otimes \cdots \otimes x_n\\ &\quad + \sum_{i=1}^{n-1}(-1)^i(1 \otimes x_1 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_n)\\ &\quad + (-1)^n 1 \otimes x_1 \otimes \cdots \otimes x_{n-1}\end{aligned}$$

while

$$\begin{aligned}s_{n-2} \circ \partial_{n-1}(x_1 \otimes \cdots \otimes x_n) &= s_{n-2}\left(\sum_{i=1}^{n-1}(-1)^{i-1}(x_1 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_n)\right)\\ &\quad + (-1)^{n-1}s_{n-2}(x_1 \otimes \cdots \otimes x_{n-1})\\ &= \sum_{i=1}^{n-1}(-1)^{i-1}(1 \otimes x_1 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots x_n)\\ &\quad + (-1)^{n-1} 1 \otimes x_1 \otimes \cdots \otimes x_{n-1}\end{aligned}$$

and adding them up, we see that $\partial_n \circ s_{n-1} + s_{n-2} \circ \partial_{n-1} = \mathrm{id}$.

Thus it follows that the Eilenberg-Mac Lane cohomology groups of a group π with coefficients in a π-module A are just $\mathrm{Ext}_{\mathbf{Z}(\pi)}(\mathbf{Z}, A)$.

In the case of associative K-algebras, the development is similar. In this paragraph, we will denote by $\otimes$, the tensor product $\otimes_K$. A two-sided Λ-module is a left $\Lambda^{\mathrm{e}} = \Lambda \otimes \Lambda^{\mathrm{op}}$-module. The Hochschild cochain complex with coefficients in a module A is shown to be the complex $\mathrm{Hom}_{\Lambda^{\mathrm{e}}}(C_\bullet, A)$ in which $C_n = \Lambda^{\otimes(n+2)}$ with boundary $\partial = \partial_n : C_n \longrightarrow C_{n-1}$ given by

$$\partial(x_0 \otimes \cdots \otimes x_{n+1}) = \sum_{i=0}^{n}(-1)^i x_0 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_{n+1}$$

By a similar formula to the group case, one shows that this chain complex augmented over Λ is linearly (actually even right Λ-linearly) contractible and therefore a projective resolution of Λ as a Λ^{e}-module. It should be observed that $\Lambda \otimes \Lambda \cong \Lambda \otimes \Lambda^{\mathrm{op}}$ as a Λ^{e}-module, although not, of course, as an algebra. The upshot is that the Hochschild cohomology of Λ with coefficients in a two-sided Λ-module A is simply $\mathrm{Ext}_{\Lambda^{\mathrm{e}}}(\Lambda, A)$.

The analysis of the Lie algebra case is somewhat more complicated (because the coboundary operator is more complicated) but the outcome is the same. If $\mathfrak{g}$ is a K-Lie algebra, the category of $\mathfrak{g}$-modules is equivalent to the category of left modules over the enveloping associative algebra $\mathfrak{g}^{\mathrm{e}}$. This is the quotient of the tensor algebra over the K-module underlying $\mathfrak{g}$ modulo the ideal generated by all

$x \otimes y - y \otimes x - [x, y]$, for $x, y \in \mathfrak{g}$. Then one can show that the cochain complex comes from a projective resolution of K—with trivial $\mathfrak{g}$-action—so that the cohomology with coefficients in the $\mathfrak{g}$-module A is $\mathrm{Ext}_{\mathfrak{g}^e}(K, A)$.

5.2 Comments. It was certainly elegant that Cartan and Eilenberg were able to find a single definition that gave cohomology theories of the previous decade in all three cases. Nonetheless there were three *ad hoc* features of their definitions that rendered their answer less than satisfactory:

1. what is a module;
2. what is X^e; and
3. what is X_J?

But the least satisfactory aspect of their definition did not become clear right away. It was those three cohomology theories ended up as the only ones for which the Cartan-Eilenberg theory was correct. This first showed up in Harrison's theory for commutative algebras.

6 The Harrison cohomology theory

Around 1960, Dave Harrison [1962] defined a cohomology theory for commutative algebras over a field. His original definition was rather obscure. As modified by the referee (who identified himself to Harrison as Mac Lane) it became somewhat less obscure, but still not obvious. Let R ba commutative K-algebra with K a field. If $0 < i < n$ define a linear map $* : R^{\otimes i} \otimes R^{\otimes (n-i)} \longrightarrow R^{\otimes n}$ by

$$(x_1 \otimes \cdots \otimes x_i) * (x_{i+1} \otimes \cdots \otimes x_n) = (x_1 \otimes (x_2 \otimes \cdots \otimes x_i) * (x_{i+1} \otimes \cdots \otimes x_n))$$
$$+ (-1)^i (x_{i+1} \otimes (x_1 \otimes \cdots \otimes x_i) * (x_{i+2} \otimes \cdots \otimes x_n))$$

which together with

$$(x_1 \otimes \cdots \otimes x_n) * () = () * (x_1 \otimes \cdots \otimes x_n) = (x_1 \otimes \cdots \otimes x_n)$$

defines inductively an operation on strings, called the shuffle.[1] If A a left R-module that is made into a two-sided R-module by having the same operation on both sides. Harrison defined a commutative n-cochain as an f in the Hochschild cochain group $C^n(R, A)$ such that

$$f((x_1 \otimes \cdots \otimes x_i) * (x_{i+1} \otimes \cdots \otimes x_n)) = 0$$

for all $0 < i < n$ and all $(x_1 \otimes \cdots \otimes x_n) \in R^n$. In other words, a commutative cochain was one that vanished on all proper shuffles. The commutative coboundary formula was the same as Hochschild's and it required a non-trivial argument to show that Harrison's cochains are invariant under it. The details are in his 1962 paper and can also be found in [Barr, 2002].

The coboundary of a 0-chain in $C(R, A)$—an element of $a \in A$ is the one chain f for which $f(x) = ax - xa$. But if A has the same action on both sides, this is 0. Thus the cochain complex breaks up into two pieces, the degree 0 piece and the rest. Hence for Harrison, $H^0(R, A) = A$ and $H^1(R, A) = \mathrm{Der}(R, A)$, the

[1] The reason it is called a shuffle is that it is the alternating sum of all possible ways of shuffling the first i cards in an n card deck with the remaining $n - i$. The inductive formula corresponds to the obvious fact that any shuffle can be thought of as consisting of taking the top card from one of the two decks, shuffling the remaining cards and then replacing that top card on top.

group of derivations of R to A. The groups $H^2(R,A)$ and $H^3(R,A)$ have the same interpretations as in the Hochschild theory, but restricted to the case that everything is commutative.

Harrison observed moreover that, provided the characteristic of the underlying ground field is not 2, the group of 2-cochains can be written as the direct sum of the commutative cochains and a complementary summand in such a way that the Hochschild cohomology splits as the direct sum of two groups, one of which is the commutative cohomology and the other is a complementary term. In my Ph.D. thesis, [1962], I pushed this splitting up to degree 4, except that now characteristic 3 also had to be excluded. What I actually showed was the Hochschild cochain complex, truncated to degree 4, could be written as a direct sum of the Harrison cochain subcomplex and a complementary subcomplex and this splits the cohomology up to degree 4. (It might be thought that you need a splitting of the cochain complex up to degree 5 to do this, but that turns out to be unnecessary.) This was used to show various facts about the Harrison cohomology group of which the perhaps the most interesting was that when R is a polynomial ring, then $H^n(R,A) = 0$ for all R-modules A and $n = 2,3,4$. Actually, Harrison had shown this for $n = 2,3$ in a different way, but it was not obvious how to extend his argument to higher dimension.

My proofs were highly computational and it was unclear how to extend any of it to higher dimensions. But I gnawed at it for five years and did eventually [Barr, 1968] find a relatively simple and non-computational construction of an idempotent in the rational group rings of symmetric groups that when applied to the Hochschild chain complex of an algebra over a field of characteristic 0 splits it into a commutative part and a complement, each invariant under the coboundary. This showed that when the ground field has characteristic 0, the Hochschild cohomology splits into two parts; one of the two is the Harrison cohomology and the other is a complement. Later, Gerstenhaber and Shack [1987], by discovering and factoring the characteristic polynomial of these idempotents, discovered a "Hodge decomposition" of the Hochschild groups:

$$H^n(R,A) = \sum_{i=1}^{n} H^{n\,i}(R,A)$$

where $H^{n\,1}(R,A)$ is the Harrison group and the other summands are part of an infinite series whose nth piece vanishes in dimensions below n, so that in each degree the sum is finite.

Harrison's original paper also contained an appendix—written by me—that used the same ideas from Mac Lane's 1958 paper to deal with the case of a commutative algebra over a general coefficient ring. However, the referee—who identified himself to Harrison as Mac Lane—insisted on an appendectomy and the paper is now lost. From this, I conclude that Mac Lane felt the approach was likely a dead end and I tend to agree.

6.1 Does the Harrison cohomology fit the Cartan-Eilenberg model? As I have already hinted, Harrison was not a fan of the Cartan-Eilenberg model of what a cohomology theory was. I do not recall that he ever mentioned it, I do not think I was even aware of it before arriving at Columbia in the fall of 1962 and I do not know if Harrison had ever considered whether his theory fit into it. It is a bit odd, since it seems clear that Cartan and Eilenberg believed that their book defined what a cohomology theory was. They scarcely mentioned the older definitions in their book and, as far as I can tell (the book lacks a bibliography), they did not even cite either Eilenberg-Mac Lane paper [1947a,b]. But Harrison's definition was in the style of the older definitions. To put his theory into the Cartan-Eilenberg framework, there would have to be, for each commutative ring R, an enveloping algebra R^e such that left R^e-modules were the same as left R-modules and a module R_J for which $H^n(R, A) = \mathrm{Ext}^n_{R^e}(R_J, A)$. Since left R-modules were the same as left R^e-modules, we would have to take R^e to be R or something Morita equivalent to it. But Ext is invariant under Morita equivalence so nothing can be gained by using anything but R. As for R_J, we leave that aside for the moment and point out that one consequence of the cohomology being an Ext is that it vanishes whenever A is injective. This is true for cohomology theories of groups, associative algebras, and Lie algebras, but is not true for commutative algebras. The easiest example is the following. Let K be any field. It is easy to see (and well-known) that the algebra of dual numbers $R = K[x]/(x^2)$ is self-injective. On the other hand, we have the non-split exact sequence

$$0 \longrightarrow x^2 \cdot K[x]/x^4 \longrightarrow K[x]/(x^4) \longrightarrow K[x]/(x^2) \longrightarrow 0$$

of commutative rings, whose kernel is, as an R-module, isomorphic to R and hence is injective. Thus $H^2(R, R) \neq 0$. See [Barr, 1968a]. This example doomed the attempted redefinition of commutative cohomology that appeared in [Barr, 1965a, 1965b].

Here is an explanation of what goes wrong. If R is a commutative K-algebra, then the chain complex that has $R^{\otimes(n+2)}$ with the Hochschild boundary operator, is a projective resolution of R as an $R \otimes R$-module. If A is a symmetric module and B is any two-sided R-module, it is evident that

$$\mathrm{Hom}_{R\otimes R}(B, A) \cong \mathrm{Hom}_R(R \otimes_{R\otimes R} B, A)$$

since $R \otimes_{R\otimes R} B$ is just the symmetrization of B. Since $R \otimes_{R\otimes R} R^{\otimes(n+2)} = R^{\otimes(n+1)}$, the Hochschild cochain complex of R with coefficients in A is

$$\mathrm{Hom}_{R\otimes R}(R^{\otimes(n+2)}, A) \cong \mathrm{Hom}_R(R^{\otimes(n+1)}, A)$$

It follows that the symmetrized chain complex is not generally acyclic; its homology is $\mathrm{Tor}^{R\otimes R}(R, R)$ and that is trivial if and only if R is separable. Even if it is, the quotient complex modulo the shuffles will not be generally acyclic.

7 Cohomology as a functor in the first argument

Although Harrison had an 8 term exact sequence involving the cohomology of R, that of R/I for an ideal I and $\mathrm{Ext}_R(I, -)$, the functoriality of cohomology in its first argument had been mostly ignored. If a connected series of functors is to look like a derived functor in a contravariant variable, it should vanish in all positive dimensions when that variable is free. But cohomology did not vanish when its first variable was free. In fact, when the group, algebra, or Lie algebra is free, the cohomology vanishes in dimensions greater than 1. This is also the case for commutative algebras, but only in characteristic 0. This, as well as other indications, suggested that if one wanted to view cohomology as a functor in the first variable, it would be best to drop the lowest degree term, renumber all the rest by -1 and also change the new lowest degree term from derivations modulo inner derivations to simply derivations. Thus was born the idea of cohomology as the derived functor of derivations, see [Barr and Rinehart, 1964].

Let us call this cohomology with the lowest term dropped and the next one modified the dimension-shifted cohomology. For the commutative cohomology, as already noted, there are no inner derivations and the lowest degree term is merely the coefficient module, so dropping it entails no loss of information. For the others, there is some cost. For associative and Lie algebras the cohomology vanishes in all dimensions if and only if the algebra is (finite-dimensional and) separable. For other purposes, the dimension-shifted cohomology seems better. For instance, the Stallings-Swan theorem states that a group is free if and only if its cohomological dimension is 1. Using the dimension-shifted theory, that number changes to 0, which is what you might expect for projectives (of course, every projective group is free).

So the first definition of the dimension-shifted is simply apply the functor $\mathrm{Der}(X,-)$ (here X is the group or algebra or whatever) to the category of X-modules and form the derived functor. This means, given a module A find an injective resolution

$$0 \longrightarrow Q_0 \longrightarrow Q_1 \longrightarrow \cdots \longrightarrow Q_n \longrightarrow \cdots$$

apply $\mathrm{Der}(X,-)$ to get the chain complex

$$0 \longrightarrow \mathrm{Der}(X,Q_0) \longrightarrow \mathrm{Der}(X,Q_1) \longrightarrow \cdots \longrightarrow \mathrm{Der}(X,Q_n) \longrightarrow \cdots$$

and define the cohomology groups to be the cohomology of that cochain complex. The trouble with this definition is that it automatically makes the cohomology vanish when the coefficients are injective and therefore can represent the cohomology only for theories that vanish on injective coefficients.

Before turning to other theories, there is one more point to be made. Since Der is like a homfunctor, it is quite easy to see that it preserves limits. The special adjoint functor theorem implies that it is representable by an object we call Diff (for module of differentials, since differentials are dual to derivations). The map $A \longrightarrow \mathrm{Der}(X,A)$ that sends an element of A to the inner derivation at A is the component at A of a natural transformation $\mathrm{Hom}_{X^e}(X^e,-) \longrightarrow \mathrm{Hom}_{X^e}(\mathrm{Diff},-)$ and is thereby induced by a homomorphism $\mathrm{Diff} \longrightarrow X^e$. In the three classic cases of groups, associative algebras and Lie algebras, this homomorphism is injective and the quotient module X^e/Diff is the heretofore mysterious X_J. Thus in those three cases, not only is the cohomology $\mathrm{Ext}_{X^e}(X_J,A)$, but the dimension-shifted cohomology is $\mathrm{Ext}_{X^e}(\mathrm{Diff},A)$. Thus we have removed one of the *ad hoc* features from the Cartan-Eilenberg theory. It can be replaced by a question as to why the map from $\mathrm{Diff} \longrightarrow X^e$ is injective in those cases, but at least we know where X_J comes from. The second special item is X^e but X^e is determined, up to Morita equivalence, by the fact that X^e-modules are the same as X-modules. Since Ext is invariant under Morita equivalence, that ambiguity is not important. The third *ad hoc* feature of the Cartan-Eilenberg definition is the definition of module and we will explain Jon Beck's surprising and elegant answer to that question next.

8 Beck modules

8.1 The definition of Beck modules. We begin Beck's theory by looking at an example in some detail. Denote by Grp the category of groups. If π is a group, the category Grp/π has as objects group homomorphisms $p : \Pi \longrightarrow \pi$. If $p' : \Pi' \longrightarrow \pi$ is another object, a map $f : p' \longrightarrow p$ is a commutative triangle

$$\begin{tikzcd} \Pi' \arrow[rr, "f"] \arrow[dr, "p'"'] & & \Pi \arrow[dl, "p"] \\ & \pi & \end{tikzcd}$$

Given a π-module A, denote by $A \rtimes \pi$ the extension corresponding to the 0 element of $H^2(\pi, A)$. This means the underlying set is $A \times \pi$ and the multiplication is given by $(a,x)(b,y) = (a + xb, xy)$. The second coordinate projection makes $A \rtimes \pi$ into an object of Grp/π.

Let us calculate the set $\mathrm{Hom}_{\mathsf{Grp}/\pi}(\Pi \xrightarrow{p} \pi, A \rtimes \pi \to \pi)$. A homomorphism $f : \Pi \longrightarrow A \rtimes \pi$ is a function $\Pi \longrightarrow A \times \pi$. But in order to be a morphism in the category, the second coordinate must be p. If we call the first coordinate d, then $f = (d, p)$. In order to be a group homomorphism, we must have, for any $x, y \in \Pi$ that

$$(d(xy), p(xy)) = (d(x), p(x))(d(y), p(y)) = (d(x) + p(x)d(y), p(x)p(y))$$

Since p is already a homomorphism, the second coordinates are equal. As for the first coordinates, the required condition is that $d(xy) = p(x)d(y) + d(x)$. Except for the p on the right hand side, this is just the formula that defines a derivation. But we can use p to make any π-module into a Π-module and then this condition is just that d is a derivation of Π to A. Thus we have shown that

$$\mathrm{Hom}_{\mathsf{Grp}/\pi}(\Pi \xrightarrow{p} \pi, A \rtimes \pi \to \pi) = \mathrm{Der}(\Pi, A)$$

Evidently, $\mathrm{Der}(\Pi, A)$ is an abelian group from the additive structure of A and it is easy to see that a morphism $(\Pi' \to \pi) \longrightarrow (\Pi \to \pi)$ induces not merely a function but a group homomorphism $\mathrm{Der}(\Pi, A) \longrightarrow \mathrm{Der}(\Pi', A)$. This means that $A \rtimes \pi \to \pi$ is actually an abelian group object in Grp/π.

The converse is also true. It is a standard fact about abelian group objects in categories with finite products that a group object G is given by a global section $1 \longrightarrow G$, an inverse map $G \longrightarrow G$ and a multiplication $G \times G \longrightarrow G$ and that the multiplication map is just the product of the two projections. If $p : \Pi \longrightarrow \pi$ is a group object in Grp/π, the terminal object of the category is $\mathrm{id} : \pi \longrightarrow \pi$ and thus the zero map is just a homomorphism $\pi \longrightarrow \Pi$ that splits p. If K is the kernel of p, it is a normal subgroup and we will write its group operation as $+$ and its identity element as 0 even though we have not yet proved it commutative. By extending the notation to the non-commutative case, we can write $\Pi = K \rtimes \pi$. In this notation, the zero morphism is given by $z(x) = (0, x)$. The product in the category Grp/π is just the fibered product over π. Since $\Pi \times_\pi \Pi = K \times K \rtimes \pi$, the multiplication is a homomorphism

$$m : K \times K \rtimes \pi \longrightarrow K \rtimes \pi$$

over π. We can write this as $m(a, b, x) = (f(a, b, x), x)$ for $a, b \in K$ and $x \in K$. Since m preserves the identity of the group object, $m(0, 0, x) = (0, x)$ from which it follows that $f(0,0,x)=0$ for any $x \in \pi$. But then

$$m(a,b,x) = m(a,b,1)m(0,0,1) = (f(a,b,1),1)(f(0,0,x),x)$$

$$= (f(a,b,1),1)(0,x) = (f(a,b,1),x)$$

so that f does not depend on x. Write $f(a,b,1) = a * b$. Then since $(0,x)$ is the identity in the fiber over x, $m(a,0,x) = (a,x) = m(0,a,x)$ so that $a*0 = 0*a = a$. But then

$$m(a,b,x) = m(a,0,1)m(0,b,1)m(0,0,x) = (a*0,1)(0*b,1)(0,x)$$

$$= (a*0+0*b,x) = (a+b,x)$$

while at the same time

$$m(a,b,x) = m(0,b,1)m(a,0,1)m(0,0,x) = (0*b,1)(a*0,1)(0,x)$$

$$= (0*b+a*0,x) = (b+a,x)$$

from which we conclude that $* = +$ and is commutative. The action of Π on K is by conjugation and since K is commutative, this induces an action of π on K. If $q : \Gamma \longrightarrow \pi$ is a group over π, then a morphism $f : (\Gamma \to \pi) \longrightarrow (\Pi \to \pi)$ has the form $f(y) = (dy, qy)$ and to be a homomorphism we must have

$$(d(yy'), q(yy')) = f(yy') = f(y)f(y') = (dy,qy)(dy',qy')$$

$$= (dy + (qy)(dy'), (qy)(qy'))$$

which means that $d(yy') = dy + (qy)(dy')$ which is the definition of a derivation (with respect to q).

This is one case of:

Theorem 8.1 (Beck) *Let $\mathscr{X}$ be one of the familiar categories (groups, algebras and rings, Lie algebras, commutative algebras and rings, Jordan algebras, ...). Then the category of modules over an object X of $\mathscr{X}$ is (equivalent to) the category of abelian group objects in the category $\mathscr{X}/X$. Moreover, let A be an X-module with $Y \longrightarrow X$ the corresponding abelian group object in $\mathscr{X}/X$. Then for an object $Z \longrightarrow X$ of that category,* $\mathrm{Hom}(Z \to X, Y \to X)$ *is canonically isomorphic to* $\mathrm{Der}(Z, A)$.

The upshot of this result is that not only do we know what a module is for an object of any category, we also know what a derivation into that module is. For example, a module over the set I is just an I-indexed family $A = \{A_i \mid i \in I\}$ of abelian groups and $\mathrm{Der}(I,A) = \prod_{i\in I} A_i$. For any $X \longrightarrow I$, which is just an I-indexed family $\{X_i \mid i \in I\}$, one can see that $\mathrm{Der}(X,A) = \prod_{i\in I} A_i^{X_i}$.

After I gave my talk, I had a private discussion with Myles Tierney and Alex Heller. Heller remarked that he first heard the definition of Beck module from Eilenberg and wondered whether the definition was originally his. Tierney, who was a student of Eilenberg's at the same time as Beck, recalled that it taken a couple months for Beck to convince Eilenberg of the correctness of his definition, but that once convinced, Eilenberg embraced it enthusiastically.

As an aside, let me say that this is the definition of bimodule. One-sided modules cannot be got as special cases of this and are apparently a different animal. Apparently one-sided modules are part of representation theory and two-sided modules are part of extension theory.

9 Enter triples

Triples (also known as monads) have been used in many places and for many reasons in category theory (as well as in theoretical computer science, especially in the theory of datatypes). They were originally invented (by Godement in [1958]) for the purpose of describing standard flabby resolutions of sheaves. They were being used for this purpose by Eckmann and his students around 1960 and by Eilenberg and Moore in their Memoir [1965a]. But this was always in additive categories (although the triples were not always additive, but Eilenberg and Moore assumed not only that but that they preserved kernels, although neither hypothesis was necessary for their purposes).

Beck was the first person who used a triple (more precisely, a cotriple) in a non-additive category to define a cohomology theory. Once more, I will illustrate what he did in the category of groups.

9.1 Triples and Eilenberg-Moore algebras. The definitions are available in many places and so I will just give a rapid sketch. If $\mathscr{X}$ is a category, a triple $\mathbf{T}$ on $\mathscr{X}$ consists of (T, η, μ) where T is an endofunctor on $\mathscr{X}$ and η and μ are natural transformations: $\eta : \mathrm{Id} \longrightarrow T$ and $\mu : T^2 \longrightarrow T$ such that the diagrams

$$
\begin{tikzcd}
T^3 \arrow[r, "T\mu"] \arrow[d, "\mu T"'] & T^2 \arrow[d, "\mu"] \\
T^2 \arrow[r, "\mu"'] & T
\end{tikzcd}
\qquad
\begin{tikzcd}
T \arrow[r, "T\eta"] \arrow[dr, "="'] & T^2 \arrow[d, "\mu"] & T \arrow[l, "\eta T"'] \arrow[dl, "="] \\
& T &
\end{tikzcd}
$$

commute. A cotriple is a triple in the dual category. Hence a cotriple $\mathbf{G}$ consists of (G, ϵ, δ) with G an endofunctor and $\epsilon : G \longrightarrow \mathrm{Id}$ and $\delta : G \longrightarrow G^2$ natural transformations such that the dual of the above diagrams commute. Given an adjoint pair $F : \mathscr{X} \longrightarrow \mathscr{Y}$ left adjoint to $U : \mathscr{Y} \longrightarrow \mathscr{X}$, with adjunction arrows $\eta : \mathrm{Id} \longrightarrow UF$ and $\epsilon : FU \longrightarrow \mathrm{Id}$, then $(UF, \eta, U\epsilon F)$ is a triple on $\mathscr{X}$ and $(FU, \epsilon, F\eta U)$ is a cotriple on $\mathscr{Y}$. Peter Huber (a student of Eckmann's) told me once that they had proved this theorem—a simple argument with naturality—because they were having so much trouble verifying the commutations in the diagram above and noticed that all their triples and cotriples were associated with adjoint pairs (see [Huber, 1961]). It is much easier to verify an adjunction than a triple. The converse is also true; every (co-)triple does arise in this way from an adjoint pair. There are two distinct proofs of this; due to Kleisli [1964, 1965] and Eilenberg-Moore [1965b]. For most purposes in algebra, the Eilenberg-Moore algebras are more interesting. But triples have also entered theoretical computer science and there the Kleisli construction is the main one.

9.2 Beck cohomology. As above, it is convenient to illustrate Beck's results in the category of groups. The adjoint pair $F \dashv U$ where $F : \mathsf{Set} \longrightarrow \mathsf{Grp}$ and $U : \mathsf{Grp} \longrightarrow \mathsf{Set}$ is the underlying set functor, gives rise to a cotriple on Grp as described. If π is a group, then $G\pi$ is the free group on the underlying set of π. As with any cotriple, there results a functor that assigns to each group π a simplicial group that has $G^{n+1}\pi$ in degree n. The face operators are constructed from ϵ by $d^i = d^i_n = G^i \epsilon G^{n-i} : G^{n+1} \longrightarrow G^n$ and $s^i = s^i_n = G^i \delta G^{n-i} : G^{n+1} \longrightarrow G^{n+2}$. The naturality and the commuting diagrams satisfied by a cotriple imply the simplicial identities:

$$\begin{array}{rcll} d_n^i \circ d_{n+1}^j & = & d_n^{j-1} \circ d_{n+1}^i & \text{if } 0 \le i < j \le n+1 \\ s_n^j \circ s_{n-1}^i & = & s_n^i \circ s_{n-1}^{j-1} & \text{if } 0 \le i < j \le n \\ d_{n+1}^i \circ s_n^j & = & \begin{cases} s_{n-1}^{j-1} \circ d_n^i & \\ 1 & \\ s_{n-1}^j \circ d_n^{i-1} & \end{cases} & \begin{array}{l} \text{if } 0 \le i < j \le n \\ \text{if } 0 \le i = j \le n \text{ or } 0 \le i-1 = j < n \\ \text{if } 0 < j < i-1 \le n \end{array} \end{array}$$

We now form, for each object X and each X-module A, the cochain complex

$$0 \longrightarrow \operatorname{Der}(GX, A) \longrightarrow \cdots \longrightarrow \operatorname{Der}(G^n X, A) \longrightarrow \operatorname{Der}(G^{n+1} X, A) \longrightarrow \cdots$$

with the coboundary map

$$\sum_{i=0}^{n} (-1)^i \operatorname{Der}(d^i X, A) : \operatorname{Der}(G^n X, A) \longrightarrow \operatorname{Der}(G^{n+1} X, A)$$

Curiously, the δ of the cotriple plays no part in this, but it is absolutely necessary in applying the acyclic models theorem mentioned below.

This gives a uniform treatment for cohomology in any equational category. Beck also showed that the group $H^1(X, A)$ did classify extensions of X with kernel A, as expected. He did not explore the second cohomology that in known cases was relevant to extensions with non-abelian kernels. Later I did in a couple cases and discovered some problems in the general case that showed that the interpretation of the dimension-shifted H^2 could not be quite the same as in the known cases (Jordan algebras supplied the first counter-example).

10 Comparison theorems

Jon Beck and I started thinking in early 1964 about the connection between the various cohomology theories that had been invented in an *ad hoc* way and the Beck (or cotriple) cohomology theories. Let us assume that we are talking about the dimension-shifted versions of the former. So we know that $H^0(X, A) = \operatorname{Der}(X, A)$ for both theories and that both versions of $H^1(X, A)$ classify the same set of extensions and hence that they are at least isomorphic (although naturality was still an issue). But we had no idea whatever about H^2 or any higher dimension. The complexes look entirely different. The Eilenberg-Mac Lane complex has, in degree n, functions of $n+1$ (because of the dimension shift) variables from π to A, while the cotriple complex has functions of one variable from $G^{n+1}\pi$ to A. We spent the fall term of 1964 working on this problem and getting exactly nowhere. For group cohomology, one approach would have been to show that the chain complex

$$\cdots \longrightarrow \operatorname{Diff} G^{n+1}\pi \longrightarrow \operatorname{Diff} G^n\pi \longrightarrow \cdots \longrightarrow \operatorname{Diff} G\pi \longrightarrow \operatorname{Diff} \pi \longrightarrow 0$$

with boundary operator $\sum_{i=1}^{n} (-1)^n \operatorname{Diff} d^i$, is exact; for then the positive part of that complex would have been a projective resolution of $\operatorname{Diff} \pi$. In retrospect, we now know how to do that directly. Fortunately for us we did not find that argument for it would have been a dead end.

Instead, Beck spoke to Harry Appelgate during the term break between 1964 and 1965. Appelgate suggested trying to solve the problem by using acyclic models, one of the subjects of his thesis [1965]. We did and within a few days we had solved the comparison problem for groups and associative algebras. For some reason, we never looked at Lie algebras, while commutative algebras, for the time being, resisted our methods.

The case of commutative algebras was settled in the late 1960s. In finite characteristic, Harrison's definition was not equivalent to the cotriple cohomology, as shown by an example due to Michel André. In characteristic 0, it was shown in [Barr, 1968] that the two theories coincided. It is shown in [Barr, 1996] that the same is true for Lie algebras.

10.1 Acyclic models. Here is a brief description of the acyclic models theorem that we proved and used to compare the Cartan-Eilenberg cohomology with the cotriple cohomology. The context of the theorem is a category $\mathscr{X}$ equipped with a cotriple $\mathbf{G} = (G, \epsilon, \delta)$ and an abelian category $\mathscr{A}$. We are given augmented chain complex functors $K_\bullet = \{K_n\}$, $L_\bullet = \{L_n\} : \mathscr{X} \longrightarrow \mathscr{A}$, for $n \geq -1$. We say that $K_\bullet$ is G-presentable if there is a natural transformation $\theta_n : K_n \longrightarrow K_nG$ for all $n \geq 0$ (note: not for $n = -1$) such that $K_n\epsilon \circ \theta_n = \mathrm{id}$ for all $n \geq 0$. We say that $L_\bullet$ is G-contractible if the complex $L_\bullet G \longrightarrow 0$ has a natural contracting homotopy (which is called s below).

Theorem 10.1 *Suppose that $K_\bullet$ is G-presentable and $L_\bullet$ is G-contractible. Then any natural transformation $f_{-1} : K_{-1} \longrightarrow L_{-1}$ can be extended to a natural transformation $f_\bullet : K_\bullet \longrightarrow L_\bullet$. Any two extensions of f_{-1} are naturally homotopic.*

Proof Here are the two diagrams required for the proof. The computations are straightforward. The map f_n is defined as the composite,

$$\begin{tikzcd}
K_n \arrow[r, "\theta_n"] & K_nG \arrow[d, "d"] & & L_nG \arrow[r, "L_n\epsilon"] & L_n \\
& K_{n-1}G \arrow[rr, "f_{n-1}"] & & L_{n-1}G \arrow[u, "s"] &
\end{tikzcd}$$

while the nth homotopy is defined as the difference of the upper and lower composite in,

$$\begin{tikzcd}
& & L_{n+1}G \arrow[r, "L_{n+1}\epsilon"] & L_{n+1} \\
K_n \arrow[r, "\theta_n"] & K_nG \arrow[r, "f_n - g_n"] \arrow[d, "d"] & L_nG \arrow[u, "s"] & \\
& K_{n-1}G \arrow[ur, "h_{n-1}"] & &
\end{tikzcd}$$

□

Of course, an immediate consequence of this is that if $K_\bullet$ and $L_\bullet$ each satisfy both hypotheses and K_{-1} is naturally isomorphic to L_{-1}, then $K_\bullet$ is naturally homotopic to $L_\bullet$. This result is augmented by

Proposition 10.2 *Suppose that $L_n = L_{-1}G^{n+1}$ with coboundary*

$$\sum_{i=0}^{n}(-1)^i G^i \epsilon G^{n-i}$$

Then $L_\bullet$ is both G-presentable and G-contractible.

In fact, δ is used to show both claims.

Before describing the modern acyclic models theorem that can be used to prove this, we will describe the issues involved. Basically, two things are required to show that a cochain complex functor is cohomologous to the cotriple complex. The first seems somewhat odd at first glance, but is satisfied in examples. That is that the terms of the complex do not depend on the structure of the object in question, but only on the underlying set (or module). It is a curious fact, but obvious from the definition that the structure is used only in the definition of the coboundary homomorphism. The second condition required is that the cohomology of a free object should vanish in positive degrees. This condition is automatic for the cotriple cohomology; hence this is a necessary condition if the two theories are to be equivalent. This is where the problem arises for Harrison's commutative cohomology in finite characteristic. Michel André showed that in characteristic p there was a non-zero cohomology class in degree $2p-1$.

Here is the argument for acyclicity in the case of free groups. We begin by observing that a free group is also free as far as derivations is concerned.

Proposition 10.3 *Suppose that Π is free on basis X and M is a Π-module. Then any function $\tau : X \longrightarrow M$ extends to a unique derivation $\Pi \longrightarrow M$.*

Proof Let $U : \mathsf{Grp} \longrightarrow \mathsf{Set}$ denote the underlying functor. This result follows from the sequence of isomorphisms

$$\begin{aligned}
\mathrm{Der}(\Pi, M) &\cong \mathrm{Hom}_{\mathsf{Grp}/\pi}(\Pi \to \pi, M \rtimes \pi \to \pi) \\
&\cong \mathrm{Hom}_{\mathsf{Set}/U\pi}(X \to U\pi, UM \times U\pi \to U\pi) \\
&\cong \mathrm{Hom}_{\mathsf{Set}}(X, UM)
\end{aligned}$$

□

This implies that $\mathrm{Diff}(\Pi)$ is the free π-module generated by S.

It is not hard to show that $C_\bullet$ is an exact chain complex and hence $C_\bullet(\Pi)$ is a free resolution of $\mathrm{Diff}(\Pi)$. In the case that Π is free, this is then a free resolution of a free module and hence necessarily split. However, we would rather get the extra information available if we know that the splitting is natural, namely that we then get a homotopy equivalence between the two chain complex functors.

We start by defining a homomorphism $\partial : C_0(\Pi) \longrightarrow \mathrm{Diff}(\Pi)$. There is a function $\tau : X \longrightarrow \mathrm{Diff}(\Pi)$ which is the inclusion of the basis. This extends to a derivation $\tau : \Pi \longrightarrow \mathrm{Diff}(\Pi)$ as above. Since $C_0(\Pi)$ is freely generated by the elements of Π, this derivation τ extends to a π-linear function $\partial : C_0(\Pi) \longrightarrow \mathrm{Diff}(\Pi)$. In accordance with the recipe above, ∂ is defined on elements of Π recursively as follows. We will denote by $\langle w \rangle$ the basis element of $C_0(\Pi)$ corresponding to $w \in \Pi$. As above, either $w = 1$ or $w = xv$ or $w = x^{-1}v$ for some $x \in X$ and some $v \in \Pi$ shorter than w. Then

$$\partial\langle w \rangle = \begin{cases} x\partial\langle v \rangle + x & \text{if } w = xv \\ x^{-1}\partial\langle v \rangle - x^{-1}x & \text{if } w = x^{-1}v \\ 0 & \text{if } w = 1 \end{cases}$$

Now define $s : \mathrm{Diff}(\Pi) \longrightarrow C_0(\Pi)$ to be the unique π-linear map such that $s(dx) = \langle x\rangle$ for $x \in X$. Since $\mathrm{Diff}(\Pi)$ is freely generated by all dx for $x \in X$, this does define a unique homomorphism. For $x \in X$, we have that $\partial \circ s(dx) = \partial\langle x\rangle = dx$ and so $\partial \circ s = \mathrm{id}$.

For each $n \geq 0$ we define a homomorphism $s : C_n \longrightarrow C_{n+1}$ as follows. We know that C_n is the free π-module generated by Π^{n+1}. We will denote a generator by $\langle w_0, \cdots, w_n\rangle$ where $w_0, \ldots, w_n$ are words in elements of X and their inverses. Then we define $s : C_n \longrightarrow C_{n+1}$ by induction on the length of the first word:

$$s\langle w_0, \ldots, w_n\rangle = \begin{cases} xs\langle w, w_1, \ldots, w_n\rangle - \langle x, w, w_1, \ldots\rangle & \text{if } w_0 = xw \\ x^{-1}s\langle w, w_1, \ldots, w_n\rangle + x^{-1}\langle x, w_0, w_1, \ldots\rangle & \text{if } w_0 = x^{-1}w \\ \langle 1, 1, w_1, \ldots, w_n\rangle & \text{if } w_0 = 1 \end{cases}$$

Proposition 10.4 *For any word w and any $x \in X$*

$$s\langle xw, w_1, \ldots, w_n\rangle = xs\langle w, w_1, \ldots, w_n\rangle - \langle x, w, w_1, \ldots, w_n\rangle$$
$$s\langle x^{-1}w, w_1, \ldots, w_n\rangle = x^{-1}s\langle w, w_1, \ldots, w_n\rangle + \langle x^{-1}, w_0, w_1, \ldots, w_n\rangle$$

Proof These are just the recursive definitions unless w begins with x^{-1} for the first equation or with x for the second. Suppose $w = x^{-1}v$. Then from the definition of s,

$$s\langle w, w_1 \ldots, w_n\rangle = x^{-1}s\langle v, w_1, \ldots, w_n\rangle + x^{-1}\langle x, w, w_1, \ldots, w_n\rangle$$

so that

$$s\langle xw, w_1, \ldots, w_n\rangle = s\langle v, w_1, \ldots, w_n\rangle$$
$$= xs\langle w, w_1, \ldots, w_n\rangle - \langle x, w, w_1, \ldots, w_n\rangle$$

The second one is proved similarly. □

Now we can prove that s is a contraction. First we will do this in dimension 0, then, by way of example, in dimension 2; nothing significant changes in any higher dimension. In dimension 0, suppose w is a word and we suppose that for any shorter word v, we have that $s \circ \partial\langle v\rangle + \partial \circ s\langle v\rangle = \langle v\rangle$. If $x = 1$, then

$$s \circ \partial\langle 1\rangle + \partial \circ s\langle 1\rangle = \partial\langle 1, 1\rangle = 1\langle 1\rangle - \langle 1\rangle + \langle 1\rangle = \langle 1\rangle$$

If $w = xv$, with $x \in X$, then

$$\partial \circ s\langle w\rangle + s \circ \partial\langle w\rangle = \partial(xs\langle v\rangle - \partial\langle x, v\rangle) + s(dw)$$
$$= x\partial \circ s\langle v\rangle - x\langle v\rangle + \langle xv\rangle - \langle x\rangle + s(x\partial(v) + dx)$$
$$= \langle w\rangle + x(\partial \circ s + s \circ \partial - 1)\langle v\rangle - \langle x\rangle + \langle x\rangle = \langle w\rangle$$

A similar argument takes care of the case that $w = x^{-1}v$. In dimension 2, the chain group $C_2(\Pi)$ is freely generated by Π^3. If we denote a generator by $\langle w_0, w_1, w_2\rangle$, we argue by induction on the length of w_0. If $w_0 = 1$, then

$$s \circ \partial\langle 1, w_1, w_2\rangle = s(\langle w_1, w_2\rangle - \langle w_1, w_2\rangle + \langle 1, w_1w_2\rangle - \langle 1, w_1\rangle)$$
$$= \langle 1, 1, w_1w_2\rangle - \langle 1, 1, w_1\rangle$$

while

$$\partial \circ s\langle 1, w_1, w_2\rangle = \partial(\langle 1, 1, w_1, w_2\rangle)$$
$$= \langle 1, w_1, w_2\rangle - \langle 1, w_1, w_2\rangle + \langle 1, w_1, w_2\rangle - \langle 1, 1, w_1w_2\rangle + \langle 1, 1, w_1\rangle$$

and these add up to $\langle 1, w_1, w_2\rangle$. Assume that $(\partial \circ s + s \circ \partial)\langle w\rangle = \langle w\rangle$ when w is shorter than w_0. Then for $w_0 = xw$,

$$\begin{aligned}\partial \circ s\langle xw, w_1, w_2\rangle &= x\partial \circ s\langle w, w_1, w_2\rangle - \partial\langle x, w, w_1, w_2\rangle \\ &= x\partial \circ s\langle w, w_1, w_2\rangle - x\langle w, w_1, w_2\rangle + \langle xw, w_1, w_2\rangle \\ &\quad - \langle x, ww_1, w_2\rangle + \langle x, w, w_1w_2\rangle - \langle x, w, w_1\rangle\end{aligned}$$

while

$$\begin{aligned}s \circ \partial\langle x, w, w_1, w_2\rangle &= xws\langle w_1, w_2\rangle - s\langle xww_1, w_2\rangle + s\langle xw, w_1w_2\rangle - s\langle xw, w_1\rangle \\ &= xws\langle w_1, w_2\rangle - xs\langle ww_1, w_2\rangle + \langle x, ww_1, w_2\rangle \\ &\quad + xs\langle w, w_1w_2\rangle - \langle x, w, w_1w_2\rangle - xs\langle w, w_1\rangle + \langle x, w, w_1\rangle \\ &= xs \circ \partial\langle w, w_1, w_2\rangle + \langle x, ww_1, w_2\rangle - \langle x, w, w_1w_2\rangle + \langle x, w, w_1\rangle\end{aligned}$$

Then,

$$\begin{aligned}(\partial \circ s + s \circ \partial)\langle xw, w_1, w_2\rangle &= x(\partial \circ s + s \circ \partial)\langle w, w_1, w_2\rangle - x\langle w, w_1, w_2\rangle \\ &\quad + \langle xw, w_1, w_2\rangle - \langle x, ww_1, w_2\rangle + \langle x, w, w_1w_2\rangle \\ &\quad - \langle x, w, w_1\rangle + \langle x, ww_1, w_2\rangle - \langle x, w, w_1w_2\rangle + \langle x, w, w_1\rangle\end{aligned}$$

Using the inductive assumption, the first two terms cancel and all the rest cancel in pairs, except for $\langle xw, w_1, w_2\rangle$, which shows that $s \circ \partial + \partial \circ s = 1$ in this case. The second case, that w_0 begins with the inverse of a letter is similar.

This completes the proof of the homotopy equivalence for group cohomology. For associative algebras the argument is quite similar. For Lie algebras and, especially for commutative algebras, it is a good deal more complicated because it is less obvious that the cochain complex of a free algebras splits. In fact, for Harrison's cochain complex, it splits only in characteristic 0 and the equivalence fails in finite characteristic. As a result, the cohomology that results from the cotriple resolution has been seen as primary.

11 Acyclic models now

Besides the acyclic models theorem quoted above, there was a weaker form due to Michel André [André, 1967, 1974] in which the conclusion was the weaker homology isomorphism and one could not infer naturality, at least as it was stated and proved. In the process of trying to settle the naturality question I discovered an acyclic models theorem that included both the version above and André's as special cases, along with at least one other interesting version. I outline the definitions and theorems. For proofs, I refer to my recent book [Barr, 2002].

In this definition, $\mathscr{C} = \mathsf{CC}(\mathscr{A})$ is the category of chain complexes of an abelian category $\mathscr{A}$.

11.1 Acyclic classes. A class Γ of objects of $\mathscr{C}$ will be called an *acyclic class* provided:

AC–1. The 0 complex is in Γ.
AC–2. The complex $C_\bullet$ belongs to Γ if and only if $SC_\bullet$ does.
AC–3. If the complexes $K_\bullet$ and $L_\bullet$ are homotopic and $K_\bullet \in \Gamma$, then $L_\bullet \in \Gamma$.

AC–4. Every complex in Γ is acyclic.

AC–5. If $K_{\bullet\bullet}$ is a double complex, all of whose rows are in Γ, then the total complex of $C_\bullet$ belongs to Γ.

Given an acyclic class Γ, let Σ denote the class of arrows f whose mapping cone is in Γ. It can be shown that this class lies between the class of homotopy equivalences and that of homology equivalences.

Suppose that $G : \mathscr{X} \longrightarrow \mathscr{X}$ is an endofunctor and that $\epsilon : G \longrightarrow \mathrm{Id}$ is a natural transformation. If $F : \mathscr{X} \longrightarrow \mathscr{A}$ is a functor, we define an augmented chain complex functor we will denote $FG^{\bullet+1} \longrightarrow F$ as the functor that has FG^{n+1} in degree n, for $n \geq -1$. Let $\partial^i = FG^i\epsilon G^{n-i} : FG^{n+1} \longrightarrow FG^n$. Then the boundary operator is $\partial = \sum_{i=0}^n (-1)^i \partial^i$. If, as usually happens in practice, G and ϵ are 2/3 of a cotriple, then this chain complex is the chain complex associated to a simplicial set built using the comultiplication δ to define the degeneracies. Next suppose that $K_\bullet \longrightarrow K_{-1}$ is an augmented chain complex functor. Then there is a double chain complex functor that has in bidegree (n, m) the term K_nG^{m+1}. This will actually commute since

$$\begin{tikzcd}
K_nG^{m+1} \arrow[r, "dG^{m+1}"] \arrow[d, "K_nG^i\epsilon G^{m-i}"'] & K_{n-1}G^{m+1} \arrow[d, "K_{n-1}G^i\epsilon G^{m-i}"] \\
K_nG^m \arrow[r, "dG^m"'] & K_{n-1}G^m
\end{tikzcd}$$

commutes by naturality for $0 \leq i \leq m$ hence so does

$$\begin{tikzcd}
K_nG^{m+1} \arrow[r, "dG^{m+1}"] \arrow[d, "\sum_{i=0}^m (-1)^i K_nG^i\epsilon G^{m-i}"'] & K_{n-1}G^{m+1} \arrow[d, "\sum_{i=0}^m (-1)^i K_{n-1}G^i\epsilon G^{m-i}"] \\
K_nG^m \arrow[r, "dG^m"'] & K_{n-1}G^m
\end{tikzcd}$$

However, the usual trick of negating every second column produces an anticommuting double complex.

This is augmented in both directions, once, using ϵ, over the single complex $K_\bullet$ and second, using the augmentation of $K_\bullet$, over the complex $K_{-1}G^{\bullet+1}$. We say that $K_\bullet$ is *ϵ-presentable with respect to* Γ if for each $n \geq 0$, the augmented chain complex $K_nG^{\bullet+1} \longrightarrow K_n \longrightarrow 0$ belongs to Γ. We say that $K_\bullet$ is *G-acyclic with respect to* Γ if the augmented complex $K_\bullet G \longrightarrow K_{-1}G \longrightarrow 0$ belongs to Γ.

It could be notes that an augmented chain complex $C_\bullet \longrightarrow C_{-1} \longrightarrow 0$ is the desuspension of the mapping cone of the map of chain complexes $C_\bullet \longrightarrow C_{-1}$ in which the latter is considered as a chain complex with C_{-1} in degree 0 and 0 in all other degrees. Thus an equivalent formulation of the two definitions above is that $K_\bullet$ is *ϵ-presentable with respect to* Γ if for each $n \geq 0$, the chain map $K_nG^{\bullet+1} \longrightarrow K_n$ belongs to Σ and that $K_\bullet$ is *G-acyclic with respect to* Γ if the chain map $K_\bullet G \longrightarrow K_{-1}G \longrightarrow 0$ belongs to Σ.

Theorem 11.1 *Let Γ be an acyclic class and Σ be the associated class of arrows. Suppose $\alpha : K_\bullet \longrightarrow K_{-1}$ and $\beta : L_\bullet \longrightarrow L_{-1}$ are augmented chain complex functors. Suppose G is an endofunctor on $\mathscr{X}$ and $\epsilon : G \longrightarrow \mathrm{Id}$ a natural transformation for which $K_\bullet$ is ϵ-presentable and $L_\bullet \longrightarrow L_{-1} \longrightarrow 0$ is G-acyclic,* *both with respect to* Γ. *Then given any natural transformation* $f_{-1} : K_{-1} \longrightarrow L_{-1}$ *there is, in* $\Sigma^{-1}\mathscr{C}$, *a unique arrow* $f_\bullet : K_\bullet \longrightarrow L_\bullet$ *that extends* f_{-1}.

Although we give no proof, here is the "magic diagram" from which it all follows:

$$\begin{tikzcd}
K_{-1}G^{\bullet+1} \arrow[d, "f_{-1}G^{\bullet+1}"'] & K_\bullet G^{\bullet+1} \arrow[l, "\alpha G^{\bullet+1}"'] \arrow[r, "K_\bullet \epsilon"] & K_\bullet \\
L_{-1}G^{\bullet+1} & L_\bullet G^{\bullet+1} \arrow[l, "\beta G^{\bullet+1}"] \arrow[r, "L_\bullet \epsilon"'] & L_\bullet
\end{tikzcd}$$

Note that each of the two hypotheses implies that one of the "wrong way" arrows belongs to Γ. When these are inverted, you get a map $K_\bullet \longrightarrow L_\bullet$.

11.2 Examples. We mention three examples of acyclic classes.

Let Γ be the class of contractible complexes. In that case, Σ is the class of homotopy equivalences. It is trivial to show that when the rows and columns of a double complex are contractible so is the total complex. It is a little more surprising—but still true—that if the rows *or* the columns are contractible, so is the total complex. The only other point to make is that in this case, the functor $\mathscr{C} \longrightarrow \Sigma^{-1}\mathscr{C}$ is surjective on arrows, so that you actually get a homotopy equivalence as conclusion.

Let Γ be the class of acyclic complexes. In that case, Σ is the class of chain maps that induce an isomorphism on homology (called homology isomorphisms). One can use a trivial spectral sequence argument to show that if all the rows or all the columns of a double complex are acyclic, then so is the total complex, but it is not hard to give a direct argument using a filtration. Since the functor $\mathscr{C} \longrightarrow \Sigma^{-1}\mathscr{C}$ is not surjective on arrows, the arrows you get are not induced by arrows between the chain complexes, but they are natural.

Let Γ be the class of chain complexes that are, at each object of $\mathscr{X}$, contractible, but not naturally so. In that case, Σ consists of arrows that induce, at each X, homotopy equivalences. This means that the arrows are natural, but the homotopy inverse and the homotopies involved are not necessarily natural. This situation arises quite naturally in topology. For example, on the category of C^∞ manifolds, the inclusion of the C^∞ chains into all the chains can be shown to induce a homotopy equivalence on any space, but there is no obviously natural way of doing this. This example answered a question raised by Rob Milson who was working with these homology groups.

References

[1] M. André (1967), *Méthode simpliciale en algèbre homologique et algèbre commutative.* Lecture Notes in Mathematics **43**, Springer-Verlag, Berlin, Heidelberg, New York.

[2] M. André (1974), *Homologie des algèbres commutatives.* Springer-Verlag, Berlin, Heidelberg, New York.

[3] H. Appelgate (1965), *Categories with models.* Dissertation, Columbia University.

[4] R. Baer (1934), *Erweiterungen von Gruppen und ihren Isomorphismem.* Math. Zeit. **38**, 375–416.

[5] M. Barr (1962), *On the cohomology of commutative algebras.* Dissertation, University of Pennsylvania.

[6] M. Barr (1965a), *A cohomology theory for commutative algebras, I.* Proc. Amer. Math. Soc. **16**, 1379–1384.

[7] M. Barr (1965b), *A cohomology theory for commutative algebras, II.* Proc. Amer. Math. Soc. **16**, 1385–1391.

[8] M. Barr (1966), *Cohomology in tensored categories. Proceedings of the Conference on Categorical Algebra*, Springer-Verlag, Berlin, Heidelberg, New York, 344–354.

[9] M. Barr (1967), *Shukla cohomology and triples.* J. Algebra **5**, 222–231.

[10] M. Barr (1968a), *A note on commutative algebra cohomology.* Bull. Amer. Math. Soc. **74**, 310–313.

[11] M. Barr (1968b), *Harrison homology, Hochschild homology and triples.* J. Algebra, **8**, 314–323.

[12] M. Barr (1969), Cohomology and obstructions: commutative algebras. In B. Eckmann, ed. *Seminar on Triples and Categorical Homology Theory*, Lecture Notes in Mathematics **80**, 357–375, Springer-Verlag, Berlin, Heidelberg, New York.

[13] M. Barr (1996a), *Cartan-Eilenberg cohomology and triples.* J. Pure Applied Algebra, **112**, 219–238.

[14] M. Barr (1996b), *Acyclic models.* Canadian J. Math., **48**, 258-273.

[15] M. Barr (2002), *Acyclic Models.* Centre de Recherche Mathématiques.

[16] M. Barr (forthcoming), *Resolutions and derived functors.* ftp.math.mcgill.ca/pub/barr/derfun.{ps,pdf}.

[17] M. Barr and J. Beck (1966), *Acyclic models and triples.* Proceedings of the Conference on Categorical Algebra, Springer, 336–343.

[18] M. Barr and J. Beck (1969), Homology and standard constructions. In M. Tierney and A. Heller, eds., *Seminar on Triples and Categorical Homology Theory*, Lecture Notes in Mathematics **80**, Springer-Verlag, Berlin, Heidelberg, New York, 245–535.

[19] M. Barr and M.-A. Knus (1971), *Extensions of Derivations.* Proc. Amer. Math. Soc. **28**, 313–314.

[20] M. Barr, and G. S. Rinehart (1966), *Cohomology as the derived functor of derivations.* Trans. Amer. Math. Soc., **122**, 416–426.

[21] D. Buchsbaum (1956), *Exact categories.* Appendix to [Cartan and Eilenberg, 1956].

[22] H. Cartan and S. Eilenberg (1956), *Homological Algebra.* Princeton University Press.

[23] C. Chevalley and S. Eilenberg (1948), *Cohomology theory of Lie groups and Lie algebras.* Trans. Amer. Math. Soc., **63**, 85–124.

[24] S. Eilenberg and S. Mac Lane (1941), *Infinite cycles and homologies.* Proc. Nat. Acad. Sci. U. S. A., **27**, 535–539.

[25] S. Eilenberg and S. Mac Lane (1942a), *Natural isomorphisms in group theory.* Proc. Nat. Acad. Sci. U. S. A., **28**, 537–543.

[26] S. Eilenberg and S. Mac Lane (1942b), *Group extensions and homology.* Ann. of Math., **43**, 757–831.

[27] S. Eilenberg and S. Mac Lane (1943), *Relations between homology and homotopy groups.* Proc. Nat. Acad. Sci. U. S. A., **29**, 155–158.

[28] S. Eilenberg and S. Mac Lane (1945a), *General theory of natural equivalences.* Trans. Amer. Math. Soc., **58**, 231–294.

[29] S. Eilenberg and S. Mac Lane (1945b), *Relations between homology and homotopy groups of spaces.* Ann. of Math., **46**, 480–509.

[30] S. Eilenberg and S. Mac Lane (1947a), *Cohomology theory in abstract groups, I.* Ann. of Math., **48**, 51–78.

[31] S. Eilenberg and S. Mac Lane (1947b), *Cohomology theory in abstract groups, II. Group extensions with a non-Abelian kernel.* Ann. of Math., **48**, 326–341.

[32] S. Eilenberg and S. Mac Lane (1951), *Homology theories for multiplicative systems.* Trans. Amer. Math. Soc., **71**, 294–330.

[33] S. Eilenberg and S. Mac Lane (1953), *Acyclic models.* Amer. J. Math., **75**, 189–199.

[34] S. Eilenberg and J. C. Moore (1965a), *Foundations of Relative Homological Algebra*, Memoirs Amer. Math. Soc. **55**.

[35] S. Eilenberg and J. C. Moore (1965b), *Adjoint functors and triples* Illinois J. Math., **9**, 381–398.

[36] Murray Gerstenhaber and S. D. Schack (1987), *A Hodge-type decomposition for commutative algebra cohomology.* J. Pure Appl. Algebra **48**, 229–247.

[37] R. Godement (1958), *Topologie algébrique et théorie des faisceaux.* Actualités Sci. Ind. No. 1252. Publ. Math. Univ. Strasbourg, **13**, Hermann, Paris.

[38] A. Grothendieck (1957), *Sur quelques points d'algèbre homologique.* Tohôku Math. Journal **2**, 119–221.

[39] M. Hall (1938), *Group rings and extensions, I.* Ann. of Math. **39**, 220–234.

[40] D. K. Harrison (1962), *Commutative algebras and cohomology.* Trans. Amer. Math. Soc., **104**, 191–204.

[41] G. Hochschild (1945), *On the cohomology groups of an associative algebra.* Ann. of Math. (2), **46**, 58–67.

[42] G. Hochschild (1946), *On the cohomology theory for associative algebras.* Ann. of Math. (2), **47**, 568–579.

[43] H. Hopf (1942, 1943), *Fundamentalgruppe und zweite Bettische Gruppen* Comment. Math. Helv. **14**, 257–309. *Nachtrag,* **15**, 27–32.

[44] H. Hopf (1945), *Ueber die Bettischen Gruppen die zu einem beliebigen Gruppe gehören.* Comment. Math. Helv. **17**, 39–79.

[45] P. Huber (1961), *Homotopy theory in general categories.* Math. Ann., **144**, 361–385.

[46] H. Kleisli (1964), *Comparaison de la résolution simpliciale à la bar-résolution.* In Catégories Non-Abéliennes, L'Université de Montréal, 85–99.

[47] H. Kleisli (1965), *Every standard construction is induced by a pair of adjoint functors.* Proc. Amer. Math. Soc., **16**, 544–546.

[48] S. Mac Lane (1950), *Duality for groups.* Bull. Amer. Math. Soc. **56**, 485–516.

[49] S. MacLane (1958), *Extensions and obstructions for rings.* Ill. J. Math. **2**, 316–345.

[50] S. MacLane (1963), *Homology.* Springer-Verlag, Berlin, Heidelberg, New York.

[51] U. Shukla (1961), *Cohomologie des algèbres associatives.* Ann. Sci. École Norm. Sup. (3), **78**, 163–209.

[52] A. M. Turing (1938), *The extensions of a group.* Compositio Math. **5**, 357–367.

[53] H. Zassenhaus (1937), *Lehrbuch der Gruppentheorie.* Hamburg Math. Einzelschriften **21**, Leipzig.

Received October 14, 2002; in revised form May 14, 2003

Fields Institute Communications
Volume **43**, 2004

A Survey of Semi-abelian Categories

Francis Borceux
Département de Mathématiques
Université de Louvain
2 chemin du Cyclotron
1348 Louvain-la-Neuve, Belgium
`borceux@math.ucl.ac.be`

Abstract. We define the notion of semi-abelian category and discuss its basic properties. Such a category is in particular a Mal'cev category. The five lemma, the nine lemma and the snake lemma hold true. Theories of semi-direct products, commutators and centrality can be developed. Every semi-abelian category contains an abelian core.

Introduction

An elementary introduction to the theory of abelian categories culminates generally with the proof of the basic diagram lemmas of homological algebra: the five lemma, the nine lemma, the snake lemma, and so on. This gives evidence of the power of the theory, but leaves the reader with the misleading impression that abelian categories constitute the most natural and general context where these results hold. This is indeed misleading, since all those lemmas are valid as well – for example – in the category of all groups, which is highly non-abelian.

This survey paper intends to give evidence, among other things, that a natural and more general context in which the diagram lemmas are valid is that of a semi-abelian category. More precisely, a finitely cocomplete exact category (for example, an algebraic variety) turns out to be semi-abelian precisely when it admits a zero object and when all the diagram lemmas of homological algebra hold true. The categories of groups, rings without unit, sheaves or presheaves of these, and so on, are semi-abelian. And a category $\mathcal{E}$ is abelian precisely when both $\mathcal{E}$ and its dual $\mathcal{E}^{\mathrm{op}}$ are semi-abelian.

We pay a special attention to the normal subobjects (=the kernels) which suffice, in a semi-abelian category, to characterize all the quotients by an equivalence relation. We prove in particular that a semi-abelian category is a Mal'cev category,

2000 *Mathematics Subject Classification.* Primary: 18E10, 18C10, 18G50; Secondary: 08B05, 08C05.

This survey paper has been written upon invitation of the editors of this volume; the author aknowledges support of FNRS grant 1.5.096.02.

that is, a category in which every reflexive relation is at once an equivalence relation. We show also the existence of semi-direct products and commutators in a semi-abelian category, extending so classical results in the case of groups. Finally, we prove that every semi-abelian category contains an abelian full subcategory of "abelian objects". Many of these results do not require the full strength of the axiomatics of semi-abelian categories: but it is a major interest of semi-abelian categories to allow recapturing all these important results from a unique and simple list of axioms.

Of course since more than 50 years, various mathematicians have tried to settle down an axiomatic context forcing the validity of the diagram lemmas of non-abelian homology. If none of these attempts imposed itself as a natural categorical context in which to work, it is probably because of their heavy technical nature, in contrast with the elegance of the theory of abelian categories. We refer the reader to the introduction of [23] for a reliable historical discussion of these pioneering works. It is in this same paper of G. Janelidze, L. Márki and W. Tholen that the definition of a semi-abelian category appears for the first time and is put in relation with the more general notion of protomodular category, due to D. Bourn (see [8]). This is the presentation adopted in this survey paper. Many other results of the paper are borrowed from [23] and subsequent work of D. Bourn, whose ideas are everywhere present in this paper. The presentation is often borrowed from [5].

1 Protomodular and semi-abelian categories

In the category Set of sets, a morphism $p\colon A \longrightarrow I$ can be seen as an I-indexed family of sets $(A_i)_{i\in I}$, simply by putting $A_i = p^{-1}(i)$. When p admits a section s, the pair (p,s) can now be seen as an I-indexed family of pointed sets $(A_i, a_i)_{i\in I}$, by putting further $a_i = s(i) \in A_i$.

More generally, given a category $\mathcal{E}$ and an object $I \in \mathcal{E}$, the category $\mathsf{Pt}_I(\mathcal{E})$ of pointed objects over I has for objects the triples

$$\big(A, p\colon A \to I, s\colon I \to A\big),\quad p \circ s = \mathrm{id}_I.$$

A morphism

$$f\colon \big(p, s\colon A \leftrightarrows I\big) \longrightarrow \big(p', s'\colon A' \leftrightarrows I\big)$$

is a morphism $f\colon A \longrightarrow A'$ making both triangles commutative, that is, $p' \circ f = p$ and $f \circ s = s'$.

When the category $\mathcal{E}$ admits pullbacks, every morphism $v\colon J \longrightarrow I$ in $\mathcal{E}$ induces by pullback a functor

$$v^*\colon \mathsf{Pt}_I(\mathcal{E}) \longrightarrow \mathsf{Pt}_J(\mathcal{E}).$$

The reader familiar with the theory of fibrations will notice that we have defined a pseudo-functor to the "category" of categories ("pseudo" meaning that $(v \circ w)^*$ is isomorphic to $w^* \circ v^*$, not equal)

$$\mathcal{E} \longrightarrow \mathsf{Cat},\ I \mapsto \mathsf{Pt}_I(\mathcal{E}),\ v \mapsto v^*,$$

thus a corresponding fibration $\mathsf{Pt}(\mathcal{E}) \longrightarrow \mathcal{E}$ of pointed objects in $\mathcal{E}$ (see [3]). This fibration has been extensively studied by D. Bourn (see [6] and [7]) and has very strong classifying properties. Its most powerful application is certainly the theory of protomodular categories (see [6], [7], [8], [10]), which allows in particular an elegant treatment of normal subobjects in categories without (necessarily) a zero object.

Definition 1.1 A category $\mathcal{E}$ is protomodular when

1. $\mathcal{E}$ admits pullbacks;
2. for every morphism $v\colon J \longrightarrow I$ in $\mathcal{E}$, the pullback functor
$$v^*\colon \mathsf{Pt}_I(\mathcal{E}) \longrightarrow \mathsf{Pt}_J(\mathcal{E})$$
reflects isomorphisms.

A zero object $\mathbf{0}$ in a category $\mathcal{E}$ is an object which is both initial and terminal. Given objects A, B in $\mathcal{E}$, we write $0\colon A \longrightarrow B$ for the unique morphism factoring through $\mathbf{0}$. The kernel $\mathsf{Ker}\, f$ of an arrow $f\colon A \longrightarrow B$ is the equalizer of f and 0, and dually for the cokernel. Equivalently, the kernel k of f is given by the following pullback:

$$\begin{tikzcd}
\mathsf{Ker}\, f \arrow[r, tail, "k"] \arrow[d] & A \arrow[d, "f"] \\
\mathbf{0} \arrow[r, tail] & B
\end{tikzcd}$$

We recall that a category $\mathcal{E}$ with pullbacks and a terminal object is finitely complete (see [4]).

Proposition 1.2 *Let $\mathcal{E}$ be a category with pullbacks and a zero object. The following conditions are equivalent:*

1. *$\mathcal{E}$ is protomodular;*
2. *the split short five lemma holds in $\mathcal{E}$, that is: given a diagram where all squares are commutative*
$$\begin{tikzcd}
\mathbf{0} \arrow[r] & A \arrow[r, tail, "k"] \arrow[d, "f"', "\cong"] & B \arrow[r, shift right, "p"'] \arrow[d, "g"'] & C \arrow[l, shift right, "s"'] \arrow[r] \arrow[d, "h", "\cong"'] & \mathbf{0} \\
\mathbf{0} \arrow[r] & A' \arrow[r, tail, "k'"] & B' \arrow[r, shift right, "p'"'] & C' \arrow[l, shift right, "s'"'] \arrow[r] & \mathbf{0}
\end{tikzcd}$$
$$k' \circ f = g \circ k,\ p' \circ g = h \circ p,\ g \circ s = s' \circ h,$$
and where moreover
$$p \circ s = \mathsf{id}_C,\ p' \circ s' = \mathsf{id}_{C'},\ k = \mathsf{Ker}\, p,\ k' = \mathsf{Ker}\, p',$$
if f and h are isomorphisms, then g is an isomorphism as well.

Proof The category $\mathsf{Pt}_{\mathbf{0}}(\mathcal{E})$ is isomorphic to the category $\mathcal{E}$. Pulling back an arrow along the morphism $0_I\colon \mathbf{0} \longrightarrow I$ is taking its kernel. Thus the split short five lemma means that the functor

$$0_I^*\colon \mathsf{Pt}_I(\mathcal{E}) \longrightarrow \mathsf{Pt}_{\mathbf{0}}(\mathcal{E}) \cong \mathcal{E}$$

reflects isomorphisms. This proves already (1) $\Rightarrow$ (2). Conversely, given an arbitrary morphism $v\colon J \longrightarrow I$, the equality $v \circ 0_J = 0_I$ implies $0_J^* \circ v^* = 0_I^*$. Since 0_J^* preserves isomorphisms and 0_I^* reflects them, v^* reflects isomorphisms. $\square$

Anticipating on 4.3, it should be observed that in opposition to the usual formulation of the short five lemma which involves normal epimorphisms p, p' and their kernels k, k', the split short five lemma is a statement involving only finite limits.

A category $\mathcal{E}$ is regular when it admits pullbacks, coequalizers of kernel pairs (= the two projections of the pullback of an arrow with itself), and when every pullback of a regular epimorphism (= a coequalizer) is still a regular epimorphism. Obviously, every regular epimorphism is then the coequalizer of its kernel pair. The category $\mathcal{E}$ is exact when it is regular and every equivalence relation in $\mathcal{E}$ is a kernel pair (see [1]).

The following definition is borrowed from G. Janelidze, L. Márki and W. Tholen (see [23]):

Definition 1.3 A category $\mathcal{E}$ is semi-abelian when

1. $\mathcal{E}$ has a zero object $\mathbf{0}$;
2. $\mathcal{E}$ has binary coproducts $A \amalg B$;
3. $\mathcal{E}$ is exact;
4. $\mathcal{E}$ is protomodular.

A semi-abelian category is thus in particular finitely complete. As the terminology suggests, the first example must be (see also 7.6 for an even more convincing result):

Example 1.4 Every abelian category is semi-abelian.

Proof An abelian category is additive, exact, finitely complete and finitely cocomplete and satisfies the short five lemma (see [19]). One concludes by 1.2. □

To produce a whole bunch of other examples, let us characterize the semi-abelian varieties of universal algebra (or equivalently, the semi-abelian Lawvere-algebraic categories). This characterization is due to D. Bourn and G. Janelidze (see [15]).

Theorem 1.5 *Let $\mathcal{E}$ be an algebraic variety. The following conditions are equivalent:*

1. *$\mathcal{E}$ is semi-abelian;*
2. *$\mathcal{E}$ is protomodular with a zero object;*
3. *the corresponding theory contains:*
   - *a unique constant* 0*;*
   - *n binary operations $\alpha_i(x,y)$ such that $\alpha_i(x,x) = 0$;*
   - *an $(n+1)$-ary operation θ such that $\theta\big(\alpha_1(x,y),\ldots,\alpha_n(x,y),y\big) = x$.*

Proof An algebraic variety is always complete, cocomplete and exact, proving (1) $\Leftrightarrow$ (2). Moreover, the existence of a unique constant in the theory is equivalent to the existence of a zero object in $\mathcal{E}$ (see [26]).

Let us first assume that $\mathcal{E}$ is semi-abelian. Writing F for the free algebra functor, consider the following diagram

$$\begin{tikzcd}
K \arrow[r, equal] \arrow[d, tail, "j'"'] & K \arrow[d, tail] \arrow[r, shift left] & \mathbf{0} \arrow[l, shift left] \arrow[d] \\
K' \arrow[r, tail, "k'"] \arrow[d, tail, "j"'] & K \vee F(y) \arrow[r, shift right, "p'"'] \arrow[d, tail, "i"] & F(y) \arrow[l, shift right, "s'"'] \arrow[d, equal] \\
K \arrow[r, tail, "k"] & F(x,y) \arrow[r, shift right, "p"'] & F(y) \arrow[l, shift right, "s"']
\end{tikzcd}$$

where $p(x) = y = p(y)$, $s(y) = y$ and $k = \mathsf{Ker}\, p$. Consider further the union $K \vee F(y)$ in the variety; p' is the restriction of p which is thus zero on K, while s factors through $F(y)$, yielding the section s'. Put $k' = \mathsf{Ker}\, p'$ and consider the corresponding factorization j. Complete this commutative diagram with the first line, where obviously $\mathsf{id}_K = \mathsf{Ker}\, 0$; this yields the factorization j'. Since k is a monomorphism, $j \circ j' = \mathsf{id}_K$ and the monomorphism j is an isomorphism. By the split short five lemma (see 1.2), the inclusion i is an isomorphism as well.

Since $F(x,y) = K \vee F(y)$, we get $x \in K \vee F(y)$. This proves the existence of n elements $\alpha_i(x,y) \in K$ and a $(n+1)$-ary operation θ such that

$$x = \theta\big(\alpha_1(x,y), \dots, \alpha_n(x,y), y\big).$$

It remains to notice that by definition of K, $\alpha_i(x,x) = 0$ for each index i.

Conversely, let us prove first that given two elements a, b in an algebra A

$$\big(\forall i\ \alpha_i(a,b) = 0\big) \Rightarrow \big(a = b\big).$$

Indeed

$$b = \theta\big(\alpha_1(b,b), \dots, \alpha_n(b,b), b\big) = \theta(0, \dots, 0, b) = \theta\big(\alpha_1(a,b), \dots, \alpha_n(a,b), b\big) = a.$$

It remains to prove the split short five lemma; we use the notation of 1.2. For the injectivity of g, consider $x, y \in B$ such that $g(x) = g(y)$. This yields

$$(h \circ p)\big(\alpha_i(x,y)\big) = (p' \circ g)\big(\alpha_i(x,y)\big) = p'\Big(\alpha_i\big(g(x), g(y)\big)\Big) = p'(0) = 0.$$

Since h is injective, $p\big(\alpha_i(x,y)\big) = 0$ and viewing k as a canonical inclusion, this yields $\alpha_i(x,y) \in A$. But then,

$$(k' \circ f)\big(\alpha_i(x,y)\big) = (g \circ k)\big(\alpha_i(x,y)\big) = \alpha_i\big(g(x), g(y)\big) = 0.$$

Since k' and f are injective, $\alpha_i(x,y) = 0$. This implies $x = y$, thus the injectivity of g.

For the surjectivity, choose now $x \in B'$ and put $y = (s \circ h^{-1} \circ p')(x) \in B$. This yields

$$p'\Big(\alpha_i\big(x, g(y)\big)\Big) = \alpha_i\big(p'(x), (p' \circ g)(y)\big) = \alpha_i\big(p'(x), p'(x)\big) = 0.$$

This proves the existence of elements $z_i \in A'$ such that $k'(z_i) = \alpha_i\bigl(x, g(y)\bigr)$. Viewing still k as a canonical inclusion, this forces

$$\begin{aligned} g\Bigl(\theta\bigl(f^{-1}(z_1), \dots, f^{-1}(z_n), y\bigr)\Bigr) &= \theta\bigl(k'(z_1), \dots, k'(z_n), g(y)\bigr) \\ &= \theta\Bigl(\alpha_1\bigl(x, g(y)\bigr), \dots, \alpha_n\bigl(x, g(y)\bigr), g(y)\Bigr) \\ &= x \end{aligned}$$

and proves the surjectivity of g. □

It should be noticed that condition 3 of theorem 1.5 appeared already in [33] and later in the appendix of [34]. To my best knowledge, it has been used only to investigate properties of congruences.

The following corollary emphasizes the case $n = 1$ in the characterization given by theorem 1.5. This choice $n = 1$ cannot always be done, as proved in [25].

Corollary 1.6 *Let $\mathbb{T}$ be an algebraic theory which possesses a unique constant 0 and two binary operations $x + y$ and $x - y$ which satisfy the axioms*

$$(x - y) + y = x, \quad x - x = 0.$$

The corresponding variety $\mathcal{V}$ is semi-abelian. In particular, when the theory $\mathbb{T}$ contains a group operation, the variety $\mathcal{V}$ is semi-abelian: this contains the cases of groups, Ω-groups, rings without unit, R-algebras, and so on.

Proof Put $n = 1$, $\alpha_1(x, y) = x - y$ and $\theta(x, y) = x + y$ in 1.5. □

Example 1.7 Let $\mathbb{T}$ be an algebraic theory as in theorem 1.5. The models of $\mathbb{T}$ in every Grothendieck topos constitute a semi-abelian category. When moreover the theory $\mathbb{T}$ admits a finite presentation, its models in every topos with Natural Number Object still constitute a semi-abelian category.

Proof When $\mathbb{T}$ admits a finite presentation, its models in an elementary topos $\mathcal{E}$ with Natural Number Object constitute a finitely complete and finitely cocomplete exact category, monadic over $\mathcal{E}$ (see [24], section D.5.3). This takes care of axioms 1, 2, 3 in definition 1.3. This allows also repeating the proof of the split short five lemma, as in 1.5, in the internal logic of the topos $\mathcal{E}$. The assumption of finite presentability can be dropped as soon as the topos is finitely complete and cocomplete. □

An abelian category admits a full and faithful exact embedding in a category of modules, yielding a highly useful corresponding metatheorem which allows developing most proofs in terms of elements. Up to now, no such theorem has been proved for semi-abelian categories. Nevertheless, Barr's metatheorem for regular categories (see [1]) can already be adapted to the present context.

Metatheorem 1.8 *Let $\mathcal{P}$ be a statement of the form $\varphi \Rightarrow \psi$, where φ and ψ can be expressed as conjunctions of properties in the following list:*

1. *some arrow is a zero arrow;*
2. *some finite diagram is commutative;*
3. *some morphism is a monomorphism;*
4. *some morphism is a regular epimorphism;*
5. *some morphism is an isomorphism;*
6. *some finite diagram is a limit diagram;*

7. *some arrow $f\colon A \longrightarrow B$ factors through some monomorphism $s\colon S \rightarrowtail B$.*

If this statement $\mathcal{P}$ is valid in the category Set_* *of pointed sets, it is valid in every regular category $\mathcal{E}$ with a zero object, thus in particular in every semi-abelian category.*

Proof If $\mathcal{E}$ is a finitely complete regular category, Barr's theorem (see [1]) indicates the existence of a full and faithful embedding

$$Z\colon \mathcal{E} \longrightarrow [\mathcal{C}^{\mathrm{op}}, \mathsf{Set}]$$

in a category of presheaves, where Z preserves and reflects finite limits and regular epimorphisms.

When $\mathcal{E}$ has a zero object, $Z(\mathbf{0})$ is the constant functor on the singleton. The natural transformation $Z(\mathbf{0}) \Rightarrow Z(A)$ defines a base point in each set $Z(A)(X)$, for all $X \in \mathcal{C}$. Thus each functor $Z(A)$ factors naturally through the category Set_* of pointed sets. This proves that the Barr embedding factors as

$$Z'\colon \mathcal{E} \longrightarrow [\mathcal{C}^{\mathrm{op}}, \mathsf{Set}_*].$$

This functor Z' is faithful since so is Z. It is also full since every natural transformation over Set_* is in particular a natural transformation over Set. Since the forgetful functor $\mathsf{Set}_* \longrightarrow \mathsf{Set}$ preserves and reflects finite limits and regular epimorphisms, Z' preserves and reflects finite limits and regular epimorphisms, since so does Z. And by construction, Z' preserves and reflects zero morphisms.

Since Z is full and faithful, it preserves and reflects isomorphisms and the commutativity of diagrams. A morphism f is a monomorphism when the kernel pair (u, v) of f is a pair of isomorphisms. Condition 7 means that the pullback of s along f is an isomorphism. So if one of the statements 1 to 7 is valid in the category of pointed sets, it is valid pointwise in $[\mathcal{C}^{\mathrm{op}}, \mathsf{Set}_*]$ and thus holds true in $\mathcal{E}$. □

We shall implicitly use this metatheorem in several proofs, when a property falls under its scope and is trivial in the case of pointed sets via an elementary "chase on the diagram".

2 Subobjects, quotients and pullbacks

First, we show that in a semi-abelian category, kernels and cokernels behave like in abelian categories.

Proposition 2.1 *In a semi-abelian category $\mathcal{E}$, pulling back along an arrow reflects monomorphisms.*

Proof Consider the following commutative diagram

$$\begin{tikzcd}
K[f'] \arrow[r, shift right, "p_1'"'] \arrow[d, "\gamma"'] & A' \arrow[l, shift right, "\Delta'"'] \arrow[r, tail, "f'"] \arrow[d, "\beta"] & X' \arrow[d, "\alpha"] \\
K[f] \arrow[r, shift right, "p_1"'] & A \arrow[l, shift right, "\Delta"'] \arrow[r, "f"'] & X
\end{tikzcd}$$

where $K[f]$ is the kernel pair of f with first projection p_1 and diagonal Δ; analogously for f'. When the right hand square is a pullback, so is the left hand square involving the projections. When f' is a monomorphism, $p_1' = \mathsf{id}_{A'} = \Delta'$, yielding the isomorphism

$$\beta^*(p_1, \Delta \colon K[f] \leftrightarrows A) \cong (\mathsf{id}_{A'}, \mathsf{id}_{A'} \colon A' \leftrightarrows A') \cong \beta^*(\mathsf{id}_A, \mathsf{id}_A \colon A \leftrightarrows A)$$

between pointed objects. By protomodularity, (p_1, Δ) is isomorphic to $(\mathsf{id}_A, \mathsf{id}_A)$ and p_1 is an isomorphism. Thus f is a monomorphism. $\square$

Theorem 2.2 *In a semi-abelian category $\mathcal{E}$, the following conditions are equivalent for a morphism $f \colon A \longrightarrow B$:*

1. *f is a monomorphism:*
2. $\mathsf{Ker}\, f = 0$.

Proof Computing the kernel is pulling back along $0 \colon \mathbf{0} \longrightarrow B$; one concludes by 2.1. $\square$

The dual result does not hold in a semi-abelian category. Of course an epimorphism has zero cokernel, but the converse is not true, not even in the category of groups. If G is a non trivial simple group with non trivial subgroup H, the cokernel of the inclusion $i \colon H \rightarrowtail G$ must be zero while i is not surjective.

We recall that in a category with finite limits, a family $(f_i \colon X_i \longrightarrow Y)_{i \in I}$ of morphisms is strongly epimorphic (and in particular epimorphic) when it does not factor through any proper subobject of Y. In a finitely complete regular category, strong epimorphisms coincide with regular ones (see [4]).

Lemma 2.3 *In a semi-abelian category $\mathcal{E}$, given a pullback diagram $p \circ u = v \circ q$ and a section $p \circ s = \mathsf{id}_Y$,*

$$\begin{tikzcd}
V \arrow[r, "u"] \arrow[d, "q"'] & X \arrow[d, shift right, "p"'] \\
W \arrow[r, "v"'] & Y \arrow[u, shift right, "s"']
\end{tikzcd}$$

the pair (u, s) is strongly epimorphic.

Proof Consider the diagram

$$\begin{tikzcd}
V' \arrow[rrr, "w"] \arrow[dr, "y"] \arrow[ddr, shift right, "q'"'] & & & X' \arrow[dl, "x"'] \arrow[ddl, shift left, "p'"] \\
& V \arrow[r, "u"] \arrow[urr, "z"] \arrow[d, shift right, "q"'] & X \arrow[d, shift right, "p"'] & \\
& W \arrow[r, "v"'] \arrow[u, shift right, "t"'] \arrow[uul, shift right, "t'"'] & Y \arrow[u, shift right, "s"'] \arrow[uur, shift left, "s'"] &
\end{tikzcd}$$

where t is the factorization of the pair $(\mathsf{id}_W, s \circ v)$ through the pullback. If u and s factor through some monomorphism x as $u = x \circ z$ and $s = x \circ s'$, we put $p' = p \circ x$ and view x as a morphism in $\mathsf{Pt}_Y(\mathcal{E})$. Pulling back this situation along v yields a corresponding morphism

$$y\colon \left(q', t'\colon V' \leftrightarrows W\right) \longrightarrow \left(q, t\colon V \leftrightarrows W\right)$$

in $\mathsf{Pt}_W(\mathcal{E})$. Since the downward directed square and the downward directed outer trapezium are pullbacks, the upper trapezium $u \circ y = x \circ w$ is a pullback as well, that is, $V' = u^{-1}(X')$. But since u factors through the subobject X', $u^{-1}(X') = V$ and y is an isomorphism. By protomodularity, x is an isomorphism as well. □

Corollary 2.4 *Every semi-abelian category $\mathcal{E}$ is unital, meaning that given two objects A and B, the pair*

$$A \xrightarrow{(\mathsf{id}_A, 0)} A \times B \xleftarrow{(0, \mathsf{id}_B)} B$$

is strongly epimorphic.

Proof In 2.3, choose $V = A$, $Y = B$, $X = A \times B$, $W = \mathbf{0}$, $u = (\mathsf{id}_A, 0)$, $p = p_B$, $s = (0, \mathsf{id}_B)$. □

Theorem 2.5 *Let $\mathcal{E}$ be a semi-abelian category $\mathcal{E}$. Every regular epimorphism $f\colon A \twoheadrightarrow B$ is the cokernel of its kernel: $f = \mathsf{Coker}\,(\mathsf{Ker}\, f)$.*

Proof Consider the following diagram, with f a regular epimorphism and k, (d_1, d_2) respectively, the kernel and the kernel pair of f.

$$\begin{tikzcd}
\mathsf{Ker}\, f \times \mathsf{Ker}\, f \arrow[r, shift left, "p_1"] \arrow[r, shift right, "p_2"'] \arrow[d, "\gamma"'] & \mathsf{Ker}\, f \arrow[r, two heads] \arrow[d, tail, "k"'] & \mathbf{0} \arrow[d, tail, "0_Y"] \\
K[f] \arrow[r, shift left, "d_1"] \arrow[r, shift right, "d_2"'] & X \arrow[r, two heads, "f"] & Y
\end{tikzcd}$$

Both left hand squares are pullbacks. Since each morphism d_i admits the diagonal $\Delta\colon X \longrightarrow K[f]$ as a section, lemma 2.3 implies that the pair (γ, Δ) is strongly epimorphic.

Choose now $h\colon X \longrightarrow Z$ such that $h \circ k = 0$; we must prove that h factors uniquely through f. Since $f = \mathsf{Coeq}(d_1, d_2)$, it suffices to prove that $h \circ d_1 = h \circ d_2$. We have indeed

$$h \circ d_1 \circ \gamma = h \circ k \circ p_1 = 0 = h \circ k \circ p_2 = h \circ d_2 \circ \gamma;$$

$$h \circ d_1 \circ \Delta = h = h \circ d_2 \circ \Delta.$$

This forces the conclusion, since the pair (γ, Δ) is epimorphic. □

Next we focus on an unusual cancellation property of pullbacks.

Lemma 2.6 *In a semi-abelian category $\mathcal{E}$, consider a diagram*

$$\begin{tikzcd}
V \arrow[r, "u"] \arrow[d, "q"'] \arrow[dr, phantom, "(1)"] & X \arrow[r, "h"] \arrow[d, shift right, "p"'] \arrow[dr, phantom, "(2)"] & A \arrow[d, "g"] \\
W \arrow[r, "v"'] & Y \arrow[r, "f"'] \arrow[u, shift right, "s"'] & B
\end{tikzcd}$$

where $p \circ s = \mathsf{id}_Y$ *and the downward directed squares are commutative. If the square* (1) *and the outer rectangle* (1)+(2) *are pullbacks, the square* (2) *is a pullback as well.*

Proof Since the square (1) is a pullback, there is a morphism $t\colon W \longrightarrow V$ such that $q \circ t = \mathsf{id}_W$ and $u \circ t = s \circ v$. Consider the diagram

$$\begin{tikzcd}
X \arrow[dr, dotted, "z"] \arrow[drr, "h"] \arrow[ddr, shift right, "p"'] & & \\
& P \arrow[r, "y"] \arrow[d, shift right, "x"'] & A \arrow[d, "g"] \\
& Y \arrow[r, "f"'] \arrow[u, shift right, "z \circ s"'] \arrow[uul, shift right, "s"'] & B
\end{tikzcd}$$

where the downward directed square is a pullback and z is the unique factorization of the outer downward directed quadrilateral through this pullback. This yields a morphism

$$z\colon \big(p, s\colon X \leftrightarrows Y\big) \longrightarrow \big(x, z \circ s\colon P \leftrightarrows Y\big)$$

in the category $\mathsf{Pt}_Y(\mathcal{E})$. The pullback assumptions in the statement imply

$$v^*\big(p, s\colon X \leftrightarrows Y\big) \cong \big(q, t\colon V \leftrightarrows W\big) \cong v^*\big(x, z \circ s\colon P \leftrightarrows Y\big).$$

Thus $v^*(z)$ is an isomorphism and therefore z is an isomorphism, by protomodularity. By construction of P, this proves that the square (2) is a pullback. □

Theorem 2.7 *In a semi-abelian category* $\mathcal{E}$, *consider a commutative diagram,*

$$\begin{tikzcd}
A \arrow[r, "u"] \arrow[d, "f_1"'] \arrow[dr, phantom, "(1)"] & B \arrow[r, "v"] \arrow[d, two heads, "f_2"'] \arrow[dr, phantom, "(2)"] & C \arrow[d, "f_3"] \\
X \arrow[r, "s"'] & Y \arrow[r, "t"'] & Z
\end{tikzcd}$$

where f_2 *is a regular epimorphism. If the outer rectangle* (1)+(2) *and the square* (1) *are both pullbacks, the square* (2) *is a pullback as well.*

Proof Let us extend the diagram of the statement with the kernel pairs (α_i, β_i) of the morphisms f_i, the corresponding diagonals Δ_i and the obvious factorizations u', v' through these.

$$\begin{tikzcd}
K[f_1] \arrow[r, "u'"] \arrow[d, shift right=3, "\alpha_1"'] \arrow[d, shift left=3, "\beta_1"] \arrow[dr, phantom, "(3)"] & K[f_2] \arrow[r, "v'"] \arrow[d, shift right=3, "\alpha_2"'] \arrow[d, shift left=3, "\beta_2"] \arrow[dr, phantom, "(4)"] & K[f_3] \arrow[d, shift right=3, "\alpha_3"'] \arrow[d, shift left=3, "\beta_3"] \\
A \arrow[r, "u"] \arrow[u, "\Delta_1" description] \arrow[d, "f_1"'] \arrow[dr, phantom, "(1)"] & B \arrow[r, "v"] \arrow[u, "\Delta_2" description] \arrow[d, two heads, "f_2"] \arrow[dr, phantom, "(2)"] & C \arrow[u, "\Delta_3" description] \arrow[d, "f_3"] \\
X \arrow[r, "s"'] & Y \arrow[r, "t"'] & Z
\end{tikzcd}$$

Using our metatheorem 1.8, it is easy to observe that since (1) and (1)+(2) are pullbacks, both downward directed squares (3) are pullbacks and both downward directed rectangles (3)+(4) are pullbacks. The morphisms α_2 and β_2 admit the section Δ_2. By 2.6, both downward directed squares (4) are pullbacks as well. By 1.8 again, this implies at once that the square (2) is a pullback. □

3 Normal subobjects and equivalence relations

As the example of groups suggests, we define:

Definition 3.1 In a semi-abelian category $\mathcal{E}$, a subobject $k\colon K \rightarrowtail A$ is normal when it is the kernel of some morphism.

Proposition 3.2 *Let $\mathcal{E}$ be a semi-abelian category $\mathcal{E}$. Every normal subobject $k\colon K \rightarrowtail A$ is the kernel of its cokernel: $k = \mathsf{Ker}\,(\mathsf{Coker}\,k)$.*

Proof Given $f\colon A \longrightarrow B$ and $k = \mathsf{Ker}\,f$, consider the image factorization $f = i \circ p$ of f. We have still $k = \mathsf{Ker}\,p$, with p a regular epimorphism. By 2.5, $p = \mathsf{Coker}\,k$. □

Proposition 3.3 *In a semi-abelian category $\mathcal{E}$, the pullback of a normal subobject along an arbitrary arrow is again a normal subobject.*

Proof Consider $f\colon A \longrightarrow B$ and a subobject $s\colon S \rightarrowtail B$ which is the kernel of $g\colon B \longrightarrow C$. Then $f^{-1}(S)$ is the kernel of $g \circ f$. □

The crucial property of normal subobjects is then:

Theorem 3.4 *In a semi-abelian category $\mathcal{E}$, the normal subobjects of a fixed object A correspond bijectively with the equivalence relations on A.*

Proof Given an equivalence relation R on A, we associate with it the kernel $\mathsf{Ker}\,q_R$ of the corresponding quotient $q_R\colon A \twoheadrightarrow A/R$.

This correspondence is surjective: every normal subobject $i\colon B \rightarrowtail A$ is the kernel of its cokernel q (see 3.2) and q itself is the quotient $q = q_R$, where R is the kernel pair of q.

By exactness of $\mathcal{E}$, every equivalence relation R on an object A is the kernel pair of its quotient $q_R\colon A \twoheadrightarrow A/R$. Pulling back the first (or the second) projection $p_1^R\colon R \longrightarrow A$ over $\mathbf{0}$ yields the kernel of p_1^R, which coincides thus with the kernel of q_R.

$$\begin{tikzcd}
K \arrow[r, tail, "k"] \arrow[d] \arrow[dr, phantom, "\text{p.b.}"] & R \arrow[r, "p_2^R"] \arrow[d, "p_1^R"'] \arrow[dr, phantom, "\text{p.b.}"] & A \arrow[d, two heads, "q_R"] \\
\mathbf{0} \arrow[r] & A \arrow[r, two heads, "q_R"'] & A/R
\end{tikzcd}$$

To prove the injectivity, consider two equivalence relations R and S on A such that the corresponding quotients $q_R\colon A \twoheadrightarrow A/R$ and $q_S\colon A \twoheadrightarrow A/S$ admit the same kernel $k\colon K \rightarrowtail A$. It follows at once that the intersection $R \cap S$ is an equivalence relation whose quotient $q_{R\cap S}\colon A \twoheadrightarrow A/(R\cap S)$ still admits k as kernel. The split short five lemma applied to the diagram

$$\begin{tikzcd}
K \arrow[r, tail, "k"] \arrow[d, equal] & R\cap S \arrow[r, shift right, "p_1^{R\cap S}"'] \arrow[d, tail] & A \arrow[l, shift right, "\Delta_{R\cap S}"'] \arrow[d, equal] \\
K \arrow[r, tail, "k"'] & R \arrow[r, shift right, "p_1^R"'] & A \arrow[l, shift right, "\Delta_R"']
\end{tikzcd}$$

where $\Delta_{R\cap S}$, Δ_R denote the diagonals, proves that $R \cap S \cong R$, that is , $R \subseteq S$. Analogously, $S \subseteq R$. $\square$

The reader is probably familiar with the property in theorem 3.4, but he should convince himself that the property fails for non semi-abelian categories, like the category of pointed sets or that of monoids. For example in the case of pointed sets, the equivalence class of 0 determines ... the equivalence class of 0, nothing else, and certainly not the whole congruence. The next corollary, whose precise meaning is recalled in the proof, emphasizes further the "good behaviour" of equivalence relations in a semi-abelian category.

Corollary 3.5 *In a semi-abelian category $\mathcal{E}$, consider an equivalence relation R on an object A. If a subobject $i\colon B \rightarrowtail A$ contains the R-equivalence class of 0, it is saturated for R.*

Proof The equivalence class of 0 is given by the left hand pullback below.

$$\begin{tikzcd}
{[0]_R} \arrow[r, "\rho"] \arrow[d, tail, "r_0"'] & R \arrow[d, tail, "r"] \\
A \arrow[r, "{(0,\mathrm{id}_A)}"'] & A\times A
\end{tikzcd}
\qquad
\begin{tikzcd}
B_1 \arrow[r, tail, "i_1"] \arrow[d, "q_1"'] & R \arrow[d, "p_1^R"] \\
B \arrow[r, tail, "i"'] & A
\end{tikzcd}
\qquad
\begin{tikzcd}
B_2 \arrow[r, tail, "i_2"] \arrow[d, "q_2"'] & R \arrow[d, "p_2^R"] \\
B \arrow[r, tail, "i"'] & A
\end{tikzcd}$$

Notice that the monomorphism $(\mathrm{id}_A, 0)$ is the inverse image of $\mathbf{0} \rightarrowtail A$ along the second projection; by composition of pullbacks, $[0]_R$ is thus also the kernel of $p_2 \circ r = p_2^R$, that is the kernel of the quotient $A \twoheadrightarrow A/R$, as observed in the proof of 3.4. The saturation of B for R means that in the other two pullbacks, the monomorphisms i_1 and i_2 determine the same subobject of R. In the language of 1.8,

$$B_1 = \{(x,y) \in R \mid x \in B\}, \; B_2 = \{(x,y) \in R \mid y \in B\}.$$

The saturation of B for R means thus

$$\text{If } (x,y) \in R, \text{ then } x \in B \Leftrightarrow y \in B,$$

which is the usual notion of R-saturation.

