DOCTOR OF PHILOSOPHY

The algebra of self-similarity and its applications

Hines, Peter Mark

Award date:
1997

Awarding institution:
University of Wales, Bangor

Link to publication

The Algebra of Self-Similarity and its Applications

A Thesis Submitted to the University of Wales
In Candidature for the Degree of
Doctor of Philosophy

by

Peter Mark Hines

1st May, 1997

School of Mathematics,
University of Wales, Bangor
Dean Street,
Bangor,
Gwynedd LL57 1UT
U.K.

Εν αρχη ην λογος,

και ο λογος ην προς τον Θεον,

και Θεος ην λογος.

Summary of Thesis

We investigate the algebra and category theory arising from the Geometry of Interaction series of papers, with the aim of abstracting the essential ideas behind these models of fragments of linear logic. The main tools used in this investigation are inverse semigroups (in particular the polycyclic monoids, and an inverse monoid of partial bijections on a term language that we call the clause semigroup) and the theory of symmetric monoidal categories, (in particular traced and compact closed categories).

Applications of the above program are given to the following

(i) Ring theory – the conditions for a (corner of a) ring R to be isomorphic to all matrix rings over R.

(ii) The construction of a composition and tensor preserving map from a category to a monoid, giving almost monoidal structures on monoids satisfying certain algebraic or categorical conditions, and a partial dual to this construction given by a restriction of the Karoubi envelope.

(iii) The Geometry of Interaction I system — the identification of the Resolution formula as a categorical trace and the cut-elimination procedure as compact closure.

(iv) Two-way automata — the identification of the composition of global transition relations as the composition in an endomorphism monoid of a compact closed category, and an explicit description of global transition relations of singleton words in terms of Girard's Resolution formula.

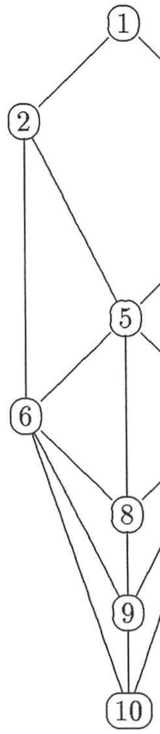
The thesis is not intended to be a study of the logical models used in the Geometry of Interaction, rather, it aims to identify the underlying algebra and category theory, and give applications.

Acknowledgments

This thesis was conducted whilst supported by a University of Wales studentship. It was typeset using the LaTeX system, and all diagrams were produced using the $XY - Pic$ package.

Thanks are due to all members — staff, students, and visitors — of the University of Wales, Bangor mathematics department. Special thanks are due to my supervisor, Dr. M. V. Lawson and my MSc. Supervisor, Prof. R. Brown, who agreed to a change of subject on my part, and indeed to all members of staff, for tolerating my frequent changes of direction.

Logical Dependencies of Chapters



1. *Introductory ideas*

2. *Representations of polycyclic monoids*

3. *Applications of polycyclic monoids to rings*

4. *Categorical self-similarity and internalising monoidal structures*

5. *The natural numbers as a self-similar object*

6. *The categorical trace, and compact closed categories*

7. *Linear logic and the Geometry of Interaction I*

8. *Analysis of the Geometry of Interaction I*

9. *The clause semigroup and its applications*

10. *Applications of the trace to automata*

Contents

Chapter 0

Introduction

0.1 Background

The background to the following thesis is far from straightforward. The main concepts arose from two independent sources — firstly was the study of various forms of self-similarity, motivated by the idea that the proper symmetry structures for objects displaying self-similarity (the canonical example being fractals, and in particular, the Cantor set) were inverse semigroups, rather than groups. Secondly, an attempt was being made to understand the algebra behind the Geometry of Interaction series of papers ([20, 21, 22]) — dynamical models of (various fragments of) linear logic, by J.Y. Girard. These use various algebraic structures, including (matrices over) partial isometries in C^*–algebras, (partial) homeomorphisms on the Cantor set, partial bijections on the natural numbers, and contracted semigroup rings of inverse monoids of actions on a term language.

It was a major surprise to discover that the same algebraic structures were arising from both lines of work. In particular, the natural symmetry structures of the Cantor set appear to be the polycyclic monoids (as demonstrated in Chapter 2), and the first two papers in the Geometry of Interaction series use matrices over an embedding of the polycyclic monoids into a C^*–algebra, the partial bijections on the natural nmbers, and the Cantor set.

The next step in the development of the thesis came from the visits of Prof. Ross Street to Bangor, and subsequent discussions with him — in particular, the theory of categorical traces, and their connections with compact closed categories. This led to the identification of the categorical structures behind the dynamics of Girard's cut-elimination processes, as presented in the Geometry of Interaction program. An unexpected spin-off was the identification of the dynamics of two-way automata in the same terms.

After that, an appropriate categorical formulation of self-similarity in terms of monoidal category theory, and its close connections with the various algebraic structures used in the remainder of the thesis (see, for example, Chapters 3, 4 and 5) led to a categorical formulation of the models used, and the construction of a one-object compact closed category that is proved in Chapter 8 to be a model of the dynamics of Geometry of Interaction 1, as used in [20].

A start was made on analysing the third Geometry of Interaction paper (Chapter 9), but this paper is very different from the other two, and we have not yet put it in the same categorical terms.

The thesis cannot be considered to be a complete analysis of the Geometry of Interaction program — a great deal of work remains to be done. However it does demonstrate that self-similarity is absolutely fundamental to these logical models, and the correct algebraic models are inverse semigroups (in particular, the polycyclic monoids). It also demonstrates that the categorical trace can be considered to be a dynamical model of computation, and the correct categorical closure required is the compact closure arising from the categorical trace on a symmetric monoidal category.

0.2 The structure of the thesis

The thesis is split up into 10 Chapters (together with this introduction, and a conclusion). These are as follows:

Chapter 1

This chapter is mainly concerned with setting up the algebraic framework for the following algebraic structures. However, it does introduce the disjoint closure, which is a method of constructing inverse semigroups from other inverse semigroups in a way that can be considered to be an inverse-semigroup theoretic version of the power set construction. The definitions of polycyclic monoids are also presented, and a method of embedding P_α into P_2, for all countable α is given.

Chapter 2

Chapter 2 considers the different representations of polycyclic monoids we are interested in. It analyses the self-similarity of the natural numbers, and shows how this is equivalent to embeddings of polycyclic monoids into the symmetric inverse monoid on the natural numbers. It also demonstrates how an embedding of an inverse monoid (with a zero) into the symmetric inverse monoid of partial bijections on a set determines a natural topology on the underlying set in which all the members of the (embedding of the) inverse monoid are partial homeomorphisms. In the case considered, the polycyclic monoids arise naturally from topological considerations on the Cantor set. It then constructs an embedding of the symmetric inverse monoid on $\mathbb{N}$ (and hence the various embeddings of P_α previously constructed) into the C^*–algebra of bounded linear operators on the Hilbert space l^2.

Chapter 3

This Chapter is concerned with consequences of polycyclic monoids being embedded into the multiplicative monoid of a ring, in terms of the symmetric monoidal structure of the set of all (finite) matrices over a ring. This leads to some very strong ring-theoretic results; in particular, the conditions for the matrix rings $M_n(R)$ over a ring R to be embedded in, or isomorphic to R itself, for all $n \in \mathbb{N}$, and the construction of a non-trivial one-object symmetric monoidal structure (apart from the units) on a ring. These results are then used to simplify the construction of the K_0 groups of rings, assuming an embedding of a polycyclic monoid.

Chapter 4

This is an attempt to put the general idea of self-similarity into a categorical setting, using (symmetric) monoidal categories. Starting with the assumption of maps $d : N \to N \otimes N$ and $c : N \otimes N \to N$ that satisfy $dc = 1_{N \otimes N}$, it demonstrates how the monoidal structure of the category is modelled in the endomorphism monoid of the object N, and constructs a map that can reasonably be considered to be a (partial) dual to the Karoubi envelope. The self-similarity considerations motivate the definition of M-monoids, which are fundamental to the remainder of the thesis, and can be considered to be weakenings, or generalisations of one-object symmetric monoidal categories.

It is also demonstrated how the construction of C–monoids (one-object cartesian closed categories, used to model the untyped lambda calculus) fits into this self-similarity framework.

Chapter 5

Chapter 5 presents the monoidal category theory of the category of relations, and its subcategory of partial bijective maps. A matrix representation of relations and partial bijective maps is presented, and given a graphical interpretation. The conditions for a matrix to represent a partial bijective map are also found.

It then puts the self-embedding results of the natural numbers from Chapter 2 into the categorical framework developed in Chapter 4. In particular, two distinct M-monoid structures are identified on $I(\mathbb{N})$ that we call the internalised tensor, and internalised direct sum (using terminology derived from their implicit construction in [20]). The canonical associativity and commutativity elements for these are specified explicitly, in terms of embeddings of polycyclic monoids, and a construction of a 'fixed point' for the direct sum, arising naturally from the tensor, is given.

Chapter 6

Chapter 6 considers the categorical trace, and proves that both the category of relations, and the category of partial bijective maps are traced. Self-similarity properties are then used to motivate the definition of traced M-monoids, and one-object traced monoidal categories, and a routine method of constructing traced M-monoids from self-similar objects of traced symmetric monoidal categories is given.

Compact closed categories are then defined, and demonstrated to be traced. The construction of compact closed categories from traced symmetric monoidal categories is presented, with particular reference to the category of relations, to give concrete examples to work with.

An alternative set of axioms for compact closed categories is given, and the equivalence with the usual definition is proved. This alternative set of axioms allows us to define compact closed M-monoids, and hence one-object compact closed categories. The self-similarity of objects in the compact closed category arising from the trace on the category of partial bijective maps is analysed, and this is used to construct self-similar objects of a compact closed category using the results of Chapter 5 on the natural numbers. Finally, these results are combined to give an explicit description of a one-object compact closed category, which is also proved to be inverse.

Chapter 7

This Chapter is entirely expository. It presents (in a non-rigorous way) Gentzen's sequent calculus, and following the approach of [23], constructs linear logic in terms of linearity considerations on Gentzen's sequent calculus, and gives the cut-elimination algorithm for the multiplicative fragment. The main part of this chapter then follows, which is an exposition of the 'Geometry of Interaction' model of multiplicative linear logic taken from [20], and a statement of the cut-elimination theorem. No new results are presented, but the description of the Geometry of Interaction 1 system is (hopefully) clarified.

Chapter 8

Chapter 8 is the description of the Geometry of Interaction 1 system in terms of the structures developed in Chapters 1 to 6. It demonstrates how this system can be represented in terms of partial bijective maps on the natural numbers, and in particular, the disjoint closure of an embedding of the polycyclic monoid on two generators.

It also shows that all the logical and structural operations can be given in terms of (variations of) the canonical elements of the M-monoid structures on $I(\mathbb{N})$ developed in Chapter 5, and how every element of $I(\mathbb{N})$ representing a proof is a partial symmetry.

Finally, the cut-elimination procedure is demonstrated to be given by the internalisation of the categorical trace at the natural numbers, and, when combined with the cut procedure, is given by the composition of the elements (together with the dual on elements) in the one-object compact closed inverse monoid, presented in Chapter 6. This allows the simplification of results from [20]; in particular, the 'essential case' of the cut-elimination procedure (a cut between two axiom links) follows directly from the monoid identity $1 \circ 1 = 1$.

Chapter 9

This Chapter is the first stage of an analysis of the 3$^{\text{rd}}$ part of the Geometry of interaction program. It proves that the algebraic structures used are in fact a (family of) inverse semigroups that we call the clause semigroups. These are defined in terms of a semilattice-theoretic approach to substitution and unification on term languages, and the concept of a 'linear pair' of terms in a term language, which is a pair that have exactly the same free variables.

The action of a term language on ground terms (as defined in [22]) is proved to be well-defined, and is extended to the semilattice structure of the term language given by substitution. The conditions on a term language required for the representation of linear logic (as given in [20]) are also considered, and demonstrated to imply an embedding of the polycyclic monoids into the clause semigroup of the term language. Finally, the Resolution formula (which is claimed to model cut-elimination in [22]) is proved to be given by the categorical trace in the category of partial bijective maps.

Chapter 10

This chapter demonstrates how composition in compact closed categories and Girard's resolution formula are both vital to the analysis of 2-way automata. In particular, the composition of global transition relations of a 2-way automaton, as defined in [3], is given by the composition in the endomorphism monoid of (two copies of) the set of states of an automaton in the compact closed categories derived from the category of relations. This allows short proofs of the facts that composition of global transition relations is associative, and (finite-state) 2-way automata can be simulated by (finite-state) classical automata.

It is then proved that the global transition relations of singleton words can be calculated directly, using Girard's resolution formula. Finally, a method of 'sticking together' two one-way automata to construct a 2-way automaton is given, and the construction of the global transition monoid of this 2-way automaton in terms of the transition monoids of the one-way automata is given.

Note that we use a slightly different model of 2-way automata to [3]; however, the equivalence of the two models is immediate.

Chapter 1

Introductory ideas

1.1 Introduction

In this chapter, we present the basic theory of categories, semigroups, and inverse semigroups. We introduce a restriction of the compatibility relation, which we call the disjointness relation, and use this to construct new inverse semigroups. We then present the theory of polycyclic monoids, and construct embeddings of the polycyclic monoid on α generators into the polycyclic monoid on 2 generators, for all countable α.

1.2 Category theory

1.2.1 Basic definitions

Definitions 1.1

A *category* $\mathbf{C}$ is defined to be a set[1] of *objects*, $Ob(\mathbf{C})$, and for every pair of objects X, Y, a set of *arrows*, or *morphisms*, denoted $\mathbf{C}(X, Y)$. For $f \in \mathbf{C}(X, Y)$, the object X is called the *domain* of f, and Y is called the *codomain* of f. These are denoted $dom(f)$ and $cod(f)$ respectively. We sometimes denote $f \in \mathbf{C}(A, B)$ by $f : A \rightarrow B$. However, this should not be taken to imply that f is a function in the set-theoretic sense. The set of all arrows of $\mathbf{C}$, written $Arr(\mathbf{C})$, has an associative partial binary operation (generally denoted by concatenation) defined on it, that satisfies, for all $a, b \in Arr(\mathbf{C})$,

1. ba is defined iff $dom(b) = cod(a)$, in which case, $dom(ba) = dom(a)$, and $cod(ba) = cod(b)$.

[1] Category theorists will note that technically, this is only the definition of a small category. The general theory, in terms of *classes*, can be found in [41].

2. For all $X \in Ob(\mathbf{C})$, there exists $1_X \in \mathbf{C}(X,X)$, satisfying, for all $f \in \mathbf{C}(X,Y)$, $f1_X = f$, and $1_Y f = f$.

A category is called *regular* if each arrow $a \in \mathbf{C}(X,Y)$ has a *generalised inverse* $a' \in \mathbf{C}(Y,X)$ satisfying $aa'a = a$ and $a'aa' = a'$. If the generalised inverse is unique, the category is called *inverse*, and the generalised inverse of an element a is denoted by a^{-1}. If the generalised inverses satisfy $aa^{-1} = 1_Y$ and $a^{-1}a = 1_X$ for all $a \in \mathbf{C}(X,Y)$, then the category $\mathbf{C}$ is called a *groupoid*.

A *functor* $\Gamma : \mathbf{C} \to \mathbf{D}$ between categories is a map from $Ob(\mathbf{C})$ to $Ob(\mathbf{D})$, and from $Arr(\mathbf{C})$ to $Arr(\mathbf{D})$, that satisfies, for all $a, b \in Arr(\mathbf{C})$,and $X \in Ob(\mathbf{C})$

- $\Gamma(1_X) = 1_{\Gamma(X)}$,
- ba is defined in $\mathbf{C}$ implies $\Gamma(b)\Gamma(a)$ is defined in $\mathbf{D}$, in which case $\Gamma(b)\Gamma(a) = \Gamma(ba)$.

A functor from a category to itself is called an *endofunctor.* A *natural transformation* between two functors $\Gamma : \mathbf{C} \to \mathbf{D}$ and $\Delta : \mathbf{C} \to \mathbf{D}$, is a function τ from $Ob(\mathbf{C})$ to $Arr(\mathbf{D})$ that satisfies $\tau(c) \in \mathbf{D}(\Gamma(c), \Delta(c))$ and $\Gamma(f)\tau(c) = \tau(c')\Delta(f)$ for each arrow $f \in \mathbf{C}(c, c')$. The members of the set (or class) $\{\tau(c) : c \in Ob(\mathbf{C})\}$ are called the *components* of τ.

A set of morphisms indexed by objects of a category, $\{\tau_{x_1,x_2,\ldots x_n} : c \to c',\ c, c' \in Ob(\mathbf{C})\}$, is said to be *natural in* $\{x_i : i = 1 \ldots n\}$ if its members are the components of a natural transformation between functors from $\mathbf{C}^n$ to $\mathbf{C}$.

1.2.2 Symmetric monoidal categories

Definitions 1.2

A category $\mathbf{M}$ is said to be *symmetric monoidal* if there exists a functor $\otimes : \mathbf{M} \times \mathbf{M} \to \mathbf{M}$, together with a unit object $I \in Ob(\mathbf{C})$, and families of isomorphisms

- $t_{A,B,C} \in \mathbf{M}(A \otimes (B \otimes C), (A \otimes B) \otimes C)$,
- $s_{A,B} \in \mathbf{M}(A \otimes B, B \otimes A)$,
- $\lambda_A \in \mathbf{M}(I \otimes A, A)$,
- $\rho_A \in \mathbf{M}(A \otimes I, A)$,

that are natural in A, B, C, and satisfy, for all arrows a, b, c,

1. $t(a \otimes (b \otimes c)) = ((a \otimes b) \otimes c)t$,

2. $s(b \otimes a) = (a \otimes b)s$,

3. $\lambda_I = \rho_I$,

(we will sometimes omit the object subscripts for clarity, when the meaning is clear) together with the following *coherence equations* for all objects A, B, C, D

1. *The MacLane Pentagon:*

$$t_{(A\otimes B),C,D}t_{A,B,(C\otimes D)} = (t_{A,B,C} \otimes 1_D)t_{A,(B\otimes C),D}(1_A \otimes t_{B,C,D}).$$

2. *The Units Triangle:*

$$(\rho_B \otimes 1_A)t_{B,I,A} = (1_B \otimes \lambda_A)$$

3. *The Commutativity Hexagon:*

$$t_{C,A,B}s_{(A\otimes B),C}t_{A,B,C} = (s_{A,C} \otimes 1_B)t_{A,C,B}(1_A \otimes s_{B,C}).$$

A monoidal category **M** is called *strictly monoidal* if the morphisms $t_{A,B,C}$, λ_A, ρ_A are identity morphisms for all $A, B, C \in Ob(\mathbf{M})$. Unless stated explicitly, we shall assume that a category that is stated to be monoidal is *not* strict.

1.3 Semigroup theory

1.3.1 Basic definitions

We refer to [30] for the basic definitions of semigroups, monoids, subsemigroups, congruences, homomorphisms, and other basic concepts.

Definitions 1.3

For any set S, its *power set*, written $P(S)$, is the set of all subsets of S. When S is a semigroup, the power set of S, together with with the induced binary operation on subsets

$$AB = \{ab : a \in A, b \in B\}$$

is also a semigroup. Also, $\emptyset \in P(S)$ for any semigroup S, and satisfies $\emptyset A = \emptyset = A\emptyset$. Therefore $P(S)$ is a semigroup with a zero, for all semigroups S. Finally, if S is a monoid, then the subset $\{1\} \in P(S)$ satisfies $\{1\}A = A = A\{1\}$. Therefore, if S is a monoid, so is $P(S)$. Note that, for any operation on an algebraic structure S, we define induced operations on its power set pointwise, so given a function $\Omega : S \to T$, we define $\Omega : P(S) \to P(T)$ by $\Omega(A) = \{\Omega(a) : a \in A\}$ for all $A \in P(A)$. (We abuse notation, and use the same symbol for an operation and the induced operation on the power set, unless the distinction is important). Note that $P(S)$ is also closed under arbitrary unions, by definition. This leads to the following result:

Lemma 1 *For all $X \in P(S)$, $\{A_i : i \in I\} \subseteq P(S)$,*

$$X \bigcup_{i \in I} A_i = \bigcup_{i \in I} X A_i.$$

Proof by definition of the induced composition,

$$X \bigcup_{i \in I} A_i = \left\{ xa : x \in X, a \in \bigcup_{i \in I} A_i \right\},$$

and

$$\bigcup_{i \in I} X A_i = \bigcup_{i \in I} \{xa : x \in X, a \in A_i\},$$

and it is immediate that these are the same. □

1.3.2 Rings and semigroup rings

We refer to [5] for the basic definitions of rings, homomorphisms, left and right ideals, quotient rings, and similar concepts.

Definitions 1.4

For any semigroup S and commutative ring with identity R, the *semigroup ring* RS is defined in [39] as follows: The elements of RS are functions $\alpha : S \to R$, where $\alpha(s) = 0$ for all but a finite number of elements of S. Addition in RS is the usual addition of maps into an abelian group. The product on RS is then defined by

$$(\alpha\beta)(s) = \sum_{uv=s} \alpha(u)\beta(v).$$

It is proved in [39] that RS is a ring, and there is a canonical embedding of S into RS, given by $x \mapsto \phi(x)$, where

$$\phi_x(s) = \begin{cases} 1 & s = x \\ 0 & s \neq x. \end{cases}$$

Also, if S is a monoid, with identity e, then there is a canonical embedding of R into RS, given by $r \mapsto \psi(r)$, where

$$\psi_r(s) = \begin{cases} r & s = e \\ 0 & s \neq e \end{cases}$$

Finally, it is proved in [39] that if we denote by ax the unique function whose value at x is a and is zero elsewhere, then every element has a unique representation as $\sum_{i=1}^{n} a_i x_i$. Therefore, we can think of RS as the set of all finite formal sums of elements of S indexed by a member of R, with the composition

$$\left(\sum_{i=1}^{n} \alpha_i s_i\right)\left(\sum_{j=1}^{m} \beta_j t_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{m} \alpha_i \beta_j s_i t_j.$$

If S has a zero, 0_S, it is immediate that the set of all formal sums of 0_S, together with their additive inverses, is an ideal of RS. In this case, the *contracted semigroup ring* is the semigroup ring RS quotiented by this ideal. By convention, the elements of RS are identified with their equivalence classes under this quotient, unless the distinction is important. An important contracted semigroup ring for our work will be the ring $\mathbb{Z}S$, for various semigroups S. See [39] for a more detailed discussion of semigroup rings, and [12] for contracted semigroup rings.

1.3.3 Inverse semigroups

Definitions 1.5

A *regular semigroup* is defined analogously to a regular category; every element $s \in S$ is required to have a *generalised inverse*, $s' \in S$ that satisfies $ss's = s$ and $s'ss' = s'$. Similarly, a regular semigroup is called *inverse* if each element has exactly one generalised inverse. This condition is equivalent to the condition that all the idempotents of an inverse semigroup commute [44]. By convention, the unique inverse of an element $s \in S$ is denoted by s^{-1}. The same reference ([44]) also proves that all idempotents of an inverse semigroup are of the form $f^{-1}f$, for some $f \in S$.

An inverse semigroup S has a partial order defined on it, called the *natural partial order*, given by $s \leq t$ iff there exists an idempotent e satisfying $s = et$. It can also be shown that $s \leq t \Leftrightarrow s = tf$, for some idempotent f.

It is easy to show that a homomorphism of inverse semigroups $f : S \to T$ satisfies $f(s)^{-1} = f(s^{-1})$ and $x \leq y \Rightarrow f(x) \leq f(y)$. If S and T both have a zero, then a *0-homomorphism* is defined to be a homomorphism that satisfies $f(0) = 0$.

The canonical example of an inverse semigroup is that of the *symmetric inverse monoid* on a set X. This is the set of all partial bijective maps on the set X, written $I(X)$. Its theory is well known; see [44]. Its elements are partial bijective functions from X to itself, which can be thought of as relations $\{(f(x), x) : x \in dom(f)\}$ with composition

$$gf = \{(a, c) : \exists b \in X; (a, b) \in g, (b, c) \in f\}.$$

The inverse of an element is given by $f^{-1} = \{(y,x) : (x,y) \in f\}$ and the partial identity on the domain (resp. image) of an element f is given by $f^{-1}f$ (resp. ff^{-1}). Also, $I(X)$ contains 1, the global identity, and 0, the empty map, for any set X. Hence, it is an inverse monoid with a zero.

See [44] for proofs of these results, and more details of the theory of symmetric inverse monoids.

1.4 The disjointness relation

We introduce a new relation on inverse semigroups that was motivated by a construction of J.-Y. Girard for partial isometries in C^*– algebras.

Definitions 1.6

Given $a, b \in S$, where S is an inverse semigroup, a and b are said to be *compatible* if both $a^{-1}b$ and ab^{-1} are idempotent. The compatibility relation is a reflexive, symmetric relation; see [44] for more details of its theory. We define the following refinement of the compatibility relation: Let S be an inverse semigroup with a zero. Then $a, b \in S$ are said to be *disjoint*[2], if they satisfy $ab^{-1} = 0 = b^{-1}a$. We denote this by $a \perp b$. Clearly, disjointness implies compatibility.

Proposition 2 *Given $a, b \in I(X)$, for some set X, then a and b are disjoint iff they have disjoint domains and images.*

Proof ($\Rightarrow$) Let $a \perp b \in I(X)$. Then $ab^{-1} = 0 = b^{-1}a$. Therefore, $a^{-1}ab^{-1}b = 0 = bb^{-1}aa^{-1}$. However, by the properties of symmetric inverse monoids, the partial identity on the domain of a is $a^{-1}a$, and the partial identity of the image of a is aa^{-1}; similarly for b. Therefore, the partial identity on the intersection of their domains is $a^{-1}ab^{-1}b$; however, this is zero. Similarly the partial identity on the intersection of their images is 0. Therefore, disjoint elements of $I(X)$ have disjoint domains / images. □

($\Leftarrow$) Let $a, b \in I(X)$ satisfy $dom(a) \cap dom(b) = \emptyset = im(a) \cap im(b)$. Then $a^{-1}ab^{-1}b = 0$. Therefore, $aa^{-1}ab^{-1}bb^{-1} = 0$, and so $ab^{-1} = 0$.

Similarly, $aa^{-1}bb^{-1} = 0$, and so $a^{-1}aa^{-1}bb^{-1}b = 0$, and so $a^{-1}b = 0$, and hence $b^{-1}a = 0$. Therefore, a and b are disjoint. □

[2] Note that we use the term *disjoint* in place of the mathematically more accurate *orthogonal*. This is because in Girard's Geometry of Interaction system [20], which this thesis is based on, the term *orthogonal* is used for an entirely different concept.

Definitions 1.7

We say that a subset of an inverse semigroup $A \subseteq S$ is *disjoint* if any two distinct elements $x, y \in A$ are disjoint.

Examples

(i) A set of elements $A \subseteq I(X)$ is a disjoint subset of $I(X)$ when the elements of A have pairwise-disjoint domains / images. Note that this implies that the set-theoretic union of the elements of A is also a partial bijective function.

(ii) Let $(R, ., +)$ be a ring, and let $S \subseteq R$ be a subsemigroup of $(R, \cdot)$ that is inverse. Then the set of all sums of disjoint subsets of S in R is also an inverse semigroup.

In what follows, we consider the properties of disjoint subsets.

Lemma 3 *Let $A, B \subseteq S$ be subsets of an inverse semigroup, and let $f : S \to T$ be a 0-homomorphism. Then*

(i) *If A is a disjoint subset, then $f(A)$ is a disjoint subset of T.*

(ii) *If A and B are disjoint (in the sense of Definition 1.7), then AB is a disjoint subset of S.*

(iii) *If A is a disjoint subset of S, then A^{-1} is a disjoint subset of S, $(A^{-1})^{-1} = A$, and $AA^{-1}A = A \cup \{0\}$.*

(The composition of subsets of S is defined in the natural way, as for the power set construction).

Proof

(i) As f is a homomorphism, $f(a)f(b) = f(ab)$ for all $a, b \in A$. Also, as f is a 0-homomorphism, $f(0) = 0$. Therefore, $f(a)^{-1}f(b) = f(a^{-1}b) = f(0) = 0$, for all $a, b \in A$. A similar proof gives that $f(a)f(b)^{-1} = 0$, and so $f(A)$ is a disjoint set.

(ii) Let $A = \{a_i : i \in I\}$, $B = \{b_k : k \in K\}$, then consider distinct elements a_ib_k and a_jb_l of AB. Then for any $i \neq j \in I$,

$$(a_ib_k)^{-1}(a_jb_l) = b_k^{-1}a_i^{-1}a_jb_l = 0,$$

since $a_i^{-1}a_j = 0$ for all $i \neq j$. When $a_i = a_j$, then $b_k \neq b_l$ (by the assumption that the two elements of AB are distinct). Therefore $(a_ib_k)^{-1}(a_ib_l)^{-1} = b_k^{-1}a_i^{-1}a_ib_l$. However, $a_i^{-1}a_ib_l = b_le$, where $e = b_l^{-1}a_i^{-1}a_ib_l$, by the commutativity of idempotents in an inverse semigroup, so $(a_ib_k)^{-1}(a_ib_l)^{-1} = b_k^{-1}b_le = 0$, by the condition that B is a disjoint subset.