To prove the corollary, consider the diagram

$$\begin{tikzcd}
{[0]_R \cap B} \arrow[r, tail] \arrow[d, equal] & B_1 \cap B_2 \arrow[r, shift left, "\pi"] \arrow[d, tail, "\beta_1"'] & B \arrow[l, shift left, "\delta"'] \arrow[d, equal] \\
{[0]_R} \arrow[r, tail] & B_1 \arrow[r, shift left, "\pi_1"'] & B \arrow[l, shift left, "\delta_1"']
\end{tikzcd}$$

where π, π_1 are the first projections and δ, δ_1 are the diagonals. An obvious diagram chase proves $\mathrm{Ker}\,\pi_1 = [0]_R$ and $\mathrm{Ker}\,\pi = [0]_R \cap B$. Since $[0]_R \subseteq B$ by assumption, $[0]_R \cap B = [0]_R$ and the split short five lemma implies that β_1 is an isomorphism. Thus $B_1 \subseteq B_2$ and analogously, $B_2 \subseteq B_1$. □

Now let us switch to another celebrated concept: the Mal'cev categories, generalizing Mal'cev varieties (see [30] and [32]). A Mal'cev operation is a ternary operation $p(x,y,z)$ which satisfies the axioms

$$p(x,y,y) = x, \; p(y,y,z) = z.$$

The most celebrated example is that of groups, where $p(x,y,z) = x - y + z$. A Mal'cev algebraic theory is one which contains a Mal'cev operation; the corresponding variety is then called a Mal'cev variety. This notion admits an elegant categorical formalization (see [17] and [18]):

Definition 3.6 A Mal'cev category $\mathcal{E}$ is a category with finite limits in which every reflexive relation is an equivalence relation.

D. Bourn has given a very elegant characterization of Mal'cev categories: they are those finitely complete categories $\mathcal{E}$ such that each category $\mathsf{Pt}_I(\mathcal{E})$ is unital in the sense of 2.4 (see [7]). Let us prove directly that:

Theorem 3.7 *Every semi-abelian category $\mathcal{E}$ is a Mal'cev category.*

Proof Let us use our metatheorem 1.8. Let $p_1, p_2 \colon R \rightrightarrows A$ be a reflexive relation on A in $\mathcal{E}$. We write $R \times_A R$ for the pullback of p_2 and p_1, which in Set_* is given by

$$R \times_A R = \{(x,y,z) | (x,y) \in R, \; (y,z) \in R\}.$$

The first and the third projections $R \times_A R \rightrightarrows A$ induce a factorization to $A \times A$ and we write T for the inverse image of R along this factorization. In Set_* this yields

$$T = \{(x,y,z) | (x,y) \in R, \; (y,z) \in R, \; (x,z) \in R\}$$

with an obvious factorization $i\colon T \rightarrowtail R \times_A R$. The kernel of $(p_1^T, p_3^T)\colon T \longrightarrow A \times A$ is given elementwise by

$$\{(0,y,0) \mid (0,y) \in R,\ (y,0) \in R,\ (0,0) \in R\}$$

while the kernel of $(p_1^{R\times_A R}, p_3^{R\times_A R})\colon R \times_A R \longrightarrow A \times A$ is

$$\{(0,y,0) \mid (0,y) \in R,\ (y,0) \in R\}.$$

Both kernels coincide since R is reflexive; let us write K for these kernels. This yields the diagram

$$\begin{tikzcd}
K \arrow[r, tail] \arrow[d, equal] & T \arrow[d, tail, "i"] \arrow[r, shift right, "{(p_1^T, p_3^T)}"'] & A \times A \arrow[l, shift right, "s^T"'] \arrow[d, equal] \\
K \arrow[r, tail] & R \times_A R \arrow[r, shift right, "{(p_1, p_3)}"'] & A \times A \arrow[l, shift right, "s"']
\end{tikzcd}$$

where elementwise, $s_T(x,z) = (x,x,z) = s(x,z)$; this makes sense again because R is reflexive. The split short five lemma applied to this diagram indicates that i is an isomorphism. This proves the transitivity of R.

We know thus already that every reflexive relation in $\mathcal{E}$ is transitive. Given a reflexive relation R on A, we construct now a reflexive relation S on R. The object $R \times R$ is provided with four projections to A; S is the inverse image of R along the morphism $(p_1^{R\times R}, p_4^{R\times R})\colon R \times R \longrightarrow A \times A$. Elementwise, this means

$$S = \Big\{ \big((a,b),(c,d)\big) \Big| (a,b) \in R,\ (c,d) \in R,\ (a,d) \in R \Big\}.$$

This relation S on R is trivially reflexive, thus it is transitive as well by the first part of the proof. Now given $(x,y) \in R$, we have both $\big((y,y),(x,y)\big) \in S$ and $\big((x,y),(x,x)\big) \in S$ because R is reflexive. By transitivity of S, $\big((y,y),(x,x)\big) \in S$, proving $(y,x) \in R$. $\square$

Proposition 3.8 *In a semi-abelian category $\mathcal{E}$, the normal subobjects of every object A constitute a modular lattice.*

Proof By 3.4, it is equivalent to work with equivalence relations. If R and S are equivalence relations on A, their composite $R \circ S$ contains both R and S, but also the diagonal of A. By 3.7, $R \circ S$ is an equivalence relation. If T is another equivalence relation containing R and S, then $R \circ S \subseteq T \circ T = T$. Thus $R \circ S = R \vee S$ in the poset of equivalence relations. Notice that this implies $R \circ S = S \circ R$, since $R \vee S = S \vee R$. On the other hand the intersection $R \cap S$ is an equivalence relation as well, the diagonal Δ_A is the smallest equivalence relation on A and $A \times A$ is the biggest one. This proves that the equivalence relations constitute a lattice with top and bottom element.

The modularity of this lattice means the restricted distributivity law

$$R \leq T \ \Rightarrow\ T \wedge (R \vee S) = (T \wedge R) \vee (T \wedge S) = R \vee (T \wedge S),$$

given three equivalence relations R, S, T on A. One inequality is obvious. To prove the converse inequality, we use our metatheorem 1.8. Take a pair (x,y) in

$T \wedge (R \vee S)$. Since $R \vee S = R \circ S = S \circ R$,

$$\exists u \in A\ (x,u) \in R\ (u,y) \in S,\ \exists v \in A\ (x,v) \in S\ (v,y) \in R.$$

Since $R \subseteq T$, $(x,y) \in T$ and $(x,u) \in R$ imply $(u,y) \in T$. But $(u,y) \in S$, thus $(u,y) \in T \wedge S$. On the other hand $(x,u) \in R$, thus $(x,y) \in R \circ (T \wedge S)$. □

Proposition 3.9 *Given an object A in a semi-abelian category $\mathcal{E}$, write* $\mathsf{Sub}(A)$ *for its poset of subobjects.*

1. *The intersection of two normal subobjects of A in* $\mathsf{Sub}(A)$ *is still a normal subobject of A.*
2. *The union of two normal subobjects of A exists in* $\mathsf{Sub}(A)$ *and is a normal subobject of A.*
3. *Given a epimorphism $f\colon A \twoheadrightarrow B$ and a normal subobject $U \subseteq A$, its image $f(U) \subseteq B$ is a normal subobject of $f(A)$.*

Proof Consider two morphisms $f\colon A \longrightarrow X$, $g\colon A \longrightarrow Y$ and the corresponding factorization $(f,g)\colon A \longrightarrow X \times Y$. It is immediate that $\mathsf{Ker}\, f \cap \mathsf{Ker}\, g = \mathsf{Ker}\,(f,g)$. This proves the first assertion.

To prove the second assertion, consider two equivalence relations R, S on A and the corresponding normal subobjects $u\colon U \rightarrowtail A$, $v\colon V \rightarrowtail A$ given by theorem 3.4. By 3.8, the union of U and V in the lattice of normal subobjects of A exists: it is the normal subobject $w\colon W \rightarrowtail A$ corresponding to the equivalence relation $R \circ S$. We shall prove that $W = U \cup V$ as ordinary subobjects of A. This means that given a subobject $z\colon Z \rightarrowtail A$ containing U and V, Z contains W as well. By 3.5 we know already that Z is saturated for both R and S. Using 1.8, $x \in W$ means $(0,x) \in R \circ S$, that is the existence of y such that $(0,y) \in R$ and $(y,x) \in S$. But $(0,y) \in R$ implies $y \in Z$, from which $(y,x) \in S$ implies $x \in Z$.

Consider now a regular epimorphism $f\colon A \twoheadrightarrow B$ and an equivalence relation R on A. The subobject $(f \times f)(R) \subseteq B \times B$ contains the diagonal of B, thus it is an equivalence relation on B by 3.7. For short, we write it simply $f(R)$. To prove assertion 3, it suffices by 3.1 to show that the image of the R-equivalence class of $0 \in A$ is the $f(R)$-equivalence class of 0 in B. We apply again our metatheorem 1.8. Trivially $(a,0) \in R$ implies $\big(f(a),0\big) \in f(R)$. Conversely given $(b,0) \in f(R)$, we have $(a,a') \in R$ such that $f(a) = b$ and $f(a') = 0$, that is, $(a',0) \in S$. This implies $(a,0) \in R \circ S = S \circ R$, thus the existence of $a'' \in A$ with $(a,a'') \in S$ and $(a'',0) \in R$. This yields finally

$$(b,0) = \big(f(a),0\big) = \big(f(a''),0\big) = (f \times f)(a'',0) \in f(R).$$

When f is an arbitrary morphism, it suffices to consider its image factorization $f = i \circ p$ and apply the argument above to the regular epimorphism g. □

Proposition 3.10 *A semi-abelian category $\mathcal{E}$ is finitely complete and finitely cocomplete.*

Proof We know already that $\mathcal{E}$ is finitely complete, since it has pullbacks and a terminal object. Since $\mathcal{E}$ has an initial object and binary coproducts, it has all finite coproducts. It remains to prove the existence of the coequalizer of a pair $f,g\colon A \rightrightarrows B$ of morphisms.

Consider the morphism $(f,g)\colon A \longrightarrow B \times B$ and its image factorization $(f,g) = r \circ q$.

$$\begin{tikzcd}
 & R \amalg B \arrow[dr, two heads, "t"] & & B \arrow[ll, "s_B"'] \arrow[dd, tail, "\Delta_B"] & \\
 & & \overline{R} \arrow[dr, tail, "\overline{r}"] & & \\
A \arrow[r, two heads, "q"] \arrow[rrr, bend right, "(f,g)"'] & R \arrow[uu, "s_R"] \arrow[rr, tail, "r"'] & & B \times B \arrow[r, shift left, "p_0"] \arrow[r, shift right, "p_1"'] & B
\end{tikzcd}$$

Since q is an epimorphism, this diagram shows at once that $\mathsf{Coeq}(f,g)$ exists if and only if $\mathsf{Coeq}(p_0 \circ r, p_1 \circ r)$ exists and these coequalizers are then equal. Thus it suffices to prove that the relation R on B has a coequalizer.

Consider now the coproduct of R and B, the factorization τ of the pair (r, Δ_B) through this coproduct and its image factorization $\tau = \overline{r} \circ t$. An arrow $x\colon B \longrightarrow X$ coequalizes $(p_0 \circ \overline{r}, p_1 \circ \overline{r})$ precisely when it coequalizes their composites with the epimorphism t. Since (s_R, s_B) are the injections of a coproduct, this is further equivalent to

$$x \circ p_0 \circ \overline{r} \circ t \circ s_R = x \circ p_1 \circ \overline{r} \circ t \circ s_R, \; x \circ p_0 \circ \overline{r} \circ t \circ s_B = x \circ p_1 \circ \overline{r} \circ t \circ s_B$$

that is

$$x \circ p_0 \circ r = x \circ p_1 \circ r, \; x \circ p_0 \circ \Delta_B = x \circ p_1 \circ \Delta_B.$$

But since $p_0 \circ \Delta_B = \mathsf{id}_B = p_1 \circ \Delta_B$, we conclude that x coequalizes $(p_0 \circ \overline{r}, p_1 \circ \overline{r})$ precisely when it coequalizes $(p_0 \circ r, p_1 \circ r)$. This proves that the relation R has a coequalizer if and only if the relation $\overline{R}$ has a coequalizer, in which case those coequalizers are equal. So it suffices to prove that the relation $\overline{R}$ has a coequalizer.

By construction, the relation $\overline{R}$ contains the diagonal of B, that is, is reflexive. Since $\mathcal{E}$ is a Mal'cev category (see 3.7), that reflexive relation $\overline{R}$ is an equivalence relation. Since the category $\mathcal{E}$ is exact, this equivalence relation is a kernel pair and therefore has a coequalizer. $\square$

4 Exact sequences and diagram lemmas

In a semi-abelian category, the notion of exact sequence is the classical one:

Definition 4.1 Let $\mathcal{E}$ be a semi-abelian category. A pair (f,g) of composable morphisms is an exact sequence when the mono-part i of the image factorization of f

$$\begin{tikzcd}
A \arrow[rr, "f"] \arrow[dr, two heads, "p"'] & & B \arrow[rr, "g"] & & C \\
 & I \arrow[ur, tail, "i"'] & & &
\end{tikzcd}$$

is also the kernel of g.

As usual, a long exact sequence is a sequence of composable morphisms such that each pair of consecutive morphisms is exact. And a short exact sequence is a long exact sequence of the form

$$0 \longrightarrow A \xrightarrow{f} B \xrightarrow{g} C \longrightarrow 0.$$

Notice nevertheless a major difference with the abelian case: not every morphism f can appear as first morphism in an exact sequence (f, g): indeed, the exactness property forces the image of f to be a normal subobject.

To avoid any ambiguity, it is thus useful to make explicit some special cases of interest:

Proposition 4.2 *Let $\mathcal{E}$ be a semi-abelian category.*

1. *a morphism $f\colon A \longrightarrow B$ is a monomorphism precisely when the sequence*
$$0 \longrightarrow A \xrightarrow{f} B$$
is exact;
2. *$k = \mathsf{Ker}\, f$ precisely when the sequence*
$$0 \longrightarrow K \xrightarrow{k} A \xrightarrow{f} B$$
is exact;
3. *a morphism $f\colon A \longrightarrow B$ is a regular epimorphism precisely when the sequence*
$$A \xrightarrow{f} B \longrightarrow 0$$
is exact;
4. *$q = \mathsf{Coker}\, f$ and the image of f is a normal subobject, precisely when the sequence*
$$A \xrightarrow{f} B \xrightarrow{q} Q \longrightarrow 0$$
is exact;
5. *a morphism $f\colon A \longrightarrow B$ is an isomorphism precisely when the sequence*
$$0 \longrightarrow A \xrightarrow{f} B \longrightarrow 0$$
is exact.

Proof Conditions 1 and 2 follow at once from 2.2, since the morphism $0 \longrightarrow A$ is its own image. Condition 3 holds because f is a regular epimorphism precisely when its image is B. The case of isomorphisms follows from assertions 1 and 3.

To prove condition 4, let us write $f = i \circ p$ for the image factorization of f. Since p is an epimorphism, $\mathsf{Coker}\, f = \mathsf{Coker}\,(i \circ p) = \mathsf{Coker}\, i$. The exactness of the given sequence means thus first that q is a regular epimorphism and second that $i = \mathsf{Im}\, f = \mathsf{Ker}\, q$. By 2.5, this implies $q = \mathsf{Coker}\, i = \mathsf{Coker}\, f$.

Conversely, let $q = \mathsf{Coker}\, f$ and $i = \mathsf{Ker}\, g$, for some morphism g. There is no restriction to suppose that g is a regular epimorphism (if not, factor it through its image). By 2.5 again, this implies

$$g = \mathsf{Coker}\, i = \mathsf{Coker}\, f = q.$$

Thus $i = \mathsf{Ker}\, g = \mathsf{Ker}\, q$ and the sequence (f, q) is exact. □

Let us now turn our attention to the classical diagram lemmas involving exact sequences (see [9]).

Theorem 4.3 (Short five lemma) *In a semi-abelian category $\mathcal{E}$, consider the following commutative diagram where the rows are exact sequences.*

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, "u"] \arrow[d, "a"'] & B \arrow[r, "p"] \arrow[d, "b"'] & C \arrow[r] \arrow[d, "c"] & 0 \\
0 \arrow[r] & A' \arrow[r, "u'"] & B' \arrow[r, "p'"] & C' \arrow[r] & 0
\end{tikzcd}$$

If a and c are isomorphisms (respectively, monomorphisms, regular epimorphisms), b is an isomorphism (respectively, monomorphism, regular epimorphism) as well.

Proof We handle first the case of isomorphisms. Consider the commutative diagram

$$\begin{tikzcd}
A' \arrow[r, "a^{-1}", "\cong"'] \arrow[d, two heads] \arrow[rrr, bend left, "u'"] \arrow[dr, phantom, "(1)"] & A \arrow[r, tail, "u"] \arrow[d, two heads] \arrow[dr, phantom, "(2)"] & B \arrow[r, "b"] \arrow[d, two heads, "p"] \arrow[dr, phantom, "(3)"] & B' \arrow[d, two heads, "p'"] \\
0 \arrow[r, equal] & 0 \arrow[r] & C \arrow[r, "\cong", "c"'] & C'
\end{tikzcd}$$

The square (2) is a pullback because $u = \operatorname{Ker} p$. The outer part of the diagram is a pullback because $u' = \operatorname{Ker} p'$. Since the left hand horizontal morphisms are isomorphisms, the rectangle (2)+(3) is a pullback. By theorem 2.7, the square (3) is a pullback, thus b is an isomorphism, since so is c.

Next let us consider the factorization $b = b_1 \circ b_2$ of b through its image. When a and c are monomorphisms, p factors through b_2 and we get the following diagram.

$$\begin{tikzcd}
0 \arrow[r] & A \arrow[r, "u"] \arrow[d, equal] & B \arrow[r, "p"] \arrow[d, two heads, "b_2"'] & C \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & A \arrow[r, "b_2 \circ u"] & B'' \arrow[r, "p''"] & C \arrow[r] & 0
\end{tikzcd}$$

An easy diagram chase, using our metatheorem 1.8, shows that the second line is still exact. By the first part of the proof, b_2 is an isomorphism. Thus $b \cong b_1$ is a monomorphism.

When a and c are regular epimorphisms, u' factors through b_1 and we obtain the diagram

$$\begin{tikzcd}
0 \arrow[r] & A' \arrow[r, "u''"] \arrow[d, equal] & B'' \arrow[r, "p' \circ b_1"] \arrow[d, tail, "b_1"'] & C' \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & A' \arrow[r, "u'"] & B' \arrow[r, "p''"] & C' \arrow[r] & 0
\end{tikzcd}$$

Again an easy diagram chase proves that the first line is still an exact sequence. The first part of the proof implies that b_1 is now an isomorphism. Thus $b \cong b_2$ is a regular epimorphism. □

Theorem 4.4 (Five lemma) *In a semi-abelian category $\mathcal{E}$, consider the commutative diagram*

$$\begin{tikzcd}
A \arrow[r, "k"] \arrow[d, "\alpha"', "\cong"] & B \arrow[r, "h"] \arrow[d, "\beta"', "\cong"] & C \arrow[r, "g"] \arrow[d, "\gamma"'] & D \arrow[r, "f"] \arrow[d, "\cong"', "\delta"] & E \arrow[d, "\cong"', "\varepsilon"] \\
A' \arrow[r, "k'"] & B' \arrow[r, "h'"] & C' \arrow[r, "g'"] & D' \arrow[r, "f'"] & E'
\end{tikzcd}$$

where the rows are exact sequences and the morphisms α, β, δ, ε are isomorphisms. Then γ is an isomorphism as well.

Proof Considering the left hand square, the images of k and k' are isomorphic, thus also the coequalizers of these images. But these coequalizers are the epimorphic parts of the image factorizations of h and h'. Working "dually" from the right, the images of g and g' are isomorphic as well. Factoring all the horizontal morphisms through their images, it remains to apply the short five lemma (see 4.3) to the central part of this new diagram, delimited by the image objects of h, h', g, g'. □

Theorem 4.5 (Nine lemma) *In a semi-abelian category $\mathcal{E}$, consider the commutative diagram 1 where the three columns and the middle row are short exact*

$$\begin{tikzcd}
 & \mathbf{0} \arrow[d] & \mathbf{0} \arrow[d] & \mathbf{0} \arrow[d] & \\
\mathbf{0} \arrow[r] & K'' \arrow[r, "k''"] \arrow[d, "u'"'] \arrow[dr, phantom, "(1)"] & X'' \arrow[r, "f''"] \arrow[d, "v'"'] & Y'' \arrow[r] \arrow[d, "w'"] & \mathbf{0} \\
\mathbf{0} \arrow[r] & K' \arrow[r, "k'"] \arrow[d, "u"'] & X' \arrow[r, "f'"] \arrow[d, "v"'] & Y' \arrow[r] \arrow[d, "w"] & \mathbf{0} \\
\mathbf{0} \arrow[r] & K \arrow[r, "k"] \arrow[d] & X \arrow[r, "f"] \arrow[d] & Y \arrow[r] \arrow[d] & \mathbf{0} \\
 & \mathbf{0} & \mathbf{0} & \mathbf{0} &
\end{tikzcd}$$

Diagram 1

sequences. The first row is a short exact sequence if and only if the last row is a short exact sequence.

Proof Notice at once that, unlike the case of abelian categories, no duality principle can be used to interchange the roles of the first and the last row. Indeed, the notion of semi-abelian category is not self-dual (see 7.6 for a convincing argument). We use freely 4.2 and our metatheorem 1.8, without recalling it every time.

Let us first assume that the last row is exact. The morphisms k' and u' are monomorphisms, thus k'' is a monomorphism and the first row is exact at K''.

A straightforward diagram chase on the first two columns indicates that the square (1) is a pullback. Another diagram chase on the first two lines concludes that $k'' = \mathsf{Ker}\, f''$. So the first row is exact at X''.

It remains to prove that f'' is a regular epimorphism. For this we consider the following diagram, where the square (2) is a pullback.

$$\begin{tikzcd}
K' \arrow[r, tail, "k'"] \arrow[d, two heads, "u"'] & X' \arrow[r, two heads, "f'"] \arrow[d, two heads, "v_1"'] & Y' \arrow[d, equal] \\
K \arrow[r, tail, "h"] \arrow[d, equal] & Z \arrow[r, two heads, "g"] \arrow[d, two heads, "v_2"'] \arrow[dr, phantom, "(2)"] & Y' \arrow[d, two heads, "w"] \\
K \arrow[r, tail, "k"'] & X \arrow[r, two heads, "f"'] & Y
\end{tikzcd}$$

The equality $f \circ v = w \circ f'$ yields the factorization v_1, with $v_2 \circ v_1 = v$ and $g \circ v_1 = f'$. The equality $f \circ k = 0 = w \circ 0$ yields h such that $g \circ h = 0$ and $v_2 \circ h = k$. The morphisms v_2 and g are regular epimorphisms, since so are w and f. The morphism h is a monomorphism, since so is k. An easy diagram chase, using $k = \mathsf{Ker}\, f$ and next the pullback (2), proves that $h = \mathsf{Ker}\, g$. Since g is a regular epimorphism, (h, g) is a short exact sequence. By the short five lemma (see 4.3), v_1 is a regular epimorphism.

Since $w \circ w' = 0 = f \circ 0$, we get a factorization $s \colon Y'' \longrightarrow Z$ such that $v_2 \circ s = 0$ and $g \circ s = w'$. Since w' is a monomorphism, so is s. A simple diagram chase proves $s = \mathsf{Ker}\, v_2$, because the square (2) is a pullback. We obtain in this way the following commutative diagram

$$\begin{tikzcd}
X'' \arrow[r, tail, "v'"] \arrow[d, "f''"'] & X' \arrow[r, two heads, "v"] \arrow[d, two heads, "v_1"'] & X \arrow[d, equal] \\
Y'' \arrow[r, tail, "s"'] & Z \arrow[r, two heads, "v_2"'] & X
\end{tikzcd}$$

where $v' = \mathsf{Ker}\, v$ and $s = \mathsf{Ker}\, v_2$. The left hand square is pullback. Indeed, if $s \circ x = v_1 \circ y$, then $v_2 \circ v_1 \circ y = 0$, from which the expected factorization through the kernel v' of $v = v_2 \circ v_1$. Since v_1 is a regular epimorphism, so is f''.

Next let us assume that the first row is an exact sequence. Since w and f' are regular epimorphisms, f is a regular epimorphism. Notice that we can "dualize" a little bit further the proof of the first part: the square $f \circ v = w \circ f'$ is a pushout and $f = \mathsf{Coker}\, k$. But this does not imply the exactness of the last row at X (see 4.2.4).

Instead, let us prove first that k is a monomorphism. Using 1.8, if $k(x) = 0$, choose $y \in K'$ such that $u(y) = x$. Then $(v \circ k')(y) = 0$, thus $k'(y) = v'(z)$ for some $z \in X''$. Since the square (1) is a pullback, we have $(y, z) \in K''$. Finally, $x = (u \circ u')(y, z) = 0$. By 2.2, k is a monomorphism.

Now in diagram 1, replace the morphism $k \colon K \longrightarrow X$ by the kernel of f, which we write $k_f \colon \mathsf{Ker}\, f \rightarrowtail X$. We get accordingly a factorization $u_f \colon K' \longrightarrow \mathsf{Ker}\, f$. The three rows (k'', f''), (k', f'), (k_f, f) and the last two columns (v', v), (w', w) of this new diagram are now exact. Thus the first column (u', u_f) is exact as well, by the first part of the proof.

Since (u', u) and (u', u_f) are short exact sequences, $\overline{k}$ is an isomorphism and $k \cong k_f = \mathsf{Ker}\, f$, which concludes the proof. □

Let us mention that the "middle nine lemma" holds as well: when the three columns, the first and the last row are exact sequences and when moreover $f' \circ k' = 0$, then the middle row is an exact sequence (see [5]).

Theorem 4.6 (Snake lemma) *In a semi-abelian category $\mathcal{E}$, consider diagram 2, where all squares of solid arrows are commutative and all sequences of solid arrows are exact. There exists an exact sequence of dotted arrows still making the diagram commutative.*

Proof Since the proof uses techniques analogous to those developed for proving the nine lemma, we simply give the construction of the various morphisms. An explicit proof can be found in [9], in the special case where f is a monomorphism and g' is a regular epimorphism. It is routine to adapt it to the general case, by factoring f and g' through their images, as we do it below.

Of course f_K and g_K are the obvious factorizations through the kernels k_u of u, k_v of v and k_w of w. And dually for f'_Q, g'_Q.

To define the diagonal morphism d, we consider first the factorizations $f = f_2 \circ f_1$ and $g' = g'_2 \circ g'_1$ of f and g' through their images. The pairs (f_2, g) and (f', g'_1) are then short exact sequences.

Next, we construct diagram 3, whose various ingredients will now be described.

The square (1) is a pullback by definition. Since k_w is a monomorphism, h is a monomorphism. Since g is a regular epimorphism, φ is a regular epimorphism.

From $g \circ k_v = k_w \circ g_K$, we obtain the factorization γ through the pullback, yielding $h \circ \gamma = k_v$ and $\varphi \circ \gamma = g_K$. Since k_v is a monomorphism, so is γ.

From the equality $g \circ f_2 = 0 = k_w \circ 0$ we obtain a factorization θ through the pullback such that $h \circ \theta = f_2$ and $\varphi \circ \theta = 0$. Since f_2 is a monomorphism, θ is a monomorphism as well. It follows at once that $\theta = \mathsf{Ker}\, \varphi$, thus (θ, φ) is a short exact sequence.

$$
\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & 0 \arrow[d] & \\
 & K_u \arrow[r, dotted, "f_K"] \arrow[d, "k_u"'] & K_v \arrow[r, dotted, "g_K"] \arrow[d, "k_v"'] & K_w \arrow[d, "k_w"] \arrow[dddll, dotted, "d"] & \\
 & X \arrow[r, "f"] \arrow[d, "u"'] & Y \arrow[r, "g"] \arrow[d, "v"'] & Z \arrow[r] \arrow[d, "w"] & 0 \\
0 \arrow[r] & X' \arrow[r, "f'"] \arrow[d, "q_u"'] & Y' \arrow[r, "g'"] \arrow[d, "q_v"'] & Z' \arrow[d, "q_w"] & \\
 & Q_u \arrow[r, dotted, "f'_Q"'] \arrow[d] & Q_v \arrow[r, dotted, "g'_Q"'] \arrow[d] & Q_w \arrow[d] & \\
 & 0 & 0 & 0 &
\end{tikzcd}
$$

Diagram 2

The commutative squares $f' \circ u = v \circ f$ and $g' \circ v = w \circ g$ yield corresponding factorizations u', w' through the images of f and g'. Since f_1 is a regular epimorphism and g_2' is a monomorphism, we have still $q_u = \operatorname{Coker} u'$ and $k_w = \operatorname{Ker} w'$.

Observe that $h = \operatorname{Ker}(w' \circ g) = \operatorname{Ker}(g_1' \circ v)$. Since $f' = \operatorname{Ker} g_1'$, this yields the factorization ψ and the "square" $f' \circ \psi = v \circ h$ is a pullback. It follows at once that $\gamma = \operatorname{Ker} \psi$. Finally the equalities $f' \circ \psi \circ \theta = v \circ h \circ \theta = v \circ f_2 = f' \circ u'$ imply $\psi \circ \theta = u'$, because f' is a monomorphism. Therefore $q_u \circ \psi \circ \theta = q_u \circ u' = 0$, which yields the factorization d through the cokernel of θ, such that $d \circ \varphi = q_u \circ \psi$. This morphism d is the connecting morphism that we wanted to construct. □

Various other diagram lemmas can be proved in a semi-abelian category. This is in particular the case for the two Noether isomorphism theorems. We refer the interested reader to [5] where detailed proofs of these results are given.

The observant reader will have noticed that the results of this section did not use the existence of binary coproducts nor the effectiveness of equivalence relations. They are thus valid in a regular protomodular category with a zero object. They are even valid in the slightly more general context of γ-categories (see [16]).

$$
\begin{tikzcd}
 & K_v \arrow[d, tail, "\gamma"'] \arrow[dr, "g_K"] \arrow[dd, bend left=60, "k_v\ (1)"] & \\
 & H \arrow[r, two heads, "\varphi"] \arrow[d, tail, "h"] \arrow[ddl, dotted, "\psi"] & K_w \arrow[d, tail, "k_w"] \arrow[dddll, dotted, "d"] \\
X'' \arrow[ur, tail, "\theta"] \arrow[r, tail, "f_2"] \arrow[d, "u'"'] & Y \arrow[r, two heads, "g"] \arrow[d, "v"] & Z \arrow[d, "w'"] \\
X' \arrow[r, tail, "f'"'] \arrow[d, "q_u"'] & Y' \arrow[r, two heads, "g'_1"'] & Z'' \\
Q_u & &
\end{tikzcd}
$$

Diagram 3

5 G-actions and semi-direct product

Another important property that semi-abelian categories share with the category of groups is the existence of semi-direct products. For the sake of clarity, we define first that notion and show afterwards that it reduces to the classical notion in the case of groups.

The crucial result is the following one; it is due to D. Bourn and G. Janelidze and uses the full strength of the axioms of a semi-abelian category (see [14]).

Theorem 5.1 *For every morphism $v: J \longrightarrow I$ of a semi-abelian category $\mathcal{E}$, the pullback functor*

$$v^*: \mathsf{Pt}_I(\mathcal{E}) \longrightarrow \mathsf{Pt}_J(\mathcal{E})$$

is monadic.

Proof We refer to [2], [28] or [4] for the theory of monads and in particular the Beck criterion for monadicity.

First of all, the functor v^* admits a left adjoint functor

$$v_!: \mathsf{Pt}_J(\mathcal{E}) \longrightarrow \mathsf{Pt}_I(\mathcal{E})$$

which is given by the pushout along v (see 3.10).

$$\begin{tikzcd} B \arrow[r, "u"] \arrow[d, shift right, "q"'] & A \arrow[d, shift right, "p"'] \\ J \arrow[r, "v"'] \arrow[u, shift right, "t"'] & I \arrow[u, shift right, "s"'] \end{tikzcd}$$

Indeed, if t admits a retraction q, its pushout s along v admits a retraction p such that $s \circ v = u \circ t$. The commutativity of this diagram implies the existence of a unique factorization

$$z\colon (q,t) \longrightarrow (h,r) = v^* v_!(q,t)$$

through the pullback (h,r) of (p,s) along v. This yields the first natural transformation of the expected adjunction. Analogously, using the same diagram but with different assumptions, let us start from $(p,s) \in \mathsf{Pt}_I(\mathcal{E})$. We can compute the pullback (q,t) of (p,s) over J; the commutativity of the diagram implies now the existence of a unique factorization

$$z'\colon (h',r') = v_! v^*(p,s) \longrightarrow (p,s),$$

from the pushout (h',r') of (q,s) along v. This yields the second natural transformation of the expected adjunction. It is routine to check the triangular identities.

Next, by protomodularity of $\mathcal{E}$, the functor v^* reflects isomorphisms (see 1.1). Thus to apply the Beck criterion, it remains to check the condition on coequalizers of reflexive pairs.

First, it is obvious to observe that pullbacks, pushouts, equalizers and coequalizers exist in $\mathsf{Pt}_I(\mathcal{E})$ and are computed as in $\mathcal{E}$. Since $\mathcal{E}$ is exact, it follows at once that each category $\mathsf{Pt}_I(\mathcal{E})$ is exact and each pullback functor v^* preserves pullbacks and regular epimorphisms. This implies that each functor v^* preserves coequalizers of kernel pairs, thus by exactness, coequalizers of equivalence relations.

Observe next that each category $\mathsf{Pt}_I(\mathcal{E})$ has binary coproducts, computed as pushouts in $\mathcal{E}$ (see 3.10). It has also a zero object, namely the pair $(\mathsf{id}_I, \mathsf{id}_I)$; each functor v^* preserves obviously these zero objects. Since pullbacks and the terminal object generate all finite limits, each category $\mathsf{Pt}_I(\mathcal{E})$ is finitely complete and each functor v^* preserves finite limits.

If $\big(p, s\colon A \leftrightarrows I\big)$ is an object in $\mathsf{Pt}_I(\mathcal{E})$, a pointed object over (p,s) in $\mathsf{Pt}_I(\mathcal{E})$ is simply a pointed object over A in $\mathcal{E}$. Since pullbacks and isomorphisms in $\mathsf{Pt}_I(\mathcal{E})$ are computed as in $\mathcal{E}$, $\mathsf{Pt}_I(\mathcal{E})$ is protomodular, since so is $\mathcal{E}$. This concludes the proof that $\mathsf{Pt}_I(\mathcal{E})$ is semi-abelian. In particular, $\mathsf{Pt}_I(\mathcal{E})$ is a Mal'cev category, by 3.7.

Consider now a reflexive pair in $\mathsf{Pt}_I(\mathcal{E})$, that is

$$\big(q, t\colon B \leftrightarrows I\big) \begin{array}{c} \xrightarrow{\;u\;} \\ \xleftarrow{\;w\;} \\ \xrightarrow[\;v\;]{} \end{array} \big(p, s\colon A \leftrightarrows I\big), \quad u \circ w = \mathsf{id}_A = v \circ w.$$

Consider the morphism (u,v) to the product $(p,s) \times (p,s)$ in $\mathsf{Pt}_I(\mathcal{E})$ and its image factorization $(u,v) = r \circ \pi$ still in $\mathsf{Pt}_I(\mathcal{E})$.

$$\begin{tikzcd}
(p,s) \arrow[r, "w"] \arrow[rrr, bend right, "\Delta"'] & (q,t) \arrow[r, two heads, "\pi"] \arrow[rr, bend left, "{(u,v)}"] & R \arrow[r, tail, "r"] & (p,s)\times(p,s) \arrow[r, shift left, "p_0"] \arrow[r, shift right, "p_1"'] & (p,s)
\end{tikzcd}$$

By assumption, the composite $(u,v)\circ w$ is the diagonal Δ. Thus R is a reflexive relation and therefore an equivalence relation, by the Mal'cev property. Since π is an epimorphism, $\mathsf{Coeq}(u,v) = \mathsf{Coeq}(p_0\circ r, p_1\circ r)$ and the coequalizer of (u,v) is that of the equivalence relation R. We know already that the coequalizer of R is preserved by the functor v^*. But since v^* preserves all the ingredients of the diagram, the coequalizer of $v^*(R)$ is also that of $\big(v^*(u), v^*(v)\big)$. □

Corollary 5.2 *For each object I of a semi-abelian category $\mathcal{E}$, the functor*

$$\mathsf{Ker} : \mathsf{Pt}_I(\mathcal{E}) \longrightarrow \mathcal{E},\ \big(p,s\colon A \leftrightarrows I\big) \mapsto \mathsf{Ker}\, p$$

is monadic and admits the functor

$$\sigma_I\colon \mathcal{E} \longrightarrow \mathsf{Pt}_I(\mathcal{E}),\ B \mapsto \big((0,\mathsf{id}_I), s_I\colon B \amalg I \leftrightarrows I\big)$$

as left adjoint.

Proof In 5.1, put $J = \mathbf{0}$. One has $\mathsf{Pt}_{\mathbf{0}}(\mathcal{E}) \cong \mathcal{E}$, while the pushout over $\mathbf{0}$ is the coproduct. □

Here is now the definition of the semi-direct product:

Definition 5.3 Let $\mathcal{E}$ be a semi-abelian category and $G \in \mathcal{E}$ an object of $\mathcal{E}$.

1. A G-algebra is an algebra for the monad $\mathbb{T}_G$ corresponding to the monadic functor $\mathsf{Ker} : \mathsf{Pt}_G(\mathcal{E}) \longrightarrow \mathcal{E}$ (see 5.2).
2. The semi-direct product $(X,\xi) \rtimes G$ of a G-algebra (X,ξ) and the object $G \in \mathcal{E}$ is the domain part H of the pointed object $\big(p,s\colon H \leftrightarrows G\big)$ corresponding to (X,ξ) via the equivalence $\mathsf{Pt}_G(\mathcal{E}) \cong \mathcal{E}^{\mathbb{T}_G}$.

Let us emphasize at once an important property of the semi-direct product:

Theorem 5.4 *Let $\mathcal{E}$ be a semi-abelian category, $G \in \mathcal{E}$ an object of $\mathcal{E}$ and (X,ξ) a G-algebra. There exists a split short exact sequence*

$$\begin{tikzcd}
\mathbf{0} \arrow[r] & X \arrow[r, tail, "k"] & (X,\xi)\rtimes G \arrow[r, shift right, "p"'] & G \arrow[l, shift right, "s"'] \arrow[r] & \mathbf{0}.
\end{tikzcd}$$

Proof The morphisms p and s are those of definition 5.3, thus $p\circ s = 0$. By commutativity of the triangle

$$\begin{tikzcd}
\mathsf{Pt}_G(\mathcal{E}) \arrow[rr, "\cong"] \arrow[dr, "\mathsf{Ker}"'] & & \mathcal{E}^{\mathbb{T}_G} \arrow[dl, "U_G"] \\
& \mathcal{E} &
\end{tikzcd}
\qquad U_G(X,\xi) = X,$$

and again with the notation of 5.3, $\mathsf{Ker}\, p \cong X$. □

It remains now to convince the reader that in the category Gp of groups, this definition recaptures the classical notions of a G-group and the semi-direct product of groups. Let us first recall the basic elements about semi-direct products of groups.

Definition 5.5 Let $(G, \cdot)$ be a group. A G-group is a triple $(X, +, m)$ consisting of a group $(X, +)$ and an action

$$m\colon G \times X \longrightarrow X, \ (g, x) \mapsto gx$$

which satisfies the axioms

$$1x = x, \ g'(gx) = (g \cdot g')x, \ g(x + x') = gx + gx'$$

for all elements $g, g' \in G$ and $x, x' \in X$.

It is immediate to observe that in the conditions of definition 5.5, the equalities $g0 = 0$ and $g(-x) = -(gx)$ hold for all elements $g \in G$ and $x \in X$.

The semi-direct product is then classically defined via the following well-known proposition:

Proposition 5.6 *Let G be a group and $(X, +, m)$ a G-group. The set $X \times G$, provided with the multiplication*

$$(x, g) \star (x', g') = (x + gx', g \cdot g')$$

is a group. This group is called the semi-direct product of $(X, +, m)$ and $(G, \cdot)$ and is written $(X, m) \rtimes G$.

Proof See [27] or any classical text on group theory. □

The key result to exhibit the link with definition 5.3 is then:

Proposition 5.7 *Let* Gp *be the category of groups and group homomorphisms and let G be a fixed group. The category of pointed objects* $\mathsf{Pt}_G(\mathsf{Gp})$ *is equivalent to the category of G-groups and their morphisms.*

Proof A morphism $f\colon (X, +, m) \longrightarrow (Y, +, n)$ of G-groups is, of course, a group homomorphism $f\colon X \longrightarrow Y$ which commutes with the actions of G, that is, $f(gx) = gf(x)$ for all elements $x \in X$ and $g \in G$.

Every G-group $(X, +, m)$ yields a pointed object in $\mathsf{Pt}_G(\mathsf{Gp})$

$$p_G, i_G\colon (X, m) \rtimes G \rightleftarrows G, \ p_G(x, g) = g, \ i_G(g) = (0, g).$$

This construction extends easily in a functor

$$\Pi\colon G\text{-}\mathsf{Gp} \longrightarrow \mathsf{Pt}_G(\mathsf{Gp}).$$

Conversely, given a group $(H, \star)$ and a pointed object

$$p, s\colon H \rightleftarrows G, \ p \circ s = \mathrm{id}_G$$

in $\mathsf{Pt}_G(\mathsf{Gp})$, we define the group $(X, +)$ to be the kernel of p. We provide X with a G-action by defining

$$m\colon G \times X \longrightarrow X, \ (g, x) \mapsto gx = s(g) \star x \star s(g)^{-1};$$

checking the axioms of definition 5.5 is routine. Again this construction extends easily in a functor

$$\Gamma\colon \mathsf{Pt}_G(\mathsf{Gp}) \longrightarrow G\text{-}\mathsf{Gp}.$$

It remains to observe that both constructions are mutually inverse, which is straightforward computation left to the reader. □

We are now ready to conclude. By 5.7 and 5.2, we have equivalences of categories

$$\begin{tikzcd} G\text{-}\mathsf{Gp} \arrow[r, "\Pi", "\cong"'] \arrow[dr, "U"'] & \mathsf{Pt}_G(\mathsf{Gp}) \arrow[r, "\Upsilon", "\cong"'] \arrow[d, "\mathsf{Ker}"] & \mathsf{Gp}^{\mathbb{T}_G} \arrow[dl, "U_G"] \\ & \mathsf{Gp} & \end{tikzcd}$$

where

$$U(X,m) = X,\ \Pi(X,m) = \big(p, s\colon (X,m) \rtimes G \leftrightarrows G\big),\ (U_G \circ \Upsilon)(p,s) = \mathsf{Ker}\, p.$$

In particular the $\mathbb{T}$-algebra $(X', \xi) = (\Upsilon \circ \Pi)(X,m)$ corresponding to the G-group (X,m) is such that $X' = \mathsf{Ker}\, p = X$. This proves that providing a group X with a G-action is equivalent to provide it with a $\mathbb{T}_G$-action. Moreover the object part of the corresponding pointed object over G is the classical semi-direct product $(X,m) \rtimes G$.

6 Commutators and centrality

Let us now switch to the study of the "commutative aspects" of a semi-abelian category.

Definition 6.1 Let $\mathcal{E}$ be a semi-abelian category. Two morphisms f, g with the same codomain commute when there exists a (necessarily unique) morphism φ making the following diagram commutative,

$$\begin{tikzcd} A \arrow[r, "s_A"] \arrow[dr, "f"'] & A \times B \arrow[d, dotted, "\varphi"] & B \arrow[l, "s_B"'] \arrow[dl, "g"] \\ & C & \end{tikzcd}$$

where $s_A = (\mathsf{id}_A, 0)$ and $s_B = (0, \mathsf{id}_B)$. We call φ the *connector* of f and g.

By 2.4, the pair (s_A, s_B) is strongly epimorphic, from which we have the uniqueness of φ, when it exists. The terminology varies according to the authors (see [10], [12], [21], [22], [31]); ours is justified by the following example.

Example 6.2 In the category of groups, two morphisms f, g as in 6.1 commute when

$$\forall a \in A\ \forall b \in B\ f(a) + g(b) = g(b) + f(a).$$

Proof If the condition of the statement holds, the formula $\varphi(a,b) = f(a) + g(b)$ defines a group homomorphism such that $\varphi(a,0) = f(a)$ and $\varphi(0,b) = g(b)$. Conversely if a connector φ exists, then

$$f(a) + g(b) = \varphi(a,0) + \varphi(0,b) = \varphi(a,b) = \varphi(0,b) + \varphi(a,0) = g(b) + f(a).$$

□

The basic result concerning the commutation of morphisms is the possibility to force it universally. The very recent and not yet published construction in the proof of 6.3 is due to D. Bourn (see [11]) and yields a definition of the commutator (see 6.4) which gains in elegance with respect to anterior definitions (see [31]).

Theorem 6.3 *In a semi-abelian category $\mathcal{E}$, consider two morphisms*

$$A \xrightarrow{\ f\ } C \xleftarrow{\ g\ } B$$

with the same codomain. There exists a morphism $\psi\colon C \longrightarrow D$, universal for making the composites $\psi \circ f$ and $\psi \circ g$ commute. This morphism ψ is a regular epimorphism.

Proof We define D to be the colimit of the outer part of the following diagram, with the dotted arrows as colimit cocone.

$$\begin{tikzcd}
 & A \arrow[dl, "s_A"'] \arrow[d, dotted, "\alpha"'] \arrow[dr, "f"] & \\
A \times B \arrow[r, dotted, "\varphi"] & D & C \arrow[l, dotted, "\psi"'] \\
 & B \arrow[ul, "s_B"] \arrow[u, dotted, "\beta"] \arrow[ur, "g"'] &
\end{tikzcd}$$

Of course the composites $\psi \circ f$ and $\psi \circ g$ commute, with connector φ.

We must prove that given another morphism $\psi'\colon C \longrightarrow D'$ such that $\psi' \circ f$ and $\psi' \circ g$ commute (let us say, with connector φ'), then ψ' factors uniquely through ψ. Putting $\alpha' = \psi' \circ f$ and $\beta' = \psi' \circ g$, we obtain a new cocone $(\alpha', \beta', \varphi', \psi')$ with vertex D' on the same diagram, from which a unique factorization $d\colon D \longrightarrow D'$ through the colimit. In particular, $d \circ \psi = \psi'$.

Next we prove that ψ is a strong (thus regular) epimorphism, from which the uniqueness of the factorization. Indeed if ψ factors through a subobject $s\colon S \rightarrowtail D$, then by commutativity of the diagram, α, β, $\varphi \circ s_A$ and $\varphi \circ s_B$ factor through s as well. Since the colimit cocone and the pair (s_A, s_B) are strongly epimorphic (see 2.4), this forces s to be an isomorphism. Thus ψ is a strong epimorphism. $\square$

Definition 6.4 Let $\mathcal{E}$ be a semi-abelian category. Given two morphisms f, g with the same codomain, their commutator $[f, g]$ is the kernel of the universal morphism ψ of theorem 6.3.

Example 6.5 In the category Gp of groups, consider two normal subgroups $h\colon H \rightarrowtail G$ and $k\colon K \rightarrowtail G$. The commutator $[h, k]$ in the sense of definition 6.4 is the usual group $[H, K]$ of commutators.

Proof The group $[H, K]$ of commutators is the subgroup of G generated by all the elements of the form $x + y - x - y$, with $x \in H$ and $y \in K$. This is a normal subgroup, by normality of H and K. The quotient $G \longrightarrow G/[H, K]$ is well-known to be the universal morphism making H and K commute. $\square$

Centrality is a special case of interest. Still with example 6.2 in mind, we define

Definition 6.6 Let $\mathcal{E}$ be a semi-abelian category. A morphism $f\colon A \longrightarrow C$ is central when it commutes with the identity on C.

Proposition 6.7 *Let $\mathcal{E}$ be a semi-abelian category. A central morphism $f\colon A \longrightarrow C$ commutes with every morphism $g\colon B \longrightarrow C$.*

Proof If $\varphi\colon A \times C \longrightarrow C$ is the connector of f and id_C, the composite $\varphi \circ (\mathsf{id}_A \times g)$ is the connector of f and g. $\square$

Finally, again in view of example 6.2, we define:

Definition 6.8 An object C of a semi-abelian category $\mathcal{E}$ is abelian when the identity on C commutes with itself (i.e. is central).

Of course example 6.2 shows that the abelian objects of the category of groups are exactly the abelian groups. This is a much more general fact, as the following theorem proves:

Theorem 6.9 *For an object C in a semi-abelian category $\mathcal{E}$, the following conditions are equivalent:*

1. *C is abelian;*
2. *the diagonal $\Delta_C\colon C \rightarrowtail C \times C$ of C is a normal subobject;*
3. *C can be provided with the structure of an internal abelian group.*

Moreover, the structure of abelian group on C is necessarily unique.

Proof An abelian group addition $+\colon C \times C \longrightarrow C$ on C must in particular satisfy the axioms $+ \circ (\mathsf{id}_C, 0) = \mathsf{id}_C$ and $+ \circ (0, \mathsf{id}_C) = \mathsf{id}_C$. By 2.4, this proves the uniqueness of $+$.

If C is abelian, write $+\colon C \times C \longrightarrow C$ for the connector of id_C with itself. We must prove that $+$ is an abelian group structure on C. Since this reduces to prove the commutativity of diagrams involving finite limits, this can be done elementwise using our metatheorem 1.8. The letters x, y, z denote "elements" of C. By definition of $+$, $x + 0 = x = 0 + x$.

Let us write $\mathsf{tw}\colon C \times C \longrightarrow C \times C$ for the twisting isomorphism. The commutativity of $+$ is the equality $+ \circ \mathsf{tw} = +$. Since the morphisms $(\mathsf{id}_C, 0)$, $(0, \mathsf{id}_C)$ constitute an epimorphic pair (see 2.4), it suffices to prove the equality after composition with them. This reduces to proving $x + 0 = 0 + x$, which holds by definition of $+$. Analogously, iterating 2.4, it suffices to prove the associativity axiom $x + (y + z) = (x + y) + z$ when two of the variables are equal to zero, which reduces again to the definition of $+$.

Thus $+$ provides already C with the structure of a commutative monoid. To prove that it is a group, consider the diagram below, where $k = \mathsf{Ker}\,+$ and $\gamma = p_2 \circ k$. The three columns and the first two rows are short exact sequences; by the nine lemma, the last row is a short exact sequence.

$$
\begin{tikzcd}
\mathbf{0} \arrow[r, tail] \arrow[d, tail] & C \arrow[r, equal] \arrow[d, tail, "(\mathrm{id}_C, 0)"] & C \arrow[d, equal] \\
K \arrow[r, tail, "k"] \arrow[d, equal] & C \times C \arrow[r, two heads, "+"] \arrow[d, two heads, "p_2"] & C \arrow[d, two heads] \\
K \arrow[r, "\gamma"'] & C \arrow[r] & \mathbf{0}
\end{tikzcd}
$$

This proves that γ is an isomorphism (see 4.2). The composite

$$\iota \colon C \xrightarrow{\gamma^{-1}} K \xrightarrow{k} C \times C \xrightarrow{p_1} C$$

yields the opposite for the addition. Indeed by commutativity of the diagram, $(k \circ \gamma^{-1})(x)$ is a pair (y, z) such that $y = \iota(x)$, $z = x$ and $y + z = 0$.

Assuming now that C is provided with a group addition $+$, the diagonal of C is the kernel of the corresponding subtraction $- \colon C \times C \longrightarrow C$.

Finally assume that the diagonal of C is the kernel of some morphism $q \colon C \times C \longrightarrow Q$. Factoring q through its image, there is no restriction to assume that q is a regular epimorphism. Consider the diagram below where $\xi = q \circ (\mathrm{id}_C, 0)$. The three rows and the first two columns are short exact sequences; by the nine lemma, the last column is a short exact sequence as well. By 4.2, this proves that ξ is an isomorphism. We define then the subtraction "$-$" of C as the composite $\xi^{-1} \circ q$.

$$
\begin{tikzcd}
\mathbf{0} \arrow[r, tail] \arrow[d, tail] & C \arrow[r, equal] \arrow[d, tail, "(\mathrm{id}_C, 0)"] & C \arrow[d, "\xi"] \\
C \arrow[r, tail, "\Delta_C"] \arrow[d, equal] & C \times C \arrow[r, two heads, "q"] \arrow[d, two heads, "p_2"] & Q \arrow[d] \\
C \arrow[r, equal] & C \arrow[r] & \mathbf{0}
\end{tikzcd}
$$

By commutativity of the diagram, we have at once $x - x = 0$ and $x - 0 = x$. From these properties, it follows at once that the equality

$$(y - z) - (x - z) = y - x$$

is satisfied when two of the variables are equal to 0. By an iterated application of 2.4, this proves the equality in full generality. It is well-known (or routine to check) that these three axioms are those defining a group in terms of its subtraction. In particular, the group addition of C is the connector of id_C with itself. □

7 Semi-abelian versus abelian

This section intends to "measure the distance" between abelian categories and semi-abelian categories. Our first comparison (see [21], [20], [13]) between both notions focuses on the construction of the "abelian core" of a semi-abelian category.

Proposition 7.1 *Let $\mathcal{E}$ be a semi-abelian category. The full subcategory $\mathsf{Ab}(\mathcal{E})$ of abelian objects is closed in $\mathcal{E}$ under finite limits, subobjects and regular quotients. In particular, this category $\mathsf{Ab}(\mathcal{E})$ is exact.*

Proof By 6.9, an object is abelian when its diagonal is a normal subobject. Given a finite limit $L = \mathsf{lim}_{i \in I}\, G_i$ in $\mathcal{E}$, the diagonal Δ_L of L is the corresponding limit of the diagonals Δ_i of the various G_i. If each G_i is abelian, its diagonal is the kernel of its cokernel $q_i\colon G_i \twoheadrightarrow Q_i$ (see 3.2). The original diagram induces factorizations between these cokernels Q_i. The limit of these morphisms q_i admits as kernel the limit of their kernels Δ_i, that is, the diagonal Δ_L of L. This proves the normality of Δ_L. Thus the limit L in $\mathcal{E}$ is an abelian object.

If G is abelian and $i\colon H \rightarrowtail G$ is an arbitrary subobject, the diagonal of H is the pullback along i of the diagonal of G. Thus H is abelian by 3.3.

Finally consider a regular epimorphism $p\colon A \twoheadrightarrow B$ with kernel $k\colon K \rightarrowtail A$ and kernel pair $R \rightrightarrows A$. Write $q\colon A \times A \twoheadrightarrow Q$ for the cokernel of $R \rightarrowtail A \times A$. This yields the following commutative diagram

$$\begin{tikzcd}
K \times K \arrow[r, tail] \arrow[d, equal] & R \arrow[r, two heads] \arrow[d, tail] \arrow[dr, phantom, "(1)"] & B \arrow[d, tail, "\Delta_B"] \\
K \times K \arrow[r, tail, "k \times k"] \arrow[d] & A \times A \arrow[r, two heads, "p \times p"] \arrow[d, two heads, "q"'] & B \times B \arrow[d, dotted, "q'"] \\
0 \arrow[r] & Q \arrow[r, equal] & Q
\end{tikzcd}$$

where Δ_B is the diagonal of B, the square (1) is a pullback, the three rows and the first two columns are short exact sequences. This forces the existence of the factorization q' and $\Delta_B = \mathsf{Ker}\, q'$, by the nine lemma (see 4.5). Thus B is abelian. □

Theorem 7.2 *Given a semi-abelian category $\mathcal{E}$, the full subcategory $\mathsf{Ab}(\mathcal{E})$ of abelian objects is abelian and regular-epi-reflective in $\mathcal{E}$.*

Proof The category $\mathsf{Ab}(\mathcal{E})$ is exact by 7.1. By 6.9, the objects of $\mathsf{Ab}(\mathcal{E})$ are the internal abelian groups. But an arbitrary morphism $f\colon A \longrightarrow B$ between abelian objects commutes with their abelian group structures. Indeed, by 2.4 this reduces to the elementwise properties

$$f(x+0) = f(x) + f(0), \quad f(0+x) = f(0) + f(x)$$

which hold trivially since $f(0) = 0$. Thus $\mathsf{Ab}(\mathcal{E})$ is equivalent to the category of abelian groups in $\mathcal{E}$, which is additive. Therefore $\mathsf{Ab}(\mathcal{E})$ is abelian, because it is both additive and exact (see [1]).

By theorem 6.3, every object $C \in \mathcal{E}$ has an abelian reflection D given by the universal morphism $\psi\colon C \twoheadrightarrow D$ forcing the commutativity of id_C with itself. Moreover, ψ is a strong epimorphism. Observe that in this special case, the colimit diagram in the proof of 6.3 reduces to the coequalizer

$$\begin{tikzcd} C \arrow[r, shift left, "(\mathsf{id}_C, 0)"] \arrow[r, shift right, "(0, \mathsf{id}_C)"'] & C \times C \arrow[r, two heads, "\psi"] & D. \end{tikzcd}$$

□

Corollary 7.3 *A semi-abelian category $\mathcal{E}$ is abelian if and only if every object of $\mathcal{E}$ is abelian.*

Let us now investigate the distance between semi-abelian and abelian categories in terms of normal subobjects.

Proposition 7.4 *In a semi-abelian category $\mathcal{E}$, every subobject $s\colon S \rightarrowtail A$ of an abelian object A is normal.*

Proof By 7.1, s is a monomorphism in $\mathsf{Ab}(\mathcal{E})$. By 7.2, this category is abelian, thus s is a kernel in $\mathsf{Ab}(\mathcal{E})$. By 7.1 again, s is the same kernel in $\mathcal{E}$. □

Theorem 7.5 *A semi-abelian category $\mathcal{E}$ is abelian if and only if every subobject is normal in $\mathcal{E}$.*

Proof In an abelian category, every subobject is normal. Conversely, if every subobject is normal, every object is abelian (see 6.9); one concludes by 7.2. □

Let us now conclude this paper with a further comparison between the notions of semi-abelian and abelian categories (see [23]) ... a comparison which justifies fully the terminology!

Theorem 7.6 *A category $\mathcal{E}$ is abelian if and only if both $\mathcal{E}$ and its dual $\mathcal{E}^{\mathsf{op}}$ are semi-abelian.*

Proof The notion of abelian category is self-dual, thus one implication follows at once from 1.4.

If $\mathcal{E}$ is semi-abelian, the two morphisms in 2.4 yield a strongly epimorphic factorization $A \amalg B \longrightarrow A \times B$. If $\mathcal{E}^{\mathsf{op}}$ is semi-abelian as well, by duality, this same morphism is also a regular monomorphism, thus an isomorphism. Therefore the codiagonal of the coproduct yields on every object A an "addition" making the following diagram commutative:

$$\begin{tikzcd} A \arrow[r, "s_1"] \arrow[dr, equal] & A \times A \arrow[d, "+"'] & A \arrow[l, "s_2"'] \arrow[dl, equal] \\ & A & \end{tikzcd}$$

This shows that every object $A \in \mathcal{E}$ is abelian (see 6.8. By 7.3, the category $\mathcal{E}$ is abelian. □

References

[1] M. Barr, *Exact categories*, Springer Lect. Notes in Math. **236** (1971) 1–120

[2] M. Barr and J. Beck, *Homology and standard constructions*, Springer Lect. Notes in Math. **80** (1969) 245–335

[3] J. Bénabou, *Fibered categories and the foundations of naive category theory*, J. of Symbolic Logic **50** (1985) 10–37

[4] F. Borceux, *Handbook of Categorical Algebra*, Cambridge Univ. Press **vol. 1–3** (1994)

[5] F. Borceux and D. Bourn, *Mal'cev, protomodular, homological and semi-abelian categories*, Mathematics and its Applications, Kluwer (2004)

[6] D. Bourn, *Normalization equivalence, kernel equivalence and affine categories*, Springer Lect. Notes in Math. **1448** (1991) 43–62

[7] D. Bourn, *Mal'cev categories and fibrations of pointed objects*, Appl. Categorical Structures **4** (1996) 302–327

[8] D. Bourn, *Normal subobjects and abelian objects in protomodular categories*, J. Algebra **228** (2000) 143–164

[9] D. Bourn, 3×3 *lemma and protomodularity*, J. Algebra **236** (2001) 778–795

[10] D. Bourn, *Intrinsic centrality and associated classifying properties*, Cahiers LMPA, Univ. du Littoral **153** (2001)

[11] D. Bourn, *Commutator theory in regular Mal'cev categories*, this volume

[12] D. Bourn and M. Gran, *Centrality and normality in protomodular categories*, Theory Appl. Categories **9** (2002) 151–165

[13] D. Bourn and M. Gran, *Central extensions in semi-abelian categories* J. Pure Appl. Alg. **175** (2002) 31–44

[14] D. Bourn and G. Janelidze, *Protomodularity, descent and semi-direct product*, Theory Appl. Categories **4** (1998) 37–46

[15] D. Bourn and G. Janelidze, *Characterization of protomodular varieties of universal algebra*, preprint (2002)

[16] M.S. Burgin, *Categories with involution and correspondences in γ-categories*, Trans. Moscow Math. Soc. **22** (1970) 181–257

[17] A. Carboni, G. M. Kelly and M. C. Pedicchio, *Some remarks on Maltsev and Goursat categories*, Appl. Categorical Structures **1** (1993) 385–421

[18] A. Carboni, J. Lambek and M. C. Pedicchio, *Diagram chasing in Mal'cev categories*, J. Pure Appl. Alg. **69** (1990) 271–284

[19] P. Freyd, *Abelian categories*, Harper and Row (1964)

[20] M. Gran, *Central extensions and internal groupoids in Maltsev categories*, J. Pure Appl. Alg. **155** (2001) 139–166

[21] S.A. Huq, *Commutator, nilpotency and solvability in categories*, Quart. J. Math. Oxford (2)**19** (1968) 363–389

[22] G. Janelidze and M. C. Pedicchio, *Pseudogroupoids and commutator theory*, Theory Appl. Categories **8** (2001) 405–456

[23] G. Janelidze, L. Márki and W. Tholen, *Semi-abelian categories*, J. Pure Appl. Alg. **168** (2002) 367–386

[24] P.T. Johnstone, *Sketches of an elephant: a topos theory compendium*, Oxford Science Publications, vol 1–2, 2002

[25] P.T. Johnstone, *A Note on the Semiabelian Variety of Heyting Semilattices*, This volume (2002)

[26] F.W. Lawvere, *Functorial semantics of algebraic theories*, Proc. Nat. Acad. Sci. U.S.A. **50** (1963) 869–873

[27] S. Mac Lane, *Homology*, Springer Verlag (1963)

[28] S. Mac Lane, *Categories for the working mathematician*, 2nd edition, Springer Verlag (1998)

[29] S. Mac Lane and I. Moerdijk, *Sheaves in geometry and logic*, Universitext, Springer Verlag (1992)

[30] A. I. Mal'cev, *On the general theory of algebraic systems*, Mat. Sbornik N. S. **35** (1954) 3–20

[31] M. C. Pedicchio, *A categorical approach to commutator theory*, J. Algebra **177** (1995) 647–657

[32] J. D. H. Smith, *Mal'cev varieties*, Springer Lect. Notes in Math. **554** (1976)

[33] A. Ursini *Osservazioni sulla varietá BIT*, Boll. U. M. I. **(4) 7** (1973) 205–211
[34] A. Ursini, *On subtractive varieties, I*, Algebra Universalis **31** (1994) 204–222

Received October 24, 2002; in revised form December 18, 2002

Fields Institute Communications
Volume **43**, 2004

Commutator Theory in Regular Mal'cev Categories

Dominique Bourn
Université du Littoral
Laboratoire de Mathématiques Pures et Appliquées
Bât. H. Poincaré
50 Rue F. Buisson BP 699
62228 Calais, France
bourn@lmpa.univ-littoral.fr

Abstract. In the categorical Mal'cev setting, we give a new construction of the commutator $[R, S]$ of two equivalence relations which extends the existence of commutators from the exact to the regular context and includes the case of topological groups. All the classical properties are satisfied, except $[R, S] \leq R \wedge S$ which requires the category $\mathbb{C}$ to be exact.

Introduction

The notion of commutator in Mal'cev varieties has been introduced by J.D.H. Smith [22]. Once the concept of Mal'cev category ([11], [12])had been established, it was quite natural to investigate the notion of commutator in this more general setting. This was first done by Pedicchio ([20] and [21]) for exact Mal'cev categories with cokernels, in a way mimicking the varietal construction of Hagemann-Herrmann [14] which, in the Mal'cev varietal context, was proved to coincide with the original construction of Smith ([22]; see also [13]).

We present here a new construction of different nature. The idea behind this comes from two directions. The first one deals with the notion of connector ([8] and [9]), which, although expressed in a seemingly easier language, is equivalent to the notion of centralizing relation and produces a means to assert that the commutator $[R, S] = 0$ in a Mal'cev categorical setting, freed of any right exactness condition. The second one deals with the notion of unital category where there is an intrinsic notion of commutativity and centrality [6], a setting in which there is a natural categorical way to force a pair of morphisms to commute, provided that regularity holds.

2000 *Mathematics Subject Classification.* Primary 18C10, 08B05, 18D35; Secondary 08A30, 18D40, 08G15.

Key words and phrases. Unital category, central morphism, Mal'cev category, commutator, abelian object.

The relationship between these two directions is that a category $\mathbb{C}$ is Mal'cev if and only if the associated fibration $\pi : Pt\mathbb{C} \to \mathbb{C}$ of pointed objects (see below) has unital fibres [4]. Consequently it is possible to translate the unital means to force commutation from the unital setting to the Mal'cev one.

One of the advantages of this new construction is to extend the notion of commutator from the exact Mal'cev context to the regular Mal'cev one, enlarging the range of examples to the Mal'cev (quasi-varieties) and the category $Gp(Top)$ of topological groups for instance. Of course, as a pointed, finitely complete, exact and protomodular category, any semi-abelian category ([16], [2]) is included in this construction. Many of the classical properties of commutators hold with our new definition (see sections 3 and 4), and some new applications are given. In particular, we obtain a left adjoint to the inclusion of the abelian objects. However the important property $[R, S] \leq R \wedge S$ requires, here, the category $\mathbb{C}$ to be exact.

1 Unital categories

We shall suppose $\mathbb{C}$ a pointed category, i.e. a finitely complete category with a zero object. We shall denote by $\alpha_X : 1 \to X$ and $\tau_X : X \to 1$ the initial and terminal maps. Clearly the class $\Omega(\mathbb{C}) = \{0_{X,Y} = \alpha_Y.\tau_X\}$ of null maps is an ideal of $\mathbb{C}$.

1.1. Definition. A *punctual span* in the pointed category $\mathcal{C}$ is a diagram of the form

$$\begin{tikzcd} X \arrow[r, shift left, "s"] & Z \arrow[l, shift left, "f"] \arrow[r, shift right, "g"'] & Y \arrow[l, shift right, "t"'] \end{tikzcd}$$

with $f.s = 1_X, g.t = 1_Y, g.s = 0, f.t = 0$ (where 0 is the zero arrow). A punctual relation is a punctual span such that the pair of maps (f, g) is jointly monic.

1.2. Examples. For any pair (X, Y) of objects in $\mathbb{C}$, there is a canonical punctual relation which is called the coarse relation:

$$\begin{tikzcd} X \arrow[r, shift left, "l_X"] & X \times Y \arrow[l, shift left, "p_X"] \arrow[r, shift right, "p_Y"'] & Y \arrow[l, shift right, "r_Y"'] \end{tikzcd}$$

where $l_X = (1_X, 0)$ and $r_Y = (0, 1_Y)$.

1.3. Definition. A pointed category $\mathbb{C}$ is called *unital*, see [6], when for each pair (X, Y) of objects in $\mathbb{C}$, the pair of maps (l_X, r_Y) is jointly strongly epic.

In any unital category, there are no other punctual relations but the coarse ones.

1.4. Examples. A variety $\mathcal{V}$ is unital if and only if it is Jonsson-Tarski, see [3], i.e such that its theory contains a unique contant 0 and a binary term + satisfying $x + 0 = x = 0 + x$. In particular, the categories Mag, Mon, CoM, Gp, Ab, Rg of respectively unitary magmas, monoids, commutative monoids, groups, abelian groups, rings are unital.

By obvious pointwise arguments, any category $\mathbb{C}^{\mathbb{E}}$ with $\mathbb{C}$ any of the previous examples are unital.

1.5. Examples. Let $U : \mathbb{C} \to \mathbb{C}'$ be any left exact conservative (i.e. reflecting isomorphisms) functor. Then if $\mathbb{C}$ is pointed, $\mathbb{C}'$ unital implies $\mathbb{C}$ unital.

When the category $\mathbb{E}$ has products, then the Yoneda embedding: $Y : \mathbb{E} \to Set^{\mathbb{E}^{op}}$ has a natural extension $Y : Mag(\mathbb{E}) \to Mag^{\mathbb{E}^{op}}$ to the category $Mag(\mathbb{E})$ of internal unitary magmas in $\mathbb{E}$ which is still left exact and conservative. Consequently $Mag(\mathbb{E})$ is unital when $\mathbb{E}$ is finitely complete. For similar reasons, the categories $Mon(\mathbb{E})$, $CoM(\mathbb{E})$, $Gp(\mathbb{E})$, $Ab(\mathbb{E})$, $Rg(\mathbb{E})$ of respectively internal monoids, commutative monoids, groups, abelian groups, rings in $\mathbb{E}$ are unital. In particular the categories $Mon(Top)$ and $Gp(Top)$ of topological monoids and topological groups are unital.

1.6. Examples. We have a non syntactical example with the dual Set_*^{op} of the category of pointed sets, and more generally with the dual of the category of pointed objects in any topos $\mathbb{E}$.

One of the main consequences of unitality is the fact that there is an intrinsic notion of commutativity and centrality. Indeed, given a unital category $\mathbb{C}$, the pair (l_X, r_Y), being jointly strongly epic, is actually jointly epic. Therefore a map $\varphi : X \times Y \to Z$ is uniquely determined by the pair of maps (f,g), $f : X \to Z$ and $g : Y \to Z$, with $f = \varphi.l_X$ and $g = \varphi.r_Y$. Accordingly the existence of such a map φ becomes a property in respect of the pair (f,g). Whence the following definitions, see [6] and also [15]:

1.7. Definition. Given a pair (f,g) of morphisms in a unital category $\mathbb{C}$, when such a map φ exists, we say that the maps f and g *cooperate* and that the map φ is the *cooperator* of the pair (f,g). A map $f : X \to Y$ is *central* when f and 1_Y cooperate. An object X is called *commutative* when the map $1_X : X \to X$ is central.

We eventually preferred the terminology "cooperate" to "commute" because of the non syntactic examples.

1.8. Examples. 1) In the category Mag, a map $f : X \to Y$ is central if and only if the following identities hold:
i) $f(x).y = y.f(x)$, for each pair $(x,y) \in X \times Y$,
ii) $f(x).(y.y') = (f(x).y).y'$ and $(y.y').f(x) = y.(y'.f(x))$ for any $(x,y,y') \in X \times Y \times Y$.
2) In the categories Mon and Gp, a map $f : X \to Y$ is central if and only if it take its values in the center of Y. In the categories CoM and Ab, any map $f : X \to Y$ is central.
3) In the category Rg, a map $f : X \to Y$ is central if and only if $f(x).y = 0$, for each pair $(x,y) \in X \times Y$.
4) In the category Set_*^{op} the only central maps are the null maps.
5) In the category Mon of monoids, two submonoids: $H \rightarrowtail G \leftarrowtail K$ cooperate precisely when they commute, that is: $\forall x \in H\ \forall y \in K\ x.y = y.x$.
6) More generally the subcategory $Com(\mathbb{C})$ of the commutative objects in the unital category $\mathbb{C}$ is always a full subcategory of $\mathbb{C}$.

We shall suppose now that the unital category $\mathbb{C}$ is moreover finitely cocomplete and regular [1], i.e. such that the regular epimorphisms are stable by pullback and any effective equivalence relation admits a quotient. This is the case of any Jonsson-Tarski variety $\mathcal{V}$ and also of the category $Gp(Top)$ of topological groups for instance. Moreover, in a regular category, the strong epis are exactly the regular epis.

In this context, we shall construct, from any pair $f : X \to Z$, $g : Y \to Z$ of coterminal maps, a map which universally makes them cooperate. Indeed consider the following diagram, where T is the colimit of the diagram made of the plain arrows:

$$
\begin{tikzcd}
 & X \arrow[dl, "l_X"'] \arrow[d, dotted, "\phi_X"] \arrow[dr, "f"] & \\
X \times Y \arrow[r, dotted, "\phi"'] & T & Z \arrow[l, dotted, "\psi"] \\
 & Y \arrow[ul, "r_Y"] \arrow[u, dotted, "\phi_Y"'] \arrow[ur, "g"'] &
\end{tikzcd}
$$

Clearly the maps ϕ_X and ϕ_Y are completely determined by the pair (ϕ, ψ), and clearly the map ϕ is the cooperator of the pair $(\psi.f, \psi.g)$. This map ψ measures the lack of cooperation between f and g.

1.9. Proposition. *Suppose $\mathbb{C}$ unital, finitely cocomplete. Then ψ is the universal arrow which, by composition, makes the pair (f,g) cooperate. The map ψ is an iso if and only if the pair (f,g) cooperates.*

Proof First let us show that the map ψ is a strong epimorphism. If $j : U \rightarrowtail T$ is a mono which, pulled back along ψ, is an isomorphism, this is also the case along ϕ_X and ϕ_Y, and thus along $\phi.l_X$ and $\phi.r_Y$. But the pair (l_X, r_Y) is jointly epic, consequently the pullback of j along ϕ is also an iso. Now the four dotted arrows form a colimit cone, and thus j is itself an iso.

Secondly suppose given a map $\chi : Z \to W$ which makes, by means of a cooperator θ, cooperate the pair (f,g) by composition. This determines a cocone on the previous diagram:

$$
\begin{tikzcd}
 & X \arrow[dl, "l_X"'] \arrow[d, dotted] \arrow[dr, "f"] & \\
X \times Y \arrow[r, dotted, "\theta"'] & W & Z \arrow[l, dotted, "\chi"] \\
 & Y \arrow[ul, "r_Y"] \arrow[u, dotted] \arrow[ur, "g"'] &
\end{tikzcd}
$$

and consequently a factorization $\bar{\chi}$ such that $\bar{\chi}.\psi = \chi$. The uniqueness of this factorization is a consequence of the fact that ψ is a strong epimorphism. The end of the lemma is straightforward. $\square$

Whence two easy applications:

1.10. Corollary. *Suppose $\mathbb{C}$ unital, finitely cocomplete and regular. Let $f : X \to Z$ be a map. Consider the following coequalizer:*

$$\begin{tikzcd} X \arrow[rr, "l_X"] \arrow[dr, "f"'] & & X \times Z \arrow[r, "\phi"'] & T \\ & Z \arrow[ur, "r_Z"'] & & \end{tikzcd}$$

then $\psi = \phi.r_Z$ is the universal map which makes f central by composition.

Let X be any object of $\mathbb{C}$. Then the associated commutative object $\gamma(X)$ is given by the following coequalizer:

$$X \overset{l_X}{\underset{r_X}{\rightrightarrows}} X \times X \xrightarrow[\phi]{} \gamma(X)$$

In other words, the full inclusion of the commutative objects $Com(\mathbb{C}) \hookrightarrow \mathbb{C}$ admits a left adjoint γ.

Proof 1) When $g = 1_Z$ the previous colimit T is reduced to that coequalizer, and $\psi = \phi.r_Z$. Accordingly, the pair $(\psi.f, \psi)$ cooperate. But ψ, as a strong epi in a regular category, is a regular epi, and then the pair $(\psi.f, 1_Z)$ cooperate, which means that $\psi.f$ is central.

2) When $f = g = 1_X$ the previous colimit T is reduced to the coequalizer in question, and the pair (ψ, ψ) cooperates. But, again, ψ is a regular epi, and then the pair $(1_X, 1_X)$ cooperates, which means that the object X is commutative. □

2 Mal'cev categories

Let us recall that a category $\mathbb{C}$ is Mal'cev when it is finitely complete and such that every reflexive relation is an equivalence relation, see [12] and [11]. Following the famous Mal'cev argument, a variety $\mathcal{V}$ is Mal'cev if and only if its theory contains a ternary term p, satisfying : $p(x, y, y) = x = p(y, y, x)$ (called a Mal'cev operation).

There is a strong connection with unital categories which is given by the following observation. Let $\mathbb{C}$ be a finitely complete category. We denote by $Pt\mathbb{C}$ the category whose objects are the split epimorphisms in $\mathbb{C}$ with a given splitting and morphisms the commutative squares between these data. We denote by $\pi : Pt\mathbb{C} \to \mathbb{C}$ the functor associating its codomain with any split epimorphism. Since the category $\mathbb{C}$ has pullbacks, the functor π is a fibration which is called the *fibration of pointed objects.*

A finitely complete category $\mathbb{C}$ is Mal'cev if and only if the fibres of the fibration π are unital, see [4].

Now consider $(d_0, d_1) : R \rightrightarrows X$ an equivalence relation on the object X in $\mathbb{C}$. We shall denote by $s_0 : X \to R$ the inclusion arising from the reflexivity of the relation, and we shall write ΔX and ∇X respectively for the smallest $(1_X, 1_X) : X \rightrightarrows X$ and the largest $(p_0, p_1) : X \times X \rightrightarrows X$ equivalence relations on X.

Since the category $\mathbb{C}$ is Mal'cev, to give the equivalence relation $(d_0, d_1) : R \rightrightarrows X$ on the object X in $\mathbb{C}$ is equivalent to give the following inclusion of the object

$((d_0, s_0) : R \rightleftarrows X)$ into the object $((p_0, s_0) : X \times X \rightleftarrows X)$ in the fibre $Pt_X(\mathbb{C})$ above X:

$$\begin{tikzcd}
R \arrow[r, "(d_0,d_1)"] \arrow[dr, shift left, "d_0"] & X \times X \arrow[d, shift left, "p_0"] \\
& X \arrow[ul, shift left, "s_0"] \arrow[u, shift left, "s_0"]
\end{tikzcd}$$

So, by abuse of notation, we shall often identify the equivalence relation R with the object $((d_0, s_0) : R \rightleftarrows X)$ of the fibre $Pt_X(\mathbb{C})$, and conversely.

Now consider $(d_0, d_1) : R \rightrightarrows X$ and $(d_0, d_1) : S \rightrightarrows X$ two equivalence relations on the same object X in $\mathbb{C}$. Then take the following pullback:

$$\begin{tikzcd}
R \times_X S \arrow[r, shift left, "p_S"] \arrow[d, shift right, "p_R"'] & S \arrow[l, shift left, "r_S"] \arrow[d, shift right, "d_{0,S}"'] \\
R \arrow[u, shift right, "l_R"'] \arrow[r, shift left, "d_{1,R}"] & X \arrow[l, shift left, "s_{0,R}"] \arrow[u, shift right, "s_{0,S}"']
\end{tikzcd}$$

where l_R and r_S are the sections induced by the maps $s_{0,R}$ and $s_{0,S}$.

Let us recall the following definition, see [8], and also [19], [18], [12], [20]:

2.1. Definition. In a Mal'cev category $\mathbb{C}$, a connector on the pair (R, S) in a Mal'cev category is a morphism

$$p : R \times_X S \to X, \ (xRySz) \mapsto p(x, y, z)$$

which satisfies the identities : $p(x, y, y) = x$ and $p(y, y, z) = z$.

This notion actually makes sense in any finitely complete category provided that some further conditions are satisfied [8], which are always fulfilled in a Mal'cev category. Moreover, in a Mal'cev category, a connector is necessarily unique when it exists (since the pair (l_R, r_S) is jointly epic), and thus the existence of a connector becomes a property. We say then that R and S are *connected.*

2.2. Examples. By Proposition 3.6, Proposition 2.12 and definition 3.1 in [20], two relations R and S in a Mal'cev variety $\mathcal{V}$ are connected if and only if $[R, S] = 0$ in the sense of Smith [22].

Accordingly we shall denote a connected pair of equivalence relations by $[R, S] = 0$. One of the advantage of the notion of connectors is that it keeps a meaning in a context freed of any right exactness condition. So, in the same way as in the varietal case, we can say that an object X in a Mal'cev category is *abelian* when $[\nabla X, \nabla X] = 0$. Similarly, a map $f : X \to Y$ is said to have *central kernel* when its kernel equivalence $R[f]$ is such that $[R[f], \nabla X] = 0$. It is said to have *abelian kernel* when $[R[f], R[f]] = 0$ (this is clearly equivalent to saying that the object $f : X \to Y$ in the slice category $\mathbb{C}/Y$ is abelian). Consequently, for instance, a naturally Mal'cev category [17], is merely a Mal'cev category in which any object is abelian or any map has a central kernel.

On the other hand, the fibre $Pt_X(\mathbb{C})$ being unital, it is natural to ask when two subobjects R and S of ∇X cooperate in this fibre.

2.3. Proposition. *Let $\mathbb{C}$ be a Mal'cev category, the subobjects R and S of ∇X cooperate in the fibre $Pt_X(\mathbb{C})$ if and only if the equivalence relations R and S are connected in $\mathbb{C}$.*

Proof Let us consider the product of R and S in $Pt_X(\mathbb{C})$. It is given by the following pullback in $\mathbb{C}$:

$$\begin{tikzcd} R \times_0 S \arrow[r] \arrow[d] & S \arrow[d, "d_0"] \\ R \arrow[r, "d_0"'] & X \end{tikzcd}$$

A cooperator between R and S in $Pt_X(\mathbb{C})$ is thus a map $\phi : R \times_0 S \to X \times X$ such that $\phi(x,y,x) = (x,y)$ and $\phi(x,x,z) = (x,z)$. But ϕ is a morphism in the fibre and necessarily is of the form $\phi(x,y,z) = (x, q(x,y,z))$. Accordingly a cooperator between R and S is just a map $q : R \times_0 S \to X$ such that $q(x,y,x) = y$ and $q(x,x,z) = z$. Consequently to set $p(u,v,w) = q(v,u,w)$ is to define a bijection between the cooperators and the connectors. $\square$

So the two points of view are equivalent. We shall prefer the one of connectors because the guiding results of [9] concerning the commutators are given in these terms.

From this observation, and the universal construction of the first section, we shall derive a new construction of the commutator. We shall suppose from now on the category $\mathbb{C}$ finitely cocomplete, Mal'cev and regular, as is the case for the Mal'cev (quasi-varieties) and for the category $Gp(Top)$ of topological groups for instance. Regular Mal'cev categories have stronger stability properties than unital regular ones since, in particular, they are stable by slice and coslice. We shall appreciate below the advantages of these stability properties.

In a regular Mal'cev category, given a regular epi $f : X \to Y$, any equivalence relation R on X has a direct image $f(R)$ along f on Y. It is given by the regular epi/mono factorization of the map $(f.d_0, f.d_1) : R \twoheadrightarrow f(R) \rightarrowtail Y \times Y$. Clearly in any regular category $\mathbb{C}$, the relation $f(R)$ is reflexive and symmetric. When moreover $\mathbb{C}$ is Mal'cev, $f(R)$ is an equivalence relation.

Now let us consider the following diagram (which we shall denote $D(R,S)$) where T is the colimit of the plain arrows:

$$\begin{tikzcd} & R \arrow[dl, "l_R"'] \arrow[dr, "d_{0,R}"] \arrow[d, dotted, "\phi_R"] & \\ R \times_X S \arrow[r, dotted, "\phi"'] & T & X \arrow[l, dotted, "\psi"] \\ & S \arrow[ul, "r_S"] \arrow[ur, "d_{1,S}"'] \arrow[u, dotted, "\phi_S"'] & \end{tikzcd}$$

Notice that, in consideration of the pullback defining $R \times_X S$, the roles of the projections d_0 and d_1 have been interchanged. As above, the maps ϕ_R and ϕ_S are completely determined by the pair (ϕ, ψ). We shall see that this map ψ measures the lack of connection between R and S.

2.4. Theorem. *Let the category $\mathbb{C}$ be finitely cocomplete, Mal'cev and regular. Then the map ψ is the universal regular epimorphism which makes the images $\psi(R)$ and $\psi(S)$ connected. The equivalence relations R and S are connected (i.e. $[R, S] = 0$) if and only if ψ is an isomorphism.*

Proof First, the map ψ is a strong epi (and thus a regular epi) for exactly the same reasons as in the previous construction in unital categories. Secondly let us denote by $\psi_R : R \to \psi(R)$ and $\psi_S : S \to \psi(S)$ the respective regular factorizations. Thanks to Proposition 4.1 in [7], the induced factorization $\bar{\psi} : R \times_X S \to \psi(R) \times_T \psi(S)$ is itself a regular epi. The aim now will be to show that $\phi : R \times_X S \to T$ factors through $\psi(R) \times_T \psi(S)$, producing a connector $p : \psi(R) \times_T \psi(S) \to T$ on $\psi(R)$ and $\psi(S)$. For that we must show that ϕ coequalizes the kernel relation $R[\bar{\psi}]$ of $\bar{\psi}$. Clearly $R[\bar{\psi}]$ is obtained by the following pullback:

$$\begin{tikzcd}
R[\bar{\psi}] \arrow[r, shift left, "R(p_S)"] \arrow[d, shift right, "R(p_R)"'] & R[\psi_S] \arrow[l, shift left, "R(r_S)"] \arrow[d, shift right, "R(d_{0,S})"'] \\
R[\psi_R] \arrow[r, shift left, "R(d_{1,R})"] \arrow[u, shift right, "R(l_R)"'] & R[\psi] \arrow[l, shift left, "R(s_{0,R})"] \arrow[u, shift right, "R(s_{0,S})"']
\end{tikzcd}$$

But the fibre $Pt_{R[\psi]}(\mathbb{C})$ is unital, and consequently the pair $(R(l_R), R(r_S))$ is jointly (strongly) epic. So it is sufficient to check the coequalization in question by composition with this pair of maps, which is straightforward. The axioms asserting that p is a connector for the pair of equivalence relations $\psi(R)$ and $\psi(S)$ are a consequence of the form of the colimit T.

Now assume that we have a regular epi $\chi : X \to Y$ such that the images $\chi(R)$ and $\chi(S)$ are connected by a map $\theta : \chi(R) \times_Y \chi(S) \to Y$. Consider the induced factorization $\bar{\chi} : R \times_X S \to \chi(R) \times_Y \chi(S)$. Then the following cocone produces the required factorization $\chi' : T \to Y$:

$$\begin{tikzcd}
 & R \arrow[dl, "l_R"'] \arrow[dr, "d_{0,R}"] \arrow[d, dotted] & \\
R \times_X S \arrow[r, dotted, "\theta.\bar{\chi}"'] & Y & X \arrow[l, dotted, "\chi"] \\
 & S \arrow[ul, "r_S"] \arrow[ur, "d_{1,S}"'] \arrow[u, dotted] &
\end{tikzcd}$$

The last point of the theorem is straightforward. □

The map ψ being a regular epi, its distance from being an isomorphism is its distance from being a monomorphism, which is exactly measured by its kernel relation $R[\psi]$. Accordingly it is meaningful to introduce the following definition:

2.5. Definition. Let the category $\mathbb{C}$ be finitely cocomplete, Mal'cev and regular. Let two equivalence relations $(d_0, d_1) : R \rightrightarrows X$ and $(d_0, d_1) : S \rightrightarrows X$ be given on the same object X in $\mathbb{C}$. The kernel relation $R[\psi]$ of the map ψ is called the commutator of R and S. It is classically denoted by $[R, S]$.

2.6. Examples. If we suppose moreover the category $\mathbb{C}$ exact [1], namely such that any equivalence relation is effective, i.e. the kernel relation of some map, then thanks to Theorem 3.9 in [20], the previous definition is equivalent to the definition of [20], and accordingly to the definition of Smith [22] in the Mal'cev varietal context.

So one of the advantage of this definition is that it extends the meaning of commutator from the exact Mal'cev context to the regular Mal'cev one, enlarging the range of examples to the Mal'cev (quasi-varieties), or to the case of the topological groups for instance.

3 Universal constructions

Let us come back to a finitely cocomplete regular Mal'cev category $\mathbb{C}$. Let us write $Rel(\mathbb{C})$ for the following category:

1. the objects are the pairs (X, R) where X is an object of $\mathbb{C}$ and R is an equivalence relation of X;
2. a morphism $f : (X, R) \to (X', R')$ is a morphism $f : X \to X'$ in $\mathbb{C}$ such that $R \leq f^{-1}(R')$.

We shall write further $Rel^2(\mathbb{C})$ for the category:

1. whose objects are the triples (X, R, S) of an object X and two equivalence relations R, S on X;
2. the morphisms $f : (X, R, S) \to (X', R', S')$ are the morphisms f such that $R \leq f^{-1}(R')$ and $S \leq f^{-1}(S')$.

Finally we write $ZRel(\mathbb{C})$ for the subcategory (actually full subcategory saturated for subobjects, see [9], and [3] for the details) of $Rel^2(\mathbb{C})$ whose objects are the triples (X, R, S) such that R and S admit a connector, and we denote by $j : ZRel(\mathbb{C}) \to Rel^2(\mathbb{C})$ the full inclusion. Then the construction T easily extends to a left adjoint $T : Rel^2(\mathbb{C}) \to ZRel(\mathbb{C})$ of j. This remark has a number of interesting consequences.

3.1. Corollary. *Let $\mathbb{C}$ be a finitely cocomplete regular Mal'cev category. Given an object X, the object $T(\nabla_X, \nabla_X)$ is the abelian object universally associated with X. This yields the left adjoint to the full inclusion $Ab(\mathbb{C}) \hookrightarrow \mathbb{C}$ of the abelian objects of $\mathbb{C}$.*

Proof Straightforward, since the object X is abelian if and only if the pair (∇_X, ∇_X) is connected. □

If we "localize" the previous result in the slice category $\mathbb{C}/Y$, we obtain:

3.2. Corollary. *Let $\mathbb{C}$ be a pointed finitely cocomplete regular Mal'cev category. Every map $f : X \to Y$ in $\mathbb{C}$ factorizes universally through a map with abelian kernel.*

Proof The slice category $\mathbb{C}/Y$ is finitely cocomplete, regular and Mal'cev. We noticed that a map $f : X \to Y$ is an abelian object in $\mathbb{C}/Y$ if and only if it has abelian kernel. By the previous corollary, the object $f : X \to Y$ has a universally associated abelian object $\tilde{X} : \tilde{f} \to Y$ in $\mathbb{C}/Y$ and the universal morphism $h : (X, f) \to (\tilde{X}, \tilde{f})$ in $\mathbb{C}/Y$ yields the factorization of f through the map $\tilde{f}$ with abelian kernel. □

3.3. Corollary. *Let $\mathbb{C}$ be a finitely cocomplete, regular Mal'cev category. Every map $f : X \to Y$ factorizes universally through a map with central kernel.*

Proof Clearly the map f in the category $\mathbb{C}$ determines a map $f : (X, R[f], \nabla_X) \to (Y, \Delta_Y, \nabla_Y)$ in $Rel^2(\mathbb{C})$. Moreover, (Y, Δ_Y, ∇_Y) lies in $ZRel(\mathbb{C})$. So, we get a factorization $f' : T(R[f], \nabla_X) \to Y$ such that $f = f'.\psi$, where $\psi : X \to T(R[f], \nabla_X)$ is the canonical strong epimorphism. Since ψ is a regular epimorphism, $\psi(\nabla_X) = \nabla_T$; moreover, the factorization $\overline{\psi} : R[f] \to R[f']$ induced by ψ is itself a regular epimorphism, proving that $\psi(R[f]) = R[f']$. By the universal property of ψ, the relations $\psi(R[f]) = R[f']$ and $\psi(\nabla_X) = \nabla_T$ are connected, proving that f' has central kernel. □

We shall also be able to associate a groupoid with any reflexive graph. Let us recall, see [9] for instance, that, given a reflexive graph on an object X in a Mal'cev category:

$$\begin{tikzcd} G \arrow[r, shift left=2, "d_1"] \arrow[r, shift right=2, "d_0"'] & X \arrow[l, "s_0"'] \end{tikzcd}$$

this graph is underlying a groupoid if and only if the commutator $[R[d_0], R[d_1]] = 0$. Consequently there is at most one structure of groupoid on a given reflexive graph. Moreover the inclusion $j : Grd(\mathbb{C}) \hookrightarrow Gph(\mathbb{C})$ of the internal groupoids into the reflexive graphs is a full inclusion saturated for subobjects, see [4], and also [12].

3.4. Proposition. *Let $\mathbb{C}$ be a finitely cocomplete, regular Mal'cev category. The full inclusion $j : Grd(\mathbb{C}) \hookrightarrow Gph(\mathbb{C})$ admits a left adjoint.*

Proof Let us start with a reflexive graph. Then consider the map $\psi : G \to X_1 = T(R[d_0], R[d_1])$. The morphism d_0 is a split epimorphism, with section s_0. The image $d_0(R[d_0])$ is simply ΔX. Thus it is connected with every equivalence relation on X, and in particular with $d_0(R[d_1])$. So we get a factorization $\overline{d_0} : X_1 \to X$ such that $\overline{d_0}.\psi = d_0$. In the same way, there is a morphism $\overline{d_1} : X_1 \to X$ such that $\overline{d_1}.\psi = d_1$. Set $\overline{s_0} = \psi.s_0$. It is then obvious that

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "\overline{d_1}"] \arrow[r, shift right=2, "\overline{d_0}"'] & X \arrow[l, "\overline{s_0}"'] \end{tikzcd}$$

is a reflexive graph and that $\psi : G \to X_1$ is a morphism of reflexive graphs. Since ψ is a regular epimorphism, we have $\psi(R[d_i]) = R[\overline{d_i}]$. Thus $R[\overline{d_0}]$ and $R[\overline{d_1}]$ are connected by definition of $X_1 = T(R[d_0], R[d_1])$. So the previous graph is provided with a structure of a groupoid. The universal property of this construction follows easily from the universal property of T. □

4 Commutator theory

We are now going to investigate the properties of the commutator.