Similarly,

$$(a_ib_k)(a_jb_l)^{-1} = a_ib_kb_l^{-1}a_j^{-1} = 0,$$

when $b_k \neq b_l$, since $b_kb_l^{-1} = 0$ for all $j \neq k$. When $b_k = b_l$, then $a_i \neq a_j$ (by the assumption that the two elements of AB are distinct). Then $(a_ib_l)(a_jb_l)^{-1} = a_ib_lb_l^{-1}a_j^{-1} = a_ia_j^{-1}f$, where $f = a_jb_lb_l^{-1}a_j^{-1}$, by the commutativity of idempotents in an inverse semigroup, and so $(a_ib_l)(a_jb_l)^{-1} = a_ia_j^{-1}f = 0$, by the condition that A is a disjoint subset. Therefore, AB is a disjoint subset of S.
(iii) For all $a \neq b \in A$, $a \perp b$; that is, $a^{-1}b = 0 = ab^{-1}$. Therefore, $(a^{-1})^{-1}(b^{-1}) = ab^{-1} = 0$, and $a^{-1}(b^{-1})^{-1} = a^{-1}b = 0$, so A^{-1} is also a disjoint subset of S. Also,

$$(A^{-1})^{-1} = \{(a^{-1})^{-1} : a \in A\} = \{a : a \in A\} = A.$$

Finally, for all $a, b, c \in A$,

$$ab^{-1}c = \begin{cases} a & a = b = c \\ 0 & \text{otherwise} \end{cases}$$

Therefore, $AA^{-1}A = \{aa^{-1}a : a \in A\} \cup \{0\} = A \cup \{0\}$ since $aa'^{-1} = 0 = a^{-1}a'$ for all $a \neq a' \in A$. □

Definitions 1.8

We define the *disjoint completion* of S, an inverse semigroup with a zero, to be the set of all disjoint subsets of an inverse semigroup S containing 0, together with the composition and inverse induced by the power set construction.

Theorem 4 *Let S be an arbitrary inverse semigroup with a zero. Then*
(i) *$DC(S)$ is an inverse semigroup,*
(ii) *S is embedded in $DC(S)$.*

Proof (i) By Lemma 3, $DC(S)$ is closed under composition and taking inverses, and the inverse map satisfies $(A^{-1})^{-1} = A$ and $AA^{-1}A = A$, for all $A \in DC(S)$. Therefore, $DC(S)$ is a regular semigroup. Now let E be an idempotent of $DC(S)$. Then for all non-zero $c \in E$, there exists $a, b \in E$ which satisfy $ab = c$, as $E^2 = E$. Therefore, $a^{-1}c$ is non-zero. Therefore, by the disjointness assumption, $a = c$, and from this, it follows trivially that $a = b$. Hence $a^2 = a$, and so all members of E are idempotent. Conversely, let $F = \{e_i : e_i^2 = e_i\} \cup \{0\}$. Then $F^2 = \{e_ie_i\} \cup \{0\}$, since $e_i = e_i^{-1}$ for all idempotents e_i, and $e_ie_j = e_i^{-1}e_j = 0$ for all $i \neq j$. Therefore $F^2 = F$, and so we have characterised the idempotents of $DC(S)$. To see that $DC(S)$

is inverse, any idempotent $E^2 = E$ must satisfy $E = \{e_i : i \in I\}$, where $e_i^2 = e_i$. Therefore, given two idempotents $E = \{e_i : i \in I\}$ and $F = \{f_j : j \in J\}$,

$$EF = \{e_i f_j : i \in I, j \in J\} = \{f_j e_i : i \in I, j \in J\} = FE.$$

Hence, the idempotents of $DC(S)$ commute, and so $DC(S)$ is inverse.
(ii) This follows immediately from the definition of the embedding $i(s) = \{s, 0\}$ for all $s \in S$; this map is clearly an injective homomorphism. □

There is an analogous definition of closure for inverse subsemigroups of symmetric inverse monoids, as follows:

Definitions 1.9

Let S be an inverse semigroup with a zero, and let $i : S \to I(X)$ be an injective 0-homomorphism. We define the *disjoint closure of* S *in* $I(X)$, denoted by $DC_X(S)$, to be the image of the map $\phi : DC(S) \to I(X)$ defined by

- $\phi(\{s, 0\}) = i(s)$,
- $\phi(A \cup B) = \phi(A) \cup \Phi(B)$, when A and B are disjoint.

We denote the set-theoretic union of disjoint partial bijections by $\vee$, and refer to this as the *disjoint join* of partial bijections. This operation then satisfies the following:

Proposition 5 *Let S be an inverse subsemigroup of $I(X)$. Then for all disjoint sets $\{a_i : i \in I\} \subseteq S$,*

$$\bigvee_{i \in I} sa_i = s \bigvee_{i \in I} a_i.$$

Proof

(i) We have seen that elements $x, y \in I(X)$ are disjoint iff they have disjoint domains and images. Therefore, any disjoint set $A = \{a_i : i \in I\} \subseteq I(X)$ consists of a set of partial bijective functions whose domains and images do not intersect. Hence, $\cup a_i$ is also a partial bijective function. Then by definition

$$\bigvee sa_i = \bigcup_{i \in I} \{(x, z) : (x, y) \in s, (y, z) \in a_i\}.$$

However,

$$\bigvee a_i = \bigcup_{i \in I} \{(y, z) : (y, z) \in a_i\},$$

so

$$s\bigvee a_i = \bigcup_{i\in I}\{(x,z) : (x,y) \in S, (y,z) \in a_i\} = \bigvee sa_i.$$

Therefore, our result follows. □

We also have the following related definition: Given an inverse subsemigroup S of $I(X)$, and an idempotent $e^2 = e$ of S, then e satisfies $e = 1_A$ for some $A \subseteq X$. Then we define its *complement* $e^\perp \in I(X)$ by $e^\perp = 1_{X\setminus A}$. Note that, for all $e^2 = e \in S$, $e \vee e^\perp = 1$.

This idea was also motivated by J.-Y. Girard's (implicit) embedding of $I(\mathbb{N})$ into $B(l^2)$ in [20]. It is also used in [9], where it was defined (for $I(\mathbb{N})$) in terms of an operation, denoted by $\overline{(\ )}$, on arbitrary elements. This definition is equivalent to $\overline{f} = (f^{-1}f)^\perp$.

The final notion of disjoint closure we introduce is the *finite disjoint closure* of S, which we define to be the collection of all finite disjoint subsets of S containing zero, together with the induced composition and inverses. It is trivial that this is an inverse semigroup, from the results given for $DC(S)$; we only need to check that the composition and inverse operations preserve finiteness, and this is immediate. We denote this inverse semigroup by $DC^{<\infty}(S)$.

Proposition 6 *$DC^{<\infty}(S)$ is embedded in the multiplicative monoid of the contracted semigroup ring $\mathbb{Z}S$.*

Proof We define a map from $DC^{<\infty}(S)$ to $\mathbb{Z}S$ by $\{s_i\}_{i=1}^n \mapsto \sum_{i=1}^n s_i$ and it is trivial from the distributivity of multiplication over addition in a ring that this is an injection that preserves composition and (multiplicative) inverses. □.

1.5 Polycyclic monoids

Definitions 1.10

Polycyclic monoids are inverse semigroups that are useful for modelling self-similarity properties in terms of partial bijections (e.g. see [27]). They were first introduced in [43], where P_X, the polycyclic monoid on the set X, was defined to be the inverse monoid (with a zero, for $n \geq 2$) generated by a set of countable cardinality, X, say $\{p_0, \ldots p_{n-1}\}$, subject to the relations $p_i p_j^{-1} = \delta_{ij}$. It will be convenient to denote the generators of P_2 by p, q.

In [43], it is proved that words in the polycyclic monoid P_X have canonical form $x^{-1}y$, where $x, y \in X^*$, the free monoid on X, and composition of canonical forms is given by

$$(v^{-1}w)(x^{-1}y) = \begin{cases} v^{-1}hy & \text{if } w = hx \text{ for some } h \\ v^{-1}k^{-1}y & \text{if } x = kw \text{ for some } k \\ 0 & \text{otherwise.} \end{cases}$$

See [43] for a proof of the above statements. The same paper also proves that the polycyclic monoids P_X are congruence-free, for $|X| \geq 2$. The monogenic polycyclic monoid is referred to as the *bicyclic monoid*, and differs from the other polycyclic monoids by having non-trivial congruences, and not having a zero.

The polycyclic monoid on n generators can also be constructed as the *inverse hull* of the free monoid on n generators. The inverse hull construction is defined on arbitrary left-cancellative monoids, as follows:

Given S, a left-cancellative monoid, consider the set of maps from S to itself defined by $\lambda_a(s) = as$ for all $a \in S$. Since S is left-cancellative, $\{\lambda_a : a \in S\}$ is a subset of the monoid of all partial bijective maps from S to itself, $I(S)$. The inverse hull of S, denoted $\Sigma(S)$, is defined to be the smallest inverse submonoid of $I(S)$ that contains $\{\lambda_a : a \in S\}$. See [42] for a fuller introduction to the theory of inverse hulls. If we take S to be the free monoid on n generators, say $S = \{p_0, p_1, \ldots p_{n-1}^{-1}\}^*$, and use the embedding $\lambda : S \to I(S)$, then it is proved in [43] that $\Sigma(S) \cong P_n$.

Proposition 7 *Let $S = \{q_1, q_2\}$, and let $F_n = \{w_1, w_2, \ldots w_n\}$ be a set of words in S^* that generate a free submonoid of S^*. Then the construction of the inverse hull of F_n^* gives an embedding of P_n into P_2.*

Proof By definition, S^* and F_n^* are free monoids on 2 and n generators respectively. However, $\Sigma(S^*) \subseteq I(S^*)$, and $\Sigma(F_n^*) \subseteq I(F_n^*)$. We can embed $\Sigma(F_n^*)$ into $I(S^*)$ by extending the action of $\lambda_a : F_n^* \to F_n^*$ in the natural way, so that $\widetilde{\lambda_a}(s) = as$ for all $a \in F_n^*$ and $s \in S^*$.

Therefore, $\{\widetilde{\lambda_f} : f \in F_n\} \subseteq \{\lambda_s : s \in S\}$. and this embedding is a monoid homomorphism. Hence $\{\widetilde{\lambda_f^{-1}} : f \in F_n\} \subseteq \{\lambda_s^{-1} : s \in S\}$ and so, by definition of the inverse hull construction, $\Sigma(F_n^*) \subseteq \Sigma(S^*)$. Therefore, since $\Sigma(S^*) \cong P_2$, and $\Sigma(F_n^*) \cong P_n$, we have constructed an embedding of P_n into P_2. □

We present an explicit example of the above construction, and derive an embedding of P_n into P_2 for all $n \geq 2$.

Theorem 8 *Let P_2 be the polycyclic monoid on two generators, p, q. The following set of maps are embeddings of P_n into P_2 for all $n \in \mathbb{N}$.*

$$\theta_n(p_i) = \begin{cases} p & i = 0 \\ pq^i & 2 \leq i \leq n-2 \\ q^{n-1} & i = n-1, \end{cases}$$

Proof The submonoid of $\{p, q\}^*$ generated by $\{\theta_n(p_i) : 0 \leq i \leq n-1\}$ is free in $\langle q_1, q_2 \rangle$, since $\{p_i : 0 \leq i \leq n-1\}$ is the set of leaves of a subtree of the prefix tree of $\{p, q\}^*$. Therefore, (by [2] p.85-88, where this example is considered) they form a prefix code, and so the semigroup generated by $\{\theta_n(p_i)\}$ is free in $\{p, q\}^*$. Hence, by the construction of P_n as the inverse hull of the free semigroup of n generators, the semigroup generated by the $\{\theta(p_i)\}$ is free in $\{p, q\}^*$ so the above construction gives an embedding of P_n into P_2 for all $n \in \mathbb{N}$. □

There is a natural extension of this to the infinite case, as follows:

Theorem 9 *Let P_2 be the polycyclic monoid on two generators, p, q. Then the map $\theta_\infty(p_i) = pq^i$ generates an embedding of P_∞ into P_2 (where we assume that $q^0 = 1$).*

Proof The submonoid of $\{p, q\}^*$ generated by $\{\theta_\infty(p_i) : i \in \mathbb{N}\}$ is a free submonoid, as before. Hence, by the construction of P_∞ as the inverse hull of the free monoid on a countably infinite set of generators, the inverse hull of $\{\theta(p_i)\}$ is isomorphic to P_∞, and is a submonoid of P_2. □

Definitions 1.11

We refer to the above embedding as the *right-associative* embedding of P_∞ into P_2.

Chapter 2

Representations of polycyclic monoids

2.1 Introduction

In this chapter, we consider the representation theory of polycyclic monoids as partial bijective maps on the natural numbers, and relate this to self-embedding properties of $\mathbb{N}$. We then demonstrate how an embedding of an inverse monoid with a zero into the symmetric inverse monoid on a set X gives a natural topology on X, which in the case of the Cantor set and the polycyclic monoid, gives the standard topology on the Cantor set. Finally, we give an embedding of $I(\mathbb{N})$ into the C^*–algebra of bounded linear operators on the Hilbert space l^2, and use this to construct an embedding of P_2 into $B(l^2)$.

2.2 The product and coproduct on the natural numbers

We establish various procedures for constructions involving, and determined by, embeddings of P_2 and P_∞ in $I(\mathbb{N})$. In the following section, we identify generators of polycyclic monoids with their images.

Definitions 2.1

An embedding θ of an inverse semigroup S into a symmetric inverse monoid $I(X)$ is called an *effective representation* when, for all $x \in X$, there exists $s \in S$ satisfying $x \in dom(\theta(s))$. See [30] for more details of the theory of effective representations. We will define a special case of this for polycyclic monoids; however we first require the following:

Lemma 1 *Consider the polycyclic monoid P_α on a countable set of generators, which we denote $\{p_i : i \in I\}$. Then for all $i \neq j \in I$, the elements $p_i^{-1}p_i$ and $p_j^{-1}p_j$ are disjoint idempotents.*

Proof Idempotency follows immediately, since both elements are of the form $a^{-1}a$ for some $a \in P_\alpha$. Also, $p_i^{-1}p_i(p_j^{-1}p_j)^{-1} = p_i^{-1}p_ip_j^{-1}p_j = 0$, since $p_ip_j^{-1} = 0$ for $i \neq j$. Similarly, $(p_i^{-1}p_i)^{-1}p_j^{-1}p_j = p_i^{-1}p_ip_j^{-1}p_j = 0$, as before. Therefore, $p_i^{-1}p_i \perp p_j^{-1}p_j$. □

We define a special case of effective representations, and say that an embedding of a polycyclic monoid P_α into a symmetric inverse monoid $I(X)$ is *strong* when it satisfies

$$\bigvee_{i \in \alpha} p_i^{-1}p_i = 1.$$

In what follows, we will relate strong embeddings of P_2 and P_∞ into $I(\mathbb{N})$ to bijections between the natural numbers and their coproduct and Cartesian product, where the *coproduct* $X \sqcup Y$ of two sets is defined by $X \sqcup Y = X \times \{0\} \cup Y \times \{1\}$.

Lemma 2 *There exists a strong embedding of P_2 into $I(\mathbb{N})$.*

Proof Define the partial bijective maps $p^{-1}, q^{-1} : \mathbb{N} \to \mathbb{N}$ by $p^{-1}(n) = 2n$ and $q^{-1}(n) = 2n+1$. Then these maps, together with their generalised inverses, satisfy $pp^{-1} = 1 = qq^{-1}$ and $qp^{-1} = pq^{-1} = 0$. Therefore, as P_2 is congruence-free, this is an embedding of P_2 into $I(\mathbb{N})$. To see that this embedding is strong, note that $dom(p) \cup dom(q) = \mathbb{N}$, and so $p^{-1}p \vee q^{-1}q = 1$. Hence we have constructed a strong embedding of P_2 into $I(\mathbb{N})$. □

Definitions 2.2

We refer to the above embedding as the *interleaving embedding* of P_2 into $I(\mathbb{N})$.

Lemma 3 *A bijection $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$, determines, and is determined by a strong embedding of P_2 in $I(\mathbb{N})$.*

Proof

Determining an embedding from a bijection

Let $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$ be a bijection. We define maps $p^{-1}, q^{-1} \in I(\mathbb{N})$ by $p^{-1}(n) = \phi(n, 0)$ and $q^{-1}(n) = \phi(n, 1)$. Using this definition, p^{-1}, q^{-1} both have domain $\mathbb{N}$, and they also have disjoint images (by the injectivity of ϕ). Hence, if we denote their (generalised) inverses by p and q respectively, then $pp^{-1} = 1 = qq^{-1}$ and $qp^{-1} = 0 = pq^{-1}$. Therefore, as p^{-1}, q^{-1} are partial bijective maps that satisfy the axioms for the generators of P_2 and are contained in $I(\mathbb{N})$, we can deduce, by the congruence-freeness of P_2, that they generate an embedding of P_2 into $I(\mathbb{N})$. Finally, to show that this embedding is strong, $dom(p) = \phi(\mathbb{N} \times \{0\})$ and $dom(q) = \phi(\mathbb{N} \times \{1\})$. Therefore, $dom(p) \cup dom(q) = \mathbb{N}$, and so $p^{-1}p \vee q^{-1}q = 1$. Hence, a bijection $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$ determines a strong embedding of P_2 into $I(\mathbb{N})$.

Determining a bijection from an embedding

Given a strong embedding of P_2 into $I(\mathbb{N})$, we have, by definition, a pair of partial bijective maps p^{-1}, q^{-1} satisfying $pp^{-1} = 1 = qq^{-1}$, $qp^{-1} = 0 = pq^{-1}$, and $p^{-1}p \vee q^{-1}q = 1$. We define a map $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$ by

$$\phi(n, i) = \begin{cases} p^{-1}(n) & i = 0 \\ q^{-1}(n) & i = 1 \end{cases}$$

This is injective, since p^{-1}, q^{-1} are injective maps, and as $qp^{-1} = 0 = pq^{-1}$, we can deduce that p^{-1}, q^{-1} have disjoint images. Also, since $p^{-1}p \vee q^{-1}q = 1$, we know that $im(p^{-1}) \cup im(q^{-1}) = \mathbb{N}$. Therefore $im(\phi) = \mathbb{N}$, and so ϕ is surjective. We can then deduce that a strong embedding of P_2 into $I(\mathbb{N})$ determines a bijection from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$. Therefore, a strong embedding uniquely determines a bijection, and vice versa. □

Lemma 4 *A bijection $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, determines and is determined by a strong embedding of P_∞ in $I(\mathbb{N})$.*

Proof

Determining an embedding from a bijection

Given a bijection $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, then we can define an infinite countable family of partial bijective maps $p_i^{-1} : \mathbb{N} \to \mathbb{N}$ $(i \in \mathbb{N})$ by $p_i(n) = \psi(i, n)$ $\forall i, n \in \mathbb{N}$. These maps are injective, since ψ is a bijection, and so we can define their partial inverses, $\{p_i : i \in \mathbb{N}\}$. By definition, $dom(p_i^{-1}) = \mathbb{N}$ for all $i \in \mathbb{N}$, and so $p_i p_i^{-1} = 1$. As ψ is injective, $im(p_i^{-1}) \cap im(p_j^{-1}) = \emptyset$ for $i \neq j$, and so $p_i p_j^{-1} = 0$. Therefore, $p_i p_j^{-1} = \delta_{ij}$, the required condition for a generating set of P_∞. Also, as P_∞ is congruence-free, the inverse subsemigroup of $I(\mathbb{N})$ generated by $\{p_i^{-1} : i \in \mathbb{N}\}$ is isomorphic to P_∞. To prove that this embedding is strong, we require $\bigcup_{i=0}^{\infty} p_i^{-1}(\mathbb{N}) = \mathbb{N}$; this follows trivially from the fact that ψ is a bijection, and from the definition of $\{p_i^{-1} : i \in \mathbb{N}\}$ in terms of ψ. Hence, any bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ determines a strong embedding of P_∞ into $I(\mathbb{N})$.

Determining a bijection from an embedding

A strong embedding of P_∞ into $I(\mathbb{N})$ specifies a countably infinite set of partial bijective maps, $\{p_i^{-1} : i \in \mathbb{N}\}$, satisfying $p_i p_j^{-1} = \delta_{ij}$ and $\bigvee_{i=0}^{\infty} p_i^{-1} p_i = 1$. We define a map $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $\psi(i, n) = p_i^{-1}(n)$. This map is injective, since p_i^{-1}, p_j^{-1} have disjoint images for $i \neq j$, and the $\{p_i^{-1}\}$ are all injective maps. Also, $Im(\psi) = \cup_{i=0}^{\infty} p_i^{-1}(\mathbb{N})$; however, we know that $\cup_{i=0}^{\infty} Im(p_i^{-1}) = \mathbb{N}$, since $\bigvee_{i=0}^{\infty} p_i^{-1}p_i = 1$. Therefore we can deduce that $Im(\psi) = \cup_{i=0}^{\infty} p_i^{-1}(\mathbb{N}) = \mathbb{N}$ and so ψ is also surjective. Therefore, a strong embedding of P_∞ into $I(\mathbb{N})$ determines a bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$, and so a strong embedding uniquely determines a bijection, and vice versa. □

Lemma 5 *A strong embedding of P_∞ into $I(\mathbb{N})$ determines a strong embedding of P_2 into $I(\mathbb{N})$.*

Proof (We will identify the generators of P_∞ with their images under this embedding). The construction is then as follows:
First note that, for any $i \neq j$,

$$p_{2i}^{-1}p_i(p_{2j}^{-1}p_j)^{-1} = p_{2i}^{-1}p_ip_j^{-1}p_{2j} = p_{2i}^{-1}\delta_{ij}p_{2j} = 0,$$

and

$$(p_{2j}^{-1}p_j)^{-1}p_{2i}^{-1}p_i = p_j^{-1}p_{2j}p_{2i}^{-1}p_i = p_j^{-1}\delta_{2i,2j}p_i = 0.$$

Therefore, $p_{2i}^{-1}p_i$ and $p_{2j}^{-1}p_j$ are disjoint; similarly, $p_{2i+1}^{-1}p_i$ and $p_{2j+1}^{-1}p_j$ are disjoint. Therefore, we can define

$$p^{-1} = \bigvee_{i=0}^{\infty} p_{2i}^{-1}p_i \ , \ q^{-1} = \bigvee_{i=0}^{\infty} p_{2i+1}^{-1}p_i,$$

and let p, q be their partial inverses. We show that these satisfy the relations for a strong embedding of P_2 into $I(\mathbb{N})$. First,

$$pp^{-1} = \left(\bigvee_{i=0}^{\infty} p_{2i}^{-1}p_i\right)^{-1}\left(\bigvee_{j=0}^{\infty} p_{2j}^{-1}p_j\right) = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}p_{2i}p_{2j}^{-1}p_j$$

$$= \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}\delta_{2i,2j}p_j = \bigvee_{i=0}^{\infty} p_i^{-1}p_i = 1,$$

since the embedding of P_∞ is strong. Similarly,

$$qq^{-1} = \left(\bigvee_{i=0}^{\infty} p_{2i+1}^{-1}p_i\right)^{-1}\left(\bigvee_{j=0}^{\infty} p_{2j+1}^{-1}p_j\right) = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}p_{2i+1}p_{2j+1}^{-1}p_j$$

$$= \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}\delta_{2i+1,2j+1}p_j = \bigvee_{i=0}^{\infty} p_i^{-1}p_i = 1,$$

again, since the embedding of P_∞ is strong. Hence $p^{-1}p = 1 = q^{-1}q$. Also,

$$pq^{-1} = \left(\bigvee_{i=0}^{\infty} p_{2i}^{-1}p_i\right)^{-1}\left(\bigvee_{j=0}^{\infty} p_{2j+1}^{-1}p_j\right)$$

$$= \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}p_{2i}p_{2j+1}^{-1}p_j = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_i^{-1}\delta_{2i,2j+1}p_j = 0,$$

since $2i$ is even, and $2j+1$ is odd, for all $i, j \in \mathbb{N}$, and

$$qp^{-1} = \left(\bigvee_{i=0}^{\infty} p_{2i+1}^{-1} p_i\right)^{-1} \left(\bigvee_{j=0}^{\infty} p_{2j}^{-1} p_j\right)$$

$$= \bigvee_{i=0}^{\infty} \bigvee_{j=0}^{\infty} p_i^{-1} p_{2i+1} p_{2j}^{-1} p_j = \bigvee_{i=0}^{\infty} \bigvee_{j=0}^{\infty} p_i^{-1} \delta_{2i+1,2j} p_j = 0,$$

since $2i+1$ is odd, and $2j$ is even. Therefore, $pq^{-1} = 0 = qp^{-1}$ and so p, q satisfy the axioms for the generators of the polycyclic monoid P_2. Hence we have constructed an embedding of P_2 into $I(\mathbb{N})$ derived from a strong embedding of P_∞ into $I(\mathbb{N})$. To see that this embedding is strong,

$$p^{-1}p \vee q^{-1}q =$$

$$\left(\bigvee_{i=0}^{\infty} p_{2i}^{-1} p_i\right) \left(\bigvee_{j=0}^{\infty} p_{2j}^{-1} p_j\right)^{-1} \vee \left(\bigvee_{k=0}^{\infty} p_{2k+1}^{-1} p_k\right) \left(\bigvee_{l=0}^{\infty} p_{2l+1}^{-1} p_l\right)^{-1}$$

$$= \left(\bigvee_{i=0}^{\infty} \bigvee_{j=0}^{\infty} p_{2i}^{-1} \delta_{ij} p_{2j}\right) \vee \left(\bigvee_{k=0}^{\infty} \bigvee_{l=0}^{\infty} p_{2k+1}^{-1} \delta_{kl} p_{2l+1}\right)$$

$$= \left(\bigvee_{i=0}^{\infty} p_{2i}^{-1} p_{2i}\right) \vee \left(\bigvee_{k=0}^{\infty} p_{2k+1}^{-1} p_{2k+1}\right)$$

$$= \bigvee_{r=0}^{\infty} p_r^{-1} p_r = 1.$$

Hence this embedding of P_2 into $I(\mathbb{N})$ is strong, and our result follows. □

Lemma 6 *A bijection bijection $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, provided the bijection $\phi_1 : \mathbb{N} \to \mathbb{N}$, defined by $\phi_1(n) = \phi(n, 1)$ satisfies the 'no fixed point' condition*

$$\bigcap_{i=0}^{\infty} \phi_1^i(\mathbb{N}) = \emptyset.$$

Also, the interleaving map from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$, constructed by applying Lemma 3 to the interleaving embedding of P_2 into $I(\mathbb{N})$, satisfies this condition.

Proof Given a bijective map $\phi : \mathbb{N} \times \{0, 1\} \to \mathbb{N}$, we construct an injective map $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, as follows: We first define injective maps $\phi_0, \phi_1 : \mathbb{N} \to \mathbb{N}$ by $\phi_0(n) = \phi(n, 0)$, $\phi_1(n) = \phi(n, 1)$. These maps are clearly injective and have disjoint images, since ϕ is bijective (they are the p^{-1}, q^{-1} generators of the polycyclic monoid derived from a bijection from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$, given previously). We then define a map $\psi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ by $\psi(x, y) = \phi_1^y(\phi_0(x))$, where ϕ_1^0 is the identity map.

We first show that ψ is injective; assume that $\psi(x, y) = \psi(a, b)$ for some $a, b \in \mathbb{N}$. That is, $\phi_1^y(\phi_0(x)) = \phi_1^b(\phi_0(a))$. Then assume (without loss of generality) that $y = b + k$ for some $k \in \mathbb{N}$, so that $\phi_1^{b+k}(\phi_0(x)) = \phi_1^b(\phi_0(a)$. However, as ϕ_1 is injective, this implies that $\phi_i^k(\phi_0(x)) = \phi_0(a)$, which implies that $k = 0$ since ϕ_0 and ϕ_1 have disjoint images. We can then deduce that $\phi_0(x) = \phi_0(a)$, and so $x = a$, by the injectivity of ϕ_0. Therefore ψ is injective.

Also, from the definition $\psi(x, y) = \phi_1^y(\phi_0(x))$, we can see that, since ϕ is bijective,

$$\bigcup_{i=1}^{n} \phi_1^i \phi_0(\mathbb{N}) = \mathbb{N} \setminus \phi_1^{n+1}(\mathbb{N}).$$

Therefore, $im(\psi) = \mathbb{N}$ if and only if

$$\bigcap_{i=0}^{\infty} \phi_1^i(\mathbb{N}) = \emptyset.$$

This result clearly holds for the 'interleaving' bijection $\phi(x, i) = 2x + i$, since $\phi_1^j(n) < \phi_1^{j+1}(n)$ for all $n, j \in \mathbb{N}$. Therefore, we have proved that a bijection ψ derived from a bijection ϕ that satisfies the condition given is bijective, and the standard interleaving embedding satisfies this condition. □

We give an example of a bijection from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$ that does not satisfy this condition, and hence does not determine a bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ by the above construction.

Example Consider the interleaving bijection ϕ given in Lemma 6, and construct a function ϕ' by $\phi'(n, 0) = \phi(n, 0)$ and

$$\phi'(n, 1) = \begin{cases} 3 & n = 0 \\ 1 & n = 1 \\ \phi(n, 1) & \text{otherwise.} \end{cases}$$

It is trivial to check that ϕ' is a bijection from $\mathbb{N} \times \{0, 1\}$ to $\mathbb{N}$, since ϕ is, but $\bigcap_{i=0}^{\infty} \phi_1^i(\mathbb{N}) = \{1\} \neq \emptyset$.

Lemma 7 *A strong embedding of P_2 into $I(\mathbb{N})$ determines a strong embedding of P_∞ into $I(\mathbb{N})$, provided the embedding of P_2 satisfies*

$$\bigcap_{i=0}^{\infty} q^{-i}(\mathbb{N}) = \emptyset.$$

Proof The result follows from Lemma 6, and the correspondence between strong embeddings of P_∞ and bijections from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$ (Lemma 4), and the correspondence between strong embeddings of P_2 and bijections from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$ (Lemma 3). The embedding of P_∞ is given by $p_i^{-1} = q^{-i}p^{-1}$, the right-associative embedding of Definitions 1.11 of Chapter 1. Hence we have constructed an embedding of P_∞ into $I(\mathbb{N})$. □

2.3 Topological spaces and inverse monoids

In this section, we demonstrate how a representation of an inverse semigroup as partial bijective maps on a set X determines a topology on X, and show how this construction gives the standard topology on the Cantor set from an embedding of the polycyclic monoid on two generators.

Definitions 2.3

Given an arbitrary set X, a *topology* on it is defined to be a set of subsets of X, $\tau \subseteq P(X)$, that satisfies the following conditions:

1. $\emptyset, X \in \tau$.
2. $A, B \in \tau$ implies that $A \cap B \in \tau$.
3. $\bigcup_{k \in \kappa} k \in \tau$ for any subset of the topology, $\kappa \subseteq \tau$.

So, a topology on X is a collection of subsets of X containing X and the empty set, that is closed under unions and finite intersections. The study of topologies on sets is an extremely wide-ranging subject; we refer to [7] for an introduction to the basic theory of continuity and homeomorphisms. The elements of a topology τ are called *open sets*, and a subset of the form $X \setminus T$, with $T \in \tau$, is called a *closed set*. Note that a subset of X may be both open and closed. In particular, X and $\emptyset$ are both open and closed.

A *basis* B for a topology τ on a set X is a set of subsets of X that satisfies

- $b \in B$ implies $b \in \tau$.
- $T \in \tau$ implies $T = \cup_{i \in I} b_i$ for some subset of B, $\{b_i : i \in I\} \subseteq B$.

If B is a basis for τ, then τ is said to be *generated by* B. Also, given a set of subsets B whose union contains X, which is closed under finite intersections, the set τ consisting of all unions of members of B (including the empty one) is the topology *generated by* the basis B. Note that, for a topology generated by a basis B, a function between topological spaces is continuous if the inverse image of every basic open set is an open set.

We give an example of a topological space, determined by a basis, that will be important in the theory of polycyclic monoids.

Definitions 2.4

The *Cantor space* Ω is defined to be the set of all countably infinite sequences of $\{0, 1\}$, written $\Omega = \{0, 1\}^{\mathbb{N}}$, and the *Cantor topology*, or *standard topology on the Cantor set*, is defined to be the topology generated by the basic open sets $C = \{C_s : s \in \{0, 1\}^*\}$, where $C_s = \{s\omega : \omega \in \Omega\}$.

Proposition 8 *The Cantor set, with the standard topology, is Hausdorff.*

Proof Consider two distinct points $\omega, \omega' \in \Omega$. Then there exists $n \in \mathbb{N}$ such that the two sequences represented by ω, ω' differ after the n^{th} place. Hence, if u, u' are the two members of $\{0,1\}^*$ given by the first $n+1$ places of ω and ω' respectively, then clearly the open set C_u contains ω but not ω', and $C_{u'}$ contains ω' but not ω. Hence the Cantor set, with this topology, is Hausdorff. □

2.3.1 Topologies generated by inverse monoids

We demonstrate how embeddings of inverse monoids into symmetric inverse monoids can be used to construct topologies on sets.

Theorem 9 *Let X denote a set, and let S be an inverse submonoid of $I(X)$ that satisfies $0 \in S$. Then the partial identities on X given by the idempotents of S form a basis for a topology on X.*

Proof Recall from Definitions 1.5 that for all $e^2 = e, f^2 = f \in I(X)$, $e = 1_A$ and $f = 1_B$ for some $A, B \subseteq X$. Also, $ef = fe = 1_{A \cap B}$. Hence, if we denote the set of all domains of elements of S by $B(S)$, then $B(S)$ is closed under finite intersections. Finally, $X \in B(S)$, as $1 = 1_X$, and $\emptyset \in B(S)$, since $0 = 1_\emptyset$. Therefore, $B(S)$ is a collection of subsets of X that contains $X, \emptyset$, and is closed under finite intersections, and so is a basis for a topology on X. □

Definitions 2.5

Let S be an inverse submonoid of $I(X)$, as above. We denote the topology generated by the basis $B(S)$ by $Top(S) \subseteq P(X)$.

Proposition 10 *Every element of S is a partial homeomorphism of the topological space $(X, Top(S))$.*

Proof Consider arbitrary $f \in S$, and $A \in B(S)$. Then $1_A = e$, for some $e \in S$. Therefore, $f^{-1}(A) = dom(f^{-1}ef)$, and $(f^{-1}ef)^2 = f^{-1}eff^{-1}ef = f^{-1}eff^{-1}f = f^{-1}ef$, as idempotents commute in an inverse semigroup. Therefore, $f^{-1}ef$ is an idempotent of S, and so $f^{-1}(A) \in B(S)$. Hence f is continuous in $(X, Top(S))$, and as every element f of an inverse semigroup is of the form $f = g^{-1}$ for some $g \in S$, then f^{-1} is also continuous. Also, as f is a partial bijective map (by definition of $I(X)$), f is a partial homeomorphism. □

Corollary 11 *Every element of the disjoint closure of S in $I(X)$ is a partial homeomorphism.*

Proof By Proposition 5 of Chapter 1, a disjoint set of elements of $I(X)$ is a set of partial bijective maps that have disjoint domains / images, and the disjoint join of a set of disjoint elements in

$I(X)$ is their set-theoretic union. Hence, any element of $DC_X(S)$ is the union of a set of disjoint partial homeomorphisms, and so is also a partial homeomorphism. □

2.3.2 The topology generated by P_2

We demonstrate how the topology generated by P_2 gives rise to the standard topology on the Cantor set.

Lemma 12 *There exists a strong embedding of P_2 into $I(\Omega)$.*

Proof We define the following partial bijective maps on Ω:

$$p^{-1}(\omega) = 0\omega \ , \ q^{-1}(\omega) = 1\omega,$$

$$p(0\omega) = \omega \ , \ q(1\omega) = \omega,$$

for all $\omega \in \Omega$. Note that p^{-1} and q^{-1} have domain Ω, and p and q have domains 0Ω and 1Ω respectively. From these definitions, $pp^{-1} = 1 = qq^{-1}$ and $pq^{-1} = 0 = qp^{-1}$, where 0 is the empty map. Hence, these elements of $I(\Omega)$ satisfy the axioms for the generators of an embedding of P_2 into $I(\Omega)$. Therefore, as they are clearly non-trivial, and as we have seen in Definitions 1.10 that polycyclic monoids are congruence-free, we have defined an embedding of P_2 into $I(\Omega)$. Also, $p^{-1}(\Omega) \cup q^{-1}(\Omega) = \Omega$; therefore, $p^{-1}p \vee q^{-1}q = 1$, and so this embedding is strong. □

We refer to this embedding of P_2 into $I(\Omega)$ as the *standard embedding* into the Cantor set.

Theorem 13 *The topology on the Cantor set generated by the standard embedding of P_2 is the standard topology on the Cantor set.*

Proof From the definition of the standard embedding, the members of $B(P_2)$ are given by $w\Omega$, for all $w \in \{0,1\}^*$, since all the domains of embeddings of elements of P_2 are of this form. Therefore, the basis of $Top(P_2)$ is the same as the basis for the standard topology on Ω. Our result then follows from the definition of the topology generated by a basis. □

2.3.3 Properties of topologies generated by P_2

Let X be a set, with the polycyclic monoid on two generators strongly embedded into $I(X)$, and consider the topology on X generated by the embedding of P_2 (we identify the elements of P_2

with their images under the embedding, for clarity). We give a characterisation of the members of $B(P_2)$.

Definitions 2.6

For each $w \in \{p^{-1}, q^{-1}\}^*$, we define a subset of X by $O_w = \{w(x) : x \in X\}$.

Lemma 14 $B(P_2) = \{O_w : w \in \{p^{-1}, q^{-1}\}^*\}$.

Proof Recall that all idempotents of an inverse semigroup are of the form ff^{-1} for some element f. Therefore, all members of $B(P_2)$ are of the form $im(f)$, for some $f \in P_2$.

Next consider an arbitrary element of P_2, written in canonical form (as given in Definitions 1.10) as $v^{-1}w$. Then the partial identity on its image is given by $v^{-1}w(v^{-1}w)^{-1} = v^{-1}ww^{-1}v = v^{-1}v$, by the composition of elements in canonical form. However, $v \in \{p, q\}^*$, and so $v^{-1}v$ is the partial identity on $v^{-1}(X)$, and so all elements of an embedding of P_2 in $I(X)$ have image in $\{p^{-1}, q^{-1}\}^*(X)$. Therefore, our result follows. □

Proposition 15 *Given $O_r, O_s \in B(P_2)$, then*

$$O_r \cap O_s = \begin{cases} O_r & s \text{ is a postfix of } r \\ O_s & r \text{ is a postfix of } s \\ \emptyset & \text{otherwise.} \end{cases}$$

Proof By definition, $O_r \cap O_s = \{r(x) : x \in X\} \cap \{s(x) : x \in X\}$. Let us assume without loss of generality that the length of r is greater than or equal to the length of s. If $r = s$, then $O_r \cap O_s = O_r = O_s$, and so the result follows trivially. Now assume that $r \neq s$. If s is not a postfix of r, then the set $O_r \cap O_s$ is empty, since p^{-1}, q^{-1} have disjoint images. On the other hand, if is s is a postfix of r, then $O_r \subseteq O_s$, and so $O_r \cap O_s = O_r$. □

Proposition 16 *Any open set of $Top(P_2)$ can be represented as a disjoint union of sets from the basis $B(P_2)$.*

Proof Let $\bigcup_{j \in J} O_{w_j}$ be an open set of $\tau(B)$ (where $w_j \in \{p^{-1}, q^{-1}\}^*$ for all $j \in J$). Assume that this cannot be written as a disjoint union. We must then be able to find 2 distinct open sets

$$O_u, O_v \in \{O_{w_j} : w_j \in \{p^{-1}, q^{-1}\}^*\}_{j \in J},$$

that satisfy $O_u \cup O_v \neq O_w$ for any $w \in \{p^{-1}, q^{-1}\}^*$, and $O_u \cap O_v \neq \emptyset$. However, by Proposition 15, $O_u \cap O_v \neq \emptyset$ implies that $O_u \cap O_v = O_u$ or $O_u \cap O_v = O_v$, and so $O_u \cup O_v = O_v$ or $O_u \cup O_v = O_u$, contradicting our assumption. Therefore, any open set of $Top(P_2)$ can be represented as a disjoint union of members of the basis $B(P_2)$. □

Proposition 17 *For any pair of basic open sets O_r and O_s, there exists a unique element of P_2 that gives a continuous bijective map that takes O_r to O_s, and is undefined elsewhere.*

Proof Given a pair of open sets O_r, O_s, then by definition, $r^{-1}(O_r) = X$, and $s(X) = O_s$. Therefore, if we compose the two maps, $sr^{-1}(O_r) = O_s$. Also, $dom(r^{-1}) = O_r$, and $im(s) = O_s$, from Definitions 2.7, so this map is undefined elsewhere.

To show that this element is unique, note that, for all $r' \neq r \in \{p^{-1}, q^{-1}\}^*$ satisfying $O_r \subseteq dom(r'^{-1})$, we have, by Proposition 15, O_r is a strictly contained in $O_{r'}$. Conversely, for all $s' \neq s \in \{p^{-1}, q^{-1}\}^*$ satisfying $O_{s'} \subseteq s(X)$, we have that, by Proposition 15, $s(X)$ is strictly contained in $s'(X)$. Therefore, uniqueness follows by the requirement that the map is undefined elsewhere. □

2.4 Embedding $I(\mathbb{N})$ into $B(l^2)$

We show how the symmetric inverse monoid on the natural numbers, $I(\mathbb{N})$, can be embedded into the C^*–algebra of bounded linear operators on the Hilbert space of square-summmable sequences of complex numbers, $B(l^2)$. This gives a method of 'converting partial bijective maps into globally defined maps', by letting the members of $I(\mathbb{N})$ act on a basis set of a countably infinite dimensional Hilbert space. There is no space to present the full definition of Hilbert spaces and C^*-algebras here; see [47] for a readable introduction. However, we do present a special case.

Definitions 2.7

The Hilbert space l^2 is defined to be the countably infinite dimensional vector space of sequences of complex numbers $(z_0, z_1, z_2, \ldots)$ that satisfy

$$\sum_{i=0}^{\infty} z_i \overline{z_i} < \infty,$$

together with the *inner product*

$$\langle (a_0, a_1, a_2, \ldots) | (b_0, b_1, b_2, \ldots) \rangle = \sum_{i=0}^{\infty} a_i \overline{b_i}.$$

See [47] for the proof that this is a Hilbert space. A function $\Phi : l^2 \to l^2$ is called *linear* if it satisfies $\Phi(\alpha\mathbf{a} + \beta\mathbf{b}) = \alpha\Phi(\mathbf{a}) + \beta\Phi(\mathbf{b})$. It is called *bounded* if, for all $\mathbf{a} \in l^2$, there exists a fixed constant $K \in \mathbb{R}$ such that $\langle\Phi(\mathbf{a})|\Phi(\mathbf{a})\rangle \leq K\langle\mathbf{a}|\mathbf{a}\rangle$. The set of all bounded linear operators of l^2 forms a C^*-algebra (see [47] for an introduction to the theory of C^*-algebras), denoted by $B(l^2)$.

An *orthogonal basis set* for l^2 is a countable spanning set of linearly independent members of l^2, which we denote $\{\mathbf{b}_i\}_{i=0}^{\infty}$, satisfying $\langle\mathbf{b}_i|\mathbf{b}_j\rangle = 0$ for all $i \neq j$. If the additional condition $\langle\mathbf{b}_i|\mathbf{b}_i\rangle = 1$ is satisfied, for all $i \in \mathbb{N}$, the basis set is called *orthonormal.* Note that, given an orthogonal basis set $\{\mathbf{b}_i\}_{i=0}^{\infty}$ for l^2, any bounded linear function $\Psi \in B(l^2)$ can be written as a countably infinite matrix M_Ψ, over $\mathbb{C}$, whose entries are given by

$$m_{ij} = \langle\mathbf{b}_j|\Psi(\mathbf{b}_i)\rangle.$$

This is enough to uniquely specify the operator Ψ, and composition of operators in this form is given by the usual matrix multiplication. There is also an additional operation $(\ )^* : B(l^2) \to B(l^2)$, given by:
Ψ^* is the operator represented by the matrix M', where $m'_{ij} = \overline{m_{ji}}$. This is called the *conjugate transpose* of the matrix M.

We consider how $I(\mathbb{N})$ can be embedded into the multiplicative structure of $B(l^2)$. Let $\{\mathbf{b}_i\}_{i=0}^{\infty}$ be an orthogonal basis set for the Hilbert space l^2. We define a map $l : I(\mathbb{N}) \to B(l^2)$ by

$$l(f)\left(\sum_{i=0}^{\infty}\alpha_i\mathbf{b}_i\right) = \sum_{j=0}^{\infty}\alpha_j\mathbf{b}_{f(j)},$$

where $\mathbf{b}_{f(x)} = \mathbf{0}$, the null vector, when $f(x)$ is undefined.

Theorem 18

(i) *l is an injective monoid homomorphism,*

(ii) *$l(f^{-1}) = l(f)^*$, for all $f \in I(\mathbb{N})$,*

(iii) *$f \perp g \in I(\mathbb{N}) \Rightarrow l(f \vee g) = l(f) + l(g)$.*

Proof

(i) From the definition,

$$l(g)l(f)\left(\sum_{i=0}^{\infty}\alpha_i\mathbf{b}_i\right) = l(g)\left(\sum_{j=0}^{\infty}\alpha_j\mathbf{b}_{f(j)}\right) = \sum_{k=0}^{\infty}\alpha_k\mathbf{b}_{gf(k)} = l(gf)\left(\sum_{i=0}^{\infty}\alpha_i\mathbf{b}_i\right).$$

Therefore, l is a homomorphism. Also,

$$l(g) = l(f) \Leftrightarrow \left(\sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{f(j)}\right) = \left(\sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{g(j)}\right),$$

and so $f(j) = g(j)$ for all $j \in \mathbb{N}$. Therefore, $f = g \in I(\mathbb{N})$, and so l is an injective homomorphism. Finally, note that

$$l(1)\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i,$$

and so $l(1) = I$, the identity of the C^*-algebra, $B(l^2)$. Therefore, l is an injective monoid homomorphism.

(ii) Recall the definition of $(\ )^*$ in $B(l^2)$ as the 'conjugate transpose'. The $(\ )^*$ operator in $B(l^2)$ then satisfies $A^* = \overline{A}^T$, where $\overline{A_{ij}}$ is the complex conjugate of a_{ij}, and so $A^*_{ij} = a^*_{ji}$. Then, from Definitions 2.8, the matrix form M of $l(f)$ satisfies

$$M_{ij} = \begin{cases} 1 & \text{if } f(i) = j \\ 0 & \text{otherwise.} \end{cases}$$

Now, $\overline{1} = 1$, and so $M^* = M^T$, and M^T satisfies

$$M^T_{ij} = \begin{cases} 1 & \text{if } f(j) = i \\ 0 & \text{otherwise,} \end{cases}$$

or rather,

$$M^T_{ij} = \begin{cases} 1 & \text{if } f^{-1}(i) = j \\ 0 & \text{otherwise,} \end{cases}$$

and so $l(f^{-1}) = l(f)^*$.

(iii) Consider arbitrary $f \perp g \in I(\mathbb{N})$; then

$$(f \vee g)(x) = \begin{cases} f(x) & x \in dom(f) \\ g(x) & x \in dom(g) \\ \text{undefined} & \text{otherwise} \end{cases}$$

and so

$$l(f \vee g)\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \left(\sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{(f\vee g)(j)}\right) = \sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{f(j)} + \sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{g(j)},$$

since $f(j)$ defined implies $g(j)$ undefined, and vice versa. Therefore,

$$l(f \vee g)\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{f(j)} + \sum_{j=0}^{\infty} \alpha_j \mathbf{b}_{g(j)} = (l(f) + l(g))\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right).$$

Hence our result follows. □

Example The above theorem gives the following strong embedding of P_2 into $B(l^2)$, as the image of the interleaving embedding of P_2 into $I(\mathbb{N})$:

$$p\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{i=0}^{\infty} \alpha_{2i} \mathbf{b}_i \ , \ q\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{i=0}^{\infty} \alpha_{2i+1} \mathbf{b}_i,$$

and

$$p^{-1}\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{i=0}^{\infty} \alpha_i \mathbf{b}_{2i} \ , \ q^{-1}\left(\sum_{i=0}^{\infty} \alpha_i \mathbf{b}_i\right) = \sum_{i=0}^{\infty} \alpha_i \mathbf{b}_{2i+1}.$$

Consider the map T from the set of analytic square-integrable functions on the complex plane, that are defineable by their Taylor series expansions, to l^2. (the map is, of course, given by taking the coefficients of z^i in the Taylor series). It is a classical result of functional analysis (see, for example [24]) that this map is well-defined and injective. Then consider the orthogonal basis set given by

$$\begin{aligned}
\mathbf{b}_0 &= (1,0,0,0,\ldots) \\
\mathbf{b}_2 &= \left(0,\tfrac{1}{2},0,0,\ldots\right) \\
\mathbf{b}_3 &= \left(0,0,\tfrac{1}{3!},0,\ldots\right) \\
\mathbf{b}_4 &= \left(0,0,0,\tfrac{1}{4!},\ldots\right) \\
&\ldots
\end{aligned}$$

and use this to construct a strong embedding of P_2 into $B(l^2)$, as in the example above (we identify P_2 with its embedding, for clarity). An example of a bounded linear operator on this basis is given by

$$T(f) \mapsto T(\frac{d}{dz}f) \ \ \forall f : \mathbb{C} \to \mathbb{C}.$$

We abuse notation, and denote this by $\frac{d}{dz}$.

Proposition 19 *With the above notation,*

1. $p^{-1}T(e^z) = T(e^z) = q^{-1}T(e^z)$,

2. $\frac{d}{dz}p^{-1} = q^{-1}$,

3. $p^{-1}pT(e^{iz}) = T(cos(z))$,

4. $q^{-1}qT(e^{iz}) = T(isin(z))$,

5. $e^{iz} = cos(z) + isin(z)$.

Proof We denote the sum $\sum_{i=0}^{\infty} a_i \mathbf{b_i}$ by $(a_0, a_1, a_2, \ldots)$ when this sum is defined. Then by the basic construction of Taylor series,

$$\begin{aligned}
T(sin(z)) &= (0,1,0,-1,0,1,0,-1,\ldots) \\
T(cos(z)) &= (1,0,-1,0,1,0,-1,0,\ldots) \\
T(e^z) &= (1,1,1,1,1,1,1,1,1,\ldots) \\
T(e^{iz}) &= (1,i,-1,-i,1,i,-1,-i,\ldots) \\
\frac{d}{dz}(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\ldots) &= (\alpha_1,\alpha_2,\alpha_3,\alpha_4,\ldots) \\
p^{-1}(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\ldots) &= (\alpha_0,0,\alpha_1,0,\alpha_2,\ldots) \\
q^{-1}(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\ldots) &= (0,\alpha_0,0,\alpha_1,0,\ldots) \\
p(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\ldots) &= (\alpha_0,\alpha_2,\alpha_4,\ldots) \\
q(\alpha_0,\alpha_1,\alpha_2,\alpha_3,\ldots) &= (\alpha_1,\alpha_3,\alpha_5,\ldots)
\end{aligned}$$

Therefore, 1, 2, 3 and 4 follow immediately. 5 follows from 3 and 4, and from the fact that the embedding of P_2 is strong, so $p^{-1}p + q^{-1}q = 1$, and from the injectivity of the map T. □

Chapter 3

Applications of polycyclic monoids to rings

3.1 Introduction

We demonstrate how matrices of different orders over a unitary ring form a symmetric monoidal category, and give applications of polycyclic monoids to ring theory. We show (as a generalisation of a construction of J-Y Girard for the C^*–algebra $B(l^2)$, as found in [20]) how an embedding of P_2 into a ring R allows us to define a composition and addition preserving map from the category of matrices over R into R, which gives isomorphisms between the matrix rings over R, and corners of the ring R, determined by some idempotent e. We use this to construct an isomorphism between a ring R and its matrix rings, when $e = 1$. We then demonstrate how these self-embedding results allow us to define a ring homomorphism from $R \times R$ to R which gives R the structure of a one-object symmetric monoidal category, apart from the unit elements. Finally, these results are applied to the construction of the K_0 group of a ring R. We prove that an embedding of P_2 into a ring R allows us to construct the K_0 group of R solely from the idempotents of R; this greatly simplifies the construction of $K_0(R)$. This construction is a ring-theoretic version of the 'Splitting Idempotents' technique used in C^*-algebra theory, which can be found in [48].

3.2 Categories arising from rings

Definitions 3.1

The category **Ring** is given by $R \in Ob(\mathbf{Ring})$ iff R is a ring, and $f \in \mathbf{Ring}(R, S)$ iff f is a ring homomorphism from R to S. We define **IRing** to be the subcategory of **Ring** consisting of all rings with an identity, together with all identity-preserving homomorphisms.

Given $R \in Ob(\mathbf{IRing})$, we define the *matrix category* of R, which we denote $\mathbf{Mat_R}$, to have objects given by $Ob(\mathbf{Mat_R}) = \mathbb{N}$ and morphisms given by $A \in \mathbf{Mat_R}(a,b)$ iff A is a $b \times a$ matrix over R, together with the usual composition of matrices. (We denote the (0×0) matrix by $()$. The *category of proper matrices* over R is defined to be the subcategory of $\mathbf{Mat_R}$ where all matrices over R are of order $i \times j$ with $i, j \geq 2$. We denote this category by $\mathbf{Mat}_{\mathbf{R}}^{\geq 2}$.

We define the *functor category*, $\mathbf{Fun_{Mat}}$, to have matrix categories of rings as objects, and functors between matrix categories as morphisms, so $\mathbf{Mat_R} \in Ob(\mathbf{Fun_{Mat}})$ iff $R \in \mathbf{IRing}$.

Proposition 1 *There exists a functor from* **IRing** *to* $\mathbf{Fun_{Mat}}$

Proof Define $M : \mathbf{IRing} \to \mathbf{Fun_{Mat}}$ by:
On objects, $M(R) = \mathbf{Mat_R}$,
On morphisms, the functor $M(f)$ is given by

- $M(f)(n) = n \in Ob(\mathbf{Mat_S})$, for all $n \in \mathbb{N}$,
- $(M(f)(A))_{ij} = f(A_{i,j})$, for all $0 \leq i < b$, $0 \leq j < a$ where $A \in \mathbf{Mat_R}(a,b)$.

As f is an identity-preserving ring homomorphism, this is clearly a functor from **IRing** to $\mathbf{Fun_{Mat}}$.
□

3.2.1 $\mathbf{Mat_R}$ as a symmetric strict monoidal category

Definitions 3.2

We define a map $\sqcup : \mathbf{Mat_R} \times \mathbf{Mat_R} \to \mathbf{Mat_R}$ as follows:

- On objects, $x \sqcup y = x + y \in Ob(\mathbf{Mat_R})$,
- On morphisms, given $A \in \mathbf{Mat_R}(a,b)$, $B \in \mathbf{Mat_R}(c,d)$, then $A \sqcup B = \begin{pmatrix} A & \mathbf{0} \\ \mathbf{0} & B \end{pmatrix}$.

Theorem 2 $\sqcup$ *is an addition-preserving functor from* $\mathbf{Mat_R} \times \mathbf{Mat_R}$ *to* $\mathbf{Mat_R}$ *that gives* $\mathbf{Mat_R}$ *the structure of a symmetric strict monoidal category.*

Proof Let X, Y, U, V be matrices over R of orders $a \times b$, $r \times s$, $b \times c$, and $s \times t$ respectively. Then it is immediate from the definition of matrix multiplication that

$$\begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & V \end{pmatrix} = \begin{pmatrix} XU & 0 \\ 0 & YV \end{pmatrix},$$

and so $(X \sqcup Y)(U \sqcup V) = (XU \sqcup YV)$. Hence $\sqcup$ is a functor from $\mathbf{Mat_R} \times \mathbf{Mat_R}$ to $\mathbf{Mat_R}$. Also it is immediate from the definition that $\sqcup$ preserves addition.

Next, the operation $\sqcup$ is associative on objects of $\mathbf{Mat_R}$, since $+$ is associative on the natural numbers. For morphisms, it also follows directly from the definition, so

$$((A \sqcup B) \sqcup C) = \begin{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} & 0 \\ 0 & C \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & \begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix} \end{pmatrix} = (A \sqcup (B \sqcup C)).$$

For every pair of objects $x, y \in OB(\mathbf{Mat_R})$, we define a morphism by

$$S_{x,y} = \begin{pmatrix} 0 & I_y \\ I_x & 0 \end{pmatrix},$$

where $I_n = \bigsqcup_{i=1}^{n}(1)$. Then from the definition of composition, for all $M \in \mathbf{Mat_R}(a, x)$ and $N \in \mathbf{Mat_R}(b, y)$, this family of morphisms satisfies $S_{x,y}(M \sqcup N)S_{b,a} = (N \sqcup M)$.