4.1 The regular context. Let us begin with the regular context. We shall later study the exact one.

4.2. Proposition. *Let two equivalence relations $(d_0, d_1) : R \rightrightarrows X$ and $(d_0, d_1) : S \rightrightarrows X$ be given on the same object X in $\mathbb{C}$. Then $[R, S] = [S, R]$.*

Proof The symmetry of the relations interchanges the d_0 and the d_1; it induces a natural isomorphism between the diagrams $D(R, S)$ and $D(S, R)$. Accordingly the colimits $T(R, S)$ and $T(S, R)$ are isomorphic, and $[R, S] = [S, R]$. □

4.3. Proposition. *Let three equivalence relations R, R' and S be given on X. Then $R' \leq R$ implies $[R', S] \leq [R, S]$.*

Proof Clearly the inclusion $R' \leq R$ determines a natural transformation from $D(R', S)$ to $D(R, S)$. Accordingly there is a factorization $T(R', S) \to T(R, S)$ which implies that $[R', S] \leq [R, S]$. □

We have also:

4.4. Proposition. *Suppose given pairs of equivalence relations (R, R') on X and (S, S') on X'. Then $[R \times S, R' \times S'] \leq [R, R'] \times [S, S']$.*

Proof This comes from the functoriality of the construction T. The projection $p_X : X \times X' \to X$ determines a map $p_X : (X \times X', R \times S, R' \times S') \to (X, R, R')$ in $Rel^2(\mathbb{C})$ which gives an image $T(p_X) : T(R \times S, R' \times S') \to T(R, R')$. The factorization $(T(p_X), T(p_{X'})) : T(R \times S, R' \times S') \to T(R, R') \times T(S, S')$ provides the required inequality. □

In a regular Mal'cev category, the equivalence relations on X can be composed, and we have moreover $R \circ S = S \circ R = R \vee S$, see [11]. The composite relation $R \circ S$ is defined in the following way: consider the following pullback:

$$\begin{array}{ccc} R \times_X S & \xrightarrow{p_S} & S \\ {\scriptstyle p_R}\downarrow & & \downarrow{\scriptstyle d_{0,S}} \\ R & \xrightarrow[d_{1,R}]{} & X \end{array}$$

and take the canonical regular-epi/mono factorization of the map $(d_0.p_0, d_1.p_1) : R \times_X S \to X \times X$, $(x, y, z) \mapsto (x, z)$, namely:

$$R \times_X S \twoheadrightarrow R \circ S \rightarrowtail X \times X$$

When R and S are two equivalence relations on X, then, in a regular category, $R \circ S$ is a reflexive relation which is not necessarily symmetric nor transitive. But this is clearly the case in a regular Mal'cev category.

4.5. Proposition. *Let three equivalence relations R, S_1 and S_2 be given on X. Then $[R, S_1 \vee S_2] = [R, S_1] \vee_{eff} [R, S_2]$, that is the supremum of $[R, S_1]$ and $[R, S_2]$ in the poset of effective equivalence relations on X.*

Proof The following square is trivially a pushout in $Rel^2(\mathbb{C})$:

$$\begin{tikzcd}
(X,R,\Delta X) \arrow[r, "1_X"] \arrow[d, "1_X"'] & (X,R,S_1) \arrow[d, "1_X"] \\
(X,R,S_2) \arrow[r, "1_X"'] & (X,R,S_1 \vee S_2)
\end{tikzcd}$$

Accordingly its image by the left adjoint functor T is still a pushout in $ZRel(\mathbb{C})$:

$$\begin{tikzcd}
T(R,\Delta X) \arrow[r, "\psi_1"] \arrow[d, "\psi_2"'] \arrow[dr, dotted, "\psi"] & T(R,S_1) \arrow[d, "\gamma_1"] \\
T(R,S_2) \arrow[r, "\gamma_2"'] & T(R,S_1 \vee S_2)
\end{tikzcd}$$

But clearly R and ΔX are connected, so that $T(R,\Delta X) = X$, and, finally, the following diagram is a pushout in $\mathbb{C}$:

$$\begin{tikzcd}
X \arrow[r, "\psi_1"] \arrow[d, "\psi_2"'] \arrow[dr, dotted, "\psi"] & T(R,S_1) \arrow[d, "\gamma_1"] \\
T(R,S_2) \arrow[r, "\gamma_2"'] & T(R,S_1 \vee S_2)
\end{tikzcd}$$

This precisely means that:

$$[R, S_1 \vee S_2] = R[\psi] = R[\psi_1] \vee_{eff} R[\psi_2] = [R,S_1] \vee_{eff} [R,S_2].$$

□

4.6. Proposition. *Let $f : X \to Y$ be a regular epi, and (R,S) a pair of equivalence relations on X. Then $[f(R), f(S)] = f(R[f] \vee_{eff} [R,S])$, and $f([R,S]) \leq [f(R), f(S)]$.*

Proof The following diagram is trivially a pushout in $Rel^2(\mathbb{C})$:

$$\begin{tikzcd}
(X,R,\Delta X) \arrow[r, "f"] \arrow[d, "1_X"'] & (Y,f(R),\Delta Y) \arrow[d, "1_Y"] \\
(X,R,S) \arrow[r, "f"'] & (Y,f(R),f(S))
\end{tikzcd}$$

Accordingly its image by the left adjoint functor T is still a pushout in $ZRel(\mathbb{C})$:

$$\begin{tikzcd}
X = T(R,\Delta X) \arrow[r, "f"] \arrow[d, "\psi"'] & T(f(R),\Delta Y) = Y \arrow[d, "\psi'"] \\
T(R,S) \arrow[r, "T(f)"'] & T(f(R),f(S))
\end{tikzcd}$$

Therefore we have:

$$f^{-1}([f(R), f(S)]) = f^{-1}(R[\psi']) = R[\psi'.f] = R[f] \vee_{eff} R[\psi] = R[f] \vee_{eff} [R,S]$$

But in a regular Mal'cev category, we have always $f(f^{-1}(\Sigma)) = \Sigma$ for any equivalence relation Σ on Y. Accordingly:

$$[f(R), f(S)] = f(f^{-1}([f(R), f(S)])) = f(R[f] \vee_{eff} [R, S])$$

On the other hand:

$$f([R, S]) = \Delta Y \vee f([R, S]) = f(R[f]) \vee f([R, S]) = f(R[f]) \vee [R, S])$$

But $R[f] \vee [R, S] \leq R[f] \vee_{eff} [R, S]$, and $f([R, S]) \leq [f(R), f(S)]$. $\square$

4.7. Corollary. *Let $\mathbb{C}$ be a cocomplete regular Mal'cev category and $f : X \to Y$ a regular epi. Given (U, V) any pair of equivalence relations on Y, we have:*

$$f^{-1}[U, V] = R[f] \vee_{eff} [f^{-1}(U), f^{-1}(V)]$$

Proof Taking $R = f^{-1}(U)$ and $S = f^{-1}(V)$, the following square is a pushout as above:

$$\begin{tikzcd} X \arrow[r, "f"] \arrow[d, "\psi"'] & Y \arrow[d, "\psi'"] \\ T(f^{-1}(U), f^{-1}(V)) \arrow[r, "T(f)"'] & T(U, V) \end{tikzcd}$$

Accordingly, $R[\psi'.f] = R[f] \vee_{eff} R[\psi] = R[f] \vee_{eff} [f^{-1}(U), f^{-1}(V)]$, while $R[\psi'.f] = f^{-1}(R[\psi']) = f^{-1}[U, V]$. $\square$

4.8 The exact context. We shall suppose from now on the category $\mathbb{C}$ finitely cocomplete, exact and Mal'cev, as is the case for the Mal'cev varieties. We already noticed that, in the exact context, our construction of the commutator coincides with the one previously given by Pedicchio, and consequently coincides in the Mal'cev varietal context with Smith's one. We shall then recover easily all the main identities.

4.9. Proposition. *We have always $[R, S_1 \vee S_2] = [R, S_1] \vee [R, S_2]$.*

Proof Straightforward, since any equivalence relation being effective, we have "$\vee = \vee_{eff}$". $\square$

4.10. Proposition. *The commutators are stable by regular direct images.*

Proof Let $f : X \to Y$ be a regular epi. We noticed that the following square is always a pushout:

$$\begin{tikzcd} X = T(R, \Delta X) \arrow[r, "f"] \arrow[d, "\psi"'] & T(f(R), \Delta Y) = Y \arrow[d, "\psi'"] \\ T(R, S) \arrow[r, "T(f)"'] & T(f(R), f(S)) \end{tikzcd}$$

But in an exact Mal'cev category, any pushout of regular epimorphisms is a regular pushout, which means that the factorization through the induced pullback is a regular epi, see [7] and also [10]. Then according to Lemma 2.1 in [7], the induced factorization

$$\tilde{f} : R[\psi] = [R, S] \to R[\psi'] = [f(R), f(S)]$$

is a regular epi, and consequently $f([R,S]) = [f(R), f(S)]$. □

Similarly:

4.11. Corollary. *Let $f : X \to Y$ be a regular epi. Given (U,V) any pair of equivalence relations on Y, we have:*

$$f^{-1}[U,V] = R[f] \vee [f^{-1}(U), f^{-1}(V)]$$

And finally:

4.12. Theorem. *We have always $[R,S] \leq R \wedge S$.*

Proof The result will follow from $[R,S] \leq R$. Take the quotient $q : X \twoheadrightarrow X/R$. Then the direct image $q(R) = \Delta(X/R)$. Thus $q(R)$ is connected to any relation, and in particular to $q(S)$. Therefore there is a factorization $T(R,S) \to X/R$ which gives $[R,S] \leq R$. □

References

[1] M. Barr, *Exact categories*, Lecture Notes in Math., **236**, Springer, 1971, 1-120.

[2] F. Borceux, *A survey of semi-abelian categories*, this volume.

[3] F. Borceux and D. Bourn, *Mal'cev, protomodular, homological and semi-abelian categories*, Kluwer, to appear 2004.

[4] D. Bourn, *Mal'cev categories and fibration of pointed objects*, Appl. Categorical Structures, **4**, 1996, 307-327.

[5] D. Bourn, *Normal subobjects and abelian objects in protomodular categories*, Journal of Algebra, **228**, 2000, 143-164.

[6] D. Bourn, *Intrinsic centrality and associated classifying properties*, Journal of Algebra, **256**, 2002, 126-145.

[7] D. Bourn, *The denormalized* 3×3 *lemma*, J. Pure Appl. Algebra, **177**, 2003, 113-129.

[8] D. Bourn and M. Gran, *Centrality and normality in protomodular categories*, Th. and Applications of Categories, **9**, 2002, 151-165.

[9] D. Bourn and M. Gran, *Centrality and connectors in Maltsev categories*, Algebra Univers., **48**, 2002, 309-331.

[10] A. Carboni, G.M. Kelly and M.C. Pedicchio, *Some remarks on Maltsev and Goursat categories*, Appl. Categorical Structures, **1**, 1993, 385-421.

[11] A. Carboni, J. Lambek and M.C. Pedicchio, *Diagram chasing in Mal'cev categories*, J. Pure Appl. Algebra, **69**, 1991, 271-284.

[12] A. Carboni, M.C. Pedicchio and N. Pirovano, *Internal graphs and internal groupoids in Mal'cev categories*, CMS Conference Proceedings, **13**, 1992, 97-109.

[13] R. Freese and R. McKenzie, *Commutator theory for congruence modular varieties*, LMS Lecture Notes, **125**, 1987.

[14] J. Hagemann and C. Hermann, *A concrete ideal multiplication for algebraic systems and its relation to congruence distributivity*, Arch. Math. , **32**, 1979, 234-245.

[15] S.A. Huq, *Commutator, nilpotency and solvability in categories*, Quart. J. Oxford, **19**, 1968, 363-389.

[16] G. Janelidze, L. Marki and W. Tholen, *Semi-abelian categories*, J. Pure Appl. Algebra, **168**, 2002, 367-386.

[17] P.T. Johnstone, *Affine categories and naturally Mal'cev categories*, J. Pure Appl. Algebra, **61**, 1989, 251-256.

[18] P.T. Johnstone, *The closed subgroup theorem for localic herds and pregroupoids*, J. Pure Appl. Algebra, **70**, 1989, 97-106.

[19] A. Kock, *The algebraic theory of moving frames*, Cahiers Top. Géom. Diff. Cat., **23**, 1982, 347-362.

[20] M.C. Pedicchio, *A categorical approach to commutator theory*, Journal of Algebra, **177**, 1995, 647-657.

[21] M.C. Pedicchio, *Arithmetical categories and commutator theory*, Appl. Categorical Structures, **4**, 1996, 297-305.

[22] J.D.H. Smith, *Mal'cev varieties*, Lecture Notes in Math., **554**, Springer, 1976.

Received December 9, 2002; in revised form February 26, 2003

Fields Institute Communications
Volume **43**, 2004

Categorical Aspects of Modularity

Dominique Bourn
Lab. Mathématiques Pures et Appliquées
Université du Littoral Côte d'Opale
50 Rue F. Buisson BP 699
62228 Calais, France
bourn@lmpa.univ-littoral.fr

Marino Gran
Lab. Mathématiques Pures et Appliquées
Université du Littoral Côte d'Opale
50 Rue F. Buisson BP 699
62228 Calais, France
gran@lmpa.univ-littoral.fr

Abstract. We investigate internal categorical structures in varieties of universal algebras and in more general categories. Special attention is paid to pseudogroupoids, which play a central role in commutator theory. We work in the general context of categories satisfying the shifting property, given by a categorical formulation of Gumm's Shifting Lemma. This general context includes in particular any congruence modular variety as well as any regular Maltsev category. Several characterizations of internal categories and internal groupoids in these categories are given.

Introduction

The purpose of this paper is to present some new results in the categorical approach to commutator theory and centrality.

The theory of commutators, first developed by Smith [22] in the context of Maltsev varieties, was then extended by Hagemann and Hermann to congruence modular varieties [14] [13] [10]. This theory can be considered as an extension of the classical notion of commutator for groups to more general varieties of universal algebras.

The categorical approach to commutator theory [20] [21] [15] [5] focuses on the connection between the properties of some internal categorical structures in varieties of universal algebras and the properties of the commutators. In many important

2000 *Mathematics Subject Classification.* Primary 08B10, 08C05, 18C05; Secondary 08A30,18D05.

Key words and phrases. Modular varieties, commutators, abelian objects, internal categories and groupoids, pseudogroupoids, shifting lemma.

algebraic varieties these internal structures are themselves of special mathematical interest: for instance, the category of internal groupoids in the category of groups, which is itself a variety of universal algebras, is important in homotopy theory [19], since it is equivalent to the variety of crossed modules [6].

Various descriptions of internal categories and internal groupoids were obtained by Janelidze and Pedicchio in the general context of congruence modular varieties [15]. These results clarified how deep the relationship between commutators and internal groupoids in modular varieties is. A good illustration of this fact is also provided by the following characterization of internal groupoids [12]: given an internal reflexive graph X in a modular variety

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"'] \end{tikzcd}$$

X has a (unique) internal groupoid structure if and only if

$$(1) \quad [R[d_0], R[d_1]] = \Delta_{X_1} \qquad \text{and} \qquad (2) \quad R[d_0] \circ R[d_1] = R[d_1] \circ R[d_0]$$

(where $R[d_0]$ and $R[d_1]$ are the congruences arising as the kernel pairs of d_0 and d_1, respectively, and Δ_{X_1} is the smallest congruence on X_1).

Important progress was made when the internal structure of *pseudogroupoid* was introduced [16]. In any modular variety it was proved that two congruences R and S on an algebra X have trivial commutator $[R, S] = \Delta_X$ if and only if there is a (unique) pseudogroupoid structure on R and S.

In this paper we investigate internal pseudogroupoids, internal categories and internal groupoids in a very general categorical context. This context includes, on the one hand, any congruence modular variety and, on the other hand, any regular Maltsev category [8], so in particular also the categories of topological groups, Hausdorff groups and torsion-free abelian groups [7].

The main axiom we require, that we call the *shifting property*, is a categorical formulation of the well-known Shifting Lemma for modular varieties: it was Gumm who proved that, for a variety of universal algebras, the validity of the Shifting Lemma is equivalent to congruence modularity [13]. In categorical language, the shifting property can be expressed very simply by the requirement that a certain kind of internal functors between internal equivalence relations are discrete fibrations (see Lemma 2.2). In the presence of this weak assumption, the internal categorical structures behave surprisingly well: the notion of pseudogroupoid can be significantly simplified, and a pseudogroupoid on two equivalence relations is necessarily unique, when it exists. Accordingly, for two given equivalence relations, having a pseudogroupoid structure becomes a property. Many other results follow from this simple axiom, and some of them improve previously known ones in the special case of modular varieties. Accordingly, this axiomatic approach allows one not only to cover a wider range of examples, but also to sharpen the results. Under many respects, the categories which satisfy the shifting property are in the same relationship with modular varieties, as Maltsev categories are with Maltsev varieties.

The paper is structured as follows:

1. Pseudogroupoids
2. The Shifting Property
3. Connectors and Abelian Objects

4. Internal Categories
5. Internal Groupoids

In the first section we recall the definition of a pseudogroupoid, and we show that the "diamond associativity" condition in its definition can be replaced by two natural and simpler conditions. In the second section we introduce the shifting property, we give some examples, and we then establish our main results on pseudogroupoids in a category satisfying the shifting property. In the third section we apply the previous results to study internal connectors [4] [5] and abelian objects in these categories. In the last two sections several characterizations of the internal categories and of the internal groupoids are given. In particular, the above-mentioned characterization of internal groupoids in modular varieties is extended to any regular category satisfying the shifting property.

Acknowledgement: The authors would like to thank M.C. Pedicchio for having drawn their attention to Gumm's work.

1 Pseudogroupoids

In this section we recall the definition of an internal pseudogroupoid introduced by Janelidze and Pedicchio in [16]. We then show that the "diamond associativity" requirement in its definition is equivalent to two simpler and natural axioms.

Let us first fix some notations. $\mathcal{C}$ will always denote a category with finite limits. Thanks to the Yoneda embedding, in order to prove finite limit properties in $\mathcal{C}$ it is sufficient to do it "elementwise" in the category of sets. We shall freely use this well-known technique through the whole paper.

For any object X in $\mathcal{C}$, we write Δ_X for the smallest equivalence relation on X, and ∇_X for the largest equivalence relation on X. For any arrow $f\colon A \to B$ let $R[f]$ denote its *kernel relation*, that is the equivalence relation arising as its kernel pair. If R and S are equivalence relations on X, we write $R\Box S$ for the double equivalence relation on R and S obtained by the following pullback:

$$\begin{tikzcd}
R\Box S \arrow[r] \arrow[d] & R\times R \arrow[d, "{[d_0\times d_0, d_1\times d_1]}"] \\
S\times S \arrow[r, "{[d_0,d_1]\times[d_0,d_1]}"'] & X\times X\times X\times X
\end{tikzcd}$$

where d_0 and d_1 represent, respectively, the first and the second projection of the equivalence relation R or S. An element in $R\Box S$ is called an R-S rectangle: it consists of four elements x, y, t, z in X with the property that xRy, tRz, xSt and ySz. An R-S rectangle will be represented equivalently by a diagram of the form

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & t \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] & z,
\end{tikzcd}$$

by a matrix $\begin{pmatrix} x & t \\ y & z \end{pmatrix}$, or by (x, y, t, z). This last "linear" notation for an element in $R\square S$ explains the following notations for the projections of $R\square S$ on R and S:

$$\begin{tikzcd} R\square S \arrow[r, shift left, "\pi_{1,3}"] \arrow[r, shift right, "\pi_{2,4}"'] \arrow[d, shift right, "\pi_{1,2}"'] \arrow[d, shift left, "\pi_{3,4}"] & S \arrow[d, shift right, "d_0"'] \arrow[d, shift left, "d_1"] \\ R \arrow[r, shift left, "d_0"] \arrow[r, shift right, "d_1"'] & X. \end{tikzcd}$$

Let us then recall the definition of a pseudogroupoid [16]:

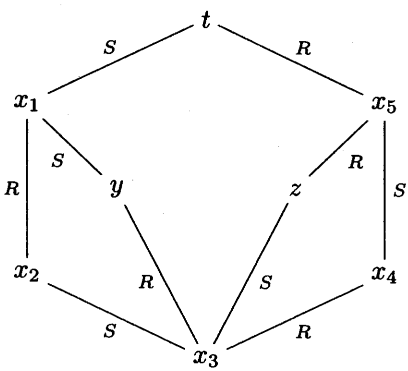
1.1. Definition. A *pseudogroupoid* on R and S is an arrow $m\colon R\square S \to X$ in $\mathcal{C}$, written as

$$m\begin{pmatrix} x & t \\ y & z \end{pmatrix} = m(x, y, t, z),$$

with the following properties:
1. $xSm(x, y, t, z)Rz$

$$\begin{tikzcd} x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & m(x, y, t, z) \arrow[d, dash, "R"] \\ y \arrow[r, dash, "S"'] & z \end{tikzcd}$$

2. $m(x, y, t, z) = m(x, y, t', z)$ (i.e. m does not depend on the third variable)
3A. $m(x, x, t, y) = y$ 3B. $m(x, y, t, y) = x$
4. $m(m(x_1, x_2, y, x_3), x_4, t, x_5) = m(x_1, x_2, t, m(x_3, x_4, z, x_5))$ for every diagram of the following form:



1.2. Examples. If R and S are two equivalence relations on X, we denote by $R \times_X S$ the pullback

$$\begin{tikzcd} R \times_X S \arrow[r, "p_1"] \arrow[d, "p_0"'] & S \arrow[d, "d_0"] \\ R \arrow[r, "d_1"'] & X. \end{tikzcd}$$

A *connector* between R and S [4] (see also [20] [21]) is an arrow $p : R \times_X S \to X$ in $\mathcal{C}$ such that

$I.\ xSp(x,y,z)Rz$

$IIA.\ p(x,x,y) = y$ $\qquad\qquad$ $IIB.\ p(x,y,y) = x$

$IIIA.\ p(x,y,p(y,u,v)) = p(x,u,v)$ $\qquad$ $IIIB.\ p(p(x,y,u),u,v) = p(x,y,v)$

Any connector p determines a pseudogroupoid: for any (x,y,t,z) in $R\Box S$ one defines $m(x,y,t,z) = p(x,y,z)$.

1.3. Examples. Let X be an *internal groupoid* in $\mathcal{C}$, represented by the diagram

$$X_1 \times_{X_0} X_1 \underset{p_0}{\overset{p_1}{\underset{m}{\rightrightarrows}}} X_1 \underset{d_1}{\overset{d_0}{\underset{s_0}{\leftrightarrows}}} X_0$$

where X_0 represents the "object of objects", X_1 is the "object of arrows", $X_1 \times_{X_0} X_1$ is the "object of composable arrows", d_0 is the domain, d_1 is the codomain and m is the groupoid composition. It determines a (unique) connector, and then a pseudogroupoid, on the kernel relations $R[d_0]$ and $R[d_1]$: the connector structure is internally defined by $p(x,y,z) = z \circ y^{-1} \circ x$. Conversely, a connector on $R[d_0]$ and $R[d_1]$ determines a groupoid structure on the underlying reflexive graph.

1.4. Examples. If R and S are two equivalence relations on X with the property that $R \cap S = \Delta_X$, then R and S have a unique pseudogroupoid structure [16]: for any (x,y,t,z) in $R\Box S$ this structure is given by $m(x,y,t,z) = t$. Indeed, the condition $R \cap S = \Delta_X$ guarantees the uniqueness of such a t, when it exists.

1.5. Proposition. *An arrow* $m\colon R\Box S \to X$ *is a pseudogroupoid on* R *and* S *if and only if it satisfies the axioms* $1, 2, 3A$, $3B$ *and the following axioms* $4A$ *and* $4B$*:*

$4A.\ m(x,y,m(x,y,x',y'),m(y,z,y',z')) = m(x,z,x',z')$ *for every diagram*

$$\begin{tikzcd} x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & x' \arrow[d, dash, "R"] \\ y \arrow[r, dash, "S"'] \arrow[d, dash, "R"'] & y' \arrow[d, dash, "R"] \\ z \arrow[r, dash, "S"] & z' \end{tikzcd}$$

$4B.\ m(m(x,y,x',y'),y',m(x',y',x'',y''),y'') = m(x,y,x'',y'')$ *for every diagram*

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & x' \arrow[r, dash, "S"] \arrow[d, dash, "R"] & x'' \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] & y' \arrow[r, dash, "S"'] & y''
\end{tikzcd}$$

Proof Let us begin by proving that the axioms $1, 2, 3A, 3B$ and 4 implies $4A$ and $4B$. We consider the situation

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & x' \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] \arrow[d, dash, "R"'] & y' \arrow[d, dash, "R"] \\
z \arrow[r, dash, "S"] & z'.
\end{tikzcd}$$

Then, thanks to axiom 1 we can form the diagram

$$\begin{tikzcd}
 & & m(x,y,x',y') \arrow[lld, dash, "S"'] \arrow[rrd, dash, "R"] & & \\
x \arrow[dd, dash, "R"'] \arrow[dr, dash, "S"] & & & & z' \arrow[dl, dash, "R"'] \arrow[dd, dash, "S"] \\
 & x \arrow[dr, dash, "R"] & & y' \arrow[dl, dash, "S"'] & \\
y \arrow[rrd, dash, "S"'] & & & & z \arrow[lld, dash, "R"] \\
 & & y & &
\end{tikzcd}$$

and then

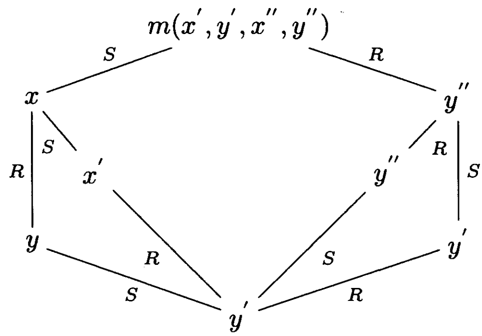
$$\begin{aligned}
& m(x,y,m(x,y,x',y'),m(y,z,y',z')) \\
= \;& m(m(x,y,x,y),z,m(x,y,x',y'),z') \\
= \;& m(x,z,m(x,y,x',y'),z') \\
= \;& m(x,z,x',z').
\end{aligned}$$

where the first equality follows from axiom 4, the second one from axiom 3B and the last one from axiom 2.

Similarly, if we start from the situation

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & x' \arrow[r, dash, "S"] \arrow[d, dash, "R"] & x'' \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] & y' \arrow[r, dash, "S"'] & y'',
\end{tikzcd}$$

we can form the diagram

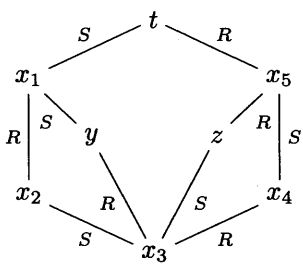


Then

$$\begin{aligned}
& m(m(x,y,x',y'),y',m(x',y',x'',y''),y'') \\
= \; & m(x,y,m(x',y',x'',y''),m(y',y',y'',y'')) \\
= \; & m(x,y,m(x',y',x'',y''),y'') \\
= \; & m(x,y,x'',y''),
\end{aligned}$$

by axioms 4, 3A and 2.

Conversely, if $m\colon R\Box S \to X$ satisfies $1, 2, 3A, 3B, 4A$ and $4B$, let us then prove that m also satisfies axiom 4. We begin with the diagram



Axiom 1 allows us to form the diagram

$$\begin{tikzcd}
x_1 \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & y \arrow[r, dash, "S"] \arrow[d, dash, "R"] & t \arrow[d, dash, "R"] \\
x_2 \arrow[r, dash, "S"'] & x_3 \arrow[r, dash, "S"'] & m(x_3, x_4, z, x_5).
\end{tikzcd}$$

Thanks to axiom $4B$ and to the fact that m does not depend on the third variable we get

$$\begin{aligned}
& m(x_1, x_2, t, m(x_3, x_4, z, x_5)) \\
= \; & m(m(x_1, x_2, y, x_3), x_3, m(y, x_3, t, m(x_3, x_4, z, x_5)), m(x_3, x_4, z, x_5)) \\
= \; & m(m(x_1, x_2, y, x_3), x_3, m(m(x_1, x_2, y, x_3), x_3, t, z), m(x_3, x_4, z, x_5)) \\
= \; & m(m(x_1, x_2, y, x_3), x_4, t, x_5),
\end{aligned}$$

where the last equality comes from axiom $4A$ applied to

$$\begin{tikzcd} m(x_1,x_2,y,x_3) \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & t \arrow[d, dash, "R"] \\ x_3 \arrow[r, dash, "S"'] \arrow[d, dash, "R"'] & z \arrow[d, dash, "R"] \\ x_4 \arrow[r, dash, "S"'] & x_5. \end{tikzcd}$$

□

2 The shifting property

In this section we introduce the main axiom of this paper, namely the shifting property. Under this rather weak assumption we are going to show that a pseudogroupoid structure on two equivalence relations is necessarily unique, when it exists. Consequently, in this context, having a pseudogroupoid structure for two equivalence relations becomes a property.

Let $Eq_X(\mathcal{C})$ be the poset of equivalence relations in $\mathcal{C}$ on a fixed object X.

Let us recall the well-known *Shifting Lemma* due to Gumm [13]: a variety $\mathcal{V}$ of universal algebras satisfies the Shifting Lemma if for any R, S, $T \in Eq_X(\mathcal{V})$ with $R \cap S \leq T$ the situation

$$\begin{tikzcd} x \arrow[r, dash, "S"] \arrow[d, dash, "R"] \arrow[d, dash, bend right=80, "T"'] & t \arrow[d, dash, "R"] \\ y \arrow[r, dash, "S"'] & z, \end{tikzcd}$$

implies that tTz.

A remarkable fact concerning the Shifting Lemma is that it characterizes congruence modular varieties, namely the varieties with the property that the lattice $Eq_X(\mathcal{V})$ is modular for any algebra X in $\mathcal{V}$.

2.1. Theorem. [13] *A variety $\mathcal{V}$ of universal algebras is congruence modular if and only if $\mathcal{V}$ satisfies the Shifting Lemma.*

In any finitely complete category the validity of the Shifting Lemma is equivalent to the following categorical property:

2.2. Lemma. *Let $\mathcal{C}$ be a finitely complete category. Then $\mathcal{C}$ satisfies the Shifting Lemma if and only if for any X in $\mathcal{C}$, for any R, S, $U \in Eq_X(\mathcal{C})$ with $R \cap S \leq U \leq R$ the canonical inclusion of equivalence relations*

$$\begin{tikzcd} U\square S \arrow[r, tail, "j"] \arrow[d, shift right, "d_0"'] \arrow[d, shift left, "d_1"] \arrow[dr, phantom, "(1)"] & R\square S \arrow[d, shift right, "d_0"'] \arrow[d, shift left, "d_1"] \\ U \arrow[r, tail, "i"'] & R \end{tikzcd}$$

is a discrete fibration.

Proof Let us first recall that an internal functor (1) as above is a discrete fibration when the commutative square involving the arrows d_1 (or, equivalently, the arrows d_0) is a pullback.

If we assume that the property here above holds, one can define $R \cap S = U$ in the assumption of the Shifting Lemma and one immediately concludes that tUz, i.e. $t(R \cap S)z$. Conversely, when the Shifting Lemma holds and $R \cap S \leq U \leq R$ the diagram

$$\begin{tikzcd} x \arrow[r, dash, "S"] \arrow[d, dash, "U"'] & t \arrow[d, dash, "R"] \\ y \arrow[r, dash, "S"'] & z \end{tikzcd}$$

allows one to form the diagram

$$\begin{tikzcd} x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] \arrow[d, dash, bend right=60, "U"'] & t \arrow[d, dash, "R"] \\ y \arrow[r, dash, "S"'] & z, \end{tikzcd}$$

so that the assumption $R \cap S \leq U$ yields tUz, proving that the internal functor (1) is a discrete fibration. □

It follows from the previous result that the Shifting Lemma is a finite limit statement:

2.3. Definition. A finitely complete category $\mathcal{C}$ satisfies the *shifting property* when it satisfies any of the two equivalent conditions of the previous Lemma.

2.4. Examples.

1. *Congruence modular varieties*: in particular any Maltsev variety [22] (groups, rings, associative algebras, Lie algebras, quasi-groups, Heyting algebras) and any distributive variety (lattices, right complemented semigroups) satisfy the shifting property.

2. Any *regular Maltsev category* satisfies the shifting property, as well as, more generally, any regular Goursat category. This fact follows easily from Proposition 3.2 in [7]. In particular the categories of topological groups, Hausdorff groups, torsion-free abelian groups, or the dual category of an elementary topos satisfy the shifting property.

Since the *shifting property* is a finite limit statement, any functor category $\mathcal{C}^{\mathcal{A}}$ of functors from any small category $\mathcal{A}$ to $\mathcal{C}$ satisfies the shifting property, provided $\mathcal{C}$ satisfies the shifting property. A useful criterion to check whether a category satisfies the shifting property is given by the following

2.5. Lemma. *Let $F\colon \mathcal{C} \to \mathcal{D}$ be a finite limit preserving functor between finitely complete categories. If F is conservative (i.e. it reflects isomorphisms) and $\mathcal{D}$ satisfies the shifting property, then $\mathcal{C}$ satisfies the shifting property.*

Proof For any X in $\mathcal{C}$, for any R, S, $U \in Eq_X(\mathcal{C})$ with $R \cap S \leq U \leq R$ the internal functor

$$\begin{tikzcd} U\Box S \arrow[r, tail, "j"] \arrow[d, shift right, "d_0"'] \arrow[d, shift left, "d_1"] & R\Box S \arrow[d, shift right, "d_0"'] \arrow[d, shift left, "d_1"] \\ U \arrow[r, tail, "i"'] & R \end{tikzcd}$$

is sent to the internal functor

$$\begin{tikzcd} F(U)\Box F(S) \arrow[r, tail, "F(j)"] \arrow[d, shift right, "F(d_0)"'] \arrow[d, shift left, "F(d_1)"] & F(R)\Box F(S) \arrow[d, shift right, "F(d_0)"'] \arrow[d, shift left, "F(d_1)"] \\ F(U) \arrow[r, tail, "F(i)"'] & F(R), \end{tikzcd}$$

which is necessarily a discrete fibration because $\mathcal{D}$ satisfies the shifting property. The functor F also reflects pullbacks, and then the first internal functor (the one in $\mathcal{C}$) is a discrete fibration as well. □

The previous result easily gives some more examples of categories satisfying the shifting property: if $\mathcal{C}$ satisfies the shifting property, so do the categories $Cat(\mathcal{C})$ of internal categories in $\mathcal{C}$ and the category $Grpd(\mathcal{C})$ of internal groupoids in $\mathcal{C}$. Indeed, if one considers the forgetful functor associating its object of morphisms with any internal category or groupoid, one easily sees that it is conservative and it preserves finite limits.

On the other hand, if $\mathcal{C}$ satisfies the shifting property, so does the comma category $\mathcal{C}/Y$ for any object Y in $\mathcal{C}$. This follows from the fact that if R and S are two equivalence relations on an object $f: X \to Y$ in $\mathcal{C}/Y$, then the object $R\Box S$ in the category $\mathcal{C}/Y$ is the same as $R\Box S$ in the category $\mathcal{C}$. Accordingly, the argument given in the proof of Lemma 2.5 and the fact that the forgetful functor from $\mathcal{C}/Y$ to $\mathcal{C}$ is conservative and preserves pullbacks allows one to conclude.

As we shall see in Proposition 2.12, the notion of pseudogroupoid can be simplified in any category satisfying the shifting property. In order to present this simplification, we first need to introduce the following definitions:

2.6. Definition. A *double zero sequence* in a pointed category $\mathcal{C}$ is a diagram of the form

$$\begin{tikzcd} X \arrow[r, shift right, "s"'] & Z \arrow[l, shift right, "f"'] \arrow[r, shift right, "g"'] & Y \arrow[l, shift right, "t"'] \end{tikzcd}$$

with $f \circ s = 1_X, g \circ t = 1_Y, g \circ s = 0, f \circ t = 0$ (where 0 is the zero arrow).

Let us then define the category $DZS(X,Y)$ of double zero sequences between X and Y, with arrows $\alpha: Z \to Z'$ making the following diagram commutative:

$$\begin{tikzcd}
X \arrow[r, shift right, "s"'] \arrow[d, equal] & Z \arrow[l, shift right, "f"'] \arrow[r, shift right, "g"'] \arrow[d, dotted, "\alpha"] & Y \arrow[l, shift right, "t"'] \arrow[d, equal] \\
X \arrow[r, shift right, "s'"'] & Z' \arrow[l, shift right, "f'"'] \arrow[r, shift right, "g'"'] & Y. \arrow[l, shift right, "t'"']
\end{tikzcd}$$

If $\mathcal{C}$ is any finitely complete category (not necessarily pointed) we denote by $Pt(\mathcal{C})$ the category whose objects (X, p, s) are the split epimorphisms $p\colon X \to Y$ in $\mathcal{C}$ with a given splitting $s\colon Y \to X$, with arrows the commutative squares between these data

$$\begin{tikzcd}
X \arrow[r, "f_1"] \arrow[d, shift right, "p"'] & X' \arrow[d, shift right, "q"'] \\
Y \arrow[r, "f_0"'] \arrow[u, shift right, "s"'] & Y' \arrow[u, shift right, "t"']
\end{tikzcd}$$

where $f_0 \circ p = q \circ f_1$ and $f_1 \circ s = t \circ f_0$.

The functor associating its codomain with any split epimorphism is denoted by $\pi\colon Pt(\mathcal{C}) \to \mathcal{C}$. This functor π is a fibration, called the *fibration of pointed objects* [3]. For any X in $\mathcal{C}$, we write $Pt_X(\mathcal{C})$ for the fiber above X, which is clearly pointed (by the trivial split epi $1_X\colon X \to X$).

Given R and S in $Eq_X(\mathcal{C})$, there is a canonical comparison arrow $\alpha\colon R\Box S \to R \times_X S$ induced by the universal property of the pullback

$$\begin{tikzcd}
R \times_X S \arrow[r, "p_1"] \arrow[d, "p_0"'] & S \arrow[d, "d_0"] \\
R \arrow[r, "d_1"'] & X,
\end{tikzcd}$$

which associates the element (x, y, z) in $R \times_X S$ with any element (x, y, t, z) in $R\Box S$. This arrow α is an arrow of double zero sequences in $DZS((R, d_1, s_0), (S, d_0, s_0))$ in the pointed category $Pt_X(\mathcal{C})$, where we write s_0 for the arrows giving the reflexivity of the relations R or S. The arrow α actually is the terminal arrow in this category:

$$\begin{tikzcd}
R \arrow[r, shift right, "s_0"'] \arrow[d, equal] & R\Box S \arrow[l, shift right, "\pi_{1,2}"'] \arrow[r, shift right, "\pi_{2,4}"'] \arrow[d, dotted, "\alpha"] & S \arrow[l, shift right, "s_0"'] \arrow[d, equal] \\
R \arrow[r, shift right, "s_0"'] & R \times_X S \arrow[l, shift right, "p_1"'] \arrow[r, shift right, "p_0"'] & S \arrow[l, shift right, "s_0"']
\end{tikzcd}$$

2.7. Definition. A *quasiconnector* on R and S is an arrow $\sigma\colon R\square S \to R\square S$ of double zero sequences in $Pt_X(\mathcal{C})$

$$
\begin{tikzcd}
R \arrow[r, shift right, "s_0"'] \arrow[d, equal] & R\square S \arrow[l, shift right, "\pi_{1,2}"'] \arrow[r, shift right, "\pi_{2,4}"'] \arrow[d, dotted, "\sigma"] & S \arrow[l, shift right, "s_0"'] \arrow[d, equal] \\
R \arrow[r, shift right, "s_0"'] & R\square S \arrow[l, shift right, "\pi_{1,2}"'] \arrow[r, shift right, "\pi_{2,4}"'] & S. \arrow[l, shift right, "s_0"']
\end{tikzcd}
$$

such that the kernel relation $R[\sigma]$ is $R[\pi_{1,2}] \cap R[\pi_{2,4}] = R[\alpha]$ (where $\alpha\colon R\square S \to R \times_X S$ is the canonical terminal arrow defined above).

2.8. Remark. The axiom $R[\sigma] = R[\alpha]$ exactly says that σ does not depend on the third variable.

A quasiconnector $\sigma\colon R\square S \to R\square S$ is always idempotent: indeed, this follows from the following simple

2.9. Lemma. *Let $\sigma\colon Z \to Z$ be an arrow in a category with pullbacks $\mathcal{C}$ such that $R[\sigma] = \nabla_Z$ (where $\nabla_Z = R[\tau_Z]$ and $\tau_Z\colon Z \to 1$ is the terminal arrow). Then $\sigma \circ \sigma = \sigma$.*

Proof Consider the kernel relation of the terminal arrow:

$$R[\sigma] = Z \times Z \underset{p_1}{\overset{p_0}{\rightrightarrows}} Z \longrightarrow 1.$$

The arrow $\sigma\colon Z \to Z$ induces the arrow $(\sigma, 1_Z)\colon Z \to Z \times Z$. Then the equality $\sigma \circ p_0 = \sigma \circ p_1$ gives

$$\sigma \circ p_0 \circ (\sigma, 1_Z) = \sigma \circ p_1 \circ (\sigma, 1_Z).$$

It follows that

$$\sigma \circ \sigma = \sigma \circ 1_Z = \sigma.$$

$\square$

2.10. Corollary. *Any quasiconnector satisfies $\sigma \circ \sigma = \sigma$.*

Proof One just needs to notice that in the category $DZS((R, d_1, s_0), (S, d_0, s_0))$ of double zero sequences between R and S in $Pt_X(\mathcal{C})$, the diagram

$$
\begin{tikzcd}
R \arrow[r, shift right, "s_0"'] & R \times_X S \arrow[l, shift right, "p_0"'] \arrow[r, shift right, "p_1"'] & S \arrow[l, shift right, "s_0"']
\end{tikzcd}
$$

is the terminal object, and the arrow $\alpha\colon R\square S \to R \times_X S$ is the terminal arrow. Then the requirement $R[\sigma] = R[\alpha]$ allows one to apply the previous Lemma to σ. $\square$

2.11. Remark. Given a quasiconnector σ, the arrow $m = d_1 \circ \pi_{1,3} \circ \sigma \colon R\Box S \to R\Box S \to S \to X$ defines an arrow $m\colon R\Box S \to X$ satisfying the axioms 1 and 2 in the definition of pseudogroupoid, and is such that

3C $\quad m(x, x, y, y) = y$

3D $\quad m(x, y, x, y) = x$.

Conversely, an arrow $m\colon R\Box S \to X$ satisfying axioms $1, 2, 3C$ and $3D$ gives rise to a quasiconnector. Accordingly, we shall refer equivalently to σ or to m when speaking of a quasiconnector.

Now, when the shifting property holds in $\mathcal{C}$, the notion of quasiconnector is equivalent to the one of pseudogroupoid:

2.12. Proposition. *If $\mathcal{C}$ is finitely complete and satisfies the shifting property, then any quasiconnector is a pseudogroupoid.*

Proof Clearly the axioms 2, $3C$ and $3D$ give

$$m(x, x, t, y) = m(x, x, y, y) = y$$

and

$$m(x, y, t, y) = m(x, y, x, y) = x,$$

so that axioms $3A$ and $3B$ always hold.

Let us then prove axiom $4B$. We write $\pi_{1,2}$ also for the kernel relation of the projection $\pi_{1,2}\colon R\Box S \to R$ and we write π_4 for the kernel relation of the projection $\pi_4\colon R\Box S \to X$ sending an element $\begin{pmatrix} x & t \\ y & z \end{pmatrix}$ in $R\Box S$ to z in X. The fact that $R[\sigma] = R[\alpha]$, so that $m\colon R\Box S \to X$ is independent of the third variable, implies that the kernel relation of m, also denoted by m, contains $R[\alpha] = \pi_{1,2} \cap \pi_4$. This remark allows us to apply the shifting property to the diagram

$$\begin{tikzcd}
\begin{pmatrix} m(x,y,x',y') & m(x,y,x',y') \\ y' & y' \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"] \arrow[d, dash, bend right, "m"'] \arrow[d, dash, "\pi_4"] & \begin{pmatrix} m(x,y,x',y') & m(x',y',x'',y'') \\ y' & y'' \end{pmatrix} \arrow[d, dash, "\pi_4"] \\
\begin{pmatrix} x & x' \\ y & y' \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"'] & \begin{pmatrix} x & x'' \\ y & y'' \end{pmatrix}
\end{tikzcd}$$

for any situation as in the following diagram

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & x' \arrow[r, dash, "S"] \arrow[d, dash, "R"] & x'' \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] & y' \arrow[r, dash, "S"'] & y''.
\end{tikzcd}$$

This implies that $m(m(x, y, x', y'), y', m(x', y', x'', y''), y'') = m(x, y, x'', y'')$ (axiom $4B$). Similarly one can prove axiom $4A$. $\square$

Given a finitely complete category $\mathcal{C}$, we write $2\text{-}Eq(\mathcal{C})$ for the category whose objects are pairs of equivalence relations (R, S, X) on the same object X:

$$
\begin{tikzcd}
R \arrow[r, shift left, "d_1"] \arrow[r, shift right, "d_0"'] & X & S, \arrow[l, shift right, "d_1"'] \arrow[l, shift left, "d_0"]
\end{tikzcd}
$$

and arrows in $2\text{-}Eq(\mathcal{C})$ are triples of arrows (f_R, f_S, f) making the following diagram commutative:

$$
\begin{tikzcd}
R \arrow[r, shift left, "d_1"] \arrow[r, shift right, "d_0"'] \arrow[d, "f_R"'] & X \arrow[d, "f"] & S \arrow[l, shift right, "d_1"'] \arrow[l, shift left, "d_0"] \arrow[d, "f_S"] \\
\overline{R} \arrow[r, shift left, "d_1"] \arrow[r, shift right, "d_0"'] & Y & \overline{S}. \arrow[l, shift right, "d_1"'] \arrow[l, shift left, "d_0"]
\end{tikzcd}
$$

Let $Psdgrd(\mathcal{C})$ be the category whose objects are pairs of equivalence relations on a given object (R, S, X, m) equipped with a pseudogroupoid structure $m\colon R\Box S \to X$. An arrow $(f_R, f_S, f)\colon (R, S, X, m) \to (\overline{R}, \overline{S}, Y, \mu)$ in $Psdgrd(\mathcal{C})$ is an arrow in $2\text{-}Eq(\mathcal{C})$ that preserves the pseudogroupoid structure, i.e. such that the diagram

$$
\begin{tikzcd}
R\Box S \arrow[r, "\tilde{f}"] \arrow[d, "m"'] & \overline{R}\Box\overline{S} \arrow[d, "\mu"] \\
X \arrow[r, "f"'] & Y
\end{tikzcd}
$$

commutes. There is a forgetful functor $U\colon Psdgrd(\mathcal{C}) \to 2\text{-}Eq(\mathcal{C})$ defined on objects by $U(R, S, X, m) = (R, S, X)$. This functor U is clearly faithful.

A remarkable consequence of the shifting property is that it forces a pseudogroupoid structure to be unique, when it exists. Moreover, the forgetful functor $U\colon Psdgrd(\mathcal{C}) \to 2\text{-}Eq(\mathcal{C})$ is full:

2.13. Theorem. *Let $\mathcal{C}$ be a finitely complete category in which the shifting property holds. Then*

1. *a pseudogroupoid structure on two equivalence relations is unique (when it exists)*
2. *the forgetful functor $U\colon Psdgrd(\mathcal{C}) \to 2\text{-}Eq(\mathcal{C})$ is a full and faithful inclusion*

Proof Let us first show that any pseudogroupoid $m\colon R\Box S \to X$ satisfies the implication

$$m(x, y, t, z) = m(x, y, t', z') \quad \Rightarrow \quad z = z'.$$

We know that $\pi_{1,2} \cap \pi_4 \leq m$, hence the shifting property applied to the situation

$$
\begin{tikzcd}
\begin{pmatrix} x & t \\ y & z \end{pmatrix} \arrow[r, dash, "\pi_4"] \arrow[d, dash, "\pi_{1,2}"] \arrow[d, dash, bend right=60, "m"'] & \begin{pmatrix} y & z \\ y & z \end{pmatrix} \arrow[d, dash, "\pi_{1,2}"] \\
\begin{pmatrix} x & t' \\ y & z' \end{pmatrix} \arrow[r, dash, "\pi_4"'] & \begin{pmatrix} y & z' \\ y & z' \end{pmatrix}
\end{tikzcd}
$$

shows that

$$z = m(y,y,z,z) = m(y,y,z',z') = z'.$$

This means that $m \cap \pi_{1,2} \leq \pi_4$.

Now, let $m\colon R\Box S \to X$ and $m'\colon R\Box S \to X$ be two pseudogroupoid structures on R and S. For any $\begin{pmatrix} x & t \\ y & z \end{pmatrix}$ in $R\Box S$, the shifting property applied to

$$\begin{tikzcd}
\begin{pmatrix} x & x \\ y & y \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"] \arrow[d, dash, "m"'] \arrow[d, dash, bend right=60, "m'"'] & \begin{pmatrix} x & t \\ y & z \end{pmatrix} \arrow[d, dash, "m"] \\
\begin{pmatrix} x & x \\ x & x \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"'] & \begin{pmatrix} x & m(x,y,t,z) \\ x & m(x,y,t,z) \end{pmatrix}
\end{tikzcd}$$

yields

$$m'(x,y,t,z) = m'(x,x,m(x,y,t,z),m(x,y,t,z)) = m(x,y,t,z).$$

Of course, the shifting property can be applied because $m\cap\pi_{1,2} = m\cap\pi_{1,2}\cap\pi_4 \leq m'$.

Let us then prove that the inclusion $U\colon Psdgrd(\mathcal{C}) \to 2\text{-}Eq(\mathcal{C})$ is full. Let $\tilde{f}\colon R\Box S \to \overline{R}\Box\overline{S}$ be the arrow induced by $(f, f_R, f_S)\colon (R,S,X,m) \to (\overline{R},\overline{S},Y,\mu)$. Consider any $\begin{pmatrix} x & t \\ y & z \end{pmatrix}$ in $R\Box S$ and the following situation

$$\begin{tikzcd}
\begin{pmatrix} x & x \\ y & y \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"] \arrow[d, dash, "m"'] \arrow[d, dash, bend right=60, "\mu\circ\tilde{f}"'] & \begin{pmatrix} x & t \\ y & z \end{pmatrix} \arrow[d, dash, "m"] \\
\begin{pmatrix} x & x \\ x & x \end{pmatrix} \arrow[r, dash, "\pi_{1,2}"'] & \begin{pmatrix} x & m(x,y,t,z) \\ x & m(x,y,t,z) \end{pmatrix}
\end{tikzcd}$$

which clearly holds since we have:

$$(\mu\circ\tilde{f})\begin{pmatrix} x & x \\ y & y \end{pmatrix} = \mu\begin{pmatrix} f(x) & f(x) \\ f(y) & f(y) \end{pmatrix} = f(x) = (\mu\circ\tilde{f})\begin{pmatrix} x & x \\ x & x \end{pmatrix}.$$

Moreover,

$$m \cap \pi_{1,2} = m \cap \pi_{1,2} \cap \pi_4 \leq \pi_{1,2} \cap \pi_4 \leq \mu\circ\tilde{f},$$

so that the shifting property gives

$$(\mu\circ\tilde{f})\begin{pmatrix} x & t \\ y & z \end{pmatrix} = (\mu\circ\tilde{f})\begin{pmatrix} x & m(x,y,t,z) \\ x & m(x,y,t,z) \end{pmatrix} = \mu\begin{pmatrix} f(x) & (f\circ m)(x,y,t,z) \\ f(x) & (f\circ m)(x,y,t,z) \end{pmatrix}$$

and then

$$(\mu\circ\tilde{f})\begin{pmatrix} x & t \\ y & z \end{pmatrix} = (f\circ m)\begin{pmatrix} x & t \\ y & z \end{pmatrix}.$$

$\square$

3 Connectors and Abelian objects

In this section we apply the previous results to the category of connectors ([4] [5]) and to the category of abelian objects in a category satisfying the shifting property. In particular it turns out that, for a given objet in such a category, being abelian becomes a property.

We begin with the following

3.1. Corollary. *Let R and S be two equivalence relations on X in a finitely complete category satisfying the shifting property. Then any arrow $p\colon R \times_X S \to X$ satisfying the axioms*
I $\quad xSp(x,y,z)Rz$
IIA $\quad p(x,x,y) = y$
IIB $\quad p(x,y,y) = x$
is a connector between R and S. Moreover, such a structure is necessarily unique.

Proof We are going to check only the axiom $IIIB$

$$p(p(x,y,u),u,v) = p(x,y,v)$$

the verification of the axiom $IIIA$ being similar.

Clearly, any arrow p satisfying the axioms I, IIA and IIB is a quasiconnector, by defining $m\colon R\Box S \to X$ as $m(x,y,t,z) = p(x,y,z)$, which is then a pseudogroupoid, thanks to Proposition 2.12. Now, for any x,y,u,v with $xRySuSv$ one forms the diagram

$$\begin{tikzcd}
x \arrow[r, dash, "S"] \arrow[d, dash, "R"'] & p(x,y,u) \arrow[r, dash, "S"] \arrow[d, dash, "R"] & p(p(x,y,u),u,v) \arrow[d, dash, "R"] \\
y \arrow[r, dash, "S"'] & u \arrow[r, dash, "S"'] & v.
\end{tikzcd}$$

Then the validity of axiom $4B$ for the pseudogroupoid structure gives:

$$\begin{aligned}
p(p(x,y,u),u,v) &= m(m(x,y,p(x,y,u),u),u,m(p(x,y,u),u,p(p(x,y,u),u,v),v),v) \\
&= m(x,y,p(p(x,y,u),u,v),v) \\
&= p(x,y,v).
\end{aligned}$$

The uniqueness of the structure follows from the uniqueness of a pseudogroupoid structure in $\mathcal{C}$.

□

Let $Conn(\mathcal{C})$ be the category (R,S,X,p) of pairs of equivalence relations on a given object equipped with a connector, with arrows $(f_R, f_S, f)\colon (R,S,X,p) \to (\overline{R},\overline{S},Y,\overline{p})$ those arrows in 2-$Eq(\mathcal{C})$ that preserve the connector, i.e. such that the diagram

$$\begin{tikzcd}
R \times_X S \arrow[r, "\tilde{f}"] \arrow[d, "p"'] & \overline{R} \times_Y \overline{S} \arrow[d, "\overline{p}"] \\
X \arrow[r, "f"'] & Y
\end{tikzcd}$$

commutes. Any connector being a pseudogroupoid, it follows from Theorem 2.13 that:

3.2. Corollary. *When C satisfies the shifting property, then the forgetful functor $V\colon Conn(C) \to$ 2-$Eq(C)$ is a full and faithful inclusion.*

It is worth mentioning that this forgetful functor V has been used to characterize Maltsev categories. As shown in [5], a finitely complete category C is Maltsev exactly when V is closed under subobjects: this means that given any monomorphism $i\colon X \to Y$ in 2-$Eq(C)$, if Y is in $Conn(C)$, then also X is in $Conn(C)$.

3.3. Definition. An object X in C is *abelian* if it is endowed with an internal Maltsev operation, i.e. there is an arrow $p\colon X \times X \times X \to X$ with $p(x,x,y) = y$ and $p(x,y,y) = x$.

In the presence of the shifting property being abelian becomes a property:

3.4. Corollary. *Any internal Maltsev operation $p\colon X \times X \times X \to X$ in a finitely complete category satisfying the shifting property is associative and commutative. This operation is necessarily unique, when it exists.*

Proof The classical associativity $p(x,y,p(z,u,v)) = p(p(x,y,z),u,v)$ follows from Corollary 3.1:

$$p(x,y,p(z,u,v)) = p(p(x,y,z),z,p(z,u,v)) = p(p(x,y,z),u,v).$$

As far as commutativity is concerned, let us write $\pi_2\colon X \times X \times X \to X$ for the projection sending the element (x,y,z) to y, and $\pi_{1,3}\colon X \times X \times X \to X \times X$ for the projection sending the element (x,y,z) to (x,z). Clearly $\Delta_{X \times X \times X} = \pi_{1,3} \cap \pi_2 \leq p$, and the shifting property applied to the diagram

$$\begin{tikzcd}
(x,x,z) \arrow[r, dash, "\pi_{1,3}"] \arrow[d, dash, "\pi_2"] \arrow[d, dash, bend right=60, "p"'] & (x,y,z) \arrow[d, dash, "\pi_2"] \\
(z,x,x) \arrow[r, dash, "\pi_{1,3}"'] & (z,y,x),
\end{tikzcd}$$

gives $p(x,y,z) = p(z,y,x)$, as desired. The uniqueness of p follows from Theorem 2.13. □

The previous Corollary shows that in a category C satisfying the shifting property an object X is abelian precisely when there is a connector between ∇_X and ∇_X. Let $Mal(C)$ be the subcategory of the abelian objects of C, with arrows those arrows in C which also respect the Maltsev operation. Now, the assumption that C satisfies the shifting property implies that any arrow in C respects the Maltsev operation (Corollary 3.2). Consequently:

3.5. Corollary. *When C satisfies the shifting property, the category of abelian objects $Mal(C)$ is naturally Maltsev.*

Proof Recall that a finitely complete category $\mathcal{A}$ is naturally Maltsev [17] if, in the category $[\mathcal{A},\mathcal{A}]$ of endofunctors and natural transformations, the identity functor $Id_{\mathcal{A}}$ is provided with a Maltsev operation. By definition of $Mal(C)$ any object X in $Mal(C)$ is provided with a Maltsev operation $p_X\colon X \times X \times X \to X$ and the naturality of p follows from the fullness of $Mal(C)$ in C. □

4 Internal categories

A description of internal categories in congruence modular varieties was obtained in [15]. We are now going to extend and improve some of the results proved in [15] thanks to a systematic use of the shifting property. Remark that an internal category, unlike an internal groupoid (Example 1.3), does not necessarily give rise to an internal pseudogroupoid. Consequently, the results of this section can not be deduced directly from the preceding ones concerning pseudogroupoids.

An internal reflexive graph X in $\mathcal{C}$

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"' description] \end{tikzcd}$$

has the property that $d_0 \circ s_0 = 1_{X_0} = d_1 \circ s_0$. X is endowed with an *internal category* structure

$$\begin{tikzcd} X_1 \times_{X_0} X_1 \arrow[r, shift left=3, "\pi_1"] \arrow[r, "m" description] \arrow[r, shift right=3, "\pi_0"'] & X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"' description] \end{tikzcd}$$

when there is a multiplication $m\colon X_1 \times_{X_0} X_1 \to X_1$, satisfying the axioms

1. $m(1_{d_0(x)}, x) = x = m(x, 1_{d_1(x)})$
2. $d_0(m(x,y)) = d_0(x)$
3. $d_1(m(x,y)) = d_1(y)$
4. $m(x, m(y,z)) = m(m(x,y),z)$

(where we write 1_X for $s_0(X)$).

4.1. Lemma. *Let X be an internal reflexive graph in a finitely complete category $\mathcal{C}$ satisfying the shifting property. If there is a multiplication $m\colon X_1 \times_{X_0} X_1 \to X_1$ on X satisfying the axiom 1, then $m \cap \pi_0 = \Delta_{X_1 \times_{X_0} X_1}$ (right cancellation property).*

Proof If we assume that $m(a,x) = m(a,y)$, then we can form the diagram

$$\begin{tikzcd} (a,x) \arrow[r, dash, "\pi_1"] \arrow[d, dash, "\pi_0"'] \arrow[d, dash, bend right=60, "m"'] & (1_{d_1(a)}, x) \arrow[d, dash, "\pi_0"] \\ (a,y) \arrow[r, dash, "\pi_1"'] & (1_{d_1(a)}, y), \end{tikzcd}$$

where each represented element is in $X_1 \times_{X_0} X_1$. Of course $\pi_0 \cap \pi_1 = \Delta_{X_1 \times_{X_0} X_1}$, and the shifting property gives

$$x = m(1_{d_1(a)}, x) = m(1_{d_1(a)}, y) = y.$$

□

4.2. Proposition. *Let X be an internal reflexive graph in a finitely complete category $\mathcal{C}$ satisfying the shifting property. Then there is at most one multiplication $m\colon X_1 \times_{X_0} X_1 \to X_1$ on X satisfying the axioms 1 and 2.*

Proof Let us consider two multiplications m and m' on X satisfying the axioms 1 and 2. For any $(x,y) \in X_1 \times_{X_0} X_1$ we can form the diagram

$$\begin{tikzcd}
(x, 1_{d_1(x)}) \arrow[r, dash, "\pi_0"] \arrow[d, dash, "m"] \arrow[d, dash, bend right, "m'"'] & (x,y) \arrow[d, dash, "m"] \\
(1_{d_0(x)}, x) \arrow[r, dash, "\pi_0"'] & (1_{d_0(m(x,y))}, m(x,y)),
\end{tikzcd}$$

where $d_0(x) = d_0(m(x,y))$ by axiom 2. Then the previous Lemma gives

$$\Delta_{X_1 \times_{X_0} X_1} = m \cap \pi_0 \leq m',$$

and clearly one has that

$$m'(x, 1_{d_1(x)}) = x = m'(1_{d_0(x)}, x).$$

Then the shifting property gives

$$m'(x,y) = m'(1_{d_0(m(x,y))}, m(x,y)) = m(x,y).$$

□

4.3. Proposition. *Let C be a finitely complete category satisfying the shifting property. Then any multiplication m on a reflexive graph satisfying the axioms* 1, 2 *and* 3 *is associative.*

Proof Given three composable arrows x, y, z in X_1

$$A \xrightarrow{x} B \xrightarrow{y} C \xrightarrow{z} D,$$

we can form the following diagram of four elements in the equivalence relation π_1 on $X_1 \times_{X_0} X_1$ (we write ∇ instead of $\nabla_{X_1 \times_{X_0} X_1}$ to save space):

$$\begin{tikzcd}
{[(m(x,y), 1_C), (y, 1_C)]} \arrow[r, dash, "\pi_0 \Box \pi_1"] \arrow[d, dash, "\nabla \times m"] \arrow[d, dash, bend right, "m \Box \pi_1"'] & {[(m(x,y), z), (y, z)]} \arrow[d, dash, "\nabla \times m"] \\
{[(x,y), (1_B, y)]} \arrow[r, dash, "\pi_0 \Box \pi_1"'] & {[(x, m(y,z)), (1_B, m(y,z))],}
\end{tikzcd}$$

where $\pi_0 \Box \pi_1$ and $m \Box \pi_1$ are thought as equivalence relations on π_1, and $\nabla \times m$ is the equivalence relation on π_1 defined by $[(a,b),(c,b)](\nabla \times m)[(d,e),(f,e)]$ if and only if $m(c,b) = m(f,e)$. Then, if $(\nabla \times m) \cap (\pi_0 \Box \pi_1) \leq m \Box \pi_1$, the shifting property implies that

$$[(m(x,y), z), (y, z)](m \Box \pi_1)[(x, m(y,z)), (1_B, m(y,z))],$$

which in particular gives

$$m(m(x,y), z) = m(x, m(y,z)),$$

as desired.

The reason why the inequality $(\nabla \times m) \cap (\pi_0 \Box \pi_1) \leq m \Box \pi_1$ holds is that $(\nabla \times m) \cap (\pi_0 \Box \pi_1) = \Delta_{\pi_1}$. For this, let us consider

$$[(a,b),(c,b)](\nabla \times m) \cap (\pi_0 \Box \pi_1)[(a,e),(c,e)].$$

The fact that $(c,b)(m\cap\pi_0)(c,e)$ implies that $b=e$ because $m\cap\pi_0=\Delta_{X_1\times_{X_0}X_1}$ (see Lemma 4.1). Accordingly, $[(a,b),(c,b)]=[(a,e),(c,e)]$. $\square$

4.4. Corollary. *For a reflexive graph X in a finitely complete category C satisfying the shifting property the following conditions are equivalent:*

1. *there is an internal category structure whose underlying reflexive graph is X*
2. *there is a unique internal category structure whose underlying reflexive graph is X*
3. *there is a multiplication m satisfying the axioms $1,2$ and 3*
4. *there is a unique multiplication m satisfying the axioms $1,2$ and 3*

4.5. Remark. This corollary extends the equivalence between conditions a,b,c,d of Theorem 4.1 in [15] from modular varieties to any category C satisfying the shifting property. Moreover, we improve that result by dropping the assumption that $R[d_0]$ and $R[d_1]$ have trivial commutator.

Let $RG(C)$ denote the category of internal reflexive graphs in C.

4.6. Proposition. *Let C be a finitely complete category satisfying the shifting property. Then the forgetful functor*

$$Cat(C)\to RG(C)$$

is a full inclusion.

Proof Thanks to Proposition 4.2 we only need to prove that any arrow (ϕ_0,ϕ_1) of internal reflexive graphs

$$\begin{tikzcd} X_1 \arrow[r, "\phi_1"] \arrow[d, shift right, "d_1"'] \arrow[d, shift left, "d_0"] & Y_1 \arrow[d, shift right, "d_1"'] \arrow[d, shift left, "d_0"] \\ X_0 \arrow[r, "\phi_0"'] & Y_0 \end{tikzcd}$$

between two internal categories (X,m) and (Y,μ) always preserves the (unique) multiplication. If $\phi\colon X_1\times_{X_0}X_1\to Y_1\times_{Y_0}Y_1$ is the arrow induced by the universal property of the pullbacks, then, given any pair of composable arrows in X_1

$$X\xrightarrow{f}Y\xrightarrow{g}Z,$$

one has that

$$(\mu\circ\phi)(f,1_Y)=\phi_1(f)=(\mu\circ\phi)(1_X,f).$$

Since $m\cap\pi_0=\Delta_{X_1\times_{X_0}X_1}$ by Lemma 4.1, we can apply the shifting property to

$$\begin{tikzcd} (f,1_Y) \arrow[r, dash, "\pi_0"] \arrow[d, dash, "m"'] \arrow[d, dash, bend right=60, "\mu\circ\phi"'] & (f,g) \arrow[d, dash, "m"] \\ (1_X,f) \arrow[r, dash, "\pi_0"'] & (1_X,m(f,g)). \end{tikzcd}$$

This implies that

$$\begin{aligned}(\mu\circ\phi)(f,g) &= (\mu\circ\phi)(1_X, m(f,g))\\ &= \mu(1_{\phi_1(X)}, (\phi_1\circ m)(f,g))\\ &= (\phi_1\circ m)(f,g).\end{aligned}$$

□

Another result can be obtained when the finitely complete category $\mathcal{C}$ is assumed to be also *regular*, i.e. $\mathcal{C}$ has coequalizers of effective equivalence relations and the regular epis are pullback stable.

4.7. Proposition. *Let $\mathcal{C}$ be a regular category satisfying the shifting property. Let X be an internal reflexive graph in $\mathcal{C}$*

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"'] \end{tikzcd}$$

equipped with a pseudogroupoid structure μ on $R[d_0]$ and $R[d_1]$. Then the following conditions are equivalent:

1. *there is a (unique) category structure on X*
2. *X_1 determines a preorder on X_0. This means that the regular image $I = \frac{X_1}{R[d_0]\cap R[d_1]}$ of the reflexive graph X is also a transitive relation on X_0.*

Proof In this proof we shall use Barr's metatheorem for regular categories. [1]. Thanks to this result, the elementwise procedure we used so far to prove finite limit properties can be augmented, in a regular category, by the fact that one may treat regular epimorphisms as if they were surjections.
It is clear that any internal category structure on X has the property that the image I is transitive. Indeed, for any (X,Y) and (Y,Z) in I there exist in X_1 at least one arrow $f\colon X\to Y$ and one arrow $g\colon Y\to Z$, and consequently the arrow $m(f,g)\colon X\to Z$ shows that (X,Z) is also in I.

Conversely, let us assume that I is transitive. Then, for any $f\colon X\to Y$ and $g\colon Y\to Z$, there exists an arrow $\phi\colon X\to Z$ with the property that $\begin{pmatrix} g & \phi\\ 1_Y & f\end{pmatrix}$ is in $R[d_0]\square R[d_1]$. We can then define $m(f,g)=\mu(g,1_Y,\phi,f)$ (μ does not depend on the third variable). We then have that

$$fR[d_0]m(f,g)R[d_1]g$$

and, moreover,

$$m(f,1_Y)=\mu(1_Y,1_Y,f,f)=f=\mu(f,1_X,f,1_X)=m(1_X,f).$$

By Corollary 4.4 the proof is then complete. □

Corollary 4.4 and Proposition 4.7 here above then extend Theorem 4.1 in [15] to any regular category satisfying the shifting property.

5 Internal groupoids

Since any internal groupoid is an internal category, the previous results give in particular a description of the internal reflexive graphs underlying a (necessarily unique) groupoid structure. Indeed, it is sufficient to require also the existence of an arrow $i\colon X_1 \to X_1$ satisfying the usual conditions of an inverse.

However, a very simple and neat description of the internal groupoids in a category $\mathcal{C}$ satisfying the shifting property can be obtained by using the fact that a groupoid always gives rise to a pseudogroupoid.

We first establish a more general result, which clarifies the relationship between the notions of pseudogroupoid and of connector:

5.1. Proposition. *Let $\mathcal{C}$ be a regular category satisfying the shifting property, and let R and S be two equivalence relations on X, for X in $\mathcal{C}$. Then the following conditions are equivalent:*

1. *there is a (unique) connector between R and S*
2. *there is a (unique) pseudogroupoid on R and S and $R \circ S = S \circ R$*
3. *there is a (unique) quasiconnector on R and S and $R \circ S = S \circ R$*
4. *the canonical arrow $\alpha\colon R\square S \to R \times_X S$ is split by a (unique) arrow $i\colon R \times_X S \to R\square S$ of double zero sequences in $Pt_X(\mathcal{C})$*

Proof The uniqueness of the various structures follows from Theorem 2.13.
$1 \Rightarrow 2$ For any $xRySz$ there exists $p(x,y,z)$ such that $xSp(x,y,z)Rz$, hence $R\circ S \leq S \circ R$, and then $R \circ S = S \circ R$.
$2 \Rightarrow 3$ Trivial.
$3 \Rightarrow 4$ By using once again Barr's metatheorem for regular categories one can prove that the canonical arrow $\alpha\colon R\square S \to R \times_X S$ is a regular epi precisely when $R \circ S = S \circ R$. Now, if $\sigma\colon R\square S \to R\square S$ is the quasiconnector on R and S, the fact that $R[\alpha] = R[\sigma]$ implies that α is the coequalizer of σ and $1_{R\square S}$. It follows that there is a unique $i\colon R \times_X S \to R\square S$ with $i \circ \alpha = \sigma$ and $\alpha \circ i = 1_{R\times_X S}$. It is then easy to check that i is an arrow of double zero sequences in $Pt_X(\mathcal{C})$.
$4 \Rightarrow 1$ By defining $p = d_1 \circ \pi_{1,3} \circ i$ one obtains an internal partial Maltsev operation satisfying also axiom I in the definition of a connector. By Corollary 3.1 this is sufficient to conclude that $p\colon R \times_X S \to X$ is a connector. □

5.2. Corollary. *Let $\mathcal{C}$ be a regular category satisfying the shifting property. For an internal reflexive graph X*

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"' description] \end{tikzcd}$$

in $\mathcal{C}$ the following conditions are equivalent:

1. *there is a (unique) groupoid structure on X*
2. *there is a (unique) connector between $R[d_0]$ and $R[d_1]$*
3. *there is a (unique) quasiconnector on $R[d_0]$ and $R[d_1]$ and*

$$R[d_0] \circ R[d_1] = R[d_1] \circ R[d_0]$$

Proof It follows by the previous Proposition. □

We can now extend Corollary 4.3 in [15] characterizing internal groupoids in modular varieties:

5.3. Corollary. *Let $\mathcal{C}$ be a regular category satisfying the shifting property. For an internal reflexive graph X in $\mathcal{C}$*

$$\begin{tikzcd} X_1 \arrow[r, shift left=2, "d_0"] \arrow[r, shift right=2, "d_1"'] & X_0 \arrow[l, "s_0"' description] \end{tikzcd}$$

the following conditions are equivalent:

1. *there is a unique groupoid structure on X*
2. *there is a unique pseudogroupoid structure on $R[d_0]$ and $R[d_1]$, and X determines an equivalence relation $I = \frac{X_1}{R[d_0] \cap R[d_1]}$ on X_0.*

Proof $1 \Rightarrow 2$ Since any groupoid X gives rise to a pseudogroupoid structure on $R[d_0]$ and $R[d_1]$, thanks to Proposition 4.7 we only need to prove that I is a symmetric relation on X_0. This is clear, because when X is a groupoid, then for any $f\colon X \to Y$ there is an arrow $f^{-1}\colon Y \to X$, so that if (X, Y) is in I, so is (Y, X). $2 \Rightarrow 1$ We already know that X is an internal category (Proposition 4.7). On the other hand, if for any (X, Y) in I, then there is also (Y, X) in I, then for any $f\colon X \to Y$ there is at least one arrow $g\colon Y \to X$. By setting $f^{-1} = \mu(1_X, f, g, 1_Y)$ and by axiom $4B$ in the definition of a pseudogroupoid one has that

$$m(f, \mu(1_X, f, g, 1_Y)) = \mu(\mu(1_X, f, g, 1_Y), 1_Y, \mu(g, 1_Y, 1_X, f), f) = \mu(1_X, f, 1_X, f),$$

so that $m(f, \mu(1_X, f, g, 1_Y)) = 1_X$. Similarly, by $4A$ one has that

$$m(\mu(1_X, f, g, 1_Y), f) = \mu(f, 1_X, \mu(f, 1_X, 1_Y, g), \mu(1_X, f, g, 1_Y)) = \mu(f, f, 1_Y, 1_Y),$$

hence $m(\mu(1_X, f, g, 1_Y), f) = 1_Y$. This proves that $\mu(1_X, f, g, 1_Y)$ is the inverse of f, and X is then an internal groupoid. □

References

[1] M. Barr, *Exact Categories*, LNM 236, Springer-Verlag, 1971, 1-120.

[2] F. Borceux, *A survey of semi-abelian categories*, to appear in this volume.

[3] D. Bourn, *Normalization, equivalence, kernel equivalence and affine categories,* LNM 1488, Springer-Verlag, 1991, 43–62.

[4] D. Bourn - M. Gran, *Centrality and normality in protomodular categories*, Theory Appl. Cat., Vol. 9, No. 8, 2002, 151-165.

[5] D. Bourn - M. Gran, *Centrality and connectors in Maltsev categories*, Algebra Universalis, 48, 2002, 309-331.

[6] R. Brown - C. Spencer, *G-Groupoids, crossed modules and the fundamental groupoid of a topological group,* Proc. Konn. Ned. Akad. Wet, 79, 1976, 296-302.

[7] A. Carboni - G.M. Kelly - M.C. Pedicchio, *Some remarks on Maltsev and Goursat categories* Appl. Categorical Structures, 1, 1993, 385-421.

[8] A. Carboni - J. Lambek - M.C. Pedicchio, *Diagram chasing in Mal'cev categories,* J. Pure Appl. Algebra, 69, 1990, 271-284.

[9] A. Carboni - M.C. Pedicchio - N. Pirovano, *Internal graphs and internal groupoids in Mal'cev categories,* Proc. Conference Montreal 1991, 1992, 97-109.

[10] R. Freese - R. McKenzie, *Commutator theory for congruence modular varieties*, Lond. Math. Soc. Lect. Notes Series, 125, Cambr. Univ. Press, 1987.

[11] M. Gran, *Internal categories in Mal'cev categories,* J. Pure Appl. Algebra, 143, 1999, 221-229.

[12] M. Gran, *Commutators and central extensions in universal algebra,* J. Pure Appl. Algebra, 174, 2002, 249-261.

[13] H. P. Gumm, *Geometrical Methods in Congruence Modular Algebras*, Mem. Amer. Math. Soc., 45 286, 1983.

[14] J. Hagemann - C. Herrmann, *A concrete ideal multiplication for algebraic systems and its relation to congruence distributivity*, Arch. Math. (Basel), 32, 1979, 234-245.

[15] G. Janelidze - M.C. Pedicchio, *Internal categories and groupoids in congruence modular varieties*, Journal of Algebra 193, 1997, 552-570.

[16] G. Janelidze - M.C. Pedicchio, *Pseudogroupoids and commutators*, Theory Appl. Cat., Vol. 8, No. 15, 2001, 408-456.

[17] P.T. Johnstone, *Affine categories and naturally Mal'cev categories,* J. Pure Appl. Algebra, 61, 1989, 251-256.

[18] E.W. Kiss, *Three remarks on the modular commutator,* Alg. Universalis 29, 1992, 455-476.

[19] J.-L. Loday, *Spaces with finitely many non-trivial homotopy groups,* J. Pure Appl. Algebra, 24, 1982, 179-202.

[20] M.C. Pedicchio, *A categorical approach to commutator theory*, Journal of Algebra, 177, 1995, 647-657.

[21] M.C. Pedicchio, *Arithmetical categories and commutator theory*, Appl. Categorical Structures, 4, 1996, 297-305.

[22] J.D.H. Smith, *Mal'cev Varieties*, LNM 554, Springer-Verlag, 1976.

Received November 28, 2002; in revised form March 14, 2003

Fields Institute Communications
Volume 43, 2004

Crossed Complexes and Homotopy Groupoids as Non Commutative Tools for Higher Dimensional Local-to-global Problems

Ronald Brown
Department of Mathematics
University of Wales, Bangor
Gwynedd LL57 1UT, U.K.
r.brown@bangor.ac.uk

With thanks to many colleagues and Bangor research students.

Abstract. We outline the main features of the definitions and applications of crossed complexes and cubical ω-groupoids with connections. These give forms of higher homotopy groupoids, and new views of basic algebraic topology and the cohomology of groups, with the ability to obtain some non commutative results and compute some homotopy types.

Contents

2000 *Mathematics Subject Classification.* Primary 01-01, 16E05, 18D05, 18D35, 55P15, 55Q05;

This is an extended account of a lecture given at the meeting on 'Categorical Structures for Descent, Galois Theory, Hopf algebras and semiabelian categories', Fields Institute, September 23-28, 2002, for which the author is grateful for support from the Fields Institute and a Leverhulme Emeritus Research Fellowship. Other support is indicated later.

Introduction

An aim is to give a survey of results obtained by R. Brown and P.J. Higgins and others over the years 1974-2002, and to point to applications and related areas. This work gives an account of some basic algebraic topology which differs from the standard account through the use of *crossed complexes*, rather than chain complexes, as a fundamental notion. In this way one obtains comparatively quickly not only classical results such as the Brouwer degree and the relative Hurewicz theorem, but also non commutative results on second relative homotopy groups, as well as higher dimensional results involving the action of and also presentations of the fundamental group. For example, the fundamental crossed complex ΠX_* of the skeletal filtration of a CW-complex X is a useful generalisation of the usual cellular chains of the universal cover of X. It also gives a replacement for singular chains by taking X to be the geometric realisation of a singular complex of a space.

A replacement for the excision theorem in homology is obtained by using cubical methods to prove a colimit theorem for the fundamental crossed complex functor on filtered spaces. This colimit theorem is a higher dimensional version of a classical example of a *non commutative local-to-global theorem*, which itself was the initial motivation for the work described here. This Seifert-Van Kampen Theorem (SVKT) determines completely the fundamental group $\pi_1(X, x)$ of a space X with base point which is the union of open sets U, V whose intersection is path connected and contains the base point x; the 'local information' is on the morphisms of fundamental groups induced by the inclusions $U \cap V \to U, U \cap V \to V$. The importance of this result reflects the importance of the fundamental group in algebraic topology, algebraic geometry, complex analysis, and many other subjects. Indeed the origin of the fundamental group was in Poincaré's work on monodromy for complex variable theory.

Essential to this use of crossed complexes, particularly for conjecturing and proving local-to-global theorems, is a construction of *higher homotopy groupoids*, with properties described by *an algebra of cubes*. There are applications to local-to-global problems in homotopy theory which are more powerful than purely classical tools, while shedding light on those tools. It is hoped that this account will increase the interest in the possibility of wider applications of these methods and results, since homotopical methods play a key role in many areas. It is relevant that Atiyah in [8] identifies some major themes in 20th century mathematics as: *local to global,* *from commutative to non commutative, increase in dimensions, homology and K-theory.* The higher categorical methods described here allow the combining of these themes, and yield steps towards a *non commutative algebraic topology.*

Higher homotopy groups

Topologists in the early part of the 20th century were well aware that: the non commutativity of the fundamental group was useful in geometric applications; for path connected X there was an isomorphism

$$H_1(X) \cong \pi_1(X,x)^{\mathrm{ab}};$$

and the abelian homology groups existed in all dimensions. Consequently there was a desire to generalise the non commutative fundamental group to all dimensions.

In 1932 Čech submitted a paper on higher homotopy groups $\pi_n(X,x)$ to the ICM at Zurich, but it was quickly proved that these groups were abelian for $n \geqslant 2$, and on these grounds Čech was persuaded to withdraw his paper, so that only a small paragraph appeared in the Proceedings [55]. We now see the reason for this commutativity as the result (Eckmann-Hilton) that a group internal to the category of groups is just an abelian group. Thus the vision of a non commutative higher dimensional version of the fundamental group has since 1932 been generally considered to be a mirage. Before we go back to the SVKT, we explain in the next section how nevertheless work on crossed modules did introduce non commutative structures relevant to topology in dimension 2.

Of course higher homotopy groups were strongly developed following on from the work of Hurewicz (1935). The fundamental group still came into the picture with its action on the higher homotopy groups, which J.H.C. Whitehead once remarked (1957) was especially fascinating for the early workers in homotopy theory. Much of Whitehead's work was intended to extend to higher dimensions the methods of combinatorial group theory of the 1930s – hence the title of his papers: 'Combinatorial homotopy, I, II' [104, 105]. The first of these two papers has been very influential and is part of the basic structure of algebraic topology. It is the development of work of the second paper which we explain here.

Whitehead's paper on 'Simple homotopy types', [106], which deals with higher dimensional analogues of Tietze transformations, has a final section using crossed complexes. We refer to this again in section 15. Related work is by R.A. Brown [15].

It is hoped also that this survey will be useful background to work on the Van Kampen Theorem for diagrams of spaces in [42], which uses a form of homotopy groupoid which is in one sense much more powerful than that given here, since it encompasses n-adic information, but in which current expositions are still restricted to the reduced (one base point) case.

1 Crossed modules

In the years 1941-50, Whitehead developed work on crossed modules to represent the structure of the boundary map of the relative homotopy group

$$\pi_2(X,A,x) \to \pi_1(A,x) \qquad (*)$$

in which both groups can be non commutative. Here is the definition.

A *crossed module* is a morphism of groups $\mu : M \to P$ together with an action $(m,p) \mapsto m^p$ of the group P on the group M satisfying the two axioms

CM1) $\mu(m^p) = p^{-1}(\mu m)p$

CM2) $n^{-1}mn = m^{\mu n}$

for all $m, n \in M, p \in P$.

Standard algebraic examples of crossed modules are:
(i) an inclusion of a normal subgroup, with action given by conjugation;
(ii) the inner automorphism map $\chi : M \to \operatorname{Aut} M$, in which χm is the automorphism $n \mapsto m^{-1}nm$;
(iii) the zero map $M \to P$ where M is a P-module;
(iv) an epimorphism $M \to P$ with kernel contained in the centre of M.

Simple consequences of the axioms for a crossed module $\mu : M \to P$ are:

1.1 $\operatorname{Im} \mu$ is normal in P.

1.2 $\operatorname{Ker} \mu$ is central in M and is acted on trivially by $\operatorname{Im} \mu$, so that $\operatorname{Ker} \mu$ inherits an action of $M / \operatorname{Im} \mu$.

Another important construction is the *free crossed P-module*

$$\partial : C(\omega) \to P$$

determined by a function $\omega : R \to P$, where R is a set. The group $C(\omega)$ is generated by $R \times P$ with the relations

$$(r,p)^{-1}(s,q)^{-1}(r,p)(s,qp^{-1}(\omega r)p)$$

the action is given by $(r,p)^q = (r,pq)$ and the boundary morphism is given by $\partial(r,p) = p^{-1}(\omega r)p$, for all $(r,p),(s,q) \in R \times P$.

A major result of Whitehead was:

Theorem W [105] *If the space $X = A \cup \{e_r^2\}_{r \in R}$ is obtained from A by attaching 2-cells by maps $f_r : (S^1, 1) \to (A, x)$, then the crossed module of (*) is isomorphic to the free crossed $\pi_1(A,x)$-module on the classes of the attaching maps of the 2-cells.*

Whitehead's proof, which stretched over three papers, 1941-1949, used transversality and knot theory – an exposition is given in [18]. Mac Lane and Whitehead [90] used this result as part of their proof that crossed modules capture all homotopy 2-types (they used the term '3-types').

The title of the paper in which the first intimation of Theorem W appeared was 'On adding relations to homotopy groups' [103]. This indicates a search for higher dimensional SVKTs.

The concept of free crossed module gives a non commutative context for *chains of syzygies*. The latter idea, in the case of modules over polynomial rings, is one of the origins of homological algebra through the notion of *free resolution*. Here is how similar ideas can be applied to groups. Pioneering work here, independent of Whitehead, was by Peiffer [94] and Reidemeister [97]. See [37] for an exposition of these ideas.

Suppose $\mathcal{P} = \langle X \mid \omega \rangle$ is a presentation of a group G, where $\omega : R \to F(X)$ is a function, allowing for repeated relators. Then we have an exact sequence

$$1 \xrightarrow{i} N(\omega R) \xrightarrow{\phi} F(X) \longrightarrow G \longrightarrow 1$$

where $N(\omega R)$ is the normal closure in $F(X)$ of the set ωR of relations. The above work of Reidemeister, Peiffer, and Whitehead showed that to obtain the next level of syzygies one should consider the free crossed $F(X)$-module $\partial : C(\omega) \to F(X)$, since this takes into account the operations of $F(X)$ on its normal subgroup $N(\omega R)$. Elements of $C(\omega)$ are a kind of 'formal consequence of the relators', so that the relation between the elements of $C(\omega)$ and those of $N(\omega R)$ is analogous to the relation between the elements of $F(X)$ and those of G. The kernel $\pi(\mathcal{P})$ of ∂ is a G-module, called the module of *identities among relations*, and there is considerable work on computing it [37, 96, 78, 67, 47]. By splicing to ∂ a free G-module resolution of $\pi(\mathcal{P})$ one obtains what is called a *free crossed resolution* of the group G. These resolutions have better realisation properties than the usual resolutions by chain complexes of G-modules, as explained later.

This notion of using crossed modules as the first stage of syzygies in fact represents a wider tradition in homological algebra, in the work of Frölich and Lue [69, 87].

Crossed modules also occurred in other contexts, notably in representing elements of the cohomology group $H^3(G, M)$ of a group G with coefficients in a G-module M [89], and as coefficients in Dedecker's theory of non abelian cohomology [60]. The notion of free crossed resolution has been exploited by Huebschmann [79, 81, 80] to represent cohomology classes in $H^n(G, M)$ of a group G with coefficients in a G-module M, and also to calculate with these.

Our results can make it easier to compute a crossed module arising from some topological situation, such as an induced crossed module [50, 51], or a coproduct crossed module [19], than the cohomology element in $H^3(G, M)$ it represents. To obtain information on such an element it is useful to work with a small free crossed resolution of G, and this is one motivation for developing methods for calculating such resolutions. However, it is not so clear what a *calculation* of such a cohomology element would amount to, although it is interesting to know whether the element is non zero, or what is its order. Thus the use of algebraic models of cohomology classes may yield easier computations than the use of cocycles, and this somewhat inverts traditional approaches.

Since crossed modules are algebraic objects generalising groups, it is natural to consider the problem of explicit calculations by extending techniques of computational group theory. Substantial work on this has been done by C.D. Wensley using the program GAP [71, 52].

2 The fundamental groupoid on a set of base points

A change in prospects for higher order non commutative invariants was derived from work of the writer published in 1967 [16], and influenced by Higgins' paper [75]. This showed that the Van Kampen Theorem could be formulated for the *fundamental groupoid* $\pi_1(X, X_0)$ *on a set* X_0 *of base points*, thus enabling computations in the non-connected case, including, as explained in [17, p.319], those in Van Kampen's original paper [83]. This use of groupoids in dimension 1 suggested the question of the use of groupoids in higher homotopy theory, and in particular the question of the existence of *higher homotopy groupoids.*

In order to see how this research programme could go it is useful to consider the statement and special features of this generalised Van Kampen Theorem for the fundamental groupoid. First, if X_0 is a set , and X is a space, then $\pi_1(X, X_0)$ denotes the fundamental groupoid on the set $X \cap X_0$ of base points. This allows the set X_0 to be chosen in a way appropriate to the geometry. For example, if the circle S^1 is written as the union of two semicircles $E_+ \cup E_-$, then the intersection $\{-1, 1\}$ of the semicircles is not connected, so it is not clear where to take the base point. Instead one takes $X_0 = \{-1, 1\}$, and so has two base points. This flexibility is very important in computations, and this example of S^1 was a motivating example for this development. As another example, you might like to consider the difference between the quotients of the actions of $\mathbb{Z}_2$ on the group $\pi_1(S^1, 1)$ and on the groupoid $\pi_1(S^1, \{-1, 1\})$ where the action is induced by complex conjugation on S^1. Relevant work on orbit groupoids has been developed by Higgins and Taylor [77, 98].

Consideration of a set of base points leads to the theorem:

Theorem 2.1 [16] *Let the space X be the union of open sets U, V with intersection W, and let X_0 be a subset of X meeting each path component of U, V, W. Then*
(C) (connectivity) *X_0 meets each path component of X and*
(I) (isomorphism) *the diagram of groupoid morphisms induced by inclusions*

$$\begin{tikzcd} \pi_1(W, X_0) \arrow[r, "i"] \arrow[d, "j"'] & \pi_1(U, X_0) \arrow[d] \\ \pi_1(V, X_0) \arrow[r] & \pi_1(X, X_0) \end{tikzcd} \tag{2.1}$$

is a pushout of groupoids.

From this theorem, one can compute a particular fundamental group $\pi_1(X, x_0)$ using combinatorial information on the graph of intersections of path components of U, V, W, but for this it is useful to develop the algebra of groupoids. Notice two special features of this result.
(i) The computation of the invariant you want, the fundamental group, is obtained from the computation of a larger structure, and so part of the work is to give methods for computing the smaller structure from the larger one. This usually involves non canonical choices, e.g. that of a maximal tree in a connected graph. The work on applying groupoids to groups gives many examples of this [75, 76, 17].
(ii) The fact that the computation can be done is surprising in two ways: (a) The fundamental group is computed *precisely*, even though the information for it uses input in two dimensions, namely 0 and 1. This is contrary to the experience in homological algebra and algebraic topology, where the interaction of several dimensions involves exact sequences or spectral sequences, which give information only up to extension, and (b) the result is a non commutative invariant, which is usually even more difficult to compute precisely.

The reason for the success seems to be that the fundamental groupoid $\pi_1(X, X_0)$ contains information in dimensions 0 and 1, and so can adequately reflect the geometry of the intersections of the path components of U, V, W and the morphisms induced by the inclusions of W in U and V.

This suggested the question of whether these methods could be extended successfully to higher dimensions.

Part of the initial evidence for this quest was the intuitions in the proof of this groupoid SVKT, which seemed to use three main ideas in order to verify the universal property of a pushout for diagram (2.1) and given morphisms f_U, f_V from $\pi_1(U, X_0), \pi_1(V, X_0)$ to a groupoid G, satisfying $f_U i = f_V j$:

• A deformation or filling argument. Given a path $a : (I, \dot{I}) \to (X, X_0)$ one can write $a = a_1 + \cdots + a_n$ where each a_i maps into U or V, but a_i will not necessarily have end points in X_0. So one has to deform each a_i to a_i' in U, V or W, using the connectivity condition, so that each a_i' has end points in X_0, and $a' = a_1' + \cdots + a_n'$

is well defined. Then one can construct using f_U or f_V an image of each a'_i in G and hence of the composite, called $F(a) \in G$, of these images. Note that we subdivide in X and then put together again in G (this uses the condition $f_U i = f_V j$ to prove that the elements of G are composable), and this part can be summarised as:

- Groupoids provided a convenient *algebraic inverse to subdivision.*

Next one has to prove that $F(a)$ depends only on the class of a in the fundamental groupoid. This involves a homotopy rel end points $h : a \simeq b$, considered as a map $I^2 \to X$; subdivide h as $h = [h_{ij}]$ so that each h_{ij} maps into U, V or W; deform h to $h' = [h'_{ij}]$ (keeping in U, V, W) so that each h'_{ij} maps the vertices to X_0 and so determines a commutative square in one of $\pi_1(Q, X_0)$ for $Q = U, V, W$. Move these commutative squares over to G using f_U, f_V and recompose them (this is possible again because of the condition $f_U i = f_V j$), noting that:

- in a groupoid, *any composition of commutative squares is commutative.*

Two opposite sides of the composite commutative square in G so obtained are identities, because h was a homotopy rel end points, and the other two are $F(a)$, $F(b)$. This proves that $F(a) = F(b)$ in G.

Thus the argument can be summarised: a path or homotopy is divided into small pieces, then deformed so that these pieces can be packaged and moved over to G, where they are reassembled. There seems to be an analogy with the processing of an email.

Notable applications of the groupoid theorem were: (i) to give a proof of a formula in Van Kampen's paper of the fundamental group of a space which is the union of two connected spaces with non connected intersection, see [17, 8.4.9]; and (ii) to show the topological utility of the construction by Higgins [76] of the groupoid $f_*(G)$ over Y_0 induced from a groupoid G over X_0 by a function $f : X_0 \to Y_0$. (Accounts of these with the notation $U_f(G)$ rather than $f_*(G)$ are given in [76, 17].) This latter construction is regarded as a 'change of base', and analogues in higher dimensions yielded generalisations of the Relative Hurewicz Theorem and of Theorem W, using induced modules and crossed modules.

There is another approach to the Van Kampen Theorem which goes via the theory of covering spaces, and the equivalence between covering spaces of a reasonable space X and functors $\pi_1(X) \to \mathsf{Set}$ [17]. See for example [61, 13] for an exposition of the relation with Galois theory. The paper [39] gives a general formulation of conditions for the theorem to hold in the case $X_0 = X$ in terms of the map $U \sqcup V \to X$ being an 'effective global descent morphism' (the theorem is given in the generality of lextensive categories). This work has been developed for toposes [53]. Analogous interpretations for higher dimensional Van Kampen theorems are not known.

The justification of the breaking of a paradigm in changing from groups to groupoids is several fold: the elegance and power of the results; the increased linking with other uses of groupoids [20, 102]; and the opening out of new possibilities in higher dimensions, which allowed for new results and calculations in homotopy theory, and suggested new algebraic constructions.

3 The search for higher homotopy groupoids

Contemplation of the proof of the SVGKT in the last section suggested that a higher dimensional version should exist, though this version amounted to an idea of a proof in search of a theorem. In the end, the results exactly encapsulated this intuition.

One intuition was that in groupoids we are dealing with a partial algebraic structure, in which composition is defined for two arrows if and only if the source of one arrow is the target of the other. This seems easily to generalise to directed squares, in which two are composable horizontally if and only if the left hand side of one is the right hand side of the other (and similarly vertically).

However the formulation of a theorem in higher dimensions required specification of the three elements of a functor

$$\Pi : (\text{topological data}) \to (\text{higher order groupoids})$$

which would allow the expression of these ideas for the proof.