Finally, the 0×0 empty matrix $()$ satisfies $A \sqcup () = A = () \sqcup A$, by definition of $\sqcup$. Therefore, for every ring R with identity, $(\mathbf{Mat_R}, \sqcup)$ is a symmetric strict monoidal category. $\square$

3.3 Maps between categories of matrices

In what follows, we will consider the category theory of matrices over rings in which polycyclic monoids are embedded. These embeddings lead to functors between different categories of matrices, and functors from a category of matrices to itself, under certain conditions. A corollary of these results is the construction of isomorphisms between (corners of) matrix rings of different orders, as a generalisation to arbitrary rings of a construction of J-Y Girard, on the C^*–algebra $B(l^2)$, found in [20]. We then show that under certain conditions, these isomorphisms lead to isomorphisms between $M_n(R)$ and R, for all $n \in \mathbb{N}$.

3.3.1 Preliminaries on embeddings of monoids in rings

Definitions 3.3

Let $e^2 = e$ be an idempotent of a semigroup S. The *local submonoid* of S determined by e, denoted eSe, is the subsemigroup of S specified by $eSe = \{ewe : w \in S\}$, together with the inherited composition. Note that e is an identity for eSe, and so eSe is a monoid. There is a similar definition for rings, where the *corner* of the ring R determined by the idempotent e is specified by $eRe = \{ere : r \in R\}$, together with the inherited composition and addition. Note that e is then a multiplicative identity for the ring eRe, and so eRe is a ring with an identity. We make a similar definition for categories. Let E be a set of idempotents of $\mathbf{C}$ indexed by the objects of the category $\mathbf{C}$, so $E = \{E_x^2 = E_x \in \mathbf{C}(x,x) : x \in Ob(\mathbf{C})\}$. Then the *corner of* $\mathbf{C}$ *determined by* E is specified by $E\mathbf{C}E = \{E_y F E_x : F \in \mathbf{C}(x,y)\}$, together with the usual composition and objects. Note that the identities at the objects of ECE are given by E_x, rather than by 1_x. Also, the endomorphism monoid of the object X in the category ECE is clearly the local submonoid $E_X\mathbf{C}(X,X)E_X$.

Let R be a ring with identity. By analogy with Definitions 1.6, Chapter 1, we say that idempotents $e, e' \in R$ are *disjoint* if they satisfy $ee' = 0 = e'e$. We denote this by $e \perp e'$. We say that P_n is *embedded in* R if there is a monoid homomorphism (preserving zeros) from P_n to the multiplicative monoid of R. By analogy with Definition 2.1, Chapter 2, we say that P_n is *strongly embedded* if it satisfies the further condition

$$\sum_{i=0}^{n-1} p_i^{-1} p_i = 1.$$

(We will present a unifying categorical structure for the different definitions of weak and strong embeddings of P_2 in Chapter 4). In what follows we shall identify embeddings of P_n with their images to simplify notation, unless the distinction is important.

Lemma 3 *Let R be a unital ring. If P_n is embedded in R, then*

(i) $e_n = \sum_{i=0}^{n-1} p_i^{-1} p_i$ *is an idempotent.*

(ii) *If P_2 is strongly embedded in R, then so is P_n for all $n \geq 2$.*

Proof

(i) The element e_n is clearly an idempotent, since it is the sum of pairwise disjoint idempotents.

(ii) We use the embedding of θ_n of P_n into P_2 given in Theorem 8 of Chapter 1. The result holds trivially for $n = 2$. If we denote $\theta_n(p_i)$ by $p_{n,i}$, for clarity, then

$$\sum_{j=0}^{n} p_{n+1,j}^{-1} p_{n+1,j} = p^{-1}p + q^{-1}\left(\sum_{i=0}^{n-1} p_{n,i}^{-1} p_{n,i}\right) q.$$

It follows that if P_n is strongly embedded in R, then P_{n+1} is as well. Hence the result follows by induction. □

3.3.2 Contracting matrices

Let R be a unital ring in which P_2 is embedded. We construct morphisms of $\mathbf{Mat_R}$ that lead to functors between matrix categories, and then to isomorphisms between matrix rings of different orders. To define these functors, we first require the following:

Definitions 3.4

We define the following matrices for all $n \geq 2$:

$$G_n = \begin{pmatrix} I_{n-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & p^{-1} & q^{-1} \end{pmatrix}, \quad U_n = \begin{pmatrix} I_{n-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & p & q \end{pmatrix}$$

and, using the embedding $\theta_n : P_n \to P_2$ of Theorem 12, Chapter 1,

$$H_n = \begin{pmatrix} \theta_n(p_0^{-1}) & \theta_n(p_1^{-1}) & \ldots & \theta_n(p_{n-1}^{-1}) \end{pmatrix}$$

and

$$V_n = \begin{pmatrix} \theta_n(p_0) & \theta_n(p_1) & \ldots & \theta_n(p_{n-1}) \end{pmatrix}$$

Lemma 4 *For all $n \geq 2$, $H_n = \prod_{i=2}^{n} G_i$ and $V_n^t = \prod_{i=n}^{2} U_i^t$.*

Proof We prove this by induction. Note that, for $i \geq 2$,

$$H_i G_{i+1} = \begin{pmatrix} \theta_i(p_0^{-1}) & \theta_i(p_1^{-1}) & \ldots & \theta_i(p_{i-1}^{-1}) & \theta_i(p_{i-1}^{-1})p^{-1} & \theta_i(p_{i-1}^{-1})q^{-1} \end{pmatrix}$$

However, by definition of θ_n, from Theorem 12, Chapter 1,

$$\theta_{n+1}(p_i^{-1}) = \begin{cases} \theta_n(p_i^{-1}) & 0 \leq i \leq n-2 \\ \theta_n(p_{n-1}^{-1})p^{-1} & i = n-1 \\ \theta_n(p_{n-1}^{-1})q^{-1} & i = n \end{cases}.$$

Therefore,

$$H_i G_{i+1} = H_{i+1}.$$

A similar calculation gives that

$$U_{i+1} V_i = V_{i+1}.$$

Finally, as $G_2 = H_2$, and $U_2 = V_2$, our result follows by induction. □

Lemma 5 *Let e_2 be as defined in Lemma 3, and define $E_n = I_{n-1} \sqcup (e_2)$ for all $n \geq 1$. Then $U_n^t G_n = I_n$ and $G_n U_n^t = E_{n-1}$.*

Proof It is immediate that

$$U_n^t G_n = \begin{pmatrix} I_{n-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & pp^{-1} & pq^{-1} \\ \mathbf{0} & qp^{-1} & qq^{-1} \end{pmatrix} = \begin{pmatrix} I_{n-2} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{pmatrix} = I_n,$$

by definition of the composition in P_2. Similarly,

$$G_n U_n^t = \begin{pmatrix} I_{n-2} & \mathbf{0} \\ \mathbf{0} & p^{-1}p + q^{-1}q \end{pmatrix} = E_{n-1}.$$

Therefore our result follows. □

Lemma 6 *Let e_n be as defined in Lemma 3. Then $H_n V_n^t = e_n$ and $V_n^t H_n = I_n$.*

Proof By definition of composition in P_n, $[V_n^t H_n]_{i,j} = p_i p_j^{-1} = \delta_{ij}$. Therefore, $V_n^t H_n = I_n$. Similarly, $H_n V_n^t = \sum_{i=0}^{n-1} p_i^{-1} p_i = e_n$, by definition of e_n. Therefore, our result follows. □

Using these results, we are able to construct maps between matrix categories, and corners of matrix categories that reduce the orders of matrices.

Definitions 3.5

Let P_2 be embedded in a unital ring R. We define a map $C : \mathbf{Mat_R}^{\geq 2} \to E\mathbf{Mat_R}E$ (where E is the family of idempotents defined in Lemma 5) by $C(A) = G_j A U_i^t$ for all $A \in \mathbf{Mat_R}^{\geq 2}(i, j)$, and $C(x) = x - 1$ for all $x \geq 2$.

Theorem 7 *C is an injective functor that preserves addition.*

Proof Consider $A, A' \in \mathbf{Mat_R}^{\geq 2}(i, j)$ and $B \in \mathbf{Mat_R}^{\geq 2}(j, k)$. Then by definition of C,

$$C(B)C(A) = G_k B U_j^t G_j A U_i^t = G_k B I_j A U_i^t = C(BA),$$

by Lemma 5. Also,

$$C(A) + C(A') = G_j A U_i^t + G_j A' U_i^t = G_j(A + A')U_i^t = C(A + A'),$$

by the distributivity of composition over addition in a ring. Therefore C preserves composition and addition. Also, $C(I_n) = G_n I_n U_n^t = E_{n-1}$, which is the identity at $n-1$ in $E\mathbf{Mat_R}E$, and so C is a functor.

Now consider $A, A' \in \mathbf{Mat_R}^{\geq 2}(i, j)$ satisfying $C(A) = C(A')$. Then by definition of C, this implies that

$$G_j A U_i^t = G_j A' U_i^t.$$

We can multiply the both sides of the above equation by U_j^t on the left, and G_i on the right, so that

$$U_j^t G_j A U_i^t G_i = U_j^t G_j A' U_i^t G_i.$$

However, $U_x^t G_x = I_x$ for all $x \geq 2$, by Lemma 5, and so $I_j A I_i = I_j A' I_i$. Therefore, C is injective. □

Corollary 8 *The restriction of C to $M_n(R)$ gives an injective ring homomorphism from $M_n(R)$ to $E_{n-1}M_{n-1}(R)E_{n-1}$.*

Proof Immediate from Theorem 7 above. □

We consider the theory given by iterating the above contraction maps, to give maps between the category of proper matrices over R and corners of the ring R.

Theorem 9 *For all $n \geq 2$, there exists an injective homomorphism from $M_n(R)$ to a corner of R.*

Proof The map $C|_{M_2(R)} \circ C|_{M_3(R)} \circ \ldots \circ C|_{M_n(R)}$ gives an injective ring homomorphism, by iterating Corollary 8. □

Proposition 10 *Consider $A \in \mathbf{Mat_R}(n, n)$, where $n \geq 2$. Then $C^{n-1}(A) = H_n A V_n^t \in R$.*

Proof For $A \in \mathbf{Mat_R}(n, n)$,

$$C^n(A) = G_2 G_3 \ldots G_{n-1} G_n A U_n^t U_{n-1}^t U_{n-2}^t \ldots U_2^t.$$

However, by Lemma 4,

$$H_n = \prod_{i=2}^{n} G_i \ , \ V_n^t = \prod_{i=n}^{2} U_i^t.$$

Therefore, $C^n(A) = H_n A V_n^t \in e_n R e_n$. □

This motivates the following definition:

Definitions 3.6

We define a map Φ from the category $\mathbf{Mat_R}$ (excluding the empty matrix) to R, as follows:
On objects

$$\Phi(n) = 1 \ \forall n \in Ob(\mathbf{Mat_R}),$$

On morphisms

$$\Phi(X) = \begin{cases} H_j X V_i^t & \forall X \in \mathbf{Mat_R}(i,j),\ i,j \geq 2, \\ X V_i^t & \forall X \in \mathbf{Mat_R}(i,1),\ i \geq 2, \\ H_j X & \forall X \in \mathbf{Mat_R}(1,j),\ j \geq 2, \\ X & \forall X \in R \end{cases}$$

Theorem 11 *Φ is a map from $\mathbf{Mat_R}$ to R that preserves the addition and multiplication operations, and is functorial when the embedding of P_2 into R is strong.*

Proof Consider $A, A' \in \mathbf{Mat_R}(i,j)$ and $B \in \mathbf{Mat_R}(j,k)$, Then, assuming $i,j \geq 2$ (otherwise this result is trivial), $\Phi(B)\Phi(A) = H_k B V_j^t H_j A V_i^t$. However, by Lemma 5, $V_j^t H_j = I_j$, and so $\Phi(B)\Phi(A) = H_k A I_j B V_j^t = H_k A B V_i^t = \Phi(BA)$. Therefore Φ preserves composition. Also, by the distributivity of composition over addition in a ring,

$$\Phi(A) + \Phi(A') = H_j A V_i^t + H_j A' V_i^t = H_j(A + A')V_i^t = \Phi(A + A').$$

Therefore Φ preserves addition. Now assume that the embedding of P_2 into R is strong. In this case, $\Phi(I_n) = H_n V_n = e_n = 1$ for all $n \geq 1$, and so Φ is functorial. □

Note that, unlike the functor, $C : \mathbf{Mat}_{\mathbf{R}}^{\geq 2} \to \mathbf{Mat_R}$, Φ is never injective.

Counterexample

$$\Phi\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \Phi\begin{pmatrix} 1 & 0 \\ 0 & p^{-1}p + q^{-1}q \end{pmatrix},$$

by definition of the map Φ. However we do have the following weaker result:

Lemma 12 *For fixed $i, j \in \mathbb{N}$, the restriction of Φ to $\mathbf{Mat_R}(i,j)$ is injective.*

Proof We assume that $i, j \geq 2$, otherwise the result is trivial. Then consider $A, A' \in \mathbf{Mat_R}(i,j)$ satisfying $\Phi(A) = \Phi(A')$. By definition of Φ, this implies that $H_j A V_i^t = H_j A' V_i^t$, so we can multiply both sides of this by V_j^t on the left, and H_i on the right, so that $V_j^t H_j A V_i^t H_i = V_j^t H_j A' V_i^t H_i$. However, $V_x^t H_x = I_x$ for all $x \geq 2$, so $I_j A I_i = I_j A' I_i$. Therefore, the restriction of Φ to the individual *Hom* sets is injective. □

Definitions 3.7

We denote the map given by restricting Φ to $\mathbf{Mat_R}^{\geq 2}(n,n)$, and corestricting Φ to $e_n R e_n$ by $\Phi_n : M_n(R) \to e_n R e_n$.

Theorem 13 *Φ_n is an injective unitary ring homomorphism for all $n \geq 2$. When the embedding of P_2 is strong, Φ_n is an injective unitary ring homomorphism from $M_n(R)$ to R.*

Proof It is immediate by Theorem 11 that Φ_n preserves addition and multiplication, and by Lemma 12 that Φ_n is injective. Finally, $\Phi_n(I_n) = H_n V_n^t = e_n$, by Lemma 6, which is the identity of $e_n R e_n$. The final part follows by noting that in the strong case, $e_n = 1$, for all $n \in \mathbb{N}$, by Lemma 3. □

3.3.3 Expanding matrices

Definitions 3.8

Let P_2 be embedded in a unital ring R. We define a map $D : E\mathbf{Mat_R}E \to \mathbf{Mat_R}$ by $D(X) = U_{j+1}^t X G_{i+1}$ for all $X \in E\mathbf{Mat_R}E(i,j)$ (where U_y and G_x are as defined in Definitions 3.4), and for all $n \in \mathbb{N}$, $D(n) = n+1$.

Theorem 14 *D is an injective functor that preserves addition.*

Proof Consider $X, X' \in E\mathbf{Mat_R}E(i,j)$ and $Y \in E\mathbf{Mat_R}E(j,k)$. Then

$$D(Y)D(X) = U_{k+1}^t Y G_{j+1} U_{j+1}^t X G_{i+1} = U_{k+1}^t Y E_j X G_{i+1} = U_{k+1}^t Y X G_{i+1} = D(YX),$$

and so D preserves composition. Also, $D(E_i) = U_{i+1}^t G_{i+1} U_{i+1}^t G_{i+1} = I_{i+1} I_{i+1} = I_{i+1}$, the identity of $i+1$ in $\mathbf{Mat_R}$. Therefore, D is a functor.

Next, $D(X) = D(X')$ implies that $D(X) = U_{j+1}^t X G_{i+1} = U_{j+1}^t X' G_{i+1}$. We can multiply both sides of this by G_{j+1} on the left and U_{i+1}^t on the right to deduce $E_j X E_i = E_j X' E_i$, and so $X = X'$, since $X, X' \in E\mathbf{Mat_R}E(i,j)$. Therefore, D is injective. □

Iterating the D functor gives the following theory:

Definitions 3.9

We define a map $\Psi_n : e_n R e_n \to \mathbf{Mat_R}(n,n)$ by $\Psi_n(r) = V_n^t r H_n$.

Theorem 15 *Ψ_n is an injective unitary ring homomorphism from $e_n R e_n$ to $M_n(R)$, and an injective unitary ring homomorphism from R to $M_n(R)$ when the embedding of P_2 into R is strong.*

Proof For arbitrary $r, s \in R$,

$$\Psi_n(s)\Psi_n(r) = V_i^t s H_i V_i^t r H_i = V_i^t s e_n r H_i = \Psi_n(sr).$$

Therefore, Ψ preserves composition. Also,

$$\Psi_n(r) + \Psi_n(s) = V_i^t r H_i + V_i^t s H_i = V_i^t (r+s) H_i = \Psi_n(r+s).$$

Therefore Ψ_n also preserves addition. Therefore, Ψ is a ring homomorphism.

Now consider $r, r' \in e_n R e_n$ satisfying $\Psi_n(r) = \Psi_n(r')$. Then $V_n^t r H_n = V_n^t r' H_n$. We can multiply by H_n on the left, and by V_n^t on the right, to deduce $e_n r e_n = e_n r' e_n$. However, $r, r' \in e_n R e_n$, and so this implies that $r = r'$, and so Ψ_n is injective.

The final part follows from Lemma 3, which shows that, when the embedding of P_2 is strong, $e_n = 1$ for all n. □

Theorem 16 *C and D are mutually inverse category isomorphisms that preserve addition.*

Proof For all $n \in Ob(\mathbf{Mat}_{\mathbf{R}}^{\geq 2})$ and $m \in Ob(E\mathbf{Mat}_{\mathbf{R}}E)$, it is immediate from the definitions of the actions of C and D on objects that $CD(m) = m$ and $DC(n) = n$. Also, for all $X \in \mathbf{Mat}_{\mathbf{R}}{}^{\geq 2}(i, j)$ and $Y \in \mathbf{Mat}_{\mathbf{R}}(r, s)$,

$$CD(Y) = C(G_s Y U_r^t) = U_s^t G_s Y U_r^t G_r = I_s Y I_r = Y,$$

and

$$DC(X) = D(U_{j+1}^t X G_{i+1}) = G_{j+1} U_{j+1}^t U G_{i+1} U_{i+1}^t = E_j X E_i = X,$$

by Lemma 5. Finally, we have already seen (Theorems 7 and 13) that C and D are functors that preserve addition. Therefore our result follows. □

We then have the following as a Corollary:

Theorem 17 *For all $n \geq 2$, Φ_n and Ψ_n are mutually inverse ring isomorphisms.*

Proof Immediate from theorem 15 above, and the definition of Φ and Ψ in terms of the iterations of the C and D functors respectively. □

In the strong case, the above results give the following Corollary:

Theorem 18 *When the embedding of P_2 into R is strong, Φ_n and Ψ_n are mutually inverse isomorphisms between $M_n(R)$ and R for all $n \in \mathbb{N}$.*

Proof Immediate from Theorem 17 above, and Lemma 3. □

3.4 Self-embeddings of rings, and monoidal structures

Let P_2 be embedded in a unital ring R. We consider what happens to the monoidal structure of $\mathbf{Mat_R}$ when we map $\mathbf{Mat_R}$ into R using the Φ map (which we have proved is a functor when the embedding of P_2 into R is strong). We will demonstrate that this gives the ring R the structure of a one-object symmetric monoidal category (apart from the units elements) when the embedding is strong, and a similar, but weaker, structure when the embedding of P_2 is not strong.

Definitions 3.10

We define a map $*$ from $\mathbf{Mat_R} \times \mathbf{Mat_R}$ (without the empty matrix) to R, as follows:

- On objects, $a * b = 1$, for all $a, b \in \mathbb{N}$,
- On morphisms, $U * V = \Phi(U \sqcup V)$.

Proposition 19

(i) *$*$ preserves addition and composition,*

(ii) *For all $a, b \in \mathbb{N}$, there exists s_{ab} satisfying $s_{jl}(U * V)s_{ki} = V * U$ for all $U \in \mathbf{Mat_R}(i,j)$ and $V \in \mathbf{Mat_R}(k,l)$.*

(iii) *For all $i, j, k \in \mathbb{N}$, there exists $t_{ijk}, t'_{ijk} \in R$ satisfying*

$$t_{x,y,z}(U * (V * W))t'_{i,j,k} = ((U * V) * W),$$

for all $U \in \mathbf{Mat_R}(i,x)$, $V \in \mathbf{Mat_R}(j,y)$, and $W \in \mathbf{Mat_R}(k,z)$.

Proof

(i) We have proved that $\sqcup$ is an addition-preserving functor in Theorem 2, and have proved that Φ preserves addition and composition in Theorem 11. Therefore, it is immediate by the definition of $*$ that it preserves addition and composition.

(ii) For all $a, b \geq 1$, define $s_{ab} = \Phi(S_{ab})$, where S_{ab} is as defined in Theorem 2. Then, as Φ preserves composition, and $S_{jl}(U \sqcup V)S_{ki} = V \sqcup U$, we can deduce that $\Phi(S_{jl}(U \sqcup V)S_{ki}) = \Phi(V \sqcup U)$, and so $s_{jl}(U * V)s_{ik} = V * U$.

(iii) For all $a, b, c \in \mathbb{N}$, define $t'_{a,b,c} = (V^t_{a+b} * I_c)(I_a * H_{b+c})$ and $t_{a,b,c} = (H_{a+b} * I_c)(I_a * V^t_{b+c})$. Then note that $(I_x \sqcup V^t_{y+z})(U \sqcup (V*W))(I_i \sqcup H_{j+k}) = U \sqcup V \sqcup W$. Also, $(H_{x+y} \sqcup I_z)(U \sqcup V \sqcup W)(V^t_{i+j} \sqcup I_k) = (U * V) \sqcup W$. Therefore,

$$(H_{x+y} \sqcup I_z)(I_x \sqcup V^t_{y+z})(U \sqcup (V * W))(I_i \sqcup H_{j+k})(V^t_{i+j} \sqcup I_k) = (U * V) \sqcup W.$$

Finally, recall that, by definition, $A * B = \Phi(A \sqcup B)$, and Φ preserves composition and addition, so we can apply Φ to the above equation, and deduce that

$$(H_{x+y} * I_z)(I_x * V^t_{y+z})(U * (V * W))(I_i * H_{j+k})(V^t_{i+j} * I_k) = (U * V) * W.$$

And this is $t_{x,y,z}(U * (V * W))t'_{i,j,k} = (U * V) * W$. □

3.4.1 The symmetric monoidal structure of a ring

We will demonstrate how a restriction of the $*$ map gives the ring R a similar but weaker structure to a one-object symmetric monoidal category (without unit elements), and the structure of a one-object symmetric monoidal category (again, without the unit elements) when the embedding of P_2 into R is strong. Note that, although the category of matrices over a ring R is strictly monoidal, the structure induced by the restriction of $*$ is not strict; it is 'monoidal up to isomorphism'.

Definitions 3.11

Let P_2 be embedded in a ring R. We define $\oplus$ to be the restriction of the map $* : \mathbf{Mat_R} \times \mathbf{Mat_R} \to R$ to members of R, considered as (1×1) matrices. So, for all $a, b \in R$,

$$a \oplus b = \Phi(a \sqcup b) = \Phi_{2,2}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = p^{-1}ap + q^{-1}bq.$$

Proposition 20

(i) *$\oplus$ is a ring homomorphism.*

(ii) *There exists elements s, t, t' of R satisfying, for all $f, g, h \in R$,*

1. $t'(f \oplus (g \oplus h)) = ((f \oplus g) \oplus h)t'$,

2. $t^2 = (t \oplus 1)t(1 \oplus t)$,

3. $s(f \oplus g) = (g \oplus f)s$,

4. $tst = (s \oplus 1)t(1 \oplus s)$.

Proof

(i) This is immediate from the fact that Φ preserves composition and addition (Theorem 11), and the fact that $\sqcup$ is an addition-preserving functor (Theorem 2).

(ii) We make the definitions $s = s_{1,1}$, $t = t_{1,1,1}$, and $t' = t'_{1,1,1}$, as defined in Proposition 9. These are given explicitly, as follows:

By definition,

$$s = \Phi\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = p^{-1}q + q^{-1}p.$$

Similarly,

$$t' = (V_2^t * I_1)(I_1 * H_2) = \Phi\begin{pmatrix} 1 & 0 & 0 \\ 0 & p^{-1} & q^{-1} \end{pmatrix}\Phi\begin{pmatrix} p & 0 \\ q & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \Phi\begin{pmatrix} p & 0 \\ p^{-1}q & q^{-1} \end{pmatrix} = p^{-1}p^2 + q^{-2}q + q^{-1}p^{-1}qp.$$

Also

$$t = (H_2 * I_1)(I_1 * V_2^t) = \Phi\begin{pmatrix} p^{-1} & q^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}\Phi\begin{pmatrix} 1 & 0 \\ 0 & p \\ 0 & q \end{pmatrix}$$

$$= \Phi\begin{pmatrix} p^{-1} & q^{-1}p \\ 0 & q \end{pmatrix} = p^{-2}p + p^{-1}q^{-1}pq + q^{-1}q^2.$$

Then our results are as follows:

1. This follows immediately, by the definition of t and t', and from Proposition 19 above.

2. We can expand t^2 as

$$t^2 = p^{-3}p + p^{-2}q^{-1}pq + p^{-1}q^{-1}pq^2 + q^{-1}q^3,$$

and, using the definition $(a \oplus b) = p^{-1}ap + q^{-1}bq$, we can expand $(t \oplus 1)t(1 \oplus t)$ as

$$(p^{-1}tp + q^{-1}q)t(p^{-1}p + q^{-1}tq) =$$

$$p^{-3}p + p^{-2}q^{-1}pq + p^{-1}q^{-1}pq^2 + q^{-1}q^3.$$

Therefore, $t^2 = (t \oplus 1)t(1 \oplus t)$. Hence our result follows.

3. By definition of $\oplus$,

$$s(f \oplus g) = (q^{-1}p + p^{-1}q)(p^{-1}fp + q^{-1}gq) = qfp^{-1} + pgq^{-1} = (g \oplus f)s.$$

4. We can expand tst and $(s\oplus 1)t(1\oplus s)$ in terms of p, q, so that

$$tst = p^{-1}q^2 + p^{-1}q^{-1}p + q^{-1}pq.$$

and

$$(s\oplus 1)t(1\oplus s) = (p^{-1}sp + q^{-1}q)t(p^{-1}p + q^{-1}sq) = p^{-1}q^2 + p^{-1}q^{-1}p + q^{-1}pq.$$

Therefore, $tst = (s\oplus 1)t(1\oplus s)$.

Hence our results follow. □

This allows us to deduce the following:

Theorem 21 *When the embedding of P_2 into R is strong, $(R,\oplus)$ has the structure of a (one-object) symmetric monoidal category, apart from the unit elements.*

Proof First note that $1\oplus 1 = p^{-1}p + q^{-1}q = e_2 = 1$ when the embedding of P_2 is strong, by the definition of a strong embedding. Therefore, $\oplus$ is a functor from $(R,\cdot)\times(R,\cdot)$ to $(R,\cdot)$ (considered as a one-object category).

Next, by definition,

$$s^2 = \Phi\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\Phi\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \Phi\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1,$$

by Theorem 18. Also,

$$t't = \Phi\begin{pmatrix} p & 0 \\ p^{-1}q & q^{-1} \end{pmatrix}\Phi\begin{pmatrix} p^{-1} & q^{-1}p \\ 0 & q \end{pmatrix} = \Phi\begin{pmatrix} 1 & 0 \\ 0 & q^{-1}q + p^{-1}p \end{pmatrix} = 1,$$

by Theorem 18. Therefore, we can write $t' = t^{-1}$ and $s = s^{-1}$. Then, by Proposition 20 above, $(R,\oplus)$ has the structure of a one-object symmetric monoidal category, apart from the unit elements. □

We will study the theory of one-object symmetric monoidal categories that do not have units elements in more detail in Chapter 4. However, we first give an application of the above results to the K_0 theory of rings.

3.5 Constructing the K_0 group of a ring using P_2

Let R be a unital ring in which P_2 is embedded (this embedding need not be strong); we will construct a semigroup from the idempotents of R. We will then show that this semigroup is isomorphic to the semigroup of (equivalence classes of) idempotents of all matrix rings over R, used in the construction of the K_0 group of R. This allows us to construct the K_0 group of R solely from the idempotents of R, rather than from the idempotents of $\mathbf{Mat_R}$.

3.5.1 The classical construction of the K_0 group of a ring

We recall standard definitions from the K_0 theory of unital rings, taken from [24], p.135-145.

Definitions 3.12

The set of idempotents of all rings of matrices over R is denoted $E(R)$, and is defined by

$$E(R) = \bigcup_{i=1}^{\infty} \{A \in M_i(R) : A^2 = A\}.$$

Note that, by the definition of an idempotent, any $A \in E(R)$ must be a square matrix. An equivalence relation $\sim$ is then defined on $E(R)$ by

$$E \sim F \Leftrightarrow \exists X, Y : XY = E \ , \ YX = F \ , \ EXF = X \ , \ FYE = Y.$$

There is an associative composition $\sqcup$ defined on $E(R)$ as follows: Given an $m \times m$ idempotent matrix e, and an $n \times n$ idempotent matrix f, over R, then $e \sqcup f$ is defined to be the $(m+n) \times (m+n)$ matrix

$$e \sqcup f = \begin{pmatrix} e & \mathbf{0} \\ \mathbf{0} & f \end{pmatrix},$$

and the relation $\sim$ is clearly a congruence with respect to this binary operation. (This operation is the restriction to idempotents of the strict monoidal functor for $\mathbf{Mat_R}$, the category of matrices over R). A binary operation is defined on the $\sim$ equivalence classes of $E(R)$ by $[e] + [f] = [e \sqcup f]$. Then $(E(R)/\sim, +)$ is clearly an abelian semigroup. We will refer to this as the K_0 *semigroup of* R, although this is not standard terminology. Finally, the group $K_0(R)$ is defined to be the abelian group given by constructing the *Grothendieck group*, or *universal group* of the commutative semigroup $(E(R)/\sim, +)$, as described in [24], p.295-297. This group has the universal property that any homomorphism from the abelian semigroup $(E(R)/\sim, +)$ into an arbitrary group G factors through $K_0(R)$.

We consider how this construction can be simplified under the assumption that P_2 is embedded in R.

3.5.2 Simplifying the K_0 construction using P_2

Let P_2 be embedded in a ring R (we identify P_2 with its image under this embedding). We consider the $\sim$ relation on the idempotents of $\mathbf{Mat_R}$, along with the $\oplus$ operation of Definitions 3.11. For our results, we first require the following:

Lemma 22 *Let $E^2 = E \in \mathbf{Mat_R}(n,n)$ be an arbitrary idempotent. Then $\Phi_n(E) \sim E$.*

Proof Denote $\Phi_n(E)$ by e. Then $e = H_nEV_n^t$ by definition of Φ_n. If we define $X = H_nE$ and $Y = EV_n^t$, then $XY = H_nE^2V_n^t = H_nEV_n^t$, by definition. Also, $YX = EV_n^tH_nE = EI_nE = E^2 = E$. Finally,

$$eXE = H_nEV_n^tH_nEE = H_nEI_nE^2 = H_nE^3 = H_nE = X,$$

and

$$EYe = EEV_n^tH_nEV_n^t = E^2I_nEV_n^t = E^3V_n^t = EV_n^t = Y.$$

Therefore $\Phi_n(E) \sim E$, by definition of the relation $\sim$. □

Definitions 3.13

We define $e(R)$ to be the set of idempotents of R, and consider the restriction of the above $\sim$ relation to $e(R)$.

Proposition 23 *$(e(R)/\sim, \oplus)$ is an abelian semigroup isomorphic to the K_0 semigroup of R.*

Proof For all $A \in E(R)$, define $a = \Phi(E)$. Then, by Lemma 22, $a \sim A$, and so $[a]_\sim = [A]_\sim$. Therefore, every idempotent of $E(R)/\sim$ has a representative in R. Next note that for all $E \sim F \in E(R)$, by definition there exists $X, Y \in \mathbf{Mat_R}$ satisfying $XY = E$, $YX = F$, $EXF = X$, $FYE = E$. Therefore, if we define $x = \Phi(X)$, and $y = \Phi(Y)$, then $\Phi(E) \sim \Phi(F)$, as Φ preserves composition and addition.

Hence the map $[A] \mapsto [\Phi(A)]$ from $E(R)/\sim$ to $e(R)/\sim$ is well-defined, and is a bijection. Finally, we show that it preserves composition. Consider $E^2 = E, F^2 = F \in E(R)$. Then $(\Phi(E) \oplus \Phi(F)) = \Phi(\Phi(E) \sqcup \Phi(F))$, and so $(\Phi(E) \oplus \Phi(F)) \sim \Phi(E) \sqcup \Phi(F)$. Also, $\Phi(E) \sim E$ and $\Phi(F) \sim F$. Therefore, $(\Phi(E) \sqcup \Phi(F)) \sim (E \sqcup F)$ as $\sim$ is a congruence on the abelian semigroup $(E(R), \sqcup)$. Next, by Lemma 22, $(E \sqcup F) \sim \Phi(E \sqcup F)$. Hence, as $\sim$ is an equivalence relation, $(\Phi(E) \oplus \Phi(F)) \sim \Phi(E \sqcup F)$ in $\mathbf{Mat_R}$, and so $\Phi(E \sqcup F) \sim \Phi(E) \oplus \Phi(F)$ in R.

Therefore, the map from $E(R)/\sim$ to $e(R)/\sim$ given by $[A] \mapsto [\Phi(A)]$ is a well-defined bijection, and maps $+$ to $\oplus$. Hence it is a semigroup isomorphism, and so our result follows. □

We can deduce the following as a corollary:

Theorem 24 *The K_0 group of R can be constructed by applying Grothendieck's construction to the abelian semigroup of Proposition 23 above.*

Proof Immediate from the isomorphism between $(\{e^2 = e \in R\}/\sim, \oplus)$ and the K_0 semigroup of R. $\square$

Hence we have shown that constructing the Grothendieck group of $(e(R)/\sim, \oplus)$, as described in [48] p.295-297, will give us the K_0 group of the ring R, and so an embedding of P_2 in a unital ring R allows us to greatly simplify the calculation of the K_0 group of R.

Chapter 4

Categorical self-similarity and internalising monoidal structures

4.1 Introduction

We generalise the results of Chapters 2 and 3 to give a categorical interpretation of self-similarity, in terms of self-similar objects of symmetric monoidal categories. This can be thought of as the non-associative case of the ring theory results of Chapter 3 — in particular, the construction of a one-object symmetric monoidal category (without unit elements). This motivates the definition of M–monoids, which have a similar, but weaker, structure to one-object symmetric monoidal categories (without unit elements). We then show how this implies (under certain conditions) the existence of embeddings of the polycyclic monoid on two generators into endomorphism monoids of self-similar objects of symmetric monoidal categories.

We also introduce several concepts which will be useful in later chapters, where we apply our categorical structures to the study of algebraic models of logic. In particular, we introduce the internal tensor product, and the concept of a fixed point for an internal tensor product — examples of these will be constructed in Chapter 5.

Finally, as an application of the above theory, we demonstrate how to construct 1-object analogues of the 'internal hom.' of a Cartesian closed category at the endomorphism monoid of a self-similar object.

4.1.1 One-object symmetric monoidal categories

Recall the definitions of symmetric monoidal categories from Definition 1.2, of Chapter 1. For much of the remainder of this thesis, we will consider various symmetric monoidal categories without unit objects. This is because of the following:

Proposition 1 *Let* $(\mathbf{M}, \otimes, t, s, I, \lambda, \rho)$ *be a symmetric monoidal category with one object. Then* M *is a set with two (strictly) commutative associative operations. However, the same does not apply to a symmetric monoidal category without units elements.*

Proof Let $(\mathbf{M}, \otimes, t, s, I, \lambda, \rho)$ be a symmetric monoidal category. Then, for all $f : A \to A$, $g : B \to B$ by naturality of the units maps, $\lambda_B^{-1} g = (1_I \otimes g)\lambda_B^{-1}$ and $\rho_A^{-1} f = (f \otimes 1_I)\rho^{-1}$. Therefore, $\rho_A^{-1} f \rho_A = (f \otimes 1_I)$ and $\lambda_B^{-1} g \lambda_B = (1_I \otimes g)$. However, as we are considering the one-object case, $A = B = I$, and from Definitions 1.2, Chapter 1, $\lambda_I = \rho_I$. Therefore $(f \otimes 1_I) = \rho_I^{-1} f \rho_I = \lambda_I^{-1} f \lambda_I = (1_I \otimes f)$, and so $\otimes$ is commutative.

Similarly, the units triangle coherence condition gives $(\rho_I \otimes 1_I)t_{I,I,I} = (1_I \otimes \lambda_I)$. However, $\lambda_I = \rho_I$ and we have proved that $\otimes$ is strictly symmetric, so $(1_I \otimes \lambda_I)t_{I,I,I} = (1_I \otimes \lambda_I)$ and it is immediate from this that t is the identity. Finally, from [41], p.161, a strict symmetric monoidal category with one object is an abelian monoid with respect to two commutative operations that satisfy the interchange law.

The second part of our proof follows by example, from Theorem 21 of Chapter 3. □

Convention In view of the above result, we make the convention that when we refer to a *one-object symmetric monoidal category*, we assume that it need not have unit elements.

4.2 Self-similarity in monoidal categories

We axiomatise our concepts of self-similarity, using the theory of symmetric monoidal categories[1] We do this by defining what we mean by a self-similar object of a symmetric monoidal category, and showing that this gives the endomorphism monoid of the self-similar object a similar structure to a one-object symmetric monoidal category.

4.2.1 Self-similar objects

Definitions 4.1

Let $(M, \otimes)$ be a symmetric monoidal category. We say that $N \in Ob(\mathbf{M})$ is a *self-similar object* if there exist morphisms $c \in \mathbf{M}(N \otimes N, N)$ and $d \in \mathbf{M}(N, N \otimes N)$ that satisfy $dc = 1_{N \otimes N}$. We

[1] In fact, the constructions of this chapter appear to work in the non-symmetric case. However, all the examples we construct will be symmetric, so we restrict our definitions accordingly.

call these morphisms the *contraction* and *division* morphisms of N. We say that N is *strongly self-similar*, or is a *strong self-similar object* if it satisfies the further condition $cd = 1_N$, so that $N \cong N \otimes N$. If N is not strongly self-similar, we say that it is *weakly self-similar*, or just *weak*, when the context is clear. If all the objects of $\mathbf{M}$ are self-similar, we say that $\mathbf{M}$ is a *self-similar category*. It is immediate that if N is a (strongly) self-similar object of a subcategory $\mathbf{S}$ of a category $\mathbf{C}$, then N is a (strongly) self-similar object of $\mathbf{C}$.

4.2.2 Free monoidal categories

In order to study the properties of self-similar objects of categories, we consider the 'simplest possible' symmetric monoidal category, and self-similarity. We also consider the properties when this self-similarity is strong.

Definitions 4.2

In [41], p.161, the freely generated symmetric monoidal category on one object, $(\mathbf{W}, \otimes)$ is defined. This has objects given by binary bracketings of the symbol $-$ (including the empty bracketing, $()$; this is the unit object), and unique natural canonical morphisms, as given in Definitions 1.2, Chapter 1. The same reference also defines, for any object B of a symmetric monoidal category $(\mathbf{M}, \otimes)$, a unique functor of monoidal categories, $(\ )_B \mathbf{W} \to \mathbf{M}$. On objects, this is given by

- $(\ -)_B = B$,
- $(u \otimes v)_B = (u)_B \otimes (v)_B$.

and we refer to [41] for its definition on morphisms.

We define the *self-similar category* $(\mathbf{S}, \otimes)$ to be the category $\mathbf{W}$, together with the assumption that the unique object (which we denote N for clarity) is self-similar. So, the morphisms of S are specified by

- $c \in \mathbf{S}(N \otimes N, N)$, $d \in \mathbf{S}(N, N \otimes N)$, so that $dc = 1_{N \otimes N}$.
- For all $f, g \in Arr(\mathbf{S})$, $f \otimes g \in Arr(\mathbf{S})$.

We call $\mathbf{S}$ the *strongly self-similar category* when N is a strongly self-similar object.

Proposition 2 *$(\mathbf{S}, \otimes)$ is a regular category.*

Proof Note that all the canonical isomorphisms have global inverses. Also, $cdc = c1_{N \otimes N} = c$ and $dcd = 1_{N \otimes N} d = d$, so c and d are mutual generalised inverses. Next, let a, b be elements of $\mathbf{S}$ that have generalised inverses a^{-1} and b^{-1}. Then $(a\otimes b)(a^{-1}\otimes b^{-1})(a\otimes b) = (aa^{-1}a\otimes bb^{-1}b) = (a\otimes b)$ and similarly, $(a^{-1}\otimes b^{-1})(a\otimes b)(a^{-1}\otimes b^{-1}) = (a^{-1}aa^{-1}\otimes b^{-1}bb^{-1}) = (a^{-1}\otimes b^{-1})$. Therefore, when a and b have generalised inverses, so does $(a\otimes b)$. Finally, if ba is defined, then so is $a^{-1}b^{-1}$, and it is immediate from the definitions of the morphisms of **S** that $a^{-1}b^{-1}$ is the generalised inverse of ba. Therefore, **S** is a regular category. □

Corollary 3 *When* **S** *is the strong self-similar category,* **S** *is a groupoid.*

Proof When N is strongly self-similar, then c and d satisfy $cd = 1_N$ and $dc = 1_{N\otimes N}$, and so, by induction, all the elements of **S** have global inverses. Hence our result follows. □

4.2.3 Tensor categories of self-similar objects

Definitions 4.3

Let N be a self-similar object of the monoidal category $(\mathbf{M},\otimes)$. We define the *tensor category* of N, which we denote $(\otimes\mathbf{N})$, to be the full subcategory of **M**, whose objects are given inductively by

- $N \in Ob(\otimes\mathbf{N})$,
- $X\otimes Y \in Ob(\otimes\mathbf{N})$ for all $X, Y \in Ob(\otimes\mathbf{N})$.

Note that if N is a strongly self-similar object, then all endomorphism monoids of objects of $(\otimes\mathbf{N})$ are isomorphic.

An important special case will be when the map $(\ )_N : \mathbf{W} \to \mathbf{M}$ is injective on objects. In this case, each object corresponds to a unique bracketing of copies of N — or equivalently, to a unique binary tree with leaves labelled by N and nodes labelled by $\otimes$ — as in the definition of **W**. When $(\otimes\mathbf{N})$ satisfies this condition, we say that it *has freely generated objects.* In this case, many of our results on self-similarity will be greatly simplified. For example, let N be a self-similar object of a symmetric monoidal category $(\mathbf{M},\otimes)$, and let $(\otimes\mathbf{N})$ be freely generated objects in **M**. For all $X \in Ob(\otimes\mathbf{N})$, we (recursively) define two families of morphisms, the *division and contraction morphisms*, $\{d_X \in \mathbf{M}(N,X)\}$, and $\{c_X \in \mathbf{M}(X,N)\}$, as follows:

- $d_N = 1_N = c_N$.
- $d_{U\otimes V} = (d_U\otimes d_V)d$,
- $c_{U\otimes V} = c(c_U\otimes c_V)$.

The fact that these morphisms are well-defined follows directly from the fact that $\otimes N$ has freely generated objects. We can then use these morphisms to define the *contraction map* ϕ from $Arr(\otimes \mathrm{N})$ to $\mathbf{M}(N, N)$ by $\phi(f) = c_Y f d_X$ for all $f \in (\otimes \mathrm{N})(X, Y)$.

Lemma 4 *Let N be a self-similar object of a symmetric monoidal category,* $(\mathbf{M}, \otimes)$*, and let* $(\otimes \mathrm{N})$ *have freely generated objects. Then, for all $X \in Ob(\otimes \mathrm{N})$, $d_X c_X = 1_X$. Also, when N is a strongly self-similar object, $c_X d_X = 1_N$.*

Proof First note that this result holds for $X = N$, by definition of d and c. Next, let U and V be objects of $(\otimes \mathrm{N})$ satisfying $d_U c_U = 1_U$ and $d_V c_V = 1_V$. Then

$$d_{U \otimes V} c_{U \otimes V} = (d_U \otimes d_V) dc (c_U \otimes c_V) = (d_U \otimes d_V) 1_{N \otimes N} (c_U \otimes c_V)$$

$$= (d_U c_U \otimes d_V c_V) = (1_U \otimes 1_V) = 1_{U \otimes V}.$$

Therefore, our result follows by induction. The second part of our proof follows by the definition of a strongly self-similar object, and a similar induction argument. □

Theorem 5 *Let N be a self-similar object of a symmetric monoidal category,* $(\mathbf{M}, \otimes)$*, and let* $(\otimes \mathrm{N})$ *have freely generated objects in* $\mathbf{M}$*. Then $\phi(g)\phi(f) = \phi(gf)$ for all $f \in (\otimes \mathrm{N})(X, Y)$, and $g \in (\otimes \mathrm{N})(Y, Z)$.*

Proof By definition, $\phi(g)\phi(f) = c_Z g d_Y c_Y f d_X$. However, by Lemma 4 above, $d_Y c_Y = 1_Y$, and so $\phi(g)\phi(f) = c_Z g f d_X = \phi(gf)$. □

Corollary 6 *When N is a strong self-similar object, ϕ defines a functor from* $(\otimes \mathrm{N})$ *to the one-object category* $\mathbf{M}(N, N)$.

Proof Define ϕ on $Ob(\otimes \mathrm{N})$ by $\phi(X) = N$ for all X. Our result is then immediate, since by the definition of a strongly self-similar object, and Theorem 5, $\phi(1_X) = 1_N$ for all X, and $\phi(g)\phi(f) = \phi(gf)$ whenever gf is defined. □

In the case where $(\otimes \mathrm{N})$ does not have freely generated objects, the definitions of c_X and d_X do not neccesarily uniquely specify elements of $(\otimes \mathrm{N})$, so we cannot define ϕ as above. For example, when $(\mathbf{M}, \otimes)$ is a strict monoidal category, then $N \otimes (N \otimes N) = (N \otimes N) \otimes N$, and so the elements $c(1_N \otimes c)$, and $c(c \otimes 1_N)$ both satisfy the definition of $c_{N \otimes N \otimes N}$.

We can however, unambiguously define a restriction of ϕ, as follows:

Definitions 4.4

Let N be a self-similar object of a symmetric monoidal category $\mathbf{M}$. We define $(\otimes_2 N)$ to be the full subcategory of M with two objects, N and $N \otimes N$. We can then define a map $\phi_2 : (\otimes_2 N) \to N$ explicitly, by

$$\phi_2(f) = \begin{cases} f & f \in \mathbf{M}(N, N) \\ cf & f \in \mathbf{M}(N, N \otimes N) \\ fd & f \in \mathbf{M}(N \otimes N, N) \\ cfd & f \in \mathbf{M}(N \otimes N, N \otimes N) \end{cases}$$

It is immediate that, when $(\otimes \mathrm{N})$ has freely generated objects, the map ϕ_2 is a restriction of ϕ to $(\otimes_2 N)$. Also, as $dc = 1_{N \otimes N}$, an identical argument to Theorem 5 gives that ϕ_2 preserves the composition of $(\otimes_2 N)$, and is a functor when N is strongly self-similar.

4.2.4 The internal tensor product

We demonstrate how endomorphism monoids of self-similar objects have a similar structure to one-object symmetric monoidal categories, and are one-object symmetric monoidal categories without units elements when the self-similarity is strong. However, we first require the following:

Lemma 7 *Let N be a self-similar object of a symmetric monoidal category* $(\mathbf{M}, \otimes, s, t)$. *Then for all* $X, Y, Z \in (\otimes_2 N)$

(i) $t_{X,Y,Z} = ((d_X \otimes d_Y) \otimes d_Z) t_{N,N,N} (c_X \otimes (c_Y \otimes c_Z))$,

(ii) $s_{X,Y} = (d_y \otimes d_X) s_{N,N} (c_X \otimes c_Y)$,

(iii) *The above two results hold for all* $X, Y, Z \in (\otimes \mathrm{N})$ *when* $(\otimes \mathrm{N})$ *has freely generated objects.*

Proof By definition, $1_A = d_A c_A$ for all $A \in Ob(\otimes_2 N)$. Then

(i) By definition, $t_{X,Y,Z} = t_{X,Y,Z}(1_X \otimes (1_Y \otimes 1_Z))$. Therefore

$$t_{X,Y,Z} = t_{X,Y,Z}(d_X c_X \otimes (d_Y c_Y \otimes d_Z c_Z)),$$

and so $t_{X,Y,Z} = ((d_X \otimes d_Y) \otimes d_Z) t_{N,N,N} (c_X \otimes (c_Y \otimes c_Z))$ by the naturality of $t_{X,Y,Z}$ in X, Y, Z.

(ii) By definition, $s_{X,Y} = s_{X,Y}(1_X \otimes 1_Y)$, and so $s_{X,Y} = s_{X,Y}(d_X c_X \otimes d_Y c_Y)$. Therefore, by the naturality of $s_{X,Y}$ in X and Y, $s_{X,Y} = (d_Y \otimes c_X) s_{N,N} (c_X \otimes d_Y)$.

(iii) This follows from the fact that $1_A = d_A c_A$, for all $A \in Ob(\otimes \mathrm{N})$ when $(\otimes \mathrm{N})$ has freely generated objects. Then the above proofs suffice. □

Definitions 4.5

Let $(\mathbf{M}, \otimes)$ be a symmetric monoidal category, and let N be a self-similar object of $\mathbf{M}$. We define the *internal tensor product at* N to be a map $\oplus : \mathbf{M}(N,N) \times \mathbf{M}(N,N) \to \mathbf{M}(N,N)$ given by $x \oplus y = c(x \otimes y)d$.

Proposition 8 *Let* N *be a self-similar object of a symmetric monoidal category* $(\mathbf{M}, \otimes)$, *and let the internal tensor product,* $\oplus$, *be as defined above. Then*

(i) $\oplus$ *is a semigroup homomorphism, and is a monoid homomorphism when* N *is strongly self-similar.*

(ii) ϕ_2 *maps* $\otimes$ *to the internal tensor* $\oplus$*; that is,* $\phi_2(f \otimes g) = \phi_2(f) \oplus \phi_2(g)$ *for all* $f, g \in \mathbf{M}(N,N)$.

(iii) *When* $(\otimes\mathrm{N})$ *has freely generated objects,* ϕ *maps* $\otimes$ *to the internal tensor* $\oplus$*; that is,* $\phi(f \otimes g) = \phi(f) \oplus \phi(g)$ *for all* $f \in (\otimes\mathrm{N})(U,X)$, $g \in (\otimes\mathrm{N})(V,Y)$.

Proof

(i) Consider arbitrary $a, b, x, y \in \mathbf{M}(N,N)$. Then

$$(a \oplus b)(x \oplus y) = c(a \otimes b)dc(x \otimes y)d = c(a \otimes b)1_{N \otimes N}(x \otimes y)d = c(ax \otimes by)d = ax \oplus by.$$

Therefore $\oplus$ is a semigroup homomorphism. Also, when N is strongly self-similar, $1 \oplus 1 = c(1 \otimes 1)d = cd = 1_N$ and so $\oplus$ is a monoid homomorphism.

(ii) By definition, $\phi_2(f \otimes g) = c(f \otimes g)d$ and $\phi_2(f) = f$, $\phi_2(g) = g$ for all $f, g \in \mathbf{M}(N,N)$. Therefore our result follows.

(iii) Assume $(\otimes\mathrm{N})$ has freely generated objects. Then by definition,

$$\phi(f \otimes g) = c_{X \otimes Y}(f \otimes g)d_{U \otimes V} = c(c_X \otimes c_Y)(f \otimes g)(d_U \otimes d_V)d.$$

However

$$\phi(f) \oplus \phi(g) = \phi(\phi(f) \otimes \phi(g)) = \phi(c_X f d_U \otimes c_Y g d_V) = c(c_X f d_U \otimes c_Y g d_V)d$$

$$= c(c_X \otimes c_Y)(f \otimes g)(d_U \otimes d_V)d,$$

by definition of $\oplus$. Therefore, $\phi(f \otimes g) = \phi(f) \oplus \phi(g)$. □

Theorem 9 *Let* N *be a self-similar object of a symmetric monoidal category* $(\mathbf{M}, \otimes, s_{X,Y}, t_{X,Y,Z})$. *Then there exists* $s, t \in \mathbf{M}(N,N)$ *satisfying*

(i) $s(u \oplus v) = (v \oplus u)s$,

(ii) $t(u \oplus (v \oplus w)) = ((u \oplus v) \oplus w)t$,

(iii) $t^2 = (t \oplus 1)t(1 \oplus t)$,

(iv) $tst = (s \oplus 1)t(1 \oplus s)$,

(v) *s and t have generalised inverses s^{-1}, t^{-1} that satisfy*

$$s = s^{-1} \ , \ s^2 = (1 \oplus 1) \ , \ tt^{-1} = (1 \oplus 1) \oplus 1 \ , \ t^{-1}t = 1 \oplus (1 \oplus 1),$$

(vi) *when N is strongly self-similar, $(\mathbf{M}(N,N), \oplus)$ is a one-object symmetric monoidal category without units elements.*

Proof Define $s = \phi_2(s_{N,N})$, where $s_{X,Y}$ is the family of commutativity morphisms for $(M, \otimes)$. Similarly, define $t = c(c \otimes 1)t_{N,N,N}(1 \otimes d)d$. (Note that, in the freely generated objects case, $t = c_{(N \otimes N) \otimes N} t_{N,N,N} d_{N \otimes (N \otimes N)} = \phi(t_{N,N,N}))$. Then

(i) This follows immediately, as ϕ_2 is a semigroup homomorphism.

(ii) Note that by definition of $\oplus$, $(u \oplus (v \oplus w)) = c(u \otimes c(v \otimes w)d)d$ and $((u \oplus v) \oplus w) = c(c(u \otimes v)d \otimes w)d$. Therefore,

$$t(u \oplus (v \oplus w)) = c(c \otimes 1)t_{N,N,N}(1 \otimes d)dc(u \otimes c(v \otimes w)d)d$$

$$= c(c \otimes 1)t_{N,N,N}(u \otimes dc(v \otimes w)d)d = c(c \otimes 1)t_{N,N,N}(u \otimes (v \otimes w)d)d$$

$$= c(c \otimes 1)((u \otimes v) \otimes w)t_{N,N,N}(1 \otimes d)d.$$

Similarly,

$$((u \oplus v) \oplus w)t = c(c(u \otimes v)d \otimes w)dc(c \otimes 1)t_{N,N,N}(1 \otimes d)d$$

$$= c(c(u \otimes v)d \otimes w)(c \otimes 1)t_{N,N,N}(1 \otimes d)d$$

$$= c(c(u \otimes v) \otimes w)(d \otimes 1)(c \otimes 1)t_{N,N,N}(1 \otimes d)d$$

$$= c(c(u \otimes v) \otimes w)t_{N,N,N}(1 \otimes d)d$$

$$= c(c \otimes 1)((u \otimes v) \otimes w)t_{N,N,N}(1 \otimes d)d.$$

Therefore $t(u \oplus (v \oplus w)) = ((u \oplus v) \oplus w)t$.

(iii) By definition of t,

$$(t \oplus 1)t(1 \oplus t) =$$

$$c(c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)d \otimes 1_N)d$$

$$c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)d)$$

$$c(1_N \otimes c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)d)d$$

$$= c(c(c \otimes 1_N)t_{N,N,N} \otimes 1_N)((1_N \otimes d) \otimes 1_N)$$

$$t_{N,N,N}$$

$$(1_N \otimes (c \otimes 1_N))(1_N \otimes t_{N,N,N}(1 \otimes d)d)d.$$

On the other hand,

$$t^2 = c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)dc(c \otimes 1_N)t_{N,N,N}(1 \otimes d)d$$

$$= c(c \otimes 1_N)t_{N,N,N}(c \otimes (1_N \otimes 1_N))$$

$$((1_N \otimes 1_N) \otimes d)t_{N,N,N}(1_N \otimes d)d$$

Which, by Lemma 7, gives

$$t^2 = c(c(c \otimes 1_N) \otimes 1_N)t_{N \otimes N,N,N}t_{N,N,N \otimes N}(1_N \otimes (1_N \otimes d)d)d.$$

However, by the MacLane Pentagon condition for a symmetric monoidal category, this is

$$t^2 = c(c(c \otimes 1_N) \otimes 1_N)(t_{N,N,N} \otimes 1_N)t_{N,N \otimes N,N}(1_N \otimes t_{N,N,N})(1_N \otimes (1_N \otimes d)d)d.$$

Again, by Lemma 7, this is

$$t^2 = c(c(c \otimes 1_N) \otimes 1_N)(t_{N,N,N} \otimes 1_N)$$

$$((1_N \otimes d) \otimes 1_N)t_{N,N,N}(1_N \otimes (c \otimes 1_N))$$

$$(1_N \otimes t_{N,N,N})(1_N \otimes (1_N \otimes d)d)d$$

$$= c(c(c \otimes 1_N)t_{N,N,N} \otimes 1_N)((1_N \otimes d) \otimes 1_N)$$

$$t_{N,N,N}$$

$$(1_N \otimes (c \otimes 1_N))(1_N \otimes t_{N,N,N}(1 \otimes d)d)d.$$

However, this is the expanded form of $(t \oplus 1)t(1 \oplus t)$, and so our result follows.