C. Ehresmann had defined double categories in [64].

Experiments were made in the years 1967-1973 to define some functor Π from spaces to some kind of double groupoid, using compositions of squares in two directions, but these proved abortive. However considerable progress was made in work with Chris Spencer in 1971-3 on investigating the algebra of double groupoids [48, 49], and showing a relation to crossed modules. Further evidence was provided when it was found [49] that group objects in the category of groupoids are not necessarily commutative objects, since they are equivalent to crossed modules. See [46] for an application of this equivalence to covering spaces of non-connected topological groups. (It turned out that this equivalence was known to the Grothendieck school in the 1960s, but not published. The equivalence should be regarded as a generalisation of the fact that congruences on a group correspond to normal subgroups of the group.)

A key discovery was that a category of double groupoids with one vertex and what we called 'special connections' [48] is equivalent to the category of crossed modules. Using these connections we could define what we called a 'commutative cube' in such a double groupoid. The key equation for this was:

$$c_1 = \begin{bmatrix} \ulcorner & a_0^{-1} & \urcorner \\ -b_0 & c_0 & b_1 \\ \llcorner & a_1 & \lrcorner \end{bmatrix}$$

which corresponded to folding flat five faces of a cube and filling in the corners with new 'canonical' elements which we called 'connections', because of a crucial 'transport law' which was borrowed from a paper on path connections in differential geometry, and can be written

$$\begin{bmatrix} \ulcorner & = \\ \shortparallel & \ulcorner \end{bmatrix} = \ulcorner .$$

The connections provide, additionally to the usual compositions, identities, and inverses, a structure which can be expressed intuitively by saying that in the 2-dimensional algebra of squares not only can you move forward and backwards, and stay still, but you can also turn left and right. For more details see for example [45, 2]. As you might imagine, there are problems in finding a formula in still higher dimensions. In the groupoid case, this is handled by a homotopy addition lemma and thin elements [30], but in the category case a formula for just a commutative 4-cube is complicated [70].

The blockage of defining a functor Π to double groupoids was resolved in 1974 in discussions with Higgins, by considering Whitehead's Theorem W. This showed that a 2-dimensional universal property was available in homotopy theory, which was encouraging; it also suggested that a theory to be any good should recover Theorem W. But this theorem was about *relative* homotopy groups. This suggested studying a relative situation $X_* : X_0 \subseteq X_1 \subseteq X$. On looking for the simplest way to get a homotopy functor from this situation using squares, the 'obvious' answer came up: consider maps $(I^2, \partial I^2, \partial\partial I^2) \to (X, X_1, X_0)$, i.e. maps of the square into X which take the edges into X_1 and the vertices into X_0, and then take homotopy classes rel vertices of such maps to form a set $\varrho_2(X_*)$. Of course this set will not inherit a group structure but the surprise is that it does inherit the structure of double groupoid with connections - the proof is not entirely trivial, and is given in [28] and the expository article [22]. In the case X_0 is a singleton, the equivalence of such double groupoids to crossed modules takes $\varrho(X_*)$ to the usual relative homotopy crossed module.

Thus a search for a *higher homotopy groupoid* was realised in dimension 2. It might be that a tendency of mathematicians to despise the notion of groupoid, as suggested in [56], contributed to such a construction not being found earlier.

Finding a good homotopy double groupoid led rather quickly, in view of the previous experience, to a substantial account of a 2-dimensional SVKT [28]. This recovers Theorem W, and also leads to new calculations in 2-dimensional homotopy theory, and in fact to some new calculations of 2-types. For a recent summary of some results and some new ones, see the paper in the J. Symbolic Computation [52] – publication in this journal illustrates that we are interested in using general methods in order to obtain specific calculations, and ones to which there seems no other route. For purely algebraic calculations, see [4].

Once the 2-dimensional case had been completed in 1975, it was easy to conjecture the form of general results for dimensions > 2, and announcements were made in [29] with full details in [30, 31]. However, these results needed a number of new ideas, even just to construct the higher dimensional compositions, and the proof of the Generalised SVKT was quite hard and intricate. Further, for applications, such as to explain how the general Π behaved on homotopies, we also needed a theory of tensor products, so that the resulting theory is quite complex. In the next section we give a summary.

4 Main results

Major features of the work over the years with Philip Higgins and others can be summarised in the following diagram of categories and functors:

Diagram 4.1

$$\begin{tikzcd}
& \text{filtered spaces} \arrow[dl, shift right, "\varrho"'] \arrow[dr, shift left, "\Pi"] & \\
\begin{array}{c}\text{cubical}\\ \omega\text{-groupoids}\\ \text{with connections}\end{array} \arrow[ur, shift right, "|\,|\circ U_*"'] \arrow[rr, shift left, "\gamma"] & & \begin{array}{c}\text{crossed}\\ \text{complexes}\end{array} \arrow[ul, shift left, "\mathcal{B}"] \arrow[ll, shift left, "\lambda"]
\end{tikzcd}$$

in which

4.1.1 the categories FTop of filtered spaces, ω-Gpd of ω-groupoids, and Crs of crossed complexes are monoidal closed, and have a notion of homotopy using $\otimes$ and a unit interval object;
4.1.2 ϱ, Π are homotopical functors (that is they are defined in terms of homotopy classes of certain maps), and preserve homotopies;
4.1.3 λ, γ are inverse adjoint equivalences of monoidal closed categories;
4.1.4 there is a natural equivalence $\gamma\varrho \simeq \Pi$, so that either ϱ or Π can be used as appropriate;
4.1.5 ϱ, Π preserve certain colimits and certain tensor products;
4.1.6 by definition, the *cubical filtered classifying space* is $\mathcal{B}^{\square} = |\,| \circ U_*$ where U_* is the forgetful functor to filtered cubical sets using the filtration of an ω-groupoid by skeleta, and $|\,|$ is geometric realisation of a cubical set;
4.1.7 there is a natural equivalence $\Pi \circ \mathcal{B}^{\square} \simeq 1$;
4.1.8 if C is a crossed complex and its cubical classifying space is defined as $B^{\square}C = (\mathcal{B}^{\square}C)_{\infty}$, then for a CW-complex X, and using homotopy as in 4.1.1 for crossed complexes, there is a natural bijection of sets of homotopy classes

$$[X, B^{\square}C] \cong [\Pi(X_*), C].$$

Here a *filtered space* consists of a (compactly generated) space X_∞ and an increasing sequence of subspaces

$$X_* : X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X_\infty.$$

With the obvious morphisms, this gives the category FTop. The tensor product in this category is the usual

$$(X_* \otimes Y_*)_n = \bigcup_{p+q=n} X_p \times Y_q.$$

The closed structure is easy to construct from the law

$$\mathsf{FTop}(X_* \otimes Y_*, Z_*) \cong \mathsf{FTop}(X_*, \mathsf{FTOP}(Y_*, Z_*)).$$

An advantage of this monoidal closed structure is that it allows an enrichment of the category FTop over either crossed complexes or ω-Gpd using Π or ϱ applied to $\mathsf{FTOP}(Y_*, Z_*)$.

The structure of *crossed complex* is suggested by the canonical example, the *fundamental crossed complex* $\Pi(X_*)$ of the filtered space X_*. So it is given by a diagram

Diagram 4.2

$$\begin{tikzcd}
\cdots \arrow[r] & C_n \arrow[r, "\delta_n"] \arrow[d, "t"] & C_{n-1} \arrow[r] \arrow[d, "t"] & \cdots \arrow[r] & C_2 \arrow[r, "\delta_2"] \arrow[d, "t"] & C_1 \arrow[d, shift right, "s"'] \arrow[d, shift left, "t"] \\
& C_0 & C_0 & & C_0 & C_0
\end{tikzcd}$$

in which in this example C_1 is the fundamental groupoid $\pi_1(X_1, X_0)$ of X_1 on the 'set of base points' $C_0 = X_0$, while for $n \geqslant 2$, C_n is the family of relative homotopy groups $\{C_n(x)\} = \{\pi_n(X_n, X_{n-1}, x) \mid x \in X_0\}$. The boundary maps are those standard in homotopy theory. There is for $n \geqslant 2$ an action of the groupoid C_1 on C_n (and of C_1 on the groups $C_1(x)$, $x \in X_0$ by conjugation), the boundary morphisms are operator morphisms, $\delta_{n-1}\delta_n = 0$, $n \geqslant 3$, and the additional axioms are satisfied that

4.3 $b^{-1}cb = c^{\delta_2 b}$, $b, c \in C_2$, so that $\delta_2 : C_2 \to C_1$ is a crossed module (of groupoids),

4.4 for $n \geqslant 3$, each group $C_n(x)$ is abelian, and with trivial action by $\delta_2(C_2(x))$, so that the family C_n becomes a $C_1/\delta_2(C_2)$-module.

Clearly we obtain a category Crs of crossed complexes; this category is not so familiar and so we give arguments for using it in the next section.

As algebraic examples of crossed complexes we have: $C = \mathbb{C}(G, n)$ where G is a group, commutative if $n \geqslant 2$, and C is G in dimension n and trivial elsewhere; $C = \mathbb{C}(G : M, n)$, where G is a group, M is a G-module, $n \geqslant 2$, and C is G in dimension 1, M in dimension n, trivial elsewhere, and zero boundary if $n = 2$; C is a crossed module (of groups) in dimensions 1 and 2 and trivial elsewhere.

A crossed complex C has a fundamental groupoid $\pi_1 C = C_1/\operatorname{Im}\delta_2$, and also for $n \geqslant 2$ a family $\{H_n(C, p) \mid p \in C_0\}$ of homology groups.

5 Why crossed complexes?

- They generalise groupoids and crossed modules to all dimensions. Note that the natural context for second relative homotopy groups is crossed modules of groupoids, rather than groups.
- They are good for modelling CW-complexes.
- Free crossed resolutions enable calculations with small CW-models of $K(G,1)$s *and their maps* (Whitehead, Wall, Baues).
- Crossed complexes give a kind of 'linear model' of homotopy types which includes all 2-types. Thus although they are not the most general model by any means (they do not contain quadratic information such as Whitehead products), this simplicity makes them easier to handle and to relate to classical tools. The new methods and results obtained for crossed complexes can be used as a model for more complicated situations. This is how a general n-adic Hurewicz Theorem was found [41].
- They are convenient for *calculation*, and the functor Π is classical, involving relative homotopy groups. We explain some results in this form later.
- They are close to chain complexes with a group(oid) of operators, and related to some classical homological algebra (e.g. *chains of syzygies*). In fact if SX is the simplicial singular complex of a space, with its skeletal filtration, then the crossed complex $\Pi(SX)$ can be considered as a slightly non commutative version of the singular chains of a space.
- The monoidal structure is suggestive of further developments (e.g. *crossed differential algebras*) see [12, 11, 10]. It is used in [23] to give an algebraic model of homotopy 3-types, and to discuss automorphisms of crossed modules.
- Crossed complexes have a good homotopy theory, with a *cylinder object, and homotopy colimits.* The homotopy classification result 4.1.8 generalises a classical theorem of Eilenberg-Mac Lane.
- They have an interesting relation with the Moore complex of simplicial groups and of simplicial groupoids (see section 18).

6 Why cubical ω-groupoids with connections?

The definition of these objects is more difficult to give, but will be indicated later. Here we explain why we need to introduce such new structures.

• The functor ϱ gives a form of *higher homotopy groupoid*, thus realising the dreams of the early topologists.

• They are equivalent to crossed complexes.

• They have a clear *monoidal closed structure*, and a notion of homotopy, from which one can deduce those on crossed complexes, using the equivalence of categories.

• It is easy to relate the functor ϱ to tensor products, but quite difficult to do this directly for Π.

• Cubical methods, unlike globular or simplicial methods, allow for a simple *algebraic inverse to subdivision*, which is crucial for our local-to-global theorems.

• The additional structure of 'connections', and the equivalence with crossed complexes, allows for the sophisticated notion of *commutative cube*, and the proof that *multiple compositions of commutative cubes are commutative.* The last fact is a key component of the proof of the GSVKT.

• They yield a construction of a *(cubical) classifying space* $B^{\square}C = (\mathcal{B}^{\square}C)_{\infty}$ of a crossed complex C, which generalises (cubical) versions of Eilenberg-Mac Lane spaces, including the local coefficient case. This has convenient relation to homotopies.

• There is a current *resurgence of the use of cubes* in for example combinatorics, algebraic topology, and concurrency. There is a Dold-Kan type theorem for cubical abelian groups with connections [36].

7 The equivalence of categories

Let Crs, ω-Gpd denote respectively the categories of crossed complexes and ω-groupoids. A major part of the work in [30] consists in defining these categories and proving their equivalence, which thus gives an example of two algebraically defined categories whose equivalence is non trivial. It is even more subtle than that because the functors $\gamma : \mathsf{Crs} \to \omega{-}\mathsf{Gpd}$, $\lambda : \omega{-}\mathsf{Gpd} \to \mathsf{Crs}$ are not hard to define, and it is easy to prove $\gamma\lambda \simeq 1$. The hard part is to prove $\lambda\gamma \simeq 1$, which shows that an ω-groupoid G may be reconstructed from the crossed complex $\gamma(G)$ it contains. The proof involves using the connections to construct a 'folding map' $\Phi : G_n \to G_n$, and establishing its major properties, including the relations with the compositions. This gives an algebraic form of some old intuitions of several ways of defining relative homotopy groups, for example using cubes or cells.

On the way we establish properties of *thin elements*, as those which fold down to 1, and show that G satisfies a strong Kan condition, namely that every box has a unique thin filler. This result plays a key role in the proof of the GSVKT for ϱ, since it is used to show an independence of choice. That part of the proof goes by showing that the two choices can be seen, since we start with a homotopy, as given by the two ends $\partial^{\pm}_{n+1}x$ of an $(n+1)$-cube x. It is then shown by induction, using the method of construction and the above result, that x is degenerate in direction $n+1$. Hence the two ends in that direction coincide.

Properties of the folding map are used also in showing in [31] that $\Pi(X_*)$ is actually included in $\varrho(X_*)$; in relating two types of thinness for elements of $\varrho(X_*)$; and in proving a *homotopy addition lemma* in $\varrho(X_*)$.

Any ω-Gpd G has an underlying cubical set UG. If C is a crossed complex, then the cubical set $U(\lambda C)$ is called the *cubical nerve* $N^{\square}C$ of C. It is a conclusion of the theory that we can also obtain $N^{\square}C$ as

$$(N^{\square}C)_n = \mathsf{Crs}(\Pi(I^n_*), C)$$

where I^n_* is the usual geometric cube with its standard skeletal filtration. The (cubical) geometric realisation $|N^{\square}C|$ is also called the *cubical classifying space* $B^{\square}C$ of the crossed complex C. The filtration C^* of C by skeleta gives a filtration $\mathbb{B}C = B^{\square}C^*$ of $B^{\square}C$ and there is (as in 4.1.7) a natural isomorphism $\Pi(B^{\square}C^*) \cong C$. Thus the properties of a crossed complex are those that are universally satisfied by $\Pi(X_*)$. These proofs use the equivalence of the homotopy categories of cubical Kan complexes and of CW-complexes. We originally took this from the Warwick Masters thesis of S. Hintze, but it is now available with a different proof from R. Antolini [5, 6].

As said above, by taking particular values for C, the classifying space $B^{\square}C$ gives cubical versions of Eilenberg-Mac Lane spaces $K(G,n)$, including the case $n = 1$ and G non commutative. If C is essentially a crossed module, then $B^{\square}C$ is called the *cubical classifying space* of the crossed module, and in fact realises the k-invariant of the crossed module.

Another useful result is that if K is a cubical set, then $\varrho(|K|_*)$ may be identified with $\varrho(K)$, the *free* ω-Gpd *on the cubical set* K, where here $|K|_*$ is the usual filtration by skeleta. On the other hand, our proof that $\Pi(|K|_*)$ is the free crossed complex on the non-degenerate cubes of K uses the generalised SVKT of the next section.

It is also possible to give simplicial and globular versions of some of the above results, because the category of crossed complexes is equivalent also to those of simplicial T-complexes [7] and of globular ∞-groupoids [32]. In fact the published paper on the classifying space of a crossed complex [35] is given in simplicial terms, in order to link more easily with well known theories. This is helpful in the equivariant results in [25, 26].

8 Main aim of the work: colimit, or local-to-global, theorems

These theorems give *non commutative tools for higher dimensional local-to-global problems* yielding a variety of new, often non commutative, calculations, which *prove* (i.e. test) the theory. We now explain these theorems in a way which strengthens the relation with descent.

We suppose given an open cover $\mathcal{U} = \{U^\lambda\}_{\lambda\in\Lambda}$ of X. This cover defines a map

$$q : E = \bigsqcup_{\lambda\in\Lambda} U^\lambda \to X$$

and so we can form an augmented simplicial space

$$\check{\mathrm{C}}(q) : \quad \cdots \qquad E \times_X E \times_X E \; \substack{\longrightarrow\\[-2pt]\longrightarrow\\[-2pt]\longrightarrow} \; E \times_X E \; \substack{\longrightarrow\\[-2pt]\longrightarrow} \; E \xrightarrow{q} X$$

where the higher dimensional terms involve disjoint unions of multiple intersections U^ν of the U^λ.

We now suppose given a filtered space X_*, a cover $\mathcal{U}$ as above of $X = X_\infty$, and so an augmented simplicial filtered space $\check{\mathrm{C}}(q_*)$ involving multiple intersections U^ν_* of the induced filtered spaces.

We still need a connectivity condition.

Definition 8.1 A filtered space X_* is *connected* if and only if for all $n > 0$ the induced map $\pi_0 X_0 \to \pi_0 X_n$ is surjective and for all $r > n > 0$ and $x \in X_0, \pi_n(X_r, X_n, x) = 0$.

Theorem 8.2 (MAIN RESULT (GSVKT)) *If U^ν_* is connected for all finite intersections U^ν of the elements of the open cover, then*
(C) (connectivity) *X_* is connected, and*
(I) (isomorphism) *the following diagram as part of $\varrho(\check{C}(q_*))$*

$$\varrho(E_* \times_{X_*} E_*) \rightrightarrows \varrho(E_*) \xrightarrow{\varrho(q_*)} \varrho(X_*). \qquad (\mathrm{c}\varrho)$$

is a coequaliser diagram. Hence the following diagram of crossed complexes

$$\Pi(E_* \times_{X_*} E_*) \rightrightarrows \Pi(E_*) \xrightarrow{\Pi(q_*)} \Pi(X_*). \qquad (\mathrm{c}\Pi)$$

is also a coequaliser diagram.

So we get calculations of the fundamental crossed complex $\Pi(X_*)$.

It should be emphasised that to get to and apply this theorem takes the two papers [30, 31], of 58 pages together. With this we deduce in the first instance:

- the usual SVKT for the fundamental groupoid on a set of base points;
- the Brouwer degree theorem ($\pi_n S^n = \mathbb{Z}$);
- the relative Hurewicz theorem;
- Whitehead's theorem that $\pi_n(X \cup \{e^2_\lambda\}, X)$ is a free crossed module;
- a more general excision result on $\pi_n(A \cup B, A, x)$ as an induced module (crossed module if $n = 2$) when $(A, A \cap B)$ is $(n-1)$-connected.

The assumptions required of the reader are quite small, just some familiarity with CW-complexes. This contrasts with some expositions of basic homotopy theory, where the proof of say the relative Hurewicz theorem requires knowledge of singular homology theory. Of course it is surprising to get this last theorem without homology, but this is because it is seen as a statement on the morphism of relative homotopy groups

$$\pi_n(X, A, x) \to \pi_n(X \cup CA, CA, x) \cong \pi_n(X \cup CA, x)$$

and is obtained, like our proof of Theorem W, as a special case of an excision result. The reason for this success is that we use algebraic structures which model the geometry and underlying processes more closely than those in common use.

Note also that these results cope well with the action of the fundamental group on higher homotopy groups.

The calculational use of the GSVKT for $\Pi(X_*)$ is enhanced by the relation of Π with tensor products (see section 15 for more details).

9 The fundamental cubical ω–groupoid $\varrho(X_*)$ of a filtered space X_*

Here are the basic elements of the construction.

I^n_*: the n-cube with its skeletal filtration.

Set $R_n(X_*) = \mathsf{FTop}(I^n_*, X_*)$. This is a *cubical set with compositions, connections, and inversions.*

For $i = 1, \dots, n$ there are standard:
face maps $\partial^\pm_i : R_n X_* \to R_{n-1} X_*$;
degeneracy maps $\varepsilon_i : R_{n-1} X_* \to R_n X_*$
connections $\Gamma^\pm_i : R_{n-1} X_* \to R_n X_*$

compositions $a \circ_i b$ defined for $a, b \in R_n X_*$ such that $\partial_i^+ a = \partial_i^- b$
inversions $-_i : R_n \to R_n$.

The connections are induced by $\gamma_i^\alpha : I^n \to I^{n-1}$ defined using the monoid structures $\max, \min : I^2 \to I$. They are essential for many reasons, e.g. to discuss the notion of *commutative cube.*

These operations have certain algebraic properties which are easily derived from the geometry and which we do not itemise here – see for example [2, 72]. These were listed first in the Bangor thesis of Al-Agl [1]. (In the paper [30] the only connections needed are the Γ_i^+, from which the Γ_i^- are derived using the inverses of the groupoid structures.)

Definition 9.1

$$p : R_n(X_*) \to \varrho_n(X_*) = (R_n(X_*)/\equiv)$$

is the quotient map, where $f \equiv g \in R_n(X_*)$ means *filter homotopic (i.e. through filtered maps) rel vertices of* I^n.

The following results are proved in [31].

9.2 The compositions on R are inherited by ϱ to give $\varrho(X_*)$ the structure of cubical multiple groupoid with connections.

9.3 The map $p : R(X_*) \to \varrho(X_*)$ is a Kan fibration of cubical sets.

The proofs of both results use methods of collapsing which are indicated in the next section. The second result is almost unbelievable. Its proof has to give a systematic method of deforming a cube with the right faces 'up to homotopy' into a cube with exactly the right faces, using the given homotopies. In both cases, the assumption that the relation $\equiv$ uses homotopies rel vertices is essential to start the induction. (In fact the paper [31] does not use homotopy rel vertices, but imposes an extra condition J_0, that each loop in X_0 is contractible X_1. A full exposition of the whole story is in preparation.)

Here is an application which is essential in many proofs.

Theorem 9.4 (Lifting multiple compositions) *Let* $[\alpha_{(r)}]$ *be a multiple composition in* $\varrho_n(X_*)$. *Then representatives* $a_{(r)}$ *of the* $\alpha_{(r)}$ *may be chosen so that the multiple composition* $[a_{(r)}]$ *is well defined in* $R_n(X_*)$.

Proof: The multiple composition $[\alpha_{(r)}]$ determines a cubical map

$$A : K \to \varrho(X_*)$$

where the cubical set K corresponds to a representation of the multiple composition by a subdivision of the geometric cube, so that top cells $c_{(r)}$ of K are mapped by A to $\alpha_{(r)}$.

Consider the diagram, in which $*$ is a corner vertex of K,

$$\begin{tikzcd} * \arrow[r] \arrow[d] & R(X_*) \arrow[d, "p"] \\ K \arrow[r, "A"'] \arrow[ur, dashed, "A'"] & \varrho(X_*) \end{tikzcd}$$

Then K collapses to $*$, written $K \searrow *$. By the fibration result, A lifts to A', which represents $[a_{(r)}]$, as required. □

So we have to explain collapsing.

10 Collapsing

We use a basic notion of collapsing and expanding due to J.H.C. Whitehead.

Let $C \subseteq B$ be subcomplexes of I^n. We say C is an *elementary collapse* of B, $B \overset{e}{\searrow} C$, if for some $s \geqslant 1$ there is an s-cell a of B and $(s-1)$-face b of a, the *free face*, such that

$$B = C \cup a, \qquad C \cap a = \partial a \setminus b$$

(where ∂a denotes the union of the proper faces of a).

We say $B_1 \searrow B_r$, B_1 *collapses* to B_r, if there is a sequence

$$B_1 \overset{e}{\searrow} B_2 \overset{e}{\searrow} \cdots \overset{e}{\searrow} B_r$$

of elementary collapses.

If C is a subcomplex of B then

$$B \times I \searrow (B \times \{0\} \cup C \times I)$$

(this is proved by induction on the dimension of $B \setminus C$).

Further I^n collapses to any one of its vertices (this may be proved by induction on n using the first example). These collapsing techniques are crucial for proving 9.2, that $\varrho(X_*)$ does obtain the structure of multiple groupoid, since it allows the construction of the extensions of filtered maps and filtered homotopies that are required.

However, more subtle collapsing techniques using partial boxes are required to prove the fibration theorem 9.3, as partly explained in the next section.

11 Partial boxes

Let C be an r-cell in the n-cube I^n. Two $(r-1)$-faces of C are called *opposite* if they do not meet.

A partial box in C is a subcomplex B of C generated by one $(r-1)$-face b of C (called a *base* of B) and a number, possibly zero, of other $(r-1)$-faces of C none of which is opposite to b.

The partial box is a box if its $(r-1)$-cells consist of all but one of the $(r-1)$-faces of C.

The proof of the fibration theorem uses a filter homotopy extension property and the following:

Proposition 11.1 Key Proposition: Let B, B' be partial boxes in an r-cell C of I^n such that $B' \subseteq B$. Then there is a chain

$$B = B_s \searrow B_{s-1} \searrow \cdots \searrow B_1 = B'$$

such that

(i) each B_i is a partial box in C
(ii) $B_{i+1} = B_i \cup a_i$ where a_i is an $(r-1)$-cell of C not in B_i;
(iii) $a_i \cap B_i$ is a partial box in a_i.

The proof is quite neat, and follows the pictures.

Induction up a such a chain of partial boxes is one of the steps in the proof of the fibration theorem 9.3.

The proof of the fibration theorem gives a program for carrying out the deformations needed to do the lifting. In some sense, it implies computing a multiple composition as in Theorem 9.4 can be done using collapsing as the guide.

Methods of collapsing are related to notions of shelling in [82], and of course any finite tree collapses to a point.

12 Thin elements

Another key concept is that of *thin element* $\alpha \in \varrho_n(X_*)$ for $n \geqslant 2$. The proofs here use strongly the algebraic results of [30].

We say α is *geometrically thin* if it has a *deficient* representative, i.e. an $a : I^n_* \to X_*$ such that $a(I^n) \subseteq X_{n-1}$.

We say α is *algebraically thin* if it is a multiple composition of degenerate elements or those coming from repeated negatives of connections. Clearly any composition of algebraically thin elements is thin.

Theorem 12.1 *(i) algebraically thin $\equiv$ geometrically thin.*
(ii) In a cubical ω–groupoid with connections, any box has a unique thin filler.

Proof The proof of the forward implication in (i) uses lifting of multiple compositions, in a stronger form than stated above.

The proofs of (ii) and the backward implication in (i) uses the full force of the algebraic relation between ω–groupoids and crossed complexes. □

These results allow one to replace arguments with commutative cubes by arguments with thin elements.

13 Sketch proof of the GSVKT

We go back to the following diagram whose top row is part of $\varrho(\check{C}(q_*))$

$$\begin{tikzcd} \varrho(E_* \times_{X_*} E_*) \arrow[r, shift left, "\partial_0"] \arrow[r, shift right, "\partial_1"'] & \varrho(E_*) \arrow[r, "\varrho(q_*)"] \arrow[dr, "f"'] & \varrho(X_*) \arrow[d, dashed, "f'"] \\ & & G \end{tikzcd} \qquad (\mathrm{c}\varrho)$$

To prove this top row is a coequaliser diagram, we suppose given a morphism $f : \varrho(E_*) \to G$ of cubical ω-groupoids with connection such that $f \circ \partial_0 = f \circ \partial_1$, and prove that there is unique $f' : \varrho(X_*) \to G$ such that $f' \circ \varrho(q_*) = f$.

To define $f'(\alpha)$ for $\alpha \in \varrho(X_*)$, you subdivide a representative a of α to give $a = [a_{(r)}]$ so that each $a_{(r)}$ lies in an element $U^{(r)}$ of $\mathcal{U}$; use the connectivity conditions and this subdivision to deform a into $b = [b_{(r)}]$ so that

$$b_{(r)} \in R(U^{(r)}_*)$$

and so obtain

$$\beta_{(r)} \in \varrho(U^{(r)}_*).$$

The elements

$$f\beta_{(r)} \in G$$

may be composed in G (by the conditions on f), to give an element

$$\theta(\alpha) = [f\beta_{(r)}] \in G.$$

So the proof of the universal property has to use an *algebraic inverse to subdivision.* Again an analogy here is with sending an email: the element you start with is subdivided, deformed so that each part is correctly labelled, the separate parts are sent, and then recombined.

The proof that $\theta(\alpha)$ is independent of the choices involved makes crucial use of properties of thin elements. The key point is: *a filter homotopy $h : \alpha \equiv \alpha'$ in $R_n(X_*)$ gives a deficient element of $R_{n+1}(X_*)$.*

The method is to do the subdivision and deformation argument on such a homotopy, push the little bits in some

$$\varrho_{n+1}(U_*^\lambda)$$

(now thin) over to G, combine them and get a thin element

$$\tau \in G_{n+1}$$

all of whose faces not involving the direction $(n+1)$ are thin *because h was given to be a filter homotopy.* An inductive argument on unique thin fillers of boxes then shows that τ *is degenerate in direction* $(n+1)$, so that the two ends in direction $(n+1)$ are the same.

This ends a rough sketch of the proof of the GSVKT for ϱ.

Note that the theory of these forms of multiple groupoids is designed to make this last argument work. We replace a formula for saying a cube h has commutative boundary by a statement that h is thin. It would be very difficult to replace the above argument, on the composition of thin elements, by a higher dimensional manipulation of formulae such as that given in section 3 for a commutative 3-cube.

14 Tensor products and homotopies

The construction of the monoidal closed structure on the category ω-Gpd is based on rather formal properties of cubical sets, and the fact that for the cubical set $\mathbb{I}^n$ we have $\mathbb{I}^m \otimes \mathbb{I}^n \cong \mathbb{I}^{m+n}$. The details are given in [33]. The equivalence of categories implies then that the category Crs is also monoidal closed, with a natural isomorphism

$$\mathsf{Crs}(A \otimes B, C) \cong \mathsf{Crs}(A, \mathsf{CRS}(B, C)).$$

Here the elements of $\mathsf{CRS}(B, C)$ are in dimension 0 the morphisms $B \to C$, in dimension 1 the *left homotopies of morphisms*, and in higher dimensions are forms of higher homotopies. The precise description of these is obtained of course by tracing out in detail the equivalence of categories. It should be emphasised that certain choices are made in constructing this equivalence, and these choices are reflected in the final formulae that are obtained.

An important result is that if X_*, Y_* are filtered spaces, then there is a natural transformation

$$\begin{aligned} \eta \,:\, \varrho(X_*) \otimes \varrho(Y_*) &\to \varrho(X_* \otimes Y_*) \\ [a] \otimes [b] &\mapsto [a \otimes b] \end{aligned}$$

where if $a : I_*^m \to X_*$, $b : I_*^n \to Y_*$ then $a \otimes b : I_*^{m+n} \to X_* \otimes Y_*$. It not hard to see, in this cubical setting, that η is well defined. It can also be shown using previous results that η is an isomorphism if X_*, Y_* are the geometric realisations of cubical sets with the usual skeletal filtration.

The equivalence of categories now gives a natural transformation of crossed complexes

$$\eta' : \Pi(X_*) \otimes \Pi(Y_*) \to \Pi(X_* \otimes Y_*). \tag{14.1}$$

It would be hard to construct this directly. It is proved in [35] that η' is an isomorphism if X_*, Y_* are the skeletal filtrations of CW-complexes. The proof uses the GSVKT, and the fact that $A \otimes -$ on crossed complexes has a right adjoint and so preserves colimits. It is proved in [10] that η is an isomorphism if X_*, Y_* are cofibred, connected filtered spaces. This applies in particular to the useful case of the filtration $B^\square C^*$ of the classifying space of a crossed complex.

It turns out that the defining rules for the tensor product of crossed complexes which follows from the above construction are obtained as follows. We first define a bimorphism of crossed complexes.

Definition 14.1 A *bimorphism* $\theta : (A, B) \to C$ of crossed complexes is a family of maps $\theta : A_m \times B_n \to C_{m+n}$ satisfying the following conditions, where $a \in A_m, b \in B_n, a_1 \in A_1, b_1 \in B_1$ (temporarily using additive notation throughout the definition):

(i)

$$\beta(\theta(a,b)) = \theta(\beta a, \beta b) \text{ for all } a \in A, b \in B\,.$$

(ii)

$$\begin{aligned}\theta(a, b^{b_1}) &= \theta(a,b)^{\theta(\beta a, b_1)} \text{ if } m \geqslant 0, n \geqslant 2\,,\\ \theta(a^{a_1}, b) &= \theta(a,b)^{\theta(a_1, \beta b)} \text{ if } m \geqslant 2, n \geqslant 0\,.\end{aligned}$$

(iii)

$$\theta(a, b+b') = \begin{cases} \theta(a,b) + \theta(a,b') & \text{if } m = 0, n \geqslant 1 \text{ or } m \geqslant 1, n \geqslant 2\,,\\ \theta(a,b)^{\theta(\beta a, b')} + \theta(a,b') & \text{if } m \geqslant 1, n = 1\,,\end{cases}$$

$$\theta(a+a', b) = \begin{cases} \theta(a,b) + \theta(a',b) & \text{if } m \geqslant 1, n = 0 \text{ or } m \geqslant 2, n \geqslant 1\,,\\ \theta(a',b) + \theta(a,b)^{\theta(a', \beta b)} & \text{if } m = 1, n \geqslant 1\,.\end{cases}$$

(iv)

$$\delta_{m+n}(\theta(a,b)) = \begin{cases} \theta(\delta_m a, b) + (-)^m \theta(a, \delta_n b) & \text{if } m \geqslant 2, n \geqslant 2\,,\\ -\theta(a, \delta_n b) - \theta(\beta a, b) + \theta(\alpha a, b)^{\theta(a, \beta b)} & \text{if } m = 1, n \geqslant 2\,,\\ (-)^{m+1}\theta(a, \beta b) + (-)^m \theta(a, \alpha b)^{\theta(\beta a, b)} + \theta(\delta_m a, b) & \text{if } m \geqslant 2, n = 1\,,\\ -\theta(\beta a, b) - \theta(a, \alpha b) + \theta(\alpha a, b) + \theta(a, \beta b) & \text{if } m = n = 1\,.\end{cases}$$

(v)

$$\delta_{m+n}(\theta(a,b)) = \begin{cases} \theta(a, \delta_n b) & \text{if } m = 0, n \geqslant 2\,,\\ \theta(\delta_m a, b) & \text{if } m \geqslant 2, n = 0\,.\end{cases}$$

(vi)

$$\begin{aligned}
\alpha(\theta(a,b)) &= \theta(a,\alpha b) \quad \text{and} \quad \beta(\theta(a,b)) = \theta(a,\beta b) \quad \text{if } m=0, n=1\,,\\
\alpha(\theta(a,b)) &= \theta(\alpha a,b) \quad \text{and} \quad \beta(\theta(a,b)) = \theta(\beta a,b) \quad \text{if } m=1, n=0\,.
\end{aligned}$$

The *tensor product* of crossed complexes A, B is given by the universal bimorphism $(A, B) \to A\otimes B$, $(a, b) \mapsto a\otimes b$. The rules for the tensor product are obtained by replacing $\theta(a, b)$ by $a \otimes b$ in the above formulae.

The conventions for these formulae for the tensor product arise from the derivation of the tensor product via the category of cubical ω-groupoids with connections, and the formulae are forced by our conventions for the equivalence of the two categories [30, 33].

The complexity of these formulae is directly related to the complexities of the cell structure of the product $E^m \times E^n$ where the n-cell E^n has cell structure e^0 if $n = 0$, $e^0_\pm \cup e^1$ if $n = 1$, and $e^0 \cup e^{n-1} \cup e^n$ if $n \geqslant 2$.

It is proved in [33] that the bifunctor $- \otimes -$ is symmetric and that if a_0 is a vertex of A then the morphism $B \to A \otimes B$, $b \to a_0 \otimes b$, is injective.

There is a standard groupoid model I of the unit interval, namely the indiscrete groupoid on two objects $0, 1$. This is easily extended trivially to either a crossed complex or an ω-Gpd. So using $\otimes$ we can define a 'cylinder object' $\mathsf{I} \otimes -$ in these categories and so a homotopy theory (cf. [84]).

15 Free crossed complexes and free crossed resolutions

Let C be a crossed complex. A *free basis* B_* for C consists of the following:
B_0 is set which we take to be C_0;
B_1 is a graph with source and target maps $s, t : B_1 \to B_0$ and C_1 is the free groupoid on the graph B_1: that is B_1 is a subgraph of C_1 and any graph morphism $B_1 \to G$ to a groupoid G extends uniquely to a groupoid morphism $C_1 \to G$;
for $n \geqslant 2$, B_n is a totally disconnected subgraph of C_n with target map $t : B_n \to B_0$; for $n = 2$, C_2 is the free crossed C_1-module on B_2 while for $n > 2$, C_n is the free $(\pi_1 C)$-module on B_n.

It may be proved using the GSVKT that if X_* is a CW-complex with the skeletal filtration, then $\Pi(X_*)$ is the free crossed complex on the characteristic maps of the cells of X_*. It is proved in [35] that the tensor product of free crossed complexes is free.

A *free crossed resolution* F_* of a groupoid G is a free crossed complex which is aspherical together with an isomorphism $\phi : \pi_1(F_*) \to G$. Analogues of standard methods of homological algebra show that free crossed resolutions of a group are unique up to homotopy equivalence.

In order to apply this result to free crossed resolutions, we need to replace free crossed resolutions by CW-complexes. A fundamental result for this is the following, which goes back to Whitehead [106] and Wall [101], and which is discussed further by Baues in [9, Chapter VI, §7]:

Theorem 15.1 *Let X_* be a CW-filtered space, and let $\phi : \pi X_* \to C$ be a homotopy equivalence to a free crossed complex with a preferred free basis. Then there is a CW-filtered space Y_*, and an isomorphism $\pi Y_* \cong C$ of crossed complexes with preferred basis, such that ϕ is realised by a homotopy equivalence $X_* \to Y_*$.*

In fact, as pointed out by Baues, Wall states his result in terms of chain complexes, but the crossed complex formulation seems more natural, and avoids questions of realisability in dimension 2, which are unsolved for chain complexes.

Corollary 15.2 If A is a free crossed resolution of a group G, then A is realised as free crossed complex with preferred basis by some CW-filtered space Y_*.

Proof We only have to note that the group G has a classifying CW-space BG whose fundamental crossed complex $\Pi(BG)$ is homotopy equivalent to A. □

Baues also points out in [9, p.657] an extension of these results which we can apply to the realisation of morphisms of free crossed resolutions.

Proposition 15.3 Let $X = K(G, 1)$, $Y = K(H, 1)$ be CW-models of Eilenberg - Mac Lane spaces and let $h : \Pi(X_*) \to \Pi(Y_*)$ be a morphism of their fundamental crossed complexes with the preferred bases given by skeletal filtrations. Then $h = \Pi(g)$ for some cellular $g : X \to Y$.

Proof Certainly h is homotopic to $\Pi(f)$ for some $f : X \to Y$ since the set of pointed homotopy classes $X \to Y$ is bijective with the morphisms of groups $A \to B$. The result follows from [9, p.657,(**)] ('if f is Π-realisable, then each element in the homotopy class of f is Π-realisable'). □

These results are exploited in [91, 44], to calculate free crossed resolutions of the fundamental groupoid of a graph of groups.

An algorithmic approach to the calculation of free crossed resolutions for groups is given in [47], by constructing partial contracting homotopies for the universal cover at the same time as constructing this universal cover inductively. This theme leads to, indeed necessitates, new ideas in rewriting, as shown in [73, 74]. A practical application in group theory is: given a presentation $\langle X \mid R\rangle$ of a group, write a given element of $N(R)$ as a consequence of the relators.

16 Classifying spaces and homotopy classification of maps

The formal relations of cubical sets and of cubical ω-groupoids with connections and the relation of Kan cubical sets with topological spaces, allow the proof of a homotopy classification theorem:

Theorem 16.1 *If K is a cubical set, and G is an ω-groupoid, then there is a natural bijection of sets of homotopy classes*

$$[|K|, |UG|] \cong [\varrho(|K|_*), G],$$

where on the left hand side we work in the category of spaces, and on the right in ω-groupoids.

Here $|K|_*$ is the filtration by skeleta of the geometric realisation of the cubical set.

We explained earlier how to define a cubical classifying space say $B^\square(C)$ of a crossed complex C as $B^\square(C) = |UN^\square C| = |U\lambda C|$. The properties already stated now give the homotopy classification theorem 4.1.7.

It is shown in [31] that if Y is a connected CW-complex, then there is a map $p : Y \to B^\square \Pi Y_*$ whose homotopy fibre is n-connected if $\pi_i Y = 0$ for $2 \leqslant i \leqslant n-1$.

It follows that if also X is a connected CW-complex with $\dim X \leqslant n$, then p induces a bijection

$$[X, Y] \to [X, B\Pi Y_*].$$

So under these circumstances we get a bijection

$$[X, Y] \to [\Pi X_*, \Pi Y_*]. \tag{16.1}$$

This result, due to Whitehead [105], translates a topological homotopy classification problem to an algebraic one. We explain below how this result can be translated to a result on chain complexes with operators.

It is also possible to define a simplicial nerve $N^\Delta(C)$ of a crossed complex C by

$$N^\Delta(C)_n = \mathsf{Crs}(\Pi(\Delta^n), C).$$

The *simplicial classifying space* of C is then defined by

$$B^\Delta(C) = |N^\Delta(C)|.$$

The properties of this simplicial classifying space are developed in [35], and in particular an analogue of 4.1.7 is proved. An application of the classifying space of a crossed module is given in [65].

The simplicial nerve and an adjointness

$$\mathsf{Crs}(\Pi(L), C) \cong \mathsf{Simp}(L, N^\Delta(C))$$

are used in [25, 26] for an equivariant homotopy theory of crossed complexes and their classifying spaces. Important ingredients in this are notions of homotopy coherence from [57] and an Eilenberg-Zilber type theorem for crossed complexes proved in Tonks' Bangor thesis [99, 100].

Labesse in [85] defines a *crossed set.* In fact a crossed set is exactly a crossed module $\delta : C \to X \rtimes G$ where G is a group acting on the set X, and $X \rtimes G$ is the associated actor groupoid; thus the simplicial construction from a crossed set described by Larry Breen in [85] is exactly the nerve of the crossed module, regarded as a crossed complex. Hence the cohomology with coefficients in a crossed set used in [85] is a special case of cohomology with coefficients in a crossed complex, dealt with in [35]. (We are grateful to Breen for pointing this out to us in 1999.)

17 Relation with chain complexes with a groupoid of operators

Chain complexes with a group of operators are a well known tool in algebraic topology, where they arise naturally as the chain complex $C_*\widetilde{X}_*$ of cellular chains of the universal cover $\widetilde{X}_*$ of a reduced CW-complex X_*. The group of operators here is the fundamental group of the space X.

J.H.C. Whitehead in [105] gave an interesting relation between his free crossed complexes (he called them 'homotopy systems') and such chain complexes. We refer later to his important homotopy classification results in this area. Here we explain the relation with the Fox free differential calculus [68].

Let $\mu : M \to P$ be a crossed module of groups, let $G = \mathrm{Coker}\mu$, and let $\phi : P \to G$ be the quotient map. Then there is an associated diagram

$$\begin{array}{ccccc} M & \xrightarrow{\mu} & P & \xrightarrow{\phi} & G \\ {\scriptstyle h_2}\downarrow & & \downarrow{\scriptstyle h_1} & & \downarrow{\scriptstyle h_0} \\ M^{\mathrm{ab}} & \xrightarrow[\partial_2]{} & D_\phi & \xrightarrow[\partial_1]{} & \mathbb{Z}[G] \end{array} \tag{17.1}$$

in which the second row consists of (right) G-modules and module morphisms. Here h_2 is simply the abelianisation map; $h_1 : P \to D_\phi$ is the universal ϕ-derivation, that is it satisfies $h_1(pq) = h_1(p)^{\phi q} + h_1(q)$, for all $p, q \in P$, and is universal for this property; and h_0 is the usual derivation $g \mapsto g - 1$. Whitehead in his Theorem 8, p. 469, of [105] gives essentially this diagram in the case P is a free group, when he takes D_ϕ to be the free G-module on the same generators as the free generators of P. Our formulation (Proposition 3.1 of [34]), which uses the derived module due to Crowell [58], includes his case. It is remarkable that diagram (17.1) is a commutative diagram in which the vertical maps are operator morphisms, and that the bottom row is defined by this property. The proof in [34] follows essentially Whitehead's proof. The bottom row is exact: this follows from results in [58], and is a reflection of a classical fact on group cohomology, namely the relation between central extensions and the Ext functor, see [89]. In the case the crossed module is the crossed module $\delta : C(\omega) \to F(X)$ derived from a presentation of a group, then $C(\omega)^{\mathrm{ab}}$ is isomorphic to the free G-module on R, D_ϕ is the free G-module on X, and it is immediate from the above that ∂_2 is the usual derivative $(\partial r/\partial x)$ of Fox's free differential calculus [68]. Thus Whitehead's results anticipate those of Fox.

It is also proved in [105] that if the restriction $M \to \mu(M)$ of μ has a section which is a morphism but not necessarily a P-map, then h_2 maps $\mathrm{Ker}\,\mu$ isomorphically to $\mathrm{Ker}\,\partial_2$. This allows calculation of the module of identities among relations by using module methods, and this is commonly exploited, see for example [67] and the references there.

Whitehead introduced the categories CW of reduced CW-complexes, HS of homotopy systems, and FCC of free chain complexes with a group of operators, together with functors

$$\mathsf{CW} \xrightarrow{\Pi} \mathsf{HS} \xrightarrow{C} \mathsf{FCC}.$$

In each of these categories he introduced notions of homotopy and he proved that C induces an equivalence of the homotopy category of HS with a subcategory of the homotopy category of FCC. Further, $C\Pi X_*$ is isomorphic to the chain complex $C_*\widetilde{X}_*$ of cellular chains of the universal cover of X, so that under these circumstances there is a bijection of sets of homotopy classes

$$[\Pi X_*, \Pi Y_*] \to [C_*\widetilde{X}_*, C_*\widetilde{Y}_*]. \tag{17.2}$$

This with the bijection (16.1) can be interpreted as an operator version of the Hopf classification theorem. It is surprisingly little known. It includes results of Olum [93] published later, and it enables quite useful calculations to be done easily, such as the homotopy classification of maps from a surface to the projective plane [66], and other cases. Thus we see once again that this general theory leads to specific calculations.

All these results are generalised in [34] to the non free case and to the non reduced case, which requires a groupoid of operators, thus giving functors

$$\mathsf{FTop} \xrightarrow{\Pi} \mathsf{Crs} \xrightarrow{\Delta} \mathsf{Chain}.$$

One utility of the generalisation to groupoids is that the functor Δ then has a right adjoint, and so preserves colimits. An example of this preservation is given in [34, Example 2.10]. The construction of the right adjoint to Δ builds on a number of constructions used earlier in homological algebra.

The definitions of the categories under consideration in order to obtain a generalisation of the bijection (17.2) has to be quite careful, since it works in the groupoid case, and not all morphisms of the chain complex are realisable.

This analysis of the relations between these two categories is used in [35] to give an account of cohomology with local coefficients. See also [46] for a relation between extension theory, crossed complexes, and covering spaces of non-connected topological groups.

18 Crossed complexes and simplicial groups and groupoids

The Moore complex NG of a simplicial group G is in general not a (reduced) crossed complex. Let D_nG be the subgroup of G_n generated by degenerate elements. Ashley showed in his thesis [7] that NG is a crossed complex if and only if $(NG)_n \cap (DG)_n = \{1\}$ for all $n \geqslant 1$.

Ehlers and Porter in [62, 63] show that there is a functor C from simplicial groupoids to crossed complexes in which $C(G)_n$ is obtained from $N(G)_n$ by factoring out

$$(NG_n \cap D_n)d_{n+1}(NG_{n+1} \cap D_{n+1}),$$

where the Moore complex is defined so that its differential comes from the last simplicial face operator.

This is one part of an investigation into the Moore complex of a simplicial group, of which the most general investigation is by Carrasco and Cegarra in [54].

An important observation in [95] is that if $N \lhd G$ is an inclusion of a normal simplicial subgroup of a simplicial group, then the induced morphism on components $\pi_0(N) \to \pi_0(G)$ obtains the structure of crossed module. This is directly analogous to the fact that if $F \to E \to B$ is a fibration sequence then the induced morphism of fundamental groups $\pi_1(F, x) \to \pi_1(E, x)$ also obtains the structure of crossed module. This is relevant to algebraic K-theory, where for a ring R the homotopy fibration sequence is taken to be $F \to B(GL(R)) \to B(GL(R))^+$.

19 Other homotopy multiple groupoids

The proof of the GSVKT outlined earlier does seem to require cubical methods, so there is still a question of the place of globular and simplicial methods in this area. A simplicial analogue of the equivalence of categories is given in [7, 92], using Dakin's notion of *simplicial T-complex* [59]. However it is difficult to describe in detail the notion of tensor product of such structures, or to formulate a proof of the colimit theorem in that context. It may be that the polyhedral methods of [82] would help here.

It is easy to define a homotopy globular set $\varrho^{\bigcirc}(X_*)$ of a filtered space X_* but it is not quite so clear how to prove directly that the expected compositions are well defined. However there is a natural graded map

$$i : \varrho^{\bigcirc}(X_*) \to \varrho(X_*) \tag{19.1}$$

and applying the folding map of [1, 2] analogously to methods in [31] allows one to prove that i of (19.1) is injective. It follows that the compositions on $\varrho(X_*)$ are inherited by $\varrho^{\bigcirc}(X_*)$ to make the latter a globular ω-groupoid. A paper is in preparation on this.

Loday in 1982 [86] defined the fundamental cat^n-group of an n-cube of spaces, and showed that cat^n-groups model all reduced weak homotopy $(n+1)$-types. Joint work [42] formulated and proved a GSVKT for the cat^n-group functor from n-cubes of spaces and this allows new local to global calculations of certain homotopy n-types [21]. This work obtains more powerful results than the purely linear theory of crossed complexes, which has however other advantages. Porter in [95] gives an interpretation of Loday's results using methods of simplicial groups. There is clearly a lot to do in this area.

Recently some absolute homotopy 2-groupoids and double groupoids have been defined, see [27] and the references there, and it is significant that crossed modules have been used in a differential topology situation by Mackaay and Picken [88]. Reinterpretations of these ideas in terms of double groupoids are started in [24].

It seems reasonable to suggest that, in the most general case, double groupoids are still somewhat mysterious objects; a generalisation of the work in [48] is given in [43], but this does not apply to the homotopy double groupoid of a map constructed in [40].

20 Conclusion and questions

• The emphasis on filtered spaces rather than the absolute case is open to question.

• *Mirroring the geometry by the algebra* is crucial for conjecturing and proving universal properties.

• *Thin elements* are crucial as modelling commutative cubes, a concept not so easy to define or handle algebraically.

• *Colimit theorems* give, when they apply, exact information even in non commutative situations. The implications of this for homological algebra and its applications could be important.

• One construction inspired eventually by this work was the *non abelian tensor product of groups*, defined in full generality with Loday in 1985, and which now has a bibliography of 79 papers (http://www.bangor.ac.uk/~mas010/nonabtens.html).

• Globular methods do fit into the scheme of higher order categorical structures, but so far have not yielded new local-to-global results in the style of the current work.

• For computations we really need strict structures (although we do want to compute invariants of homotopy colimits, cf, [44]).

• In homotopy theory, identifications in low dimensions can affect high dimensional homotopy. So we need structure in a range of dimensions to model homotopical identifications algebraically. The idea of identifications in low dimensions is reflected in the algebra by 'induced constructions'.

• In this way we calculate some crossed modules modelling homotopy 2-types, whereas the corresponding k-invariant is often difficult to calculate.

• The use of crossed complexes in Čech theory is a current project with Tim Porter.

• **Question:** Are there applications of higher homotopy groupoids in other contexts where the fundamental groupoid is currently used, such as algebraic geometry?

• **Question:** There are uses of double groupoids in differential geometry, for example in Poisson geometry, and in 2-dimensional holonomy [38]. Is there a non abelian De Rham Theory, using an analogue of crossed complexes?

• **Question:** Is there a truly non commutative integration theory based on limits of multiple compositions of elements of multiple groupoids?

Acknowledgements

This work has been partially supported by the SRC, EPSRC, British Council, ARC, and INTAS, the latter under grants INTAS 97-31961 'Algebraic Homotopy, Galois Theory and Descent'; 93-367 'Algebraic K-theory, groups and categories'; 'Algebraic K-theory, Groups and Algebraic Homotopy Theory'.

References

[1] F. Al-Agl, 1989, *Aspects of multiple categories*, Ph.D. thesis, University of Wales, Bangor.

[2] F. Al-Agl, R. Brown and R. Steiner, *Multiple categories: the equivalence between a globular and cubical approach*, Advances in Math., 170 (2002) 71–118.

[3] M. Alp, *GAP, Crossed modules, Cat^1-groups: Applications of computational group theory* Ph.D. thesis, University of Wales, Bangor, (1997),
`http://www.informatics.bangor.ac.uk/public/math/research/ftp/theses/alp.ps.gz`.

[4] M. Alp and C.D. Wensley, *Enumeration of cat^1-groups of low order*, Int. J. Math. Computation, 10 (2000) 407–424.

[5] R. Antolini, *Cubical structures and homotopy theory*, Ph.D. thesis, Univ. Warwick, Coventry, (1996).

[6] R. Antolini, *Cubical structures, homotopy theory*, Ann. Mat. Pura Appl., 178 (2000) 317–324.

[7] N. Ashley, *Simplicial T-complexes*, University of Wales, Bangor, Ph.D. thesis (1978), published as *Simplicial T-Complexes: a non abelian version of a theorem of Dold-Kan*, Dissertationes Math., 165 (1988) 11–58.

[8] M.F. Atiyah, *Mathematics in the 20th Century*, Bull. London Math. Soc., 34 (2002) 1–15.

[9] H.J. Baues, *Algebraic Homotopy*, volume 15 of *Cambridge Studies in Advanced Mathematics*, Cambridge University Press (1989).

[10] H.J. Baues and R. Brown, *On relative homotopy groups of the product filtration, the James construction, and a formula of Hopf*, J. Pure Appl. Algebra, 89 (1993) 49–61.

[11] H.J. Baues and D. Conduché, *On the tensor algebra of a non-abelian group and applications*, K-theory, 5 (1992) 531–554.

[12] H.-J. Baues and A. Tonks, *On the twisted cobar construction*, Math. Proc. Cambridge Philos. Soc., 121 (1997) 229–245.

[13] F. Borceux and G. Janelidze, *Galois theories*, Cambridge Studies in Advanced Mathematics, 72, Cambridge University Press, Cambridge, 2001.

[14] L. Breen, *Théorie de Schreier supérieure*, Ann. Sci. École. Norm. Sup., 25 (1992) 465–514.

[15] R.A. Brown, *Generalized group presentation and formal deformations of CW complexes*, Trans. Amer. Math. Soc., 334 (1992) 519–549.

[16] R. Brown, *Groupoids and van Kampen's Theorem*, Proc. London Math. Soc., 3 (1967) 385–340.

[17] R. Brown, *Topology:a geometric account of general topology, homotopy types and the fundamental groupoid*, Ellis Horwood Series: Mathematics and its Applications, Ellis Horwood Ltd., Chichester, second edition (1988) (first edition, McGraw Hill, 1968).

[18] R. Brown, *On the second relative homotopy group of an adjunction space: an exposition of a theorem of J. H. C. Whitehead*, J. London Math. Soc. (2), 22 (1980) 146–152.

[19] R. Brown, *Coproducts of crossed P-modules: applications to second homotopy groups and to the homology of groups*, Topology, 23 (1984) 337–345.
[20] R. Brown, *From groups to groupoids: a brief survey*, Bull. London Math. Soc., 19 (1987) 113–134.
[21] R. Brown, 1992, *Computing homotopy types using crossed n-cubes of groups*, in *Adams Memorial Symposium on Algebraic Topology, 1 (Manchester, 1990)*, volume 175 of *London Math. Soc. Lecture Note Ser.*, 187–210, Cambridge University Press, Cambridge.
[22] R. Brown, *Groupoids and crossed objects in algebraic topology*, Homology, homotopy and applications, 1 (1999) 1–78.
[23] R. Brown and N. D. Gilbert, *Algebraic models of 3-types and automorphism structures for crossed modules*, Proc. London Math. Soc. (3), 59 (1989) 51–73.
[24] R. Brown and J. F. Glazebrook, *Connections, local subgroupoids, and a holonomy Lie groupoid of a line bundle gerbe*, Univ. Iagellonicae Acta Math. (to appear), math.DG/02/0322..
[25] R. Brown, M. Golasiński, T. Porter and A. Tonks, *Spaces of maps into classifying spaces for equivariant crossed complexes*, Indag. Math. (N.S.), 8 (1997) 157–172.
[26] R. Brown, M. Golasiński, T. Porter and A. Tonks, *Spaces of maps into classifying spaces for equivariant crossed complexes. II. The general topological group case*, K-Theory, 23 (2001) 129–155.
[27] R. Brown, K. A. Hardie, K. H. Kamps and T. Porter, *A homotopy double groupoid of a Hausdorff space*, Theory Appl. Categ., 10 (2002) 71–93.
[28] R. Brown and P. J. Higgins, *On the connection between the second relative homotopy groups of some related spaces*, Proc.London Math. Soc., (3) 36 (1978) 193–212.
[29] R. Brown and P. J. Higgins, *Sur les complexes croisés d'homotopie associés à quelques espaces filtrés*, C. R. Acad. Sci. Paris Sér. A-B, 286 (1978) A91–A93.
[30] R. Brown and P. J. Higgins, *The algebra of cubes*, J. Pure Appl. Algebra , 21 (1981) 233–260.
[31] R. Brown and P. J. Higgins, *Colimit theorems for relative homotopy groups*, J. Pure Appl. Alg, 22 (1981) 11–41.
[32] R. Brown and P. J. Higgins, *The equivalence of crossed complexes and ω-groupoids,*, Cahiers Top. Géom. Diff., 22 (1981) 370–386.
[33] R. Brown and P. J. Higgins, *Tensor products and homotopies for ω-groupoids and crossed complexes,*, J. Pure Appl. Alg, 47 (1987) 1–33.
[34] R. Brown and P. J. Higgins, *Crossed complexes and chain complexes with operators*, Math. Proc. Camb. Phil. Soc., 107 (1990) 33–57.
[35] R. Brown and P. J. Higgins, *The classifying space of a crossed complex*, Math. Proc. Cambridge Philos. Soc., 110 (1991) 95–120.
[36] R. Brown and P. J. Higgins, *Cubical abelian groups with connections are equivalent to chain complexes*, Homology, Homotopy and Applications, 5 (2003) 49–52.
[37] R. Brown and J. Huebschmann, *Identities among relations*, in R.Brown and T.L.Thickstun, eds., *Low Dimensional Topology*, London Math. Soc Lecture Notes, Cambridge University Press (1982) 153–202.
[38] R. Brown and I. Icen, *Towards two dimensional holonomy*, Advances in Math., 178 (2003) 141–175.
[39] R. Brown and G. Janelidze, *Van Kampen theorems for categories of covering morphisms in lextensive categories*, J. Pure Appl. Algebra, 119 (1997) 255–263.
[40] R. Brown and G. Janelidze, *A new homotopy double groupoid of a map of spaces*, Appl. Cat. Struct. (to appear).
[41] R. Brown and J.-L. Loday, *Homotopical excision, and Hurewicz theorems, for n-cubes of spaces*, Proc. London Math. Soc., (3) 54 (1987) 176–192.
[42] R. Brown and J.-L. Loday, *Van Kampen Theorems for diagrams of spaces*, Topology, 26 (1987) 311–337.
[43] R. Brown and K.C.H. Mackenzie, *Determination of a double Lie groupoid by its core diagram*, J. Pure Appl. Algebra, 80 (1992) 237–272.
[44] R. Brown, E. Moore, T. Porter and C. Wensley, *Crossed complexes, and free crossed resolutions for amalgamated sums and HNN-extensions of groups*,Georgian Math. J., 9 (2002) 623–644.

[45] R. Brown and G. H. Mosa, *Double categories, 2-categories, thin structures and connections*, Theory Appl. Categ., 5 (1999) 163–175.

[46] R. Brown and O. Mucuk, *Covering groups of non-connected topological groups revisited*, Math. Proc. Camb. Phil. Soc, 115 (1994) 97–110.

[47] R. Brown and A. Razak Salleh, *Free crossed resolutions of groups and presentations of modules of identities among relations*, LMS J. Comput. Math., 2 (1999) 28–61.

[48] R. Brown and C. B. Spencer, *Double groupoids and crossed modules*, Cah. Top. Géom. Diff., 17 (1976) 343–362.

[49] R. Brown and C. B. Spencer, *G-groupoids, crossed modules and the fundamental groupoid of a topological group*, Proc. Kon. Ned. Akad. v. Wet, 79 (1976) 296–302.

[50] R. Brown and C. D. Wensley, *On finite induced crossed modules, and the homotopy 2-type of mapping cones*, Theory Appl. Categ., 1 (1995) 54–70.

[51] R. Brown and C. D. Wensley, *Computing crossed modules induced by an inclusion of a normal subgroup, with applications to homotopy 2-types*, Theory Appl. Categ., 2 (1996) 3–16.

[52] R. Brown and C. D. Wensley, *Computations and homotopical applications of induced crossed modules*, J. Symb. Comp., 35 (2003) 59–72.

[53] M. Bunge and S. Lack, *Van Kampen theorems for toposes*, Advances in Math., 179 (2003) 291–317.

[54] P. Carrasco and A. M. Cegarra, *Group-theoretic algebraic models for homotopy types*, J. Pure Appl. Algebra, 75 (1991) 195–235.

[55] E. Ĉech, 1933, *Höherdimensionale Homotopiegruppen*, in *Verhandlungen des Internationalen Mathematiker-Kongresses Zurich 1932*, volume 2, 203, International Congress of Mathematicians (4th : 1932 : Zurich, Switzerland, Walter Saxer, Zurich, reprint Kraus, Nendeln, Liechtenstein, 1967.

[56] A. Connes, *Noncommutative geometry*, Academic Press Inc., San Diego, CA (1994).

[57] J.-M. Cordier and T. Porter, *Homotopy coherent category theory*, Trans. Amer. Math. Soc., 349 (1997) 1–54.

[58] R.H. Crowell, *The derived module of a homomorphism*, Advances in Math., 5 (1971) 210–238.

[59] K. Dakin, *Kan complexes and multiple groupoid structures*, Ph.D. thesis, University of Wales, Bangor (1977).

[60] P. Dedecker, *Sur la cohomologie non abélienne. II*, Canad. J. Math., 15 (1963) 84–93.

[61] A. Douady and R. Douady, *Algebres et théories Galoisiennes*, volume 2, CEDIC, Paris (1979).

[62] P. J. Ehlers and T. Porter, *Varieties of simplicial groupoids. I. Crossed complexes*, J. Pure Appl. Algebra, 120 (1997) 221–233.

[63] P. J. Ehlers and T. Porter, *Erratum to: "Varieties of simplicial groupoids. I. Crossed complexes" [J. Pure Appl. Algebra 120 (1997) no. 3, 221–233;* , J. Pure Appl. Algebra, 134 (1999) 207–209.

[64] C. Ehresmann, *Catégories et structures*, Dunod, Paris, (1964).

[65] G.J. Ellis, *Homology of 2-types*, J. London Math. Soc., (2) 46 (1992) 1–27.

[66] G. J. Ellis, *Homotopy classification the J. H. C. Whitehead way*, Exposition. Math., 6 (1988) 97–110.

[67] G. Ellis and I. Kholodna, *Three-dimensional presentations for the groups of order at most 30*, LMS J. Comput. Math., 2 (1999) 93–117 +2 appendixes (HTML and source code).

[68] R.H. Fox, *Free differential calculus I: Derivations in the group ring*, Ann. Math., 57 (1953) 547–560.

[69] A. Fröhlich, *Non-Abelian homological algebra. I. Derived functors and satellites.*, Proc. London Math. Soc. (3) 11 (1961) 239–275.

[70] P. Gaucher, *Combinatorics of branchings in higher dimensional automata*, Theory Appl. Categ., 8 (2001) 324–376.

[71] The GAP Group, 2002, *Groups, Algorithms, and Programming, version 4.3*, Technical report, http://www.gap-system.org.

[72] M. Grandis and L. Mauri, *Cubical sets and their site*, Theory and Apl. Categories, (to appear).

[73] A. Heyworth, *Applications of rewriting systems and Groebner bases to computing Kan extensions and identities among relations*, U.W. Bangor Ph.D. thesis (1998) (Preprint 98.23).

[74] A. Heyworth and C.D. Wensley, *Logged rewriting and identities among relators*, Groups St Andrews 2001 in Oxford Vol I, eds. C. Campbell, E. Robertson, G. S. Smith, C.U.P., (2003) 256–276.

[75] P.J. Higgins, *Presentations of groupoids, with applications to aroups*, Proc. Camb. Phil. Soc., 60 (1964) 7–20.

[76] P.J. Higgins, *Categories and Groupoids*, van Nostrand, New York (1971).

[77] P.J. Higgins and J. Taylor, *The fundamental groupoid and homotopy crossed complex of an orbit space*, in K. H. Kamps et al., ed., *Category Theory: Proceedings Gummersbach 1981*, SLNM 962 (1982) 115–122.

[78] C. Hog-Angeloni and W. Metzler, eds., *Two-dimensional homotopy and combinatorial group theory*, volume 197 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge (1993).

[79] J. Huebschmann, *Crossed n-fold extensions of groups and cohomology*, Comment. Math. Helv., 55 (1980) 302–313.

[80] J. Huebschmann, *Automorphisms of group extensions and differentials in the Lyndon-Hochschild-Serre spectral sequence*, J. Algebra 72 (1981) 296–334.

[81] J. Huebschmann, *Group extensions, crossed pairs and an eight term exact sequence*, Jour. fur. reine. u. ang. Math., 321 (1981) 150–172.

[82] D.W. Jones, *Polyhedral T-complexes*, Ph.D. thesis, University of Wales, Bangor, (1984). Published as *A general theory of polyhedral sets and their corresponding T-complexes*, Diss. Math., 266 (1988).

[83] E.H.v. Kampen, *On the connection between the fundamental groups of some related spaces*, Amer. J. Math., 55 (1933) 261–267.

[84] K.H. Kamps and T. Porter, *Abstract homotopy and simple homotopy theory*, World Scientific Publishing Co. Inc., River Edge, NJ, (1997).

[85] J.-P. Labesse, *Cohomologie, stabilisation et changement de base*, Astérisque, appendix A by Laurent Clozel and Labesse, and Appendix B by Lawrence Breen, 257 (1999).

[86] J.-L. Loday, *Spaces with finitely many homotopy groups*, J.Pure Appl. Algebra , 24 (1982) 179–202.

[87] A.S.T. Lue, *Cohomology of groups relative to a variety*, J. Algebra , 69 (1981) 155–174.

[88] M. Mackaay and R. Picken, *Holonomy and parallel transport for abelian gerbes*, Advances in Math., 170 (2002) 287–339.

[89] S. MacLane, *Homology*, number 114 in Grundlehren, Springer (1967).

[90] S. MacLane and J.H.C. Whitehead, *On the 3-type of a complex*, Proc. Nat. Acad. Sci. U.S.A., 36 (1950) 41–48.

[91] E.J. Moore, , *Graphs of Groups: Word Computations and Free Crossed Resolutions*, Ph.D. thesis, University of Wales, Bangor (2001) UWB Preprint 01.02.

[92] G. Nan Tie, *A Dold-Kan theorem for crossed complexes*, J. Pure Appl. Algebra, 56 (1989) 177–194.

[93] P. Olum, *Mappings of manifolds and the notion of degree*, Ann. of Math., 58 (1953) 458–480.

[94] R. Peiffer, *Über Identitäten zwischen Relationen*, Math. Ann., 121 (1949) 67 – 99.

[95] T. Porter, *n-types of simplicial groups and crossed n-cubes*, Topology, 32 (1993) 5–24.

[96] S. Pride, 1991, *Identities among relations*, in A. V. E.Ghys, A.Haefliger, ed., *Proc. Workshop on Group Theory from a Geometrical Viewpoint*, 687–716, International Centre of Theoretical Physics, Trieste, 1990, World Scientific.

[97] K. Reidemeister, *Über Identitäten von Relationen*, Abh. Math. Sem. Hamburg, 16 (1949) 114–118.

[98] J. Taylor, *Quotients of groupoids by the action of a group*, Math. Proc. Camb. Phil. Soc., 103 (1988) 239–249.

[99] A. P. Tonks, 1993, *Theory and applications of crossed complexes*, Ph.D. thesis, University of Wales, Bangor (UWB Math Preprint 93.17).

[100] A.P. Tonks, *On the Eilenberg-Zilber theorem for crossed complexes*, J. Pure Appl. Algebra, 179 (2003) 199–220.

[101] C.T.C. Wall, *Finiteness conditions for CW-complexes II*, Proc. Roy. Soc. Ser. A, 295 (1966) 149–166.

[102] A. Weinstein, *Groupoids: unifying internal and external symmetry*, Notices. Amer. Math. Soc., 43 (1996) 744–752. math.RT/9602220.

[103] J.H.C. Whitehead, *On adding relations to homotopy groups*, Ann. of Math., 42 (1941) 409–428.

[104] J.H.C. Whitehead, *Combinatorial Homotopy I*, Bull. Amer. Math. Soc., 55 (1949) 213–245.

[105] J.H.C. Whitehead, *Combinatorial Homotopy II*, Bull. Amer. Math. Soc., 55 (1949) 453–496.

[106] J.H.C. Whitehead, *Simple homotopy types*, Amer. J. Math., 72 (1950) 1–57.

Received December 16, 2002; in revised form May 10, 2003

Fields Institute Communications
Volume 43, 2004

Galois Groupoids and Covering Morphisms in Topos Theory

Marta Bunge
Department of Mathematics and Statistics
McGill University,
805 Sherbrooke St West
Montréal QC, Canada H3A 2K6
bunge@math.mcgill.ca

Abstract. The goals of this paper are (1) to compare the Galois groupoid that appears naturally in the construction of the coverings fundamental groupoid topos given by Bunge (1992) with the formal Galois groupoid defined by Janelidze (1990) in a very general setting given by a pair of adjoint functors, and (2) to discuss a good notion of covering morphism of a topos that is general enough to include, in addition to the covering projections determined by the locally constant objects, also the unramified morphisms of topos theory given by those local homeomorphisms which are at the same time complete spreads in the sense of Bunge-Funk (1996, 1998). We also (3) introduce and study a notion of a Galois topos that generalizes in different ways those of Grothendieck (1971) and Moerdijk (1989), (4) explain the role of stack completions in distinguishing Galois groupoids from fundamental groupoids when the base topos is not Sets but arbitrary, (5) extend to the case of an arbitrary base topos results of Bunge-Moerdijk (1997) concerning the comparison between the coverings and the paths fundamental groupoid toposes, and (6) discuss pseudofunctorialy of the fundamental groupoid constructions and apply it to give a simple version of the van Kampen theorem for toposes of Bunge-Lack (2003).

1 Introduction

One of the purposes of this paper is to compare, given a locally connected and locally simply connected topos $\mathscr{E}$ over a base topos $\mathscr{S}$, the Galois groupoid of automorphisms of the canonical "point" (in effect, a "bag of points") of the

---

2000 *Mathematics Subject Classification.* Primary 18B, 18D, 18F; Secondary 57M.
Key words and phrases. topos, covering, locally constant object, Galois groupoid, fundamental groupoid, stack completion, Galois topos, complete spread, van Kampen theorems.
Invited lecture for the Workshop on Categorical Structures for Descent and Galois Theory, Hopf Algebras and Semiabelian Categories, The Fields Institute for Research in the Mathematical Sciences, University of Toronto, September 23-28, 2002. Research partially supported by the Natural Sciences and Engineering Research Council of Canada.

coverings fundamental groupoid topos of $\mathscr{E}$ as defined in [9], with the formal Galois groupoid of Janelidze [23, 24] in this setting.

The notion of a locally constant object is central to these considerations; we begin then by investigating in section 2 (Locally constant objects in toposes) the connection between (rather, identification of) the notion of a *locally constant object* given in [14] (in turn inspired by [9]), with that of Janelidze [24, 25] and, in passing, also with that of Barr and Diaconescu [3], with which both agree if the topos $\mathscr{E}$ is connected and the base topos $\mathscr{S}$ is Set.

In section 3 (Stack completions and the fundamental groupoid of a topos) we revisit the construction of the coverings fundamental groupoid of a locally connected topos $\mathscr{E}$ over a base topos $\mathscr{S}$ given in [9], an account of which is briefly given in [15] under the implicit assumption that the base topos $\mathscr{S}$ is Set. A new ingredient here is the observation, not previously made explicit in either [9] or [15], that there are *two* groupoids involved, to wit, for each cover U in $\mathscr{E}$, there is on the one hand the *internal* Galois groupoid G_U of automorphisms of the universal cover of the topos of U-split objects in $\mathscr{E}$, and on the other hand its stack completion π_U which classifies U-split torsors and which in principle need not be internal to $\mathscr{S}$ but only indexed (fibered) over $\mathscr{S}$. When the base topos is Set, this distinction dissapears and the two groupoids are usually identified in practice [21, 15]. This brings into consideration the possible advantages of assuming that the base (elementary) topos $\mathscr{S}$ satisfies an *axiom of stack completions* (ASC), which was suggested by Lawvere in 1974 and which is known to hold at least of all Grothendieck toposes $\mathscr{S}$, as shown in [18, 27]. This axiom guarantees that the (coverings) fundamental groupoid of a locally connected and locally simply connected topos $\mathscr{E}$ over $\mathscr{S}$, which is the stack completion of the internal Galois groupoid, is represented, as an $\mathscr{S}$-indexed category, by an internal groupoid, weakly equivalent to the Galois groupoid.

We prove in section 4 (Galois groupoids and Galois toposes) that there is an equivalence (and not just a Morita equivalence) between the Galois groupoid of automorphisms of a universal cover of a locally connected and locally simply connected topos $e : \mathscr{E} \to \mathscr{S}$, and the Galois groupoid of the pure theory associated with the first pair of adjoints in the 3-tuple $e_! \dashv e^* \dashv e_*$ given by the (locally connected) geometric morphism e. We also show that in topos theory, and just as in the pure Galois theory of [24, 5], the Galois theory is implicit in the construction of the Galois groupoid; in the case of toposes, we obtain it from the presence of the third adjoint in the sequel of three determined by e, but a different argument that employs only the cartesian closed structure is available, as shown by Janelidze.

Also in section 4 we introduce and study the notion of a *Galois topos* over an arbitrary base topos $\mathscr{S}$. Although these relative Galois toposes are not assumed to be either connected or pointed, they come naturally equipped with a bag of points indexed by the connected components of a (non-connected) universal cover; this is in line with the view advocated by Grothendieck [22] and Brown [6] that, rather than fixing a single base point, one ought work with a suitable "paquet des points", for instance one that is invariant under the symmetries in the given situation. This idea was naturally and independently incorporated in topos theory both by Kennison [29] and myself [9], by discussing the fundamental groupoid of an *unpointed* (and possibly pointless) locally connected topos.

Pointed connected Galois toposes over $\mathscr{S}$ have been investigated by Moerdijk [31] following Grothendieck [20] (see also [27]). We obtain here characterization theorems in a manner parallel to [31]; in particular we show that our Galois toposes over a base topos $\mathscr{S}$ satisfying (ASC) (and which correspond, modulo the intervention of locales in the subject, to the multi-Galois toposes of Grothendieck [21]) are precisely the classifying toposes of prodiscrete (localic) groupoids in $\mathscr{S}$.

In section 5 (Locally paths simply connected toposes over an arbitrary base) we recall the paths version of the fundamental group topos [33, 34], give a constructive version of the existence of a comparison map from the paths to the coverings fundamental groupoid toposes, and then prove the equivalence of the comparison map under an assumption of the locally paths simply connected type. We also investigate pseudonaturality of the fundamental groupoid topos constructions and their universal properties, and apply this to give a simplified version of the van Kampen theorem for locally paths simply connected toposes, in connection with [14].

In section 6 (Generalized covering morphisms and a van Kampen theorem) we turn to the topic of covering morphisms, also in connection with the Galois and fundamental groupoids. The notion we give here has good enough properties so that (a pushout version of) the van Kampen theorem from [14] holds for these coverings and is also general enough to include, not only the familiar covering projections determined by the locally constant objects, but also the unramified coverings introduced and studied in [11, 12, 19], namely those local homeomorphisms which are also complete spreads.

2 Locally constant objects in toposes

Let $\mathscr{E}$ be an elementary topos bounded over an elementary topos $\mathscr{S}$ by means of a geometric morphism $e : \mathscr{E} \to \mathscr{S}$.

Before proceeding, we explain this assumption roughly in simple terms. The notion of an *elementary topos* is at the same time a generalization of the notion of a topological space via sheaves, and a universe for set theory that is not necessarily a classical one. The base topos $\mathscr{S}$ above is interpreted as a choice of set theory, whereas the topos $\mathscr{E}$ is thought of as sheaves on a site in $\mathscr{S}$, that is, on a category in $\mathscr{S}$ equipped with a Lawvere-Tierney topology, explicitated by the *bounded* nature of the geometric morphism $e : \mathscr{E} \to \mathscr{S}$. If $\mathscr{S}$ is the category of sets, what this amounts to is simply that $\mathscr{E}$ is a *Grothendieck topos*, with $e : \mathscr{E} \to \mathscr{S}$ the unique geometric morphism whose direct image part $e_* : \mathscr{E} \to \mathscr{S}$ is the global sections functor, and whose inverse image part $e^* : \mathscr{S} \to \mathscr{E}$ assigns to a set S the associated sheaf the presheaf whose value is constantly S. In particular, we may think of $\mathscr{E}$ as a generalized space defined within the set theory $\mathscr{S}$. It is of course of interest, in the setting of generalized spaces within $\mathscr{S}$, to consider geometric morphisms $\mathscr{F} \to \mathscr{E}$ over $\mathscr{S}$, for instance covering morphisms.

One can also express, in the setting of elementary toposes and geometric morphisms, generalizations of familiar topological notions, such as that of a locally connected space, by making special assumptions about the given (in general not unique) "structure morphism" $e : \mathscr{E} \to \mathscr{S}$. For instance, one says that $\mathscr{E}$ is *locally connected* over $\mathscr{S}$ by means of e if there is given an $\mathscr{S}$-indexed left adjoint $e_! \dashv e^*$, where e is the geometric morphism given by the (necessarily $\mathscr{S}$-indexed) adjoint pair $e^* \dashv e_*$. We think of $e_! : \mathscr{E} \to \mathscr{S}$ as the (set of) connected components functor.

We now recall a *fundamental pushout* construction in the definition of the fundamental groupoid of $\mathscr{E}$ as given in [9]. We call it "fundamental" both on account of its role in the fundamental groupoid, and since this construction contains basically all of the information we use to develop Galois theory in the context of toposes.

For an epimorphism $U \to 1$ in $\mathscr{E}$ (a *cover* in $\mathscr{E}$) denote by $\mathscr{G}_U$ the topos defined by the following pushout in $\mathbf{Top}_{\mathscr{S}}$

$$\begin{tikzcd}
\mathscr{E}/U \arrow[r, "\varphi_U"] \arrow[d, "\rho_U"] & \mathscr{E} \arrow[d, "\sigma_U"] \\
\mathscr{S}/e_!U \arrow[r, "p_U"'] & \mathscr{G}_U
\end{tikzcd} \tag{$*$}$$

where $\varphi_U : \mathscr{E}/U \to \mathscr{E}$ is the canonical local homeomorphism and where ρ_U is the connected locally connected part in the (unique) factorization of the composite $e\,\varphi_U : \mathscr{E}/U \to \mathscr{S}$, which is the composite of two locally connected hence locally connected, into a connected locally connected morphism followed by a (surjective) local homeomorphism.

We point out that in the 2-category $\mathbf{Top}_{\mathscr{S}}$ of toposes bounded over $\mathscr{S}$, geometric morphisms over $\mathscr{S}$, and iso 2-cells, pushouts exist and are calculated simply as pullbacks of the inverse image parts in **CAT**, so that anything that depends solely on pushouts has, in principle, "nothing to do with toposes". However, the fact that, by performing a pullback in **CAT** of this kind to a diagram of toposes and (inverse images of) geometric morphisms, one obtains again toposes and (inverse images of) geometric morphisms is what makes this construction so powerful in topos theory. Without this general fact, one may need to show that a certain category (respectively, function) is a topos (respectively, geometric morphism) "by hand" in each such situation, which may be rather involved.

The fundamental pushout construction $(*)$ was what motivated the "family" definition of a locally constant object formally introduced in [14] and given below in the special context of a locally connected topos $\mathscr{E}$ over $\mathscr{S}$. It is, in fact, the natural notion to consider in topology in the case of a non-connected cover; in the context of Grothendieck toposes it also appears in [20]. This was seen independently also by Janelidze [23, 24]. There is, however, a non constructively given equivalence with a "single object" description when the topos $\mathscr{E}$ is connected and defined over *Set*, as shown by Barr and Diaconescu [3], but it is not clear to us at present whether this equivalence (which is of course reasonable to expect) can be made constructive.

Definition 2.1 An object A of (a locally connected topos) $\mathscr{E}$ (with structure map $e : \mathscr{E} \to \mathscr{S}$) is said to be *locally constant* if there exists a cover U in $\mathscr{E}$ which splits A in the sense that there is a morphism $\alpha : S \to e_!U$ in $\mathscr{S}$ and an isomorphism $\theta : e^*S \times_{e^*e_!U} U \to A \times U$ over U, where U is equipped with the morphism $\eta_U : U \to e^*e_!U$ given by the unit of the adjointness $e_! \dashv e^*$ evaluated at U.

The following result is easily proven and appears in [14].

Proposition 2.2 *If $e : \mathscr{E} \to \mathscr{S}$ is locally connected, A an object and U a cover in $\mathscr{E}$, the following are equivalent:*

1. *A is locally constant and split by U in the sense of Definition 2.1.*
2. *A is U-locally trivial with respect to the adjoint pair $e_! \dashv e^*$ in the sense of [24], that is, there is a morphism $\alpha : S \to e_!U$ and a morphism $\zeta : A \to$*

$U \to e^*S$ such that the square

$$\begin{tikzcd} A \times U \arrow[r, "\pi_2"] \arrow[d, "\zeta"'] & U \arrow[d, "\eta"] \\ e^*S \arrow[r, "e^*\alpha"'] & e^*e_!U \end{tikzcd}$$

is a pullback.

3. *There is a morphism $\alpha : J \to I$ in $\mathscr{S}$, a morphism $\eta : U \to e^*I$ in $\mathscr{E}$, and an isomorphism $\theta : e^*J \times_{e^*I} U \to A \times U$ over U.*

Remark 2.3 The equivalence between (1) and (2) in Proposition 2.2 shows that an object A of a locally connected topos $\mathscr{E}$ is locally constant (relative to its structure map $e : \mathscr{E} \to \mathscr{S}$) in the sense of Definition 2.1 if and only if the unique morphism $A \to 1$ is locally trivial with respect to the adjoint pair $e_! \dashv e^*$ in the sense of Janelidze [24]. The equivalence between (1) and (3) in Proposition 2.2 strips the notion to its barest terms and in particular has the desirable consequence of its *stability* under pullback along arbitrary geometric morphisms. Note that (3) is meaningful in an arbitrary topos $\mathscr{E}$ bounded over $\mathscr{S}$ and that it agrees with (1) (and so also with (2)) if $\mathscr{E}$ is locally connected over $\mathscr{S}$. In [14] we gave (3) as the notion of a locally constant object in a topos $\mathscr{E}$ defined over $\mathscr{S}$.

Corollary 2.4 *Let $\varphi : \mathscr{F} \to \mathscr{E}$ be a geometric morphism in* $\mathbf{Top}_{\mathscr{S}}$. *If A is a locally constant object in $\mathscr{E}$ split by the cover U in $\mathscr{E}$ in the sense of* (2) *in Proposition 2.2, then φ^*A is a locally constant object in $\mathscr{F}$, split by the cover φ^*U in $\mathscr{F}$.*

Proof The conclusion is easily checked. □

Denote by $\mathrm{Spl}(U)$ the full subcategory of $\mathscr{E}$ determined by the U-split objects of $\mathscr{E}$ in the sense of Definition 2.1.

Lemma 2.5 [9] *Let*

$$\begin{tikzcd} \mathscr{C} \arrow[r, "q"] \arrow[d, "p"'] & \mathscr{B} \arrow[d, "h"] \\ \mathscr{A} \arrow[r, "k"'] & \mathscr{D} \end{tikzcd}$$

be a pushout in $\mathbf{Top}_{\mathscr{S}}$ *where p and q are both locally connected. Then, each of h and k is locally connected. Furthermore, if p (respectively, q) is connected, then so is h (respectiveky k).*

Proposition 2.6 *Let $\mathscr{G}_U$ be the topos in the fundamental pushout $(*)$ applied to $\mathscr{E}$ and U. Then there exists an equivalence functor $\Phi : \mathscr{G}_U \to \mathrm{Spl}(U)$. Under this equivalence, the fully faithful functor ${\sigma_U}^* : \mathscr{G}_U \to \mathscr{E}$ corresponds to the inclusion of* $\mathrm{Spl}(U)$ *into $\mathscr{E}$.*

Proof An object of $\mathscr{G}_U$ in the fundamental pushout $(*)$ of toposes is precisely a locally constant object split by U in the sense of Definition 2.1. This is a consequence of the well-known fact that pushouts in $\mathbf{Top}_{\mathscr{S}}$ are calculated as pullbacks in $\mathbf{CAT}$ of the inverse image parts. A morphism from $\langle X , S \to e_!U , \theta\rangle$ to $\langle X' , S' \to e_!U , \theta'\rangle$, both objects in $\mathscr{G}_U$, is a pair of morphisms $f : X \to X'$ and $\alpha : S \to S'$, the latter over $e_!U$, compatible with the isomorphisms θ and θ'.

It follows from Lemma 2.5 that σ_U is connected and locally connected. Since the forgetful functor $\sigma_U{}^* : \mathcal{G}_U \to \mathcal{E}$ is then fully faithful and factors through the inclusion $\mathrm{Spl}(U) \hookrightarrow \mathcal{E}$, it follows from the equivalence of (1) and (3) in Proposition 2.2 that there is an equivalence $\Phi : \mathcal{G}_U \to \mathrm{Spl}(U)$. $\square$

Remark 2.7 The equivalence functor $\Phi : \mathcal{G}_U \to \mathrm{Spl}(U)$ forgets the splitting of the U-split objects in the sense of Definition 2.1, as well as the compatibility with the splittings of any morphism between objects in $\mathcal{G}_U$. In the absence of the axiom of choice for $\mathcal{S}$, there is no inverse to Φ. The importance of this remark will emerge from the developments in the next section.

We give now a brief review of the pure Galois theory of Janelidze [23, 24, 5]. There is given a pair of adjoint functors $I \dashv H$ with $I : \mathcal{E} \to \mathcal{S}$ and $H : \mathcal{S} \to \mathcal{E}$ (here $\mathcal{E}$ and $\mathcal{S}$ are *categories*) such that for every object $E \in \mathcal{E}$ the counit $I^E H^E \to \mathrm{Id}_{\mathcal{S}/IE}$ of the induced adjoint pair $I^E \dashv H^E$ in the diagram

$$\mathcal{E}/E \underset{I^E}{\overset{H^E}{\leftrightarrows}} \mathcal{S}/IE$$

is an iso. In this context, a morphism $\alpha : A \to B$ in $\mathcal{E}$ is said to be *trivial* if the diagram

$$\begin{tikzcd} A \arrow[r, "\alpha"] \arrow[d, "\eta_A"] & B \arrow[d, "\eta_B"] \\ HIA \arrow[r, "HI\alpha"'] & HIB \end{tikzcd}$$

is a pullback; $\alpha : A \to B$ is a *locally trivial* morphism relative to the adjoint pair $I \dashv H$ if there exists $p : E \to B$ of effective descent such that $p^*(A, \alpha)$ is trivial in $\mathcal{E}/E$.

The category $\mathrm{Spl}(p)$ of locally trivial morphisms split by p is easily seen to be given as the pullback

$$\begin{tikzcd} \mathcal{E}/E & \mathcal{E}/B \arrow[l, "p^*"'] \\ \mathcal{S}/IE \arrow[u, "H^E"] & \mathrm{Spl}(p) \arrow[l] \arrow[u] \end{tikzcd}$$

.

It is shown in [23] that if in the pure setting one lets $\mathrm{Gal}(p)$ be given by the canonical diagram

$$I((E \times_B E) \times_E (E \times_B E)) \rightrightarrows\!\!\!\to I(E \times_B E) \;\overrightarrow{\underleftarrow{\longrightarrow}}\; I(E)$$

then, although $\mathrm{Gal}(p)$ is not in general an internal category in $\mathcal{S}$, it makes sense to consider its actions and to form a category $\mathcal{S}^{\mathrm{Gal}(p)}$. Furthermore, there is an equivalence

$$\mathrm{Spl}(p) \simeq \mathcal{S}^{\mathrm{Gal}(p)}$$

to be interpreted as the *fundamental theorem of Galois theory.*

In order to get an actual groupoid, it is assumed in [24]of the cover $p : E \to B$ considered in [23] that it is a normal (or regular) cover. As we show in the next section, the existence of a normal cover (actually of the *universal cover*) in the topos setting can be obtained using general properties of the fundamental pushout construction.

If $p : E \to B$ is an I-*normal* object of $\mathscr{E}$ in the sense that $p \in \mathrm{Spl}(p)$, then as shown in [24], the canonical morphism

$$I((E \times_B E) \times_E (E \times_B E)) \to I(E \times_B E) \times_{I(E)} I(E \times_B E)$$

is an isomorphism, from which the desired groupoid structure can be extracted. The reader ought to consult [23, 24, 5] for details.

We prove next that there is also a Galois groupoid in the topos context, associated with a cover $p : U \to 1$, which we obtain from the pushout topos $\mathscr{G}_U$ with the aid of the Joyal-Tierney theorem [28] and additional topos-theoretic considerations [9].

Proposition 2.8 *Let $\mathscr{E}$ be a locally connected topos over $\mathscr{S}$, and let U be a cover in $\mathscr{E}$. Then, the topos $\mathscr{G}_U$ in the fundamental pushout* $(*)$ *is the classifying topos $\mathscr{B}G_U$ of an etale complete discrete localic groupoid G_U in $\mathscr{S}$, by an equivalence*

$$\mathscr{G}_U \simeq \mathscr{B}G_U$$

which identifies p_U with the canonical bag of points of $\mathscr{B}G_U$.

Proof It follows from properties of the fundamental pushout $(*)$ that the geometric morphism $p_U : \mathscr{S}/e_!U \to \mathscr{G}_U$ is a local homeomorphism. Indeed, since both φ_U and ρ_U are locally connected, so is p_U by Lemma 2.5, and since p_U is also totally disconnected in the sense of [9], it must be a local homeomorphism; it is surjective since $U \to 1$ is an epi, hence φ_U is surjective. From general facts about descent toposes [28, 33] it is derived that there is determined an (etale complete) discrete localic groupoid G_U whose discrete locale of objects is $e_!U$ and is such that $\mathscr{G}_U$ is equivalent to the classifying topos $\mathscr{B}G_U$. □

Recall that one of our goals is to compare, in the case of a *locally simply connected* topos $\mathscr{E}$ over $\mathscr{S}$, the Galois groupoid of Janelidze [24] associated with the universal cover, with the Galois groupoid G_U that arises as in Proposition 2.8 above. In section 4 we shall prove their equivalence as groupoids, but prior to that we will, in section 3, make explicit its connection with the fundamental groupoid $\pi_1^c(\mathscr{E})$ of $\mathscr{E}$ as constructed in [9].

3 Stack completions and the fundamental groupoid of a topos

In this section we update the definition of the coverings fundamental groupoid $\pi_1^c(\mathscr{E})$ given in [9] (see also [15]), where $\mathscr{E}$ is a locally connected topos over $\mathscr{S}$ by means of a geometric morphism $e : \mathscr{E} \to \mathscr{S}$. A new ingredient here is making explicit what was only implicit in those sources, namely, the role of stacks (for the topology of regular epimorphisms) of groupoids in $\mathscr{S}$ in the definition of the fundamental groupoid of $\mathscr{E}$. To this end, we require that the (elementary) topos $\mathscr{S}$ satisfy an additional assumption, to wit, that the stack completions of groupoids in $\mathscr{S}$ be (the externalizations of) groupoids in $\mathscr{S}$. Without it, the main result about the fundamental groupoid is stated in the form "is weakly equivalent to a prodiscrete groupoid" (as in [9]) rather than more simply in the form "is a prodiscrete groupoid".

The notion of a stack ("champ" in the work of Grothendieck and Giraud) can be made relative to the regular epimorphism topology of a given topos $\mathscr{S}$. Stacks of category objects (for the regular epimorpisms topology of $\mathscr{S}$) are discussed in [16], and stack completions shown to exist (and described) in [17]. Roughly speaking, stacks are *good* ($\mathscr{S}$-indexed)categories in the sense that "locally in it" implies "in it".

We are here interested only in groupoids. We begin by making some remarks about stack completions in the case of groupoids in $\mathscr{S}$. For G a groupoid in $\mathscr{S}$ (regarded as a category object), its stack completion is given (up to equivalence) by any pair $< \mathscr{A}, F >$, with $\mathscr{A}$ an $\mathscr{S}$-indexed category, $F : [G] \to \mathscr{A}$ an $\mathscr{S}$-indexed functor, and where $\mathscr{A}$ is a stack (of groupoids) and $F : [G] \to \mathscr{A}$ a weak equivalence. The stack completion of a groupoid G in $\mathscr{S}$, as a category object in which idempotents split, is identified in [17] with the $\mathscr{S}$-indexed category $\mathrm{Pointess}_{\mathscr{S}}(\mathscr{B}G)$ of $\mathscr{S}$-essential points of $\mathscr{B}G$. Also the $\mathscr{S}$-indexed category $\mathrm{Tors}(G)$, given by assigning to an object I of $\mathscr{S}$ the category $\mathrm{Tors}(G)^I$ of G-torsors in $\mathscr{S}/I$, and letting transition morphisms given by pullback, is a stack completion of G. Diaconescu's theorem [27], stating that for a groupoid G the topos $\mathscr{B}G$ classifies G-torsors, suggests yet another identification of the stack completion of $[G]$, as in the following theorem.