(iv) By definition of s and t above,

$$(s \oplus 1)t(1 \oplus s) = c(cs_{N,N}d \otimes 1_N)dc(c \otimes 1_N)t_{N,N,N}t_{N,N,N}(1_N \otimes d)dc(1_N \otimes cs_{N,N}d)d.$$

Similarly, $tst = c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)dcs_{N,N}dc(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)d$ and so, by Lemma 7,

$$tst = (c(c \otimes_N 1)t_{N,N,N}s_{N \otimes N,N}(d \otimes 1_N)(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)d$$

$$= c(c \otimes 1_N)t_{N,N,N}s_{N \otimes N,N}t_{N,N,N}(1_N \otimes d)d$$

Therefore, by the hexagon condition for a symmetric monoidal category

$$tst = c(c \otimes 1_N)(s_{N,N} \otimes 1_N)t_{N,N,N}(1_N \otimes s_{N,N}d)d$$

$$= c(c \otimes 1)(s_{N,N} \otimes 1_N)t_{N,N,N}(1_N \otimes s_{N,N})(1_N \otimes d)d$$

and this is the expanded version of $(s \otimes 1)t(1 \oplus s)$. Hence our result follows.

(v) We define $s^{-1} = \phi_2(s^{-1}_{N,N})$, and $t^{-1} = c(1_N \otimes c)t^{-1}_{N,N,N}(d \otimes 1_N)d$. (Note that, in the freely generated objects case, $t^{-1} = \phi(t^{-1}_{N,N,N})$). It is then immediate that $s^{-1} = s$, as $s^{-1}_{N,N} = s_{N,N}$, and it is immediate by **(i)** that $s^2 = (1 \oplus 1)$.

Also, $tt^{-1} = c(c \otimes 1_N)t_{N,N,N}(1_N \otimes d)dc(1_N \otimes c)t^{-1}_{N,N,N}(d \otimes 1_N)d$ and so

$$tt^{-1} = c(c \otimes 1_N)t_{N,N,N}t^{-1}_{N,N,N}(d \otimes 1_N)d = c(c \otimes 1_N)((1_N \otimes 1_N) \otimes 1_N)(d \otimes 1_N)d = ((1 \oplus 1) \oplus 1),$$

by definition of $\oplus$. The identity $t^{-1}t = (1 \oplus (1 \oplus 1))$ follows similarly.

(vi) We have seen that strong self-similarity implies that $1 \oplus 1 = 1$, in which case $\oplus$ defines a functor, and the axioms for a one-object symmetric monoidal category then follow immediately from **(i)** to **(v)** above. □

4.3 Definitions and theory of M-monoids

We axiomatise the above properties of endomorphism monoids of self-similar objects.

Definitions 4.6

We define an *M-monoid* to be a monoid, M, together with a semigroup homomorphism $\oplus : M \times M \to M$ that we call the *internal tensor*, and special elements s, t that we call the *commutativity* and *associativity* elements, that satisfy

1. $t(u \oplus (v \oplus w)) = ((u \oplus v) \oplus w)t$.
2. $s(u \oplus v) = (v \oplus u)s$.
3. $t^2 = (t \oplus 1)t(1 \oplus t)$.
4. $tst = (s \oplus 1)t(1 \oplus s)$.
5. s and t have generalised inverses satisfying $s^{-1} = s$, $s^2 = (1 \oplus 1)$, $tt^{-1} = (1 \oplus 1) \oplus 1$ and $t^{-1}t = 1 \oplus (1 \oplus 1)$.

We call an M-monoid *strong* if it satisfies the further condition that $\oplus$ is a monoid homomorphism. This implies, and is implied by the condition that the generalised inverses of s, t are global inverses (i.e. $s^{-1}s = 1$ and $t^{-1}t = 1 = tt^{-1}$). Of course, strong M-monoids are one-object symmetric monoidal categories, apart from the units conditions. The following results are then immediate from Theorem 9:

Proposition 10 *Every endomorphism monoid of a self-similar object in a symmetric monoidal category is an M-monoid.* □

Corollary 11 *Let N be a strongly self-similar object of a symmetric monoidal category. Then the endomorphism monoid of N is a strong M-monoid.* □

We can also deduce the following from Chapter 3:

Theorem 12 *Let P_2 be embedded in a unital ring R. Then the monoid of R is an M-monoid, and is strong when the embedding of P_2 is strong.*

Proof We identify the generators of P_2 under the embedding claimed. Then the internal tensor is defined by $(r \oplus s) = p^{-1}rp \oplus q^{-1}sq$ for all $r, s \in R$. The M-monoid structure of the monoid of R then follows by Proposition 20 of Chapter 3, and the strongness result is a restatement of Theorem 21 of Chapter 3. □

Note that this proof does not use the additive inverses of the ring R, and hence also holds for semirings.

4.3.1 Constructing embeddings of polycyclic monoids

We relate the categorical approach to self-similarity given above to the inverse semigroup theoretic approach given in Chapters 1 to 3. Specifically, we give sufficient conditions for the existence of an embedding of the polycyclic monoid on two generators into the endomorphism monoid of a self-similar object.

Definitions 4.7

Let N be a self-similar object of a symmetric monoidal category $(\mathbf{M}, \otimes)$, and let $\mathbf{M}(N, N)$ have a zero, which we denote 0. We say that maps $\pi_1, \pi_2 \in \mathbf{M}(N \otimes N, N)$ are left and right *projection maps* and $i_1, i_2 \in \mathbf{M}(N, N \otimes N)$ are their *inclusion maps*, if they satisfy

1. $\pi_1 i_1 = 1_N$ and $\pi_2 i_2 = 1_N$,
2. $i_1\pi_1 = 1_N \otimes 0$ and $i_2\pi_2 = 0 \otimes 1_N$.

Theorem 13 *Let N be a self-similar object of a symmetric monoidal category $(\mathbf{M}, \otimes)$. If N has projection and inclusion maps, then there exists an embedding of P_2 into $\mathbf{M}(N, N)$.*

Proof As ϕ_2 preserves composition, $\phi_2(\pi_1)\phi_2(i_1) = 1_N = \phi_2(\pi_2)\phi_2(i_2)$ and $\phi_2(i_1)\phi_2(\pi_1) = 1 \oplus 0$, and $\phi_2(i_2)\phi_2(\pi_2) = 0 \oplus 1$. Hence, as $\oplus$ is a semigroup homomorphism, $\phi_2(i_1)\phi_2(\pi_1)\phi_2(i_2)\phi_2(\pi_2) = 0 = \phi_2(i_2)\phi_2(\pi_2)\phi_2(i_1)\phi_2(\pi_1)$. Therefore, $\phi_2(\pi_1)\phi_2(i_2) = 0 = \phi_2(\pi_2)\phi_2(i_1)$, and so $\phi_2(\pi_1), \phi_2(\pi_2)$ satisfy the axioms for the generators of P_2, and $\phi_2(i_1), \phi_2(i_2)$ satisfy the conditions for their generalised inverses. Hence, by the definition of ϕ_2, the embedding of P_2 into $\mathbf{M}(N, N)$ is generated by

$$p^{-1} \mapsto ci_1 \ , \ q^{-1} \mapsto ci_2,$$

and their generalised inverses

$$p \mapsto \pi_1 d \ , \ q \mapsto \pi_2 d.$$

Therefore, we have constructed an embedding of P_2 into $\mathbf{M}(N, N)$. □

Definitions 4.8

We refer to $\phi_2(\pi_1)$ and $\phi_2(\pi_2)$ as the left and right *internal projections* of $\oplus$. Similarly, we refer to $\phi_2(i_1)$ and $\phi_2(i_2)$ as the left and right *internal inclusions* of $\oplus$. We identify the generators of P_2 and their generalised inverses with their images under this embedding, for clarity.

Proposition 14 *$p^{-1}(f \oplus g)p = f$ and $q^{-1}(f \oplus g)q = g$ for all $f, g \in \mathbf{M}(N, N)$.*

Proof By definition, $p^{-1}(f \oplus g)p = \phi_2(\pi_1(f \otimes g)i_1) = \phi_2(f) = f$. Similarly, $q^{-1}(f \oplus g)q = \phi_2(\pi_2(f \otimes g)i_2) = \phi_2(g) = g$. Hence our result follows. □

Definitions 4.9

Let $(M, \oplus)$ be an M-monoid with a zero. We say that it has *internal projections* if there exists an embedding of P_2 into M satisfying $p(f \oplus g)p^{-1} = f$ and $q(f \oplus g)q^{-1} = g$ for all $f, g \in M$. Similarly, we say that it has *internal inclusions* if there exists an embedding of P_2 into M satisfying $p^{-1}fp = f \oplus 0$ and $q^{-1}gq = 0 \oplus g$ for all $f, g \in M$.

From the above, we are in a position to prove the converse to Theorem 13, Chapter 3.

Theorem 15 *Let R be a unital ring, and assume there exists $c \in \mathbf{Mat}_{\mathbf{R}}(2,1)$, $d \in \mathbf{Mat}_{\mathbf{R}}(1,2)$ such that the map $\phi : M_2(R) \to R$, defined by $\phi(X) = cXd$, is an injective ring homomorphism. Then P_2 is embedded in R.*

Proof First note that the condition on R is equivalent to stating that 1 is a self-similar object of the category $\mathbf{Mat_R}$. We demonstrate that $\mathbf{Mat_R}$ has inclusions and projections. Define

$$\pi_1 = (1\ 0)\ ,\ \pi_2 = (0\ 1),$$

and

$$i_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\ ,\ i_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then it is immediate from the definition of composition in $\mathbf{Mat_R}$ that $\pi_1 i_1 = 1_R = \pi_2 i_2$, and

$$i_1\pi_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\ , i_2\pi_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore, $i_1\pi_1 = 1 \sqcup 0$ and $i_2\pi_2 = 0 \sqcup 1$, and so $\mathbf{Mat_R}$ has projections and inclusions. Therefore, by Theorem 13 above, there exists an embedding of P_2 into R. □

4.3.2 Fixed points of internal tensors

We formalise the definition of a 'fixed point' for an internal tensor. This concept will be very useful in the construction of algebraic models of logical systems, presented in Chapters 7 and 8, and we will give a concrete example in Chapter 5. This definition was motivated by the properties J.-Y. Girard requires for a model of the exponential operator of linear logic [17, 20, 21].

Definitions 4.10

Let $(M, \oplus)$ be an M-monoid. We define a *right fixed point* of an element $f \in M$ to be an element F of M that satisfies $f \oplus F = F$. Similarly, a *left fixed point* of f is an element F' that satisfies $F' \oplus f = F'$. Note that these two definitions imply, for an M-monoid with internal projections, $p^{-1}Fp = f$ and $q^{-1}Fq = F$ when F is a right fixed point of f. Similarly, for a left fixed point, $p^{-1}F'p = F'$ and $q^{-1}F'q = f$. If every element of an M-monoid M has a right (resp. left) fixed point, we call M a *right (resp. left) recursive* M-monoid. Also, if there exists a semigroup homomorphism $Y : M \to M$ where $Y(f)$ is a right (resp. left) fixed point for f, for all $f \in M$, we call Y a right (resp. left) *fixed point* or *recursion* homomorphism.

4.4 Expanding M-monoids into categories

Consider a self-similar object N of a symmetric monoidal category $(\mathbf{M}, \otimes)$. From the description of the contraction of $(\otimes \mathrm{N})$ into the endomorphism monoid of N, we can see that identities of objects of $(\otimes N)$ are mapped to partial identities (i.e. idempotents) of the endomorphism monoid of N — in the case when N is a weak self-similar object, and $(\otimes N)$ has freely generated objects, this is an injective correspondence. We consider how reversible this contraction process is, by constructing categories from the idempotents of M–monoids. The standard technique for doing this is the Karoubi envelope, and we define various restrictions of this, to give an inverse process to the contraction map ϕ. This allows us to consider ϕ to be a (partial) dual to the Karoubi envelope.

4.4.1 The Karoubi envelope of an M-monoid

Definitions 4.11

Let M be a monoid. The *Karoubi envelope* of M is defined to be the category, which we denote $\mathbf{K_M}$, specified by, for all $a, e, f \in M$,

- $e \in Ob(\mathbf{K_M})$ iff $e^2 = e$,
- $a \in \mathbf{K_M}(e, f)$ iff $fa = a = ae$.

Note that the Karoubi envelope is in fact defined for arbitrary categories (see [38] for details of the general construction). However, we are only interested in the one-object, or monoid case.

Proposition 16 *Let $(M, \oplus)$ be an M-monoid. Then $\mathbf{K_M}$ is a symmetric monoidal category without the unit object.*

Proof Define the monoidal functor $\otimes$ on $\mathbf{K_m}$, as follows:

- $e \otimes f = e \oplus f$, for all $e, f \in Ob(\mathbf{K_M})$.
- $a \otimes b = a \oplus b$ for all $a \in \mathbf{K_M}(e, f)$, $b \in \mathbf{K_M}(e', f')$.

As $\oplus$ is a semigroup homomorphism, $(f \otimes f')(a \otimes b)(e \otimes e') = (a \otimes b)$ for all $a \in \mathbf{K_M}(e, f)$, $b \in \mathbf{K_M}(e', f')$, by definition of $\mathbf{K_M}$. Also, as $\oplus$ is a semigroup homomorphism, $(a \otimes b)(c \otimes d) = (ac \otimes bd)$ when ac and bd are defined in $\mathbf{K_M}$. Therefore, $\otimes$ is a functor from $\mathbf{K_M} \times \mathbf{K_M}$ to $\mathbf{K_M}$. We can then define, for all $e, f, g \in Ob(\mathbf{K_M})$,

- $s_{e,f} = (f \oplus e)s(e \oplus f)$,
- $t_{e,f,g} = ((e \oplus f) \oplus g)t(e \oplus (f \oplus g))$,
- $t^{-1}_{e,f,g} = (e \oplus (f \oplus g))t^{-1}((e \oplus f) \oplus g)$,

where these composites are taken in M. From these definitions,

$$s_{e,f} \in \mathbf{K_M}(e \oplus f, f \oplus e) \ , \ t_{e,f,g} \in \mathbf{K_M}(e \oplus (f \oplus g), (e \oplus f) \oplus g),$$

and the axioms for a symmetric monoidal category (apart from the units conditions) follow immediately from the axioms for an M-monoid (Definitions 4.6). Therefore, $\mathbf{K_M}$ has the structure of a symmetric monoidal category apart from the units conditions. □

Proposition 17 *The identity, 1, of an M-monoid* $(M, \oplus)$ *is a self-similar object of the symmetric monoidal category* $(\mathbf{K_M}, \otimes)$.

Proof We require elements $c \in \mathbf{K_M}(1 \otimes 1, 1)$ and $d \in \mathbf{K_M}(1, 1 \otimes 1)$ that satisfy $dc = 1_{1\otimes 1}$. Note that $1(1 \oplus 1) = (1 \oplus 1) = (1 \oplus 1)1$ in M, so we can define $c = (1 \oplus 1) \in \mathbf{K_M}(1 \otimes 1, 1)$ and $d = (1 \oplus 1) \in \mathbf{K_M}(1, 1 \otimes 1)$. Then it is immediate that $dc = 1 \oplus 1 = 1_{1\otimes 1}$. Therefore, 1 is a self-similar object of $\mathbf{K_M}$. □

The above result is trivial when $(M, \oplus)$ is a strong M-monoid, in which case $1 \otimes 1$ is the same object of $\mathbf{K_M}$ as 1.

Lemma 18 *Let* N *be a self-similar object of a symmetric monoidal category* $(\mathbf{M}, \otimes)$, *and assume* $(\otimes \mathrm{N})$ *has freely generated objects. Then we can define* ϕ *as in Definitions 4.3, and for all* $a \in (\otimes \mathrm{N})(X, Y)$,

$$\phi(a) \in \mathbf{K}_{\mathbf{M}(N,N)}(\phi(1_X), \phi(1_Y)).$$

Proof As $a \in (\otimes \mathrm{N})(X, Y)$, $1_Y a 1_X = a$,and so, by Theorem 5, $\phi(1_Y)\phi(a)\phi(1_X) = \phi(a)$. Therefore, our result follows, by definition of $\mathbf{K}_{\mathbf{M}(N,N)}$. □

4.4.2 The tensor envelope of an M-monoid

In what follows we will use a restriction of the Karoubi envelope, to demonstrate the connection between (weak) self-similarity and the idempotents of an M-monoid. This is defined as follows:

Definitions 4.12

Let $(M, \oplus)$ be an M-monoid. We define its *Karoubi tensor envelope*, or just *tensor envelope*, which we denote $\mathbf{K}^{\oplus}_{\mathbf{M}}$, to be the subcategory of $\mathbf{K_M}$ specified by

- $1 \in Ob(\mathbf{K}^{\oplus}_{\mathbf{M}})$.

- For all $e, f \in Ob(\mathbf{K}_{\mathbf{M}}^{\oplus})$, $(e \oplus f) \in Ob(\mathbf{K}_{\mathbf{M}}^{\oplus})$.

- $a \in \mathbf{K}_{\mathbf{M}}^{\oplus}(e, f)$ iff $fa = a = ae$, as before.

The following is then immediate from this definition:

Proposition 19 *Let $M, \oplus$ be a strong M-monoid. Then $\mathbf{K}_{\mathbf{M}}^{\oplus}$ is a one-object category isomorphic to M.*

Proof Immediate from the definition of a strong M-monoid, where $1 \oplus 1 = 1$, and from the definition of $\mathbf{K}_{\mathbf{M}}^{\oplus}$. □

Definitions 4.13

Let N be a weakly self-similar object of a symmetric monoidal category $(\mathbf{M}, \otimes)$, where $(\otimes\mathbf{N})$ has freely generated objects, and denote the internalisation of $\otimes$ at N by $\oplus$. We define a map $\Phi : (\otimes\mathbf{N}) \to \mathbf{K}_{\mathbf{M(N,N)}}^{\oplus}$ as follows:
Let X and Y be arbitrary objects of $(\otimes\mathbf{N})$, and let $a \in (\otimes\mathbf{N})(X, Y)$ be an arbitrary morphism. Then

- $\Phi(X) = \phi(1_X) = c_X d_X$,

- $\Phi(a) = \phi(a) = c_Y a d_X \in \mathbf{K}_{\mathbf{M(N,N)}}(\phi(1_Y), \phi(1_X))$.

We also define a map $\Psi : \mathbf{K}_{\mathbf{M(N,N)}}^{\oplus} \to (\otimes\mathbf{N})$, as follows:
On objects,

- $\Psi(1) = N$,

- $\Psi(X \otimes Y) = \Psi(X) \otimes \Psi(Y)$, for all $X, Y \in Ob(\mathbf{K}_{\mathbf{M(N,N)}}^{\oplus})$.

Note that for all $e \in Ob(\mathbf{K}_{\mathbf{M}}^{\oplus})$ this implies that $\Psi(e) = c_X d_X$, for some unique $X \in Ob(\otimes\mathbf{N})$, by the freely generated objects condition. We can then define Ψ on morphisms by

- $\Psi(a) = d_Y a c_X$ for all $a \in (\otimes\mathbf{N})(\Phi(X), \Phi(Y))$.

Theorem 20 *Φ and Ψ are mutually inverse $\otimes$-preserving category isomorphisms.*

Proof For all $X \in Ob(\otimes\mathbf{N})$, $\Phi(1_X) = d_X c_X$, the identity of $\Phi(X)$. Also, for all b, a composable in $(\otimes\mathbf{N})$, $\Phi(b)\Phi(a) = \phi(b)\phi(a) = \phi(ba) = \Phi(ba)$. Therefore, Φ is a functor.
Similarly, $\Psi(1_e) = 1_X$ for all $e \in Ob(\mathbf{K}_{\mathbf{M(N,N)}}^{\oplus})$, where $\Phi(X) = e$. Also, for all

$$a \in \mathbf{K}_{\mathbf{M(N,N)}}^{\otimes}(\Phi(X), \Phi(Y)) \ , \ b \in \mathbf{K}_{\mathbf{M(N,N)}}^{\oplus}(\Phi(Y), \Phi(Z)),$$

$\Psi(b)\Psi(a) = d_Z b c_Y d_Y a c_X = d_Z d b 1_Y a c_X = d_Z b a c_X = \Psi(ba)$ by definition of Ψ. Therefore, Ψ is also a functor.

Next, for arbitrary $X, Y \in Ob(\otimes\mathbf{N})$,

$$\Phi(X \otimes Y) = \phi(1_{X \otimes Y}) = \phi(1_X \oplus 1_Y) = \phi(1_X) \oplus \phi(1_Y) = \Phi(X) \otimes \Phi(Y),$$

by definition of $\otimes$ in $\mathbf{K}^{\oplus}_{\mathbf{M(N,N)}}$. Similarly, for arbitrary $a, b \in Hom(\otimes\mathbf{N})$,

$$\Phi(a) \otimes \Phi(b) = \phi(a) \oplus \phi(b) = \phi(a \otimes b) = \Phi(a \otimes b).$$

Therefore, Φ preserves $\otimes$.

The definition of Ψ gives $\Psi(e \otimes f) = \Psi(e) \otimes \Psi(f)$ for arbitrary $e, f \in Ob(\mathbf{K}^{\oplus}_{\mathbf{M(N,N)}})$. Also, for arbitrary $a \in \mathbf{K}^{\oplus}_{\mathbf{M(N,N)}}(e, f)$, $b \in \mathbf{K}^{\oplus}_{\mathbf{M(N,N)}}(e', f')$,

$$\Psi(a \otimes b) = d_{(Y \otimes Y')}(a \otimes b) c_{(X \otimes X')},$$

where

$$Y = \Psi(f)\ ,\ Y' = \Psi(f'),$$

$$X = \Psi(e)\ ,\ X' = \Psi(e').$$

However, $d_{Y \otimes Y'} = (d_Y \otimes d_{Y'})d$, and $c_{X \otimes X'} = c(c_X \otimes c_{X'})$, by Definition 4.3, so

$$\Psi(a \otimes b) = (d_Y \otimes d_{Y'})d(a \oplus b)c(c_X \otimes c'_X) = (d_Y \otimes d_{Y'})dc(a \otimes b)dc(c_X \otimes c'_X)$$

$$= (d_Y \otimes d_{Y'})(a \otimes b)(c_X \otimes c_{X'}) = (d_Y a c_X \otimes d_{Y'} b c_{X'}) = \Psi(a) \otimes \Psi(b).$$

Therefore Ψ also preserves $\otimes$.

Hence Ψ and Φ are both $\otimes$ preserving functors.

Next, let X be an arbitrary object of $(\otimes\mathbf{N})$. Then $\Psi(\Phi(X)) = \Psi(1_X) = X$. Conversely, let e be an arbitrary object of $\mathbf{K}^{\oplus}_{\mathbf{M(N,N)}}$. Then $\Psi(e) = X$, for some $X \in Ob(\otimes\mathbf{N})$ satisfying $\Phi(X) = e$. Therefore, $\Phi(\Psi(e)) = e$.

Finally, consider arbitrary $a \in (\otimes\mathbf{N})(X, Y)$. By definition,

$$\Psi(\Phi(a)) = \Psi(c_Y a d_X) = d_Y c_Y a d_X c_X = 1_Y a 1_X = a \in (\otimes\mathbf{N})(X, Y).$$

Conversely, consider arbitrary $b \in \mathbf{K}^{\oplus}_{\mathbf{M(N,N)}}(e,f)$, where $\Phi(X) = e$ and $\Phi(Y) = f$. Then

$$\Phi(\Psi(b)) = \Psi(d_Y b c_X) = c_Y d_Y b c_X d_X = fbe = e.$$

Therefore, Φ and Ψ are mutually inverse $\otimes$-preserving functors, and so $K^{\oplus}_{\mathbf{M}(N,N)} \cong (\otimes \mathbf{N})$. □

Hence, we are justified in considering the map ϕ from $(\otimes\mathbf{N})$ to the M-monoid $(\mathbf{M}(N,N),\oplus)$ to be a partial dual to the Karoubi envelope when $(\otimes\mathbf{N})$ has freely generated objects.

4.5 Self-similar objects in Cartesian closed categories

It is well known ([1, 23, 38]) that Cartesian closed categories are models of typed lambda calculi. In a similar way, one-object Cartesian closed categories without the terminal object are models of untyped lambda calculus. One-object Cartesian closed categories, or monoids, of this form are called C-monoids, and it is a classical result that the Karoubi envelope of a C-monoid is a Cartesian closed category. See [38] for details of the theory of C-monoids.

We consider the theory of endomorphism monoids of strongly self-similar objects of Cartesian closed categories. This allows us to construct 1-object analogues of the 'internal hom.', which is the essential part of the categorical closure.

Definitions 4.14

A *Cartesian closed category* is a category $\mathbf{C}$, where the Cartesian product $\mathbf{C}\times\mathbf{C} \to \mathbf{C}$ is a functor that satisfies, for all $A, B, C \in Ob(\mathbf{C})$ and $f \in \mathbf{C}(C \times A, B)$,

1. There exists a terminal object I; that is, for all $A \in Ob(\mathbf{C})$ there exists a unique arrow $t_A : A \to I$.
2. For all $A, B \in Ob(\mathbf{C})$ there exist projections $\pi_1 : A \times B \to A$ and $\pi_2 : A \times B \to B$.
3. For all objects A, B, C and morphisms $f \in \mathbf{C}(C \times A, B)$ there exists:
   - An object, which we denote $[A \to B]$,
   - A morphism $\epsilon \in \mathbf{C}([A \to B] \times A, B)$ (the *evaluation morphism*),
   - A unique morphism $\lambda f \in \mathbf{C}(C, [A \to B])$ (the *abstraction morphism*)

   that satisfy $\epsilon(\lambda f \times 1_A) = f$.

(We are following the conventions and notation of [1]; this is based on the description given in [38], apart from the assumption that there is a 'natural numbers object').

We demonstrate how the existence of a strongly self-similar object (satisfying certain extra conditions) in a Cartesian closed category allows us to construct a one-object analogue of axiom 3 above. Let us denote $(\otimes N)$ by $(\times N)$ when $\otimes$ is the cartesian product $\times$. Then the following theorem holds:

Theorem 21 *Let N be a strongly self-similar object of a Cartesian closed category, $\mathbf{C}$, where $Ob(\times \mathbf{N})$ is closed under the $[\ \rightarrow\ ]$ operation. Then the endomorphism monoid of N in $\mathbf{C}$ has elements α and η that satisfy the one-object analogues of the axioms for the evaluation and abstraction maps.*

Proof We denote the internalisation of the Cartesian product by $\oplus$. First note that, as the Cartesian product has projections, $A \times B = A' \times B'$ implies that $\pi_1(A \times B) = \pi_1(A' \times B')$, and so $A = A'$. Similarly, $B = B'$, and so $(\times N)$ has freely generated objects. Therefore, we can uniquely define c_X and d_X, for all $X \in Ob(\times \mathbf{N})$ as shown in Definitions 4.3, and construct a map ϕ from $(\times \mathbf{N})$ to N that preserves composition, and maps $\times$ to $\oplus$, as shown in Theorem 5 and Proposition 8. Then for all $f \in \mathbf{C}(N, N)$, we define $\eta = \phi(\epsilon)$ and $\alpha(f) = \phi(\lambda(fc))$, and this gives, by definition, $\eta(\alpha(f) \oplus 1) = \phi(\epsilon)(\phi(\lambda(fc) \oplus 1)) = \phi(\epsilon)\phi(\lambda(fc) \otimes 1) = \phi(\epsilon(\lambda(fc) \otimes 1))$ which is $\phi(fc)$, by definition of a Cartesian closed category. Finally, $\phi(fc) = fcd = f$, by definition of ϕ, and the strong self-similarity of N. Therefore, $\eta(\alpha(f) \oplus 1) = f$, for all $f \in \mathbf{C}(N, N)$, which is the one-object analogue of axiom 3 for a Cartesian closed category. □

Chapter 5

The natural numbers as a self-similar object

5.1 Introduction

We present the theory of the category of relations, with particular emphasis on the subcategory of partial bijective maps. We show how the disjoint union of sets (a coproduct) is a symmetric monoidal functor, and give a matrix representation of morphisms, along with the conditions for a matrix to represent a partial bijective map.

We then show that the set of natural numbers is a self-similar object with respect to both the disjoint union monoidal functor and the usual Cartesian product of sets. These two distinct self-similar structures give rise to two distinct internal tensors on the inverse semigroup of partial bijective maps on the natural numbers, with the internalisation of the disjoint union having internal projections / inclusions, and the internalisation of the Cartesian product having internal projections. We relate these properties to the embeddings of polycyclic monoids used to characterise bijections from $\mathbb{N}$ to $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N} \sqcup \mathbb{N}$ in Chapter 2. Finally we demonstrate how under certain conditions relating the Cartesian product to the coproduct, the internalised Cartesian product can be used to construct an injective right fixed point homomorphism for the internalised disjoint union.