Theorem 3.1 *Let G be a groupoid in $\mathscr{S}$ given by a diagram*

$$G_1 \times_{G_0} G_1 \;\substack{\longrightarrow \\ \longrightarrow \\ \longrightarrow}\; G_1 \;\substack{\longrightarrow \\ \longleftarrow \\ \longrightarrow}\; G_0 \,.$$

Then, the stack completion of G is given by the $\mathscr{S}$-indexed category of points of $\mathscr{B}G$.

Proof There is a canonical bag of points $p_G : \mathscr{S}/G_0 \to \mathscr{B}G$ such that composing with it defines an $\mathscr{S}$-indexed functor

$$F : [G] \to \mathrm{Points}_{\mathscr{S}}(\mathscr{B}G).$$

Explicitly, given an object I of $\mathscr{S}$, $F^I : [I, G] \to \mathbf{Top}_{\mathscr{S}}(\mathscr{S}/I, \mathscr{B}G)$ assigns, to an object $x : I \to G_0$ of the category $[I, G]$, the composite $\mathscr{S}/I \to \mathscr{S}/G_0 \to \mathscr{B}G$ of the geometric morphism induced by x with p_G, and to a morphism $f : x \to y$ in $[I, G]$, given by a morphism $f : I \to G_1$ in $\mathscr{S}$ such that $d_0 f = x$ and $d_1 f = y$, the natural transformation $f : p_G x \to p_G y$. This assignment is easily checked to be $\mathscr{S}$-indexed.

It follows just as in [8] (Proposition 3.1) that $\mathrm{Points}_{\mathscr{S}}(\mathscr{B}G)$ is a stack, using for this now that the $\mathscr{S}$-essential surjections (a particular case of open surjections) are of effective descent in $\mathbf{Top}_{\mathscr{S}}$ [28].

It remains to verify that F is a weak equivalence. That for each I the functor F^I is fully faithful is a direct consequence of the etale completeness of G, which means that the square in $\mathbf{Top}_{\mathscr{S}}$ given below

$$\begin{tikzcd} \mathscr{S}/G_1 \arrow[r, "d_1"] \arrow[d, "d_0"'] & \mathscr{S}/G_0 \arrow[d, "p_G"] \\ \mathscr{S}/G_0 \arrow[r, "p_G"'] & \mathscr{B}G \end{tikzcd} \tag{1}$$

is a pullback, and which holds for any discrete (localic) groupoid G. We show next that F is essentially surjective. Let $q : \mathscr{S}/I \to \mathscr{B}G$ be a geometric morphism, that is, an object of $\mathrm{Points}_{\mathscr{S}}(\mathscr{B}G)^I$. Consider the pullback of q along p_G, which we claim looks like this:

$$\begin{tikzcd} \mathscr{S}/J \arrow[r, "r"] \arrow[d, "\alpha"] & \mathscr{S}/G_0 \arrow[d, "p_G"] \\ \mathscr{S}/I \arrow[r, "q"'] & \mathscr{B}G \end{tikzcd} \tag{2}$$

where $\alpha : J \to I$ is an epimorphism in $\mathscr{S}$ and where $r : \mathscr{S}/J \to \mathscr{S}/G_0$ is induced by some $y : J \to G_0$, so that $F^J(y) \cong \alpha^*(q)$, or F is indeed essentially surjective. The claim is true because the canonical p_G is a surjective local homeomorphism, a fact already noticed in [9] and used in the proof of Proposition 2.8 in the previous section, and that therefore so is the geometric morphism opposite it in the pullback; it follows that the latter is induced by an epimorphism $\alpha : J \to I$ in $\mathscr{S}$, as claimed. Furthermore, the geometric morphism $r : \mathscr{S}/J \to \mathscr{S}/G_0$ opposite q in the pullback is necessarily induced by a morphism $y : J \to G_0$ in $\mathscr{S}$, hence the result. □

Corollary 3.2 *Let G be a groupoid in $\mathscr{S}$ and let $\mathscr{G} = \mathscr{B}G$ be its classifying topos. Then, the following hold:*

1. *Any geometric morphism $p : \mathscr{S}/I \to \mathscr{G}$ of $\mathscr{G}$ is $\mathscr{S}$-essential.*
2. *With p as above, the inverse image part $p^* : \mathscr{G} \to \mathscr{S}/I$ is naturally equivalent to a functor*

$$\mathrm{Hom}_{\mathscr{G}/g^*(I)}(A \to g^*(I), (-) \times g^*(I)) : \mathscr{G} \to \mathscr{S}/I$$

where $A \to g^*(I)$ has the structure of a G-torsor in $\mathscr{G}/g^*(I)$.

Proof 1. Since the $\mathscr{S}$-indexed category $\mathrm{Points}_{\mathscr{S}}(\mathscr{B}G)$ and its sub-indexed category $\mathrm{Pointess}_{\mathscr{S}}(\mathscr{B}G)$ are both stack completions of (the externalization $[G]$ of) G, they must be equivalent.

2. The $\mathscr{S}$-indexed category $\mathrm{Tors}(G)$ is also a stack completion of $[G]$ and a full subcategory of $\mathscr{B}G = \mathscr{S}^{G^{op}}$, hence equivalent to $\mathrm{Pointess}_{\mathscr{S}}(\mathscr{B}G)$ by an equivalence which gives the desired representability of inverse image parts of $\mathscr{S}$-essential (bags of) points by torsors.

□

Definition 3.3 A topos $\mathscr{S}$ is said to satisfy the *(ASC)* (or the axiom of stack completions) if for any groupoid G in $\mathscr{S}$ there exists (i) a groupoid $\tilde{G}$ in $\mathscr{S}$ whose externalization $[\tilde{G}]$ is a stack, and (ii) a weak equivalence internal functor (or homomorphism of groupoids) $G \to \tilde{G}$. (It is known that any Grothendieck topos $\mathscr{S}$ satisfies this axiom on account of the existence of a small generating class [18, 27].)

The *Morita equivalence theorem* from [17] takes on the following form in the case of groupoids.

Theorem 3.4 *Let $\mathscr{S}$ be a topos which satisfies the axiom of stack completions in the sense of Definition 3.3. Let G and H be groupoids in $\mathscr{S}$. Then there exists an equivalence of categories*

$$\mathrm{Hom}_{\mathscr{S}}(\tilde{G}, \tilde{H}) \simeq \mathbf{Top}_{\mathscr{S}}(\mathscr{B}G, \mathscr{B}H)$$

where $\tilde{G}$ and $\tilde{H}$ are the stack completions of G and H.

We now turn to *topos cohomology* of a locally connected topos $\mathscr{E}$ defined over a base topos $\mathscr{S}$, of which we shall furthermore assume satisfies (ASC) in the sense of Definition 3.3. For each discrete group K, the cohomology of $\mathscr{E}$ with coefficients in

K is denoted by $H^1(\mathscr{E}; K)$ and defined as the object in $\mathscr{S}$ of connected components of the category $\mathrm{Tors}(\mathscr{E}; K)$, that is, as the "set" of isomorphism classes of K-torsors in $\mathscr{E}$.

Consider now the equivalence

$$H^1(\mathscr{E}; K) \simeq \Pi_0(\mathrm{colim}\{\mathrm{Tors}(\mathscr{E}; K)^U\})$$

where the indexing is taken over a (cofinal set in a) generating poset of covers U in $\mathscr{E}$, and where $\mathrm{Tors}(\mathscr{E}; K)^U$ is the groupoid of U-split K-torsors in $\mathscr{E}$. It is now *a priori* clear that the question of representing the functor $H^1(\mathscr{E}; -)$ by some suitable (prodiscrete localic) groupoid $\pi_1^c(\mathscr{E})$ reduces to that of obtaining discrete groupoids π_U representing the functors $\mathrm{Tors}(\mathscr{E}; -)^U : \mathrm{Gps}(\mathscr{S}) \to \mathrm{Gpds}(\mathscr{S})$, since then the desired groupoid $\pi_1^c(\mathscr{E})$ would be definable as a limit of a filtered system of the π_U.

The central idea in definining the fundamental groupoid of $\mathscr{E}$ is contained in the pushouts $(*)$ of the previous section and in the Galois groupoids G_U associated with them as in Proposition 2.8. The G_U are etale complete groupoids (hence Galois groupoids), and the toposes $\mathscr{B}(G_U)$ their classifying toposes.

Let $U \leq V$, where U and V are covers in $\mathscr{E}$, and let $\alpha : U \to V$ be a given morphism in $\mathscr{E}$ witnessing the relation. Then there is an induced geometric morphism $\varphi_\alpha : \mathscr{G}_U \to \mathscr{G}_V$ over $\mathscr{S}$ which commutes with the bags of points p_U and p_V, and which therefore may be interpreted as an object of $\mathbf{Top}_{\mathscr{S}}[\mathscr{B}G_U, \mathscr{B}G_V]_+$, the full subcategory of $\mathbf{Top}_{\mathscr{S}}[\mathscr{B}G_U, \mathscr{B}G_V]$ whose objects are geometric morpisms commuting with the canonical (bags of) points. The square brackets in both cases indicate that the morphisms in those Hom-categories are to be taken to be iso 2-cells in $\mathbf{Top}_{\mathscr{S}}$.

This is clear from the cube below, using the pushout property $(*)$ of the front face:

$$\begin{tikzcd}
& \mathscr{E}/V \arrow[rr] \arrow[dd] & & \mathscr{E} \arrow[dd] \\
\mathscr{E}/U \arrow[ur] \arrow[rr, crossing over] \arrow[dd] & & \mathscr{E} \arrow[ur, equal] & \\
& \mathscr{S}/e_!V \arrow[rr, dashed] & & \mathscr{G}_V \\
\mathscr{S}/e_!U \arrow[ur] \arrow[rr] & & \mathscr{G}_U \arrow[ur] \arrow[from=uu, crossing over] &
\end{tikzcd}$$

As shown in [8] there is a strong equivalence of categories:

$$\mathrm{Hom}(G_U, G_V) \simeq \mathbf{Top}_{\mathscr{S}}[\mathscr{B}(G_U), \mathscr{B}(G_V)]_+.$$

We derive from it that there are induced groupoid homomorphisms $g_\alpha : G_U \to G_V$, for each α. Furthermore, since the geometric morphisms φ_α are connected (and locally connected), the homomorphisms g_α are *full and essentially surjective* by [32].

From the above diagram follows, by virtue of the connectedness of $\sigma_U : \mathscr{E} \to \mathscr{G}_U$, that the geometric morphisms φ_α depend on α up to unique iso 2-cell. In particular, the groupoid homomorphisms g_α, which are determined by the geometric morphisms φ_α up to iso 2-cells, also depend on the α up to unique iso 2-cells. This and the cofinality of $\mathrm{Cov}(\mathscr{E})$ implies that the 2-system is bifiltered (and biordered) in the sense of the following definition [29].

Definition 3.5 A 2-system $\{G_i\}$ of discrete groupoids, groupoid homomorphisms and 2-cells between them, indexed by a category $\mathscr{C}$, is said to be *bifiltered and biordered* if

1. For any two groupoids G_i , G_j in the system, there is $k \in \mathscr{C}$ and morphisms $\alpha : i \to k$ and $\beta : j \to k$, inducing homomorphisms $g_\alpha : G_i \to G_k$, $g_\beta : G_j \to G_k$.
2. For any other morphism $\alpha' : i \to j$, and corresponding $g'_\alpha : G_i \to G_j$, there exists a unique (iso) 2-cell $\lambda_{i,i'} : g_\alpha \to g'_\alpha$.

We have also that for any group K in $\mathscr{S}$ there is an equivalence

$$\mathbf{Top}_{\mathscr{S}}(\mathscr{B}(G_U), \mathscr{B}(K)) \simeq \mathrm{Tors}(\mathscr{E}; K)^U$$

natural in K. Indeed, by Diaconescu's theorem [27], the topos $\mathscr{B}(K)$ classifies K-torsors in $\mathscr{E}$, and under this correspondence, the U-split K-torsors in $\mathscr{E}$ are in bijection with those geometric morphisms $\varphi : \mathscr{E} \to \mathscr{B}(K)$ whose inverse image part φ^* factors through the inclusion of $\mathrm{Spl}(U)$ in $\mathscr{E}$ and the latter is equivalent to $\mathscr{B}(G_U)$); notice that the canonical $p_G : \mathscr{S}/G_0 \to \mathscr{B}(G_U)$ has no bearing in this equivalence.

It follows that the Galois groupoids G_U only *weakly* represent cohomology. In order to get a strong representation, we must consider their stack completions.

Lemma 3.6 *Let $\mathscr{S}$ be a topos satisfying (ASC) in the sense of Definition 3.3. Let $\mathscr{E}$ be a locally connected topos over $\mathscr{S}$, U a cover in $\mathscr{E}$, and G_U the Galois groupoid obtained in Proposition 2.8. Then, if $F_U : G_U \to \tilde{G_U}$ is the stack completion of G_U, the discrete groupoid $\pi_U = \tilde{G_U}$ represents the functor*

$$\mathrm{Tors}(\mathscr{E}; -)^U : \mathrm{Gps}(\mathscr{S}) \to \mathrm{Gpds}(\mathscr{S}).$$

Further, the weak equivalence $F_U : G_U \to \tilde{G_U}$ induces an equivalence $\tilde{F_U} : \mathscr{B}(G_U) \to \mathscr{B}(\tilde{G_U})$.

Proof It follows from Theorem 3.4 that for the groupoid G_U and any discrete group K, there is a strong equivalence

$$\mathrm{Hom}(\tilde{G_U}, K) \simeq \mathbf{Top}_{\mathscr{S}}[\mathscr{B}(G_U), \mathscr{B}(K)]$$

and therefore, also a strong equivalence

$$\mathrm{Hom}(\tilde{G_U}, K) \to \mathrm{Tors}(\mathscr{E}; K)^U.$$

The equivalence of the classifying toposes is a consequence of the fact that $\mathscr{S}$ is a stack and therefore the functor $\mathscr{B}(-) \simeq \mathscr{S}^{(-)^{op}}$ carries weak equivalence functors into (strong) equivalence functors. □

Definition 3.7 Denote by $\Pi_1^c(\mathscr{E})$ the limit in $\mathbf{Top}_{\mathscr{S}}$ of the filtered system of toposes $\mathrm{Spl}(U)$ and connected locally connected geometric morphisms $h_{UV} : \mathrm{Spl}(U) \to \mathrm{Spl}(V)$ for each pair of covers U, V with $U \leq V$, indexed by a cofinal generating poset of covers in $\mathscr{E}$. The topos $\Pi_1^c(\mathscr{E})$ is said to be the *coverings fundamental group topos* of $\mathscr{E}$ over $\mathscr{S}$. It is equivalent, as a category, to the full subcategory of $\mathscr{E}$ generated by the locally constant covers in $\mathscr{E}$, in the sense of Definition 2.1. There is a connected locally connected geometric morphism $\sigma_{\mathscr{E}} : \mathscr{E} \to \Pi_1^c(\mathscr{E})$.

The main theorem of [9] may now be stated as follows, under the assumption that the base topos $\mathscr{S}$ satisfies (ASC).

Theorem 3.8 *Let $\mathscr{E}$ be a locally connected topos over a base topos $\mathscr{S}$, where $\mathscr{S}$ is assumed to satisfy (ASC) in the sense of Definition 3.3. Then the coverings fundamental group topos $\Pi_1^c(\mathscr{E})$ defined in Definition 3.7 is the classifying topos of a prodiscrete localic groupoid $\pi_1^c(\mathscr{E})$ which represents first-degree cohomology of $\mathscr{E}$ with coefficients in discrete groups.*

Proof Let $\{\pi_U\}$ be system of groupoids obtained from the system of groupoids $\{G_U\}$ as in Lemma 3.6, that is, by taking stack completions, and where the transition homomorphisms are homomorphisms $g_{\tilde{U}V} : \pi_U \to \pi_V$ induced by the $g_\alpha : G_U \to G_V$ for *any* α witnessing $U \leq V$, as follows from the Morita equivalence theorem for discrete groupoids. Since the $g_\alpha : G_U \to G_V$ are full and essentially surjective homomorhisms, the same is true of the homomorphisms $g_{UV} : \pi_{\tilde{U}} \to \pi_V$. Furthermore, since the original 2-system is *bi*filtered and *bi*ordered solely on account of the dependency of the transition homomorphisms g_α on the witnessing morphisms $\alpha : U \to V$, the 1-system of the stack completions, whose transition homomorphisms $g_{\tilde{U}V}$ only depend on $U \leq V$, is simply filtered.

Let $\pi_1^c(\mathscr{E}) = \lim\{\pi_U\}$ be the limit of the filtered system of discrete groupoids, taken in the category of *localic groupoids*. We now prove that the prodiscrete groupoid $\pi_1^c(\mathscr{E})$ represents cohomology of $\mathscr{E}$ with coefficients in discrete groups. Explicitly, the claim is that for a discrete group K, there is an isomorphism

$$H^1(\mathscr{E}; K) \simeq [\pi_1^c(\mathscr{E}), K].$$

This follows from Lemma 3.6 and various canonical isomorphisms:

$$\begin{aligned}
[\pi_1^c(\mathscr{E}), K] &\simeq \Pi_0(\mathrm{Hom}(\pi(\mathscr{E}), K)) = \Pi_0(\mathrm{Hom}(\lim\{\pi_U\}, K)) \\
&\simeq \Pi_0(\mathrm{colim}\{\mathrm{Hom}(\pi_U, K)\}) \\
&\simeq \Pi_0(\mathrm{colim}\{\mathrm{Tors}(\mathscr{E}; K)^U\}) \simeq \Pi_0(\mathrm{Tors}(\mathscr{E}; K)) \simeq H^1(\mathscr{E}; K)
\end{aligned}$$

□

In the rest of the paper we shall be dealing for the most part with locally simply connected toposes. We recall the definition.

Definition 3.9 A locally connected topos $\mathscr{E}$ over $\mathscr{S}$ is said to be *locally simply connected* if there is a single cover U in $\mathscr{E}$ which splits all locally constant objects in $\mathscr{E}$ in the sense of Definition 2.1.

4 Galois groupoids and Galois toposes

In this section we compare, in the case of a (locally connected and) locally simply connected topos $\mathscr{E}$ over $\mathscr{S}$, the Galois groupoid of [24] given any $e_!$-normal cover in $\mathscr{E}$, with the fundamental groupoid $\pi_1^c(\mathscr{E})$; in particular, this applies to the universal cover in $\mathscr{E}$, easily seen to be $e_!$-normal cover in $\mathscr{E}$.

Recall (e.g., from [8]) the notion of a G-torsor in $\mathscr{E}$, for G a groupoid in $\mathscr{E}$ given by means of a diagram

$$G_1 \times_{G_0} G_1 \xrightarrow[\longrightarrow]{\longrightarrow} G_1 \underset{\longrightarrow}{\overset{\longrightarrow}{\longleftarrow}} G_0 \,.$$

A (right) G-object $\mathbf{T}$ is given by the data $< \gamma : T \to G_0 \;,\; \theta >$ with T an object of $\mathscr{E}$, $p : T \to 1$ is the unique morphism into the terminal object 1, and $\theta : T \times_{G_0} G_1 \to T$

an action (unitary and associative). We say that $\mathbf{T}$ is a *G-torsor* in $\mathscr{E}$ if it is a right G-object for which $p : T \to 1$ is an epimorphism and

$$< \mathrm{proj}_1 \ , \ \theta >: T \times_{G_0} G_1 \to T \times T$$

is an isomorphism.

Definition 4.1 Let $e : \mathscr{E} \to \mathscr{S}$ be a locally connected geometric morphism. By an *$\mathscr{S}$-Galois family* of $\mathscr{E}$ we mean an object X of $\mathscr{E}$ of global support, equipped with a morphism $\delta : X \to e^*I$, such that the pair $(X \ , \ \delta)$ is both connected and a $\mathrm{Aut}(\delta)$-torsor in $\mathscr{E}/e^*I$ over $\mathscr{S}$. Explicitly, for the the canonical action θ of $e^*(\mathrm{Aut}(\delta))$ (a discrete groupoid in $\mathscr{E}$ with object of objects e^*I) on δ, the induced

$$< \mathrm{proj}_1 \ , \ \theta >: X \times_{e^*I} e^*(Aut(\delta)) \to X \times X$$

is an isomorphism. In particular, X is a locally constant object in the sense of Definition 2.1 and a cover in $\mathscr{E}$.

We omit the subscripts U in what follows for the geometric morphisms σ_U and p_U which occur in the pushout diagram $(*)$ defining $\mathscr{G}_U$.

Denote by $g : \mathscr{G}_U \to \mathscr{S}$ the structure geometric morphism. There is a natural isomorphism $g\sigma \simeq e$. Since $\sigma : \mathscr{E} \to \mathscr{G}_U$ is connected and locally connected and since $e : \mathscr{E} \to \mathscr{S}$ is locally connected, it follows that $g : \mathscr{G}_U \to \mathscr{S}$ is locally connected. We will use this in what follows.

Theorem 4.2 *Let $\mathscr{E}$ be a locally connected topos bounded over $\mathscr{S}$ by means of a geometric morphism $e : \mathscr{E} \to \mathscr{S}$. Let U be a cover in $\mathscr{E}$. Let $p : \mathscr{S}/e_!U \to \mathscr{G}_U$ be the discrete point that arises in the pushout definition $(*)$. Then, The morphism $p : \mathscr{S}/e_!U \to \mathscr{G}_U$ is a surjective local homeomorphism, its inverse image part is represented (uniquely up to isomorphism) by an $\mathscr{S}$-Galois family $(\sigma^*A, \sigma^*\zeta)$ and the square*

$$\begin{tikzcd}
\mathscr{E}/\sigma^*A \arrow[r, "\varphi_{\sigma^*A}"] \arrow[d, "\rho_{\sigma^*\zeta}"'] & \mathscr{E} \arrow[d, "\sigma"] \\
\mathscr{S}/e_!U \arrow[r, "p"'] & \mathscr{G}_U
\end{tikzcd}$$

is a pullback. In addition, $\mathscr{G}_U$ is generated by a single $\mathscr{S}$-Galois family $(\sigma^*A, \sigma^*\zeta : \sigma^*A \to e^*e_!U)$ and is the classifying topos of the discrete localic groupoid $\mathrm{Aut}(\sigma^*\zeta)$.

Proof It follows from Lemma 2.5 that p is locally connected therefore $\mathscr{S}$-essential; it is also surjective. Therefore, by [7] Proposition 1.2, its inverse image p^* is represented by a family $\zeta\colon A \to g^*e_!U$ in the sense that there is a natural isomorphism

$$p^* \simeq \hom_{\mathscr{G}_U(\mathscr{E})/g^*e_!U}(\zeta \ , \ \Delta_{e_!(U)}(-))$$

where $\Delta_{e_!(U)}(X) = X \times g^*e_!(U)$, equipped with the appropriate projection. The morphism $\zeta\colon A \to g^*e_!U$ represents p^* which has a right adjoint; in particular, $\zeta\colon A \to g^*e_!U$ is connected in $\mathscr{G}_U/g^*e_!U$. From the remark that there is a natural isomorphism

$$\hom_{\mathscr{G}_U/g^*e_!U}(\zeta \ , \ \Delta_{e_!(U)}(-)) \simeq \hom_{\mathscr{G}_U}(A, -)$$

and the faithfulness of p follows that A has global support. An explicit description of the representing family is given in the proof of [8] Theorem 1.2; let $\zeta = (p^{e_!U})_!(\delta_{e_!U} : e_!U \to e_!U \times e_!U)$.

We shall also regard this family as one in $\mathscr{E}/e^*e_!U$ by applying the fully faithful σ^*.

That p is a local homeomorphism is proven in [9] by recourse to the notion of a totally disconnected geometric morphism. In particular, there is induced an equivalence $\mathscr{G}_U/A \simeq \mathscr{S}/e_!U$ which identifies the localic point p with the local homeomorphism $\varphi_A : \mathscr{G}_U/A \to \mathscr{G}_U$. The statement about the mentioned square being a pullback is a consequence of this.

Also a consequence of this is the claim that $(\sigma^*A, \sigma^*\zeta)$ is a torsor, as can be seen by evaluating the naturally isomorphic functors p^* and $\varphi^*\sigma^*$ at σ^*A; this gives the desired isomorphism

$$< \mathrm{proj}_1 \ , \ \theta >: \sigma^*A \times_{e^*e_!U} e^*(Aut(\sigma^*\zeta)) \to \sigma^*A \times \sigma^*A$$

as in Definition 4.1.

The family $\zeta : A \to g^*e_!U$ representing p^* is generating for $\mathscr{G}_U$ since p^* is faithful. Hence $\mathscr{G}_U$ is generated by a single $\mathscr{S}$-Galois family.

Finally, since p is a local homeomorphism hence open, and since it is surjective, it is of effective descent by [28]. It follows that there is an equivalence $\mathscr{G}_U \simeq \mathscr{B}(G_U)$ where $G_U = \mathrm{Aut}(p)$ is a localic groupoid, discrete since p is (not just open but) a local homeomorphism.

From the fact that ζ represents p^* now follows that $\mathrm{Aut}(p)$ and $\mathrm{Aut}(\sigma^*\zeta)$ are equivalent groupoids in $\mathscr{S}$. □

The following proposition states that the universal cover $\sigma^*\zeta : \sigma^*A \to e^*e_!U$ in the locally simply connected topos $\mathscr{E}$ is $e_!$-normal in the sense of [24]. This is immediate from the properties of universal covers; we give below an alternative and direct proof of the assertion in question using only part of the data supplied by the pushout square $(*)$ in $\mathbf{Top}_{\mathscr{S}}$.

Lemma 4.3 *The $\mathscr{S}$-Galois family ζ representing p^* as in Theorem 4.2 may be alternatively obtained as the morphism*

$$\sigma^*(u_{\sigma_!U}) : \sigma^*\sigma_!U \to \sigma^*g^*g_!\sigma_!U \simeq e^*e_!U,$$

where u is the unit of the adjointness $g_! \dashv g^*$.

Proof This follows readily from Theorem 4.2 and the commutativity of the pushout square $(*)$, in particular, from the natural isomorphism $p_!\rho_! \simeq \sigma_!\varphi_{U!}$. □

Proposition 4.4 *Let $\mathscr{E}$ be locally simply connected with U as a single cover splitting all locally constant objects. Let (A, ζ) be the $\mathscr{S}$-Galois family of $\mathscr{E}$ with ζ representing $p : \mathscr{S}/e_!U \to \mathscr{G}_U = \mathscr{B}(\pi_1^c(\mathscr{E}))$. Then (A, ζ), which we identify with $(\sigma^*A, \sigma^*\zeta)$ if regarded in $\mathscr{E}$, is a universal cover in the sense that σ^*A is a cover, $(\sigma^*A, \sigma^*\zeta)$ is a normal object for the adjoint pair $e_! \dashv e^*$, and σ^*A splits every locally constant object of $\mathscr{E}$.*

Proof It is shown in Theorem 4.2 that the discrete groupoid G_U is the groupoid of automorphisms of the localic point p and, since the latter is represented by an $\mathscr{S}$-Galois family $\zeta : A \to g^*e_!U$, G_U is also realized as the discrete groupoid $\mathrm{Aut}(\zeta)$. Furhermore, a concrete instance of such a family is $\sigma^*(u_{\sigma_!U}) : \sigma^*\sigma_!U \to e^*e_!U$.

We now show that $\sigma^*A = \sigma^*\sigma_!U$ has global support and is a normal object (for the adjoint pair $e_! \dashv e^*$). For the former, notice that since σ is connected and locally connected (in particular $\sigma_!1 \simeq 1$) and U has global support, $\sigma^*A = \sigma^*\sigma_!U \to \sigma^*\sigma_!1 \simeq 1$ is an epimorphism.

For the latter, observe that since σ^* is fully faithful and $A = \sigma_! U \in \mathrm{Spl}(U)$, $\sigma^* A \simeq \sigma^*\sigma_! U$ is U-split. Use now Lemma 5.3 of [14] applied to σ (since it is locally connected, and in particular the adjoint pair $\sigma_! \dashv \sigma^*$ is $\mathcal{S}$-indexed); it says in this case that for any object X of $\mathcal{G}_U$, U splits $\sigma^* X$ if and only if $\sigma_! U$ splits X. In particular, since U splits $\sigma^* A = \sigma^*\sigma_! U$, it follows that $\sigma_! U$ splits $\sigma_! U$, hence (by stability) $\sigma^* A = \sigma^*\sigma_! U$ splits $\sigma^* A = \sigma^*\sigma_! U$. □

Corollary 4.5 *Let $\mathcal{E}$ be a locally connected and locally simply connected topos over $\mathcal{S}$. Let U be a single cover which splits all locally constant objects of $\mathcal{E}$. Let A be the universal cover. Let $H(\mathcal{E}) = \mathrm{Gal}(\sigma^*\zeta)$ be the Galois groupoid obtained as in [24] using that $\sigma^*\zeta$ is an $e_!$-normal object, and let $G = \pi_1^c(\mathcal{E})$. Then, there is an equivalence $H \simeq G$, explicitly,*

$$\mathrm{Gal}(\sigma^*\zeta) \simeq \pi_1^c(\mathcal{E}).$$

Proof This follows readily from the isomorphism

$$\sigma^* A \times_{e^* e_! U} e^*(Aut(\sigma^*\zeta)) \to \sigma^* A \times \sigma^* A$$

applying the functor $e_!$ to both sides to get

$$e_!(\sigma^* A \times_{e^* e_! U} e^*(Aut(\sigma^*\zeta))) \simeq e_!(\sigma^* A \times \sigma^* A)$$

and then using that

$$e_! \sigma^* A \simeq g_! \sigma_! \sigma^* \sigma_! U \simeq g_! \sigma_! U \simeq e_! U$$

and *Frobenius Reciprocity* for the $\mathcal{S}$-indexed adjoint pair $e_! \dashv e^*$ to obtain

$$e_!(\sigma^* A \times_{e^* e_! U} e^*(Aut(\sigma^*\zeta))) \simeq e_! U \times_{e_! U} Aut(\sigma^*\zeta) \simeq Aut(\sigma^*\zeta)$$

hence the desired isomorphism

$$Aut(\sigma^*\zeta) \simeq e_!(\sigma^* A \times \sigma^* A).$$

□

Remark 4.6 We now sum up our argument leading to the Galois groupoid in the rich context of a locally connected topos, while comparing it with the purely categorical Galois theory of Janelidze.

In the case of a locally connected topos $e : \mathcal{E} \to \mathcal{S}$, we also start with a pure set-up given by an adjoint pair $e_! \dashv e^*$ satisfying certain conditions explicitated in [23]. In addition, we use, *primo*, that from the $\mathcal{S}$-indexedness of the adjoint pair $e_! \dashv e^*$ follows the existence of a normal cover (X, δ), and that in consequence one gets (as in [24]) a groupoid $\mathrm{Gal}(X, \delta)$ for which the "fundamental theorem of Galois theory" holds; *secundo*, that using the further right adjoint e_* and everything that derives from it in the pushout $(*)$, one has that (X, δ) is σ^* applied to a family which represents (the inverse image part of) the surjective $\mathcal{S}$-essential point p_U of the topos $\mathrm{Spl}(U)$ and is therefore an $\mathcal{S}$-ATO in the appropriate slice category, and *tertio*, from the additional observation that p_U is a local homeomorphism follows that (X, δ) is a Galois family (a torsor, not just an $\mathcal{S}$-ATO) in $\mathcal{E}$, and that $\mathrm{Gal}(X, \delta)$ *is* equivalent to the (Galois) groupoid of automorphisms of the universal cover.

In categorical Galois theory there is another way to justify these steps given the assumption of normality, as pointed out to me by Janelidze. To give a concrete situation consider the "most classical" example, namely the topos $\mathrm{Sh}(B)$ of sheaves on a connected and locally (simply) connected space B admitting a universal covering $p : E \to B$. According to the categorical Galois theory, the Galois group is defined as $e_!(E \times_B E)$, but classically (Chevalley-Grothendieck) it is to be defined as the automorphism group $\mathrm{Aut}_B(E)$. From the fundamental theorem of Galois theory, Janelidze obtains (i) the category equivalence $\mathrm{Spl}(E,p) \cong \mathrm{Set}^G$ for $G = e_!(E \times_B E)$; (ii) under this equivalence, (E,p) corresponds to G acting on itself via its multiplication; (iii) therefore the automorphism group $\mathrm{Aut}(E,p) = \mathrm{Aut}_B(E)$ is isomorphic to $\mathrm{Aut}(G)$, the automorphism group of G considered as a G-set as above, and (iv) since $\mathrm{Aut}(G)$ is isomorphic to the group G, it can be concluded that $\mathrm{Aut}_B(E) \cong e_!(E \times_B E)$. The same argument can be used when the base category $\mathscr{S}$ is not necessarily Set but is cartesian closed. in conclusion, the full topos setting (and in particular, the right adjoint e_* and the Joyal-Tierney theorem) is certainly not *needed*, contrary to what I had previously implied.

We will now consider an unpointed version of a notion of a Galois topos introduced by Moerdijk [31] following Grothendieck [1].

Definition 4.7 A topos $\mathscr{E}$ bounded over $\mathscr{S}$ will be said to be a *$\mathscr{S}$-Galois topos* if $\mathscr{E}$ is locally connected and has an internal site $\mathbf{C}$ of definition determined by objects A of $\mathscr{E}$ for which there exists a morphism $\zeta : A \to e^*I$ which is an $\mathscr{S}$-Galois family in $\mathscr{E}$ in the sense of Definition 4.1.

We set out to *characterize* $\mathscr{S}$-Galois toposes. For this we need to observe the behaviour of $\mathscr{S}$-Galois objects under the transition morphisms in the system whose limit gives the fundamental groupoid.

Lemma 4.8 *Let $\mathscr{E}$ and $\mathscr{F}$ be locally connected toposes over $\mathscr{S}$ with corresponding structure maps e and f. Let $\varphi : \mathscr{F} \to \mathscr{E}$ be a connected and locally connected geometric morphism over $\mathscr{S}$ and let $\zeta : A \to e^*I$ be any $\mathscr{S}$-Galoisfamily in $\mathscr{E}$. Then,*

$$\varphi^*A \xrightarrow{\varphi^*\zeta} \varphi^*e^*I \cong f^*I$$

is an $\mathscr{S}$-Galois family in $\mathscr{F}$.

Proof Observe that $\varphi^*(Aut(\zeta)) \simeq Aut(\varphi^*(\zeta))$ since φ^* is a fully faithful left exact functor which has a right adjoint. Observe also that, since φ is connected and locally connected, if $\zeta : A \to e^*I$ is connected in $\mathscr{E}/e^*I$, then $\varphi^*(\zeta) : \varphi^*(A) \to \varphi^*(e^*I) = f^*I$ is connected in $\mathscr{F}/f^*I$. $\square$

Proposition 4.9 *The limit of an inversely filtered system of $\mathscr{S}$-Galois toposes of the form $\mathscr{B}(G_\alpha)$ with G_α a discrete groupoid in $\mathscr{S}$, and connected locally connected transition morphisms between them, is an $\mathscr{S}$-Galois topos.*

Proof For each $\mathscr{G}_\alpha$ there is a single generating $\mathscr{S}$-Galois family $\zeta_\alpha : A_\alpha \to e^*(I_\alpha)$ by assumption. Moreover, for each α one has the canonical internal site given by the discrete groupoid $\mathrm{Aut}(\zeta_\alpha)$. By the construction of filtered inverse limits of toposes in terms of (internal) sites [33], it follows that these families determine families in $\mathscr{G}$ which form an internal site for $\mathscr{G}$. Since the transition morphisms are connected locally connected, it follows from Lemma 4.8 that these are $\mathscr{S}$-Galois objects of $\mathscr{G}$. $\square$

The following theorem is the analogue of [31] Theorem 3.2 in the unpointed case.

Theorem 4.10 *Let $\mathscr{E}$ be a locally connected bounded over $\mathscr{S}$. Then, (1) and (3) are equivalent. If furthermore the base topos $\mathscr{S}$ is assumed to satisfiy the* *axiom of stack completions in the sense of Definition 3.3, then all four conditions are equivalent.*

1. *$\mathscr{E}$ is an $\mathscr{S}$-Galois topos.*
2. *The canonical morphism $\sigma_{\mathscr{E}} : \mathscr{E} \to \Pi_1^c(\mathscr{E})$ is an equivalence.*
3. *$\mathscr{E}$ is generated by its locally constant covers.*
4. *$\mathscr{E}$ is equivalent to a topos $\mathscr{B}(G)$ for G a prodiscrete localic groupoid.*

Proof The implication (1) $\Rightarrow$ (3) follows from Proposition 4.4. The converse (3) $\Rightarrow$ (1) amounts to the existence of a Galois closure in this constructive setting and follows immediately from Theorem 4.2. The implication (2) $\Rightarrow$ (1) follows from the conjunction of Theorem 3.8 and Proposition 4.9. The converse (1) $\Rightarrow$ (2) holds because, from the construction of filtered inverse limits of toposes in terms of sites [33], it follows that the inverse image of $\sigma_{\mathscr{E}}$ establishes an equivalence of sites, hence of the respective toposes. The implications (2) $\Rightarrow$ (4) and (4) $\Rightarrow$ (2) are part of Theorem 3.8. □

Corollary 4.11 *Let $\mathscr{E}$ be a locally connected topos bounded over $\mathscr{S}$. Then, the following are equivalent*

1. *$\mathscr{E}$ is a locally simply connected $\mathscr{S}$-Galois topos.*
2. *$\mathscr{E}$ is locally simply connected and the canonical morphism $\sigma_{\mathscr{E}} : \mathscr{E} \to \Pi_1^c(\mathscr{E})$ is an equivalence.*
3. *$\mathscr{E}$ is locally connected and $\mathscr{E} \simeq \mathscr{E}_{lc}$ as a category, where $\mathscr{E}_{lc}$ is the full subcategory of $\mathscr{E}$ whose objects are the locally constant objects in $\mathscr{E}$.*
4. *$\mathscr{E}$ is the classifying topos $\mathscr{B}(G)$ for G a discrete groupoid in $\mathscr{S}$.*

Remark 4.12 Grothendieck [21] *defines* the fundamental groupoid of a (locally simply connected) Galois (Grothendieck) topos $\mathscr{G}$ as the groupoid of points of $\mathscr{G}$. This is consistent with the results of the last section, particularly Theorem 3.1 and Theorem 3.8, since Set satisfies the axiom of choice, hence trivially the axiom of stack completions. This definition is then extended in [21] to what are called therein "multi-Galois toposes" and which classify the actions of a progroup. In the case of $\mathscr{S}$-Galois toposes for any topos $\mathscr{S}$ (including Set), we are dealing with classifying toposes $\mathscr{B}(\pi)$ of prodiscrete groupoids π, which are localic but not discrete; as proven in [8], the stack completion of such a π (for the topology of open surjections of locales) is obtained by considering the *localic* points of $\mathscr{B}(\pi)$. In the case of $\pi_1^c(\mathscr{E}) = lim\{\pi_U\}$, there is a canonical localic point p of the classifying topos $\Pi_1^c(\mathscr{E}) = lim\{\mathscr{B}(\pi_U)\}$ and this localic point may be said to be a *prodiscrete localic point*, as it is obtained as a limit of (discrete) bags of points [9].

5 Locally paths simply connected toposes over an arbitrary base

Recall the definition of the paths fundamental group topos $\Pi_1^p(\mathscr{E})$ for $\mathscr{E}$ a topos bounded over $\mathscr{S}$ [31]. It is given as the colimit $\tau_{\mathscr{E}} : \mathscr{E} \to \Pi_1^p(\mathscr{E})$ in $\mathbf{Top}_{\mathscr{S}}$ for the descent diagram

$$\mathscr{E}^{\Delta} \;\substack{\longrightarrow\\ \longrightarrow\\ \longrightarrow}\; \mathscr{E}^{I} \;\underset{\varepsilon_1}{\overset{\varepsilon_0}{\rightrightarrows}}\; \mathscr{E}$$

obtained from the cosimplicial locale by taking its exponential into $\mathscr{E}$. The assignment of $\Pi_1^p(\mathscr{E})$ to $\mathscr{E}$ is evidently pseudofunctorial and the geometric morphisms $\tau_{\mathscr{E}} : \mathscr{E} \to \Pi_1^p(\mathscr{E})$ are the components of a natural transformation. In the case of a *connected* locally connected topos, $\tau_{\mathscr{E}} : \mathscr{E} \to \Pi_1^p(\mathscr{E})$ is an open surjection, hence

$\Pi_1^p(\mathscr{E})$ is the classifying topos of an open localic groupoid $\pi_1^p(\mathscr{E})$; we shall not make this assumption here.

The (unique) path-lifting property can be expressed in topos theory. Let Δ_n denote the standard n-simplex locale in the topos $\mathscr{S}$, constructed from the unit interval locale I in the usual manner.

Definition 5.1 (i) An object Y of $\mathscr{E}$ is said to have the *path-lifting property* if the induced geometric morphism

$$(\varphi_Y)^{\Delta_n} \colon (\mathscr{E}/Y)^{\Delta_n} \to (\mathscr{E})^{\Delta_n}$$

is an open surjection.

(ii) An object Y of $\mathscr{E}$ is said to have the *unique path-lifting property* if the commutative diagram

$$\begin{tikzcd}
(\mathscr{E}/Y)^{\Delta_n} \arrow[r, "\varphi_Y^{\Delta_n}"] \arrow[d, "\varepsilon_0"] & \mathscr{E}^{\Delta_n} \arrow[d, "\varepsilon_0"] \\
\mathscr{E}/Y \arrow[r, "\varphi_Y"'] & \mathscr{E}
\end{tikzcd}$$

is a pullback.

Proposition 5.2 *Any cover Y in $\mathscr{E}$ with the unique path-lifting property also satisfies the path-lifting property.*

Proof Open surjections are pullback stable. □

Proposition 5.3 *Let A be a locally constant object in $\mathscr{E}$. Then A has the unique path-lifting property.*

Proof The Lemma 6.1 of [15], established for Grothendieck toposes, is valid over an arbitrary base topos. Indeed, locally constant sheaves in the sense of Definition 2.1 (or "constructible sheaves" in the sense of Grothendieck [1]) are constant on contractible locales, the argument is just as in topology. Furthermore, the required stability holds by Lemma 4.8. □

The *comparison lemma* between the paths and the coverings fundamental group toposes was established in [15] for a *connected* locally connected topos $\mathscr{E}$ over $\mathscr{S}$. In the non-connected case, we need to work with the locally constant objects of Definition 2.1 rather than with those of Barr and Diaconescu [3]. The proof we give here uses Galois closure and the fact that locally constant objects have the (unique) path-lifting property.

Theorem 5.4 *Let $\mathscr{E}$ be a locally connected topos bounded over $\mathscr{S}$. Then there exists a geometric morphism*

$$\kappa_{\mathscr{E}} : \Pi_1^p(\mathscr{E}) \to \Pi_1^c(\mathscr{E})$$

for which the triangle

$$\begin{tikzcd}
\mathscr{E} \arrow[r, "\sigma_{\mathscr{E}}"] \arrow[d, "\tau_{\mathscr{E}}"'] & \Pi_1^c(\mathscr{E}) \\
\Pi_1^p(\mathscr{E}) \arrow[ur, "\kappa_{\mathscr{E}}"']
\end{tikzcd}$$

commutes.

Proof Let U be a cover in $\mathscr{E}$ and let $X \in \mathrm{Spl}_U(\mathscr{E})$. We wish to show that X has a (canonical) action by paths. We resort to Galois closure in $\mathscr{G}_U(\mathscr{E}) \simeq \mathrm{Spl}_U(\mathscr{E})$. Let $\zeta : A \to e^*e_!U$ be a generating $\mathscr{S}$-Galois family for $\mathscr{G}_U(\mathscr{E})$ as obtained in Theorem 4.2. Recall from Theorem 4.2 that the diagram

$$\begin{tikzcd} \mathscr{E}/A \arrow[r, "\varphi_A"] \arrow[d] & \mathscr{E} \arrow[d] \\ \mathscr{S}/e_!U \arrow[r, "p_U"'] & \mathscr{G}_U \end{tikzcd}$$

is a pullback (in particular, it commutes). Consider now the diagram

$$\begin{tikzcd} (\mathscr{E}/A)^\Delta \arrow[r] \arrow[d, shift left=2] \arrow[d] \arrow[d, shift right=2] & (\mathscr{E})^\Delta \arrow[d, shift left=2] \arrow[d] \arrow[d, shift right=2] \\ (\mathscr{E}/A)^I \arrow[r] \arrow[d, shift left=1] \arrow[d, shift right=1] & \mathscr{E}^I \arrow[d, shift left=1] \arrow[d, shift right=1] \\ \mathscr{E}/A \arrow[r] & \mathscr{E} \end{tikzcd}$$

It follows now from Proposition 5.3 and Proposition 5.2 that the three horizontal arrows in the above descent diagram are open surjections, hence of effective descent. Therefore, it is sufficient to prove that $(\varphi_A)^*X$, as an object of $\mathscr{E}/A$ which arises as $\rho^*(\zeta)$ for some $\zeta : S \to e_!U$, has an action by paths. In turn, we claim that is enough to prove that $(\varphi_A)^*X$, as an object of $\mathscr{E}/A$ regarded as a topos over $\mathscr{S}/e_!U$, has an action by paths. This is because, since $\mathscr{S}/e_!U \to \mathscr{S}$ is a surjective local homeomorphism, the top arrow in the pullback square below is also one:

$$\begin{tikzcd} {[(\mathscr{E}/A)^I]_U} \arrow[r] \arrow[d] & \mathscr{E}/A^I \arrow[d, "\varepsilon_0"] \\ \mathscr{S}/e_!U \arrow[r, "\varphi_{e_!U}"'] & \mathscr{S} \end{tikzcd}$$

and similarly for $\Delta = \Delta_1$ and the other simplices locales.

The next observation is that for $i = 0, 1$ the triangle

$$\begin{tikzcd} ((\mathscr{E}/A)^I)_U \arrow[r, "\varepsilon_i"] \arrow[d, "\varepsilon"'] & \mathscr{E}/A \\ \mathscr{E}/A \times_{\mathscr{S}/e_!U} \mathscr{E}/A \arrow[ur, "\pi_i"'] & \end{tikzcd}$$

commutes up to iso 2-cell. By the theorem of Moerdijk and Wraith [34], ε is an open surjection since $\mathscr{E}/A \to \mathscr{S}/e_!U$ is connected locally connected. Therefore, since (for the same reason) the latter is the *coequalizer* of the pair π_0, π_1, it coequalizes the pair $\varepsilon_0, \varepsilon_1$ considered above. It follows easily from this that the object $(\varphi_U)^*X = \rho^*\zeta$ has an action by paths.

The preceding argument shows the existence of a canonical geometric morphism $\kappa_U : \Pi_1^p(\mathscr{E}) \to \mathscr{G}_U$, for each cover U, from which the existence of the desired morphism $\kappa_{\mathscr{E}} : \Pi_1^p(\mathscr{E}) \to \mathscr{G}(\mathscr{E})$ follows. □

An obvious problem with the definition of $\Pi_1^c(\mathscr{E})$ for a locally connected topos $\mathscr{E}$ is the lack of pseudofunctoriality, by contrast with the pseudofunctoriality of $\Pi_1^p(\mathscr{E})$, as noted in [15]. As in topology, it is primarily the *paths* fundamental groupoid that is of interest, with the *coverings* version as a means for calculating the latter for spaces where the two are equivalent.

We have seen in the previous section that there is always a comparison map

$$\kappa_{\mathscr{E}} : \Pi_1^p(\mathscr{E}) \to \Pi_1^c(\mathscr{E})$$

and that this is a connected locally connected geometric morphism. For $\mathscr{E}$ a Galois topos, $\kappa_{\mathscr{E}}$ is an equivalence, in which case (that is defined on $\mathscr{S}$-Galois toposes) the assignment of $\Pi_1^c(\mathscr{E})$ to $\mathscr{E}$ is pseudofunctorial. But this case is not too interesting if what we are seeking is some sort of reflection of a *comprehensive* 2-category of toposes including the $\mathscr{S}$-Galois toposes into the 2-category of $\mathscr{S}$-Galois toposes, both as full sub 2-categories of $\mathbf{LTop}_{\mathscr{S}}$. Since the assignment of $\mathscr{E}_{lc}$ (the full subcategory of $\mathscr{E}$ consisting of its locally constant objets) to $\mathscr{E}$ is an idempotent operation, as is more generally the assignment of the full subcategory of $\mathscr{E}$ *generated* by the locally constant objects, the reflection would immediately result from the pseudofunctoriality and from it, in turn, a *van Kampen* theorem (fundamental group topos version) in the form of the preservation of pushouts could be directly inferred.

Let $\mathbf{GTop}_{\mathscr{S}}$ the full sub 2-category of $\mathbf{LTop}_{\mathscr{S}}$ with objects the $\mathscr{S}$-Galois toposes and let $I : \mathbf{GTop}_{\mathscr{S}} \to \mathbf{LTop}_{\mathscr{S}}$ be the inclusion 2-functor.

Remark 5.5 Denote by $\mathrm{lsc}(\mathbf{LTop}_{\mathscr{S}})$ (respectively $\mathrm{lsc}(\mathbf{GTop}_{\mathscr{S}})$) the full sub 2-category of $\mathbf{LTop}_{\mathscr{S}}$ (respectively of $\mathbf{GTop}_{\mathscr{S}}$) whose objects are locally simply connected toposes. The inclusion I above restricts to an inclusion $I : \mathrm{lsc}\mathbf{GTop}_{\mathscr{S}} \to \mathrm{lsc}\mathbf{LTop}_{\mathscr{S}}$. The only apparent way to render pseudofunctorial this assignment is (as done in [14]) to replace $\mathrm{lsc}(\mathbf{LTop}_{\mathscr{S}})$ by the non-full sub 2-category of $\mathbf{LTop}_{\mathscr{S}}$ whose objects are pairs $(\mathscr{E}\ , A)$ with $\mathscr{E}$ locally connected and A a *universal cover* in $\mathscr{E}$ (by which we mean a locally constant object A of $\mathscr{E}$ with global support which splits all of the locally constant objects in $\mathscr{E}$, as implicitly in Proposition 4.4), and where a morphism $(\mathscr{E}, A) \to (\mathscr{F}, B)$ is given by a pair (φ, s) consisting of a geometric morphism $\varphi : \mathscr{E} \to \mathscr{F}$ and a morphism $s : A \to \varphi^*(B)$ in $\mathscr{E}$. However, this results in a loss of the reflection property, so that a simple-minded van Kampen theorem is not a consequence. In this non-full setting, however, a van Kampen theorem can be derived by imposing certain restrictions on the pushouts, as in [14] Theorem 6.5; the same restrictions reappear naturally in the coverings version of a van Kampen theorem, as in [14] or, more generally as in the next section.

Notice now that if $\mathscr{E}$ is a locally simply connected (and locally connected) topos with universal cover A, then $\mathscr{E}/A$ has the property that $\Pi_1^c(\mathscr{E}/A) \simeq \mathscr{S}/e_!A$, in such a way that the equivalence is compatible with the given morphisms with domain $\mathscr{E}/A$. From the existence of the comparison map follows then that to require that $\Pi_1^p(\mathscr{E}/A) \simeq \mathscr{S}/e_!A$ is a stronger condition. We thus make the following definition (see [15] for a related notion in the case of Grothendieck toposes).

Definition 5.6 A locally connected topos $\mathscr{E}$ is said to be *locally paths simply connected* if $\mathscr{E}$ is locally simply connected and the universal cover A has the property that $\Pi_1^p(\mathscr{E}/A) \simeq \mathscr{S}/e_!A$ by an equivalence which commutes with $\tau_{\mathscr{E}/A} : \mathscr{E}/A \to \Pi_1^p(\mathscr{E}/A)$ and $\rho_{\eta_A} : \mathscr{E}/A \to \mathscr{S}/e_!A$.

Theorem 5.7 *For $\mathscr{E}$ a locally connected and locally paths simply connected topos, the comparison map $\kappa_{\mathscr{E}} : \Pi_1^p(\mathscr{E}) \to \Pi_1^c(\mathscr{E})$ is an equivalence.*

Proof It is enough to prove that there exists a unique geometric morphism $\lambda_{\mathscr{E}} : \Pi_1^c(\mathscr{E}) \to \Pi_1^p(\mathscr{E})$ for which the triangle

$$\begin{tikzcd} \mathscr{E} \arrow[r, "\tau_{\mathscr{E}}"] \arrow[d, "\sigma_{\mathscr{E}}"'] & \Pi_1^p(\mathscr{E}) \\ \Pi_1^c(\mathscr{E}) \arrow[ur, "\lambda_{\mathscr{E}}"'] & \end{tikzcd}$$

commutes and is such hat $\lambda_{\mathscr{E}}\kappa_{\mathscr{E}} = \mathrm{id}_{\Pi_1^p(\mathscr{E})}$.

To this end, consider the diagram

$$\begin{tikzcd} \mathscr{E}/A \arrow[r, "\varphi_A"] \arrow[d, "\rho_A"] \arrow[dd, bend right=60, "\tau_{\mathscr{E}/A}"'] & \mathscr{E} \arrow[d, "\sigma_{\mathscr{E}}"'] \arrow[dd, bend left=60, "\tau_{\mathscr{E}}"] \\ \mathscr{S}/e_!A \arrow[r, "p_A"] \arrow[d, "\simeq"] & \Pi_1^c(\mathscr{E}) \arrow[d, dotted, "\lambda_A"'] \\ \Pi_1^p(\mathscr{E}/A) \arrow[r, "\Pi_1^p(\varphi_A)"'] & \Pi_1^p(\mathscr{E}) \end{tikzcd}$$

where A is a universal cover of $\mathscr{E}$. Since A splits every locally constant object in $\mathscr{E}$, the top square is a pushout. The desired morphism $\lambda_{\mathscr{E}}$ arises from the pseudofunctoriality of Π_1^p and pseudonaturality of τ, using now that A is a universal cover and $\mathscr{E}$ is locally paths simply connected, so that $\Pi_1^p(\mathscr{E}/A) \simeq \mathscr{S}/e_!A$ by an equivalence which commutes with $\tau_{\mathscr{E}/A} : \mathscr{E}/A \to \Pi_1^p(\mathscr{E}/A)$. □

With obvious notations, we now have the following.

Proposition 5.8 *The inclusion I : lpsc$\mathbf{GTop}_{\mathscr{S}}$ $\to$ lpsc$\boldsymbol{LTop}_{\mathscr{S}}$ has a left biadjoint given by the 2-functor Π_1^c : lpsc$\boldsymbol{LTop}_{\mathscr{S}}$ $\to$ lpsc$\mathbf{GTop}_{\mathscr{S}}$. The unit of adjointness has as components the (connected locally connected) geometric morphisms $\sigma_{\mathscr{E}} : \mathscr{E} \to \Pi_1^c(\mathscr{E})$.*

Proof The pseudofunctoriality of the assigment of the topos $\Pi_1^c(\mathscr{E})$ to a locally connected and locally paths simply connected topos $\mathscr{E}$ in the sense of Definition 5.6 is easily deduced from Theorem 5.7 and the stability of locally constant objects under pullback along arbitrary geometric morphisms. That $\Pi_1^c(\mathscr{E})$ is a (locally paths simply connected) $\mathscr{S}$-Galois topos for each locally connected (and locally paths simply connected) topos $\mathscr{E}$ follows in turn (since $\Pi_1^c(\mathscr{E})$ is again locally connected) from the claim that the induced geometric morphism

$$\Pi_1^c(\sigma_{\mathscr{E}}) : \Pi_1^c(\mathscr{E}) \to \Pi_1^c(\Pi_1^c(\mathscr{E}))$$

is an equivalence. To see this it is sufficient to observe that the operation of carving out $\mathscr{E}_{lc}$ from $\mathscr{E}$ is an idempotent one. It is now obvious that Π_1^c is a reflection onto lpsc$\mathbf{GTop}_{\mathscr{S}}$ with σ the unit of the adjointness. This proves the first assertion; the second assertion is a consequence of the first and Corollary 4.11. □

It follows from Proposition 5.8 that pushouts in lpsc$\mathbf{LTop}_{\mathscr{S}}$ are taken by $\Pi_1^c(-)$ into pushouts in $\mathbf{Top}_{\mathscr{S}}$ in which all toposes in the pushout are locally paths simply connected Galois toposes, a statement which may be interpreted as a van Kampen theorem for (locally paths simply connected) toposes in terms of fundamental groupoid toposes [14]. We now derive a version of the van Kampen theorem in terms of fundamental groupoids.

Proposition 5.9 *Let*

$$\begin{tikzcd} \mathscr{E}_1 \arrow[r, "\alpha_1"] & \mathscr{E} \\ \mathscr{E}_0 \arrow[u, "\beta_1"] \arrow[r, "\beta_2"'] & \mathscr{E}_2 \arrow[u, "\alpha_2"'] \end{tikzcd}$$

be a pullback in $\mathbf{Top}_{\mathscr{S}}$ *in which all four geometric morphisms are inclusions, with the induced* $\alpha : \mathscr{E}_1 + \mathscr{E}_2 \to \mathscr{E}$ *a locally connected surjection, and where all four toposes are locally connected and locally paths simply connected. Then, the induced diagram*

$$\begin{tikzcd} \mathscr{B}(G_1) \arrow[r, "\alpha_1"] & \mathscr{B}(G) \\ \mathscr{B}(G_0) \arrow[u, "\beta_1"] \arrow[r, "\beta_2"'] & \mathscr{B}(G_2) \arrow[u, "\alpha_2"'] \end{tikzcd}$$

is a pushout of classifying toposes of discrete groupoids, where all four morphisms commute with the canonical localic points.

Proof The first remark to make is that, under the assumptions made (enough for this that α be of effective descent, for instance, an open surpjection), the given pullback is also a pushout ([14]). In particular, the pseudofunctor Π_1^p carries it into a pushout which, under the further assumption that all four toposes are locally simply connected, translates into the pullback of classifying toposes of discrete groupoids as in the statement of the theorem.

It remains to prove that the morphisms in that pushout commute with the base points. Notice first that all four morphisms in the original pullback diagram are locally connected. Indeed, since α is locally connected, also α_1 and α_2 (composites of α with locally connected coproduct injections) are locally connected and, in turn (by the pullback condition), also β_1 and β_2 are locally connected. Thus, is enough to prove that, if $\varphi : \mathscr{E} \to \mathscr{F}$ is locally connected, with $\mathscr{E}$ and $\mathscr{F}$ locally paths simply connected, then the induced $\Pi_1^c(\varphi) : \Pi_1^c(\mathscr{E}) \to \Pi_1^c(\mathscr{F})$, alternatively viewed as a geometric morphism $\mathscr{B}(\varphi) : \mathscr{B}(G(\mathscr{E})) \to \mathscr{B}(G(\mathscr{F}))$, is a +-geometric morphism of classifying toposes in the sense that it commutes with the canonical localic points.

To see this, observe that since φ is locally connected, it follows that if we let U be a cover in$\mathscr{E}$ splitting all locally constant objects in $\mathscr{E}$, then $W = \varphi_!(U)$ splits all locally constant objects in $\mathscr{F}$. However, it does not follow from U a cover in $\mathscr{E}$ that also W is a cover in $\mathscr{F}$. Without loss of generality, however, we may replace W by $V = \varphi_!(U) + K$, where K is any cover splitting all locally constant objects in $\mathscr{F}$; V is now both a cover and splits all locally constant objects in $\mathscr{F}$. Moreover, there is a morphism $e_!U \to f_!V$, determined by the first coproduct injection and the isomorphism $f_!(\varphi_!(U) + K) \cong e_!U + f_!K$.

Consider now the universal cover $\tilde{V}$ associated with V, taken together with the canonical morphisms $\delta : \tilde{V} \to f^*f_!V$ which represents (the inverse image part of) the canonical localic point p_V of $\mathscr{F}$.

We have a cube

$$\begin{tikzcd}[row sep=small, column sep=small]
 & \mathscr{F}/\tilde{V} \arrow[rr] \arrow[dd] & & \mathscr{F} \arrow[dd] \\
\mathscr{E}/U \arrow[ur] \arrow[rr] \arrow[dd] & & \mathscr{E} \arrow[ur] \arrow[dd] & \\
 & \mathscr{S}/f_!V \arrow[rr, dashed] & & \mathscr{G}_V(\mathscr{F}) \\
\mathscr{S}/e_!U \arrow[ur] \arrow[rr] & & \mathscr{G}_U(\mathscr{E}) \arrow[ur] &
\end{tikzcd}$$

where the front square is a pushout and the back square is (a pushout and) a pullback.

Since there is given a morphism $e_!U \to f_!V$ in $\mathscr{S}$, as pointed out above, there is first of all a unique geometric morphism $\mathscr{E}/U \to \mathscr{F}/\tilde{V}$ making the left side and the top squares commute, using for this that the back square is a pullback (universal cover $\tilde{V}$ in $\mathscr{F}$). Since the front square is a pushout, we now obtain a unique geometric morphism $\mathscr{G}_U(\mathscr{E}) \to \mathscr{G}_V(\mathscr{F})$ which makes the right side and the bottom squares commute. In particular, this is a morphism $\mathscr{B}(G_U(\mathscr{E})) \to \mathscr{B}(G_V(\mathscr{F}))$ which commutes with the canonical localic points p_U and p_V as claimed.

It remains to see that the above argument can applied consistently to the four geometric morphisms in the pushout square, but this can be done just as in [14] (section 6). □

Denote by **Gpd**($\mathscr{S}$) the 2-category of (discrete localic) groupoids in $\mathscr{S}$.

Corollary 5.10 *Let*

$$\begin{tikzcd}
\mathscr{E}_1 \arrow[r, "\alpha_1"] & \mathscr{E} \\
\mathscr{E}_0 \arrow[u, "\beta_1"] \arrow[r, "\beta_2"'] & \mathscr{E}_2 \arrow[u, "\alpha_2"']
\end{tikzcd}$$

be a diagram in $\mathbf{Top}_{\mathscr{S}}$ *satisfying the conditions of Proposition 5.9. Then the diagram*

$$\begin{tikzcd}
\pi_1^c(\mathscr{E}_1) \arrow[r, "\pi_1^c(\alpha_1)"] & \pi_1^c(\mathscr{E}) \\
\pi_1^c(\mathscr{E}_0) \arrow[u, "\Pi_1^c(\beta_1)"] \arrow[r, "\pi_1^c(\beta_2)"'] & \pi_1^c(\mathscr{E}_2) \arrow[u, "\pi_1^c(\alpha_2)"']
\end{tikzcd}$$

is a pushout in **Gpd**.

Proof This is an immediate consequence of Proposition 5.9 and the Morita theorem for localic groupoids from [8]. □

6 Generalized covering morphisms and a van Kampen theorem

In topology, covering morphisms (in a more general sense than that of covering projections) are usually required to be local homeomorphisms which furthermore satisfy a path-lifting property. The latter should be thought of as topological coverings. Since morphisms not satisfying either condition, such as certain complete spreads in topos theory [11, 12], have also been thought of as covering morphisms "with singularities", we want to also make precise in what (generalized) sense are these still to be considered as coverings.

We shall restrict, as we have done so far in this paper, to locally connected toposes over $\mathscr{S}$, even if this assumption is not always strictly required. Denote by $\mathbf{LTop}_{\mathscr{S}}$ the full sub 2-category of $\mathbf{Top}_{\mathscr{S}}$ whose objects are the locally connected toposes.

A *notion* $\mathscr{C}$ *of covering morphism* on $\mathbf{LTop}_{\mathscr{S}}$ will be given by an assignment, for each $\mathscr{E} \in \mathbf{LTop}_{\mathscr{S}}$, of a category $\mathscr{C}(\mathscr{E}) \subseteq \mathbf{LTop}_{\mathscr{S}}/\mathscr{E}$, where the slice category $\mathbf{LTop}_{\mathscr{S}}/\mathscr{E}$ has as objects, geometric morphisms with codomain $\mathscr{E}$, with a morphism from $\varphi : \mathscr{F} \to \mathscr{E}$ to $\psi : \mathscr{G} \to \mathscr{E}$ given by a geometric morphism $\chi : \mathscr{F} \to \mathscr{G}$ for which there is a natural isomorphism $\alpha : \varphi \to \psi\chi$.

Denote by $\mathscr{L} : \mathscr{K}^{\mathrm{op}} \to \mathbf{CAT}$ the pseudofunctor which assigns to a topos $\mathscr{E}$ the slice category $\mathbf{LTop}_{\mathscr{S}}/\mathscr{E}$ and to a geometric morphism $\varphi : \mathscr{F} \to \mathscr{E}$ the functor given by pulling back along φ.

We will say (following [16]) that a subpseudofunctor $\mathscr{C}$ of $\mathscr{L}: \mathscr{K}^{\mathrm{op}} \to \mathbf{CAT}$ is a *stack* for a class Φ of morphisms of effective descent in $\mathbf{Top}_{\mathscr{S}}$, assumed to be closed under composition and pullbacks, if for any object $\mathscr{E}$ of $\mathscr{K}$, given $\alpha \in \mathscr{L}(\mathscr{E})$ such that for some $\varphi : \mathscr{F} \to \mathscr{E}$ with $\varphi \in \Phi$, it is the case that $\varphi^*(\alpha) \in \mathscr{C}(\mathscr{F})$, then it follows that $\alpha \in \mathscr{C}(\mathscr{E})$.

Definition 6.1 Let $\mathscr{K}$ be a sub 2-category of $\mathbf{LTop}_{\mathscr{S}}$. A notion $\mathscr{C}$ of covering morphisms is said to define *a fibration of covering morphisms for* $\mathscr{K}$ with respect to a class Φ as above if,

1. $\mathscr{C}$ extends to a pseudofunctor $\mathscr{C} : \mathscr{K}^{\mathrm{op}} \to \mathbf{CAT}$,
2. there is a fully faithful pseudonatural transformation $H : \mathscr{C} \to \mathscr{L}$, and
3. $\mathscr{C}$ is a stack for the class Φ.

We say that a pseudofunctor $\mathscr{C} : \mathscr{K}^{\mathrm{op}} \to \mathbf{CAT}$ *preserves binary products* if for each pair of objects E and F in $\mathscr{K}$, the canonical functors

$$\mathscr{C}(E+F) \to \mathscr{C}(E) \times \mathscr{C}(F)$$

induced by the injections $i : E \to E+F$ and $j : F \to E+F$ is an equivalence. In the terminology of [30], $\mathscr{C}$ is an *intensive quantity* on $\mathscr{K}$.

The assignment to a topos $\mathscr{E}$ of the class of *local homeomorphisms* over it constitutes a pseudofunctor $\mathscr{A}: \mathbf{LTop}_{\mathscr{S}}^{\mathrm{op}} \to \mathbf{CAT}$; given any geometric morphism $\varphi : \mathscr{F} \to \mathscr{E}$ in $\mathbf{LTop}_{\mathscr{S}}$, pulling back along φ restricts to a functor $A(\varphi) : \mathscr{A}(\mathscr{E}) \to \mathscr{A}(\mathscr{F})$. The pseudofunctor $\mathscr{A}$ is trivially a stack for any (morphism of effective descent) class Φ. That $\mathscr{A}$ preserves binary products follows from the fact that colimits in $\mathbf{Top}_{\mathscr{S}}$ (or $\mathbf{LTop}_{\mathscr{S}}$) are calculated as limits in $\mathbf{CAT}$ of the corresponding diagrams of inverse image parts [33].

Notice that the 2-extensivity of $\mathscr{K}$ (for $\mathscr{K} = \mathbf{Top}_{\mathscr{S}}$, or $\mathscr{K} = \mathbf{LTop}_{\mathscr{S}}$) says that $\mathscr{L}$ preserves binary products. Notice also that there is a fully faithful pseudonatural transformation $J : \mathscr{A} \to \mathscr{L}$. Indeed, given any composite $\zeta = \psi\varphi$, where both ζ and ψ are local homeomorphisms, so is φ.

The following is a version of the *van Kampen theorem* in terms of covering morphisms. It is an abstraction of Theorem 5.7 of [14], given therein for covering projections on toposes; the proof is analogous.

Theorem 6.2 *Let $\mathscr{C}$ be a fibration of covering morphisms on $\mathscr{K}$ with respect to a class Φ of morphisms of effective descent in $\mathscr{K}$. Assume furthermore that $\mathscr{C}$ preserves binary products and that $H : \mathscr{C} \to \mathscr{L}$ factors through the canonical* *inclusion* $J : \mathscr{A} \to \mathscr{L}$. *Let*

$$\begin{tikzcd} \mathscr{E}_1 \arrow[r, "\alpha_1"] & \mathscr{E} \\ \mathscr{E}_0 \arrow[u, "\beta_1"] \arrow[r, "\beta_2"'] & \mathscr{E}_2 \arrow[u, "\alpha_2"'] \end{tikzcd}$$

be a pushout diagram in $\mathbf{LTop}_{\mathscr{S}}$ *in which the induced map* $\alpha : \mathscr{E}_1 + \mathscr{E}_2 \to \mathscr{E}$ *is a morphism in* Φ. *Then*

$$\begin{tikzcd} \mathscr{C}(\mathscr{E}_1) \arrow[r, "\beta_1^*"] & \mathscr{C}(\mathscr{E}_0) \\ \mathscr{C}(\mathscr{E}) \arrow[u, "\alpha_1^*"] \arrow[r, "\alpha_2^*"'] & \mathscr{C}(\mathscr{E}_2) \arrow[u, "\beta_2^*"'] \end{tikzcd}$$

is a pullback in **CAT**.

Another example of a fibration of covering morphisms is given by letting $\mathscr{C}(\mathscr{E}) = \mathscr{E}_{lc}$ be the full subcategory of $\mathscr{E}$ determined by the (locally trivial) covering morphisms in terms of the adjunction $e_! \dashv e^*$ in the sense of [24]; equivalently $\mathscr{C}(\mathscr{E})$ is the category of local homeomorphisms over $\mathscr{E}$ defined by a locally constant object in the sense of Definition 2.1. The pseudofunctoriality of $\mathscr{C}$ on $\mathbf{LTop}_{\mathscr{S}}$ follows from the stability of locally constant objects under pullbacks along arbitrary geometric morphisms, as remarked in Section 2. There is a pseudonatural transformation $H : \mathscr{C} \to \mathscr{A}$ (with fully faithful components) by the very definition of $\mathscr{C}$. It is shown in [14] that H is φ-cartesian for every surjective locally connected φ. This means that the pseudonaturality square

$$\begin{tikzcd} \mathscr{A}(X) \arrow[r, "\mathscr{A}(\varphi)"] & \mathscr{A}(Y) \\ \mathscr{C}(X) \arrow[u, "HX"] \arrow[r, "\mathscr{C}(\varphi)"'] & \mathscr{C}(Y) \arrow[u, "HY"'] \end{tikzcd}$$

is a pullback in $\mathscr{K}$. This readily implies that $\mathscr{C}$ is a stack for the class Φ of locally connected surjections. It is also shown in [14] that $\mathscr{C}$ preserves binary products.

In particular, there is a *van Kampen theorem* for $\mathscr{C}$ as an application of Theorem 6.2, already shown in [14].

We shall now consider a different fibration of covering morphisms for which we need to recall some basic facts about complete spreads from [11, 12]. Let $\mathscr{E}$ be a (locally connected) topos presented over $\mathscr{S}$ by a site $\mathbf{C}$. A geometric morphism $\varphi : \mathscr{F} \to \mathscr{E}$ in $\mathbf{LTop}_{\mathscr{S}}$ is a *complete spread* if, for the discrete opfibration $D : \mathbf{D} \to \mathbf{C}$ which is associated with the functor $D : \mathbf{C} \to \mathscr{S}$ which is the composite

$$\mathbf{C} \xrightarrow{\varepsilon} \mathscr{E} \xrightarrow{\varphi^*} \mathscr{F} \xrightarrow{f_!} \mathscr{S},$$

the square

$$\begin{tikzcd} \mathscr{F} \arrow[r] \arrow[d, "\varphi"] & \mathscr{S}^{\mathbf{D}^{\mathrm{op}}} \arrow[d, "u"] \\ \mathscr{E} \arrow[r] & \mathscr{S}^{\mathbf{C}^{\mathrm{op}}} \end{tikzcd}$$

where u is induced by D, is a pullback.

Complete spreads are related to distributions in the sense of Lawvere [30]. A distribution on a (locally connected) topos $\mathscr{E}$ over $\mathscr{S}$ is any $\mathscr{S}$-cocontinuous functor $\mu : \mathscr{E} \to \mathscr{S}$. Denote by $\mathbf{Dist}(\mathscr{E})$ the category of distributions on $\mathscr{E}$. Explicitly, $\mathbf{Dist}(\mathscr{E}) = \mathrm{Coc}_{\mathscr{S}}(\mathscr{A}(\mathscr{E})\ ,\ \mathscr{S})$ and each geometric morphism $\varphi : \mathscr{F} \to \mathscr{E}$ induces, by composition with the $\mathscr{S}$-cocontinuous functor $\mathscr{A}(\varphi) = \varphi^* : \mathscr{E} \to \mathscr{F}$, a functor $\mathbf{Dist}(\varphi) : \mathbf{Dist}(\mathscr{F}) \to \mathbf{Dist}(\mathscr{E})$.

For each $\mathscr{E} \in \mathbf{Top}_{\mathscr{S}}$ there is a topos $\mathscr{M}(\mathscr{E})$ [10] (the "symmetric topos") and an $\mathscr{S}$-essential geometric morphism $\delta : \mathscr{E} \to \mathscr{M}(\mathscr{E})$ such that the pair $\langle \mathscr{M}(\mathscr{E})\ ,\ \delta\rangle$ classifies distributions on $\mathscr{E}$.

If $\varphi : \mathscr{F} \to \mathscr{E}$ is $\mathscr{S}$-essential, then composition with the $\mathscr{S}$-cocontinuous functor $\varphi_! : \mathscr{F} \to \mathscr{E}$ is a functor $\mathscr{R}_\varphi : \mathbf{Dist}(\mathscr{E}) \to \mathbf{Dist}(\mathscr{F})$ which is right adjoint to the functor $\mathbf{Dist}(\varphi) : \mathbf{Dist}(\mathscr{F}) \to \mathbf{Dist}(\mathscr{E})$, as the latter is given by composition with φ^*. Under the equivalence between distributions and complete spreads, it is the right adjoint $\mathscr{R}_\varphi$ which corresponds to pulling back complete spreads over $\mathscr{E}$ along the $\mathscr{S}$-essential morphism $\varphi : \mathscr{F} \to \mathscr{E}$ to give a complete spread over $\mathscr{F}$.

For objects $\mathscr{E}$ and $\mathscr{F}$ in $\mathbf{Top}_{\mathscr{S}}$, their coproduct $\mathscr{E} + \mathscr{F}$ exists in $\mathbf{Top}_{\mathscr{S}}$ and is given by the product category $\mathscr{E} \times \mathscr{F}$; the injections $\mathrm{i} : \mathscr{E} \to \mathscr{E} + \mathscr{F}$ and $\mathrm{j} : \mathscr{F} \to \mathscr{E} + \mathscr{F}$ have as inverse image parts the left adjoints to the projections. As geometric morphisms over $\mathscr{S}$, the injections into the coproduct are locally connected, since the projections have an $\mathscr{S}$-indexed left adjoint as well as a right adjoint. It is shown in [14] that $\mathbf{Top}_{\mathscr{S}}$ is an extensive 2-category; this says that for the 2-categorical versions of the slices, the coproduct pseudofunctor $F : \mathbf{Top}_{\mathscr{S}} \times \mathbf{Top}_{\mathscr{S}} \to \mathbf{Top}_{\mathscr{S}}$ induces biequivalences

$$F_{\langle \mathscr{E}, \mathscr{F}\rangle} : \mathbf{Top}_{\mathscr{S}}/\mathscr{E} \times \mathbf{Top}_{\mathscr{S}}/\mathscr{F} \to \mathbf{Top}_{\mathscr{S}}/(\mathscr{E} + \mathscr{F})$$

for all $\mathscr{E}$ and $\mathscr{F}$ in $\mathbf{Top}_{\mathscr{S}}$. Similar statements hold for $\mathbf{LTop}_{\mathscr{S}}$ [14].

Denote by $\mathscr{D}(\mathscr{E})$ the category of complete spreads (with locally connected domain) over $\mathscr{E}$. There is an functor $B : \mathscr{D}(\mathscr{E}) \to \mathbf{Dist}(\mathscr{E})$ which assigns, to a complete spread (in fact to any geometric morphism) $\varphi : \mathscr{F} \to \mathscr{E}$ in $\mathbf{LTop}_{\mathscr{S}}$ the distribution $\mu : \mathscr{E} \to \mathscr{S}$ given by the composite $\mu = f_! \varphi^*$. This functor is an equivalence. This gives rise to a (unique) factorization of any given geometric morphism $\varphi : \mathscr{F} \to \mathscr{E}$ with locally connected domain into a composite of a *pure* morphism $\pi : \mathscr{F} \to \mathscr{D}$ followed by a complete spread $\psi : \mathscr{D} \to \mathscr{E}$ with locally connected domain.

All of the above considerations remain true if $\mathbf{Top}_{\mathscr{S}}$ is replaced by $\mathbf{LTop}_{\mathscr{S}}$ [14].

Proposition 6.3 *Let $\mathscr{K}$ be $\mathbf{LTop}_{\mathscr{S}}$. The assignment of $\mathscr{D}(\mathscr{E})$ to $\mathscr{E}$ preserves binary products, in the sense that for any locally connnected toposes $\mathscr{E}$ and $\mathscr{F}$, there is a (canonical) equivalence $\mathscr{D}(\mathscr{E} + \mathscr{F}) \simeq \mathscr{D}(E) \times \mathscr{D}(F)$ of categories.*

Proof Consider the functor

$$F : \mathscr{D}(\mathscr{E} + \mathscr{F}) \to \mathscr{D}(\mathscr{E}) \times \mathscr{D}(\mathscr{F})$$

induced by pulling back along the coproduct injections. Since the latter are $\mathscr{S}$-essential geometric morphisms (in fact, locally connected), this is well defined [13].

Consider now the coproduct pullback diagram

$$\begin{tikzcd} \mathscr{A} \arrow[r, "u"] \arrow[d, "\psi"'] & \mathscr{A}+\mathscr{B} \arrow[d, "\psi+\zeta"] & \mathscr{B} \arrow[l, "v"'] \arrow[d, "\zeta"] \\ \mathscr{E} \arrow[r, "i"'] & \mathscr{E}+\mathscr{F} & \mathscr{F} \arrow[l, "j"] \end{tikzcd}$$

with ψ and ζ both complete spreads with locally connected domain.

We claim that the assignment of $\psi + \zeta$ to $\langle \psi, \zeta \rangle$ defines a functor

$$G : \mathscr{D}(\mathscr{E}) \times \mathscr{D}(\mathscr{F}) \to \mathscr{D}(\mathscr{E} + \mathscr{F})$$

To see that G is well defined we must prove that $\psi + \zeta$ is a complete spread (with locally connected domain). Consider the factorization $\psi + \zeta \simeq \varphi\rho$, with ρ pure and φ a complete spread with locally connected domain $\mathscr{C}$. By extensivity of $\mathscr{K}$ it follows first that $\mathscr{C} \simeq \mathscr{C}_0 + \mathscr{C}_1$ and $\varphi \simeq \varphi_0 + \varphi_1$, and second, that $\rho = \rho_0 + \rho_1$ From the stability of the pure/complete spread factorization under pullbacks along locally connected geometric morphisms, and the fact that ψ and ζ are complete spreads, it follows that ρ_0 and ρ_1 are (pure) equivalences and hence that $\varphi \simeq \psi + \zeta$, so that $\psi + \zeta$ is a complete spread as claimed.

The pair $\langle F, G \rangle$ consitutes an equivalence by the extensivity of $\mathbf{LTop}_{\mathscr{S}}$.

□

Let $\mathscr{U}(\mathscr{E})$ be the full subcategory of $\mathscr{E}$ determined by its complete spread objects [12, 19], alternatively viewed as the "unramified morphisms" over $\mathscr{E}$, consisting of those complete spreads over $\mathscr{E}$ which are also local homeomorphisms. Denote by $\mathscr{U}$ the intersection of $\mathscr{A}$ and $\mathscr{D}$.

Proposition 6.4 *$\mathscr{U}$ is a fibration of covering morphisms on $\boldsymbol{LTop}_{\mathscr{S}}$ for the class Φ of surjective local homeomorphisms. Furthermore, $\mathscr{U}$ preserves binary products.*

Proof The stability of complete spread objects (or of the unramified morphisms they determine) under pullback in $\mathbf{LTop}_{\mathscr{S}}$ has been shown in [19]. Hence, $\mathscr{U} : \mathbf{LTop}_{\mathscr{S}}^{\mathrm{op}} \to \mathbf{CAT}$ may be regarded as a sub pseudofunctor of $\mathscr{A}$ (not just of $\mathscr{L}$) . The remaining properties are a consequence Proposition 6.3. □

Corollary 6.5 *Let*

$$\begin{tikzcd} \mathscr{E}_1 \arrow[r, "\alpha_1"] & \mathscr{E} \\ \mathscr{E}_0 \arrow[u, "\beta_1"] \arrow[r, "\beta_2"'] & \mathscr{E}_2 \arrow[u, "\alpha_2"'] \end{tikzcd}$$

be a pushout diagram (or a pullpack diagram with α_1 and α_2 are pseudomonic) in $\boldsymbol{LTop}_{\mathscr{S}}$ in which the induced $\alpha : \mathscr{E}_1 + \mathscr{E}_2 \to \mathscr{E}$ is a surjective local homeomorphism. Then

$$\begin{tikzcd} \mathscr{U}(\mathscr{E}_1) \arrow[r, "\beta_1^*"] & \mathscr{U}(\mathscr{E}_0) \\ \mathscr{U}(\mathscr{E}) \arrow[u, "\alpha_1^*"] \arrow[r, "\alpha_2^*"'] & \mathscr{U}(\mathscr{E}_2) \arrow[u, "\beta_2^*"'] \end{tikzcd}$$

is a pullback in $\mathbf{CAT}$.

Proof It follows directly from Theorem 6.2 and Proposition 6.4. □

A topological notion of covering morphism on $\mathbf{LTop}_{\mathscr{S}}$ ought to restrict at least to all local homeomorphisms satisfying a path-lifting property.

Definition 6.6 A fibration $\mathscr{C}$ of covering morphisms in the sense of Definition 6.1 is said to be *topological* if:

1. $\mathscr{C}$ is a subpseudofunctor of $\mathscr{A} : \mathbf{LTop}_{\mathscr{S}}^{\mathrm{op}} \to \mathbf{CAT}$, and
2. for every (locally connected) topos $\mathscr{E}$, every cover $X \in \mathscr{C}(\mathscr{E})$ has the unique path-lifting property in the sense of Definition 5.1.

Let us examine how the three examples of fibrations of covering morphisms (given so far) fare from the topological viewpoint expressed in Definition 6.6. The *second condition* in the above definition excludes $\mathscr{A}$ as a topological fibration of covering morphisms since not every cover X in an arbitrary (locally connected) topos $\mathscr{E}$ has the unique path-lifting property, or not every local homeomorphism is a (discrete) fibration. Thus, arbitrary local homeomorphisms, though covering morphisms, are not in general topological covering morphisms. The *first condition* in the above definition excludes also $\mathscr{D}$ as a topological fibration of covering morphisms since $\mathscr{D}$ is not a pseudofuctor on $\mathbf{LTop}_{\mathscr{S}}^{\mathrm{op}}$ but also it is not the case that for every $\mathscr{E}$, $\mathscr{D}(\mathscr{E})$ is included in $\mathscr{A}(\mathscr{E})$, as not all complete spreads are local homeomorphisms. By contrast, the fibration $\mathscr{C}$ of locally trivial covering morphisms satisfies the required conditions.

Proposition 6.7 *Let $\mathscr{C}(\mathscr{E}) = \mathscr{E}_{lc}$ be the full subcategory of $\mathscr{E}$ determined by the locally constant objects in the sense of Definition 2.1. Then $\mathscr{C}$ defines a topological fibration of covering morphisms in the sense of Definition 6.6.*

Proof That $\mathscr{C}$ is a subpseudofunctor of $\mathscr{A}$ is part of the definition of $\mathscr{C}$. That every object of every $\mathscr{C}(\mathscr{E})$ satisfies the unique path-lifting property was established in Proposition 5.3. □

It is natural to ask whether the topological notion expressed in Definition 6.6 includes examples other than the locally trivial covering morphisms.

The following result was shown in [12] for the locally constant objects in the sense of Barr and Diaconescu [3]; we need to establish it here for the locally constant objects in the sense of Definition 2.1.

Proposition 6.8 *Let Y be a locally constant object in a locally connected topos $\mathscr{E}$. Then, Y is a complete spread object in $\mathscr{E}$.*

Proof Let Y be a locally constant object in $\mathscr{E}$ in the sense of Definition 2.1. By Proposition 2.2 an equivalent condition (expressed therein as condition (3)) is that there is a cover U in $\mathscr{E}$, a morphism $\alpha : J \to I$ in $\mathscr{S}$, a morphism $\eta : U \to e^*I$, and a morphism $\zeta : Y \times U \to e^*J$, for which the square

$$\begin{tikzcd} Y \times U \arrow[r, "\pi_2"] \arrow[d, "\zeta"] & U \arrow[d, "\eta"] \\ e^*J \arrow[r, "e^*\alpha"'] & e^*I \end{tikzcd}$$

is a pullback. By the above and the local character of complete spreads along surjective local homeomorphisms [12](Proposition 7.1), it is sufficient to prove that any geometric morphism $\mathscr{E}/e^*J \to \mathscr{E}/e^*I$ (induced by a morphism $\alpha : J \to I$ in $\mathscr{S}$) is a complete spread. We know [12](proof of Theorem 7.3) that constant objects are complete spreads. Therefore the composite $\mathscr{E}/e^*J \to \mathscr{E}/e^*I \to \mathscr{E}$ as well as the second morphisms in the composite are both complete spreads. By [11] (Proposition 1.2.3) this implies that $\mathscr{E}/e^*J \to \mathscr{E}/e^*I$ is a complete spread. □

Theorem 6.9 *Let X be a complete spread cover (i.e., a complete spread object with full supoort) in a locally connected topos $\mathscr{E}$. Then X has the unique path-lifting property in the sense of Definition 5.1.*

Proof The following is a characterization of complete spread covers [12] Proposition 7.5. Let $\mathscr{E}$ be a locally connected topos with pure inclusion $\mathscr{E} \hookrightarrow \mathscr{S}^{\mathbf{C}^{\mathrm{op}}}$. Assume that X is a cover in $\mathscr{E}$ with corresponding (surjective) discrete opfibration $X : \mathbf{X} \to \mathbf{C}$ and associated discrete fibration $D : \mathbf{D} \to \mathbf{C}$. Then X is a complete spread object if and only if the square

$$\begin{tikzcd}
\mathscr{E}/X \arrow[r] \arrow[d] & \mathscr{S}^{\mathbf{D}^{\mathrm{op}}} \arrow[d] \\
\mathscr{E} \arrow[r] & \mathscr{S}^{\mathbf{C}^{\mathrm{op}}}
\end{tikzcd}$$

is a pullback, where the right vertical arrow is induced by D and therefore is a surjective discrete fibration (and opfibration); the geometric morphism that it induces has the unique paths-lifting property.

We now show that the local homeomorphism $\mathscr{E}/X \to \mathscr{E}$ has the unique path-lifting property. Consider the cube:

$$\begin{tikzcd}
& \mathscr{D} \arrow[rr] \arrow[dd] & & \mathscr{S}^{\mathbf{D}^{\mathrm{op}}} \arrow[dd] \\
\mathscr{D}^{\Delta_n} \arrow[rr] \arrow[ur] \arrow[dd] & & \mathscr{S}^{\mathbf{D}^{\mathrm{op}}\Delta_n} \arrow[ur] \arrow[dd] & \\
& \mathscr{E} \arrow[rr] & & \mathscr{S}^{\mathbf{C}^{\mathrm{op}}} \\
\mathscr{E}^{\Delta_n} \arrow[rr] \arrow[ur] & & \mathscr{S}^{\mathbf{C}^{\mathrm{op}}\Delta_n} \arrow[ur] &
\end{tikzcd}$$

In it, the front, back, and right side faces are pullbacks, hence so is the left side square. This gives the unique paths-lifting property for the geometric morphism $\mathscr{D} \to \mathscr{E}$, as claimed. □

Proposition 6.10 *$\mathscr{U}$ is a topological fibration of covering morphisms.*

Proof By Proposition 6.4 $\mathscr{U}$ is a fibration of covering morphisms. By definition, $\mathscr{U}$ is a subfibration of $\mathscr{A}$. The path-lifting property is shown in Theorem 6.9. □

We note that, as shown explicitly in [19], the inclusion $\mathscr{C} \hookrightarrow \mathscr{U}$ is proper. Hence this is a genuine example of a notion of covering morphism having analogous topological properties as that of the locally trivial covering morphisms, yet it is not (seemingly) an instance of the Janelidze notion. Although $\mathscr{U}$ is a stack and contains all trivial morphisms, it is not the stack completion of it; indeed the stack completion of the trivial morphisms is $\mathscr{C}$. Although not discussed in this paper, it can be shown, just as in topology, that in the case of locally paths simply connected toposes, the two classes $\mathscr{C}$ and $\mathscr{U}$ agree.

Acknowledgements. I am very grateful to George Janelidze for directing my attention to an interesting connection between his work and mine, and to Eduardo Dubuc for valuable comments on Galois toposes and on my construction of the fundamental groupoid of a topos. Thanks are due to Steve Vickers, Steve Lack and George Janelidze for useful concrete remarks on a preliminary version of this paper, and to Bill Lawvere, Ronnie Brown and Claudio Hermida for motivating general discussions.

References

[1] M. Artin, A. Grothendieck, and J.-L. Verdier, *Théorie des Topos et Cohomologie Etale des Schemas (SGA4), Lecture Notes in Mathematics 269, 270, 305*, Springer-Verlag, Berlin-Heidelberg-New York, 1972.

[2] M. Barr and R. Diaconescu, Atomic Toposes, *J. Pure Appl. Alg.* 17: 1–24, 1980.

[3] M. Barr and R. Diaconescu, On locally simply connected toposes and their fundamental groups, *Cahiers Top. Geo. Diff. Cat.* 22:301–314, 1981.

[4] M. Barr and R. Paré, Molecular Toposes, *J. Pure Appl. alg.* 17: 265–280, 1980.

[5] F. Borceux and G. Janelidze, *Galois theories*, Cambridge Studies in Advanced Mathematics 72, Cambridge Univ. Press, 2001.

[6] R. Brown, Groupoids and van Kampen theorems, *Proc. London Math. Soc.* 3-17:385–401, 1967.

[7] M. Bunge, Internal Presheaf Toposes, *Cahiers de Top. et Géom. Diff.* 18-3:291–330, 1977.

[8] M. Bunge, An application of descent to a classification theorem for toposes, *Math. Proc. Camb. Phil. Soc.* 107:59–79, 1990.

[9] M. Bunge, Classifying toposes and fundamental localic groupoids, in R.A.G.Seely, ed., *Category Theory '91, CMS Conference Proceedings* 13:73–96, 1992.

[10] M. Bunge and A. Carboni, The symmetric topos, *J. Pure Appl. Alg.* **105** (1995) 233-249.

[11] M. Bunge and J. Funk, Spreads and the symmetric topos, *J. Pure Appl. Alg.* 113:1–38, 1996.

[12] M. Bunge and J. Funk, Spreads and the symmetric topos II, *J. Pure Appl. Alg.* 130:49–84, 1998.

[13] M. Bunge and J. Funk, On a bicomma object condition for KZ-doctrines, *J. Pure Appl. Alg.* **143** (1999) 69–105.

[14] M. Bunge and S. Lack, Van Kampen theorems for toposes, *Advances in Mathematics* 179-2:291–317, 2003.

[15] M. Bunge and I. Moerdijk, On the construction of the Grothendieck fundamental group of a topos by paths, *J. Pure Appl. Alg.* 116:99–113, 1997.

[16] M. Bunge and R. Paré, Stacks and equivalences of categories, *Cah. de Top. et Géom. Diff.* 20-4:373–399, 1979.

[17] M. Bunge, Stack completions and Morita equivalence for categories in a topos, *Cah. de Top. et Géom. Diff.* 20-4:401–436, 1979.

[18] J. Duskin, An outline of non-abelian cohomology in a topos: (1) The theory of bouquets and gerbes, *Cahiers de Top. et Géo. Diff. Catégoriques* 23:165–191, 1982.

[19] J. Funk and E .D. Tymchatyn, Unramified maps, *JP Jour. Geom. Topol.* 1:249–280, 2001.

[20] A. Grothendieck, *Revetements étales et groupe fondamental (SGA1), Lecture Notes in Mathematics* 224, Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[21] A. Grothendieck, *La longue marche a travers la théorie de Galois, Tome I*, Transcription d'un manuscrit inédit, par J. Malgoire, Université Montpellier II, 7–48, 1995.

[22] A. Grothendieck, Esquisse d'un programme, in: L. Scnepps and P. Lochak, *Geometric Galois Actions. 1. Around Grothendieck's Esquisse d'un Programme London Mathematical Society Lecture Notes Series 242*, Cambridge University Press, 1997.

[23] G. Janelidze, Precategories and Galois theory, in: A. Carboni et al., *Category Theory. proceedings Como 1990* Lecture Notes in Mathematics 1488, 157–173, 1991.

[24] G. Janelidze, Pure Galois theory in categories, *J. Alg.* 132:270–286, 1990.

[25] G. Janelidze, A note on Barr-Diaconescu covering theory, *Contemp. Math* 131:121–124, 1992.

[26] G. Janelidze, D. Schumacher, and R. Street, Galois theory in variable categories, *Applied Categorical Structures* 1:103–110, 1993.

[27] P. T. Johnstone, *Topos Theory*, Academic Press, 1977.
[28] A. Joyal and M. Tierney, An extension of the Galois theory of Grothendieck, *Mem. Amer. Math. Soc.* 309, 1984.
[29] J. Kennison, The fundamental localic groupoid of a topos, *J. Pure Appl. Algebra* 77:67–86, 1992.
[30] F. W. Lawvere, *Extensive and intensive quantities*, Lectures given at the Workshop on Categorical Methods in Geometry, Aarhus, 1983.
[31] I. Moerdijk, Prodiscrete groups and Galois toposes, *Proc. Koninklijke Ned. Akad. van Wetenschappen* 51:219–234, 1989.
[32] I. Moerdijk, Continuous fibrations and inverse limits of toposes, *Compositio Math.* 58:45–72, 1986.
[33] I. Moerdijk, The classifying topos of a continuous groupoid I, *Trans. Amer. Math. Soc.* 310:629-668, 1988.
[34] I. Moerdijk and G. C. Wraith, Connected locally connected toposes are path-connected, *Trans. Amer. Math. Soc.* 295:849–859, 1986.

Received August 28, 2002; in revised form October 21,2003

Fields Institute Communications
Volume **43**, 2004

Galois Corings From the Descent Theory Point of View

S. Caenepeel
Faculty of Applied Sciences
Vrije Universiteit Brussel, VUB
Pleinlaan 2
B-1050 Brussels, Belgium
scaenepe@vub.ac.be

Abstract. We introduce Galois corings, and give a survey of properties that have been obtained so far. The Definition is motivated using descent theory, and we show that classical Galois theory, Hopf-Galois theory and coalgebra Galois theory can be obtained as a special case.