5.2 The category of relations

Definitions 5.1

The *category of relations*, **Rel**, is the category with all sets as objects, and relations between sets as morphisms. So $R \in \mathbf{Rel}(X,Y)$ is a relation $R \subseteq Y \times X$. Given $R \in \mathbf{Rel}(X,Y)$ and $S \in \mathbf{Rel}(Y,Z)$, their composite, $SR \in \mathbf{Rel}(X,Z)$ is defined by $(z,x) \in SR$ iff there exist $y \in Y$ satisfying $(y,x) \in R$ and $(z,y) \in S$. The identities at objects are the diagonal relations $\Delta_X = \{(x,x) : x \in X\}$.

Lemma 1 *Composition in* **Rel** *distributes over union of sets. That is, given* $R \in \mathbf{Rel}(X,Y)$, $S,T \in \mathbf{Rel}(Y,Z)$, *and* $U \in \mathbf{Rel}(Z,T)$, *then* $(S \cup T)R = SR \cup TR$ *and* $U(S \cup T) = US \cup UT$.

Proof By definition of composition,

$$(z,x) \in (S \cup T)R \Rightarrow \exists y \in Y : \ (y,x) \in R \ , \ (z,y) \in (S \cup T).$$

However, this is equivalent to

$$(z,x) \in (S \cup T)R \Rightarrow (y,x) \in R \ , \ (z,y) \in S \text{ or } (z,y) \in T \text{ for some } y \in Y.$$

Therefore, $(S \cup T)R = SR \cup TR$. A similar proof gives that $U(S \cup T) = US \cup UT$, and so composition distributes over union. □

Definitions 5.2

The endomorphism monoid of a set $X \in Ob(\mathbf{Rel})$ is the monoid of relations from X to itself, denoted $B(X)$. Given arbitrary $R \in B(X)$, we define R^n, for $n \in \mathbb{N}$, by $R^0 = \Delta_X$ and $R^{i+1} = RR^i$. Given $R \in B(X)$, its *transitive closure*, R^∞, is defined in [30] to be $R^\infty = \bigcup_{i=1}^\infty R^n$, where it is proved to be the smallest transitive relation containing R. Of more interest to us is the *Kleene star* operation, denoted $(\ )^*$, defined (in [30]) on $B(X)$ by $R^* = \bigcup_{i=0}^\infty R^i$.

Proposition 2 R^* *is the smallest reflexive transitive relation containing* R.

Proof Clearly, R^* contains R. Also, $\Delta_X \subseteq R$, so R is reflexive. A relation $\rho \in B(X)$ is transitive iff it satisfies $(z,y),(y,x) \in \rho \Rightarrow (z,x) \in \rho$ or equivalently, $\rho^2 \subseteq \rho$. However, R^∞ is transitive, and $R^* = I \cup R^\infty$, and so R^* is clearly transitive.

Conversely, given any reflexive transitive relation S containing R, then $\Delta_X \in S$, since S is reflexive, and by transitivity of S, $R^i \subseteq S$ for all $i \in \mathbb{N}$, and so $R^* \subseteq S$. Therefore, R^* is the smallest reflexive transitive relation on X containing R. □

5.2.1 The monoidal structure of the category of relations

Definitions 5.3

We define the coproduct on **Rel** to be the map $\sqcup : \mathbf{Rel} \times \mathbf{Rel} \to \mathbf{Rel}$ given by $X \sqcup Y = X \times \{0\} \cup Y \times \{1\}$ on objects, and $R \sqcup S = \{((y,0),(x,0))\} \cup \{((v,1),(u,1))\}$ for all $(y,x) \in R$, $(v,u) \in S$ on morphisms.

Lemma 3 *$\sqcup$ is a functor from* $\mathbf{Rel} \times \mathbf{Rel}$ *to* **Rel**.

Proof Consider arbitrary $R \in \mathbf{Rel}(X,Y)$, $S \in \mathbf{Rel}(Y,Z)$, $T \in \mathbf{Rel}(A,B)$, $U \in \mathbf{Rel}(B,C)$. We will demonstrate that $SR \sqcup UT = (S \sqcup U)(R \sqcup T)$.

First note that

$$SR = \{(z,x) : \exists y \in Y \text{ satisfying } (y,x) \in R, (z,y) \in S\}$$

and

$$UT = \{(c,a) : \exists b \in B \text{ satisfying } (b,a) \in T, (c,b) \in U\}.$$

Then by definition of $\sqcup$,

$$(SR \sqcup UT) = \{((z,0),(x,0)) : (z,x) \in SR\} \cup \{((c,1),(a,1)) : (c,a) \in TU\}.$$

Conversely,

$$R \sqcup T = \{((y,0),(x,0)) : (y,x) \in R\} \cup \{((b,1),(a,1)) : (b,a) \in T\}$$

and

$$S \sqcup U = \{((z,0),(y,0)) : (z,y) \in S\} \cup \{((c,1),(b,1)) : (c,b) \in U\},$$

so

$$(S \sqcup U)(R \sqcup T) = \{((z,0),(x,0)) : \exists y \in Y \text{ satisfying } (y,x) \in R, (z,y) \in S\} \cup$$

$$\{((c,1),(a,1)) : \exists b \in B \text{ satisfying } (b,a) \in T, (c,b) \in U\}.$$

Therefore, $SR \sqcup UT = (S \sqcup U)(R \sqcup T)$. Also, it is immediate that $\Delta_X \sqcup \Delta_Y = \Delta_{X \sqcup Y}$ and so $\sqcup$ is a functor. □

Definitions 5.4

For all sets X, Y, Z, we define relations that play the rôle of commutativity, associativity, and units morphisms respectively. We also choose the empty set $\{\}$, to play the rôle of the identity object. The special morphisms are denoted

- $S_{X,Y} \in \mathbf{Rel}(X \sqcup Y, Y \sqcup X)$
- $T_{X,Y,Z} \in \mathbf{Rel}(X \sqcup (Y \sqcup Z), (X \sqcup Y) \sqcup Z)$
- $\rho_X \in \mathbf{Rel}(X \sqcup \{\}, X)$
- $\lambda_X \in \mathbf{Rel}(\{\} \sqcup X, X)$.

and these are defined as follows:

- $S_{X,Y} = \{((x,1),(x,0)) : x \in X\} \cup \{((y,0),(y,1)) : y \in Y\}$
- $T_{X,Y,Z} = \{(((x,0),0),(x,0)) : x \in X\} \cup \{(((y,1),0),((y,0),1)) : y \in Y\} \cup \{((z,1),((z,1),1)) : z \in Z\}$
- $\rho_X = \{(x,(x,0)) : x \in X\}$
- $\lambda_X = \{(x,(x,1)) : x \in X\}$.

These morphisms are all isomorphisms; their inverses are as follows:

- $S^{-1}_{X,Y} = \{((x,0),(x,1)) : x \in X\} \cup \{(y,1),(y,0) : y \in Y\} = S_{Y,X}$
- $T^{-1}_{X,Y,Z} = \{((x,0),((x,0),0)) : x \in X\} \cup \{(((y,0),1),((y,1),0)) : y \in Y\} \cup \{(((z,1),1),(z,1)) : z \in Z\}$
- $\rho^{-1}_X = \{((x,0),x) : x \in X\}$
- $\lambda^{-1}_X = \{((x,1),x) : x \in X\}$

Theorem 4 $(\mathbf{Rel}, \sqcup, S, T, \rho, \lambda, \{\})$ *is a symmetric monoidal category.*

Proof We have already proved that $\sqcup$ is a functor. Recall the axioms for a symmetric monoidal category from Definitions 1.2, Chapter 1; then, for all sets X, Y, Z, by definition of composition of relations, and the definitions of S, T, λ, ρ above,

1. $S_{X,Y}S^{-1}_{X,Y} = \{((x,0),(x,0)) : x \in X)\} \cup \{((y,1),(y,1)) : y \in Y)\} = \Delta_{X \sqcup Y}$. Similarly, $S^{-1}_{X,Y}S_{X,Y} = \{((x,1),(x,1)) : x \in X)\} \cup \{((y,0),(y,0)) : y \in Y)\} = \Delta_{Y \sqcup X}$.
2.

$$T^{-1}_{X,Y,Z}T_{X,Y,Z} =$$

$$\{((x,0),(x,0)) : x \in X\} \cup$$

$$\{(((y,0),1),((y,0),1)) : y \in Y\} \cup$$

$$\{(((z,1),1),((z,1),1)) : z \in Z\} =$$

$$\Delta_{X \sqcup (Y \sqcup Z)},$$

$$T_{X,Y,Z} T^{-1}_{X,Y,Z} =$$

$$\{(((x,0),0),((x,0),0)) : x \in X\} \cup$$

$$\{(((y,1),0),((y,1),0)) : y \in Y\} \cup$$

$$\{((z,1),(z,1)) : z \in Z\} =$$

$$\Delta_{(X \sqcup Y) \sqcup Z}.$$

3.

$$\lambda_X^{-1} \lambda_X = \{((x,1),(x,1)) : x \in X\} = \Delta_{\{\} \sqcup X},$$

$$\lambda_X \lambda_X^{-1} = \{(x,x) : x \in X\} = \Delta_X.$$

4.

$$\rho_X^{-1} \rho_X = \{((x,0),(x,0)) : x \in X\} = \Delta_{X \sqcup \{\}},$$

$$\rho_X \rho_X^{-1} = \{(x,x) : x \in X\} = \Delta_X.$$

Hence the defining conditions for a symmetric monoidal category are satisfied. Next note that, for arbitrary sets U, V, W, X,

1.

$$T_{(U \sqcup V),W,X} T_{U,V,(W \sqcup X)} =$$

$$\{(\alpha, \beta) : u \in U, v \in V, w \in W, x \in X\} =$$

$$(T_{U,V,W} \sqcup \Delta_X) T_{U,(V \sqcup W),X} (\Delta_U \sqcup T_{V,W,X}),$$

where

$$\beta = (((u,0),((v,0),((w,0),(x,1)),1),1)$$

and

$$\alpha = (((u,0),(v,1)),0),((w,1),0),(x,1))).$$

Therefore, the MacLane Pentagon condition holds.

2. $(\rho_V \sqcup \Delta_U) T_{V,\{\},U} = \{(\epsilon, \delta) : u \in U, v \in V\} = (\Delta_V \sqcup \lambda_U)$, where $\delta = ((v,0),((u,1),1)$ and $\epsilon = ((v,0),(u,1))$. Therefore the units triangle condition holds.

3. $T_{W,U,V}S_{(U\sqcup V),W}T_{U,V,W} = \{(\gamma,\beta) : u \in U, v \in V, w \in W\} = (S_{U,W} \sqcup \Delta_V)T_{U,W,V}(\Delta_U \sqcup S_{V,W})$, where $\beta = ((u,0),((v,0),(w,1)),1)$ and $\gamma = ((((w,0),(u,1)),0),(v,1))$, so the commutativity hexagon condition holds.

Therefore, the coherence conditions are also satisfied.

Finally, to prove that the canonical morphisms are natural, consider arbitrary relations $f \in \mathbf{Rel}(X,A)$, $g \in \mathbf{Rel}(Y,B)$ and $g \in \mathbf{Rel}(Z,C)$. Then

$$f \sqcup g = \{((a,0),(x,0)) : (a,x) \in f\} \cup \{((b,1),(y,1)) : (b,y) \in g\}.$$

Similarly,

$$g \sqcup f = \{((a,1),(x,1)) : (a,x) \in f\} \cup \{((b,0),(y,0)) : (b,y) \in g\}.$$

Therefore,

$$(g \sqcup f)S_{X,Y} = \{((a,1),(x,0)) : (a,x) \in f\} \cup \{((b,0),(y,1)) : (b,y) \in g\}.$$

Conversely,

$$S_{A,B}(f \sqcup g) = \{((a,1),(x,0)) : (a,x) \in f\} \cup \{((b,0),(y,1)) : (b,y) \in g\},$$

and so $(g \sqcup f)S_{X,Y} = S_{A,B}(f \sqcup g)$, and hence X_{XY} is natural in X and Y.

Also,

$$f \sqcup (g \sqcup h) = \{((a,0),(x,0)) : (a,x) \in f\} \cup$$
$$\{(((b,0),1),((y,0),1)) : (b,y) \in g\} \cup$$
$$\{(((c,1),1),((z,1),1)) : (c,y) \in h\}.$$

Similarly,

$$(f \sqcup g) \sqcup h = \{(((a,0),),((x,0),0)) : (a,x) \in f\} \cup$$
$$\{(((b,1),0),((y,1),0)) : (b,y) \in g\} \cup$$
$$\{((c,1),(z,1)) : (c,y) \in h\}.$$

Therefore,

$$((f \sqcup g) \sqcup h)T_{X,Y,Z} = \{(((a,0),0),(x,0)) : (a,x) \in f\} \cup$$
$$\{(((b,1),0),((y,0),1) : (b,y) \in g\} \cup$$
$$\{((c,1),((z,1),1)) : (c,z) \in h\}.$$

Conversely,

$$T_{A,B,C}(f \sqcup (g \sqcup h)) = \{(((a,0),0),(x,0)) : (a,x) \in f\} \cup$$

$$\{(((b,1),0),((y,0),1) : (b,y) \in g\} \cup$$

$$\{((c,1),((z,1),1)) : (c,z) \in h\}.$$

Therefore, $T_{A,B,C}(f \sqcup (g \sqcup h)) = ((f \sqcup g) \sqcup h)T_{X,Y,Z}$ and so $T_{X,Y,Z}$ in natural in X and Y and Z.

Finally, $f \sqcup \{\} = \{((a,0),(x,0)) : (a,x) \in f\}$ and so

$$\rho_A(f \sqcup 1_{\{\}}) = \{(a,(x,0)) : (a,x) \in f\}.$$

Conversely,

$$f\rho_X = \{(a,(x,0)) : (a,x) \in f\}.$$

Therefore, $\rho_A(f \sqcup 1_{\{\}}) = f\rho_X$ and so ρ_X is natural in X. A similar proof suffices to show that λ_X is natural in X.

Therefore $(\mathbf{Rel}, \sqcup, T, S, \lambda, \rho, \{\})$ is a symmetric monoidal category. □

5.2.2 The matrix form of relations

We study the properties of relations of the form $R \in \mathbf{Rel}(U \sqcup V, W \sqcup X)$, and demonstrate a close connection with the algebra of matrices over a ring.

Proposition 5 *Any relation in* $\mathbf{Rel}(U \sqcup V, W \sqcup X)$ *determines four relations, in* $\mathbf{Rel}(U,W)$, $\mathbf{Rel}(U,X)$, $\mathbf{Rel}(V,W)$, $\mathbf{Rel}(V,X)$ *respectively; conversely, any four relations of this form determine a relation in* $\mathbf{Rel}(U \sqcup V, W \sqcup X)$.

Proof Given an arbitrary relation $R \in \mathbf{Rel}(U \sqcup V, W \sqcup X)$, then

$$R \subseteq (W \times \{0\} \cup X \times \{1\}) \times (U \times \{0\} \cup V \times \{1\}),$$

or, equivalently, by using the distributivity of $\times$ over $\cup$,

$$R \subseteq (W \times \{0\}) \times (U \times \{0\}) \cup (W \times \{0\}) \times (V \times \{1\})$$

$$\cup (X \times \{1\}) \times (U \times \{0\}) \cup (X \times \{1\}) \times (V \times \{1\}).$$

So the relation R can be split up into the (disjoint) union of four relations, $R = r_{00} \cup r_{10} \cup r_{01} \cup r_{11}$, where

$$r_{00} \in \mathbf{Rel}(U \times \{0\}, W \times \{0\}) \ , \ r_{10} \in \mathbf{Rel}(V \times \{1\}, W \times \{0\}),$$

$$r_{01} \in \mathbf{Rel}(U \times \{0\}, X \times \{1\}) \ , \ r_{11} \in \mathbf{Rel}(V \times \{1\}, X \times \{1\}).$$

However, there exists an isomorphism β_{00} from $\mathbf{Rel}(U \times \{0\}, W \times \{0\})$ to $\mathbf{Rel}(U, W)$, given by $((w,0),(u,0)) \in R$ iff $(w,u) \in \beta_{00}(R)$. Similarly, we can construct isomorphisms

$$\beta_{10} : \mathbf{Rel}(V \times \{1\}, W \times \{0\}) \to \mathbf{Rel}(V, W),$$

$$\beta_{01} : \mathbf{Rel}(U \times \{0\}, X \times \{1\}) \to \mathbf{Rel}(U, X),$$

$$\beta_{11} : \mathbf{Rel}(V \times \{1\}, X \times \{1\}) \to \mathbf{Rel}(V, X).$$

Hence, the relation R uniquely determines four relations $\beta_{00}(r_{00})$, $\beta_{10}(r_{10})$, $\beta_{01}(r_{01})$, and $\beta_{11}(r_{11})$ in $\mathbf{Rel}(U, W)$, $\mathbf{Rel}(V, W)$, $\mathbf{Rel}(U, X)$, and $\mathbf{Rel}(V, X)$ respectively. So, we can refer to the function β from $R \in \mathbf{Rel}(U \sqcup V, W \sqcup X)$ to the 4-tuple of relations $(\beta_{00}(r_{00}), \beta_{10}(r_{10}), \beta_{01}(r_{01}), \beta_{11}(r_{11}))$. Conversely, given four relations $a \in \mathbf{Rel}(U, W)$, $b \in \mathbf{Rel}(V, W)$, $c \in \mathbf{Rel}(U, X)$ and $d \in \mathbf{Rel}(V, X)$, we can define $r_{00} = \beta_{00}^{-1}(a)$, $r_{10} = \beta_{10}^{-1}(b)$, $r_{01} = \beta_{01}^{-1}(c)$ and $r_{11} = \beta_{11}^{-1}(d)$, and these satisfy

$$r_{00} \in \mathbf{Rel}(U \times \{0\}, W \times \{0\}) \ , \ r_{10} \in \mathbf{Rel}(V \times \{1\}, W \times \{0\}),$$

$$r_{01} \in \mathbf{Rel}(U \times \{0\}, X \times \{1\}) \ , \ r_{11} \in \mathbf{Rel}(V \times \{1\}, X \times \{1\}).$$

Therefore, we can define $R = r_{00} \cup r_{10} \cup r_{01} \cup r_{11}$, and say that $R = \beta^{-1}(a, b, c, d)$.

Hence, any relation in $\mathbf{Rel}(U \sqcup V, W \sqcup X)$ uniquely determines four relations in $\mathbf{Rel}(U, W)$, $\mathbf{Rel}(U, X)$, $\mathbf{Rel}(V, W)$, $\mathbf{Rel}(V, X)$, and any four relations of this form uniquely determine a relation in $\mathbf{Rel}(U \sqcup V, W \sqcup X)$. Finally, note that $\beta^{-1}\beta(R) = R$, and $\beta\beta^{-1}(a, b, c, d) = (a, b, c, d)$. □

We write the 4-tuples of relations in matrix form, so that $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is shorthand for $R = \beta^{-1}(a, b, c, d) \in \mathbf{Rel}(U \sqcup V, W \sqcup X)$, as defined above, and consider how composition is defined on 4-tuples of relations constructed in this way.

Theorem 6 *Given $R \in \mathbf{Rel}(U \sqcup V, W \sqcup X)$, $S \in \mathbf{Rel}(W \sqcup X, Y \sqcup Z)$, defined by*

$$R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ , \ S = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

then their composite, $SR \in \mathbf{Rel}(U \sqcup V, Y \sqcup Z)$ is given by

$$SR = \begin{pmatrix} ea \cup fc & eb \cup fd \\ ga \cup hc & gb \cup hd \end{pmatrix},$$

that is, the usual definition of matrix multiplication, with addition interpreted by union, and multiplication interpreted by the composition of relations.

Proof We have proved that, in **Rel**, composition distributes over union. So, given

$$R = r_{00} \cup r_{10} \cup r_{01} \cup r_{11} \ , \ S = s_{00} \cup s_{01} \cup s_{10} \cup s_{11},$$

where $R = \beta^{-1}(a,b,c,d)$, $S = \beta^{-1}(e,f,g,h)$, then $SR =$

$$s_{00}r_{00} \cup s_{00}r_{10} \cup s_{00}r_{01} \cup s_{00}r_{11} \cup$$

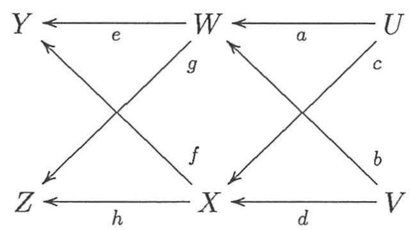
$$s_{01}r_{00} \cup s_{01}r_{10} \cup s_{01}r_{01} \cup s_{01}r_{11} \cup$$

$$s_{10}r_{00} \cup s_{10}r_{10} \cup s_{10}r_{01} \cup s_{10}r_{11} \cup$$

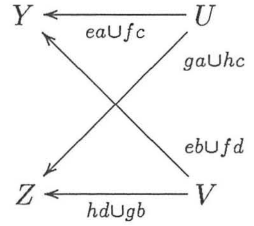
$$s_{11}r_{00} \cup s_{11}r_{10} \cup s_{11}r_{01} \cup s_{11}r_{11}.$$

However, the members of r_{00} and r_{10} are of the form $((w,0),q)$, whereas the members of s_{01} and s_{11} are of the form $(q,(m,1))$. Therefore, $s_{01}r_{00} = s_{11}r_{00} = \emptyset = s_{01}r_{10} = s_{11}r_{10}$. Conversely, members of r_{01} and r_{11} are of the form $((v,1),q)$, whereas member of s_{00} and s_{10} are of the form $(q,(m,0))$. Therefore, $s_{00}r_{01} = s_{00}r_{11} = \emptyset = s_{00}r_{11} = s_{10}r_{11}$. Hence the product is given by $SR = s_{00}r_{00} \cup s_{01}r_{01} \cup s_{00}r_{10} \cup s_{01}r_{11} \cup s_{10}r_{00} \cup s_{11}r_{01} \cup s_{10}r_{10} \cup s_{11}r_{11}$ and splitting this up into matrix form gives $SR = \begin{pmatrix} ea \cup fc & eb \cup fd \\ ga \cup hc & gb \cup hd \end{pmatrix}$, the required result. □

The relations between U, V, W, X, Y, Z, and their composition can be represented diagramatically, as follows:



becomes



So, intuitively, composition can be thought of as taking the union of all paths from U to Y, U to Z, V to Y, and V to Z.

5.3 The category of partial bijective maps

Definitions 5.5

We define the category **Inj** to have sets as objects, and partial bijective maps between sets as morphisms. Composition is the usual composition of partial bijective maps, and for any sets X, Y there is a zero map 0_{XY} satisfying $f0_{XY} = 0_{XA}$, and $0_{XY}g = 0_{BY}$ for all $f \in \mathbf{Inj}(Y, A)$, $g \in \mathbf{Inj}(B, X)$. The map 0_{XY} is the map from X to Y with empty domain and image.

Proposition 7 **Inj** *is an inverse category, and the monoid* $\mathbf{Inj}(X, X)$ *is the symmetric inverse monoid on* X.

Proof Given $f \in \mathbf{Inj}(X, Y)$, define $f^{-1} \in \mathbf{Inj}(Y, X)$ by

$$f^{-1}(y) = \begin{cases} \text{the unique } x \text{ satisfying } f(x) = y & \text{when defined,} \\ \text{undefined} & \text{otherwise.} \end{cases}$$

Then clearly $ff^{-1}f = f$, $f^{-1}ff^{-1} = f^{-1}$, and f^{-1} is the only morphism that will satisfy these two identities. Therefore **Inj** is an inverse category. Also, $\mathbf{Inj}(X, X) = I(X)$ follows trivially from the definition of the category of partial bijective maps and the definition of the symmetric inverse monoid on a set (Definitions 1.5, Chapter 1). □

Proposition 8 **Inj** *is isomorphic to a subcategory of* **Rel** *that has the same objects as* **Rel**.

Proof Define a functor Γ from **Inj** to **Rel** as follows: Γ is the identity map on objects, and given $f \in \mathbf{Inj}(X, Y)$, then $\Gamma(f) = \{(f(x), x) : x \in X, f(x) \text{ is defined}\}$. Clearly $\Gamma(f) \subseteq Y \times X$, and $\Gamma(g)\Gamma(f) = \Gamma(gf)$. Therefore Γ is an embedding of **Inj** into **Rel**. □

Convention We will refer to **Inj** as a subcategory of **Rel**, rather than referring to the existence of an embedding of **Inj** into **Rel**.

5.3.1 The monoidal structure of the category of partial bijective maps

Theorem 9 **Inj** *is a symmetric monoidal category.*

Proof Given $f \in \mathbf{Inj}(A, B)$ and $g \in \mathbf{Inj}(X, Y)$, then $f \sqcup g \in \mathbf{Rel}(A \sqcup X, A \sqcup Y)$ is also a partial bijective relation. It can be written as a function, as follows:

$$(f \sqcup g)(u) = \begin{cases} (f(a), 0) & u = (a, 0), \\ (g(x), 1) & u = (x, 1). \end{cases}$$

Hence **Inj** is closed under $\sqcup$. Also, for all sets X, Y, Z, the special elements $T_{X,Y,Z}$, $S_{X,Y}$, λ_X, ρ_X are isomorphisms, and hence are contained in the subcategory of partial bijective maps. Therefore, the proof that **Rel** is a monoidal category (Theorem 4) also suffices to show that **Inj** is a symmetric monoidal category. □

5.3.2 The matrix form of partial bijective maps

As **Inj** is a subcategory of **Rel**, any morphism of the form $F \in \mathbf{Inj}(U \sqcup V, W \sqcup X)$ can be written in matrix form as $F = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for some $a \in \mathbf{Inj}(U,W)$, $b \in \mathbf{Inj}(V,W)$, $c \in \mathbf{Inj}(U,X)$, $d \in \mathbf{Inj}(V,X)$. Note that composition of partial bijective maps written in this form interprets as matrix multiplication, as before; however, all unions considered must be disjoint, as all morphisms are partial bijections.

Theorem 10 *Given partial bijective maps $a \in \mathbf{Inj}(U,W)$, $b \in \mathbf{Inj}(V,W)$, $c \in \mathbf{Inj}(U,X)$ and $d \in \mathbf{Inj}(V,X)$, then a necessary and sufficient condition for the matrix $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ to represent a partial bijective map is that partial bijective maps in the same row of the matrix have disjoint images, and partial bijective maps in the same column of the matrix have disjoint domains.*

Proof

Explicitly, this condition can be written as

1. $dom(a) \cap dom(c) = \emptyset$
2. $dom(b) \cap dom(d) = \emptyset$
3. $im(a) \cap im(b) = \emptyset$
4. $im(c) \cap im(d) = \emptyset$

Our proofs are then as follows:

($\Rightarrow$) Assume $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{Inj}(U \sqcup V, W \sqcup X)$. Then

1. If we can find $u \in dom(a) \cap dom(c)$, then $a(u) \in W$ and $c(u) \in Y$, so R is not well-defined as a function. Therefore, we must have $dom(a) \cap dom(c) = \emptyset$
2. If we can find $v \in dom(b) \cap dom(d)$, then $b(v) \in W$, and $d(v) \in X$, so R is not well-defined as a function. Therefore, we must have $dom(b) \cap dom(d) = \emptyset$.

3. If we can find $w \in im(a) \cap im(b)$, then there must be some $u \in U$ and $v \in V$ satisfying $a(v) = w = b(x)$, so R is not a partial bijective function. Therefore, we must have $im(a) \cap im(b) = \emptyset$

4. If we can find $x \in im(c) \cap im(d)$, then there must be some $u \in U$ and $v \in V$ satisfying $c(u) = x = d(v)$, and so R is not a partial bijective function.

Therefore the above conditions are necessary for the map R to be a member of $\mathbf{Inj}(U \sqcup V, W \sqcup X)$.
($\Leftarrow$) It is clear that $U \times \{0\} \cap V \times \{1\} = \emptyset$ and $W \times \{0\} \cap X \times \{1\} = \emptyset$. Therefore conditions 1 – 4 are enough to ensure that $\beta^{-1}(a)$, $\beta^{-1}(b)$, $\beta^{-1}(c)$, and $\beta^{-1}(d)$, as defined in Proposition 5, all have disjoint domains and images. Hence $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{Inj}(U \sqcup V, W \sqcup X)$, and so the above conditions are also sufficient. □

A characterisation of conditions 1 – 4 above solely in terms of the partial bijective maps of **Inj** is given by $ac^{-1} = 0_{XW} = bd^{-1}$ and $a^{-1}b = 0_{VU} = c^{-1}d$.

5.4 The natural numbers as a self-similar object

Consider the category **Inj**. We demonstrate that the set of natural numbers (or indeed, any countable set; however, we use the natural numbers for clarity) is a self-similar object, with respect to both the coproduct $\sqcup$, which we have just proved in Theorem 4 is a monoidal fuctor, and the Cartesian product of sets. We then consider the M-monoid structures of $I(\mathbb{N})$, the endomorphism monoid of $\mathbb{N}$ in **Inj**, generated by internalising the coproduct and the Cartesian product, and show how these relate to embeddings of P_2 and P_∞ respectively.