Introduction

Galois descent theory [36] has many applications in several branches of mathematics, such as number theory, commutative algebra and algebraic geometry; to name such one example, it is an essential tool in computing the Brauer group of a field. In the literature, several generalizations have appeared. Galois theory of commutative rings has been studied by Auslander and Goldman [3] and by Chase, Harrison and Rosenberg [16], see also [21]. The group action can be replaced by a Hopf algebra (co)action, leading to Hopf-Galois theory, see [17] (in the case where the Hopf algebra is finitely generated projective), and [24], [35] in the general case. More recently, coalgebra Galois extensions were introduced by Brzeziński and Hajac [10]. It became clear recently that a nice unification of all these theories can be formulated using the language of corings. Let us briefly sketch the history.

During the nineties, several unifications of the various kinds of Hopf modules that had appeared in the literature have been proposed. Doi [22] and Koppinen [30] introduced Doi-Hopf modules. A more general concept, entwined modules, was proposed by Brzeziński and Majid [11]. Böhm introduced Doi-Hopf modules over a weak bialgebra ([6]), and the author and De Groot proposed weak entwined modules [12]. Takeuchi [39] observed that all types of modules can be viewed as comodules over a coring, a concept that was already introduced by Sweedler [38], but then more or less forgotten, at least by Hopf algebra theorists; the idea was further investigated by Brzeziński [9]. He generalized several properties that had been studied in special cases to the situation where one works over a general coring, such

2000 *Mathematics Subject Classification.* Primary 16W30; Secondary 13B05.
Key words and phrases. Coring, Galois extension, descent theory.

as separability and Frobenius type properties, and it turned out that computations sometimes become amazingly simple if one uses the language of corings, indicating that this is really the right way to look at the problem. Brzeziński also introduces the notion of Galois coring: to a ring extension $i: B \to A$, one can associate the so-called canonical coring; a morphism from the canonical coring to another coring $\mathcal{C}$ is determined completely by a grouplike element x; if this morphism is an isomorphism, then we say that $(\mathcal{C}, x)$ is a Galois coring.

The canonical coring leads to an elegant formulation of descent theory: the category of descent data associated to the extension $i: B \to A$ is nothing else then the category of comodules over the canonical coring. This is no surprise: to an A-coring, we can associate a comonad on the category of A-modules, and the canonical coring is exactly the comonad associated to the adjoint pair of functors, given by induction and restriction of scalars. Thus, if a coring is isomorphic to the canonical coring, and if the induction functor is comonadic, then the category of descent data is isomorphic to the category of comodules over this coring, and equivalent to the category of B-modules. This unifies all the versions of descent theory that we mentioned at the beginning of this introduction.

In this paper, we present a survey of properties of Galois corings that have been obtained so far. We have organized it as follows: in Section 1, we recall definition, basic properties and examples of comodules over corings; in Section 2, we explain how descent theory can be formulated using the canonical coring. We included a full proof of Proposition 2.3, which is the noncommutative version of the fact that the induction functor is comonadic if and only if the ring morphism is pure as a map of modules. In Section 3, we introduce Galois corings, and discuss some properties, taken from [9] and [41]. In Section 4, Morita theory is applied to find some equivalent properties for a progenerator coring to be Galois; in fact, Theorem 4.7 is a new result, and generalizes results of Chase and Sweedler [17]. In Section 5, we look at special cases, and we show how to recover the "old" Galois theories. In Section 6, we present a recent generalization, due to El Kaoutit and Gómez Torrecillas [25].

1 Corings

Let A be a ring (with unit). The category of A-bimodules is a braided monoidal category, and an A-coring is by definition a coalgebra in the category of A-bimodules. Thus an A-coring is a triple $\mathcal{C} = (\mathcal{C}, \Delta_{\mathcal{C}}, \varepsilon_{\mathcal{C}})$, where

- $\mathcal{C}$ is an A-bimodule;
- $\Delta_{\mathcal{C}}: \mathcal{C} \to \mathcal{C} \otimes_A \mathcal{C}$ is an A-bimodule map;
- $\varepsilon_{\mathcal{C}}: \mathcal{C} \to A$ is an A-bimodule map

such that

$$(\Delta_{\mathcal{C}} \otimes_A I_{\mathcal{C}}) \circ \Delta_{\mathcal{C}} = (I_{\mathcal{C}} \otimes_A \Delta_{\mathcal{C}}) \circ \Delta_{\mathcal{C}}, \tag{1.1}$$

and

$$(I_{\mathcal{C}} \otimes_A \varepsilon_{\mathcal{C}}) \circ \Delta_{\mathcal{C}} = (\varepsilon_{\mathcal{C}} \otimes_A I_{\mathcal{C}}) \circ \Delta_{\mathcal{C}} = I_{\mathcal{C}}. \tag{1.2}$$

Sometimes corings are considered as coalgebras over noncommutative rings. This point of view is not entirely correct: a coalgebra over a commutative ring k is a k-coring, but not conversely: it could be that the left and and right action of k on the coring are different.

The Sweedler-Heyneman notation is also used for a coring $\mathcal{C}$, namely

$$\Delta_{\mathcal{C}}(c) = c_{(1)} \otimes_A c_{(2)},$$

where the summation is implicitely understood. (1.2) can then be written as

$$\varepsilon_{\mathcal{C}}(c_{(1)})c_{(2)} = c_{(1)}\varepsilon_{\mathcal{C}}(c_{(2)}) = c.$$

This formula looks like the corresponding formula for usual coalgebras. Notice however that the order matters in the above formula, since $\varepsilon_{\mathcal{C}}$ now takes values in A which is noncommutative in general. Even worse, the expression $c_{(2)}\varepsilon_{\mathcal{C}}(c_{(1)})$ makes no sense at all, since we have no well-defined switch map $\mathcal{C} \otimes_A \mathcal{C} \to \mathcal{C} \otimes_A \mathcal{C}$. A morphism between two corings $\mathcal{C}$ and $\mathcal{D}$ is an A-bimodule map $f:\ \mathcal{C} \to \mathcal{D}$ such that

$$\Delta_{\mathcal{D}}(f(c)) = f(c_{(1)}) \otimes_A f(c_{(2)}) \text{ and } \varepsilon_{\mathcal{D}}(f(c)) = \varepsilon_{\mathcal{C}}(c),$$

for all $c \in \mathcal{C}$. A right $\mathcal{C}$-comodule $M = (M, \rho)$ consists of a right A-module M together with a right A-linear map $\rho:\ M \to M \otimes_A \mathcal{C}$ such that:

$$(\rho \otimes_A I_{\mathcal{C}}) \circ \rho = (I_M \otimes_A \Delta_{\mathcal{C}}) \circ \rho, \tag{1.3}$$

and

$$(I_M \otimes_A \varepsilon_{\mathcal{C}}) \circ \rho = I_M. \tag{1.4}$$

We then say that $\mathcal{C}$ coacts from the right on M. Left $\mathcal{C}$-comodules and $\mathcal{C}$-bicomodules can be defined in a similar way. We use the Sweedler-Heyneman notation also for comodules:

$$\rho(m) = m_{[0]} \otimes_A m_{[1]}.$$

(1.4) then takes the form $m_{[0]}\varepsilon_{\mathcal{C}}(m_{[1]}) = m$. A right A-linear map $f:\ M \to N$ between two right $\mathcal{C}$-comodules M and N is called right $\mathcal{C}$-colinear if $\rho(f(m)) = f(m_{[0]}) \otimes m_{[1]}$, for all $m \in M$.

Corings were already considered by Sweedler in [38]. The interest in corings was revived after a mathematical review written by Takeuchi [39], in which he observed that entwined modules can be considered as comodules over a coring. This will be discussed in the examples below.

Example 1.1 As we already mentioned, if A is a commutative ring, then an A-coalgebra is also an A-coring.

Example 1.2 Let $i:\ B \to A$ be a ring morphism; then $\mathcal{D} = A \otimes_B A$ is an A-coring. We define

$$\Delta_{\mathcal{D}}:\ \mathcal{D} \to \mathcal{D} \otimes_A \mathcal{D} \cong A \otimes_B A \otimes_B A$$

and

$$\varepsilon_{\mathcal{D}}:\ \mathcal{D} = A \otimes_B A \to A$$

by

$$\Delta_{\mathcal{D}}(a \otimes_B b) = (a \otimes_B 1_A) \otimes_A (1_A \otimes_B b) \cong a \otimes_B 1_A \otimes_B b$$

and

$$\varepsilon_{\mathcal{D}}(a \otimes_B b) = ab.$$

Then $\mathcal{D} = (\mathcal{D}, \Delta_{\mathcal{D}}, \varepsilon_{\mathcal{D}})$ is an A-coring. It is called the canonical coring associated to the ring morphism i. We will see in the next section that this coring is crucial in descent theory.

Example 1.3 Let k be a commutative ring, G a finite group, and A a G-module algebra. Let $\mathcal{C} = \oplus_{\sigma\in G} Av_\sigma$ be the left free A-module with basis indexed by G, and let $p_\sigma : \mathcal{C} \to A$ be the projection onto the free component Av_σ. We make $\mathcal{C}$ into a right A-module by putting

$$v_\sigma a = \sigma(a)v_\sigma.$$

A comultiplication and counit on $\mathcal{C}$ are defined by putting

$$\Delta_{\mathcal{C}}(av_\sigma) = \sum_{\tau\in G} av_\tau \otimes_A v_{\tau^{-1}\sigma} \text{ and } \varepsilon_{\mathcal{C}} = p_e,$$

where e is the unit element of G. It is straightforward to verify that $\mathcal{C}$ is an A-coring. Notice that, in the case where A is commutative, we have an example of an A-coring, which is not an A-coalgebra, since the left and right A-action on $\mathcal{C}$ do not coincide.

Let us give a description of the right $\mathcal{C}$-comodules. Assume that $M = (M, \rho)$ is a right $\mathcal{C}$-comodule. For every $m \in M$ and $\sigma \in G$, let $\overline{\sigma}(m) = m_\sigma = (I_M \otimes_A p_\sigma)(\rho(m))$. Then we have

$$\rho(m) = \sum_{\sigma\in G} m_\sigma \otimes_A v_\sigma.$$

$\overline{e}$ is the identity, since $m = (I_M \otimes_A \varepsilon_{\mathcal{C}}) \circ \rho(m) = m_e$. Using the coassociativity of the comultiplication, we find

$$\sum_{\sigma\in G} \rho(m_\sigma) \otimes v_\sigma = \sum_{\sigma,\tau\in G} m_\sigma \otimes_A v_\tau \otimes_A v_{\tau^{-1}\sigma} = \sum_{\rho,\tau\in G} m_{\tau\rho} \otimes_A v_\tau \otimes_A v_\rho,$$

hence $\rho(m_\sigma) = \sum_{\tau\in G} m_{\tau\sigma} \otimes_A v_\tau$, and $\overline{\tau}(\overline{\sigma}(m)) = m_{\tau\sigma} = \overline{\tau\sigma}(m)$, so G acts as a group of k-automorphisms on M. Moreover, since ρ is right A-linear, we have that

$$\rho(ma) = \sum_{\sigma\in G} \overline{\sigma}(ma) \otimes_A v_\sigma = \sum_{\sigma\in G} \overline{\sigma}(m) \otimes_A v_\sigma a = \sum_{\sigma\in G} \overline{\sigma}(m)\sigma(a) \otimes_A v_\sigma$$

so $\overline{\sigma}$ is A-semilinear (cf. [29, p. 55]): $\overline{\sigma}(ma) = \overline{\sigma}(m)\sigma(a)$, for all $m \in M$ and $a \in A$. Conversely, if G acts as a group of right A-semilinear automorphims on M, then the formula

$$\rho(m) = \sum_{\sigma\in G} \overline{\sigma}(m) \otimes_A v_\sigma$$

defines a right $\mathcal{C}$-comodule structure on $\mathcal{M}$.

Example 1.4 Now let k be a commutative ring, G an arbitrary group, and A a G-graded k-algebra. Again let $\mathcal{C}$ be the free left A-module with basis indexed by G:

$$\mathcal{C} = \oplus_{\sigma\in G} Au_\sigma$$

Right A-action, comultiplication and counit are now defined by

$$u_\sigma a = \sum_{\tau\in G} a_\tau u_{\sigma\tau} \ ; \ \Delta_{\mathcal{C}}(u_\sigma) = u_\sigma \otimes_A u_\sigma \ ; \ \varepsilon_{\mathcal{C}}(u_\sigma) = 1.$$

$\mathcal{C}$ is an A-coring; let $M = (M, \rho)$ be a right $\mathcal{C}$-comodule, and let $M_\sigma = \{m \in M \mid \rho(m) = m \otimes_A u_\sigma\}$. It is then clear that $M_\sigma \cap M_\tau = \{0\}$ if $\sigma \neq \tau$. For any $m \in M$, we can write in a unique way:

$$\rho(m) = \sum_{\sigma\in G} m_\sigma \otimes_A u_\sigma.$$

Using the coassociativity, we find that $m_\sigma \in M_\sigma$, and using the counit property, we find that $m = \sum_\sigma m_\sigma$. So $M = \oplus_{\sigma\in G} M_\sigma$. Finally, if $m \in M_\sigma$ and $a \in A_\tau$, then it follows from the right A-linearity of ρ that

$$\rho(ma) = (m \otimes_A u_\sigma)a = ma \otimes_A u_{\sigma\tau},$$

so $ma \in M_{\sigma\tau}$, and $M_\sigma A_\tau \subset M_{\sigma\tau}$, and M is a right G-graded A-module. Conversely, every right G-graded A-module can be made into a right $\mathcal{C}$-comodule.

Example 1.5 Let H be a bialgebra over a commutative ring k, and A a right H-comodule algebra. Now take $\mathcal{C} = A \otimes H$, with A-bimodule structure

$$a'(b \otimes h)a = a'ba_{[0]} \otimes ha_{[1]}.$$

Now identify $(A\otimes H)\otimes_A(A\otimes H) \cong A\otimes H\otimes H$, and define the comultiplication and counit on $\mathcal{C}$, by putting $\Delta_\mathcal{C} = I_A \otimes \Delta_H$ and $\varepsilon_\mathcal{C} = I_A \otimes \varepsilon_H$. Then $\mathcal{C}$ is an A-coring. The category $\mathcal{M}^\mathcal{C}$ is isomorphic to the category of relative Hopf modules. These are k-modules M with a right A-action and a right H-coaction ρ, such that

$$\rho(ma) = m_{[0]}a_{[0]} \otimes_A m_{[1]}a_{[1]}$$

for all $m \in M$ and $a \in A$.

Example 1.6 Let k be a commutative ring, A a k-algebra, and C a k-coalgebra, and consider a k-linear map $\psi: C \otimes A \to A \otimes C$. We use the following Sweedler type notation, where the summation is implicitely understood:

$$\psi(c \otimes a) = a_\psi \otimes c^\psi = a_\Psi \otimes c^\Psi.$$

(A, C, ψ) is called a (right-right) entwining structure if the four following conditions are satisfied:

$$(ab)_\psi \otimes c^\psi = a_\psi b_\Psi \otimes c^{\psi\Psi}; \tag{1.5}$$
$$(1_A)_\psi \otimes c^\psi = 1_A \otimes c; \tag{1.6}$$
$$a_\psi \otimes \Delta_C(c^\psi) = a_{\psi\Psi} \otimes c^\Psi_{(1)} \otimes c^\psi_{(2)}; \tag{1.7}$$
$$\varepsilon_C(c^\psi)a_\psi = \varepsilon_C(c)a. \tag{1.8}$$

Let $\mathcal{C} = A \otimes C$ as a k-module, with A-bimodule structure

$$a'(b \otimes c)a = a'ba_\psi \otimes c^\psi.$$

Comultiplication and counit on $A \otimes C$ are defined as in Example 1.5. $\mathcal{C}$ is a coring, and the category $\mathcal{M}^\mathcal{C}$ is isomorphic to the category $\mathcal{M}(\psi)^C_A$ of entwined modules. These are k-modules M with a right A-action and a right C-coaction ρ such that

$$\rho(ma) = m_{[0]}a_\psi \otimes_A m^\psi_{[1]},$$

for all $m \in M$ and $a \in A$.

Actually Examples 1.3, 1.4 and 1.5 are special cases of Example 1.6

- Example 1.3: take $C = (kG)^* = \oplus_{g\in G} kv_\sigma$, the dual of the group ring kG, and $\psi(v_\sigma \otimes a) = \sigma(a) \otimes v_\sigma$;
- Example 1.4: take $C = kG = \oplus_{g\in G} ku_\sigma$, the group ring, and $\psi(u_\sigma \otimes a) = \sum_{\tau\in G} a_\tau \otimes u_{\sigma\tau}$;
- Example 1.5: take $C = H$, and $\psi(h \otimes a) = a_{[0]} \otimes ha_{[1]}$.

If $\mathcal{C}$ is an A-coring, then its left dual ${}^*\mathcal{C} = {}_A\mathrm{Hom}(\mathcal{C}, A)$ is a ring, with (associative) multiplication given by the formula

$$f\#g = g \circ (I_{\mathcal{C}} \otimes_A f) \circ \Delta_{\mathcal{C}} \text{ or } (f\#g)(c) = g(c_{(1)}f(c_{(2)})), \tag{1.9}$$

for all left A-linear $f, g : \mathcal{C} \to A$ and $c \in \mathcal{C}$. The unit is $\varepsilon_{\mathcal{C}}$. We have a ring homomorphism $i : A \to {}^*\mathcal{C}$, $i(a)(c) = \varepsilon_{\mathcal{C}}(c)a$. We easily compute that

$$(i(a)\#f)(c) = f(ca) \text{ and } (f\#i(a))(c) = f(c)a, \tag{1.10}$$

for all $f \in {}^*\mathcal{C}$, $a \in A$ and $c \in \mathcal{C}$. We have a functor $F : \mathcal{M}^{\mathcal{C}} \to \mathcal{M}_{{}^*\mathcal{C}}$, where $F(M) = M$ as a right A-module, with right ${}^*\mathcal{C}$-action given by $m \cdot f = m_{[0]}f(m_{[1]})$, for all $m \in M$, $f \in {}^*\mathcal{C}$. If $\mathcal{C}$ is finitely generated and projective as a left A-module, then F is an isomorphism of categories: given a right ${}^*\mathcal{C}$-action on M, we recover the right $\mathcal{C}$-coaction by putting $\rho(m) = \sum_j (m \cdot f_j) \otimes_A c_j$, where $\{(c_j, f_j) \mid j = 1, \cdots, n\}$ is a finite dual basis of $\mathcal{C}$ as a left A-module. ${}^*\mathcal{C}$ is a right A-module, by (1.10): $(f \cdot a)(c) = f(c)a$, and we can consider the double dual $({}^*\mathcal{C})^* = \mathrm{Hom}_A({}^*\mathcal{C}, A)$. We have a canonical morphism $i : \mathcal{C} \to ({}^*\mathcal{C})^*$, $i(c)(f) = f(c)$, and we call $\mathcal{C}$ reflexive (as a left A-module) if i is an isomorphism. If $\mathcal{C}$ is finitely generated projective as a left A-module, then $\mathcal{C}$ is reflexive. For any $\varphi \in ({}^*\mathcal{C})^*$, we then have that $\varphi = i(\sum_j \varphi(f_j)c_j)$.

Corings with a grouplike element. Let $\mathcal{C}$ be an A-coring, and suppose that $\mathcal{C}$ coacts on A. Then we have a map $\rho : A \to A \otimes_A \mathcal{C} \cong \mathcal{C}$. The fact that ρ is right A-linear implies that ρ is completely determined by $\rho(1_A) = x$: $\rho(a) = xa$. The coassociativity of the coaction yields that $\Delta_{\mathcal{C}}(x) = x \otimes_A x$ and the counit property gives us that $\varepsilon_{\mathcal{C}}(x) = 1_A$. We say that x is a *grouplike element* of $\mathcal{C}$ and we denote $G(\mathcal{C})$ for the set of all grouplike elements of $\mathcal{C}$. If $x \in G(\mathcal{C})$ is grouplike, then the associated $\mathcal{C}$-coaction on A is given by $\rho(a) = xa$.

If $x \in G(\mathcal{C})$, then we call $(\mathcal{C}, x)$ a *coring with a fixed grouplike element.* For $M \in \mathcal{M}^{\mathcal{C}}$, we call

$$M^{\mathrm{co}\mathcal{C}} = \{m \in M \mid \rho(m) = m \otimes_A x\}$$

the submodule of coinvariants of M; note that this definition depends on the choice of the grouplike element. Also observe that

$$A^{\mathrm{co}\mathcal{C}} = \{b \in A \mid bx = xb\}$$

is a subring of A. Let $i : B \to A$ be a ring morphism. i factorizes through $A^{\mathrm{co}\mathcal{C}}$ if and only if

$$x \in G(\mathcal{C})^B = \{x \in G(\mathcal{C}) \mid xb = bx, \text{ for all } b \in B\}.$$

We then have two pairs of adjoint functors (F, G) and (F', G'), respectively between the categories $\mathcal{M}_B$ and $\mathcal{M}^{\mathcal{C}}$ and the categories ${}_B\mathcal{M}$ and ${}^{\mathcal{C}}\mathcal{M}$. For $N \in \mathcal{M}_B$ and $M \in \mathcal{M}^{\mathcal{C}}$,

$$F(N) = N \otimes_B A \text{ and } G(M) = M^{\mathrm{co}\mathcal{C}}.$$

The unit and counit of the adjunction are

$$\nu_N : N \to (N \otimes_B A)^{\mathrm{co}\mathcal{C}}, \ \nu_N(n) = n \otimes_B 1;$$

$$\zeta_M : M^{\mathrm{co}\mathcal{C}} \otimes_B A \to M, \ \zeta_M(m \otimes_B a) = ma.$$

The other adjunction is defined in a similar way. We want to discuss when (F, G) and (F', G') are category equivalences. In Section 2, we will do this for the canonical coring associated to a ring morphism; we will study the general case in Section 3.

2 The canonical coring and descent theory

Let $i:\ B \to A$ be a ring morphism. The problem of descent theory is the following: suppose that we have a right A-module M. When do we have a right B-module N such that $M = N \otimes_B A$? The same problem can be stated for modules with an additional structure, for example algebras. In the case where A and B are commutative, this problem has been discussed in a purely algebraic context in [29]. In fact the results in [29] are the affine versions of Grothendieck's descent theory for schemes, see [28]. In the situation where A and B are arbitrary, descent theory has been discussed by Cipolla [18], and, more recently, by Nuss [33]. For a purely categorical treatment of the problem, making use of monads, we refer to [7]. Here we will show that the results in [29] and [18] can be restated elegantly in terms of comodules over the canonical coring.

Let $\mathcal{D} = A \otimes_B A$ be the canonical coring associated to the ring morphism $i:\ B \to A$, and let $M = (M, \rho)$ be a right $\mathcal{D}$-comodule. We will identify $M \otimes_A \mathcal{D} \cong M \otimes_B A$ using the natural isomorphism. The coassociativity and the counit property then take the form

$$\rho(m_{[0]}) \otimes m_{[1]} = m_{[0]} \otimes_B 1_A \otimes_B m_{[1]} \text{ and } m_{[0]}m_{[1]} = m.$$

$1_A \otimes_B 1_A$ is a grouplike element of $\mathcal{D}$. As we have seen at the end of Section 1, we have two pairs of adjoint functors, respectively between $\mathcal{M}_B$ and $\mathcal{M}^{\mathcal{D}}$, and ${}_B\mathcal{M}$ and ${}^{\mathcal{D}}\mathcal{M}$, which we will denote by (K, R) and (K', R'). The unit and counit of the adjunction will be denoted by η and ε. K is called the comparison functor. If (K, R) is an equivalence of categories, then the "descent problem" is solved: $M \in \mathcal{M}_A$ is isomorphic to some $N \otimes_B A$ if and only if we can define a right $\mathcal{D}$-coaction on M.

Recall that a morphism of left B-modules $f:\ M \to M'$ is called pure if and only if $f_N = I_N \otimes_B f:\ N \otimes_B M \to N \otimes_B M'$ is monic, for every $N \in \mathcal{M}_B$. $i:\ B \to A$ is pure as a morphism of left B-modules if and only if η_N being injective, for all $N \in \mathcal{M}_B$, since η_N factorizes through i_N.

Proposition 2.1 *The comparison functor K is fully faithful if and only if $i:\ B \to A$ is pure as a morphism of left B-modules.*

Proof The comparison functor K is fully faithful if and only if η_N is bijective, for all $N \in \mathcal{M}_B$. From the above observation, it follows that it suffices to show that left purity of i implies that η_N is surjective. Since η_N is injective, we have that $N \subset (N \otimes_B A)^{\mathrm{co}\mathcal{D}} \subset N \otimes_B A$. Take $q = \sum_i n_i \otimes_B a_i \in (N \otimes_B A)^{\mathrm{co}\mathcal{D}}$. Then

$$\rho(\sum_i n_i \otimes_B a_i) = \sum_i n_i \otimes_B a_i \otimes_B 1 = \sum_i n_i \otimes_B 1 \otimes_B a_i. \tag{2.1}$$

Consider the right B-module $P = (P \otimes_B A)/N$, and let $\pi:\ N \otimes_B A \to P$ be the canonical projection. Applying $\pi \otimes_B I_A$ to (2.1), we obtain

$$\pi(q) \otimes_B 1 = \sum_i \pi(n_i \otimes_B 1) \otimes_B a_i = 0 \text{ in } P \otimes_B A,$$

hence $\pi(q) = 0$, since i_P is an injection. This means that $q \in N$, which is exactly what we needed. $\square$

We also have an easy characterization of the fact that R is fully faithful.

Proposition 2.2 *The right adjoint R of the comparison functor K is fully faithful if and only if $\bullet \otimes_B A$ preserves the equalizer of ρ and i_M, for every $(M,\rho) \in \mathcal{M}^{\mathcal{D}}$. In particular, if A is flat as a left B-module, then R is fully faithful.*

Proof $M^{\mathrm{co}\mathcal{D}}$ is the equalizer of the maps

$$0 \longrightarrow M^{\mathrm{co}\mathcal{D}} \xrightarrow{j} M \underset{i_M}{\overset{\rho}{\rightrightarrows}} M \otimes_B A.$$

First assume that $M^{\mathrm{co}\mathcal{D}} \otimes_B A$ is the equalizer

$$0 \longrightarrow M^{\mathrm{co}\mathcal{D}} \otimes_B A \xrightarrow{j \otimes_B I_A} M \otimes_B A \underset{i_M \otimes_B I_A}{\overset{\rho \otimes_B I_A}{\rightrightarrows}} M \otimes_B A \otimes_B A. \tag{2.2}$$

From the coassociativity of ρ, it now follows that $\rho(m) \in M^{\mathrm{co}\mathcal{D}} \otimes_B A \subset M \otimes_B A \cong M \otimes_A (A \otimes_B A)$, for all $m \in M$, and we have a map $\rho: M \to M^{\mathrm{co}\mathcal{D}} \otimes_B A$. From the counit property, it follows that $\varepsilon_M \circ \rho = I_M$. For $m \in M^{\mathrm{co}\mathcal{D}}$ and $a \in A$, we have

$$\rho(\varepsilon_M(m \otimes_B a)) = \rho(ma) = \rho(m)a = (m \otimes_B 1)a = m \otimes_B a.$$

Thus the counit ε_M has an inverse, for all M, and R is fully faithful.

Conversely, assume that ε_M is bijective. Take $\sum_i m_i \otimes_B a_i \in M^{\mathrm{co}\mathcal{D}} \otimes_B A$, and put $m = \sum_i m_i a_i \in M$. Then $\rho(m) = m_{[0]} \otimes_B m_{[1]} = \sum_i m_i \otimes_B a_i \in M \otimes_B A$. Consequently, if $\sum_i m_i \otimes_B a_i = 0 \in M \otimes_B A$, then $m = m_{[0]}m_{[1]} = 0$, so $\sum_i m_i \otimes_B a_i = 0 \in M^{\mathrm{co}\mathcal{D}} \otimes_B A$, and we have shown that the canonical map $M^{\mathrm{co}\mathcal{D}} \otimes_B A \to M \otimes_B A$ is injective. □

If A and B are commutative, and $i: B \to A$ is pure as a morphism of B-modules, then $\bullet \otimes_B A$ preserves the equalizer of ρ and i_M, for every $(M,\rho) \in \mathcal{M}^{\mathcal{D}}$, and therefore R is fully faithful. This result is due to Joyal and Tierney (unpublished); an elementary proof was given recently by Mesablishvili [32]. We will now adapt Mesablishvili's proof to the noncommutative situation. In view of Proposition 2.1, one would expect that a sufficient condition for the fully faithfulness of R is the fact that i is pure as a morphism of left B-modules. It came as a surprise to the author that we need right purity instead of left purity.

We consider the contravariant functor $C = \mathrm{Hom}_{\mathbb{Z}}(\bullet, \mathbb{Q}/\mathbb{Z}) : \underline{\underline{\mathrm{Ab}}} \to \underline{\underline{\mathrm{Ab}}}$. $\mathbb{Q}/\mathbb{Z}$ is an injective cogenerator of $\underline{\underline{\mathrm{Ab}}}$, and therefore C is exact and reflects isomorphisms. If B is a ring, then C induces functors

$$C: \mathcal{M}_B \to {}_B\mathcal{M} \text{ and } {}_B\mathcal{M} \to \mathcal{M}_B.$$

For example, if $M \in \mathcal{M}_B$, then $C(M)$ is a left B-module, by putting $(b \cdot f)(m) = f(mb)$. For $M \in \mathcal{M}_B$ and $P \in {}_BM$, we have the following isomorphisms, natural in M and P:

$$\mathrm{Hom}_B(M, C(P)) \cong {}_B\mathrm{Hom}(P, C(M)) \cong C(M \otimes_B P) \tag{2.3}$$

If $P \in {}_B\mathcal{M}_B$, then $C(P) \in {}_B\mathcal{M}_B$, and the above isomorphisms are isomorphisms of left B-modules.

Proposition 2.3 *Let $i: B \to A$ be a ring morphism, and assume that i is pure as a morphism of right B-modules. Then the adjoint R of the comparison functor is fully faithful.*

Proof We have to show that (2.2) is exact, for all $(M,\rho) \in \mathcal{M}^{\mathcal{D}}$. If i is pure in $\mathcal{M}_B$, then $i_{C(B)}: B \otimes_B C(B) \to A \otimes_B C(B)$ is a monomorphism in $\mathcal{M}_B$, hence

$$C(i_{C(B)}): C(A \otimes_B C(B)) \to C(B \otimes_B C(B))$$

is an epimorphism in ${}_B\mathcal{M}$. Applying (2.3), we find that

$$C(i) \circ \bullet: {}_B\mathrm{Hom}(C(B), C(A)) \to {}_B\mathrm{Hom}(C(B), C(B))$$

is also an epimorphism. This implies that $C(i): C(A) \to C(B)$ is a split epimorphism in ${}_B\mathcal{M}$, and then it follows that for every $M \in \mathcal{M}_B$,

$$C(i) \circ \bullet: \mathrm{Hom}_B(M, C(A)) \to \mathrm{Hom}_B(M, C(B))$$

is a split epimorphism in ${}_B\mathcal{M}$. Applying (2.3) again, we find that

$$C(i_M): C(M \otimes_B A) \to C(M)$$

is a split epimorphism in ${}_B\mathcal{M}$.
In $\mathcal{M}_B$, we have the following commutative diagram with exact rows:

$$\begin{tikzcd}
0 \arrow[r] & M^{\mathrm{co}\mathcal{D}} \arrow[r, "j"] \arrow[d, "j"] & M \arrow[r, shift left, "\rho"] \arrow[r, shift right, "i_M"'] \arrow[d, "i_M"] & M \otimes_B A \arrow[d, "i_{M\otimes_B A}"] \\
0 \arrow[r] & M \arrow[r, "\rho"] & M \otimes_B A \arrow[r, shift left, "\rho \otimes I_A"] \arrow[r, shift right, "i_M \otimes I_A"'] & M \otimes_B A \otimes_B A
\end{tikzcd}$$

Applying the functor C to this diagram, we obtain the following commutative diagram with exact rows. We also know that $C(i_M)$ and $C(i_{M\otimes_B A})$ have right inverses h and h'.

$$\begin{tikzcd}
C(M \otimes_B A \otimes_B A) \arrow[r, shift left, "C(\rho \otimes I_A)"] \arrow[r, shift right, "C(i_M \otimes I_A)"'] \arrow[d, shift left, "C(i_{M\otimes_B A})"] & C(M \otimes_B A) \arrow[r, "C(\rho)"] \arrow[d, shift left, "C(i_M)"] & C(M) \arrow[r] \arrow[d, shift left, "C(j)"] & 0 \\
C(M \otimes_B A) \arrow[r, shift left, "C(\rho)"] \arrow[r, shift right, "C(i_M)"'] \arrow[u, shift left, "h'"] & C(M) \arrow[r, "C(j)"] \arrow[u, shift left, "h"] & C(M^{\mathrm{co}\mathcal{D}}) \arrow[r] \arrow[u, shift left, "k"] & 0
\end{tikzcd}$$

Diagram chasing leads to the existence of a right inverse k of $C(j)$, such that $k \circ C(j) = C(\rho) \circ h$. But this means that the bottom row of the above diagram is a split fork in ${}_B\mathcal{M}$, split by the morphisms

$$C(M \otimes_B A) \xleftarrow{h} C(M) \xleftarrow{k} C(M^{\mathrm{co}\mathcal{D}})$$

(see [31, p.149] for the definition of a split fork). Split forks are preserved by arbitrary functors, so applying ${}_B\mathrm{Hom}(A, \bullet)$, we obtain a split fork in ${}_B\mathcal{M}$; using (2.3), we find that this split fork is isomorphic to

$$\begin{tikzcd}
C(M \otimes_B A \otimes_B A) \arrow[r, shift left, "C(\rho\otimes I_A)"] \arrow[r, shift right, "C(i_M\otimes I_A)"'] & C(M \otimes_B A) \arrow[r, "C(j\otimes I_A)"] & C(M^{\mathrm{co}\mathcal{D}} \otimes_B A)
\end{tikzcd}$$

The functor C is exact and reflects isomorphisms, hence it also reflects coequalizers. It then follows that (2.2) is an equalizer in $\mathcal{M}_B$, as needed. □

The converse of Proposition 2.3 is not true in general: the natural inclusion $i: \mathbb{Z} \to \mathbb{Q}$ is not pure in $\mathcal{M}_{\mathbb{Z}}$, but the functor R is fully faithful. Indeed, if $(M, \rho) \in \mathcal{M}^{\mathcal{D}}$, then M is a $\mathbb{Q}$-vector space, and $\rho: M \to M \otimes_{\mathbb{Z}} \mathbb{Q} \cong M$ is the identity, $M^{\mathrm{co}\mathcal{D}} = M$, and $\varepsilon_M: M^{\mathrm{co}\mathcal{D}} \otimes_{\mathbb{Z}} \mathbb{Q} \cong M \to M$ is also the identity.

It would be interesting to know if there exists a ring morphism $i: B \to A$ which is pure in ${}_B\mathcal{M}$, but not in $\mathcal{M}_B$, and such that (K, R) is an equivalence of categories.

Consider $K' = A \otimes_B \bullet: {}_B\mathcal{M} \to {}^{\mathcal{D}}\mathcal{M}$ and $R' = {}^{\mathrm{co}\mathcal{D}}(\bullet): {}^{\mathcal{D}}\mathcal{M} \to {}_B\mathcal{M}$. The next result is an immediate consequence of Propositions 2.2 and 2.3 and their left handed versions, and can be viewed as the noncommutative version of the Joyal-Tierney Theorem.

Theorem 2.4 *Let $i: B \to A$ be a morphism of rings. Then the following assertions are equivalent.*

1. *(K, R) and (K', R') are equivalences of categories;*
2. *K and K' are fully faithful;*
3. *i is pure in $\mathcal{M}_B$ and ${}_B\mathcal{M}$.*

We have seen in Proposition 2.2 that R is fully faithful if A is flat as a left B-module.

Proposition 2.5 *Let $i: B \to A$ be a morphism of rings, and assume that A is flat as a left B-module. Then (K, R) is an equivalence of categories if and only if A is faithfully flat as a left B-module.*

Proof First assume that A is faithfully flat as a left B-module. It follows from Proposition 2.1 that it suffices to show that A is pure as a left B-module. For $N \in \mathcal{M}_B$, the map

$$f = i_N \otimes_B I_A: N \otimes_B A \to N \otimes_B A \otimes_B A$$

is injective: if $f(\sum_i n_i \otimes_B a_i) = \sum_i n_i \otimes_B 1 \otimes_B a_i = 0$, then, multiplying the second and third tensor factor, we find that $\sum_i n_i \otimes_B a_i = 0$. Since A is faithfully flat as a left B-module, it follows that i_N is injective.

Conversely assume that (K, R) is an equivalence of categories. Then the functor R is exact. Let

$$0 \to N' \to N \to N'' \to 0 \tag{2.4}$$

be a sequence of right B-modules such that the sequence

$$0 \to N' \otimes_B A \to N \otimes_B A \to N'' \otimes_B A \to 0$$

is exact. Applying the functor R to the sequence, and using the fact that η is an isomorphism, we find that (2.4) is exact, so it follows that A is faithfully flat as a left B-module. □

The descent data that are considered in [18] are nothing else then comodules over the canonical coring (although the author of [18] was not aware of this). The descent data in [29] are different, so let us indicate how to go from descent data to comodules over the canonical coring.

Let $i: B \to A$ be a morphism of commutative rings. A *descent datum* consists of a pair (M, g), with $M \in \mathcal{M}_A$, and $g: A \otimes_B M \to M \otimes_B A$ an $A \otimes_B A$-module homorphism such that

$$g_2 = g_3 \circ g_1: A \otimes_B A \otimes_B M \to A \otimes_B M \otimes_B A \tag{2.5}$$

and

$$\mu_M(g(1 \otimes_B m)) = m, \tag{2.6}$$

for all $m \in M$. Here g_i is obtained by applying I_A to the i-th tensor position, and g to the two other ones. It can be shown that (2.6) can be replaced by the condition that g is a bijection. A morphism of two descent data (M,g) and (M',g') consists of an A-module homomorphism $f: M \to M'$ such that

$$(f \otimes_B I_A) \circ g = g' \circ (I_A \otimes_B f).$$

$\underline{\underline{\mathrm{Desc}}}(A/B)$ will be the category of descent data.

Proposition 2.6 *Let $i: B \to A$ be a morphism of commutative rings. We have an isomorphism of categories*

$$\underline{\underline{\mathrm{Desc}}}(A/B) \cong \mathcal{M}^{A \otimes_B A}$$

Proof (sketch) For a right $\mathcal{D}$-comodule (M,ρ), we define $g: A \otimes_B M \to M \otimes_B A$ by $g(a \otimes_B m) = m_{[0]} a \otimes_B m_{[1]}$. Then (M,g) is a descent datum. Conversely, given a descent datum (M,g), the map $\rho: M \to M \otimes_B A$, $\rho(m) = g(1 \otimes_B m)$ makes M into a right $\mathcal{D}$-comodule. □

3 Galois corings

Let A be a ring, $(\mathcal{C},x)$ a coring with a fixed grouplike element, and $i: B \to A^{\mathrm{co}\mathcal{C}}$ a ring morphism. We have seen at the end of Section 1 that we have two pairs of adjoint functors (F,G) and (F',G'). We also have a morphism of corings

$$\mathrm{can}: \mathcal{D} = A \otimes_B A \to \mathcal{C},\ \mathrm{can}(a \otimes_B a') = axa'.$$

Proposition 3.1 *With notation as above, we have the following results.*

1. *If F is fully faithful, then $i: B \to A^{\mathrm{co}\mathcal{C}}$ is an isomorphism;*
2. *if G is fully faithful, then* can $: \mathcal{D} = A \otimes_B A \to \mathcal{C}$ *is an isomorphism.*

Proof 1. F is fully faithful if and only if ν is an isomorphism; it then suffices to observe that $i = \nu_B$.
2. G is fully faithful if and only if ζ is an isomorphism. $\mathcal{C} \in \mathcal{M}^{\mathcal{C}}$, the right coaction being induced by the comultiplication. The map $f: A \to \mathcal{C}^{\mathrm{co}\mathcal{C}}$, $f(a) = ax$, is an isomorphism of (A,B)-bimodules; the inverse of f is the restriction of $\varepsilon_{\mathcal{C}}$ to $\mathcal{C}^{\mathrm{co}\mathcal{C}}$. Indeed, if $c \in \mathcal{C}^{\mathrm{co}\mathcal{C}}$, then $\Delta_{\mathcal{C}}(c) = c \otimes_A x$, hence $c = \varepsilon(c)x = f(\varepsilon(c))$. It follows that $\mathrm{can} = \zeta_{\mathcal{C}} \circ (f \otimes_B I_A)$ is an isomorphism. □

Proposition 3.1 leads us to the following Definition.

Definition 3.2 Let $(\mathcal{C},x)$ be an A-coring with a fixed grouplike, and let $B = A^{\mathrm{co}\mathcal{C}}$. We call $(\mathcal{C},x)$ a Galois coring if the canonical coring morphism can $: \mathcal{D} = A \otimes_B A \to \mathcal{C}$, $\mathrm{can}(a \otimes_B b) = axb$ is an isomorphism.

Let $i: B \to A$ be a ring morphism. If $x \in G(\mathcal{C})^B$, then we can define a functor

$$\Gamma: \mathcal{M}^{\mathcal{D}} \to \mathcal{M}^{\mathcal{C}},\ \Gamma(M,\rho) = (M,\widetilde{\rho})$$

with $\widetilde{\rho}(m) = m_{[0]} \otimes_A x m_{[1]} \in M \otimes_A \mathcal{C}$ if $\rho(m) = m_{[0]} \otimes_B m_{[1]} \in M \otimes_B A$. It is easy to see that $\Gamma \circ K = F$, and therefore we have the following result.

Proposition 3.3 *Let $(\mathcal{C},x)$ be a Galois A-coring. Then Γ is an isomorphism of categories. Consequently R (resp. K) is fully faithful if and only if G (resp. F) is fully faithful.*

Let us now give some alternative characterizations of Galois corings; for the proof, we refer to [41, 3.6].

Proposition 3.4 *Let $(\mathcal{C}, x)$ be an A-coring with fixed grouplike element, and $B = A^{\mathrm{co}\mathcal{C}}$. The following assertions are equivalent.*

1. *$(\mathcal{C}, x)$ is Galois;*
2. *if $(M, \rho) \in \mathcal{M}^{\mathcal{C}}$ is such that $\rho: M \to M \otimes_A \mathcal{C}$ is a coretraction, then the evaluation map*
$$\varphi_M: \mathrm{Hom}^{\mathcal{C}}(A, M) \otimes_B A \to M, \ \varphi_M(f \otimes_B m) = f(m)$$
is an isomorphism;
3. *$\varphi_{\mathcal{C}}$ is an isomorphism.*

From Theorem 2.4 and Proposition 3.3, we immediately obtain the following result.

Theorem 3.5 *Let $(\mathcal{C}, x)$ be a Galois A-coring, and put $B = A^{\mathrm{co}\mathcal{C}}$. Then the following assertions are equivalent.*

1. *(F, G) and (F', G') are equivalences of categories;*
2. *the functors F and F' are fully faithful;*
3. *$i: B \to A$ is pure in ${}_B\mathcal{M}$ and $\mathcal{M}_B$.*

Remark 3.6 Let us make some remarks about terminology. In the literature, there is an inconsistency in the use of the term "Galois". An alternative definition is to require that $(\mathcal{C}, x)$ satisfies the equivalent definitions of Theorem 3.5, so that (F, G) and (F', G') are category equivalences. In Section 5, we will discuss special cases that have appeared in the literature before. In some cases, there is an agreement with Definition 3.2 (see e.g. [8], [35]), while in other cases, category equivalence is required (see e.g. [17], [21]).

In the particular situation where $\mathcal{C} = A \otimes H$, as in Example 1.5, the property that (F, G) is an equivalence (resp. G is fully faithful) is called the Strong (resp. Weak) Structure Theorem (see [24]). Let $i: B \to A$ be a ring morphism, and $\mathcal{D}$ the canonical coring. In the situation where A and B are commutative, i is called a descent morphism (resp. an effective descent morphism) if K is fully faithful (resp. (K, R) is an equivalence). In the general situation, (K, R) is an equivalence if and only if the functor $\bullet \otimes_B A: \mathcal{M}_B \to \mathcal{M}_A$ is comonadic (see e.g. [7, Ch. 4]).

Let us next look at the case where A is flat as a left B-module. Wisbauer [41] calls the following two results the *Galois coring Structure Theorem.*

Proposition 3.7 *Let $(\mathcal{C}, x)$ be an A-coring with fixed grouplike element, and $B = A^{\mathrm{co}\mathcal{C}}$. Then the following statements are equivalent.*

1. *$(\mathcal{C}, x)$ is Galois and A is flat as a left B-module;*
2. *G is fully faithful and A is flat as a left B-module;*
3. *$\mathcal{C}$ is flat as a left A-module, and A is a generator in $\mathcal{M}^{\mathcal{C}}$.*

Proof 1) $\Rightarrow$ 2) follows from Propositions 2.2 and 3.3. 2) $\Rightarrow$ 1) follows from Proposition 3.1. For the proof of 1) $\Leftrightarrow$ 3), we refer to [41, 3.8]. □

Proposition 3.8 *Let $(\mathcal{C}, x)$ be an A-coring with fixed grouplike element, and $B = A^{\mathrm{co}\mathcal{C}}$. Then the following statements are equivalent.*

1. *$(\mathcal{C}, x)$ is Galois and A is faithfully flat as a left B-module;*
2. *(F, G) is an equivalence and A is flat as a left B-module;*

3. *$\mathcal{C}$ is flat as a left A-module, and A is a projective generator in $\mathcal{M}^{\mathcal{C}}$.*

Proof The equivalence of 1) and 2) follows from Propositions 2.5 and 3.3. For the remaining equivalence, we refer to [41]. □

A right $\mathcal{C}$-comodule N is called semisimple (resp. simple) in $\mathcal{M}^{\mathcal{C}}$ if every $\mathcal{C}$-monomorphism $U \to N$ is a coretraction (resp. an isomorphism). Similar definitions apply to left $\mathcal{C}$-comodules and $(\mathcal{C},\mathcal{C})$-bicomodules. $\mathcal{C}$ is said to be right (left) semisimple if it is semisimple as a right (left) $\mathcal{C}$-comodule. $\mathcal{C}$ is called a simple coring if it is simple as a $(\mathcal{C},\mathcal{C})$-bicomodule. For the proof of the following result, we refer to [26].

Proposition 3.9 *For an A-coring $\mathcal{C}$, the following assertions are equivalent:*

1. *$\mathcal{C}$ is right semisimple;*
2. *$\mathcal{C}$ is projective as a left A-module and $\mathcal{C}$ is semisimple as a left ${}^*\mathcal{C}$-module;*
3. *$\mathcal{C}$ is projective as a right A-module and $\mathcal{C}$ is semisimple as a right $\mathcal{C}^*$-module;*
4. *$\mathcal{C}$ is left semisimple.*

The connection to Galois corings is the following, due to Wisbauer [41, 3.12], and to El Kaoutit, Goméz Torrecillas and Lobillo [26, Theorem 4.4]:

Proposition 3.10 *For an A-coring with a fixed grouplike element $(\mathcal{C}, x)$, the following assertions are equivalent:*

1. *$\mathcal{C}$ is a simple and left (or right) semisimple coring;*
2. *$(\mathcal{C}, x)$ is Galois and $\mathrm{End}^{\mathcal{C}}(A)$ is simple and left semisimple;*
3. *$(\mathcal{C}, x)$ is Galois and B is a simple left semisimple subring of A;*
4. *$\mathcal{C}$ is flat as a right A-module, A is a projective generator in ${}^{\mathcal{C}}\mathcal{M}$, and $\mathrm{End}^{\mathcal{C}}(A)$ is simple and left semisimple (the left $\mathcal{C}$-coaction on A being given by $\rho^l(a) = ax$).*

4 Galois corings and Morita theory

Let $(\mathcal{C}, x)$ be a coring with a fixed grouplike element, $B = A^{\mathrm{co}\mathcal{C}}$, and $\mathcal{D} = A \otimes_B A$. We can consider the left dual of the map can:

$$^*\mathrm{can}: \ {}^*\mathcal{C} \to {}^*\mathcal{D} \cong {}_B\mathrm{End}(A)^{\mathrm{op}}, \ {}^*\mathrm{can}(f)(a) = f(xa).$$

The following result is obvious.

Proposition 4.1 *If $(\mathcal{C}, x)$ is Galois, then ${}^*\mathrm{can}$ is an isomorphism. The converse property holds if $\mathcal{C}$ and A are finitely generated projective, respectively as a left A-module, and a left B-module.*

Let $Q = \{q \in {}^*\mathcal{C} \mid c_{(1)}q(c_{(2)}) = q(c)x, \text{ for all } c \in \mathcal{C}\}$. A straightforward computation shows that Q is a $({}^*\mathcal{C}, B)$-bimodule. Also A is a left $(B, {}^*\mathcal{C})$-bimodule; the right ${}^*\mathcal{C}$-action is induced by the right $\mathcal{C}$-coaction: $a \cdot f = f(xa)$. Now consider the maps

$$\tau: \ A \otimes_{{}^*\mathcal{C}} Q \to B, \ \tau(a \otimes_{{}^*\mathcal{C}} q) = q(xa);$$

$$\mu: \ Q \otimes_B A \to {}^*\mathcal{C}, \ \mu(q \otimes_B a) = q\#i(a).$$

With this notation, we have the following property (see [14]).

Proposition 4.2 *$(B, {}^*\mathcal{C}, A, Q, \tau, \mu)$ is a Morita context.*

Properties of this Morita context are studied in [1], [2], [14] and [15]. It generalizes (and unifies) Morita contexts discussed in [5], [17], [19], [20] and [23]. We recall the following properties from [14] and [15].

Proposition 4.3 [14, Th. 3.3 and Cor. 3.4] *If τ is surjective, then $M^{\mathrm{co}\mathcal{C}} = M^{{}^*\mathcal{C}} = \{m \in M \mid m \cdot f = mf(x),$ for all $f \in {}^*\mathcal{C}\}$, for all $M \in \mathcal{M}^{\mathcal{C}}$.*
The following assertions are equivalent:

1. *τ is surjective;*
2. *there exists $q \in Q$ such that $q(x) = 1$;*
3. *for every $M \in \mathcal{M}^{{}^*\mathcal{C}}$, the map*

$$\omega_M : M \otimes_{{}^*\mathcal{C}} Q \to M^{{}^*\mathcal{C}}, \ \omega_M(m \otimes_{{}^*\mathcal{C}} q) = m \cdot q$$

is bijective.

Proposition 4.4 [15], [14, Th. 3.5] *The following assertions are equivalent:*

1. *μ is surjective;*
2. *$\mathcal{C}$ is finitely generated and projective as a left A-module and G is fully faithful.*

As an application of Proposition 4.3, we have the following result.

Corollary 4.5 *Assume that $\mathcal{C}$ is finitely generated projective as a left A-module. Consider the adjoint pair $(F = \bullet \otimes_B A, G = (\bullet)^{\mathrm{co}\mathcal{C}})$, and the functors $\widetilde{F} = \bullet \otimes_B A$ and $\widetilde{G} = \bullet \otimes_{{}^*\mathcal{C}} Q$ coming from the Morita context of Proposition 4.2. Then $F \cong \widetilde{F}$ and $G \cong \widetilde{G}$ if τ is surjective.*

Proof Take $N \in \mathcal{M}_B$. $F(N)$ corresponds to $\widetilde{F}$ under the isomorphism $\mathcal{M}^{\mathcal{C}} \cong \mathcal{M}_{{}^*\mathcal{C}}$. If τ is surjective, then it follows from Proposition 4.3 that $\omega : \widetilde{G} \to G$ is bijective. □

Let us now compute the Morita context associated to the canonical coring.

Proposition 4.6 *Let $i : B \to A$ be a ring morphism, and assume that i is pure as a morphism of left B-modules. Then the Morita context associated to the canonical coring $(\mathcal{D} = A \otimes_B A, 1 \otimes_B 1)$ is the Morita context $(B, {}_B\mathrm{End}(A)^{\mathrm{op}}, A, {}_B\mathrm{Hom}(A, B), \varphi, \psi)$ associated to A as a left B-module (see [4, II.4]).*

Proof From Proposition 2.1, it follows that

$$A^{\mathrm{co}\mathcal{D}} = \{b \in A \mid b \otimes_B 1 = 1 \otimes_B b\} = B.$$

Take $q \in Q \subset {}_A\mathrm{Hom}(A \otimes_B A, A)$ and the corresponding $\tilde{q} \in {}_B\mathrm{Hom}(A, A)$, given by $\tilde{q}(a) = q(1 \otimes_B a)$. Then

$$q(a' \otimes_B a)(1 \otimes_B 1) = (a' \otimes_B 1)q(1 \otimes_B a).$$

Taking $a' = 1$, we find

$$\tilde{q}(a) \otimes_B 1 = 1 \otimes_B \tilde{q}(a)$$

hence $\tilde{q}(a) \in B$, and it follows that $Q \subset {}_B\mathrm{Hom}(A, B)$. The converse inclusion is proved in a similar way. A straightforward verification shows that $\varphi = \tau$ and $\psi = \mu$. □

Recall that the context associated to the left A-module B is strict if and only if A is a left B-progenerator. We are now ready to prove the following result. In Section 5, we will see that it is a generalization of [17, Th. 9.3 and 9.6].

Theorem 4.7 *Let $(\mathcal{C},x)$ be a coring with fixed grouplike element, and assume that $\mathcal{C}$ is a left A-progenerator. We take a subring B' of $B = A^{\mathrm{co}\mathcal{C}}$, and consider the map*

$$\mathrm{can}' : \mathcal{D}' = A \otimes_{B'} A \to \mathcal{C},\ \mathrm{can}'(a \otimes_{B'} a') = axa'$$

Then the following statements are equivalent:

1. • can' *is an isomorphism;*
   • A *is faithfully flat as a left B'-module.*
2. • ${}^*\mathrm{can}'$ *is an isomorphism;*
   • A *is a left B'-progenerator.*
3. • $B = B'$;
   • *the Morita context $(B, {}^*\mathcal{C}, A, Q, \tau, \mu)$ is strict.*
4. • $B = B'$;
   • (F, G) *is an equivalence of categories.*

Proof 1) $\Leftrightarrow$ 2). Obviously ${}^*\mathrm{can}'$ is an isomorphism if can' is an isomorphism, and the converse holds if $\mathcal{C}$ is a left A-progenerator and A is a left B'-progenerator. If can' is an isomorphism, then $A \otimes_{B'} A = \mathcal{D}' \cong \mathcal{C}$ is a left A-progenerator, hence A is a left B'-progenerator.
1) $\Rightarrow$ 3). Since A is faithfully flat as a left B'-module, $A^{\mathrm{co}\mathcal{D}'} = B'$. Since can' is an isomorphism, it follows that $B = A^{\mathrm{co}\mathcal{C}} = A^{\mathrm{co}\mathcal{D}'} = B'$. Then $\mathrm{can} = \mathrm{can}'$ is an isomorphism, hence the Morita contexts associated to $(\mathcal{C},x)$ and $(\mathcal{D}, 1 \otimes_B 1)$ are isomorphic. From the equivalence of 1) and 2), we know that A is a left B-progenerator, so the context associated to $(\mathcal{D}, 1 \otimes_B 1)$ is strict, see the remark preceding Theorem 4.7. Therefore the Morita context $(B, {}^*\mathcal{C}, A, Q, \tau, \mu)$ associated to $(\mathcal{C},x)$ is also strict.
3) $\Rightarrow$ 1). and 3) $\Rightarrow$ 4). If $(B, {}^*\mathcal{C}, A, Q, \tau, \mu)$ is strict, then A is a left B-progenerator, and a fortiori faithfully flat as a left B-module. τ is surjective, so it follows from Corollary 4.5 that $F \cong \widetilde{F}$ and $G \cong \widetilde{G}$. $(\widetilde{F}, \widetilde{G})$ is an equivalence, so (F, G) is also an equivalence. Then $(\mathcal{C},x)$ is Galois by Proposition 3.1.
4) $\Rightarrow$ 1). can is an isomorphism, by Proposition 3.1, and we have already seen that this implies that A is a left B-progenerator, so A is faithfully flat as a left B-module. □

5 Application to particular cases

5.1 Coalgebra Galois extensions. From [10], we recall the following Definition.

Definition 5.1 Let $i : B \to A$ be a morphism of k-algebras, and C a k-coalgebra. A is called a C-Galois extension of B if the following conditions hold:

1. A is a right C-comodule;
2. $\mathrm{can} : A \otimes_B A \to A \otimes C$, $\mathrm{can}(a \otimes_B a') = aa'_{[0]} \otimes_B a'_{[1]}$ is an isomorphism;
3. $B = \{a \in A \mid \rho(a) = a\rho(1)\}$.

Proposition 5.2 *Let $i : B \to A$ be a morphism of k-algebras, and C a k-coalgebra. A is called a C-Galois extension of B if and only if there exists a right-right entwining structure (A, C, ψ) and $x \in G(A \otimes C)$ such that $A^{\mathrm{co}A\otimes C} = B$ and $(A \otimes C, x)$ is a Galois coring.*

Proof Let (A, C, ψ) be an entwining structure. We have seen in Example 1.6 that $\mathcal{C} = A \otimes C$ is a coring. Given a grouplike element $x = \sum_i a_i \otimes c_i$, we have a right $\mathcal{C}$-coaction on A, hence A is an entwined module (see Example 1.6), and therefore a C-comodule. The C-coaction is given by the formula

$$\rho(a) = a_{[0]} \otimes a_{[1]} = \sum_i a_i a_\psi \otimes c_i^\psi.$$

Then the conditions of Definition 5.1 are satisfied, and A is C-Galois extension of B.

Conversely, let A be a C-Galois extension of B. can is bijective, so the coring structure on $A \otimes_B A$ induces a coring structure on $A \otimes C$. We will show that this coring structure comes from an entwining structure (A, C, ψ).

It is clear that the natural left A-module structure on $A \otimes C$ makes can into a left A-linear map. The right A-module structure on $A \otimes C$ induced by can is given by

$$(b \otimes c)a = \mathrm{can}(\mathrm{can}^{-1}(b \otimes c)a).$$

Since $\mathrm{can}^{-1}(1_{[0]} \otimes 1_{[1]}) = 1 \otimes_B 1$, we have

$$(1_{[0]} \otimes 1_{[1]})a = \mathrm{can}(1 \otimes a) = a_{[0]} \otimes a_{[1]}. \tag{5.1}$$

The comultiplication Δ on $A \otimes C$ is given by

$$\Delta(a \otimes c) = (\mathrm{can} \otimes_A \mathrm{can})\Delta_{\mathcal{C}}(\mathrm{can}^{-1}(a \otimes c)) \in (A \otimes C) \otimes_A (A \otimes C),$$

for all $a \in A$ and $c \in C$. can is bijective, so we can find $a_i, b_i \in A$ such that

$$\mathrm{can}(\sum_i a_i \otimes_B b_i) = \sum_i a_i b_{i[0]} \otimes_B b_{i[1]} = a \otimes c,$$

and we compute that

$$\begin{aligned}
\Delta(a \otimes c) &= (\mathrm{can} \otimes_A \mathrm{can})\Delta_{\mathcal{C}}(\sum_i a_i \otimes_B b_i) \\
&= \sum_i \mathrm{can}(a_i \otimes_B 1) \otimes_A \mathrm{can}(1 \otimes_B b_i) \\
&= \sum_i (a_i 1_{[0]} \otimes 1_{[1]}) \otimes_A (b_{i[0]} \otimes b_{i[1]}) \\
&= \sum_i (a_i 1_{[0]} \otimes 1_{[1]}) b_{i[0]} \otimes_A (1 \otimes b_{i[1]}) \\
(5.1) \quad &= \sum_i (a_i b_{i[0]} \otimes b_{i[1]}) \otimes_A (1 \otimes b_{i[2]}) \\
&= (a \otimes c_{(1)}) \otimes_A (1 \otimes c_{(2)}).
\end{aligned}$$

Finally

$$\begin{aligned}
\varepsilon_{\mathcal{C}} c(a \otimes c) &= \varepsilon_{\mathcal{C}}(\sum_i a_i \otimes b_i) = \sum_i a_i b_i \\
&= \sum_i a_i b_{i[0]} \varepsilon_C(b_{i[1]}) = a\varepsilon_C(c).
\end{aligned}$$

Now define $\psi : C \otimes A \to A \otimes C$ by $\psi(c \otimes a) = (1_A \otimes c)a$. It follows from [9] that (A, C, ψ) is an entwining structure. $\square$

Let (A, C, ψ) be an entwining structure, and consider $g \in C$ grouplike. Then $x = 1_A \otimes g$ is a grouplike element of $A \otimes C$. Let us first describe the Morita context from the previous Section.

Observe that ${}^*\mathcal{C} = {}_A\mathrm{Hom}(A \otimes C, A) \cong \mathrm{Hom}(C, A)$ as a k-module. The ring structure on ${}^*\mathcal{C}$ induces a k-algebra structure on $\mathrm{Hom}(C, A)$, and this k-algebra is denoted $\#(C, A)$. The product is given by the formula

$$(f\#g)(c) = f(c_{(2)})_\psi g(c_{(1)}^\psi). \tag{5.2}$$

We have a natural algebra homomorphism $i:\ A \to \#(C, A)$, $i(a)(c) = \varepsilon_C(c)a$, and we have, for all $a \in A$ and $f:\ C \to A$:

$$(i(a)\#f)(c) = a_\psi f(c^\psi) \text{ and } (f\#i(a))(c) = f(c)a. \tag{5.3}$$

$\mathrm{Hom}(C, A)$ will denote the k-algebra with the usual convolution product, that is

$$(f * g)(c) = f(c_{(1)})g(c_{(2)}). \tag{5.4}$$

The ring of coinvariants is

$$B = A^{\mathrm{co}C} = \{b \in A \mid b_\psi \otimes g^\psi = b \otimes g\}, \tag{5.5}$$

and the bimodule Q is naturally isomorphic to

$$Q = \{q \in \#(C, A) \mid q(c_{(2)})_\psi \otimes c_{(1)}^\psi = q(c) \otimes g\}.$$

We have maps

$$\mu:\ Q \otimes_{B'} A \to \#(C, A),\ \mu(q \otimes_B a)(c) = q(c)a,$$
$$\tau:\ A \otimes_{\#(C,A)} Q \to B,\ \tau(a \otimes q) = a_\psi q(x^\psi),$$

and $(B, \#(C, A), A, Q, \tau, \mu)$ is a Morita context.

Proposition 5.3 [14, Prop. 4.3] *Assume that $\lambda:\ C \to A$ is convolution invertible, with convolution inverse λ^{-1}. Then the following assertions are equivalent:*

1. $\lambda \in Q$;
2. *for all $c \in C$, we have*
$$\lambda^{-1}(c_{(1)})\lambda(c_{(3)})_\psi \otimes c_{(2)}^\psi = \varepsilon(c)1_A \otimes g; \tag{5.6}$$
3. *for all $c \in C$, we have*
$$\lambda^{-1}(c_{(1)}) \otimes c_{(2)} = \lambda^{-1}(c)_\psi \otimes g^\psi. \tag{5.7}$$

Notice that condition 3) means that λ^{-1} is right C-colinear. If such a $\lambda \in Q$ exists, then we call (A, C, ψ, g) cleft.

Proposition 5.4 [14, Prop. 4.4] *Assume that (A, C, ψ, g) is a cleft entwining structure. Then the map τ in the associated Morita context is surjective.*

We say that the entwining structure (A, C, ψ, g) satisfies the *right normal basis property* if there exists a left B-linear and right C-colinear isomorphism $B \otimes C \to A$. The following is one of the main results in [14]. As before, we consider the functor $F = \bullet \otimes_B A:\ \mathcal{M}_B \to \mathcal{M}(\psi)_A^C$ and its right adjoint $G = (\bullet)^{\mathrm{co}C}$.

Theorem 5.5 [14, Theorem 4.5] *Let (A, C, ψ, g) be an entwining structure with a fixed grouplike element. The following assertions are equivalent:*

1. *(A, C, ψ, g) is cleft;*
2. *(F, G) is a category equivalence and (A, C, ψ, g) satisfies the right normal basis property;*

3. *(A, C, ψ, g) is Galois, and satisfies the right normal basis property;*
4. *the map* $^*\mathrm{can} : \#(C, A) \to \mathrm{End}_B(A)^{\mathrm{op}}$ *is bijective and (A, C, ψ, g) satisfies the right normal basis property.*

5.2 Hopf-Galois extensions. Let H be a Hopf algebra over a commutative ring k with bijective antipode, and A a right H-comodule algebra (cf. Example 1.5). Then $\mathcal{C} = A \otimes H$ is an A-coring, and $1_A \otimes 1_H \in G(\mathcal{C})$. Let $B = A^{\mathrm{co}H}$. The canonical map is now the following:

$$\mathrm{can} :\ A \otimes_B A \to A \otimes H,\ \mathrm{can}(a' \otimes_B a) = a' a_{[0]} \otimes a_{[1]}$$

Definition 5.6 (see e.g. [24, Def. 1.1]) A is a Hopf-Galois extension of B if and only if can is an isomorphism.

Obviously A is a Hopf-Galois extension of B if and only if $(A \otimes H, 1_A \otimes 1_H)$ is a Galois coring.

Assume now that H is a progenerator as a k-module, i.e. H is finitely generated, faithful, and projective as a k-module. Then $A \otimes H$ is a left A-progenerator, so we can apply the results of Section 4. We will show that we recover results from [17]. To this end, we will describe the Morita context associated to $(A \otimes H, 1_A \otimes 1_H)$. First we compute $^*\mathcal{C}$. We have already seen in Section 5.1 that $^*\mathcal{C} \cong \#(H, A)$. As a module, $\mathrm{Hom}(H, A) \cong H^* \otimes A$, since H is finitely generated and projective. The multiplication on $\#(H, A)$ can be transported into a multiplication on $H^* \otimes A$, giving us a k-algebra denoted by $H^*\#A$. A straightforward computation shows that this multiplication is given by the following formula. H^* is a coalgebra, since H is finitely generated projective, and H^* acts on A from the left: $h^* \rightharpoonup a = \langle h^*, a_{[1]} \rangle a_{[0]}$. Then we can compute that

$$(h^*\#a)(k^*\#b) = (k^*_{(1)} * h^*)\#(k^*_{(2)} \rightharpoonup a_{[0]})b. \tag{5.8}$$

Consider the map $\mathrm{can}' :\ A \otimes A \to A \otimes H$; its dual $^*\mathrm{can}' :\ H^*\#A \to \mathrm{End}(A)^{\mathrm{op}}$ is given by

$$^*\mathrm{can}'(h^*\#a)(b) = (h^* \rightharpoonup b)a. \tag{5.9}$$

Take $y = \sum_i h_i^*\#a_i \in H^*\#A$. $y \in Q$ if and only if

$$\sum_i \langle h_i^*, h_{(2)} \rangle a_{i[0]} \otimes h_{(1)} a_{i[1]} = \sum_i \langle h_i^*, h \rangle a_i \otimes 1,$$

for $h \in H$. Since H is finitely generated and projective, this is also equivalent to

$$\sum_i \langle h_i^*, h_{(2)} \rangle a_{i[0]} \langle h^*, h_{(1)} a_{i[1]} \rangle = \sum_i \langle h_i^*, h \rangle a_i \langle h^*, 1 \rangle,$$

for all $h \in H$ and $h^* \in H^*$. The left hand side equals

$$\sum_i \langle h^*_{(1)} * h_i^*, h \rangle \langle h^*_{(2)}, a_{i[1]} \rangle a_{i[0]},$$

so we find that $y \in Q$ if and only if

$$\sum_i (h^*_{(1)} * h_i^*)\#(h^*_{(2)} \rightharpoonup a_i) = \langle h^*, 1 \rangle \sum_i h_i^*\#a_i,$$

or

$$y(h^*\#1) = \langle h^*, 1 \rangle y.$$

for all $h^* \in H^*$. Thus

$$Q = \{y \in H^*\#A \mid y(h^*\#1) = \langle h^*, 1\rangle y, \text{ for all } h^* \in H^*\}. \tag{5.10}$$

Elementary computations show that the maps μ and τ from the Morita context are the following:

$$\tau: \ A \otimes_{H^*\#A} Q \to B, \ \tau(a \otimes y) = {}^*\mathrm{can}(y)(a);$$
$$\mu: \ Q \otimes_B A \to H^*\#A, \ \mu(y \otimes a) = y(\varepsilon_C\#a),$$

where $B = A^{\mathrm{co}H}$, as usual. Theorem 4.7 now takes the following form.

Proposition 5.7 *Let H be a k-progenerator Hopf algebra over a commutative ring k, and A a right H-comodule algebra. Then the following statements are equivalent (with notation as above):*

1. • $\mathrm{can}': \ A \otimes A \to A \otimes H, \ \mathrm{can}'(a' \otimes a) = a'a_{[0]} \otimes a_{[1]}$ *is bijective;*
   • *A is faithfully flat as a k-module.*
2. • ${}^*\mathrm{can}': \ H^*\#A \to \mathrm{End}(A)^{\mathrm{op}}, \ {}^*\mathrm{can}(h^*\#a)(b) = (h^* \rightharpoonup b)a$ *is an isomorphism;*
   • *A is a k-progenerator.*
3. • $A^{\mathrm{co}H} = k$;
   • *the Morita context $(k, H^*\#A, A, Q, \tau, \mu)$ is strict.*
4. • $A^{\mathrm{co}H} = k$;
   • *the adjoint pair of functors $(F = \bullet \otimes A, G = (\bullet)^{\mathrm{co}H})$ is an equivalence between the categories $\mathcal{M}_k$ and $\mathcal{M}_A^H$.*

[17, Theorems 9.3 and 9.6] follow from Proposition 5.7.

5.3 Classical Galois Theory. As in Example 1.3, let G be a finite group, and A a G-module algebra. We have seen that $\mathcal{C} = A \otimes (kG)^* = \oplus_{\sigma\in G} Av_\sigma$ is an A-coring. $\sum_\sigma v_\sigma$ is a grouplike element. Since $(kG)^*$ is finitely generated and projective, we can apply Proposition 5.7. We have

$$\mathrm{can}': \ A \otimes A \to \oplus_{\sigma\in G} Av_\sigma, \ \mathrm{can}'(a \otimes b) = \sum_\sigma a\sigma(b)v_\sigma.$$

$${}^*\mathcal{C} = \oplus_\sigma u_\sigma A,$$

with multiplication

$$(u_\sigma a)(u_\tau b) = u_{\tau\sigma}\tau(a)b,$$

and

$${}^*\mathrm{can}': \ \oplus_\sigma u_\sigma A \to \mathrm{End}(A)^{\mathrm{op}}, \ {}^*\mathrm{can}'(u_\sigma a)(b) = \sigma(b)a.$$

We also have

$$Q = \{\sum_\sigma u_\sigma\sigma(a) \mid a \in A\} \cong A,$$

which is not surprising since $(kG)^*$ is a Frobenius Hopf algebra (see [14] and [20]). If $A^G = k$, then we have a Morita context $(k, {}^*\mathcal{C}, A, A, \tau, \mu)$, where the connecting maps are the following:

$$\tau: \ A \otimes_{{}^*\mathcal{C}} A \to k, \ \tau(a \otimes b) = \sum_\sigma \sigma(ab);$$

$$\mu: \ A \otimes A \to {}^*\mathcal{C}, \ \mu(a \otimes b) = \sum_\sigma u_\sigma\sigma(b)a.$$

Proposition 5.7 now takes the following form (compare to [21, Prop. III.1.2]).

Proposition 5.8 *Let G be a finite group, k a commutative ring and A a G-module algebra. Then the following statements are equivalent:*

1. • can' *is an isomorphism;*
   • *A is faithfully flat as a k-module.*
2. • $^*\mathrm{can}'$ *is an isomorphism;*
   • *A is a k-progenerator.*
3. • $A^G = k$;
   • *the Morita context $(k, {}^*\mathcal{C}, A, A, \tau, \mu)$ is strict.*
4. • $A^G = k$;
   • *the adjoint pair of functors $(F = \bullet \otimes A, G = (\bullet)^G)$ is an equivalence between the categories of k-modules and right A-modules on which G acts as a group of right A-semilinear automorphisms.*

In the case where A is a commutative G-module algebra, we have some more equivalent conditions.

Proposition 5.9 *Let G be a finite group, k a commutative ring and A a commutative G-module algebra. Then the statements of Proposition 5.8 are equivalent to*

5. • $A^G = k$;
   • *for each non-zero idempotent $e \in A$ and $\sigma \neq \tau \in G$, there exists $a \in A$ such that $\sigma(a)e \neq \tau(a)e$;*
   • *A is a separable k-algebra (i.e. A is projective as an A-bimodule).*
6. • $A^G = k$;
   • *there exist $x_1, \cdots, x_n, y_1, \cdots y_n \in A$ with*

$$\sum_{j=1}^{n} x_j \sigma(y_j) = \delta_{\sigma,e}$$

   for all $\sigma \in G$.
7. • $A^G = k$;
   • *for each maximal ideal m of A, and for each $\sigma \neq e \in G$, there exists $x \in A$ such that $\sigma(x) - x \notin m$.*

Proof We refer to [16, Th 1.3] and [21, Prop. III.1.2]. □

If $A = l$ is a field, then the second part of condition 7. is satisfied. Let l be a finite field extension of a field k, and G the group of k-automorphisms of l. Then $l^G = k$ if and only if l is a normal and separable (in the classical sense) extension of k (see e.g. [37, Th. 10.8 and 10.10]). Thus we recover the classical definition of a Galois field extension.

5.4 Strongly graded rings. As in Example 1.4, let G be a group, and A a G-graded ring, and $\mathcal{C} = \oplus_{\sigma\in G} Au_\sigma$. Fix $\lambda \in G$, and take the grouplike element $u_\lambda \in G(\mathcal{C})$. Then $M^{\mathrm{co}\mathcal{C}} = M_\lambda$, for any right G-graded A-module, and $B = A^{\mathrm{co}\mathcal{C}} = A_e$. Since B is a direct factor of A, A is flat as a left and right B-module, and $i : B \to A$ is pure in $\mathcal{M}_B$ and ${}_B\mathcal{M}$. Also

$$\mathrm{can} :\ A \otimes_B A \to \oplus_{\sigma\in G} Au_\sigma,\ \ \mathrm{can}(a' \otimes a) = \sum_{\sigma\in G} a' a_\sigma u_{\lambda\sigma}.$$

Proposition 5.10 *With notation as above, the following assertions are equivalent.*

1. *A is strongly G-graded, that is,* $A_\sigma A_\tau = A_{\sigma\tau}$*, for all* $\sigma, \tau \in G$*;*
2. *the pair of adjoint functors* $(F = \bullet \otimes_B A, G = (\bullet)_\lambda)$ *is an equivalence between* $\mathcal{M}_B$ *and* $\mathcal{M}_A^G$*, the category of G-graded right A-modules;*
3. $(\mathcal{C}, u_\lambda)$ *is a Galois coring.*

In this case A is faithfully flat as a left (or right) B-module.

Proof 1) $\Rightarrow$ 2) is a well-known fact from graded ring theory. We sketch a proof for completeness sake. The unit of the adjunction between $\mathcal{M}_B$ and $\mathcal{M}_A^G$ is given by

$$\eta_N : N \to (N \otimes_B A)_\lambda, \ \eta_N(n) = n \otimes_B 1_A.$$

η_N is always bijective, even if A is not strongly graded. Let us show that the counit maps $\zeta_M : M_\lambda \otimes_B A \to M$, $\zeta_M(m \otimes_B a) = ma$ are surjective. For each $\sigma \in G$, we can find $a_i \in A_{\sigma^{-1}}$ and $a'_i \in A_\sigma$ such that $\sum_i a_i a'_i = 1$. Take $m \in M_\tau$ and put $\sigma = \lambda_{-1}\tau$. Then $m = \zeta_M(\sum_i m a_i \otimes_B a'_i)$, and ζ_M is surjective.
If $m_j \in M_\lambda$ and $c_j \in A$ are such that $\sum_j m_j c_j = 0$, then for each $\sigma \in G$, we have

$$\sum_j m_j \otimes c_{j\sigma} = \sum_{i,j} m_j \otimes c_{j\sigma} a_i a'_i = \sum_{i,j} m_j c_{j\sigma} a_i \otimes a'_i = 0.$$

hence $\sum_j m_j \otimes c_j = \sum_{\sigma \in G} = \sum_j m_j \otimes c_{j\sigma} = 0$, so ζ_M is also injective.
2) $\Rightarrow$ 3) follows from Proposition 3.1.
3) $\Rightarrow$ 1) follows from Theorem 3.5 and the fact that $i : B \to A$ is pure in $\mathcal{M}_B$ and ${}_B\mathcal{M}$.

The final statement follows from Proposition 2.5 and the fact that A is flat as a B-module. □

Notice that, in this situation, the fact that $(\mathcal{C}, u_\lambda)$ is Galois is independent of the choice of λ.

6 A more general approach: comatrix corings

Let $\mathcal{C}$ be an A-coring. We needed a grouplike $x \in \mathcal{C}$ in order to make A into a right $\mathcal{C}$-comodule. In [25], the following idea is investigated. A couple $(\mathcal{C}, \Sigma)$, consisting of a coring $\mathcal{C}$ and a right $\mathcal{C}$-comodule Σ which is finitely generated and projective as a right A-module, will be called a coring with a fixed finite comodule. Let $T = \mathrm{End}^{\mathcal{C}}(\Sigma)$. Then we have a pair of adjoint functors

$$F = \bullet \otimes_T \Sigma : \mathcal{M}_T \to \mathcal{M}^{\mathcal{C}} \ ; \ G = \mathrm{Hom}^{\mathcal{C}}(\Sigma, \bullet) : \mathcal{M}^{\mathcal{C}} \to \mathcal{M}_T,$$

with unit ν and counit ζ given by

$$\nu_N : N \to \mathrm{Hom}^{\mathcal{C}}(\Sigma, N \otimes_T \Sigma), \ \nu_N(n)(u) = n \otimes_T u;$$

$$\zeta_M : \mathrm{Hom}^{\mathcal{C}}(\Sigma, M) \otimes_T \Sigma \to M, \ \zeta_M(f \otimes_T u) = f(u).$$

In the situation where $\Sigma = A$, we recover the adjoint pair discussed at the end of Section 1. A particular example is the *comatrix coring*, generalizing the *canonical coring*. Let A and B be rings, and $\Sigma \in {}_B\mathcal{M}_A$, with Σ finitely generated and projective as a right A-module. Let

$$\{(e_i, e_i^*) \mid i = 1, \cdots, n\} \subset \Sigma \times \Sigma^*$$

be a finite dual basis of Σ as a right A-module. $\mathcal{D} = \Sigma^* \otimes_B \Sigma$ is an (A, A)-bimodule, and an A-coring, via

$$\Delta_{\mathcal{D}}(\varphi \otimes_B u) = \sum_i \varphi \otimes_B e_i \otimes_A e_i^* \otimes_B u \text{ and } \varepsilon_{\mathcal{D}}(\varphi \otimes_B u) = \varphi(u).$$

Furthermore $\Sigma \in \mathcal{M}^{\mathcal{D}}$ and $\Sigma^* \in {}^{\mathcal{D}}\mathcal{M}$. The coactions are given by

$$\rho^r(u) = \sum_i e_i \otimes_A e_i^* \otimes_B u \ ; \ \rho^l(\varphi) = \sum_i \varphi \otimes_B e_i \otimes_A e_i^*.$$

We also have that ${}^*\mathcal{D} \cong {}_B\mathrm{End}(\Sigma)^{\mathrm{op}}$. El Kaoutit and Gómez Torrecillas proved the following generalization of the Faithfully Flat Descent Theorem.

Theorem 6.1 *Let $\Sigma \in {}_B\mathcal{M}_A$ be finitely generated and projective as a right A-module, and $\mathcal{D} = \Sigma^* \otimes \Sigma$. Then the following are equivalent*

1. *Σ is faithfully flat as a left B-module;*
2. *Σ is flat as a left B-module and $(\bullet \otimes_B \Sigma, \mathrm{Hom}^{\mathcal{D}}(\Sigma, \bullet))$ is a category equivalence between $\mathcal{M}_B$ and $\mathcal{M}^{\mathcal{D}}$.*

Let $(\mathcal{C}, \Sigma)$ be a coring with a fixed finite comodule, and $T = \mathrm{End}^{\mathcal{C}}(\Sigma)$. We have an isomorphism $f : \Sigma^* \to \mathrm{Hom}^{\mathcal{C}}(\Sigma, \mathcal{C})$ given by

$$f(\varphi) = (\varphi \otimes_A I_{\mathcal{C}}) \circ \rho \text{ and } f^{-1}(\phi) = \varepsilon_{\mathcal{C}} \circ \phi,$$

for all $\varphi \in \Sigma^*$ and $\phi \in \mathrm{Hom}^{\mathcal{C}}(\Sigma, \mathcal{C})$. Consider the map

$$\mathrm{can} = \zeta_{\mathcal{C}} \circ (f \otimes_B I_{\Sigma}) : \ \mathcal{D} = \Sigma^* \otimes_T \Sigma \to \mathrm{Hom}^{\mathcal{C}}(\Sigma, \mathcal{C}) \otimes_T \Sigma \to \mathcal{C}.$$

We compute easily that $\mathrm{can}(\varphi \otimes_B u) = \varphi(u_{[0]})u_{[1]}$. can is a morphism of corings, and can is an isomorphism if and only if $\zeta_{\mathcal{C}}$ is an isomorphism.

Definition 6.2 [25, 3.4] Let $(\mathcal{C}, \Sigma)$ be a coring with a fixed finite comodule, and $T = \mathrm{End}^{\mathcal{C}}(\Sigma)$. $(\mathcal{C}, \Sigma)$ is termed Galois if can : $\Sigma^* \otimes_T \Sigma \to \mathcal{C}$ is an isomorphism.

Theorem 6.3 [25, 3.5] *If $(\mathcal{C}, \Sigma)$ is Galois, and Σ is faithfully flat as a left T-module, then (F, G) is an equivalence of categories.*

For further results, we refer to [25].

ACKNOWLEDGEMENTS

The author thanks George Janelidze for stimulating discussions about descent theory, and the referee for pointing out that the proper assumption in Proposition 2.3 is right purity of i, instead of left purity.

References

[1] Abuhlail J. *Morita contexts for corings and equivalences*, in "Hopf algebras in non-commutative geometry and physics", Caenepeel S. and Van Oystaeyen, F. (eds.), Lecture Notes Pure Appl. Math., Dekker, New York, to appear.
[2] Abuhlail, J. *Rational modules for corings*, Comm. Algebra **31** (2003), 5793–5840.
[3] Auslander, M. and Goldman, O. *The Brauer group of a commutative ring.* Trans. Amer. Math. Soc. **97** (1960), 367–409.
[4] Bass, H. "Algebraic K-theory", Benjamin, New York, 1968.
[5] Beattie, M., Dăscălescu, S. and Raianu, Ş. *Galois extensions for co-Frobenius Hopf algebras*, J. Algebra **198** (1997), 164–183.
[6] Böhm, G. *Doi-Hopf modules over weak Hopf algebras*, Comm. Algebra **28** (2000), 4687–4698.
[7] Borceux, F. Handbook of categorical algebra 2, Encyclopedia Math. Appl. **51**, Cambridge University Press, Cambridge, 1994.

[8] Brzeziński, T. *Coalgebra-Galois extensions from the extension point of view*, in "Hopf algebras and quantum groups", Caenepeel S. and Van Oystaeyen, F. (eds.), Lecture Notes in Pure and Appl. Math. **209**, Marcel Dekker, New York, 2000.

[9] T. Brzeziński, *The structure of corings. Induction functors, Maschke-type theorem, and Frobenius and Galois properties*, Algebr. Representat. Theory **5** (2002), 389–410.

[10] Brzeziński, T. and Hajac, P.M. *Coalgebra extensions and algebra coextensions of Galois type*, Comm. Algebra **27** (1999), 1347-1367.

[11] Brzeziński, T. and Majid, S. *Coalgebra bundles*, Comm. Math. Phys. **191** (1998), 467–492.

[12] Caenepeel, S. and De Groot, E. *Modules over weak entwining structures*, Contemp. Math. **267** (2000), 31–54.

[13] Caenepeel, S., Militaru, G. and Zhu, Shenglin "Frobenius and separable functors for generalized module categories and nonlinear equations", Lecture Notes in Math. **1787**, Springer Verlag, Berlin, 2002.

[14] Caenepeel, S., Vercruysse, J. and Wang, Shuanhong *Morita Theory for corings and cleft entwining structures*, J. Algebra, to appear.

[15] Caenepeel, S., Vercruysse, J. and Wang, Shuanhong *Rationality properties for Morita contexts associated to corings*, preprint.

[16] Chase, S., Harrison, D. and Rosenberg, A. *Galois theory and Galois cohomology of commutative rings*, Mem. Amer. Math. Soc. **52** (1965), 1–19.

[17] Chase, S. and Sweedler, M. E. "Hopf algebras and Galois theory", Lect. Notes in Math. **97**, Springer Verlag, Berlin, 1969.

[18] Cipolla, M. *Discesa fedelmente piatta dei moduli*, Rendiconti del Circolo Matematico di Palermo, Serie II **25** (1976).

[19] Cohen, M. and Fischman, D. *Semisimple extensions and elements of trace 1*, J. Algebra **149** (1992), 419–437.

[20] Cohen, M., Fischman, D. and Montgomery, S. *Hopf Galois extensions, smash products, and Morita equivalence*, J. Algebra **133** (1990), 351–372.

[21] DeMeyer F. and Ingraham, E. "Separable algebras over commutative rings", Lecture Notes in Math. **181**, Springer Verlag, Berlin, 1971.

[22] Doi, Y. *Unifying Hopf modules*, J. Algebra **153** (1992), 373–385.

[23] Doi, Y. *Generalized smash products and Morita contexts for arbitrary Hopf algebras*, in "Advances in Hopf algebras", Bergen, J. and Montgomery, S. (eds.), Lect. Notes Pure Appl. Math. **158**, Dekker, New York, 1994.

[24] Doi, Y. and Takeuchi, M. *Hopf-Galois extensions of algebras, the Miyashita-Ulbrich action and Azumaya algebras*, J. Algebra **121** (1989), 488–516.

[25] El Kaoutit, L. and Gómez Torrecillas, J. *Comatrix corings: Galois corings, descent theory, and a structure Theorem for cosemisimple corings*, Math. Z., **244** (2003), 887–906.

[26] El Kaoutit, L., Gómez Torrecillas, J. and Lobillo, F. J. *Semisimple corings*, preprint 2001.

[27] Gómez Torrecillas, J. *Separable functors in corings*, Int. J. Math. Math. Sci. **30** (2002), 203–225.

[28] Grothendieck, A. *Technique de Descente I*, Sém. Bourbaki, exp. **190** (1959-1960).

[29] Knus, M. A. and Ojanguren, M. "Théorie de la descente et algébres d'Azumaya", Lecture Notes in Math. **389**, Springer Verlag, Berlin, 1974.

[30] Koppinen, M. *Variations on the smash product with applications to group-graded rings*, J. Pure Appl. Algebra **104** (1995), 61–80.

[31] Mac Lane, S. "Categories for the working mathematician", second edition, Graduate Texts in Mathematics **5**, Springer Verlag, Berlin, 1997.

[32] Mesablishvili, B. *Pure morphisms of commutative rings are effective descent morphisms for modules - a new proof*, Theory Appl. Categories **7** (2000), 38–42.

[33] Nuss, P. *Noncommutative descent and nonabelian cohomology*, K-Theory, **12** (1997), 23–74.

[34] Schauenburg, P. *Doi-Koppinen modules versus entwined modules*, New York J. Math., **6** (2000), 325–329.

[35] Schneider, H.J. *Principal homogeneous spaces for arbitrary Hopf algebras*, Israel J. Math. **70** (1990), 167–195.

[36] Serre, J.P. "Cohomologie Galoisienne", Lect. Notes in Math. **5**, Springer Verlag, Berlin, 1965.

[37] Stewart, I. "Galois theory", second edition, Chapman and Hall, London, 1989.

[38] Sweedler, M. E. *The predual Theorem to the Jacobson-Bourbaki Theorem*, Trans. Amer. Math. Soc. **213** (1975), 391–406.

[39] Takeuchi, M., as referred to in MR 2000c 16047, by A. Masuoka.
[40] Wisbauer, R. *On the category of comodules over corings*, in "Mathematics and mathematics education (Bethlehem, 2000)", World Sci. Publishing, River Edge, NJ, 2002, 325–336.
[41] Wisbauer, R. *On Galois corings*, in "Hopf algebras in non-commutative geometry and physics", Caenepeel S. and Van Oystaeyen, F. (eds.), Lecture Notes Pure Appl. Math., Dekker, New York, to appear.

Received January 27, 2003; in revised form June 10, 2003

Fields Institute Communications
Volume 43, 2004

Quantum Categories, Star Autonomy, and Quantum Groupoids

Brian Day
Centre of Australian Category Theory
Macquarie University
N.S.W. 2109, Australia

Ross Street
Centre of Australian Category Theory
Macquarie University
N.S.W. 2109, Australia
street@math.mq.edu.au

Abstract. A useful general concept of bialgebroid seems to be resolving itself in recent publications; we give a treatment in terms of modules and enriched categories. Generalizing this concept, we define the term "quantum category" in a braided monoidal category with equalizers distributed over by tensoring with an object. The definition of antipode for a bialgebroid is less resolved in the literature. Our suggestion is that the kind of dualization occurring in Barr's star-autonomous categories is more suitable than autonomy (= compactness = rigidity). This leads to our definition of quantum groupoid intended as a "Hopf algebra with several objects".

1 Introduction

This paper has several purposes. We wish to introduce the concept of quantum category. We also wish to generalize the theory of $*$-autonomous categories in the sense of [1]. The connection between these two concepts is that they lead to our notion of quantum groupoid.

It was shown in [32] that $*$-autonomous categories provide models of the linear logic described in [19]. This suggests an interesting possibility of interactions between computer science and quantum group theory. Perhaps it will be possible, in future papers, to exploit the dichotomy between categories as structures and categories of structures. For example, what is the quantum category of finite sets, or the quantum category of finite dimensional vector spaces?

It is well known that ordinary categories are not models of an ordinary algebraic (Lawvere) theory; rather, they are models of a finite-limit theory, requiring

2000 *Mathematics Subject Classification.* Primary 81R50, 18D10; Secondary 03F52, 20L05.

operations to be defined in stages since some of them are defined on finite limits of earlier operations. Quantum categories, in a braided monoidal category with equalizers distributed over by tensoring, similarly involve operations defined on objects created by tensoring and taking equalizers of previously defined objects and operations.

The section headings are as follows:

1. Introduction
2. Ordinary categories revisited
3. Takeuchi bialgebroids
4. The lax monoidal operation $\times_R$
5. Monoidal star autonomy
6. Modules and promonoidal enriched categories
7. Forms and promonoidal star autonomy
8. The star and Chu constructions
9. Star autonomy in monoidal bicategories
10. Ordinary groupoids revisited
11. Hopf bialgebroids
12. Quantum categories and quantum groupoids

Before looking at quantum categories we will develop, in this introduction, a definition of "category" which suggests the definition of "quantum category". We will then relate this definition to the literature.

We use the terminology of Eilenberg-Kelly [18] for monoidal categories and monoidal functors; so we use the adjective "*strong* monoidal" for a functor which preserves tensor and unit up to coherent natural isomorphisms. A comonoidal category would have, instead of a tensor product, a tensor coproduct $\mathcal{A} \to \mathcal{A}\times\mathcal{A}$ and a counit with appropriately coherent constraints; this concept is not so interesting for ordinary categories but becomes more so for enriched categories. Comonoidal functors would go between comonoidal categories. So, for monoidal categories $\mathcal{A}$ and $\mathcal{X}$, like [27], we use the term *opmonoidal functor* for a functor

$$F : \mathcal{A} \to \mathcal{X}$$

equipped with a natural family of morphisms $\delta_{A,B} : F(A \otimes B) \to FA \otimes FB$ and a morphism $\varepsilon : FI \to I$ that are coherent.

For any set X, consider the monoidal category $\mathrm{Set}/X \times X$ of sets over $X \times X$ with the tensor product defined by

$$\left(A \xrightarrow{(s,t)} X \times X \right) \otimes \left(B \xrightarrow{(u,v)} X \times X \right) = \left(P \xrightarrow{(sop,voq)} X \times X \right)$$

where P is the pullback of $t : A \to X$ and $u : B \to X$ with projections $p : P \to A$ and $q : P \to B$. The objects of $\mathrm{Set}/X \times X$ are directed graphs with vertex-set X and the monoids are the categories with object-set X; this is well known (see [25]) and easy. Less well known, but also easy, is the fact that category structures on the graph $A \xrightarrow{(s,t)} X \times X$ amount to monoidal structures on the category Set/A of sets over A together with a strong monoidal structure on the functor

$$\Sigma_{(s,t)} : \mathrm{Set}/A \to \mathrm{Set}/X \times X$$

defined on objects by composing the function into A with (s,t).

To see this, notice that every object of a slice category Set$/C$ is a coproduct of elements $c : 1 \to C$ of C (here 1 is a chosen set with precisely one element) so that any tensor product on Set$/C$, which preserves coproducts in each variable, will be determined by its value on elements (which may not be another element in general). The tensor product on Set$/X \times X$ is such, and its value on elements is given by $(x, y) \otimes (u, v) = (x, v)$ when $y = u$ (which is in fact another element) but is the unique function $\emptyset \to X \times X$ when $y \neq u$. Since $\Sigma_{(x,t)}$ is conservative and coproduct preserving, and is to be strong monoidal, the tensor product on Set$/A$ preserves coproducts in each variable. An object of Set$/A$ has the same source set as its value under $\Sigma_{(x,t)}$. So, for elements a and b of A, the tensor product $a \otimes b$ is an element of A if and only if $t(a) = s(b)$; in this case, $s(a \otimes b) = s(a)$ and $t(a \otimes b) = t(b)$; otherwise, $a \otimes b$ is the unique function $\emptyset \to A$. The unit for the monoidal category Set$/X \times X$ is the diagonal $X \to X \times X$, so the unit for Set$/A$ has the form $i : X \to A$ with $s(i(x)) = t(i(x)) = x$ for all $x \in X$. Thus we have a reflexive graph X, A, s, t, i with a composition operation. We leave the reader to check that the associativity and unity constraints of the monoidal category Set$/A$ give associativity for the composition and that each $i(x)$ is an identity.

Conversely, suppose we have a category $\mathbf{A}$ with underlying graph $A \xrightarrow{(s,t)} X \times X$. Notice that Set$/A$ is canonically equivalent to the category $[A,\text{Set}]$ of functors from the discrete category A to Set. We can define a promonoidal structure (in the sense of [9]) on A by

$$P(a,b;c) = \begin{cases} 1 & \text{when } c \text{ is the composite of } a \text{ and } b; \\ \emptyset & \text{otherwise;} \end{cases} \qquad \text{and}$$

$$J(a) = \begin{cases} 1 & \text{when } a \text{ is an identity;} \\ \emptyset & \text{otherwise.} \end{cases}$$

Then $[A,\text{Set}]$ becomes a monoidal category under convolution; this transports to a monoidal structure on Set$/A$ for which $\Sigma_{(s,t)}$ is strong monoidal.

When our category $\mathbf{A}$ is actually a groupoid (that is, every arrow is invertible), there is a bijection $S : A \to A$ defined by $Sa = a^{-1}$. We draw attention to the isomorphisms

$$P(a, b; Sc) \cong P(b, c; Sa)$$

noting here that S is its own inverse and that the diagram

$$\begin{tikzcd} A \arrow[r, "S"] \arrow[d, "(s,t)"'] & A \arrow[d, "(s,t)"] \\ X \times X \arrow[r, "S"'] & X \times X \end{tikzcd}$$

commutes, where the lower S is the switch map — which is inversion for X as a chaotic category (meaning, the category whose object set is X and each homset has exactly one element). We will relate this kind of "antipode" structure to $*$-autonomy.

Now suppose we have a category $\mathbf{A} : A \xrightarrow{(s,t)} X \times X$ and suppose we regard Set$/A$ as monoidal in the manner described above. The functor $\Sigma_{(s,t)}$ has a right adjoint $(s,t)^*$ defined by pulling back along (s,t). The strong monoidal structure on $\Sigma_{(s,t)}$ is obviously both monoidal and opmonoidal; the opmonoidal structure transforms to a monoidal structure on the right adjoint $(s,t)^*$ in such a way that the unit and counit for the adjunction are monoidal natural transformations. The composite of monoidal functors is monoidal; so the endofunctor $G_A = \Sigma_{(s,t)}(s,t)^*$ is also monoidal. The adjunction also generates a comonad structure on G_A in such a way that the counit and comultiplication are monoidal natural transformations; we have a monoidal comonad G_A on $\mathrm{Set}/X \times X$. Remember the term "monoidal comonad"!

It is also important to notice that $(s,t)^*$ has a right adjoint $\Pi_{(s,t)}$; so the endofunctor G_A has a right adjoint $(s,t)^*\Pi_{(s,t)}$. By Beck's Theorem (see [25] for example), $\Sigma_{(s,t)}$ is comonadic since it is obviously conservative (that is, reflects isomorphisms) and preserves equalizers. On the other hand, any monoidal comonad on a monoidal category leads to a monoidal structure on the category of Eilenberg-Moore coalgebras in such a way that the forgetful functor is strong monoidal (see [28] or [27] for example). Any cocontinuous endofunctor of $\mathrm{Set}/X \times X$ has the form $\Sigma_{(s,t)}(s,t)^*$ for some graph $A \xrightarrow{(s,t)} X \times X$. Assembling all this, we obtain:

Proposition 1.1 *Categories with underlying graph $A \xrightarrow{(s,t)} X \times X$ are in bijection with monoidal comonad structures on the endofunctor $\Sigma_{(s,t)}(s,t)^*$ of $\mathrm{Set}/X \times X$.*

Let us compare the combinatorial context of Proposition 1.1 with the linear algebra context. Szlachányi [37] has shown that, for a k-algebra R, the $\times_R$-bialgebras of Takeuchi [38] are opmonoidal monads on the monoidal category $\mathrm{Vect}_k^{R\otimes R^\circ}$ of left R-, right R-bimodules over R where the underlying endofunctor of the monad is a left adjoint. These $\times_R$-bialgebras of Takeuchi have been convincingly proposed (see [41, 24, 31]) as the good concept of "bialgebroid" based on R (that is, with "object of objects R").

Here we face the usual dilemma. Given a k-bialgebra H, is it better to consider the category of modules for the underlying algebra with the monoidal structure coming from the comultiplication, or, the category of comodules for the underlying coalgebra with the monoidal structure coming from the multiplication? Our preference is definitely the latter since the obvious linearization of the group case leads to this decision; also see [21]. When H is finite dimensional (as a vector space over a field k) there is essentially no difference. We feel that the functor from the category of sets to the category of k-vector spaces should provide the mechanism for regarding classical categories as quantum categories. For this we need to dualize the $\times_R$-bialgebras of Takeuchi to be based on a k-coalgebra C rather than a k-algebra R; indeed, Brzeziński-Militaru [8] have already made this dualization of the $\times_R$-bialgebras of Takeuchi based on a k-coalgebra C rather than a k-algebra R. We take this as our concept of quantum category; it involves a monoidal comonad. Actually, our general setting of a monoidal bicategory formalizes this duality.

The basic examples of quantum groups are Hopf algebras with braidings (also called quasitriangular elements or R-matrices) or cobraidings, depending how the dilemma is resolved. Indeed, these basic quantum groups are cotortile bialgebras (see [21]). We leave it to a future paper to define and discuss braidings and twists on quantum categories.

So what is a quantum groupoid? It should be a quantum category with an "antipode". We first develop a notion of antipode for the $\times_R$-bialgebras of Takeuchi. We are influenced by the chaotic example $R^\circ \otimes R$ itself where we believe the antipode should be the switch isomorphism $(R^\circ \otimes R)^\circ \to R^\circ \otimes R$. This is not a dualization in the sense of [14] but a dualization of the kind that arises in Barr's $*$-autonomous monoidal categories [1].

Consequently we are led to study $*$-autonomy for enriched categories. In fact, we define $*$-autonomous promonoidal $\mathcal{V}$-categories and show this notion is preserved under convolution. There is always the canonical promonoidal structure on $\mathcal{A}^{\mathrm{op}} \otimes \mathcal{A}$. (see the concluding remarks of [9]) which is $*$-autonomous (as remarked by Luigi Santocanale after the talk [12]) and leads under convolution to the tensor product of bimodules. The Chu construction as described in [3] and [36] is purely for ordinary categories: it needs the repetition and deletion of variables that are available in a cartesian closed base category such as Set. We vastly extend the notion of $*$-autonomy to include enriched categories and other contexts. We provide a general star-construction which leads to the Chu construction as a special case.

Equipped with this we can define when a Takeuchi $\times_R$-bialgebra is "Hopf". Then, by dualizing from k-algebras to k-coalgebras, we define quantum groupoids to be $*$-autonomous quantum categories.

2 Ordinary categories revisited

Let us consider Proposition 1.1 from a slightly different viewpoint. A left adjoint (or cocontinuous) functor F: $\mathrm{Set}/X \to \mathrm{Set}/Y$ between slice categories is determined by its restriction to the elements $x : 1 \to X$ of X, and so, by a functor

$$X \to \mathrm{Set}/Y \xrightarrow{\sim} [Y, \mathrm{Set}],$$

where we regard the sets X and Y as discrete categories and write $[\mathcal{A}, \mathcal{B}]$ for the category of functors and natural transformations from $\mathcal{A}$ to $\mathcal{B}$. However, the functors $X \to [Y, \mathrm{Set}]$ are in bijection with functors $S : X \times Y \to \mathrm{Set}$ which we think of as matrices

$$S = (S(x; y))_{(x,y) \in X \times Y} \,.$$

This gives us an equivalent (actually "biequivalent") way of looking at the 2-category whose objects are (small) sets, whose morphisms $F : X \to Y$ are cocontinuous functors $\mathrm{Set}/X \to \mathrm{Set}/Y$, and whose 2-cells are natural transformations; however, rather than a 2-category we only have a bicategory which we call Mat(Set) (compare [6] for example). Again, the objects are sets, the morphisms $S : X \to Y$ are matrices, and the 2-cells $\theta : S \Rightarrow T$ are matrices of functions

$$\theta = (\theta(x; y) : S(x; y) \to T(x; y))_{(x,y) \in X \times Y} \,;$$

vertical composition of 2-cells is defined by entrywise composition of functions, horizontal composition of morphisms $S : X \to Y$ and $T : Y \to Z$ is defined by matrix multiplication

$$(T \circ S)(x; z) = \sum_{y \in Y} S(x; y) \times T(y; z),$$

and horizontal composition is extended in the obvious way to 2-cells. We write X: $X \to X$ for the identity matrix (or Kronecker delta):

$$X(x; y) = \begin{cases} 1 & \text{for } x = y, \\ \emptyset & \text{otherwise.} \end{cases}$$

Of course Mat(Set) is also biequivalent to the bicategory Span(Set) of spans (in the sense of Bénabou [5]) in the category Set of sets.

In fact, Mat(Set) is an autonomous monoidal bicategory in the sense of the authors [15]. That is, there is a reasonably well behaved tensor product pseudo-functor

$$\mathrm{Mat(Set)} \times \mathrm{Mat(Set)} \to \mathrm{Mat(Set)}$$

which is simply defined on objects by cartesian product of sets and likewise, by cartesian product entrywise, on morphisms and 2-cells. Each object Y is actually self-dual since a matrix $X \times Y \to Z$ can be identified with a matrix $X \to Y \times Z$. This means that $Y \times Z$ is the internal hom in Mat(Set) of Y and Z (mimicking the fact that in finite-dimensional vector spaces the vector space of linear functions from V to W is isomorphic to $V^* \otimes W$). In particular, $X \times X$ is the internal endohom of X; and so we expect it to be a pseudomonoid in Mat(Set) (mimicking the fact that the internal endohom of an object in a monoidal category is an internal monoid).

Let us be more specific about this pseudomonoid structure on $X \times X$ in Mat(Set). The multiplication

$$P : (X \times X) \times (X \times X) \to X \times X$$

is defined by $P(y_2, x_2, y_1, x_1; x, y) = X(y; x_1) \times X(y_1; x_2) \times X(y_2; x)$. The unit $J : 1 \to X \times X$ is defined by $J(\bullet; x, y) = X(x; y)$. One easily checks the canonical associativity and unit isomorphisms

$$P \circ (P \times (X \times X)) \cong P \circ ((X \times X) \times P)$$
$$P \circ (J \times (X \times X)) \cong X \times X \cong P \circ ((X \times X) \times J).$$

Thinking of the set $X \times X$ as a discrete category, we see that P, J and these isomorphisms form a promonoidal structure on $X \times X$. Noting that, under the equivalence of categories

$$[X \times X, \mathrm{Set}] \xrightarrow{\sim} \mathrm{Set}/X \times X,$$

the convolution monoidal structure for $X \times X$ transports across the equivalence to the monoidal structure on Set$/X \times X$ described in the Introduction, the following result becomes a corollary of Proposition 1.1.

Proposition 2.1 *Categories with object set X are equivalent to monoidal comonads on the internal endohom pseudomonoid $X \times X$ in the monoidal bicategory Mat(Set).*

It may be instructive to sketch a direct proof of this result. A monoidal comonad G on $X \times X$ comes equipped with 2-cells

$$\delta : G \to G \circ G,\ \varepsilon : G \to X \times X,\ \mu : P \circ (G \times G) \to G \circ P \ \text{ and } \ \eta : J \to G \circ J,$$

subject to appropriate axioms. The mere existence of ε is quite a strong condition since $X(x; u) \times X(y; v)$ is empty unless $x = u$ and $y = v$; so $G(x, y; u, v)$ is empty unless $x = u$ and $y = v$. This leads us to put

$$A(x, y) = G(x, y; x, y)$$

which defines the homsets of our category A. It is then easy to check that μ defines composition and η provides the identities for the category A. We note finally that δ is forced to be a genuine diagonal morphism: we are dealing here with the categories of "commutative geometry".

3 Takeuchi bialgebroids

We are now ready to move from set theory to linear algebra. Let k be any commutative ring and write $\mathcal{V}$ for the monoidal category of k-modules; we write $\otimes$ for the tensor product of k-modules. Monoids R in $\mathcal{V}$ will be called k-algebras and we write $\mathcal{V}^R$ for the category of left R-modules; we can think of R as a one-object $\mathcal{V}$-category [18] so that $\mathcal{V}^R$ is the category of $\mathcal{V}$-functors from R to $\mathcal{V}$. From this viewpoint the k-algebra R^0, which is just R with opposite multiplication, is the opposite $\mathcal{V}$-category of R.

We briefly recall the preliminaries of Morita theory starting with Watts' Theorem [39] characterizing cocontinuous functors between categories of modules. For k-algebras R and S, a left adjoint (or cocontinuous) functor $F : \mathcal{V}^R \to \mathcal{V}^S$ between module categories is, up to isomorphism, determined by its restriction to the $\mathcal{V}$-dense (see [13]) full subcategory of $\mathcal{V}^R$ consisting of R itself as a left R-module. This full subcategory is isomorphic to R^0. So the left S-module $F(R) = M$ is also a right R-, left S-bimodule which we call a *module* from R to S and use the arrow notation $M : R \to S$. (The fact that R is actually on the left of the arrow and S on the right, rather than the other way around, has to do with our convention to compose functions in the usual order.) We also identify M with an object of $\mathcal{V}^{R^0 \otimes S}$.

There is a 2-category whose objects are k-algebras, whose morphisms $R \to S$ are left adjoint functors $F : \mathcal{V}^R \to \mathcal{V}^S$, and whose 2-cells are natural transformations between such functors F; the compositions are the usual ones for functors and natural transformations. This 2-category is biequivalent to the bicategory $\mathrm{Mod}(\mathcal{V})$ whose objects are k-algebras, whose morphisms are modules $M : R \to S$, and whose 2-cells are 2-sided module morphisms; the horizontal composite $N \circ M : R \to T$ of $M : R \to S$ and $N : S \to T$ is the tensor product $N \otimes_S M$ of the modules M and N over S ; vertical composition of 2-cells is the usual composition of module morphisms.

Indeed, like Mat(Set), the bicategory $\mathrm{Mod}(\mathcal{V})$ is autonomous monoidal. The tensor product is that of $\mathcal{V}$: k-algebras R and S are taken to the k-algebra $R \otimes S$, modules $M : R \to S$ and $M' : R' \to S'$ are taken to the module $M \otimes M' : R \otimes R' \to S \otimes S'$, and module morphisms are tensored using the functoriality of $M \otimes M'$ in the two variables. The opposite k-algebra S^0 acts as a dual for S since the category of modules $R \otimes X \to T$ is equivalent to the category of modules $R \to S^0 \otimes T$.

It follows that $R^0 \otimes R$ is an internal endohom for R and, as such, is a pseudomonoid in $\mathrm{Mod}(\mathcal{V})$. The multiplication

$$P : (R^0 \otimes R) \otimes (R^0 \otimes R) \to R^0 \otimes R$$

is $P = R \otimes R \otimes R$ as a k-module, with the further actions defined by

$$(x \otimes y)(a \otimes b \otimes c)(y_1 \otimes x_1 \otimes y_2 \otimes x_2) = (yax_1) \otimes (y_1 b x_2) \otimes (y_2 c x)$$

for $a \otimes b \otimes c \in P$, $x \otimes y \in R^0 \otimes R$ and $x_1 \otimes y_1 \otimes x_2 \otimes y_2 \in R^0 \otimes R \otimes R^0 \otimes R$. The unit

$$J : k \to R^0 \otimes R$$

is just $J = R$ as *a* k-module, with the further action $(x \otimes y)a = yax$. One easily checks that there are canonical isomorphisms

$$\begin{aligned} P \otimes_{R^e \otimes R^e} (R^e \otimes P) &\cong P \otimes_{R^e \otimes R^e} (P \otimes R^e) \qquad \text{and} \\ P \otimes_{R^e \otimes R^e} (R^e \otimes J) &\cong R^e \cong P \otimes_{R^e \otimes R^e} (J \otimes R^e) \end{aligned}$$

where we have used the traditional notation $R^e = R^0 \otimes R$ for this pseudomonoid; the "e" superscript could be thought to stand for "endo" as well as the usual "envelope".

Definition 3.1 A *Takeuchi bialgebroid* is a k-module R together with an opmonoidal monad on R^e in the monoidal bicategory $\mathrm{Mod}(\mathcal{V})$.

To see that this definition agrees with that of $\times_R$-bialgebra as defined by Takeuchi [38] (and developed by [24, 41, 31, 8, 37]) we shall be more explicit about what an opmonoidal monad A on any pseudomonoid E involves.

In any monoidal bicategory $\mathcal{B}$ (with tensor product $\otimes$ and unit k) we use the term *pseudomonoid* (or "monoidal object") for an object E equipped with a binary multiplication $P : E \otimes E \to E$ and a unit $J : k \to E$ which are associative and unital up to coherent invertible 2-cells. A *monoidal morphism* $f : E \to E'$ is a morphism equipped with coherent 2-cells $P \circ (f \otimes f) \Rightarrow f \circ P$ and $J \Rightarrow f \circ J$. A *monoidal 2-cell* is a 2-cell compatible with these last coherent 2-cells. With the obvious compositions, this defines a bicategory $\mathrm{Mon}\mathcal{B}$ of pseudomonoids in $\mathcal{B}$. For example, if $\mathcal{B}$ is the cartesian-monoidal 2-category Cat of categories, functors and natural transformations then $\mathrm{Mon}\mathcal{B}$ is the 2-category MonCat of monoidal categories, monoidal functors and monoidal natural transformations as in [18].

We write $\mathcal{B}^{\mathrm{co}}$ for the bicategory obtained from $\mathcal{B}$ on reversing 2-cells. We put

$$\mathrm{Opmon}\mathcal{B} = (\mathrm{Mon}\mathcal{B}^{\mathrm{co}})^{\mathrm{co}};$$

the objects are again pseudomonoids, the morphisms are *opmonoidal morphisms*, and the 2-cells are *opmonoidal 2-cells*. An *opmonoidal monad* in $\mathcal{B}$ is a monad in $\mathrm{Opmon}\mathcal{B}$.

A monoidal morphism $f : E \to E'$ is called strong when the 2-cells $J \Rightarrow f \circ J$ and $P \circ (f \otimes f) \Rightarrow f \circ P$ are invertible. The inverses for these 2-cells equip such a strong f with the structure of opmonoidal morphism.

Now we return to the case of opmonoidal monads in $\mathcal{B} = \mathrm{Mod}(\mathcal{V})$. First of all, we have a module $A : E \to E$. The monad structure consists of module morphisms

$$\mu : A \otimes_E A \to A \qquad \text{and} \qquad \eta : E \to A$$

satisfying the usual conditions of associativity and unitality:

$$\mu \circ (\mu \otimes_E 1_A) = \mu \circ (1_A \otimes_E \mu), \qquad \mu \circ (\eta \otimes_E 1_A) = 1_A = \mu \circ (1_A \otimes_E \eta).$$

The opmonoidal structure consists of module morphisms

$$\delta : A \otimes_E P \to P \otimes_{E \otimes E} (A \otimes A) \quad \text{and} \quad \varepsilon : A \otimes_E J \to J$$

satisfying the following conditions:

$$\begin{tikzcd}
A \otimes_E P \otimes_{E^{\otimes 2}} (E \otimes P) \arrow[dddd, "\delta\otimes 1"'] \arrow[dr, "\cong"] & & & P \otimes_{E^{\otimes 2}} ((A \otimes_E P) \otimes A) \arrow[dddd, "1\otimes(\delta\otimes 1)"] \\
& A \otimes_E P \otimes_{E^{\otimes 2}} (P \otimes E) \arrow[r, "\delta\otimes 1"] & P \otimes_{E^{\otimes 2}} A^{\otimes 2} \otimes_{E^{\otimes 2}} (P \otimes E) \arrow[ur, "\cong"] & \\
& P \otimes_{E^{\otimes 2}} (A \otimes (A \otimes_E P)) \arrow[r, "1\otimes(1\otimes\delta)"'] & P \otimes_{E^{\otimes 2}} (A \otimes (P \otimes_{E^{\otimes 2}} A^{\otimes 2})) \arrow[dr, "\cong"] & \\
P \otimes_{E^{\otimes 2}} A^{\otimes 2} \otimes_{E^{\otimes 2}} (E \otimes P) \arrow[ur, "\cong"] & & & P \otimes_{E^{\otimes 2}} ((P \otimes_{E^{\otimes 2}} A^{\otimes 2}) \otimes A)
\end{tikzcd}$$

$$
\begin{tikzcd}
A \otimes_E P \otimes_{E^{\otimes 2}} (E \otimes J) \arrow[r, "\delta\otimes 1"] \arrow[dr, "\cong"'] & P \otimes_{E^{\otimes 2}} A^{\otimes 2} \otimes_{E^{\otimes 2}} (E \otimes J) \cong P \otimes_{E^{\otimes 2}} (A \otimes (A \otimes_E J)) \arrow[d, "1\otimes(1\otimes\varepsilon)"] \\
& A \cong P \otimes_{E^{\otimes 2}} (A \otimes J)
\end{tikzcd}
$$

$$
\begin{tikzcd}
A \otimes_E P \otimes_{E^{\otimes 2}} (J \otimes E) \arrow[r, "\delta\otimes 1"] \arrow[dr, "\cong"'] & P \otimes_{E^{\otimes 2}} A^{\otimes 2} \otimes_{E^{\otimes 2}} (J \otimes E) \cong P \otimes_{E^{\otimes 2}} ((A \otimes_E J) \otimes A) \arrow[d, "1\otimes(1\otimes\varepsilon)"] \\
& A \cong P \otimes_{E^{\otimes 2}} (J \otimes A)
\end{tikzcd}
$$

$$
\begin{tikzcd}
A \otimes_E A \otimes_E P \arrow[rr, "\mu\otimes 1"] \arrow[dd, "1\otimes\delta"'] & & A \otimes_E P \arrow[dr, "\delta"] & \\
& & & P \otimes_{E^{\otimes 2}} A^{\otimes 2} \\
A \otimes_E P \otimes_{E^{\otimes 2}} A^{\otimes 2} \arrow[r, "\delta\otimes 1"'] & P \otimes_{E^{\otimes 2}} A^{\otimes 2} \otimes_{E^{\otimes 2}} A^{\otimes 2} \arrow[r, "\cong"'] & P \otimes_{E^{\otimes 2}} (A \otimes_E A)^{\otimes 2} \arrow[ur, "1\otimes\mu^{\otimes 2}"'] &
\end{tikzcd}
$$

$$
\begin{tikzcd}
A \otimes_E A \otimes_E J \arrow[r, "\mu\otimes 1"] \arrow[d, "1\otimes\varepsilon"'] & A \otimes_E J \arrow[d, "\varepsilon"] \\
A \otimes_E J \arrow[r, "\varepsilon"'] & J
\end{tikzcd}
\quad
\begin{tikzcd}
P \arrow[r, "\eta\otimes 1"] \arrow[dr, "1\otimes\eta^{\otimes 2}"'] & A \otimes_E P \arrow[d, "\delta"] \\
& P \otimes_{E^{\otimes 2}} A^{\otimes 2}
\end{tikzcd}
\quad
\begin{tikzcd}
J \arrow[r, "\eta\otimes 1"] \arrow[dr, "1"'] & A \otimes_E J \arrow[d, "\varepsilon"] \\
& J
\end{tikzcd}
$$

Notice in particular that A becomes a k-algebra with multiplication defined by composing μ with the quotient morphism $A \otimes A \to A \otimes_E A$ and with unit $\eta(1)$. Indeed, $\eta : E \to A$ becomes a k-algebra morphism. Moreover, the structure on A as a module $A : E \to E$ is induced by $\eta : E \to A$ via $eae' = \eta(e)a\eta(e')$.

From time to time we will require special properties of bicategories such as $\mathrm{Mod}(\mathcal{V})$. In particular, at this moment, we need to point out that $\mathrm{Mod}(\mathcal{V})$ admits both the Kleisli and Eilenberg-Moore constructions for monads. For monads in 2-categories rather than bicategories, the universal nature of these constructions was defined in [33]; however, for the kind of phenomenon for modules we are about to explain, a better reference is [34]. To be explicit, a *monad* in a bicategory $\mathcal{B}$ is an object A of $\mathcal{B}$ together with a monoid t in the monoidal category $\mathcal{B}(A, A)$ in which the tensor product is horizontal composition in $\mathcal{B}$. An *Eilenberg-Moore* object for (A, t) is an object denoted A^t for which there is an equivalence of categories

$$\mathcal{B}(X, A^t) \simeq \mathcal{B}(X, A)^{\mathcal{B}(X,t)}$$

pseudonatural in objects X of $\mathcal{B}$, where the right-hand side is the category of Eilenberg- Moore algebras for the monad $\mathcal{B}(X, t)$ on the category $\mathcal{B}(X, A)$ in the familiar sense of say [25]. The existence of Eilenberg-Moore objects is a completeness condition on $\mathcal{B}$; that condition on $\mathcal{B}^{\mathrm{op}}$ is the Kleisli construction, the notion of monad being invariant under this kind of duality. That is, a *Kleisli object* for (A,t) is an object denoted A_t for which there is an equivalence of categories

$$\mathcal{B}(A_t, X) \simeq \mathcal{B}(A, X)^{\mathcal{B}(t,X)}$$

pseudonatural in objects X of $\mathcal{B}$.

Now we move more explicitly to the bicategory Mod($\mathcal{V}$). Notice that each k-algebra morphism $f : R \to S$ leads to two modules $f_* : R \to S$ and $f^* : S \to R$ which are both equal to S as k-modules but with the module actions defined by

$$sxr = sxf(r) \qquad \text{and} \qquad rys = f(r)ys$$

for $x \in f_*$, $y \in f^*$, $r \in R$ and $s \in S$. What is more, there are module morphisms

$$R \to f^* \otimes_S f_* \qquad \text{and } f_* \otimes_R f^* \to S,$$

the former defined by f and the latter defined by multiplication in S, forming the unit and counit of an adjunction in which f^* is right adjoint to f_*.

Suppose $A : E \to E$ is a monad on the k-algebra E in the bicategory Mod($\mathcal{V}$). The multiplication $\mu : A \otimes_E A \to A$ and unit $\eta : E \to A$ morphisms compose with the quotient morphism $A \otimes A \to A \otimes_E A$ and the unit $k \to E$, respectively, to provide the k-module A with a k-algebra structure with $\eta : E \to A$ becoming a morphism of k-algebras. Then μ can be regarded as a 2-cell

$$\begin{tikzcd}
E \arrow[rr, "A"] \arrow[dr, "\eta_*"'] & \arrow[d, phantom, "\overset{\mu}{\Leftarrow}"] & E \arrow[dl, "\eta_*"] \\
& A &
\end{tikzcd}$$

in Mod($\mathcal{V}$); it is a right action of the monoid A on η_*. Indeed, this is the universal right action of A on modules out of E; that is, the above triangle exhibits A as the Kleisli construction for the monad A on E. Since the homcategories of Mod($\mathcal{V}$) are cocomplete and composition with a given module preserves these colimits, the triangle

$$\begin{tikzcd}
E \arrow[rr, "A"] & \arrow[d, phantom, "\overset{\mu'}{\Rightarrow}"] & E \\
& A \arrow[ul, "\eta^*"] \arrow[ur, "\eta^*"'] &
\end{tikzcd}$$

in which μ' is the mate of μ under the adjunction $\eta_* \dashv \eta^*$, exhibits A as the Eilenberg-Moore construction for the monad A on E. That is, μ' is the universal left action of A on modules into E.

The following result abstracts Proposition 2.16 of [27].

Lemma 3.2 *If the monoidal bicategory $\mathcal{B}$ admits the Eilenberg-Moore construction for monads then so does* Opmon$\mathcal{B}$. *Furthermore, the forgetful morphism*

$$\text{Opmon}\mathcal{B} \to \mathcal{B}$$

preserves the Eilenberg-Moore construction.

In particular, this means that OpmonMod($\mathcal{V}$) admits the Eilenberg-Moore construction. (That the Kleisli construction exists for promonoidal monads was remarked in Section 3 of [10].)

Proposition 3.3 *Suppose E is a pseudomonoid in Mod($\mathcal{V}$) and $\eta : E \to A$ is a k-algebra morphism. There is an equivalence between the category of opmonoidal monad structures μ, δ, ε on $A : E \to E$ inducing η and the category of pseudomonoid structures on A for which $\eta^* : A \to E$ is a strong monoidal morphism.*

Proof In one direction, given the opmonoidal monad A on E inducing the given η, Lemma 3.2 lifts the triangle involving μ' to a triangle in OpmonMod($\mathcal{V}$) where it is again the Eilenberg-Moore construction. In particular, the adjunction $\eta_* \dashv \eta^*$ lifts to OpmonMod($\mathcal{V}$) and so, for general reasons explained in [22], $\eta^* : A \to E$ is strong monoidal. In the other direction, any k-algebra morphism $\eta : E \to A$ always has the property that η_* is opmonadic in Mod($\mathcal{V}$); that is, it supplies the Kleisli construction for the opmonoidal monad $\eta^* \otimes_A \eta_*$ on E generated by the adjunction $\eta_* \dashv \eta^*$. This opmonoidal monad has the form $A, \mu, \delta, \varepsilon, \eta$ as required. These two directions are the object functions for an obvious equivalence of categories. □

It follows that a Takeuchi bialgebroid can equally be defined as consisting of a k-algebra R, a k-algebra morphism $\eta : R^e \to A$, and a pseudomonoid structure on A for which η^* is strong monoidal.

In preparation for interpreting Takeuchi bialgebroids in terms of module categories, we need to clarify further some monoidal terminology. The concepts are not new but the terminology is inconsistent in the literature.

We say that a monoidal $\mathcal{V}$-category $\mathcal{A}$ is *left closed* when, for all pairs of objects B, C, there is an object $[B, C]_\ell$, called the *left internal hom of* B and C, for which there are isomorphisms

$$\mathcal{A}(A, [B, C]_\ell) \cong \mathcal{A}(A \otimes B, C),$$

$\mathcal{V}$-natural in A. A *right internal hom* $[B, C]$ satisfies

$$\mathcal{A}(A, [B, C]_r) \cong \mathcal{A}(B \otimes A, C).$$

We call a monoidal $\mathcal{V}$-category *closed* when it is both left and right closed. (This differs from Eilenberg-Kelly [18] who use "closed" for left closed. However, they were mainly interested in the symmetric case where left closed implies right closed.)

As pointed out in [18], if $\mathcal{A}$ and $\mathcal{X}$ are closed monoidal, a monoidal $\mathcal{V}$-functor $F : \mathcal{A} \to \mathcal{X}$, with its (lax) constraints

$$\phi_0 : I \to FI \qquad \text{and} \qquad \phi_{2;A,B} : FA \otimes FB \to F(A \otimes B)$$

subject to axioms, could equally be called a *left closed $\mathcal{V}$-functor* since these constraints are in bijection with pairs

$$\phi_0 : I \to FI \qquad \text{and} \qquad \phi^\ell_{2;B;C} : F[B, C]_\ell \to [FB, FC]_\ell$$

satisfying corresponding axioms. Equally F could be called a *right closed $\mathcal{V}$-functor* since the constraints are in bijection with pairs

$$\phi_0 : I \to FI \qquad \text{and} \qquad \phi^r_{2;A;;C} : F[A, C]_r \to [FA, FC]_r$$

satisfying corresponding axioms. We call a monoidal $\mathcal{V}$-functor F *normal* when ϕ_0 is invertible. As usual we call F *strong monoidal* when it is normal and each $\phi_{2;A,B}$ is invertible. We define F to be *strong left closed* when it is normal and each $\phi^\ell_{2;B;C}$ is invertible; it is *strong right closed* when it is normal and each $\phi^r_{2;A;C}$ is invertible; and it is *strong closed* when it both strong left and strong right closed.

Pseudomonoid structures on A in Mod($\mathcal{V}$) are equivalent to closed monoidal structures on the $\mathcal{V}$-category $\mathcal{V}^A = \text{Mod}(\mathcal{V})(k, A)$ of left A-modules; this is a special case of convolution in the sense of [9]. In fact, since k is a comonoid in $\mathrm{Mod}(\mathcal{V})$, we have a monoidal pseudofunctor

$$\mathrm{Mod}(\mathcal{V})(k,-): \ \mathrm{Mod}(\mathcal{V}) \to \mathcal{V}\text{-Cat},$$

which, as such, takes pseudomonoids to pseudomonoids. Since it is representable by k, it also preserves Eilenberg-Moore constructions (and all weighted limits for that matter). This means that when we apply $\mathrm{Mod}(\mathcal{V})(k,-)$ to a Takeuchi bialgebroid $\eta : R^e \to A$, we obtain a strong monoidal monadic functor

$$\mathcal{V}^A \to \mathcal{V}^{R^e}.$$

Conversely, given a k-algebra morphism $\eta : R^e \to A$, a $\mathcal{V}$-monoidal structure on $\mathcal{V}^A$, and a strong monoidal structure on the functor $\mathcal{V}^A \to \mathcal{V}^{R^e}$, we obtain a Takeuchi bialgebroid structure on $\eta : R^e \to A$. This is because $\mathcal{V}^A \to \mathcal{V}^{R^e}$ has both adjoints and is conservative (= reflects isomorphisms), so is monadic; but being strong monoidal and colimit preserving, any monoidal structure on $\mathcal{V}^A$ will be automatically closed, reflecting the fact that the monoidal $\mathcal{V}$-category $\mathcal{V}^{R^e}$ is closed. Consequently, by [9], the monoidal structure on $\mathcal{V}^A$ is obtained by convolution of a pseudomonoid structure on A.

By Theorem 5.1 of [30] (also see Theorem 3.1 of [8]) characterizing the $\times_R$-bialgebras of Takeuchi as monoidal structures on $\mathcal{V}^A$ for which $\mathcal{V}^A \to \mathcal{V}^{R^e}$ is strong monoidal, we have shown that our Takeuchi bialgebroids are the $\times_R$-bialgebras. We will see this in another way in the next section.

4 The lax monoidal operation $\times_R$

In order to define a bimonoid (or bialgebra) in a monoidal category, the monoidal category requires some kind of commutativity of the tensor product such as a braiding. A braiding can be regarded as a second monoidal structure on the category for which the new tensor is strongly monoidal with respect to the old. The so-called Eckmann-Hilton argument forces the new tensor to be isomorphic to the old and forces a braiding to appear (see [20]).

Ah, but what if the second tensor is only a lax multitensor and is only monoidal with respect to the old monoidal structure? Then there is certainly no need for the two structures to coincide. However, it is still possible to speak of a bimonoid: there is sufficient structure to express compatibility between a monoid structure for one tensor and a comonoid structure (on the same object) for the other tensor. After some preliminaries about right extensions in bicategories, we shall describe in detail just such a situation.

On top of the already discussed diverse properties and rich structure enjoyed by $\mathrm{Mod}(\mathcal{V})$, we also have the property that all right liftings and right extensions exist. Despite the terminology (from [33] for example), these concepts are very familiar in the usual theory of modules.

Suppose M and M' are modules $R \to S$. We put

$$\mathrm{Hom}_R^S(M, M') = \ \mathrm{Mod}(\mathcal{V})(R,S)(M, M');$$

that is, traditionally, it is the k-module of left S, right R-bimodule morphisms from M to M'. Now consider three modules as in the triangle

$$\begin{tikzcd} R \arrow[rr, "M"] \arrow[dr, "N"'] & & S \arrow[dl, "L"] \\ & T & \end{tikzcd}$$

Let $\mathrm{Hom}_R(M,N) : S \to T$ denote the k-module of right R-module morphisms with right S- and left T-actions defined by $(tfs)(m) = tf(sm)$ for

$$s \in S,\ t \in T,\ f \in \mathrm{Hom}_R(M,N) \quad \text{and} \quad m \in M.$$

Let $\mathrm{Hom}^T(L,N) : R \to S$ denote the k-module of left T-module morphisms with right R- and left S-actions defined by $(sgr)(\ell) = g(\ell s)r$ for

$$r \in R,\ s \in S,\ g \in \mathrm{Hom}^T(L,N) \quad \text{and } \ell \in L.$$

There are natural isomorphisms

$$\mathrm{Hom}_S^T(L, \mathrm{Hom}_R(M,N)) \cong \mathrm{Hom}_R^T(L \otimes_S M, N) \cong \mathrm{Hom}_R^S(M, \mathrm{Hom}^T(L,N)).$$

induced by evaluation morphisms

$$ev_N^M : \mathrm{Hom}_R(M,N) \otimes_S M \to N \quad \text{and} \quad ev_N^L : L \otimes_S \mathrm{Hom}^T(L,N)) \to N.$$

In bicategorical terms, $\mathrm{Hom}_R(M,N)$ is the right extension of N along M, while $\mathrm{Hom}^T(L,N)$ is the right lifting of N through L.

We require *normal lax monoidal categories* in the sense of [16] and [17]. These structures have been considered by Michael Batanin; they are the algebras for the categorical operad defined on page 88 of [4]. A lax monoidal structure on a category $\mathcal{E}$ amounts to a sequence of functors

$$\underset{n}{\bullet} : \underbrace{\mathcal{E} \times \cdots \times \mathcal{E}}_{n} \to \mathcal{E}$$

(thought of as multiple tensor products) together with substitution operations μ_ξ in the direction we will give below in our main example, and a unit $\eta : X \to \bullet_1 X$, satisfying three axioms. This is called *normal* when η is invertible (and so can be replaced by an identity).

Consider any pseudomonoid E, with multiplication P and unit J, in a monoidal bicategory $\mathcal{B}$ which admits all right extensions (where we have in mind $\mathcal{B} = \mathrm{Mod}(\mathcal{V})$). Then the endohom category $\mathrm{End}(E) = \mathcal{B}(E,E)$ becomes a lax monoidal category as follows. We define

$$P_n : E^{\otimes n} \to E$$

to be the composite

$$E^{\otimes n} \xrightarrow{P \otimes E^{\otimes(n-2)}} E^{\otimes(n-1)} \xrightarrow{P \otimes E^{\otimes(n-3)}} \cdots \xrightarrow{P \otimes E} E^{\otimes 2} \xrightarrow{P} E$$

for $n \geq 2$, to be the identity of $n = 1$, and to be J when $n = 0$. The coherence conditions for a pseudomonoid ensure that $P_m \cong P_n \circ (P_{m_1} \otimes \cdots \otimes P_{m_n})$ for each partition $\xi : m_1 + \cdots m_n = m$.

We define the multiple tensor $\bullet_n(M_1, \ldots, M_n)$ of objects $M_1, \ldots, M_n$ of $\mathrm{End}(E)$ to be the right extension of $P_n \circ (M_1 \otimes \cdots \otimes M_n)$ along P_n; that is,

$$\underset{n}{\bullet}(M_1, \ldots, M_n) = \mathrm{Hom}_{E^{\otimes n}}(P_n, P_n \otimes_{E^{\otimes n}} (M_1 \otimes \cdots \otimes M_n)).$$

The lax associativity constraint

$$\mu_\xi : \underset{n}{\bullet}(\underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1}),\dots,\underset{m_n}{\bullet}(M_{n1},\dots,M_{nm_n})) \to \underset{m}{\bullet}(M_{11},\dots,M_{nm_n})$$

for each partition $\xi : m_1 + \cdots m_n = m$ is, by using the right extension property of the target, induced by the morphism

$$\underset{n}{\bullet}(\underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1}),\dots,\underset{m_n}{\bullet}(M_{n1},\dots,M_{nm_n})) \circ P_m \to P_m \circ (M_{11},\dots,M_{nm_n})$$

which, after "conjugation" with $P_m \cong P_n \circ (P_{m_1} \otimes \cdots \otimes P_{m_n})$, is the composite

$$\begin{aligned}
&\underset{n}{\bullet}(\underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1}),\dots,\underset{m_n}{\bullet}(M_{n1},\dots,M_{nm_n})) \circ P_n \\
&\qquad \circ \ (P_{m_1} \otimes \cdots \otimes P_{m_n}) \xrightarrow{ev\circ 1} \\
&P_n \circ (\underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1}),\dots,\underset{m_n}{\bullet}(M_{n1},\dots,M_{nm_n})) \\
&\qquad \circ \ (P_{m_1} \otimes \cdots \otimes P_{m_n}) \xrightarrow{\cong} \\
&P_n \circ ((\underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1}) \circ P_{m_1}) \otimes \cdots \otimes \\
&\qquad (\bullet_{m_n}(M_{n1},\dots,M_{nm_n})) \circ P_{m_n}) \xrightarrow{1\circ(ev\otimes\cdots\otimes ev)} \\
&P_n \circ ((P_{m_1} \circ \underset{m_1}{\bullet}(M_{11},\dots,M_{1m_1})) \otimes \cdots \otimes \\
&\qquad P_{m_n} \circ \underset{m_n}{\bullet}(M_{n1},\dots,M_{nm_n}))) \xrightarrow{\cong} \\
&P_n \circ (P_{m_1} \otimes \cdots \otimes P_{m_n}) \circ (M_{11} \otimes \cdots \otimes M_{nm_n}).
\end{aligned}$$

The three axioms for a lax monoidal category can be verified. Since $P_1 : E \to E$ is the identity, we see that $\bullet_1 M = M$; so the lax monoidal structure on $\mathrm{End}(E)$ is normal.

As an endomorphism category $\mathrm{End}(E)$ is also a monoidal category for which the tensor product is composition. So $\mathrm{End}(E)$ is an object of the 2-category MonCat. Now MonCat is a monoidal 2-category with cartesian product as tensor. We will now see that $\mathrm{End}(E)$ is a lax monoid in MonCat.

Proposition 4.1 *Regard End(E) as a monoidal category under composition. The functors $\bullet_n$: End(E)^n $\to$ End(E) are equipped with canonical monoidal structures such that the substitutions μ_ξ are monoidal natural transformations.*

Proof The structure in question is the family of morphisms

$$\underset{n}{\bullet}(N_1,\dots,N_n) \circ \underset{n}{\bullet}(M_1,\dots,M_n) \to \underset{n}{\bullet}(N_1 \circ M_1,\dots,N_n \circ M_n)$$

which, using the right extension property of the target, are induced by the composites

$$\begin{aligned}
&\bullet_n(N_1,\dots,N_n) \circ \bullet_n(M_1,\dots,M_n) \circ P_n \xrightarrow{1\circ ev} \\
&\bullet_n(N_1,\dots,N_n) \circ P_n \circ (M_1,\dots,M_n) \xrightarrow{ev\circ 1}
\end{aligned}$$

$$P_n \circ (N_1 \otimes \cdots \otimes N_n) \circ (M_1 \otimes \cdots \otimes M_n) \xrightarrow{\cong} P_n \circ ((N_1 \circ M_1) \otimes \cdots \otimes (N_n \circ M_n)).$$

The compatibility of these morphisms with the lax associativity morphisms is readily verified. □

A monoid for composition in $\mathrm{End}(E)$ is a monad on E in $\mathcal{B}$. We write $\mathrm{MonEnd}(E)$ for the category of monads on E; the morphisms are 2-cells between the endofunctors of the monads that are compatible with the units and multiplications. It follows from Proposition 4.1 that the lax monoidal structure on $\mathrm{End}(E)$ lifts to the category $\mathrm{MonEnd}(E)$.

The concept of comonoid makes sense in any lax monoidal category.

Proposition 4.2 *A Takeuchi bialgebroid can equally be defined as a k-algebra R together with a comonoid in the lax monoidal category MonEnd(R^e).*

Proof Both a Takeuchi bialgebroid $A : R^e \to R^e$ and a comonoid in $\mathrm{MonEnd}(R^e)$ start with a monad $A : R^e \to R^e$ on R^e in $\mathrm{Mod}(\mathcal{V})$. To make this a comonoid in $\mathrm{MonEnd}(R^e)$ we need a comultiplication $\delta' : A \to \underset{2}{\bullet}(A, A)$ and a counit $\varepsilon' : A \to \underset{0}{\bullet}$ satisfying axioms. By the right extension properties of their targets, these morphisms are determined by morphisms $\delta : A \circ P_2 \to P_2 \circ (A \otimes A)$ and $\varepsilon : A \circ P_0 \to P_0$, exactly as for a Takeuchi bialgebroid. The condition that δ' and ε' should form a comonoid translates to the first three diagrams on δ and ε describing an opmonoidal monad (as in Section 3) while the conditions that δ' and ε' should respect the monad structure translate to the last four diagrams on δ and ε. So the comonoid is equivalently a Takeuchi bialgebroid. □

The operation $\underset{2}{\bullet}$ on $\mathrm{MonEnd}(R^e)$ is precisely the operation $\times_R$ of Takeuchi [38]; also compare Section 2 of [31][1] whose $\alpha : (M \times_R P) \times_R N \to M \times_R P \times_R N$, for example, is our substitution $\mu_\xi : \underset{2}{\bullet}(\underset{2}{\bullet}(M, P), N) \to \underset{3}{\bullet}(M, P, N)$ for $\xi : 2 + 1 = 3$. To help the reader make these identifications explicit, let $E = R^e = R^0 \otimes R$ take left-E, right E-bimodules M and N, and recall that $P_2 = R \otimes R \otimes R$ with the actions explained in Section 3. There is a canonical isomorphism

$$P_2 \otimes_{E^{\otimes 2}} (M \otimes N) \cong M \otimes_R N$$

where $M \otimes_R N = M \otimes N / \langle ((x \otimes 1)m) \otimes n \sim m \otimes ((1 \otimes x)n) \rangle$. Then we have the following calculation where the third isomorphism is obtained by evaluating the homomorphisms at $1 \otimes 1 \otimes 1 \in R \otimes R \otimes R$.

$$\begin{aligned}
\underset{2}{\bullet}(M, N) &\cong \mathrm{Hom}_{E^{\otimes 2}}(P_2, P_2 \otimes_{E^{\otimes 2}} (M \otimes N)) \cong \mathrm{Hom}_{E^{\otimes 2}}(P_2, M \otimes_R N) \\
&\cong \left\{ \sum_i m_i \otimes_R n_i \in M \otimes_R N \,\middle|\, \sum_i m_i(x \otimes 1) \otimes_R n_i \right. \\
&\qquad \left. = \sum_i m_i \otimes_R n_i(1 \otimes x) \ \forall\, x \in R \right\} \\
&= M \times_R N.
\end{aligned}$$

5 Monoidal star autonomy

In this section we extend the theory of $*$-autonomous categories in the sense of Barr (see [1], and, for the non-symmetric case, see [3]) to enriched categories in the sense of Eilenberg-Kelly [18]. The kind of duality present in a $*$-autonomous

[1] In its basic form the integral notation attributed to Mac Lane in [31] is originally due to Yoneda; see page 546 of [42]. It was adopted by [13] for their concept of "end" and "coend" in the general enriched context; however, their use of subscripts and superscripts on the integral (adopted by [25]) is the reverse of [31]. This reversal is reproduced in [8].

category is closer than compactness (also called rigidity or autonomy) to what is needed for an antipode in a bialgebroid or quantum category, and so for a concept of Hopf bialgebroid or quantum groupoid (see Example 7.4).

A $\mathcal{V}$-functor $F : \mathcal{A} \to \mathcal{B}$ is called *eso* (for "essentially surjective on objects") when every object of $\mathcal{B}$ is isomorphic to one of the form FA for some object A of $\mathcal{A}$.

A *left star operation* for a monoidal $\mathcal{V}$-category $\mathcal{A}$ is an eso $\mathcal{V}$-functor

$$S_\ell : \mathcal{A} \to \mathcal{A}^{\mathrm{op}}$$

together with a $\mathcal{V}$-natural family of isomorphisms (called the *left star constraint*)

$$\mathcal{A}(A \otimes B, S_\ell C) \cong \mathcal{A}(A, S_\ell(B \otimes C)).$$

It follows that $\mathcal{A}$ is left closed with $[B, C]_\ell \cong S_\ell(B \otimes D)$ where $S_\ell D \cong C$.

A *right star operation* for a monoidal $\mathcal{V}$-category $\mathcal{A}$ is an eso $\mathcal{V}$-functor

$$S_r : \mathcal{A}^{\mathrm{op}} \to \mathcal{A}$$

together with a $\mathcal{V}$-natural family of isomorphisms (called the *right star constraint*)

$$\mathcal{A}(A \otimes B, S_r C) \cong \mathcal{A}(B, S_r(C \otimes A)).$$

It follows that $\mathcal{A}$ is then right closed with $[A, C]_r \cong S_r(E \otimes A)$ where $S_r E \cong C$.

A monoidal $\mathcal{V}$-category $\mathcal{A}$ is called $*$-*autonomous* when it is equipped with a left star operation which is fully faithful. Since it follows that S_ℓ is then an equivalence of $\mathcal{V}$-categories, we write S_r for its adjoint equivalence so that the left star constraint can be written as

$$\mathcal{A}(A \otimes B, S_\ell C) \cong \mathcal{A}(B \otimes C, S_r A).$$

We see from this that S_r is a right star operation and $*$-autonomy can equally be defined in terms of a fully faithful right star operation. It follows that $*$-autonomous monoidal $\mathcal{V}$-categories are closed, with internal homs given by the formulas

$$[B, C]_\ell \cong S_\ell(B \otimes S_r C) \text{ and } [A, C]_r \cong S_r(S_\ell C \otimes A).$$

Notice that

$$\mathcal{A}(A, S_\ell I) \cong \mathcal{A}(I \otimes A, S_\ell I) \cong \mathcal{A}(A \otimes I, S_r I) \cong \mathcal{A}(A, S_r I),$$

so that $S_\ell I \cong S_r I$ (by the Yoneda Lemma). The object $S_\ell I$ is called the *dualizing object* and determines the left star operation via $[B, S_\ell I]_\ell \cong S_\ell B$.

For the reader interested in checking that our $*$-autonomous monoidal categories agree with Michael Barr's $*$-autonomous categories, we recommend Definition 2.3 of [2] as the appropriate one for comparison. Also see [35].

A monoidal $\mathcal{V}$-category is autonomous if and only if there exists a left star operation S_ℓ and a family of $\mathcal{V}$-natural isomorphisms

$$S_\ell(A \otimes B) \cong S_\ell B \otimes S_\ell A.$$

If $\mathcal{A}$ is autonomous then taking the left dual provides a left star operation with isomorphisms as required which *a fortiori* satisfy the conditions for a strong monoidal $\mathcal{V}$-functor. To see the less obvious implication, suppose we have an S_ℓ and the isomorphisms. Then $[B, C]_\ell \cong S_\ell(B \otimes D) \cong S_\ell D \otimes S_\ell B \cong C \otimes S_\ell B$ where $S_\ell D \cong C$, so $S_\ell B$ is a left dual for B. So every object B has a left dual $S_\ell B$. However, every object B is isomorphic to $S_\ell D$ for some D. This implies that D is a right dual for B.

6 Modules and promonoidal enriched categories

An important part of our goal is to extend star autonomy from monoidal categories to promonoidal categories. In preparation, in this section we shall discuss some basic facts about enriched categories and modules between them. Then we will review promonoidal categories and promonoidal functors in the enriched context [9]. We obtain a result about restriction along a promonoidal functor.

Let $\mathcal{V}$ denotes any complete and cocomplete symmetric monoidal closed category. We write $\mathcal{V}$-Mod for the symmetric monoidal bicategory (in the sense of [15]) whose objects are $\mathcal{V}$-categories and whose hom-categories are defined by

$$\mathcal{V}\text{-Mod}(\mathcal{A},\mathcal{B}) = [\mathcal{A}^{\mathrm{op}} \otimes \mathcal{B}, \mathcal{V}].$$

The objects $M : \mathcal{A} \to \mathcal{B}$ of $\mathcal{V}$-Mod$(\mathcal{A},\mathcal{B})$ are called *modules from $\mathcal{A}$ to $\mathcal{B}$*. The composite of modules $M : \mathcal{A} \to \mathcal{B}$ and $N : \mathcal{B} \to \mathcal{C}$ is defined by the equation

$$(N \circ M)(A, C) = \int^{B} N(B, C) \otimes M(A, B);$$

the integral here is the "coend" in the sense of [13] (also see [23]). The tensor product for $\mathcal{V}$-Mod is the usual tensor product of $\mathcal{V}$-categories in the sense of [18] (also see [23]); explicitly, an object of $\mathcal{A} \otimes \mathcal{B}$ is a pair (A, B) where A is an object of $\mathcal{A}$ and B is an object of $\mathcal{B}$, and the homs are defined by

$$(\mathcal{A} \otimes \mathcal{B})((A, B), (A', B')) = \mathcal{A}(A, A') \otimes \mathcal{B}(B, B').$$

Actually $\mathcal{V}$-Mod is autonomous since we have

$$\mathcal{V}\text{-Mod}(\mathcal{A} \otimes \mathcal{B}, \mathcal{C}) \cong \mathcal{V}\text{-Mod}(\mathcal{B}, \mathcal{A}^{\mathrm{op}} \otimes \mathcal{C})$$

since both sides are isomorphic to $[\mathcal{B}^{\mathrm{op}} \otimes \mathcal{A}^{\mathrm{op}} \otimes \mathcal{C}, \mathcal{V}]$.

We have reversed the direction of modules from that in [?] so that a promonoidal $\mathcal{V}$-category $\mathcal{A}$ is precisely a pseudomonoid (monoidal object) of $\mathcal{V}$-Mod (rather than $\mathcal{A}^{\mathrm{op}}$ being such). The multiplication module $P : \mathcal{A} \otimes \mathcal{A} \to \mathcal{A}$ and the unit module $J : \mathcal{I} \to \mathcal{A}$ are equally $\mathcal{V}$-functors

$$P : \mathcal{A}^{\mathrm{op}} \otimes \mathcal{A}^{\mathrm{op}} \otimes \mathcal{A} \to \mathcal{V} \quad \text{and} \quad J : \mathcal{A} \to \mathcal{V},$$

and we have associativity constraints

$$\int^{X} P(X, C; D) \otimes P(A, B; X) \cong \int^{Y} P(A, Y; D) \otimes P(B, C; Y)$$

and unital constraints

$$\int^{X} P(X, A; B) \otimes JX \cong \mathcal{A}(A, B) \cong \int^{Y} P(A, Y; B) \otimes JY,$$

satisfying the usual two axioms (see [9]) which yield coherence. It is convenient to introduce the $\mathcal{V}$-functors

$$P_n : \underbrace{\mathcal{A}^{\mathrm{op}} \otimes \cdots \otimes \mathcal{A}^{\mathrm{op}} \otimes \mathcal{A}}_{n} \to \mathcal{V},$$

for all natural numbers n, which we define as follows:

$$P_0 A = JA, \; P_1(A_1; A) = \mathcal{A}(A_1, A), \; P_2(A_1, A_2; A) = P(A_1, A_2; A)$$

and

$$P_{n+1}(A_1, \ldots, A_{n+1}; A) = \int^{X} P(X, A_{n+1}; A) \otimes P(A_1, \ldots, A_n; X).$$

We think of $P_n(A_1, \ldots, A_n; A)$ as the object of multimorphisms from $A_1, \ldots, A_n$ to A in $\mathcal{A}$. For example, when $\mathcal{A}$ is a monoidal $\mathcal{V}$-category, we have a promonoidal structure on $\mathcal{A}$ with

$$P_n(A_1, \ldots, A_n; A) \cong \mathcal{A}(A_1 \otimes \cdots \otimes A_n, A),$$

where the multitensor product is, say, bracketed from the left.

It will also be convenient to define a *multimorphism structure* on a $\mathcal{V}$-category $\mathcal{A}$ to be a sequence of $\mathcal{V}$-functors

$$P_n : \underbrace{\mathcal{A}^{\mathrm{op}} \otimes \cdots \otimes \mathcal{A}^{\mathrm{op}} \otimes \mathcal{A}}_{n} \to \mathcal{V}$$

subject to no constraints. So a promonoidal structure is an example where all the P_n are obtained from the particular ones for $n = 0, 1, 2$. A *multitensor structure* on $\mathcal{A}$ is a multimorphism structure for which each $P_n(A_1, \ldots, A_n; -)$ is representable; so we have objects $\underset{n}{\otimes}(A_1, \ldots, A_n)$ of $\mathcal{A}$ and a $\mathcal{V}$-natural family of isomorphisms

$$P_n(A_1, \ldots, A_n; A) \cong \mathcal{A}(\underset{n}{\otimes}(A_1, \ldots, A_n), A).$$

For example, when $\mathcal{A}$ is monoidal, we obtain $\underset{n}{\otimes}(A_1, \ldots, A_n)$ inductively from the cases $n = 0, 1$, and 2 where it is the unit, the identity functor, and the binary tensor product, respectively.

Suppose $\mathcal{A}$ and $\mathcal{E}$ are promonoidal $\mathcal{V}$-categories. A $\mathcal{V}$-functor $H : \mathcal{E} \to \mathcal{A}$ is called promonoidal when it is equipped with $\mathcal{V}$-natural families of morphisms

$$\phi_{2;U,V;W} : P(U,V;W) \to P(HU, HV; HW) \quad \text{and} \quad \phi_{0;U} : JU \to JHU$$

that are compatible in the obvious way with the associativity and unital constraints [10]. For any such promonoidal H, we can inductively define $\mathcal{V}$-natural families of morphisms

$$\phi_{n;U_1,\ldots,U_n;U} : P_n(U_1, \ldots, U_n; U) \to P_n(HU_1, \ldots, HU_n; HU)$$

using the inductive definition of P_n. In particular, $\phi_{1;U;V} : \mathcal{E}(U,V) \to \mathcal{A}(HU, HV)$ is the effect of H on homs. We say that H is *promonoidally fully faithful* when each $\phi_{n;U_1,\ldots,U_n;U}$ is invertible. We say F is *normal* when each $\phi_{0;U}$ is invertible.

A promonoidal $\mathcal{V}$-functor $H : \mathcal{E} \to \mathcal{A}$ also gives rise in the obvious way to $\mathcal{V}$-natural families of morphisms

$$\bar{\phi}_{2;A,B;W} : \int^{U,V} P(U,V;W) \otimes \mathcal{A}(A, HU) \otimes \mathcal{A}(B, HV) \to P(A, B; HW),$$

$$\phi^{\ell}_{2;U,B;C} : \int^{B,C} P(U,V;W) \otimes \mathcal{A}(B, HV) \otimes \mathcal{A}(HW, C) \to P(HU, B; C),$$

and

$$\phi^{r}_{2;A,V;C} : \int^{U,W} P(U,V;W) \otimes \mathcal{A}(A, HU) \otimes \mathcal{A}(HW, C) \to P(A, HV; C).$$

We need to say a little bit about convolution (see [9, 11, 17]). For $\mathcal{V}$-categories $\mathcal{A}$ and $\mathcal{X}$ equipped with multimorphism structures, the *convolution multimorphism structure* on the $\mathcal{V}$-functor $\mathcal{V}$-category $[\mathcal{A}, \mathcal{X}]$ is defined by

$$P_n(M_1, \ldots, M_n; M) = \int_{A_1,\ldots,A_n} [P_n(A_1, \ldots, A_n; A), P_n(M_1 A_1, \ldots, M_n A_n; MA)]$$

whenever these ends all exist (for example, when $\mathcal{A}$ is small). In the case where $\mathcal{X}$ is multitensored, the convolution is also multitensored by the formula

$$\underset{n}{\star}(M_1,\ldots,M_n)(A) = \int^{A_1,\ldots,A_n} P_n(A_1,\ldots,A_n;A)\otimes\underset{n}{\otimes}(M_1A_1,\ldots,M_nA_n),$$

provided the appropriate weighted colimits (expressed here by coends and tensors) exist in $\mathcal{X}$. In the case where $\mathcal{A}$ is promonoidal, if $\mathcal{X}$ is cocomplete closed monoidal then so is $[\mathcal{A},\mathcal{X}]$ (see [9]).

Proposition 6.1 *Suppose $H : \mathcal{E} \to \mathcal{A}$ is a normal promonoidal $\mathcal{V}$-functor. The restriction $\mathcal{V}$-functor*

$$[H,1] : [\mathcal{A},\mathcal{V}] \to [\mathcal{E},\mathcal{V}]$$

is a normal monoidal $\mathcal{V}$-functor. It is strong monoidal if and only if each $\bar{\phi}_{2;A,B;W}$ is invertible. It is strong left (respectively, strong right) closed if and only if each $\phi^{\ell}_{2;U,B;C}$ (respectively, $\phi^{r}_{2;A,V;C}$) is invertible.

Proof The monoidal unital constraint for $[H,1]$ is $\phi_{0;U} : JU \to JHU$. To obtain the associativity constraint, we use the Yoneda Lemma to replace

$$(MH\star NH)W = \int^{U,V} P(U,V;W)\otimes MHU\otimes NHV$$

by the isomorphic expression

$$\int^{U,V,A,B} P(U,V;W)\otimes\mathcal{A}(A,HU)\otimes\mathcal{A}(B,HV)\otimes MHU\otimes NHV$$

and take the morphism into

$$(M\star N)HW = \int^{A,B} P(A,B;HW)\otimes MA\otimes NB$$

of the form $\int^{A,B}\bar{\phi}_{2;A,B;W}\otimes 1\otimes 1$ which is clearly invertible if $\bar{\phi}_{2;A,B;W}$ is. The converse comes by taking M and N to be representable and using Yoneda.

Similarly, the left closed constraint for $[H,1]$ is obtained by composing the morphism $\int_{B,C}[\phi^{\ell}_{2;U,B;C}\otimes 1,1]$ from

$$[N,L]_{\ell}HU = \int_{B,C}[P(HU,B;C)\otimes NB, LC]$$

to

$$\int_{V,W,B,C}[P(U,V;W)\otimes\mathcal{A}(B,HV)\otimes\mathcal{A}(HW,C)\otimes NB, LC]$$

with the Yoneda isomorphism between this last expression and

$$[NH,LH]_{\ell}U = \int_{V,W}[P(U,V;W)\otimes NHV, LHW];$$

this constraint is clearly invertible if $\phi^{\ell}_{2;U,V;C}$ is, and the converse comes by taking N and L to be representable. The right closed case is dual. □

7 Forms and promonoidal star autonomy

A problem with $*$-autonomy is that the common base categories (like the category of sets and the category of vector spaces) are not themselves $*$-autonomous. So we do not expect the convolution monoidal structure on $[\mathcal{A}, \mathcal{V}]$ to be $*$-autonomous even when $\mathcal{A}$ is. We introduce the notion of *form* to address this problem: forms do exist on base categories and carry over to convolutions, while $*$-autonomy is to be equipped with a special kind of form. The definition of a $*$-autonomous promonoidal $\mathcal{V}$-category will be expressed in terms of forms.

A *form* for a promonoidal $\mathcal{V}$-category $\mathcal{A}$ is a module $\sigma : \mathcal{A} \otimes \mathcal{A} \to \mathcal{I}$ (where $\mathcal{I}$ is the usual one-object $\mathcal{V}$-category) together with an isomorphism $\sigma \circ (P \otimes 1) \cong \sigma \circ (1 \otimes P)$. In other words, a form is a $\mathcal{V}$-functor

$$\sigma : \mathcal{A}^{\mathrm{op}} \otimes \mathcal{A}^{\mathrm{op}} \to \mathcal{V}$$

together with a $\mathcal{V}$-natural family of isomorphisms

$$\int^X \sigma(X, C) \otimes P(A, B; X) \cong \int^Y \sigma(A, Y) \otimes P(B, C; Y)$$

called *form constraints.* Indeed, we can inductively obtain isomorphisms

$$\int^X \sigma(X, A_{n+1}) \otimes P_n(A_1, \ldots, A_n; X) \cong \int^Y \sigma(A_1, Y) \otimes P_n(A_2, \ldots, A_{n+1}; Y)$$

called the *generalized form constraints.* A promonoidal $\mathcal{V}$-category with a chosen form is called *formal.*

For example, every object K of any promonoidal $\mathcal{V}$-category $\mathcal{A}$ defines a form $\sigma(A, B) = P(A, B; K)$; the form constraints are provided by the promonoidal associativity and unit constraints. Other examples are $*$-autonomous monoidal categories, as we shall soon discover. Moreover, we will also see that forms carry over to various constructions such as tensor products and general convolutions of $\mathcal{V}$-categories.

If $\mathcal{A}$ is monoidal, using Yoneda, the form constraints become

$$\sigma(A \otimes B, C) \cong \sigma(A, B \otimes C).$$

A form is called *continuous* when $\sigma(A, -)$ and $\sigma(-, B) : \mathcal{A}^{\mathrm{op}} \to \mathcal{V}$ are small (weighted) limit preserving for all objects A and B of $\mathcal{A}$.

Proposition 7.1 *Let $\mathcal{A}$ and $\mathcal{X}$ be formal promonoidal $\mathcal{V}$-categories.*

(a) If $\mathcal{A}$ and $\mathcal{X}$ are formal then the tensor product $\mathcal{A} \otimes \mathcal{X}$ with promonoidal structure

$$P_n((A_1, X_1), \ldots, (A_n, X_n); (A, X)) = P_n(A_1, \ldots, A_n; A) \otimes P_n(X_1, \ldots, X_n; X)$$

admits the form $\sigma((A, X), (B, Y)) = \sigma(A, B) \otimes \sigma(X, Y)$.

(b) If $\mathcal{A}$ is small and $\mathcal{X}$ is cocomplete closed monoidal with a continuous form then the convolution monoidal $\mathcal{V}$-category $[\mathcal{A}, \mathcal{X}]$ admits the continuous form

$$\sigma(M, N) = \int_{A,B} [\sigma(A, B), \sigma(MA, NB)].$$

Proof (a) This is trivial.

(b) We have the calculation

$$\begin{aligned}
\sigma(M \star N, L) &= \int_{U,C} [(M \star N)U \otimes LC, \sigma(U,C)] \\
&\cong \int_{U,C} \left[\sigma(U,C), \sigma\left(\left(\int^{A,B} P(A,B;U) \otimes MA \otimes NB, LC\right)\right]\right. \\
&\cong \int_{U,A,B,C} [\sigma(U,C) \otimes P(A,B;U), \sigma(MA \otimes NB, LC)] \\
&\cong \int_{(,A,B,C} [\sigma(A,U) \otimes P(B,C;U), \sigma(MA, NB \otimes LC)] \\
&\cong \int_{U,A} \left[\sigma(A,U), \sigma\left(MA, \int^{B,C} P(B,C;U) \otimes NB \otimes LC\right)\right] \\
&\cong \int_{U,A} [\sigma(A,U), \sigma(MA, (N \star L)U)] \cong \sigma(M, N \star L).
\end{aligned}$$

□

A form $\sigma : \mathcal{A} \otimes \mathcal{A} \to \mathcal{I}$ transforms under the duality of $\mathcal{V}$-modules to a $\mathcal{V}$-module $\hat{\sigma} : \mathcal{A} \to \mathcal{A}^{\mathrm{op}}$. We say the form σ is *non-degenerate* when $\hat{\sigma}$ is an equivalence as a $\mathcal{V}$-module (a Morita equivalence if you prefer). A form σ is said to be representable when there exists a $\mathcal{V}$-functor $S_\ell : A \to \mathcal{A}^{\mathrm{op}}$ and a $\mathcal{V}$-natural isomorphism

$$\sigma(A,B) \cong \mathcal{A}(A, S_\ell B).$$

A promonoidal $\mathcal{V}$-category is $*$-*autonomous* when it is equipped with a representable non-degenerate form. In fact, if $\mathcal{A}$ satisfies a minimal completeness condition ("Cauchy completeness") then "representable" is redundant. Notice that S_ℓ is necessarily an equivalence, with adjoint inverse S_r, say, and the form constraints have the cyclic appearance

$$P(A,B;S_\ell C) \cong P(B,C;S_r A).$$

More generally, using Yoneda, the generalized form constraints become

$$\begin{aligned}
&P_n(A_1, \ldots, A_n; S_\ell A_{n+1}) \\
&\cong \int^X \mathcal{A}(X, S_\ell A_{n+1}) \otimes P_n(A_1, \ldots, A_n; X) \\
&\cong \int^X \sigma(X, A_{n+1}) \otimes P_n(A_1, \ldots, A_n; X) \cong \int^Y \sigma(A_1, Y) \otimes P_n(A_2, \ldots, A_{n+1}; Y) \\
&\cong \int^Y \mathcal{A}(Y, S_r A_1) \otimes P_n(A_2, \ldots, A_{n+1}; Y) \cong P_n(A_2, \ldots, A_{n+1}; S_r A_1).
\end{aligned}$$

A monoidal category is $*$-autonomous in the monoidal sense if and only if it is $*$-autonomous in the promonoidal sense.

Corollary 7.2 *In Proposition 7.1, if $\mathcal{A}$ and $\mathcal{X}$ are $*$-autonomous then so are*

$$(a)\ \mathcal{A} \otimes \mathcal{X} \qquad and \qquad (b)\ [\mathcal{A}, \mathcal{X}].$$

Proof (a)

$$\begin{aligned}
\sigma((A,X),(B,Y)) &= \sigma(A,B) \otimes \sigma(X,Y) \cong \mathcal{A}(A, S_\ell B) \otimes \mathcal{X}(X, S_\ell Y) \\
&\cong (\mathcal{A} \otimes \mathcal{X})((A,X),(S_\ell B, S_\ell Y)).
\end{aligned}$$

(b)

$$\begin{aligned}\sigma(M,N) &= \int_{A,B}[\sigma(A,B),\sigma(MA,NB)] \cong \int_{A,B}[\mathcal{A}(A,S_\ell B),\mathcal{X}(MA,S_\ell NB)]\\ &\cong \int_B \mathcal{X}(MS_\ell B, S_\ell NB) \cong [\mathcal{A},\mathcal{X}](MS_\ell, S_\ell N) \cong [\mathcal{A},\mathcal{X}](M, S_\ell N S_r).\end{aligned}$$

□

Example 7.3 As noted in the final remarks of [9], for any $\mathcal{V}$-category $\mathcal{C}$, there is a canonical promonoidal structure on $\mathcal{C}^{\mathrm{op}} \otimes \mathcal{C}$. It is explicitly defined by

$$P_0(C,D) = J(C,D) = \mathcal{C}(C,D)$$

and

$$P_2((D_1,C_1),(D_2,C_2);(C_3,D_3)) = \mathcal{C}(C_3,D_1)\otimes\mathcal{C}(C_1,D_2)\otimes\mathcal{C}(C_2,D_3).$$

More generally,

$$\begin{aligned}&P_n((D_1,C_1),\ldots,(D_n,C_n);(C_{n+1},D_{n+1}))\\ &\quad = \mathcal{C}(C_{n+1},D_1)\otimes\mathcal{C}(C_1,D_2)\otimes\cdots\otimes\mathcal{C}(C_n,D_{n+1}).\end{aligned}$$

After the lecture [12], Luigi Santocanale observed that $\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C}$ is $*$-autonomous. To be precise, define $S : (\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C})^{\mathrm{op}} \to \mathcal{C}^{\mathrm{op}}\otimes\mathcal{C}$ by $S(D,C) = (C,D)$. Clearly

$$\begin{aligned}&P_n((D_1,C_1),\ldots,(D_n,C_n);(C_{n+1},D_{n+1}))\\ &\quad = P_n((D_2,C_2),\ldots,(D_{n+1},C_{n+1});(C_1,D_1)),\end{aligned}$$

so that $S_r = S_\ell = S$ for $*$-autonomy. To relate this to our discussion of bialgebroids (Section 3), note that a k-algebra $\mathcal{C} = R$ is a one-object $\mathcal{V}$-category (for $\mathcal{V}$ the category of k-modules) and so the "chaotic bialgebroid" $\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C} = R^e$ is $*$-autonomous.

Example 7.4 The notion of Hopf $\mathcal{V}$-algebroid appearing in Definition 21 of [15] is an example of a $*$-autonomous promonoidal $\mathcal{V}$-category. Suppose that the $\mathcal{V}$-category $\mathcal{C}$ is comonoidal [9]; that is, $\mathcal{C}$ is a pseudomonoid (or monoidal object) in $(\mathcal{V}\text{-Cat})^{\mathrm{op}}$: this means we have $\mathcal{V}$-functors $\Delta : \mathcal{C} \to \mathcal{C}\otimes\mathcal{C}$ and $E : \mathcal{C} \to \mathcal{I}$, coassociative and counital up to coherent $\mathcal{V}$-natural isomorphisms. It is easy to see that Δ must be given by the diagonal $\Delta C = (C,C)$ on objects. A multimorphism structure Q on $\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C}$ is then defined by

$$Q_n(C;C_1,\ldots,C_n) = \mathcal{C}(C,C_1)\otimes\cdots\otimes\mathcal{C}(C,C_n);$$

the actions on hom-objects require the $\mathcal{V}$-functors Δ and E. Indeed, Q defines a promonoidal structure (compare Section 5 of [9]). If this promonoidal $\mathcal{V}$-category is $*$-autonomous then the condition $Q(A,B;S_\ell C) \cong Q(B,C;S_r A)$ becomes

$$\mathcal{C}(A,S_\ell C)\otimes\mathcal{C}(B,S_\ell C) \cong \mathcal{C}(B,S_r A)\otimes\mathcal{C}(C,S_r A) \cong \mathcal{C}(B,S_r A)\otimes\mathcal{C}(A,S_\ell C),$$

which precisely gives the condition

$$\mathcal{C}(A,C)\otimes\mathcal{C}(B,C) \cong \mathcal{C}(B,S_r,A)\otimes\mathcal{C}(A,C)$$

for the authors' concept of Hopf $\mathcal{V}$-algebroid [15].

A promonoidal functor $H : \mathcal{E} \to \mathcal{A}$ between $*$-autonomous promonoidal $\mathcal{V}$-categories is called *$*$-autonomous* when it is equipped with a $\mathcal{V}$-natural transformation

$$\tau^{\ell} : HS_{\ell} \to S_{\ell}H$$

such that the following diagram commutes

$$\begin{tikzcd}
P(U,V;S_{\ell}W) \arrow[r, "\phi_{2;U,V;S_{\ell}W}"] \arrow[d, "\cong"'] & P(HU,HV;SH_{\ell}W) \arrow[r, "P(1,1;\tau^{\ell})"] & P(HU,HV;S_{\ell}HW) \arrow[d, "\cong"] \\
P(V,W;S_rU) \arrow[r, "\phi_{2;V,W;S_rU}"'] & P(HV,HW;HS_r,U) \arrow[r, "P(1,1;\tau^{r})"'] & P(HV,HW;S_rHU)
\end{tikzcd}$$

where $\tau^r : HS_r \to S_rH$ is the mate of τ^{ℓ} under the adjunction between S_{ℓ} and S_r. We call H *strong $*$-autonomous* when τ^{ℓ} is invertible; it follows that τ^r is invertible.

Proposition 7.5 *Suppose $H : \mathcal{E} \to \mathcal{A}$ is a strong $*$-autonomous promonoidal $\mathcal{V}$-functor. If the restriction $\mathcal{V}$-functor $[H,1] : [\mathcal{A},\mathcal{V}] \to [\mathcal{E},\mathcal{V}]$ is strong monoidal then it is strong closed.*

Proof The idea of the proof is to use $*$-autonomy to cycle the criterion of Proposition 6.1 for $[H,1]$ to be strong monoidal into the criteria for it to be strong closed. The precise calculation for strong left closed is as follows:

$$\begin{aligned}
&\int^{V,W} P(U,V;W) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(HW,C) \\
\cong\ &\int^{V,W} P(U,V;W) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(HW,S_{\ell}S_rC) \\
\cong\ &\int^{V,W} P(U,V;W) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(S_rC,S_rHW) \\
\cong\ &\int^{V,W} P(U,V;W) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(S_rC,HS_rW) \\
\cong\ &\int^{V,W} P(U,V;S_{\ell}W) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(S_rC,HW) \\
\cong\ &\int^{V,W} P(U,V;S_rU) \otimes \mathcal{A}(B,HV) \otimes \mathcal{A}(S_rC,HW) \\
\cong\ &P(B,S_rC;HS_rU) \cong P(P,S_rC;S_rHU) \cong P(HU,B;S_{\ell}S_rC) \cong P(HU,B;C).
\end{aligned}$$

$\square$

The next simple observation can be useful in this context.

Proposition 7.6 *Suppose $U : \mathcal{A} \to \mathcal{X}$ is any $\mathcal{V}$-functor with a left adjoint F, and suppose there are equivalences $S : \mathcal{A} \to \mathcal{A}^{\mathrm{op}}$ and $S : \mathcal{X} \to \mathcal{X}^{\mathrm{op}}$ such that $S \circ U \cong U \circ S$. Then U has a right adjoint $S^{-1} \circ F \circ S$ and the monad $T = U \circ F$ generated by the original adjunction has a right adjoint comonad $G = U \circ S^{-1} \circ F \circ S$. Dually, F has a left adjoint $S^{-1} \circ U \circ S$. A doubly infinite string of adjunctions is thereby created.*

Proof Clearly $U : \mathcal{A}^{\mathrm{op}} \to \mathcal{X}^{\mathrm{op}}$ has F as right adjoint whereas the mutually inverse equivalences S and S^{-1} are adjoint to each other on both sides. The results now follow by composing adjunctions. $\square$

An *opform* for a promonoidal $\mathcal{V}$-category $\mathcal{A}$ is a $\mathcal{V}$-functor

$$\sigma : \mathcal{A} \otimes \mathcal{A} \to \mathcal{V}$$

and $\mathcal{V}$-natural isomorphisms

$$\int_X [P(A,B;X), \sigma(X,C)] \cong \int_Y [P(B,C;Y), \sigma(A,Y)],$$

called *opform constraints.* For a monoidal $\mathcal{V}$-category, we see by Yoneda's Lemma that an opform on $\mathcal{A}$ is the same as a form on $\mathcal{A}^{\mathrm{op}}$. Moreover, in general, if σ is a form on $\mathcal{A}$ and K is any object of $\mathcal{V}$ then an opform σ_K on $\mathcal{A}$ is defined by the equation

$$\sigma_K(A,B) = [\sigma(A,B), K].$$

Proposition 7.7 *Let $\mathcal{A}$ be a small promonoidal $\mathcal{V}$-category. Each opform σ for $\mathcal{A}$ determines a continuous form for the convolution monoidal $\mathcal{V}$-category $[\mathcal{A}, \mathcal{V}]$ via the formula*

$$\sigma(M,N) = \int_{A,B} [MA \otimes NB, \sigma(A,B)].$$

Furthermore, every continuous form on $[\mathcal{A}, \mathcal{V}]$ arises thus from an opform on $\mathcal{A}$.

Proof We have the calculation

$$\begin{aligned}
\sigma(M \star N, L) &= \int_{U,C} [(M \star N)U \otimes LC, \sigma(U,C)] \\
&\cong \int_{U,C} \left[\int^{A,B} P(A,B;U) \otimes MA \otimes NB \otimes LC, \sigma(U,C) \right] \\
&\cong \int_{U,A,B,C} [MA \otimes NB \otimes LC, [P(A,B;U), \sigma(U,C)]] \\
&\cong \int_{U,A,B,C} [MA \otimes NB \otimes LC, [P(B,C;U), \sigma(A,U)]] \\
&\cong \int_{U,A,B,C} [MA, [P(B,C;U) \otimes NB \otimes LC, \sigma(A,U)]] \\
&\cong \int_{U,A} \left[MA, \left[\int^{B,C} P(B,C;U) \otimes NB \otimes LC, \sigma(A,U) \right] \right] \\
&\cong \sigma(M, N \star L).
\end{aligned}$$

Conversely, any continuous form σ on $[\mathcal{A}, \mathcal{V}]$ will have

$$\begin{aligned}
\sigma(M,N) &\cong \sigma\left(\int^A MA \otimes \mathcal{A}(A,-), \int^B NB \otimes \mathcal{A}(B,-) \right) \\
&\cong \int_{A,B} [MA \otimes MB, \sigma(\mathcal{A}(A,-), \mathcal{A}(B,-))].
\end{aligned}$$

so that σ will be determined by its value on representables. We define σ for $\mathcal{A}$ by

$$\sigma(A,B) = \sigma(\mathcal{A}(A,-), \mathcal{A}(B,-)).$$

We have the calculation

$$\begin{aligned}
&\int_U [P(A,B;U), \sigma(\mathcal{A}(U,-),\mathcal{A}(C,-))] \\
&\cong \sigma\left(\left(\int^U P(A,B;U)\otimes \mathcal{A}(U,-), \mathcal{A}(C,-)\right)\right) \\
&\cong \sigma(P(A,B;-),\mathcal{A}(C,-)) \cong \sigma(\mathcal{A}(A,-)\star\mathcal{A}(B,-),\mathcal{A}(C,-)) \\
&\cong \sigma(\mathcal{A}(A,-),\mathcal{A}(B,-)\star\mathcal{A}(C,-)) \\
&\cong \sigma(\mathcal{A}(A,-),P(B,C;-)) \cong \int_V [P(B,C;V),\sigma(\mathcal{A}(A,-),\mathcal{A}(V,-))].
\end{aligned}$$

□

8 The star and Chu constructions

We adhere to the spirit of the review [36] where the Chu construction is defined at the multimorphism level. The star construction on a multimorphism structure yields one that is $*$-autonomous. When applied to a promonoidal $\mathcal{V}$-category, the result may not be promonoidal — hence the need to work at the more general level.

For that, we define a general multimorphism structure to be *$*$-autonomous* when there exists an equivalence $S_\ell : \mathcal{A} \to \mathcal{A}^{\mathrm{op}}$ of $\mathcal{V}$-categories and a sequence of $\mathcal{V}$-natural isomorphisms

$$P_n(A_1,\dots,A_n;S_\ell A_{n+1}) \cong P_n(A_2,\dots,A_{n+1};S_r A_1)$$

where S_r is an adjoint inverse for S_ℓ.

In this section we will show how to modify a multimorphism structure, with a prescribed S_ℓ, to obtain a $*$-autonomous one with the same S_ℓ. We first need a natural definition: an *equivalence* $F : \mathcal{A} \to \mathcal{B}$ of multimorphism structures is an equivalence F of $\mathcal{V}$-categories together with natural isomorphisms

$$P_n(A_1,\dots,A_n;A) \cong P_n(FA_1,\dots,FA_n;FA);$$

the inverse equivalence of F is obviously also a multimorphism equivalence.

Notice that, for any $*$-autonomous multimorphism structure, $S_\ell \circ S_\ell : \mathcal{A} \to \mathcal{A}$ is a multimorphism equivalence: for we have the calculation

$$\begin{aligned}
&P_n(A_1,\dots,A_n;A) \\
&\cong P_n(A_1,\dots,A_n;S_rS_\ell A) \\
&\cong P_n(S_\ell A,A_1,\dots,A_{n-1};S_\ell A_n) \cong P_n(S_\ell A,A_1,\dots,A_{n-1};S_rS_\ell S_\ell A_n) \\
&\cong \cdots \cong P_n(S_\ell S_\ell A_2,\dots,S_\ell S_\ell A_n,S_\ell A;S_\ell A_1) \cong P_n(S_\ell S_\ell A_1,\dots,S_\ell S_\ell A_n;S_\ell S_\ell A).
\end{aligned}$$

Now to our construction. Suppose we have a multimorphism structure P on any $\mathcal{V}$-category $\mathcal{A}$ equipped with a contravariant $\mathcal{V}$-functor $S_\ell : \mathcal{A}^{\mathrm{op}} \to \mathcal{A}$ such that $S_\ell \circ S_\ell : \mathcal{A} \to \mathcal{A}$ is an equivalence of multimorphism structures. It follows that S_ℓ is an equivalence; we write S_r for the adjoint equivalence. The *starring* of this situation is the multimorphism structure P^* on $\mathcal{A}$ defined by the formula

$$\begin{aligned}
&P^*_n(X_1,\dots,X_n;S_\ell X_{n+1}) \\
&= \int^{U_{ij}(1\le i<j\le n+1)} \bigotimes_{m=1}^{n+1} P_n(U_{m\,m+1},\dots U_{m\,n+1},S_rU_{1\,m},\dots,S_rU_{m-1\,m};S_rX_m).
\end{aligned}$$

Proposition 8.1 *The starring P^* produces a $*$-autonomous multimorphism structure on $\mathcal{A}$ with the given S_ℓ.*

Proof Extend the definition of the U_{ij} and X_i by putting $U_{ji} = S_r U_{ij}$ and $X_{n+i+1} = S_r S_r X_i$. From the definition, we have

$$P_n^*(X_2, \ldots, X_{n+1}; S_r X_1)$$
$$= \int^{V_{ij}(1 \leq i < j \leq n+1)} \bigotimes_{m=1}^{n+1} P_n(V_{m\,m+1}, \ldots V_{m\,n+1}, S_r V_{1\,m}, \ldots, S_r V_{m-1\,m}; S_r X_{m+1}),$$

which we notice is isomorphic to the formula for $P_n^*(X_1, \ldots, X_n; S_\ell X_{n+1})$ on making the change of variables $V_{ij} = U_{i+1\,j+1}$ and using the isomorphisms

$$P_n(U_{12}, \ldots, U_{1\,n+1}; S_r X_1) \cong P_n(S_r S_r U_{12}, \ldots, S_r S_r U_{1\,n+1}; S_r S_r S_r X_1).$$

□

Let $\mathcal{C}$ be a $\mathcal{V}$-category with a multimorphism structure P and a multimorphism equivalence $T : \mathcal{C} \to \mathcal{C}$. We suppose furthermore that $\mathcal{C}$ is a comonoidal $\mathcal{V}$-category with derived promonoidal structure Q as made explicit in Example 7.4. We require that $T : \mathcal{C}^{\mathrm{op}} \to \mathcal{C}^{\mathrm{op}}$ is an equivalence for the multimorphism structure Q (that is, that $T : \mathcal{C} \to \mathcal{C}$ is a comonoidal equivalence).

We want to apply the star construction to $\mathcal{A} = \mathcal{C}^{\mathrm{op}} \otimes \mathcal{C}$ with $S_\ell(C, D) = (D, T^{-1}C)$, so that $S_r(C, D) = (TD, C)$, and with the tensor product multimorphism structure $Q \otimes P$ for the P and Q as described in the last paragraph. Notice $S_\ell S_\ell(C, D) = (T^{-1}C, T^{-1}D)$ so that $S_\ell \circ S_\ell : \mathcal{C}^{\mathrm{op}} \otimes \mathcal{C} \to \mathcal{C}^{\mathrm{op}} \otimes \mathcal{C}$ is indeed a multimorphism equivalence.

Let us calculate the star R^* of $R = Q \otimes P$:

$$R_n^*\left((X_1, Y_1), \ldots, (X_n, Y_n); (Y_{n+1}, T^{-1}X_{n+1})\right)$$
$$= \int^{(U_{ij}, V_{ij})} \bigotimes_{m=1}^{n+1} \left(\begin{array}{l} Q_n(TY_m; U_{m\,m+1}, \ldots, U_{m\,n+1}, TV_{1\,m}, \ldots, TV_{m-1\,m}) \\ \otimes P_n(V_{m\,m+1}, \ldots, V_{m\,n+1}, U_{1\,m}, \ldots, U_{m-1\,m}; X_m) \end{array} \right)$$
$$\cong \int^{(U_{ij}, V_{ij})} \bigotimes_{m=1}^{n+1} \left(\begin{array}{l} \mathcal{C}(TY_m, U_{m\,m+1}) \otimes \cdots \otimes \mathcal{C}(TY_m, U_{m\,n+1}) \\ \otimes \mathcal{C}(Y_m, V_{1\,m}) \otimes \cdots \otimes \mathcal{C}(Y_m, V_{m-1\,m}) \\ \otimes P_n(V_{m\,m+1}, \ldots, V_{m\,n+1}, U_{1\,m}, \ldots, U_{m-1\,m}; X_m) \end{array} \right)$$
$$\cong \int^{(U_{ij}, V_{ij})} \bigotimes_{r<s} (\mathcal{C}(TY_r, U_{rs}) \otimes \mathcal{C}(Y_s, V_{rs}))$$
$$\otimes \bigotimes_{m=1}^{n+1} P_n(V_{m\,m+1}, \ldots, V_{m\,n+1}, U_{1\,m}, \ldots, U_{m-1\,m}; X_m)$$
$$\cong \bigotimes_{m=1}^{n+1} P_n(Y_{m+1}, \ldots, Y_{n+1}, TY_1, \ldots, TY_{m-1}; X_m)$$

which has the same shape as the multimorphism structure described in [36].

Proposition 8.2 *In the situation just described, if P is actually a monoidal structure on $\mathcal{C}$, then R^* is a $*$-autonomous promonoidal structure on $\mathcal{C}^{\mathrm{op}} \otimes \mathcal{C}$.*

Proof After Proposition 8.1, it suffices to show that R^* is promonoidal. We need to see that each R_n^* is determined by the $n = 0$ and $n = 2$ cases. The general calculation is by induction so we trust that the following exemplary step will be sufficient indication for the reader:

$$\begin{aligned}
&\int^{A_1,B_1} R_2^*\left((X_1,Y_1),(X_2,Y_2);(B_1,T^{-1}A_1)\right) \\
&\qquad\qquad \otimes R_2^*\left((TB_1,A_1),(X_3,Y_3);(B_4,T^{-1}A_4)\right) \\
&\cong \int^{A_1,B_1} \left(\begin{array}{l} P_2(Y_2,B_1;X_1)\otimes P_2(B_1,TY_1;X_2)\otimes P_2(TY_1,TY_2;A_1) \\ \otimes P_2(Y_3,Y_4;TB_1)\otimes P_2(Y_4,A_1;X_3)\otimes P_2(A_1,TY_3;X_4) \end{array} \right) \\
&\cong \int^{A_1,B_1} \left(\begin{array}{l} \mathcal{C}(Y_2\otimes B_1,X_1)\otimes \mathcal{C}(B_1\otimes TY_1,X_2)\otimes \mathcal{C}(TY_1\otimes TY_2,A_1) \\ \otimes \mathcal{C}(Y_3\otimes Y_4,TB_1)\otimes \mathcal{C}(Y_4\otimes A_1,X_3)\otimes \mathcal{C}(A_1\otimes TY_3,X_4) \end{array} \right) \\
&\cong \left(\begin{array}{l} \mathcal{C}(Y_2\otimes T(Y_3\otimes Y_4),X_1)\otimes \mathcal{C}(T(Y_3\otimes Y_4)\otimes TY_1,X_2) \\ \otimes \mathcal{C}(Y_4\otimes TY_1\otimes TY_2,X_3)\otimes \mathcal{C}(TY_1\otimes TY_2\otimes TY_3,X_4) \end{array} \right) \\
&\cong P_3(Y_2,TY_3,TY_4;X_1)\otimes P_3(TY_3,TY_4,TY_1;X_2)\otimes P_3(Y_4,TY_1,TY_2;X_3) \\
&\qquad\qquad \otimes P_3(TY_1,TY_2,TY_3;X_4) \\
&\cong R_3^*\left((X_1,Y_1),(X_2,Y_2),(X_3,Y_3);(Y_4,T^{-1}X_4)\right).
\end{aligned}$$

□

Proposition 8.3 *In the situation of the Proposition 8.2, further suppose that P is closed monoidal and that the comonoidal structure on $\mathcal{C}$ is representable by an object K, an operation $B \bullet C$, and $\mathcal{V}$-natural isomorphisms*

$$\mathcal{C}(A,K)\cong I \qquad \textit{and} \qquad \mathcal{C}(A,B\bullet C)\cong (\mathcal{C}(A,B)\otimes \mathcal{C}(A,C))$$

where the right-hand sides require the counit and comultiplication for their effects on homs. Then R^* is a $*$-autonomous monoidal structure on $\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C}$.

Proof We have the calculations:

$$\begin{aligned}
&R_2^*(X_1,Y_1),(X_2,Y_2);(Y_3,T^{-1}X_3)) \\
&\quad\cong\quad P_2(Y_2,Y_3;X_1)\otimes P_2(Y_3,TY_1;X_2)\otimes P_2(TY_1,TY_2;X_3) \\
&\quad\cong\quad \mathcal{C}(Y_2\otimes Y_3,X_1)\otimes \mathcal{C}(Y_3\otimes TY_1,X_2)\otimes \mathcal{C}(TY_1\otimes TY_2,X_3) \\
&\quad\cong\quad \mathcal{C}\left(Y_3,[Y_2,X_1]_r\right)\otimes \mathcal{C}(Y_3,[TY_1,X_2]_\ell)\otimes \mathcal{C}(TY_1\otimes TY_2,X_3) \\
&\quad\cong\quad \mathcal{C}\left(Y_3,[Y_2,X_1]_r\bullet[TY_1,X_2]_\ell\right)\otimes \mathcal{C}(Y_1\otimes Y_2,T^{-1}X_3) \\
&\quad\cong\quad (\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C})\left(([Y_2,X_1]_r\bullet[TY_1,X_2]_\ell,Y_1\otimes Y_2),(Y_3,T^{-1}X_3)\right)
\end{aligned}$$

and

$$\begin{aligned}
R_0^*(Y,T^{-1}X) &\cong P_0(X)\cong\mathcal{C}(I,X)\cong\mathcal{C}(Y,K)\otimes\mathcal{C}(T^{-1}I,T^{-1}X) \\
&\cong (\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C})((K,I),(Y,T^{-1}X)).
\end{aligned}$$

so that $\mathcal{C}^{\mathrm{op}}\otimes\mathcal{C}$ is monoidal with unit (K,I) and tensor product

$$(X_1,Y_1)\otimes(X_2,Y_2)=([Y_2,X_1]_r\bullet[TY_1,X_2]_\ell,Y_1\otimes Y_2).$$

□

A particular case of Proposition 8.3 is the Chu construction of [3]. Here $\mathcal{V}$ is the category of sets with cartesian monoidal structure (although any cartesian closed base would do). Then every $\mathcal{V}$-category $\mathcal{C}$ is comonoidal via the diagonal functor Δ. The representability of this structure as required in Proposition 8.3 amounts to $\mathcal{C}$ having finite limits; so K is the terminal object and $B \bullet C = B \times C$ is the product of B and C. Then R^* is the $*$-autonomous monoidal structure on $\mathcal{C}^{\mathrm{op}} \otimes \mathcal{C}$ arising from any monoidal closed category $\mathcal{C}$ with finite products and a monoidal endoequivalence T.

However, the case of finite products for ordinary categories is not the only example where the representable comonoidal structure can be found. For any $\mathcal{V}$, such structure exists for example on any $\mathcal{C}$ which is a free $\mathcal{V}$-category on an ordinary category with finite products.

9 Star autonomy in monoidal bicategories

In order to exploit duality, we need to generalise the notion of star autonomy to pseudomonoids in a monoidal bicategory $\mathcal{B}$. The work of Sections 5 to 8 is a special case taking place in the autonomous monoidal bicategory $\mathcal{V}$-Mod of $\mathcal{V}$-categories and $\mathcal{V}$-modules as defined in Section 6.

As mentioned in Section 3, for pseudomonoids A and E in $\mathcal{B}$, where we write p and j for the multiplications and units, *a monoidal morphism* $g : A \to E$ is a morphism equipped with coherent 2-cells

$$\phi_2 : p \circ (g \otimes g) \Rightarrow g \circ p \text{ and } \phi_0 : j \Rightarrow g \circ j.$$

The morphism is called *strong monoidal* when ϕ_2 and ϕ_0 are invertible. When g has a left adjoint h, there are 2-cells

$$\phi_2^{\ell} : h \circ p \circ (1 \otimes g) \Rightarrow p \circ (h \otimes 1) \text{ and } \phi_2^{r} : h \circ p \circ (g \otimes 1) \Rightarrow p \circ (1 \otimes h)$$

obtained from ϕ_2 as mates under adjunction. We say g is *strong left [right] closed* when ϕ_2^{ℓ} [respectively, ϕ_2^{r}] is invertible; it is strong closed when it is both.

For a pseudomonoid A in $\mathcal{B}$, the category $\mathcal{B}(I, A)$ is monoidal with tensor product defined by

$$m \star n = p \circ (m \otimes n).$$

The internal homs, provided $\mathcal{B}$ has the relevant right liftings, are defined as follows: $[n, r]_{\ell}$ is the right lifting of r through $p \circ (1_A \otimes n)$ while $[m, r]_r$ is the right lifting of r through $p \circ (m \otimes 1_A)$.

Proposition 9.1 *If $g : A \to E$ is a strong monoidal morphism between pseudomonoids then $\mathcal{B}(I, g) : \mathcal{B}(I, A) \to \mathcal{B}(I, E)$ is a strong monoidal functor. If g has a left adjoint h and is strong closed then the functor $\mathcal{B}(I, g)$ is strong closed.*

Proof For the first sentence we have

$$\begin{aligned} \mathcal{B}(I,g)(m \star n) &= g \circ p \circ (m \otimes n) \cong p \circ (g \otimes g) \circ (m \otimes n) \\ &\cong p \circ (g \circ m) \otimes (g \circ n) \cong (g \circ m) \star (g \circ n) \\ &\cong \mathcal{B}(I,g)(m) \star \mathcal{B}(I,g)(n). \end{aligned}$$

For the second sentence consider the diagram

$$\begin{tikzcd} & I \arrow[dl, "{[gon,gor]_\ell}"'] \arrow[dr, "gor"] \arrow[drr, "r"] \arrow[d, phantom, "\Rightarrow"] \arrow[drr, phantom, shift right=4, "\Rightarrow"] & & \\ X \arrow[r, "1\otimes(gon)"'] & X \otimes X \arrow[r, "p"'] & X \arrow[r, "h"'] & A \end{tikzcd}.$$

The right-hand triangle is a right lifting since h is left adjoint to g. The left-hand triangle is a right lifting by definition of the left internal hom. So the outside triangle exhibits $[g \circ n, g \circ r]_\ell$ as a right lifting of r along the bottom composite. However, if g is strong left closed, the bottom composite is isomorphic to

$$h \circ p \circ (1 \otimes g) \circ (1 \otimes n) \cong p \circ (h \otimes 1) \circ (1 \otimes n) \cong p \circ (1 \otimes n) \circ h.$$

However, the right lifting of r through $p \circ (1 \otimes n)$ is $[n, r]_\ell$, and the right lifting of $[n, r]_\ell$ through h is $g \circ [n, r]_\ell$. So we have $g \circ [n, r]_\ell \cong [g \circ n, g \circ r]_\ell$ proving $\mathcal{B}(I, g)$ strong left closed. Right closedness is dual. □

A *form* for a pseudomonoid A in $\mathcal{B}$ is a morphism $\sigma : A \otimes A \to I$ together with an isomorphism

$$\begin{tikzcd} A^{\otimes 3} \arrow[r, "p \otimes 1_A"] \arrow[d, "1_A \otimes p"'] \arrow[dr, phantom, "\gamma \atop \cong"] & A^{\otimes 2} \arrow[d, "\sigma"] \\ A^{\otimes 2} \arrow[r, "\sigma"'] & I \end{tikzcd}$$

called the *form constraint.*

In the bicategories $\mathcal{B}$ that we have in mind there are special morphisms (as abstracted by Wood [40]). The special morphisms h have right adjoints h^* and, in some cases, are precisely the morphisms with right adjoints, sometimes called *maps* in $\mathcal{B}$. For example, in Mat(Set) the maps are precisely the matrices arising from functions, and these are the special morphisms we want. For Mod($\mathcal{V}$), the special morphisms are those modules isomorphic to h_* for some algebra morphism h. In the bicategory of $\mathcal{V}$-categories and $\mathcal{V}$-modules the special modules are those arising from $\mathcal{V}$-functors.

Suppose $\mathcal{B}$ has selected special maps and that $\mathcal{B}$ is autonomous. Each form $\sigma : A \otimes A \to I$ corresponds to a morphism $\hat{\sigma} : A \to A^0$. We say that the form σ is *representable* when $\hat{\sigma}$ is isomorphic to a special map. We say that σ is *non-degenerate* when $\hat{\sigma}$ is an equivalence.

A pseudomonoid in $\mathcal{B}$ is defined to be $*$-*autonomous* when it is equipped with a non-degenerate representable form. For example, for any object R of $\mathcal{B}$ and any equivalence $v : R \to R^{00}$ the canonical endohom pseudomonoid $R^e = R^0 \otimes R$ becomes $*$-autonomous when equipped with the form $\sigma : R^e \otimes R^e \to I$ defined by

$$\hat{\sigma} = 1_{R^0} \otimes v : R^e = R^0 \otimes R \to R^0 \otimes R^{00} = R^{e0}.$$

An opmorphism $h : E \to A$ between $*$-autonomous pseudomonoids is called $*$-autonomous when there is an isomorphism

$$\begin{tikzcd} E \otimes E \arrow[d, "h \otimes h"'] \arrow[dr, "\sigma"] \arrow[d, phantom, shift left=4ex, "\overset{\tau}{\Leftarrow}"] & \\ A \otimes A \arrow[r, "\sigma"'] & I \end{tikzcd}$$

such that the following equation holds:

$$
\begin{tikzcd}
E^{\otimes 3} \arrow[r, "p\otimes 1"] \arrow[dd, "h\otimes h\otimes h"'] \arrow[ddr, phantom, "\psi_2\otimes 1\Downarrow"] & E^{\otimes 2} \arrow[r, "\sigma"] \arrow[d, "h\otimes h"'] \arrow[r, phantom, shift right=4, "\tau\Downarrow"] & I \\
& A^{\otimes 2} \arrow[ur, "\sigma"] \arrow[d, phantom, "\cong\gamma"] & \\
A^{\otimes 3} \arrow[ur, "p\otimes 1"] \arrow[r, "1\otimes p"'] & A^{\otimes 2} \arrow[uur, "\sigma"'] &
\end{tikzcd}
\quad = \quad
\begin{tikzcd}
E^{\otimes 3} \arrow[r, "p\otimes 1"] \arrow[dr, "1\otimes p"'] \arrow[dd, "h\otimes h\otimes h"'] \arrow[ddr, phantom, "1\otimes\psi_2\Downarrow"] & E^{\otimes 2} \arrow[r, "\sigma"] \arrow[d, phantom, "\gamma\cong"] & I \\
& E^{\otimes 2} \arrow[ur, "\sigma"'] \arrow[d, "h\otimes h"'] \arrow[r, phantom, "\tau\Downarrow"] & \\
A^{\otimes 3} \arrow[r, "1\otimes p"'] & A^{\otimes 2} \arrow[uur, "\sigma"'] &
\end{tikzcd}
$$

We are particularly interested in opmorphisms h that are maps. Then the right adjoint h^* is a morphism of pseudomonoids. Under these circumstances we define h to be *strong $*$-autonomous* when the mate

$$\tau^\ell : \sigma \circ (h^* \otimes 1) \Rightarrow \sigma \circ (1 \otimes h)$$

of τ is invertible. It follows that $\tau^r : \sigma \circ (1 \otimes h^*) \Rightarrow \sigma \circ (h \otimes 1)$ is also invertible.

Proposition 9.2 *Suppose $h : E \to A$ is a strong $*$-autonomous special opmorphism between $*$-autonomous pseudomonoids in $\mathcal{B}$. If h^* is strong monoidal then h^* is strong closed.*

Proof We have the calculation

$$
\begin{aligned}
&\sigma \circ (h \otimes 1) \circ (p \otimes 1) \circ (1 \otimes h^* \otimes 1) \\
&\cong \ \sigma \circ (1 \otimes h^*) \circ (p \otimes 1) \circ (1 \otimes h^* \otimes 1) \cong \sigma \circ (p \otimes 1) \circ (1 \otimes h^*) \circ (1 \otimes h^* \otimes 1) \\
&\cong \ \sigma \circ (1 \otimes p) \circ (1 \otimes h^* \otimes h^*) \cong \sigma \circ (1 \otimes h^*) \circ (1 \otimes p) \\
&\cong \ \sigma \circ (h \otimes 1) \circ (1 \otimes p) \cong \sigma \circ (1 \otimes p) \circ (h \otimes 1 \otimes 1) \\
&\cong \ \sigma \circ (1 \otimes p) \circ (h \otimes 1 \otimes 1).
\end{aligned}
$$

It follows that $\hat{\sigma} \circ h \circ p \circ (1 \otimes h^*) \cong \hat{\sigma} \circ p \circ (h \otimes 1)$. Left strong closedness follows since σ is non-degenerate. Right closedness is dual. □

Motivated by Proposition 3.3, we define *basic data* in an autonomous monoidal bicategory $\mathcal{B}$ to consist of an object R equipped with a special opmorphism $h : R^0 \otimes R \to A$ into a pseudomonoid A such that h^* is strong monoidal. Here $R^0 \otimes R$ has the canonical endohom pseudomonoid structure. Suppose further that $R^e = R^0 \otimes R$ is $*$-autonomous via a form arising as above from an equivalence $v : R \to R^{00}$. The basic data is called *Hopf* when A is equipped with a $*$-autonomous structure and h is strong $*$-autonomous.

From basic data, by applying the pseudofunctor $\mathcal{B}(I, -) : \mathcal{B} \to \mathrm{Cat}$, we obtain an adjunction

$$\mathcal{B}(I, h) \dashv \mathcal{B}(I, h^*) : \mathcal{B}(I, A) \to \mathcal{B}(I, R^0 \otimes R)$$

which transports via the equivalence $\mathcal{B}(I, R^0 \otimes R) \xrightarrow{\sim} \mathcal{B}(R, R)$ to an adjunction between $\mathcal{B}(I, A)$ and $\mathcal{B}(R, R)$. The pseudomonoid structure on A induces a monoidal structure on $\mathcal{B}(I, A)$ and the canonical endohom pseudomonoidal structure on $R^0 \otimes R$ induces the monoidal structure on $\mathcal{B}(R, R)$ whose tensor product is composition in $\mathcal{B}$. Since h^* is strong monoidal, the right adjoint $\mathcal{B}(I, A) \to \mathcal{B}(R, R)$ is strong monoidal. By Propositions 9.1 and 9.2, this right adjoint is also strong closed in the Hopf case.

Since basic and Hopf basic data are expressible purely in terms of the monoidal bicategory structure and the special maps of $\mathcal{B}$, the next result is clear.

Proposition 9.3 *Strong monoidal pseudofunctors that preserve special maps also preserve basic and Hopf basic data.*

Remark 9.4 The day after we submitted this paper to the Fields Workshop organizers, the preprint [7] appeared on math.arXiv. We contacted Dr. Gabriella Böhm who pointed out that, in our original preprint, we had not been specific about the $*$-autonomous structure on $R^e = R^0 \otimes R$ in our definition of Hopf basic data. This was indeed an omission and we had in mind the symmetric case where we had the opportunity to take $R^{00} = R$ and $\hat{\sigma} : R^0 \otimes R \to R^0 \otimes R^{00}$ the identity.

10 Ordinary groupoids revisited

Let us return to the definition of ordinary category as formulated in Propositions 1.1 and 2.1. Let G be a monoidal comonad on the internal endohom pseudomonoid $X \times X$ in the monoidal bicategory Mat(Set). Recall that $G(x, y; u, v)$ is empty unless $x = u$ and $y = v$, and we put

$$\mathbf{A}(x, y) = G(x, y; x, y)$$

which defines the homsets of our category $\mathbf{A}$. Let A denote the set of arrows of the category $\mathbf{A}$; we have the triangle

$$\begin{tikzcd} X \times X \arrow[rr, "G"] & & X \times X \\ & A \arrow[ul, "(s,t)_*"] \arrow[ur, "(s,t)_*"'] \arrow[u, phantom, "\overset{\delta}{\Leftarrow}"] & \end{tikzcd}$$

which is the universal coaction of G on a morphism into $X \times X$; it is the Eilenberg-Moore construction for the comonad G. By a dual of Lemma 3.2, there is a pseudomonoid structure on A such that the whole triangle lifts to the Eilenberg-Moore construction in the bicategory MonMat(Set).

We already pointed out in the Introduction what the pseudomonoidal structure on A is; that on $X \times X$ is the special case of a chaotic category. Referring to the definition of basic data at the end of Section 9, we have:

Proposition 10.1 *An equivalent definition of ordinary small categories is that they are basic data in the autonomous monoidal bicategory Mat(Set)$^{\mathrm{co}}$ where the special morphisms are all the maps.*

Proof Reversing 2-cells interchanges left and right adjunctions. So for a morphism to have a right adjoint in Mat(Set)$^{\mathrm{co}}$ is to be a right adjoint in Mat(Set); that is, to be the reverse of a function. Basic data in Mat(Set)$^{\mathrm{co}}$ therefore consists of a set X , a pseudomonoid A in Mat(Set), and a function $(s, t) : A \to X \times X$ that is strong monoidal. The functor $\mathcal{B}(I, h^*)$ as at the end of Section 9 transports to the left-adjoint functor

$$\Sigma_{(s,t)} : \ \mathrm{Set}/A \to \ \mathrm{Set}/X \times X$$

of the Introduction, which by Section 9 is strong monoidal. So we have a category $\mathbf{A}$. Conversely, if $\mathbf{A}$ is a category, clearly (s, t) is strong monoidal. □

The discussion of the Introduction already shows that, if $\mathbf{A}$ is a groupoid, then it is $*$-autonomous in Mat(Set) with $Sa = a^{-1}$. In particular (the chaotic case), the endohom $X \times X$ is $*$-autonomous with $S(x, y) = (y, x)$. For $\mathbf{A}$ a groupoid,

$(s,t)^* : X \times X \to A$ is a strong $*$-autonomous map in $\mathrm{Mat(Set)}^{\mathrm{co}}$. So we have Hopf basic data in $\mathrm{Mat(Set)}^{\mathrm{co}}$. The converse almost holds.

Proposition 10.2 *Consider a category as basic data in Mat(Set)$^{\mathrm{co}}$. The category is a groupoid iff the basic data are Hopf.*

Proof The characterizing property of $S = S_\ell$ is that

$$b \circ a = Sc \qquad \text{iff} \qquad c \circ b = S^{-1}a.$$

For each object x, put $e_x = S1_x$. Taking $c = 1_x$ and $b = S^{-1}a$ to ensure $c \circ b = S^{-1}a$, we deduce that $S^{-1}a \circ a = e_x$ for all $a : x \to y$. Taking $a = 1_x$ we see that $e_x = S^{-1}1_x$ so $Se_x = 1_x$. Now go back to the characterizing property with $c = e_x$, b arbitrary, and $a = S(e_x \circ b)$ to ensure $c \circ b = S^{-1}a$: so we deduce that $b \circ S(e_x \circ b) = Se_x = 1_x$. It follows that every morphism b has a right inverse. So the category is a groupoid. □

Remark 10.3 (This arose in lunchtime conversation with John Baez and Isar Stubbe.) The operation S_ℓ of $*$-autonomy is not unique. For a groupoid $\mathbf{A}$ as we have been considering, we can choose any endomorphism e_x of each object x and define $S_\ell a = a^{-1} \circ e_x$ so that $S_r a = e_x \circ a^{-1}$. This defines another $*$-autonomous structure on our pseudomonoid A.

Remark 10.4 The argument of this section can be internalized to any finitely complete category $\mathcal{E}$. In particular, groupoids internal to $\mathcal{E}$ can be identified with Hopf basic data in the monoidal bicategory $\mathrm{Span}(\mathcal{E})^{\mathrm{co}}$. More details will be provided in Example 12.3.

11 Hopf bialgebroids

A bialgebroid A based on a k-algebra R is an opmonoidal monad on R^e in $\mathrm{Mod}(\mathcal{V})$ (see Section 3). We have already seen that A becomes a k-algebra and that $\eta^* : A \to R^e$ provides the Eilenberg-Moore object for the monad, thereby lifting to the bicategory of pseudomonoids in $\mathrm{Mod}(\mathcal{V})$.

In the terminology of Section 9, a bialgebroid is precisely basic data $\eta : R^e \to A$ in $\mathcal{B} = \mathrm{Mod}(\mathcal{V})$. We define a bialgebroid $\eta : R^e \to A$ to be *Hopf* when this basic data in $\mathrm{Mod}(\mathcal{V})$ is Hopf; that is, A should be $*$-autonomous and $\eta^* : A \to R^e$ should be strong $*$-autonomous. It follows from Section 9 that $\mathrm{Mod}(\mathcal{V})(k, \eta^*)$ is strong monoidal and strong closed; this is none other than the functor

$$\mathcal{V}^A \to \mathcal{V}^{R^e}$$

defined by restriction along $\eta : R^e \to A$; compare Proposition 7.5 in the case of one-object $\mathcal{V}$-categories.

Preservation of internal homs was taken as paramount in the Hopf algebroid notions of [15] and [31]. Example 7.4 explains the connection between our work here and that of [15] while we see from the last paragraph that our Hopf bialgebroids are more restrictive than the Hopf algebroids of [31].

Remark 11.1 In the correspondence mentioned in Remark 9.4 Dr. Böhm advised us that her notion of Hopf bialgebroid in [7] fits our setting, where $\mathcal{V}$ is the category of vector spaces, and that she has examples where the $*$-autonomous structure S_ℓ on $R^e = R^0 \otimes R$ is defined by

$$S_\ell(x \otimes y) = x \otimes u(y)$$

with u a non-identity k-algebra automorphism of R. This kind of perturbation fits well with our treatment of the Chu construction in Section 8.

Example 11.2 Let $\mathcal{V}$ continue to be the category of k-vector spaces and let $\mathcal{A}$ denote the category of commutative k-algebras. The category $\mathcal{A}$ is finitely cocomplete; the pushout of two morphisms out of an object A is given by tensoring over A the codomains of the two morphisms. Definition B.3.7 of [29] labels groupoids internal to $\mathcal{A}^{\mathrm{op}}$ as "Hopf algebroids" (generalizing the idea that a commutative Hopf algebra is exactly a group in $\mathcal{A}^{\mathrm{op}}$). In fact, these are examples of Hopf bialgebroids in our sense. To see this we make use of the strong monoidal pseudofunctor

$$\mathrm{Span}(\mathcal{A}^{\mathrm{op}})^{\mathrm{co}} \to \mathrm{Mod}(\mathcal{V})$$

which takes each commutative algebra A to itself as an algebra and each cospan C from A to B in $\mathcal{A}$ to C with actions of A and B coming from the morphisms into C. By Remark 10.4, each Ravenel "Hopf algebroid" is Hopf basic data in $\mathrm{Span}(\mathcal{A}^{\mathrm{op}})^{\mathrm{co}}$. Then, by Proposition 9.3 our pseudofunctor applies to give Hopf basic data in $\mathrm{Mod}(\mathcal{V})$; that is, to give a Hopf bialgebroid. We are grateful to Terry Bisson for pointing out the book [29] which features good examples of groupoids internal to $\mathcal{A}^{\mathrm{op}}$ occurring in algebraic topology.

12 Quantum categories and quantum groupoids

It remains to state the main definitions of the paper. We now have the motivation and concepts readily at hand.

Let $\mathcal{V}$ be a braided monoidal category with coreflexive equalizers (that is, equalizers of pairs of morphisms with a common left inverse). We begin by recalling the definition of the right autonomous monoidal bicategory $\mathrm{Comod}(\mathcal{V})$ as appearing in [14]. We assume the condition:

each of the functors $X \otimes - : \mathcal{V} \to \mathcal{V}$ preserves coreflexive equalizers.

Briefly, $\mathrm{Comod}(\mathcal{V}) = \mathrm{Mod}(\mathcal{V}^{\mathrm{op}})^{\mathrm{coop}}$. To make calculations we will need to make the definition more explicit.

The objects of $\mathrm{Comod}(\mathcal{V})$ are comonoids C in $\mathcal{V}$; the comultiplication and counit are denoted by $\delta : C \to C \otimes C$ and $\varepsilon : C \to I$. The hom-category $\mathrm{Comod}(\mathcal{V})(C, D)$ is the category of Eilenberg-Moore coalgebras for the comonad $C \otimes - \otimes D$ on the category $\mathcal{V}$. This implies that the morphisms $M : C \to D$ in $\mathrm{Comod}(\mathcal{V})$ are comodules from C to D; that is, left C-, right D-comodules. So M is an object of $\mathcal{V}$ together with a coaction $\delta : M \to C \otimes M \otimes D$ satisfying the expected equations. It is sometimes useful to deal with the left and right actions $\delta_\ell : M \to C \otimes M$ and $\delta_r : M \to M \otimes D$ which are obtained from δ using the counit. The 2-cells $f : M \Rightarrow M' : C \to D$ in $\mathrm{Comod}(\mathcal{V})$ are morphisms $f : M \to M'$ in $\mathcal{V}$ respecting the coactions.

Composition of comodules $M : C \to D$ and $N : D \to E$ is given by the equalizer

$$N \circ M = M \underset{D}{\otimes} N \to M \otimes N \underset{1\otimes\delta_\ell}{\overset{\delta_r \otimes 1}{\rightrightarrows}} M \otimes D \otimes N.$$

The identity comodule $C \to C$ is C with the obvious coaction. We point out that the pair of morphisms being equalized here have a common left inverse $1 \otimes \varepsilon \otimes 1$; so the equalizer is coreflexive.

The remaining details describing $\mathrm{Comod}(\mathcal{V})$ as a bicategory should now be clear.

Remark 12.1 (a) When $\mathcal{V}$ = Set, it is readily checked that Comod($\mathcal{V}$) is biequivalent to Mat(Set).

(b) The main case that should be kept in mind is when $\mathcal{V}$ is the category of vector spaces over a field k; then the objects of Comod($\mathcal{V}$) are precisely k-coalgebras.

(c) If $\mathcal{V}$ itself is a $*$-autonomous monoidal category then the distinction between Mod($\mathcal{V}$) and Comod($\mathcal{V}$) evaporates.

(d) By the Chu construction, any complete cocomplete closed monoidal $\mathcal{V}$ can be embedded into a complete cocomplete $*$-autonomous monoidal $\mathcal{E} = \mathcal{V}^{\mathrm{op}} \otimes \mathcal{V}$ taking V to $(1, V)$ where 1 is the terminal object of $\mathcal{V}$. The embedding is strong monoidal and preserves colimits and connected limits. So we can take full advantage of remark (c) by working in $\mathcal{E}$-Mod and deducing results for both Mod($\mathcal{V}$) and Comod($\mathcal{V}$).

Returning to general $\mathcal{V}$, we note that each comonoid morphism $f : C \to D$ determines a comodule $f_* : C \to D$ defined to be C together with the coaction

$$C \xrightarrow{\delta} C \otimes C \xrightarrow{\delta\otimes f} C \otimes C \otimes D,$$

and a comodule $f^* : D \to C$ defined to be C together with the coaction

$$C \xrightarrow{\delta} C \otimes C \xrightarrow{f\otimes\delta} D \otimes C \otimes C.$$

Notice that we have $\gamma_f : f_* \circ f^* \Rightarrow 1_D$ which is defined to be $f : C \to D$ since $f_* \circ f^* = f^* \underset{C}{\otimes} f_* = C$ with coaction $C \xrightarrow{\delta} C \otimes C \xrightarrow{1\otimes\delta} C \otimes C \otimes C \xrightarrow{f\otimes 1\otimes f} D \otimes C \otimes D$. Also, $f_* \underset{D}{\otimes} f^* = f^* \circ f_*$ is the equalizer

$$f_* \underset{D}{\otimes} f^* \to C \otimes C \underset{(C\otimes f\otimes C)\circ(C\otimes\delta)}{\overset{(C\otimes f\otimes C)\circ(\delta\otimes C)}{\rightrightarrows}} C \otimes D \otimes C \quad ;$$

and, since

$$C \longrightarrow C \otimes C \underset{C\otimes\delta}{\overset{\delta\otimes C}{\rightrightarrows}} C \otimes C \otimes C$$

is an (absolute) equalizer, we have a unique morphism $C \to f_* \underset{D}{\otimes} f^*$ commuting with the morphisms into $C \otimes C$; this gives us $\omega_f : 1_C \Rightarrow f^* \circ f_*$. Indeed, γ_f, ω_f are the counit and unit for an adjunction $f_* \dashv f^*$ in the bicategory Comod($\mathcal{V}$).

The comodules f^* provide the special maps for the bicategory Comod$(\mathcal{V})^{\mathrm{co}}$. Suppose C, D are comonoids. Then $C \otimes D$ becomes a comonoid with coaction

$$C \otimes D \xrightarrow{\delta\otimes\delta} C \otimes C \otimes D \otimes D \xrightarrow{C\otimes c_{C,D}\otimes D} C \otimes D \otimes C \otimes D$$

where c is the braiding and, as justified by coherence theorems, we ignore associativity in $\mathcal{V}$. For comodules $M : C \to C'$ and $N : D \to D'$, we obtain a comodule $M \otimes N : C \otimes D \to C' \otimes D'$ where the coaction is given in the obvious way using the braiding. This extends to a pseudofunctor $\otimes :$ Comod($\mathcal{V}$) $\times$ Comod($\mathcal{V}$) $\to$ Comod($\mathcal{V}$). The remaining structure required to obtain Comod($\mathcal{V}$) as a monoidal bicategory should be obvious.

Write C^0 for C with the comultiplication

$$C \xrightarrow{\delta} C \otimes C \xrightarrow{c_{C,C}} C \otimes C.$$

There is a pseudonatural equivalence between the category of comodules $M : C \otimes D \to E$. and the category of comodules $\hat{M} : D \to C^0 \otimes E$, where $M = \hat{M}$ as objects. It follows that C^0 is a right bidual for C. This defines the structure of right autonomous monoidal bicategory on $\mathrm{Comod}(\mathcal{V})$.

Each $C^0 \otimes C$ has the canonical structure of a pseudomonoid in $\mathrm{Comod}(\mathcal{V})$ because it is an endohom in the autonomous monoidal bicategory.

A *quantum category* in $\mathcal{V}$ is basic data $C, h : C^0 \otimes C \to A$ in $\mathrm{Comod}(\mathcal{V})^{\mathrm{co}}$.

A *quantum groupoid* in $\mathcal{V}$ is Hopf basic data in $\mathrm{Comod}(\mathcal{V})^{\mathrm{co}}$.

Our referee has sensibly recommended that we unpackage these definitions for the utility of the reader and for comparison with the definition of "bicoalgebroid" in [8].

A *quantum graph* $\mathbf{A}$ in $\mathcal{V}$ consists of

- a comonoid C, called the *object object* of $\mathbf{A}$,
- a comonoid A, called the *arrow object* of $\mathbf{A}$, and
- comonoid morphisms $s : A \to C^0$ and $t : A \to C$, called *source* and *target morphisms* of $\mathbf{A}$,

such that the following diagram commutes.

$$\begin{tikzcd}
A \arrow[r, "\delta"] \arrow[d, "\delta"'] & A \otimes A \arrow[rr, "t \otimes s"] & & C \otimes C \\
A \otimes A \arrow[r, "s \otimes t"'] & C \otimes C \arrow[urr, "c_{C,C}"'] & &
\end{tikzcd}$$

It follows that $r : A \xrightarrow{\delta} A \otimes A \xrightarrow{s \otimes t} C^0 \otimes C$ is a comonoid morphism. Therefore we have a comodule $I \xrightarrow{\varepsilon^*} A \xrightarrow{r_*} C^0 \otimes C$ which corresponds, under $C \dashv C^0$, to a comodule $C \to C$; explicitly, it is $A : C \to C$ with coactions

$$\delta_\ell : A \xrightarrow{\delta} A \otimes A \xrightarrow{1 \otimes s} A \otimes C \xrightarrow{c_{C,A}^{-1}} C \otimes A$$

$$\delta_r : A \xrightarrow{\delta} A \otimes A \xrightarrow{1 \otimes t} A \otimes C.$$

Then we can define the *composable pairs object* $P = A \underset{C}{\otimes} A$ as the composite comodule $C \xrightarrow{A} C \xrightarrow{A} C$; explicitly, it is the equalizer

$$\begin{tikzcd}
P \arrow[r, "\iota"] & A \otimes A \arrow[r, shift left, "\delta_r \otimes 1"] \arrow[r, shift right, "1 \otimes \delta_\ell"'] & A \otimes C \otimes A
\end{tikzcd}$$

which becomes a comodule $P : C \to C$ via right and left coactions induced by

$$A \otimes A \xrightarrow{1 \otimes \delta_r} A \otimes A \otimes C \quad \text{and} \quad A \otimes A \xrightarrow{\delta_\ell \otimes 1} C \otimes A \otimes A.$$

Although in general P is not a comonoid with ι a comonoid morphism, there is a unique morphism $\delta_\ell : P \to A \otimes A \otimes P$ such that the following diagram commutes.

$$\begin{tikzcd}
P \arrow[r, "\iota"] \arrow[d, "\delta_\ell"'] & A \otimes A \arrow[r, "\delta \otimes \delta"] & A \otimes A \otimes A \otimes A \arrow[d, "1 \otimes c_{A,A} \otimes 1"] \\
A \otimes A \otimes P \arrow[rr, "1 \otimes 1 \otimes \iota"'] & & A \otimes A \otimes A
\end{tikzcd}$$

This is because the diagram

$$P \xrightarrow{\iota} A \otimes A \xrightarrow{\delta \otimes \delta} A \otimes A \otimes A \otimes A \xrightarrow{1 \otimes c_{A,A} \otimes 1} A \otimes A \otimes A \otimes A \underset{1 \otimes 1 \otimes 1 \otimes \delta_\ell}{\overset{1 \otimes 1 \otimes \delta_r \otimes 1}{\rightrightarrows}} A \otimes A \otimes A \otimes C \otimes A$$

commutes, and $1 \otimes 1 \otimes \iota$ is the equalizer of $1 \otimes 1 \otimes \delta_r \otimes 1$ and $1 \otimes 1 \otimes 1 \otimes \delta_\ell$. A small calculation (four steps using string diagrams) proves that $\delta_\ell : P \to A \otimes A \otimes P$ is a left coaction of the comonoid $A \otimes A$ on P.

A *composition morphism* for a quantum graph $\mathbf{A}$ is a comodule morphism

$$\mu : P \to A : C \to C$$

that satisfies the axioms CM0, CM1 and CM2 below.

CM0. $\mu : A \underset{C}{\otimes} A \to A$ is associative in the monoidal category $\mathrm{Comod}(\mathcal{V})(C, C)$.

CM1. The following diagram commutes:

$$P \xrightarrow{\delta_\ell} A \otimes A \otimes P \underset{\varepsilon \otimes s \otimes 1}{\overset{t \otimes \varepsilon \otimes 1}{\rightrightarrows}} C \otimes P \xrightarrow{1 \otimes \mu} C \otimes A.$$

Before stating CM2 we need to notice, using CM1, that there exists a unique morphism $\delta_r : P \to P \otimes A$ such that the following diagram commutes.

$$\begin{tikzcd} P \arrow[r, "\delta_\ell"] \arrow[d, "\delta_r"'] & A \otimes A \otimes P \arrow[d, "1 \otimes 1 \otimes \mu"] \\ P \otimes A \arrow[r, "\iota \otimes 1"'] & A \otimes A \otimes A \end{tikzcd}$$

This is because the diagram

$$P \xrightarrow{\delta_\ell} A \otimes A \otimes P \xrightarrow{1 \otimes 1 \otimes \mu} A \otimes A \otimes A \underset{1 \otimes \delta_\ell \otimes 1}{\overset{\delta_r \otimes 1 \otimes 1}{\rightrightarrows}} A \otimes C \otimes A \otimes A$$

commutes, and $\iota \otimes 1$ is the equalizer of $\delta_r \otimes 1 \otimes 1$ and $1 \otimes \delta_\ell \otimes 1$. Now we can state:

CM2. The following diagram commutes.

$$\begin{tikzcd} P \arrow[r, "\mu"] \arrow[d, "\delta_r"'] & A \arrow[d, "\delta"] \\ P \otimes A \arrow[r, "\mu \otimes 1"'] & A \otimes A \end{tikzcd}$$

It can now be shown that $P : A \otimes A \to A$ is a comodule with coactions δ_ℓ and δ_r as above.

An *identities morphism* for $\mathbf{A}$ is a comodule morphism $\eta : C \to A : C \to C$ satisfying the axioms

IM0. η is a unit for μ in $\mathrm{Comod}(\mathcal{V})(C, C)$.

IM1. The following diagram commutes.

$$\begin{tikzcd} C \arrow[rr, "\eta"] \arrow[dr, "\varepsilon"'] & & A \arrow[dl, "\varepsilon"] \\ & I & \end{tikzcd}$$

IM2. The following diagram commutes.

$$\begin{tikzcd}
 & & & C \otimes A \arrow[dr, "\eta\otimes 1"] & \\
C \arrow[r, "\eta"] & A \arrow[r, "\delta"] & A \otimes A \arrow[ur, "t\otimes 1"] \arrow[dr, "s\otimes 1"'] \arrow[rr, "1_{A\otimes A}"] & & A \otimes A \\
 & & & C \otimes A \arrow[ur, "\eta\otimes 1"'] &
\end{tikzcd}$$

It follows that A becomes a pseudomonoid in $\mathrm{Comod}(\mathcal{V})$ when equipped with the multiplication P, the unit $J = \eta_*$, and the canonical associativity and unit constraints. Furthermore, $r_* : A \to C^0 \otimes C$ becomes strong monoidal.

Notice that we obtain a morphism $\varsigma : P \to C \otimes C \otimes C$ by taking either of the equal routes in the diagram

$$\begin{tikzcd}
P \arrow[r, "\iota"] & A \otimes A \arrow[r, shift left, "\delta_r \otimes 1"] \arrow[r, shift right, "1\otimes\delta_\ell"'] & A \otimes C \otimes A \arrow[r, "s\otimes 1\otimes t"] & C \otimes C \otimes C.
\end{tikzcd}$$

A quantum category is the same as a quantum graph equipped with a composition morphism and an identities morphism. The basic data in $\mathrm{Comod}(\mathcal{V})^{\mathrm{co}}$ is the comodule $r_* : A \to C^0 \otimes C$.

When $\mathcal{V}$ is the monoidal category of vector spaces over a field k, our quantum graph corresponds to BC1 of [8] while our axioms CM0-CM2 amount to BC2 of [8] and our axioms IM0-IM2 amount to BC3 of [8].

The *chaotic quantum category* $\mathbf{A} = \mathbf{C}_{\mathrm{ch}}$ *on* C is defined by $A = C^0 \otimes C$, $s = 1_{C^0} \otimes \varepsilon$ and $t = \varepsilon \otimes 1_C$. Thus $P = C \otimes C \otimes C$ with $\iota = 1_C \otimes \delta \otimes 1_C$, $\delta_r = 1_C \otimes 1_C \otimes \delta$ and $\delta_\ell = \delta \otimes 1_C \otimes 1_C$. Finally, $\mu = 1_C \otimes \varepsilon \otimes 1_C$, $\eta = \delta$ and $\varsigma = 1_{C\otimes C\otimes C}$.

A quantum groupoid is a quantum category $\mathbf{A}$ equipped with comonoid equivalences

$$v : C \to C^{00} \qquad \text{and} \qquad \nu : A \to A^0$$

such that $s\nu \cong t$ and $t\nu \cong vs$, and for which there is a left $A \otimes A \otimes A$-comodule isomorphism $\gamma : P_\ell \cong P_r$, where P_ℓ is P with the left coaction

$$P \xrightarrow{\ \delta\ } A \otimes A \otimes P \otimes A \xrightarrow{1\otimes 1\otimes 1\otimes \nu} A \otimes A \otimes P \otimes A \xrightarrow{1\otimes 1\otimes c_{P,A}} A \otimes A \otimes P$$

and P_r is P with the left coaction

$$P \xrightarrow{\ \delta\ } A \otimes A \otimes P \otimes A \xrightarrow{1\otimes 1\otimes 1\otimes \nu'} A \otimes A \otimes P \otimes A \xrightarrow{c_{A\otimes A\otimes P,A}} A \otimes A \otimes P$$

in which ν' is an inverse equivalence for ν and δ is the coaction associated with the comodule $P : A \otimes A \to A$, such that the following square commutes.

$$\begin{tikzcd}
P \arrow[r, "\varsigma"] \arrow[d, "\gamma"'] & C \otimes C \otimes C \arrow[d, "c_{C,C\otimes C}"] \\
P \arrow[r, "\varsigma"'] & C \otimes C \otimes C
\end{tikzcd}$$

Example 12.2 Let $\mathcal{V}$ be the symmetric monoidal category of vector spaces over a field k. For any set X, let FX be the vector space with X as basis. This F is the object function for a strong monoidal functor $F : \mathrm{Set} \to \mathcal{V}$ that preserves coreflexive equalizers (exercise!). It therefore induces a strong monoidal pseudofunctor

$$\bar{F} : \ \mathrm{Comod}(\mathrm{Set})^{\mathrm{co}} \to \ \mathrm{Comod}(\mathcal{V})^{\mathrm{co}}.$$

Special maps are preserved by $\bar{F}$. It follows from Proposition 9.3 that $\bar{F}$ takes each category to a quantum category and each groupoid to a quantum groupoid.

Example 12.3 Following up on Remark 10.4 where $\mathcal{E}$ is a category with finite limits, we shall lead the reader into showing how quantum categories and quantum groupoids in $\mathcal{V} = \mathcal{E}$ (where the tensor product is cartesian product) are precisely categories and groupoids in $\mathcal{E}$. Every object of $\mathcal{E}$ has a unique comonoid structure defined by the diagonal morphism, every morphism of $\mathcal{E}$ is a comonoid morphism, and the only 2-cells between morphisms are equalities. Also each object C has $C^0 = C$. So a quantum graph $\mathbf{A}$ in $\mathcal{E}$ is just a pair of morphisms

$$s, t : A \to C;$$

that is, $\mathbf{A}$ is a (directed) graph in $\mathcal{E}$. The equalizer $P = A \underset{C}{\otimes} A$ is now easily seen to be the pullback of s and t; that is, P is the usual object of composable pairs in the graph. A composition morphism μ and an identities morphism η are precisely what is required to make $\mathbf{A}$ a category in $\mathcal{E}$. If $\mathbf{A}$ is a quantum groupoid then, because of the absence of 2-cells, $v : C \to C$ and $\nu : A \to A$ are isomorphisms while $s\nu = t$ and $t\nu = vs$. Arguing as for Proposition 10.2, we see that v is actually the identity and ν makes $\mathbf{A}$ a groupoid in $\mathcal{E}$.

References

[1] M. Barr, $*$-Autonomous categories, with an appendix by Po Hsiang Chu. Lecture Notes in Mathematics 752 (Springer, Berlin, 1979).

[2] M. Barr, Nonsymmetric $*$-autonomous categories, Theoretical Computer Science 139 (1995) 115-130.

[3] M. Barr, The Chu construction, Theory Appl. Categories 2 (1996) 17-35.

[4] M. A. Batanin, Homotopy coherent category theory and A_∞-structures in monoidal categories, J. Pure Appl. Algebra 123 (1998) 67-103.

[5] J. Bénabou, Introduction to bicategories, Lecture Notes in Math. 47 (Springer, Berlin, 1967) 1- 77.

[6] R. Betti, A. Carboni, R. Street and R. Walters, Variation through enrichment, J. Pure Appl. Algebra 29 (1983) 109-127.

[7] G. Böhm, An alternative notion of Hopf algebroid, arXiv:math.QA/0301169 (16 Jan 2003).

[8] T. Brzezinski and G. Militaru, Bialgebroids, $\times_R$-bialgebras and duality, arXiv:math.QA/0012164 v2 (7 Nov 2001).

[9] B.J. Day, On closed categories of functors, Lecture Notes in Math. 137 (Springer- Verlag, Berlin 1970) 1-38.

[10] B.J. Day, Note on monoidal monads, J. Austral. Math. Soc. Ser. A 23 (1977) 292-311.

[11] B.J. Day, Promonoidal functor categories, J. Austral. Math. Soc. Ser. A 23 (1977) 312-328.

[12] B.J. Day, $*$-autonomous convolution (Talk to the Australian Category Seminar on 5 March 1999).

[13] B.J. Day and G.M. Kelly, Enriched functor categories, Lecture Notes in Math. 106 (Springer, Berlin, 1969) 178-191.

[14] B.J. Day, P. McCrudden and R. Street, Dualizations and antipodes, Appl. Categorical Structures 11 (2003) 229-260.

[15] B.J. Day and R. Street, Monoidal bicategories and Hopf algebroids, Advances in Math. 129 (1997) 99-157.

[16] B.J. Day and R. Street, Lax monoids, pseudo-operads, and convolution, in: "Diagrammatic Morphisms and Applications", Contemporary Mathematics 318 (AMS; April 2003) 75-96.

[17] B.J. Day and R. Street, Abstract substitution in enriched categories, J. Pure Appl. Algebra 179 (2003) 49-63.

[18] S. Eilenberg and G.M. Kelly, Closed categories, Proceedings of the Conference on Categorical Algebra at La Jolla (Springer, 1966) 421-562.

[19] J.-Y. Girard, Linear logic, Theoretical Computer Science 50 (1987) 1-102.

[20] A. Joyal and R. Street, Braided tensor categories, Advances in Math. 102 (1993) 20-78.

[21] A. Joyal and R. Street, An introduction to Tannaka duality and quantum groups, Lecture Notes in Math. 1488 (Springer-Verlag Berlin, Heidelberg 1991) 411-492.
[22] G.M. Kelly, Doctrinal adjunction, Lecture Notes in Math. 420 (Springer, Berlin, 1974) 257-280.
[23] G.M. Kelly, Basic Concepts of Enriched Category Theory, London Math. Soc. Lecture Notes Series 64 (Cambridge University Press 1982).
[24] J.-H. Lu, Hopf algebroids and quantum groupoids, Int. J. Math. 7 (1996) 47-70.
[25] S. Mac Lane, Categories for the Working Mathematician, Graduate Texts in Math. 5 (Springer-Verlag, Berlin 1971).
[26] G. Maltsiniotis, Groupoï des quantiques, Comptes Rendus Acad. Sci. Paris 314, Série I (1992) 249-252.
[27] P. McCrudden, Opmonoidal monads, Theory Appl. Categories 10 (2002) 469-485.
[28] I. Moerdijk, Monads on tensor categories, J. Pure Appl. Algebra 168 (2002) 189-208.
[29] D.C. Ravenal, Nilpotence and periodicity in stable homotopy theory, Annals of Mathematics Studies 128 (Princeton University Press 1992).
[30] P. Schauenburg, Bialgebras over non-commutative rings and a structure theorem for Hopf bimodules, Applied Categorical Structures 6 (1998) 193-222.
[31] P. Schauenburg, Duals and doubles of quantum groupoids ($\times_R$-Hopf algebras), Contemporary Math. 267 (2000) 273-299.
[32] R.A.G. Seely, Linear logic, $*$-autonomous categories and cofree coalgebras, Contemporary Math. 92 (Amer. Math. Soc., Providence, RI, 1989) 371-382.
[33] R. Street, The formal theory of monads, J. Pure Appl. Algebra 2 (1972) 149-168.
[34] R. Street, Cauchy characterization of enriched categories, Rendiconti del Seminario Matematico e Fisico di Milano 51 (1981) 217-233.
[35] R. Street, Quantum Groups: an entrée to modern algebra, <http://www-texdev.mpce.mq.edu.au/Quantum/Quantum.html>
[36] R. Street, Review of [3], Math. Rev. 97f:18004 (1997).
[37] K. Szlachányi, The monoidal Eilenberg-Moore construction and bialgebroids, J. Pure Appl. Algebra 182 (2003) 287-315.
[38] M. Takeuchi, Groups of algebras over $A \otimes \bar{A}$, J. Math. Soc. Japan 29 (1977) 459-492.
[39] C. Watts, Intrinsic characterizations of some additive functors, Proc. American Math. Soc. 11 (1960) 5-8.
[40] R.J. Wood, Abstract proarrows I, Cahiers Topologie Géom. Différentielle 23 (1982) 279-290.
[41] P. Xu, Quantum groupoids, Commun. Math. Phys. 216 (2001) 539-581.
[42] N. Yoneda, On Ext and exact sequences, J. Fac. Sci. Univ. Tokyo Sect. I 8 (1960) 507-576.

Received January 16, 2003; in revised form September 2,2003

Fields Institute Communications
Volume 43, 2004

Morphisms of 2-groupoids and Low-dimensional Cohomology of Crossed Modules

John W. Duskin
Department of Mathematics
244 Mathematics Building
State University of New York at Buffalo
Buffalo, NY 14260-2900 U.S.A.
duskin@math.buffalo.edu

Rudger W. Kieboom
Department of Mathematics
Vrije Universiteit Brussel
Pleinlaan 2 F 10
1050 Brussels, Belgium
rkieboom@vub.ac.be

Enrico M. Vitale
Département de Mathématique
Université catholique de Louvain
Chemin du Cyclotron 2
1348 Louvain-la-Neuve, Belgium
vitale@math.ucl.ac.be

Abstract. Given a morphism $P\colon \mathcal{G} \to \mathcal{H}$ of 2-groupoids, we construct a 6-term 2-exact sequence of cat-groups and pointed groupoids. We use this sequence to obtain an analogue for cat-groups (and, in particular, for crossed modules) of the fundamental exact sequence of non-abelian group cohomology. The link with simplicial topology is also explained.

Introduction

The aim of this paper is to obtain a basic result in low-dimensional cohomology of crossed modules. Homology and cohomology of crossed modules have been studied extensively, and a satisfactory theory has been developed (see [7] and the references therein, [14, 19, 20, 26]). The existing literature on this subject considers crossed modules and their morphisms as a category. Our point of view is that crossed modules are in a natural way the objects of a 2-category, and therefore they should be studied in a 2-dimensional context. This different point of view leads to

2000 *Mathematics Subject Classification.* Primary 18G50; Secondary 18B40, 18D05, 18D35, 20L05.
Third author supported by FNRS grant 1.5.116.01.

a choice of limits and colimits which are the natural ones in our 2-categorical setting, that is bilimits, but which do not have a universal property in the underlying category of crossed modules. Accordingly, the notions of exactness and extension we consider are not equivalent to those studied in the previous papers devoted to this subject.

The result we look for to test our theory is a generalization to crossed modules of the fundamental exact sequence in non-abelian cohomology of groups [25]. To get this result, we adapt to crossed modules the method developed by Brown in [5], where the fundamental sequence is obtained as a special case of an exact sequence associated to a fibration of groupoids. In fact, to follow in a more transparent way the analogy with groups, we work with cat-groups instead of crossed modules, since the 2-category of (strict and small) cat-groups is biequivalent to the 2-category of crossed modules [6, 23].

The paper is organized as follows. In the first section we recall, for the reader's convenience and in a way convenient to be generalized to 2-groupoids, the result due to Brown. Section 2 is devoted to the construction of a 6-term 2-exact sequence of strict cat-groups and pointed groupoids from any morphism of 2-groupoids. For basic facts on cat-groups and 2-exact sequences we refer to [18, 27] and the bibliography therein; we recall in Section 2 the definitions we need. The idea of an higher-dimensional version of Brown's exact sequence comes from the paper [17] by Hardie, Kamps and the second author. The precise link between the main result of [17] and our 2-exact sequence is explained in Remark 2.7. In the third section we fix a cat-group $\mathbb{G}$ and an extension

$$\mathbb{A} \xrightarrow{i} \mathbb{B} \xrightarrow{j} \mathbb{C}$$

of $\mathbb{G}$-cat-groups. From such an extension we obtain, as a particular case of the sequence in Section 2, a 6-term 2-exact sequence of cat-groups and pointed groupoids

$$\mathbb{A}^{\mathbb{G}} \to \mathbb{B}^{\mathbb{G}} \to \mathbb{C}^{\mathbb{G}} \to H^1(\mathbb{G},\mathbb{A}) \to H^1(\mathbb{G},\mathbb{B}) \to H^1(\mathbb{G},\mathbb{C})$$

which is the 2-dimensional generalization of the fundamental sequence in non-abelian group cohomology. As a corollary of the main result of Section 2, we also get a 9-term exact sequence of groups and pointed sets. Instead of exploiting the homological notion of 2-exactness, this 9-term sequence can also be obtained using classical results from simplicial topology. This is the content of Section 4.

1 Brown's exact sequence

As in the topological case, it is better to work with the homotopy fibre instead of the "set-theoretical" fibre. In this way, we can obtain an exact sequence from any functor between groupoids (and not only from a fibration). Moreover, we avoid some choices which would be quite hard to handle in the higher dimensional analogue developed in Section 2.

Recall that a groupoid $\mathbb{G}$ is a category in which each arrow is an isomorphism. Consider now a functor between groupoids

$$P\colon \mathbb{G} \to \mathbb{H}$$

and fix an object H in $\mathbb{H}$. The homotopy fibre $\mathbb{F}_H$ of P at the point H is the following comma groupoid:

- objects of $\mathbb{F}_H$ are pairs $(Y, y\colon P(Y) \to H)$, with Y an object of $\mathbb{G}$ and y an arrow in $\mathbb{H}$;

- an arrow $f\colon (Y_1, y_1) \to (Y_2, y_2)$ in $\mathbb{F}_H$ is an arrow $f\colon Y_1 \to Y_2$ in $\mathbb{G}$ such that $P(f) \cdot y_2 = y_1$ (composition denoted from left to right).

There is an obvious faithful functor $j\colon \mathbb{F}_H \to \mathbb{G}$. If X is an object of $\mathbb{G}$, we write $\mathbb{F}_X$ for $\mathbb{F}_{P(X)}$.

Now fix an object X in $\mathbb{G}$ and consider the following groups and pointed sets:

- $\pi_0(\mathbb{G})$, the set of isomorphism classes of objects of $\mathbb{G}$, pointed by the class of X; $\pi_0(\mathbb{H})$, pointed by the class of $P(X)$; $\pi_0(\mathbb{F}_X)$, pointed by the class of $(X, 1_{P(X)})$;
- $\mathbb{G}(X) = \mathbb{G}(X, X)$, the group of automorphisms of the object X in $\mathbb{G}$; $\mathbb{H}(X) = \mathbb{H}(P(X), P(X))$, the group of automorphisms of the object $P(X)$ in $\mathbb{H}$; $\mathbb{F}_X(X) = \mathbb{F}_X((X, 1_{P(X)}), (X, 1_{P(X)}))$, the group of automorphisms of $(X, 1_{P(X)})$ in $\mathbb{F}_X$.

They can be connected by the following morphisms (square brackets are isomorphism classes):

- $j_X\colon \mathbb{F}_X(X) \to \mathbb{G}(X) \quad j_X(f\colon X \to X) = f$;
- $P_X\colon \mathbb{G}(X) \to \mathbb{H}(X) \quad P_X(f\colon X \to X) = P(f)$;
- $\pi_0(P)\colon \pi_0(\mathbb{G}) \to \pi_0(\mathbb{H}) \quad [X] \mapsto [P(X)]$;
- $\pi_0(j)\colon \pi_0(\mathbb{F}_X) \to \pi_0(\mathbb{G}) \quad [Y, y\colon P(Y) \to P(X)] \mapsto [Y]$;
- $\delta\colon \mathbb{H}(X) \to \pi_0(\mathbb{F}_X) \quad \delta(x\colon P(X) \to P(X)) = [X, x\colon P(X) \to P(X)]$.

Proposition 1.1 *With the previous notations, the sequence*

$$0 \to \mathbb{F}_X(X) \xrightarrow{j_X} \mathbb{G}(X) \xrightarrow{P_X} \mathbb{H}(X) \xrightarrow{\delta} \pi_0(\mathbb{F}_X) \xrightarrow{\pi_0(j)} \pi_0(\mathbb{G}) \xrightarrow{\pi_0(P)} \pi_0(\mathbb{H})$$

is exact.

Proof Consider an element $[Y, y]$ in $\pi_0(\mathbb{F}_X)$ and assume that $[Y] = [X]$ in $\pi_0(\mathbb{G})$. Then there is an arrow $y'\colon Y \to X$ in $\mathbb{G}$ and therefore $[Y, y] = [X, P(y')^{-1} \cdot y] = \delta(P(y')^{-1} \cdot y)$ because $y'\colon (Y, y) \to (X, P(y')^{-1} \cdot y)$ is an arrow in $\mathbb{F}_X$. The rest of the proof is straightforward. □

Now consider the strict fibre $\mathbb{S}_H$ of P at the point H:

- objects of $\mathbb{S}_H$ are the objects Y of $\mathbb{G}$ such that $P(Y) = H$;
- an arrow $f\colon Y_1 \to Y_2$ of $\mathbb{G}$ is in $\mathbb{S}_H$ if $P(f) = 1_H$.

There is, for each object H in $\mathbb{S}_H$, a full and faithful functor $i_H\colon \mathbb{S}_H \to \mathbb{F}_H$. Clearly, P is a fibration of groupoids [5] if and only if for each H the functor i_H is essentially surjective on objects. Therefore, if P is a fibration, we can replace $\mathbb{F}_X(X)$ and $\pi_0(\mathbb{F}_X)$ by $\mathbb{S}_X(X)$ and $\pi_0(\mathbb{S}_X)$ and we obtain Brown's exact sequence associated to a fibration of groupoids (Theorem 4.3 in [5]).

2 The 2-exact sequence

In this section we fix a morphism of 2-groupoids

$$P\colon \mathcal{G} \to \mathcal{H}$$

that is a 2-functor between 2-categories in which each arrow is an equivalence and each 2-cell is an isomorphism.

Fix an object H in $\mathcal{H}$; the homotopy fibre $\mathcal{F}_H$ of P at the point H is the following 2-groupoid:

- objects are pairs $(Y, y\colon P(Y) \to H)$, with Y an object in $\mathcal{G}$ and y an arrow in $\mathcal{H}$;

- arrows $(f,\varphi)\colon (Y_1,y_1) \to (Y_2,y_2)$ are pairs with $f\colon Y_1 \to Y_2$ an arrow in $\mathcal{G}$ and $\varphi\colon y_1 \Rightarrow P(f)\cdot y_2\colon P(Y_1) \to H$ a 2-cell in $\mathcal{H}$;
- a 2-cell $\alpha\colon (f,\varphi) \Rightarrow (g,\psi)\colon (Y_1,y_1) \to (Y_2,y_2)$ is a 2-cell $\alpha\colon f \Rightarrow g$ in $\mathcal{G}$ such that the following diagram commutes

$$\begin{tikzcd} P(f)\cdot y_2 \arrow[rr, Rightarrow, "P(\alpha)\cdot y_2"] & & P(g)\cdot y_2 \\ & y_1 \arrow[ul, Rightarrow, "\varphi"] \arrow[ur, Rightarrow, "\psi"'] & \end{tikzcd}$$

There is a morphism $j\colon \mathcal{F}_H \to \mathcal{G}$ which sends $\alpha\colon (f,\varphi) \Rightarrow (g,\psi)\colon (Y_1,y_1) \to (Y_2,y_2)$ to $\alpha\colon f \Rightarrow g\colon Y_1 \to Y_2$. The morphism j is faithful on arrows and on 2-cells. If X is an object of $\mathcal{G}$, we write $\mathcal{F}_X$ for $\mathcal{F}_{P(X)}$.

We recall now the notion of 2-exact sequence for pointed groupoids (and, in particular, for cat-groups, i.e. monoidal groupoids in which each object is invertible, up to isomorphisms, w.r.t. the tensor product). Morphisms of pointed groupoids (cat-groups) are pointed functors (monoidal functors). A natural transformation between pointed (monoidal) functors is always assumed to be pointed (monoidal). Let $F\colon \mathbb{G} \to \mathbb{H}$ be a morphism of pointed groupoids; its homotopy kernel $kF\colon KerF \to \mathbb{G}$ is the homotopy fibre (in the sense of Section 1) of F on the base object I of $\mathbb{H}$. There is a natural transformation $\kappa F\colon kF\cdot F \Rightarrow 0$ (0 is the morphism which sends each arrow to the identity of I) given, for each object (Y,y) of $KerF$, by $y\colon F(Y) \to I$.

$$\begin{tikzcd} & \mathbb{G} \arrow[dr, "F"] \arrow[d, phantom, "\kappa F\Downarrow"] & \\ KerF \arrow[ur, "kF"] \arrow[rr, "0"'] & {} & \mathbb{H} \end{tikzcd} \qquad \begin{tikzcd} & \mathbb{G} \arrow[dr, "F"] \arrow[d, phantom, "\varphi\Downarrow"] & \\ \mathbb{K} \arrow[ur, "G"] \arrow[rr, "0"'] & {} & \mathbb{H} \end{tikzcd}$$

Moreover, given a pointed groupoid $\mathbb{K}$, a morphism G and a natural transformation φ as in the previous diagram, there is a unique comparison morphism $G'\colon \mathbb{K} \to KerF$, $G'(g\colon A_1 \to A_2) = G(g)\colon (G(A_1),\varphi_{A_1}) \to (G(A_2),\varphi_{A_2})$, such that $G'\cdot kF = G$ and $G'\cdot \kappa F = \varphi$ (compare with [15]). The universal property of $(KerF, kF, \kappa F)$ as a bilimit, discussed in [18, 27], determines it uniquely, up to equivalence.

Definition 2.1 Consider two morphisms G, F and a natural transformation φ of pointed groupoids as in the previous diagram; we say that the triple (G,φ,F) is 2-exact if the comparison $G'\colon \mathbb{K} \to KerF$ is full and essentially surjective on objects.

Now come back to the 2-functor between 2-groupoids $P\colon \mathcal{G} \to \mathcal{H}$ and fix an object X of $\mathcal{G}$. We can consider the following three hom-categories, which are in fact strict cat-groups: $\mathcal{G}(X) = \mathcal{G}(X,X)$, $\mathcal{H}(X) = \mathcal{H}(P(X),P(X))$ and $\mathcal{F}_X(X) = \mathcal{F}_X((X,1_{P(X)}),(X,1_{P(X)}))$. Moreover, we can consider the classifying groupoid $cl(\mathcal{G})$ of the 2-groupoid $\mathcal{G}$: $cl(\mathcal{G})$ has the same objects as $\mathcal{G}$ and 2-isomorphism classes of arrows of $\mathcal{G}$ as arrows. The groupoid $cl(\mathcal{G})$ is pointed by the object X. Similarly, we have the groupoid $cl(\mathcal{H})$ pointed by $P(X)$ and the groupoid $cl(\mathcal{F}_X)$ pointed by $(X,1_{P(X)})$. These cat-groups and pointed groupoids can be connected by the following morphisms (square brackets are 2-isomorphism classes of arrows):

- $j_X \colon \mathcal{F}_X(X) \to \mathcal{G}(X)$
  $\alpha \colon (f,\varphi) \Rightarrow (g,\psi) \colon (X,1_{P(X)}) \to (X,1_{P(X)}) \;\mapsto\; \alpha \colon f \Rightarrow g \colon X \to X$
- $P_X \colon \mathcal{G}(X) \to \mathcal{H}(X)$
  $\alpha \colon f \Rightarrow g \colon X \to X \;\mapsto\; P(\alpha) \colon P(f) \Rightarrow P(g) \colon P(X) \to P(X)$
- $cl(j) \colon cl(\mathcal{F}_X) \to cl(\mathcal{H})$
  $[f,\varphi] \colon (Y_1,y_1) \to (Y_2,y_2) \;\mapsto\; [f] \colon Y_1 \to Y_2$
- $cl(P) \colon cl(\mathcal{G}) \to cl(\mathcal{H})$
  $[f] \colon Y_1 \to Y_2 \;\mapsto\; [P(f)] \colon P(Y_1) \to P(Y_2)$
- $\delta \colon \mathcal{H}(X) \to cl(\mathcal{F}_X)$
  $\beta \colon h \Rightarrow k \colon P(X) \to P(X) \;\mapsto\; [1_X,\beta] \colon (X,h) \to (X,k)$.

Proposition 2.2 *With the previous notations, the sequence*

$$\mathcal{F}_X(X) \xrightarrow{j_X} \mathcal{G}(X) \xrightarrow{P_X} \mathcal{H}(X) \xrightarrow{\delta} cl(\mathcal{F}_X) \xrightarrow{cl(j)} cl(\mathcal{G}) \xrightarrow{cl(P)} cl(\mathcal{H})$$

with the obvious natural transformations $j_X \cdot P_X \Rightarrow 0$, $P_X \cdot \delta \Rightarrow 0$, $\delta \cdot cl(j) \Rightarrow 0$, $cl(j) \cdot cl(P) \Rightarrow 0$, *is 2-exact.*

Proof 1) 2-exactness in $\mathcal{G}(X)$: it is straightforward to verify that the functor $j_X \colon \mathcal{F}_X(X) \to \mathcal{G}(X)$ is exactly the kernel of $P_X \colon \mathcal{G}(X) \to \mathcal{H}(X)$.

2) 2-exactness in $\mathcal{H}(X)$: consider the comparison $P'_X \colon \mathcal{G}(X) \to Ker\delta$
$\alpha \colon f \Rightarrow g \colon X \to X \;\mapsto\; P(\alpha) \colon (P(f),[f,1_{P(f)}]) \Rightarrow (P(g),[g,1_{P(g)}])$
- P'_X is essentially surjective: given an object $(h,[f,\varphi])$ in $Ker\delta$, we obtain an arrow $\varphi \colon (h,[f,\varphi]) \Rightarrow \delta(f)$ in $Ker\delta$;
- P'_X is full: given an arrow $\beta \colon \delta(f) \Rightarrow \delta(g)$ in $Ker\delta$, then $\delta(\beta) \cdot [g,1_{P(g)}] = [f,1_{P(f)}]$, but this means that there exists a 2-cell $\alpha \colon f \Rightarrow g$ such that $P(\alpha) = \beta$.

3) 2-exactness in $cl(\mathcal{G})$: consider the comparison $j' \colon cl(\mathcal{F}_X) \to Ker(cl(P))$
$[f,\varphi] \colon (Y_1,y_1) \to (Y_2,y_2) \;\mapsto\; [f] \colon (Y_1,[y_1]) \to (Y_2,[y_2])$
- j' is essentially surjective: obvious;
- j' is full: let $[f] \colon j'(Y_1,y_1) \to j'(Y_2,y_2)$ be an arrow in $Ker(cl(P))$, this means that there exists a 2-cell $\varphi \colon y_1 \Rightarrow P(f) \cdot y_2$ and then $[f,\varphi] \colon (Y_1,y_1) \to (Y_2,y_2)$ is an arrow in $cl(\mathcal{F}_X)$.

4) 2-exactness in $cl(\mathcal{F}_X)$: consider the comparison $\delta' \colon \mathcal{H}(X) \to Ker(cl(j))$
$\beta \colon h \Rightarrow k \colon P(X) \to P(X) \;\mapsto\; [1_X,\beta] \colon (X,h,[1_X]) \to (X,k,[1_X])$
- δ' is full: let $[f,\varphi] \colon \delta'(h) \to \delta'(k)$ be an arrow in $Ker(cl(j))$, then $[f] \cdot [1_X] = [1_X]$, that is there exists a 2-cell $\alpha \colon 1_X \Rightarrow f$. We obtain $\beta \colon h \Rightarrow k$ in $\mathcal{H}(X)$ in the following way :

$$\beta = (h \xRightarrow{\varphi} P(f) \cdot k \xRightarrow{P(\alpha)^{-1} \cdot k} k)\,;$$

- δ' is essentially surjective: consider an object

$$(Y, y \colon P(Y) \to P(X), [x] \colon Y \to X)$$

in $Ker(cl(j))$, then $P(x)^{-1} \cdot y \colon P(X) \to P(X)$ is an object in $\mathcal{H}(X)$ and $[x,c] \colon (Y,y,[x]) \to \delta'(P(x)^{-1} \cdot y)$ is an arrow in $Ker(cl(j))$, where c is the canonical 2-cell $c \colon y \Rightarrow P(x) \cdot P(x)^{-1} \cdot y$.

□

As in Section 1, if $(\mathbb{G}, I)$ is a pointed groupoid (a cat-group), we write $\pi_0(\mathbb{G})$ for the pointed set (the group) of isomorphism classes of objects and $\pi_1(\mathbb{G})$ for the (abelian) group of automorphisms $\mathbb{G}(I,I)$. π_0 and π_1 extend to morphisms and carry

2-exact sequences on exact sequences of pointed sets or groups. Finally, observe that if $\mathcal{G}$ is a 2-groupoid and X is a chosen object in $\mathcal{G}$, then $\pi_1(cl(\mathcal{G})) = \pi_0(\mathcal{G}(X))$. In a similar way, if $P\colon \mathcal{G} \to \mathcal{H}$ is a 2-functor, then $\pi_1(cl(P)) = \pi_0(P_X)$.

Corollary 2.3 *Let $P\colon \mathcal{G} \to \mathcal{H}$ be a 2-functor between 2-groupoids and fix an object X in $\mathcal{G}$; the following is an exact sequence of groups and pointed sets (the last three terms)*

$$0 \to \pi_1(\mathcal{F}_X(X)) \to \pi_1(\mathcal{G}(X)) \to \pi_1(\mathcal{H}(X)) \to$$

$$\pi_1(cl(\mathcal{F}_X)) = \pi_0(\mathcal{F}_X(X)) \to \pi_1(cl(\mathcal{G})) = \pi_0(\mathcal{G}(X)) \to \pi_1(cl(\mathcal{H})) = \pi_0(\mathcal{H}(X))$$

$$\to \pi_0(cl(\mathcal{F}_X)) \to \pi_0(cl(\mathcal{G})) \to \pi_0(cl(\mathcal{H}))\,.$$

Proof As far as exactness in $\pi_1(\mathcal{F}_X(X))$ is concerned, observe that j_X is the kernel of P_X, so it is faithful and then $\pi_1(j_X)$ is injective. The rest follows from the 2-exactness of the sequence in Proposition 2.2 and the previous remarks on π_0 and π_1. □

Remark 2.4 If $P\colon \mathbb{G} \to \mathbb{H}$ is a functor between groupoids, we can look at it as a 2-functor between discrete 2-groupoids (2-groupoids with no non-trivial 2-cells). The exact sequence of Corollary 2.3 reduces then to the exact sequence of Proposition 1.1, because the first non-trivial term is $\pi_0(\mathcal{F}_X(X))$.

Remark 2.5 Brown's exact sequence of Proposition 1.1 satisfies a strong exactness condition in $\pi_0(\mathbb{F}_X)$, which is the transition point between groups and pointed groupoids. The 2-dimensional analogue of strong exactness has been formulated in [16]. It is not difficult to prove that the sequence of Proposition 2.2 is strongly 2-exact in $cl(\mathcal{F}_X)$, let us just observe that the needed action $\mathcal{H}(X)\times cl(\mathcal{F}_X) \to cl(\mathcal{F}_X)$ sends $(f\colon P(X) \to P(X), (Y, y\colon P(Y) \to P(X)))$ into $(Y, y\cdot f\colon P(Y) \to P(X) \to P(X))$.

Remark 2.6 Proposition 2.2 and Corollary 2.3 hold also for a morphism $P\colon \mathcal{G} \to \mathcal{H}$ of bigroupoids, that is a pseudo-functor between bicategories [2, 3] in which each arrow is an equivalence and each 2-cell is an isomorphism. The generalization is strightforward: just observe that the homotopy fibre $\mathcal{F}_H$ inherits a structure of bicategory from that of $\mathcal{G}$. Clearly, if $\mathcal{G}$ and $\mathcal{H}$ are bigroupoids, the cat-groups $\mathcal{F}_X(X), \mathcal{G}(X)$ and $\mathcal{H}(X)$ of Proposition 2.2 are no longer strict. In Section 3 we will use this more general version of Proposition 2.2.

Remark 2.7 Recall that a morphism of bigroupoids $P\colon \mathcal{G} \to \mathcal{H}$ is a fibration if the functor $cl(P)\colon cl(\mathcal{G}) \to cl(\mathcal{H})$ is a fibration of groupoids and for each Y_1, Y_2 in $\mathcal{G}$ the functor $P_{Y_1,Y_2}\colon \mathcal{G}(Y_1, Y_2) \to \mathcal{H}(P(Y_1), P(Y_2))$ is a fibration of groupoids [17, 22]. This is equivalent to ask that for each object X in $\mathcal{G}$, the induced functor $St_P(X)\colon St_{\mathcal{G}}(X) \to St_{\mathcal{H}}(P(X))$ (where the "star-groupoid" $St_{\mathcal{G}}(X)$ is the groupoid having morphisms $y\colon X \to Y$ as objects and 2-cells $\alpha\colon y_1 \Rightarrow y_2\colon X \to Y$ as arrows) is an essentially surjective fibration. When P is a fibration, one easily checks that for each object H of $\mathcal{H}$, the homotopy fibre $\mathcal{F}_H$ is biequivalent to the strict fibre $\mathcal{S}_H$ (i.e. the sub-bigroupoid of $\mathcal{G}$ having as 2-cells the 2-cells $\alpha\colon f \Rightarrow g\colon Y_1 \to Y_2$ such that $P(\alpha)$ is the identity 2-cell of 1_H). Therefore, if P is a fibration, $\mathcal{F}_X(X)$ and $cl(\mathcal{F}_X)$ are equivalent to $\mathcal{S}_X(X)$ and $cl(\mathcal{S}_X)$ and the sequence of Corollary 2.3 is exactly the Hardie-Kamps-Kieboom 9-term exact sequence associated to a fibration of bigroupoids (Theorem 2.4 in [17]).

Remark 2.8 Proposition 2.2 can be also used to construct a Picard-Brauer 2-exact sequence from a homomorphism of unital commutative rings. In fact such a morphism induces a pseudo-functor between the bigroupoids having Azumaya algebras as objects, invertible bimodules as arrows and bimodule isomorphisms as 2-cells. Compare with [27], where a similar 2-exact sequence is obtained using homotopy cokernels instead of homotopy fibres.

3 The cohomology sequence

Let us fix a cat-group $\mathbb{G}$. A $\mathbb{G}$-cat-group is a pair $(\mathbb{C},\gamma)$ where $\mathbb{C}$ is a cat-group and $\gamma\colon \mathbb{G} \to Aut\mathbb{C}$ is a monoidal functor with codomain the cat-group of monoidal auto-equivalences of $\mathbb{C}$. $\mathbb{G}$-cat-groups are the objects of a 2-category, having equivariant monoidal functors as arrows and compatible monoidal transformations as 2-cells (see [9, 12] for more details and for an equivalent definition of $\mathbb{G}$-cat-group in terms of an action $\mathbb{G}\times\mathbb{C} \to \mathbb{C}$). Observe that homotopy kernels in the 2-category of $\mathbb{G}$-cat-groups are computed as in the 2-category of cat-groups (in other words, if $j\colon (\mathbb{B},\beta) \to (\mathbb{C},\gamma)$ is a morphism of $\mathbb{G}$-cat-groups and $i\colon \mathbb{A} \to \mathbb{B}$ is its kernel as a morphism of cat-groups, then $\mathbb{A}$ inherits from $\mathbb{B}$ a structure $\alpha\colon \mathbb{G} \to Aut\mathbb{A}$ of $\mathbb{G}$-cat-group such that $i\colon \mathbb{A} \to \mathbb{B}$ is a morphism of $\mathbb{G}$-cat-groups).

If $(\mathbb{C},\gamma)$ is a $\mathbb{G}$-cat-group, a *derivation* is a pair $\langle M\colon \mathbb{G} \to \mathbb{C}, \mathbf{m}\rangle$ where M is a functor and

$$\mathbf{m} = \{m_{X,Y}\colon M(X)\otimes\gamma(X)(M(Y)) \to M(X\otimes Y)\}_{X,Y\in\mathbb{G}}$$

is a natural family of coherent isomorphisms (for more details, see [13], where $\mathbb{C}$ is assumed to be braided, or [11], where $\mathbb{G}$ is discrete). Derivations are the objects of a bigroupoid $\mathcal{Z}^1(\mathbb{G},\mathbb{C})$:

- an arrow is a pair $\langle C,\mathbf{c}\rangle\colon \langle M,\mathbf{m}\rangle \to \langle N,\mathbf{n}\rangle$ with $C\in\mathbb{C}$ and
$$\mathbf{c} = \{c_X\colon M(X)\otimes\gamma(X)(C) \to C\otimes N(X)\}_{X\in\mathbb{G}}$$
is a natural family of isomorphisms, compatible with $\mathbf{m}$ and $\mathbf{n}$;
- a 2-cell $f\colon \langle C,\mathbf{c}\rangle \Rightarrow \langle C',\mathbf{c}'\rangle$ is an arrow $f\colon C\to C'$ in $\mathbb{C}$ compatible with $\mathbf{c}$ and $\mathbf{c}'$.

In $\mathcal{Z}^1(\mathbb{G},\mathbb{C})$ there is a trivial derivation

$$\theta_{\mathbb{C}} = \langle 0\colon \mathbb{G}\to\mathbb{C}, \{I\otimes\gamma(X)(I)\simeq I\}\rangle$$

and the cat-group $\mathcal{Z}^1(\mathbb{G},\mathbb{C})(\theta_{\mathbb{C}},\theta_{\mathbb{C}})$ is the cat-group $\mathbb{C}^{\mathbb{G}}$ of $\mathbb{G}$-*invariant objects.* Explicitly:

- an object of $\mathbb{C}^{\mathbb{G}}$ is a pair $\langle C,\mathbf{c}\rangle$ with $C\in\mathbb{C}$ and
$$\mathbf{c} = \{c_X\colon \gamma(X)(C)\to C\}_{X\in\mathbb{G}}$$
a natural family of isomorphisms compatible with the monoidal structure of $\mathbb{G}$;
- an arrow $f\colon \langle C,\mathbf{c}\rangle \to \langle D,\mathbf{d}\rangle$ in $\mathbb{C}^{\mathbb{G}}$ is an arrow $f\colon C\to D$ in $\mathbb{C}$ such that the following diagram commutes for each $X\in\mathbb{G}$

$$\begin{tikzcd}
\gamma(X)(C) \arrow[r, "c_X"] \arrow[d, "\gamma(X)(f)"'] & C \arrow[d, "f"] \\
\gamma(X)(D) \arrow[r, "d_X"'] & D
\end{tikzcd}$$

A morphism $j\colon (\mathbb{B},\beta) \to (\mathbb{C},\gamma)$ of $\mathbb{G}$-cat-groups induces a pseudo-functor

$$j_*\colon \mathcal{Z}^1(\mathbb{G},\mathbb{B}) \to \mathcal{Z}^1(\mathbb{G},\mathbb{C}).$$

This pseudo-functor j_* sends a derivation $\langle H\colon \mathbb{G} \to \mathbb{B}, \mathbf{h}\rangle$ into the derivation $\langle H \cdot j\colon \mathbb{G} \to \mathbb{B} \to \mathbb{C}, j(\mathbf{h})\rangle$, where $j(\mathbf{h})$ is defined by the following composition

$$\begin{array}{c} j(H(X)) \otimes \gamma(X)(j(H(Y))) \\ \big\downarrow{\scriptstyle\simeq} \\ j(H(X)) \otimes j(\beta(X)(H(Y))) \\ \big\downarrow{\scriptstyle\simeq} \\ j(H(X) \otimes \beta(X)(H(Y))) \\ \big\downarrow{\scriptstyle j(h_{X,Y})} \\ j(H(X \otimes Y)) \end{array}$$

(the first isomorphism is the equivariant structure of j, the second one is its monoidal structure) and is defined in an obvious way on arrows and 2-cells.

In the next lemma we need the homotopy fibre $\mathbb{F}$ of j_* at the point $\theta_{\mathbb{B}}$. Let us describe explicitly the objects of $\mathbb{F}$ (without loosing in generality, we can assume that $j(I) = I$, so that $j_*(\theta_{\mathbb{B}}) = \theta_{\mathbb{C}}$). An object of $\mathbb{F}$ is a 4-tuple

$$\langle\, \mathcal{D} = \langle H\colon \mathbb{G} \to \mathbb{B}, \mathbf{h}\rangle \in \mathcal{Z}^1(\mathbb{G},\mathbb{B}),\ \langle \overline{H} \in \mathbb{C}, \overline{\mathbf{h}}\rangle \in \mathcal{Z}^1(\mathbb{G},\mathbb{C})(j_*(\mathcal{D}), \theta_{\mathbb{C}})\,\rangle$$

with

$$\mathbf{h} = \{h_{X,Y}\colon H(X) \otimes \beta(X)(H(Y)) \to H(X \otimes Y)\}_{X,Y \in \mathbf{G}}$$
$$\overline{\mathbf{h}} = \{\overline{h}_X\colon j(H(X)) \otimes \gamma(X)(\overline{H}) \to \overline{H}\}_{X \in \mathbf{G}}$$

Lemma 3.1 *Consider an essentially surjective morphism $j\colon \mathbb{B} \to \mathbb{C}$ of $\mathbb{G}$-cat-groups and its homotopy kernel*

$$\begin{tikzcd} & \mathbb{B} \arrow[dr, "j"] & \\ \mathbb{A} = Ker j \arrow[ur, "i"] \arrow[rr, "0"'] & \arrow[u, phantom, "\kappa j \Downarrow"] & \mathbb{C} \end{tikzcd}$$

The homotopy fibre $\mathbb{F}$ of $j_*\colon \mathcal{Z}^1(\mathbb{G},\mathbb{B}) \to \mathcal{Z}^1(\mathbb{G},\mathbb{C})$ at the point $\theta_{\mathbb{B}}$ is biequivalent to the bigroupoid of derivations $\mathcal{Z}^1(\mathbb{G},\mathbb{A})$.

Proof Given an object in $\mathcal{Z}^1(\mathbb{G},\mathbb{A})$

$$\langle F\colon \mathbb{G} \to \mathbb{A}, \mathbf{f} = \{f_{X,Y}\colon F(X) \otimes \alpha(X)(F(Y)) \to F(X \otimes Y)\}_{X,Y \in \mathbf{G}}\rangle$$

we get an object in $\mathbb{F}$

$$\langle\, \langle F \cdot i\colon \mathbb{G} \to \mathbb{A} \to \mathbb{B}, i(\mathbf{f})\rangle, \langle I \in \mathbb{C}, \{\kappa j_{F(X)}\colon j(i(F(X))) \to I\}_{X \in \mathbf{G}}\rangle\,\rangle$$

This construction extends to a 2-functor $\epsilon\colon \mathcal{Z}^1(\mathbb{G},\mathbb{A}) \to \mathbb{F}$ which is always locally an equivalence (even if $j\colon \mathbb{B} \to \mathbb{C}$ is not essentially surjective). Let us check that ϵ is surjective on objects up to equivalence. Let $\langle\, \langle H\colon \mathbb{G} \to \mathbb{B}, \mathbf{h}\rangle, \langle \overline{H} \in \mathbb{C}, \overline{\mathbf{h}}\rangle\,\rangle$ be an object of $\mathbb{F}$. Since $j\colon \mathbb{B} \to \mathbb{C}$ is essentially surjective, there is an object $Z \in \mathbb{B}$ and an arrow $z\colon \overline{H} \to j(Z)$. Now we can construct a functor

$$D\colon \mathbb{G} \to \mathbb{B} \quad X \mapsto Z^* \otimes H(X) \otimes \beta(X)(Z)$$

(Z^* is a dual of Z in the cat-group $\mathbb{B}$) which has a structure of derivation $\mathbf{d} = \{d_{X,Y}\colon D(X)\otimes\beta(X)(D(Y)) \to D(X\otimes Y)\}$ obtained from that of H in the following way

$$\begin{array}{c}
D(X) \otimes \beta(X)(D(Y)) \\
\Big\downarrow{=} \\
Z^* \otimes H(X) \otimes \beta(X)(Z) \otimes \beta(X)(Z^* \otimes H(Y) \otimes \beta(Y)(Z)) \\
\Big\downarrow{\simeq} \\
Z^* \otimes H(X) \otimes \beta(X)(Z) \otimes \beta(X)(Z^*) \otimes \beta(X)(H(Y) \otimes \beta(Y)(Z)) \\
\Big\downarrow{\simeq} \\
Z^* \otimes H(X) \otimes \beta(X)(H(Y)) \otimes \beta(X)(\beta(Y)(Z)) \\
\Big\downarrow{\simeq} \\
Z^* \otimes H(X) \otimes \beta(X)(H(Y)) \otimes \beta(X \otimes Y)(Z) \\
\Big\downarrow{1\otimes h_{X,Y}\otimes 1} \\
Z^* \otimes H(X \otimes Y) \otimes \beta(X \otimes Y)(Z) = D(X \otimes Y)
\end{array}$$

Observe now that the functor $D\colon \mathbb{G} \to \mathbb{B}$ factors through the kernel of j. Indeed, if $X \in \mathbb{G}$, we have

$$\begin{array}{c}
j(D(X)) = j(Z^* \otimes H(X) \otimes \beta(X)(Z)) \\
\Big\downarrow{\simeq} \\
j(Z^*) \otimes j(H(X)) \otimes j(\beta(X)(Z)) \\
\Big\downarrow{\simeq} \\
j(Z^*) \otimes j(H(X)) \otimes \gamma(X)(j(Z)) \\
\Big\downarrow{z^*\otimes 1\otimes z^{-1}} \\
\overline{H}^* \otimes j(H(X)) \otimes \gamma(X)(\overline{H}) \\
\Big\downarrow{1\otimes \overline{h}_X} \\
\overline{H}^* \otimes \overline{H} \simeq I
\end{array}$$

Let us call $\widetilde{D}\colon \mathbb{G} \to \mathbb{A}$ the factorization of $D\colon \mathbb{G} \to \mathbb{B}$ through the kernel $\mathbb{A}$. The structure $\mathbf{d}$ of the derivation D pass to $\widetilde{D}$ because $\overline{\mathbf{h}}$ is compatible with $j(\mathbf{h})$ and with the structure of the trivial derivation $\theta_{\mathbb{C}}$. In this way, we have built up an object $\langle \widetilde{D}\colon \mathbb{G} \to \mathbb{A}, \tilde{\mathbf{d}}\rangle$ of $\mathcal{Z}^1(\mathbb{G}, \mathbb{A})$. Finally, an arrow

$$\langle\, \langle H\colon \mathbb{G} \to \mathbb{B}, \mathbf{h}\rangle,\, \langle \overline{H} \in \mathbb{C}, \overline{\mathbf{h}}\rangle \,\rangle \to \epsilon\langle \widetilde{D}\colon \mathbb{G} \to \mathbb{A}, \tilde{\mathbf{d}}\rangle$$

in $\mathbb{F}$ is provided by $Z \in \mathbb{B}$, $z\colon \overline{H} \to j(Z)$ and by the family of canonical isomorphisms

$$\{H(X) \otimes \beta(X)(Z) \simeq Z \otimes Z^* \otimes H(X) \otimes \beta(X)(Z) = Z \otimes D(X)\}_{X\in\mathbb{G}}$$

□

An essentially surjective morphism with its homotopy kernel

$$\begin{tikzcd} & \mathbb{B} \arrow[dr, "j"] & \\ \mathbb{A} \arrow[ur, "i"] \arrow[rr, "0"'] & \arrow[u, phantom, "\kappa j\Downarrow"] & \mathbb{C} \end{tikzcd}$$

is called in [4, 8, 24] an *extension.* Putting together Proposition 2.2 and the previous lemma, we get our generalization of the fundamental sequence in non-abelian group cohomology. We write $H^1(\mathbb{G},\mathbb{C})$ for $cl(\mathcal{Z}^1(\mathbb{G},\mathbb{C}))$.

Corollary 3.2 *Consider an extension of* $\mathbb{G}$*-cat-groups*

$$\begin{tikzcd} & \mathbb{B} \arrow[dr, "j"] & \\ \mathbb{A} \arrow[ur, "i"] \arrow[rr, "0"'] & \arrow[u, phantom, "\kappa j\Downarrow"] & \mathbb{C} \end{tikzcd}$$

There is a 2-exact sequence of cat-groups and pointed groupoids

$$\mathbb{A}^{\mathbb{G}} \to \mathbb{B}^{\mathbb{G}} \to \mathbb{C}^{\mathbb{G}} \to H^1(\mathbb{G},\mathbb{A}) \to H^1(\mathbb{G},\mathbb{B}) \to H^1(\mathbb{G},\mathbb{C})\,.$$

Remark 3.3 Observe that if the cat-groups $\mathbb{G},\mathbb{B},\mathbb{C}$ and the monoidal functor $j\colon \mathbb{B} \to \mathbb{C}$ are strict, then $j_*\colon \mathcal{Z}^1(\mathbb{G},\mathbb{B}) \to \mathcal{Z}^1(\mathbb{G},\mathbb{C})$ is a 2-functor between 2-groupoids, and the cat-groups involved in the previous corollary are strict, that is they are crossed modules.

To end, we sketch an equivalent description of the bigroupoid $\mathcal{Z}^1(\mathbb{G},\mathbb{C})$ of derivations using the semi-direct product. Let us start with a general construction: if $\mathcal{G}$ and $\mathcal{H}$ are bicategories, $[\mathcal{G},\mathcal{H}]$ is the bicategory of pseudo-functors $\mathcal{G} \to \mathcal{H}$, pseudo-natural transformations and modifications [2, 3]. If $\mathbb{G}$ and $\mathbb{H}$ are cat-groups, we can see them as bicategories with only one object, and $[\mathbb{G},\mathbb{H}]$ is now a bigroupoid. Explicitly:

- an object of $[\mathbb{G},\mathbb{H}]$ is a monoidal functor $F\colon \mathbb{G} \to \mathbb{H}$;
- an arrow $(H,\varphi)\colon F \to G\colon \mathbb{G} \to \mathbb{H}$ is an object H of $\mathbb{H}$ and a natural transformation

$$\begin{tikzcd} & \mathbb{H} \arrow[dr, "-\otimes H"] & \\ \mathbb{G} \arrow[ur, "F"] \arrow[dr, "G"'] & \varphi\Downarrow & \mathbb{H} \\ & \mathbb{H} \arrow[ur, "H\otimes -"'] & \end{tikzcd}$$

making commutative the following diagrams

$$\begin{tikzcd} F(X)\otimes F(Y)\otimes H \arrow[r, "1\otimes\varphi_Y"] \arrow[d, "\simeq"'] & F(X)\otimes H\otimes G(Y) \arrow[r, "\varphi_X\otimes 1"] & H\otimes G(X)\otimes G(Y) \arrow[d, "\simeq"] \\ F(X\otimes Y)\otimes H \arrow[rr, "\varphi_{X\otimes Y}"'] & & H\otimes G(X\otimes Y) \end{tikzcd}$$

$$
\begin{tikzcd}
F(I) \otimes H \arrow[d, "\varphi_I"'] & I \otimes H \simeq H \arrow[l, "\simeq"'] \arrow[d, "1"] \\
H \otimes G(I) & H \otimes I \simeq H \arrow[l, "\simeq"]
\end{tikzcd}
$$

Observe that composition of arrows and parallel composition of 2-cells are defined using the tensor product in $\mathbb{H}$. Finally, a morphism of cat-groups $P\colon \mathbb{H} \to \mathbb{K}$ induces a pseudo-functor $P_*\colon [\mathbb{G}, \mathbb{H}] \to [\mathbb{G}, \mathbb{K}]$.

Now we apply the previous construction to a particular case. Consider a cat-group $\mathbb{G}$ and a $\mathbb{G}$-cat-group $(\mathbb{C}, \gamma\colon \mathbb{G} \to Aut\mathbb{C})$. Following [12], we can construct the semi-direct product $\mathbb{G} \times_\gamma \mathbb{C}$ together with the projection $P\colon \mathbb{G} \times_\gamma \mathbb{C} \to \mathbb{G}$, which is a monoidal functor. Therefore, we have a pseudo-functor

$$P_*\colon [\mathbb{G}, \mathbb{G} \times_\gamma \mathbb{C}] \to [\mathbb{G}, \mathbb{G}]$$

and it is possible to construct a biequivalence from $\mathcal{Z}^1(\mathbb{G}, \mathbb{C})$ to the homotopy fibre of P_* at the point $Id_{\mathbb{G}} \in [\mathbb{G}, \mathbb{G}]$. This biequivalence is another way to formulate the universal property of the semi-direct product studied in [13].

4 The simplicial topological point of view

As with the original paper [5] of Brown, we are motivated by homotopy theory and the classical exact sequences of homotopy groups and pointed sets which occur there, although in the previous sections the linkage to 2-types of topological spaces is far in the background. Indeed, because our bicategorical notion of 2-exact sequence, the presentation has been more homological/algebraic in feeling.

Nevertheless, the link to simplicial homotopy theory as pioneered by Daniel Kan, John Moore, and John Milnor in the fifties is quite direct. To briefly review the relevant parts of this theory [1], recall that as observed by Kan, the property of the singular complex of a topological space which permits one to combinatorially define *all* of the homotopy groups at any base point is that the singular complex has a simple simplicial horn-lifting property which makes it a “Kan complex”, and that a corresponding “Kan fibration” property (that corresponds essentially to the lifting properties of a fibration of spaces) is all that is needed to associate to a pointed simplicial Kan fibration f with fiber F :

$$F \subset E \longrightarrow B$$

[1]For more detail see [21], [10], or [1]

a long exact sequence of pointed sets, groups and abelian groups:

$$\begin{tikzcd}
\pi_n(F) \arrow[r, "\pi_n(i)"] & \pi_n(E) \arrow[r, "\pi_n(f)"] & \pi_n(B) \arrow[dll, dotted, no head] \\
\pi_2(F) \arrow[r, "\pi_2(i)"'] & \pi_2(E) \arrow[r, "\pi_2(f)"] & \pi_2(B) \arrow[dll, "\delta"'] \\
\pi_1(F) \arrow[r, "\pi_1(i)"'] & \pi_1(E) \arrow[r, "\pi_1(f)"] & \pi_1(B) \arrow[dll, "\delta"'] \\
\pi_0(F) \arrow[r, "\pi_0(i)"'] & \pi_0(E) \arrow[r, "\pi_0(f)"] & \pi_0(B).
\end{tikzcd}$$

Now associated with any pointed simplicial complex X is the contractible complex $\mathcal{P}(X)$ of "based paths" of X which has as its 0-simplices the 1-simplices of X of the form $x \longrightarrow 0$, where $0 \in X_0$ is the base point. It is supplied with a canonical pointed simplicial map (last face) $d_n(X)\colon \mathcal{P}(X) \longrightarrow X$ which is a Kan fibration provided that X is a Kan complex. The fiber of this simplicial map is then also a Kan complex and is, of course, the complex $\Omega(X)$ of loops of X at the base point of X.

$$\Omega(X) \subset \mathcal{P}(X) \longrightarrow X$$

The long exact sequence associated with this fibration (since $\pi_i(\mathcal{P}(X)) = \{0\}$) just recapitulates the familiar sequence of isomorphisms

$$\pi_i(\Omega(X)) \simeq \pi_{i+1}(X) \quad i \geq 0.$$

Thus if $f\colon X \longrightarrow Y$ is a pointed simplicial map of Kan complexes, one can form the pullback along f of the fibration $d_n(Y)\colon \mathcal{P}(Y) \longrightarrow Y$

$$\begin{tikzcd}
\Omega(Y) \arrow[r, "\cong"] \arrow[d, hook] & \Omega(Y) \arrow[d, hook] \\
\Gamma(f) \arrow[r, "pr_2"] \arrow[d, "pr_1"'] & \mathcal{P}(Y) \arrow[d] \\
X \arrow[r, "f"'] & Y
\end{tikzcd}$$

The fibers at the base point are then isomorphic and pr_1, as a pullback of a fibration, is itself a fibration. We then obtain a long exact sequence associated with the fibration pr_1,

$$\begin{tikzcd}
\pi_n(\Omega(Y)) \arrow[r, "\pi_n(i)"] & \pi_n(\Gamma(f)) \arrow[r, "\pi_n(f)"] & \pi_n(X) \arrow[dll, dotted] \\
\pi_2(\Omega(Y)) \arrow[r, "\pi_2(i)"'] & \pi_2(\Gamma(f)) \arrow[r, "\pi_2(pr_1)"] & \pi_2(X) \arrow[dll, "\delta"] \\
\pi_1(\Omega(Y)) \arrow[r, "\pi_1(i)"'] & \pi_1(\Gamma(f)) \arrow[r, "\pi_1(pr_1)"] & \pi_1(X) \arrow[dll, "\delta"] \\
\pi_0(\Omega(Y)) \arrow[r, "\pi_0(i)"'] & \pi_0(\Gamma(f)) \arrow[r, "\pi_0(pr_1)"] & \pi_0(X)
\end{tikzcd}$$

which then becomes the *long exact sequence of the pointed simplicial mapping* $f\colon X \longrightarrow Y$:

$$\begin{tikzcd}
\pi_{n+1}(Y) \arrow[r, "\pi_n(i)"] & \pi_n(\Gamma(f)) \arrow[r, "\pi_n(f)"] & \pi_n(X) \arrow[dll, dotted] \\
\pi_3(Y) \arrow[r, "\pi_2(i)"'] & \pi_2(\Gamma(f)) \arrow[r, "\pi_2(pr_1)"] & \pi_2(X) \arrow[dll, "\delta"] \\
\pi_2(Y) \arrow[r, "\pi_1(i)"'] & \pi_1(\Gamma(f)) \arrow[r, "\pi_1(pr_1)"] & \pi_1(X) \arrow[dll, "\delta"] \\
\pi_1(Y) \arrow[r, "\pi_0(i)"'] & \pi_0(\Gamma(f)) \arrow[r, "\pi_0(pr_1)"] & \pi_0(X) \arrow[dll, "\pi_0(f)"] \\
\pi_0(Y) & &
\end{tikzcd}$$

or, equivalently,

$$\begin{tikzcd}
\pi_n(\Gamma(f)) \arrow[r, "\pi_n(pr_1)"] & \pi_n(X) \arrow[r, "\pi_n(f)"] & \pi_n(Y) \arrow[dll, dotted] \\
\pi_2(\Gamma(f)) \arrow[r, "\pi_2(pr_1)"'] & \pi_2(X) \arrow[r, "\pi_2(f)"] & \pi_2(Y) \arrow[dll, "\pi_1(i)"] \\
\pi_1(\Gamma(f)) \arrow[r, "\pi_1(pr_1)"'] & \pi_1(X) \arrow[r, "\pi_1(f)"] & \pi_1(Y) \arrow[dll, "\pi_0(i)"] \\
\pi_0(\Gamma(f)) \arrow[r, "\pi_0(pr_1)"'] & \pi_0(X) \arrow[r, "\pi_0(f)"] & \pi_0(Y)
\end{tikzcd}$$

and thus another justification for calling $\Gamma(f)$ the *homotopy fiber* of $f\colon X \longrightarrow Y$. If f is already a fibration with fiber F, then $pr_2\colon \Gamma(f) \longrightarrow \mathcal{P}(Y)$ is a fibration and its long exact sequence combined with the contractibility of $\mathcal{P}(Y)$ then gives that $\pi_n(F) \simeq \pi_n(\Gamma(f))$, as expected. If f is an inclusion $f\colon X \subseteq Y$ then $\Gamma(f)$ defines the *homotopy groups of* Y *relative to* X with $\pi_{n-1}(\Gamma(f)) = \pi_n(Y; X)$.