5.5 Internalising the coproduct on the natural numbers

Lemma 11 $\mathbb{N}$ *is a strongly self-similar object of* $(\mathbf{Inj}, \sqcup)$

Proof Recall the definition of a strongly self-similar object, from Chapter 4, Definition 4.1. Hence we require a bijective map from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$. We have seen in Chapter 2, Lemma 2, that the existence of a bijective function from $\mathbb{N} \sqcup \mathbb{N}$ to $\mathbb{N}$ (and hence its inverse) is equivalent to the existence of a strong embedding of P_2 into $I(\mathbb{N})$. Therefore, the embedding of Lemma 2, Chapter 2 (generated by the 'interleaving map', $\phi(n, i) = 2n + i$, for $n \in \mathbb{N}$, $i \in \{0, 1\}$) is enough to demonstrate that $\mathbb{N}$ is a strongly self-similar object of **Inj**. □

We can deduce from the above that the endomorphism monoid of $\mathbb{N}$ has an M-monoid structure, with internal inclusions / projections. We will give an explicit description, in terms of the embedding of P_2, of the internalised coproduct, and commutativity and symmetry elements. In what follows, we will identify the generators of P_2 with their images under this embedding, unless the distinction is important.

We will denote the internalisation of the coproduct by $\oplus$, and its internal symmetry and associativity morphisms by s and t respectively.

Theorem 12

(i) *The internalisation of* $\sqcup$ *in* $I(\mathbb{N})$ *is given by* $a \oplus b = p^{-1}ap \vee q^{-1}bq$.

(ii) *The commutativity morphism is given by* $s = p^{-1}q \vee q^{-1}p$, *and the associativity morphism is given by* $t = p^{-2}p \vee p^{-1}q^{-1}pq \vee q^{-1}q^2$

(iii) $(I(\mathbb{N}), \oplus)$ *has internal projections / inclusions.*

Proof

(i) Recall the definition of the internalisation of a monoidal functor from Definitions 4.5, Chapter 4, as $a \oplus b = c(a \otimes b)d$, where c is the contraction morphism in $\mathbf{M}(N \otimes N, N)$, and d is the division morphism in $\mathbf{M}(N, N \otimes N)$. In this case, the contraction morphism is the bijection ϕ, and the division morphism is its inverse ϕ^{-1}. Hence, the internalised monoidal functor is given by $(a \oplus b) = \phi(a \sqcup b)\phi^{-1}$. However, by the construction of a strong embedding of P_2 into $I(\mathbb{N})$ from a bijection $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$, (Lemma 3, Chapter 2), $p^{-1}(n) = \phi(n, 0)$, $q^{-1}(n) = \phi(n, 1)$, and similarly, $\phi^{-1}(n) = p(n) \sqcup q(n)$. (Note that this map is bijective, as p and q have disjoint domains). Therefore, we can conclude that $\phi(a \sqcup b)\phi^{-1} = p^{-1}ap \vee q^{-1}bq$. Hence, the internalisation of $\sqcup$ in $I(\mathbb{N})$ is given by $\oplus$, as defined above.

(ii) By definition, $s^2 = p^{-1}p \vee q^{-1}q = 1$, so s is a self-inverse element of $I(\mathbb{N})$, and

$$s(f \oplus g) = (q^{-1}p \vee p^{-1}q)(p^{-1}fp \vee q^{-1}gq) = q^{-1}fp \vee p^{-1}gq = (g \oplus f)s.$$

Next, by definition of t and $\oplus$,

$$t(a \oplus (b \oplus c)) = t(p^{-1}ap \vee q^{-1}(p^{-1}bp \vee q^{-1}cq)q)$$

$$= p^{-2}ap \vee p^{-1}q^{-1}bpq \vee q^{-1}cq^2 = ((a \oplus b) \oplus c)t.$$

Also, $tt^{-1} = p^{-1}p \vee q^{-1}q = t^{-1}t$, and since the embedding of P_2 into $I(\mathbb{N})$ is strong, $t^{-1}t = 1 = tt^{-1}$. Finally, $tst = p^{-1}q^2 \vee p^{-1}q^{-1}p \vee q^{-1}pq$ and $(s \oplus 1)t(1 \oplus s) = (p^{-1}sp \vee q^{-1}q)t(p^{-1}p \vee q^{-1}sq) =$

$p^{-1}q^2 \vee p^{-1}q^{-1}p \vee q^{-1}pq$. Therefore $tst = (s \oplus 1)t(1 \oplus s)$. Hence t and s satisfy the conditions for the associativity and commutativity elements of the strong M-monoid, $(I(\mathbb{N}), \oplus)$.

(iii) This is immediate from the definition of $\oplus$ in terms of the embedding of P_2. □

Corollary 13 *The disjoint closure of P_2 in $I(\mathbb{N})$ is a strong M-monoid with internal projections and inclusions.*

Proof Note that the internal tensor product, the associativity and commutativity elements, and internal projections / inclusions are all defined in $DC_{\mathbb{N}}(P_2)$. Therefore our result follows. □

Corollary 14 *Let a morphism $F \in I(\mathbb{N} \sqcup \mathbb{N})$ be represented by the matrix $F = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$. Then the image of F under the isomorphism from $I(\mathbb{N} \sqcup \mathbb{N})$ to $I(\mathbb{N})$ is given by*

$$\phi(F) = p^{-1}rp \vee p^{-1}sq \vee q^{-1}tp \vee q^{-1}uq.$$

Proof Immediate from the definition of the map $\phi : \mathbb{N} \sqcup \mathbb{N} \to \mathbb{N}$, and its inverse, $\phi^{-1} : \mathbb{N} \to \mathbb{N} \sqcup \mathbb{N}$ in terms of the embedding of P_2 into $I(\mathbb{N})$. □

5.6 Internalising the Cartesian product on the natural numbers

We demonstrate that $\mathbb{N}$ is a strongly self-similar object of $(\mathbf{Inj}, \times)$, and use this to construct an alternative M-monoid structure on $I(\mathbb{N})$.

Lemma 15 *$\mathbb{N}$ is a strongly self-similar object of $(\mathbf{Inj}, \times)$.*

Proof Recall from Lemma 4, Chapter 2 that a strong embedding of P_∞ into $I(\mathbb{N})$ uniquely determines, and is determined by, a bijection $[\ ,\ ]$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. So we can take $[\ ,\ ]$ to be the bijection derived from applying the construction of Lemma 7, Chapter 2, to the interleaving embedding of P_2, found in Definitions 2.2, Chapter 2. Then, as $[\ ,\ ]$ is a bijection, it has a global inverse, and so $c = [\ ,\ ] \in Inj(\mathbb{N} \times \mathbb{N}, \mathbb{N})$ and $d = [\ ,\ ]^{-1} \in Inj(\mathbb{N}, \mathbb{N} \times \mathbb{N})$ satisfy the conditions for division and contraction maps of a strongly self-similar object. □

We identify the generators of P_∞ with their images under this embedding, unless the distinction is important.

Proposition 16 *The internalisation of $\times$ in $I(\mathbb{N})$, which we denote by $\otimes$, is given in terms of the embedding of P_∞ by*

$$(u \otimes v) = \bigvee_{i=0}^{\infty} p_{u(i)}^{-1} v p_i.$$

Proof We first prove that this disjoint join is well-defined. Note that the $\{p_i\}_{i=0}^{\infty}$ have disjoint domains, so the $\{p_i^{-1}\}_{i=0}^{\infty}$ have disjoint images. Hence the elements $\{p_{u(i)}^{-1}\}_{i=0}^{\infty}$ have disjoint images, as u is a partial bijective map. Therefore, the infinite join given above is well-defined as a partial bijective map on $\mathbb{N}$, as the individual terms in the join have disjoint domains / images.

Next, note that $[a, b]$ can be written in terms of the embedding of P_∞, by $[a, b] = p_a^{-1}(b)$, so we just need to check that $(u \otimes v)(p_a^{-1}(b)) = p_{u(a)}^{-1}v(b)$, for arbitrary $a, b \in \mathbb{N}$. We can expand the left hand side of this as

$$(u \otimes v)p_a^{-1}(b) = \left(\bigvee_{i=0}^{\infty} p_{u(i)}^{-1} v p_i\right)(p_a^{-1}(b)) = \left(\bigvee_{i=0}^{\infty} p_{u(i)}^{-1} v \delta_{i,a}\right)(b) = p_{u(a)}^{-1}v(b).$$

Therefore, as a and b were arbitrary natural numbers, $(u \otimes v)([a, b]) = [u(a), v(b)]$ for all $a, b \in \mathbb{N}$. □

It is then immediate that $(I(\mathbb{N}), \otimes)$ is a strong M–monoid (that is, a one-object symmetric monoidal category without units). We consider the associativity and commutativity elements.

Proposition 17 *The element τ of $I(\mathbb{N})$, defined by*

$$\tau = \bigvee_{i=0}^{\infty} (p_i^{-1} \otimes 1)p_i,$$

satisfies $\tau([a, [b, c]]) = [[a, b], c]$ and $\tau^{-1}\tau = \tau\tau^{-1} = 1$.

Proof We first show that the above disjoint join is well-defined. The definition of $\otimes$ gives us

$$(p_i^{-1} \otimes 1) = \left(\bigvee_{j=0}^{\infty} p_{p_i^{-1}(j)}^{-1} p_j\right),$$

and so

$$\tau = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p_{p_i^{-1}(j)}^{-1} p_j p_i.$$

Note that $p_j p_i p_{i'}^{-1} p_{j'}^{-1} = \delta_{jj'}\delta_{ii'}\delta_{jj'}$ by the axioms for P_∞. Hence, $dom(p_j p_i) \cap dom(p_{j'} p_i) = \emptyset$ when $i \neq i'$, $j \neq j'$. Therefore, the separate terms in the double join have disjoint domains. To see that they also have disjoint images, note that $p_i^{-1}(j) \neq p_{i'}^{-1}(j')$ for all $i \neq i'$, $j \neq j'$. Hence $im(p_{p_i^{-1}(j)}) \cap im(p_{p_{i'}^{-1}(j')}) = \emptyset$ for all $i \neq i'$, $j \neq j'$. Therefore, the individual terms in this join have disjoint domains / images, and so the join is well-defined.

To prove that τ has full image, note that

$$\tau\tau^{-1} = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} (p_i^{-1} \otimes 1)p_i p_j^{-1}(p_j^{-1} \otimes 1)$$

and by the composition of generators of P_∞

$$\tau\tau^{-1} = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty}(p_i^{-1}\otimes 1)\delta_{ij}(p_j^{-1}\otimes 1) = \bigvee_{i=0}^{\infty}(p_i^{-1}\otimes 1)(p_i^{-1}\otimes 1),$$

and since $\otimes$ is a monoid homomorphism, $\tau\tau^{-1} = \bigvee_{i=0}^{\infty}(p_i^{-1}p_i\otimes 1)$. By definition of $\otimes$, we can write this as

$$\tau\tau^{-1} = \bigvee_{i=0}^{\infty}\bigvee_{k=0}^{\infty} p^{-1}_{p_i^{-1}p_i(k)}p_k,$$

and as the embedding of P_∞ into $I(\mathbb{N})$ is strong, for eack $k\in\mathbb{N}$ there exists a unique $i\in\mathbb{N}$ satisfying $p_i^{-1}p_i(k)=k$, and $p_j^{-1}p_j(k)$ is undefined for $j\neq i$. Therefore,

$$\tau\tau^{-1} = \bigvee_{k=0}^{\infty} p_k^{-1}p_k = 1,$$

as the embedding of P_∞ into $I(\mathbb{N})$ is strong. Therefore, τ has full image.

To prove that τ has full domain, note that

$$\tau^{-1}\tau = \bigvee_{j=0}^{\infty}\bigvee_{i=0}^{\infty} p_j^{-1}(p_j\otimes 1)(p_i^{-1}\otimes 1)p_i$$

and as $\otimes$ is a monoid homomorphism,

$$\tau^{-1}\tau = \bigvee_{j=0}^{\infty}\bigvee_{i=0}^{\infty} p_j^{-1}(\delta_{ij}\otimes 1)p_i = \bigvee_{j=0}^{\infty}\bigvee_{i=0}^{\infty} p_j^{-1}\delta_{ij}p_i = \bigvee_{i=0}^{\infty} p_i^{-1}p_i = 1$$

as $\otimes$ is a monoid homomorphism. Therefore, τ has full domain.

Finally, we show that τ has the given action on $\mathbb{N}$. First note that for all $a,b,c\in\mathbb{N}$,

$$[a,[b,c]] = p_a^{-1}(p_b^{-1}(c))\ ,\ [[a,b],c] = p^{-1}_{p_a^{-1}(b)}(c).$$

So, we require $\tau(p_a^{-1}(p_b^{-1}(c))) = p^{-1}_{p_a^{-1}(b)}(c)$. However,

$$\tau p_a^{-1}p_b^{-1}(c) = \left(\bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p^{-1}_{p_i^{-1}(j)}p_jp_i\right) p_a^{-1}p_b^{-1}(c)$$

$$= \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p^{-1}_{p_i^{-1}(j)}p_j\delta_{ia}p_b^{-1}(c) = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{\infty} p^{-1}_{p_i^{-1}(j)}\delta_{bj}\delta_{ia}(c) = p^{-1}_{p_a^{-1}(b)}(c).$$

Hence, as a,b,c were arbitrary, τ has the required action on $\mathbb{N}$. □

Corollary 18 *Let $\tau\in I(\mathbb{N})$ be as defined above. Then*

(i) $\tau(u\otimes(v\otimes w)) = ((u\otimes v)\otimes w)\tau$ *for arbitrary $u,v,w\in I(\mathbb{N})$.*

(ii) $\tau^2 = (\tau\otimes 1)\tau(1\otimes\tau)$.

Proof

(i) Consider arbitrary $n \in \mathbb{N}$, and write it uniquely as $n = [a, [b, c]]$. Then

$$\tau(u \otimes (v \otimes w))(n) = \tau(u \otimes (v \otimes w))([a, [b, c]]) = \tau[u(a), [v(b), w(c)]] = [[u(a), v(b)], w(c)].$$

Conversely,

$$((u \otimes v) \otimes w)\tau(n) = ((u \otimes v) \otimes w)\tau([a, [b, c]])$$

$$= ((u \otimes v) \otimes w)[[a, b], c] = [[u(a), v(b)], w(c)] = \tau(u \otimes (v \otimes w))(n).$$

Hence $\tau(u \otimes (v \otimes w)) = ((u \otimes v) \otimes w)\tau$.

(ii) Consider arbitrary $n = [a, [b, [c, d]]] \in \mathbb{N}$. Then $\tau^2(n) = \tau([[a, b], [c, d]]) = [[[a, b], c], d]$ and

$$(\tau \otimes 1)\tau(1 \otimes \tau)([a, [b, [c, d]]]) = (\tau \otimes 1)\tau([a, [[b, c], d]])$$

$$= (\tau \otimes 1)([[a, [b, c]], d]) = [[[a, b], c], d].$$

Therefore, as n was chosen arbitrarily, $\tau^2 = (\tau \otimes 1)\tau(1 \otimes \tau)$. □

Proposition 19 *There is a unique map $\sigma \in I(\mathbb{N})$ that satisfies $\sigma([a, b]) = [b, a]$. Moreover, this map is not a member of $DC_{\mathbb{N}}(P_\infty)$.*

Proof As the map $\psi = [\ ,\ ]$ determined by the embedding of P_∞ is a bijection from $\mathbb{N}^2$ to $\mathbb{N}$, and the map $S_{\mathbb{N},\mathbb{N}} : (a, b) \mapsto (b, a)$ of $I(\mathbb{N} \times \mathbb{N})$ is a bijection, we can define the map $\sigma = [\ ,\ ]S_{\mathbb{N},\mathbb{N}}[\ ,\ ]^{-1} : \mathbb{N} \to \mathbb{N}$, and from the definition, we can see that $\sigma([a, b]) = [b, a]$ for all $a, b \in \mathbb{N}$.

To prove that σ is not a member of the disjoint closure of P_∞ in $I(\mathbb{N})$, we use a topological argument. Recall the construction of a topology on a set derived from an effective representation of an inverse monoid as partial bijections (Theorem 9, Chapter 2). We apply this to our embedding of P_∞ into $I(\mathbb{N})$ to deduce that $T = \{w(\mathbb{N}) : w \in \{p_i^{-1} : 0 \le i < \infty\}^*\} \cup \{\emptyset\}$ forms the basis of a topology on $\mathbb{N}$ in which all the elements of P_∞ are continuous.

Consider $X = x(\mathbb{N})$ and $Y = y(\mathbb{N})$ in T. Then if x is a prefix of y, $X \cap Y = Y$. Conversely, if y is a prefix of x, then $X \cap Y = X$. Otherwise, since the p_i^{-1} have full domains and disjoint images, $X \cap Y = \emptyset$. Hence, T is closed under finite intersections, and so does indeed form the basis of a topology on $\mathbb{N}$. Also, the members of $DC(P_\infty)$ map open sets to open sets, since for an arbitrary set of pairwise disjoint elements $\{a_i : i \in I\}$, and an arbitrary basic open set X,

$$\left(\bigvee_{i \in I} a_i\right)(X) = \left(\bigvee_{i \in I} a_i x\right)(\mathbb{N}) = \bigcup_{i \in I} a_i x(\mathbb{N}).$$

Hence, the disjoint join of pairwise disjoint elements maps open sets to open sets, and so is continuous. Therefore, T is the basis for a topology on $\mathbb{N}$ in which all members of $DC(P_\infty)$ are continuous.

Recall that the action of σ on $\mathbb{N}$ is $\sigma([a,b]) = [b,a]$. So, in terms of the generators of the embedding of P_∞, $\sigma p_a^{-1}(b) = p_b^{-1}(a)$. Consider the action of σ on an open set of $\mathbb{N}$. For arbitrary $a \in \mathbb{N}$, $p_a^{-1}(\mathbb{N}) = \{[a,n] : n \in \mathbb{N}\}$ is open, by definition. Also, $\sigma p_a^{-1}(\mathbb{N}) = \{[n,a] : n \in \mathbb{N}\}$, and so $\sigma p_a^{-1}(\mathbb{N}) = \{p_n^{-1}(a) : n \in \mathbb{N}\}$. However, consider the intersection of $\sigma p_a^{-1}(\mathbb{N})$ with some basic open set $p_x^{-1}(\mathbb{N})$. Then $\{p_n^{-1}(a) : n \in \mathbb{N}\} \cap \{p_x^{-1}(m) : m \in \mathbb{N}\} = \{p_x^{-1}(a)\}$, and this is a singleton element of $\mathbb{N}$. However, any open set in the above topology is infinite (or empty), so we can deduce that σ is not a continuous map in this topology. Therefore, σ is not a member of $DC(P_\infty)$. □

Corollary 20 *σ satisfies $\sigma(u \otimes v) = (v \otimes u)\sigma$, for all $u, v \in I(\mathbb{N})$.*

Proof For arbitrary $n = [a,b] \in \mathbb{N}$,

$$\sigma(u \otimes v)(n) = \sigma(u \otimes v)([a,b]) = \sigma([u(a), v(b)]) = [v(b), u(a)]$$

However, $(v \otimes u)\sigma([a,b]) = (v \otimes u)([b,a]) = [v(b), u(a)]$, and so $\sigma(u \otimes v) = (v \otimes u)\sigma$, for all $u, v \in I(\mathbb{N})$. □

Corollary 21 *σ and τ satisfy the condition $\tau\sigma\tau = (\sigma \otimes 1)\tau(1 \otimes \sigma)$.*

Proof For arbitrary $n = [a, [b,c]] \in \mathbb{N}$,

$$\tau\sigma\tau(n) = \tau\sigma([[a,b],c,]) = \tau([c,[a,b]]) = [[c,a],b],$$

and

$$(\sigma \otimes 1)\tau(1 \otimes \sigma)(n) = (\sigma \otimes 1)\tau([a,[c,b]] = (\sigma \otimes 1)([[a,c],b] = [[c,a],b].$$

Hence, as n was chosen arbitrarily, $\tau\sigma\tau = (\sigma \otimes 1)\tau(1 \otimes \sigma)$. □

Theorem 22 *The elements τ and σ of $I(\mathbb{N})$, as defined above, are associativity and commutativity elements of the M-monoid $(I(\mathbb{N}), \otimes)$.*

Proof Recall the definition of an M-monoid from definitions 4.6 of Chapter 4. Then τ satisfies conditions 1 and 3, by Corollary 18, σ satisfies condition 2 by Corollary 20, and σ and τ satisfy condition 4 by Corollary 21. □

Proposition 23 *The M-monoid $(I(\mathbb{N}), \otimes)$ has internal projections, as defined in Definitions 4.9, Chapter 4.*

Proof For arbitrary fixed $i \in \mathbb{N}$,

$$p_i(1 \otimes g)p_i^{-1} = p_i \left(\bigvee_{j=0}^{\infty} p_j^{-1} g p_j \right) p_i^{-1} = g,$$

and $p_i\sigma(f \otimes 1)\sigma^{-1}p_i^{-1} = p_i(1 \otimes f)p_i^{-1} = f$. Therefore, we have constructed maps that satisfy the conditions for internal projections. □

5.7 Properties of the internalisation of the Cartesian product

We consider some algebraic properties of the $\otimes$ homomorphism.

Definitions 5.6

Let $[\ ,\ ]$ and $\otimes$ be as defined in Lemma 15 and Proposition 16 respectively. We define two maps from $I(\mathbb{N})$ to $I(\mathbb{N})$ by $?(u) = (u \otimes 1)$ and $!(v) = (1 \otimes v)$.

Proposition 24

(i) *! and ? are monoid homomorphisms that commute with each other.*

(ii) *The map ? is given explicitly by*

$$?(u) = \bigvee_{i=0}^{\infty} p_{u(i)}^{-1} p_i.$$

and is an embedding of $I(\mathbb{N})$ into $DC_{\mathbb{N}}(P_\infty)$.

(iii) *The map ! is given explicitly by*

$$!(v) = \bigvee_{i=0}^{\infty} p_i^{-1} v p_i,$$

and hence satisfies $p_n!(v)p_n^{-1} = v$, for all $n \in \mathbb{N}$.

Proof

(i) As $\otimes$ is a monoid homomorphism, it is immediate that the maps given by $u \mapsto (u \otimes 1)$ and $v \mapsto (1 \otimes v)$ are monoid homomorphisms. Also, for arbitrary $n \in \mathbb{N}$, $(a \otimes b)(n) = (a \otimes b)[x, y]$, for some unique $x, y \in \mathbb{N}$, and so $[a(x), b(y)] = (1 \otimes b)[a(x), y] = (a \otimes 1)[x, b(y)]$. Hence, as n was arbitrary, our result follows.

(ii) By Proposition 16, the definition of ? in terms of the generators of P_∞ is

$$?(u) = (u \otimes 1) = \bigvee_{i=0}^{\infty} p_{u(i)}^{-1} p_i.$$

Note that u is an arbitrary member of $I(\mathbb{N})$, but $(u \otimes 1)$ is always a member of $DC(P_\infty)$. Also we have already seen that ? is a monoid homomorphism. Finally, this map is injective, since

$$?(u) = ?(v) \Leftrightarrow \bigvee_{i=0}^{\infty} p_{u(i)}^{-1} p_i = \bigvee_{i=0}^{\infty} p_{v(i)}^{-1} p_i,$$

and multiplying on the right hand side by some fixed p_j^{-1} will give us $p_{u(j)}^{-1} = p_{v(j)}^{-1}$, which is equivalent to $u(j) = v(j)$, and since j was chosen arbitrarily, we can deduce that $u = v$. Hence ?

is an injective homomorphism from $I(\mathbb{N})$ to $DC_{\mathbb{N}}(P_\infty)$.

(iii) By definition of $\otimes$ (Proposition 16),

$$!(v) = (1 \otimes v) = \bigvee_{i=0}^{\infty} p_i^{-1} v p_i.$$

Therefore, for arbitrary $n \in \mathbb{N}$,

$$p_n !(v) p_n^{-1} = p_n \left(\bigvee_{i=0}^{\infty} p_i^{-1} v p_i \right) p_n^{-1} = \bigvee_{i=0}^{\infty} \delta_{n,i} v \delta_{n,i} = v.$$

Hence our result follows. □

We can consider the ! homomorphism to be 'constructing an infinite number of copies of the action of a function', since $p_i!(a)p_i^{-1} = a$ for all $i \in \mathbb{N}$, and $dom(p_i) \cap dom(p_j) = \emptyset$ for all $i \neq j$, and so there are an infinite number of disjoint internal projections.

Conversely, the ? operator takes a map u on the natural numbers, and 'lifts' it to a map that acts on the (set-theoretic) domains of the $\{p_i\}$ by $(u \otimes 1)(dom(p_i)) = dom(p_{u(i)})$.

5.7.1 Constructing fixed points for the coproduct using the Cartesian product

In the following, we take an embedding of P_2 into $I(\mathbb{N})$ that satisfies the 'no fixed point' condition for the construction of a strong embedding of P_∞ into $I(\mathbb{N})$, found in Lemma 6 of Chapter 2.

We show that the ! homomorphism can be thought of as an infinitary form of the direct sum construction, and deduce from this that ! is a right fixed point homomorphism for the M-monoid $(I(\mathbb{N}), \oplus)$.

Theorem 25 *Let the internal tensor $\otimes$ be defined in terms of an embedding of P_∞ into $I(\mathbb{N})$, derived from the right-associative embedding of P_∞ into P_2 (as given in Definitions 2.3, Chapter 2). Then the homomorphism ! is a right fixed point homomorphism (as given in Definitions 4.10, Chapter 4) for the M-monoid $(I(\mathbb{N}), \oplus)$.*

Proof We have already seen that !, defined by $!(f) = (1 \otimes f)$ is an injective monoid homomorphism. Also, as we are using the right-associative embedding of P_∞ into P_2,

$$f \oplus !(f) = p^{-1} f p \vee q^{-1} \left(\bigvee_{i=0}^{\infty} p_i^{-1} f p_i \right) q,$$

and as $p_i^{-1} = q^{-i} p^{-1}$,

$$f \oplus !(f) = p^{-1} f p \vee \left(\bigvee_{i=0}^{\infty} q^{-1} p_i^{-1} f p_i q \right) = p^{-1} f p \vee \bigvee_{i=1}^{\infty} p_i^{-1} f p_i = \bigvee_{i=0}^{\infty} p_i^{-1} f p_i = (1 \otimes f) = !(f).$$

Therefore $f\oplus!(f)=!(f)$, and so $!: I(\mathbb{N}) \to I(\mathbb{N})$ is a right fixed point homomorphism for $(I(\mathbb{N}),\oplus)$.
□

We also have the following identity connecting ? and $\oplus$.

Proposition 26 *Given* $f, a_0, a_1, a_2, \ldots, \in I(\mathbb{N})$, *then*

$$?(f)^{-1}(a_0 \oplus a_1 \oplus a_2 \ldots \oplus a_n \oplus 0)?(f) = (a_{f(0)} \oplus a_{f(1)} \oplus \ldots \oplus a_{f(n)} \oplus 0),$$

where

$$a_{f(x)} = \begin{cases} a_{f(x)} & x \in dom(f) \ , \ f(x) \leq n \\ 0 & otherwise. \end{cases}$$

Proof By definition, $(f \otimes 1)^{-1}(a_0 \oplus a_1 \oplus a_2 \ldots \oplus a_n \oplus 0)(f \otimes 1) =$

$$\left(\bigvee_{i=0}^{\infty} p_i^{-1} p_{f(i)}\right)\left(\bigvee_{j=0}^{n} p_j^{-1} a_j p_j\right)\left(\bigvee_{k=0}^{\infty} p_{f(k)}^{-1} p_k\right) = \bigvee_{i=0}^{\infty}\bigvee_{j=0}^{n}\bigvee_{k=0}^{\infty} p_i^{-1}\delta_{f(i),j} a_j \delta_{j,f(k)} p_k$$

$$= \bigvee_{i=0}^{\infty} p_i^{-1} a_{f(i)} p_i = (a_{f(0)} \oplus a_{f(1)} \oplus \ldots \oplus a_{f(n)} \oplus 0).$$

Hence our result follows. □

Chapter 6

The categorical trace, and compact closed categories

6.1 Introduction

The stated aim of this thesis was to understand the algebra and category theory behind the geometry of interaction series of papers. In the previous 5 chapters, we have developed concepts that we claim will be enough to model the logical operations, as they appear in [20, 21]. We next present what we claim will be the correct model of the dynamical part of the system. This is the categorical trace on symmetric monoidal categories, as presented in [35], and the related concept of compact closed categories.

In particular, we consider the categorical trace on the category of relations, and demonstrate how the category of partial bijective maps is also closed under this operation. We then show how the trace can be internalised in the same way as the monoidal structure, as demonstrated in Chapter 4 and use this to motivate the definition of a traced M-monoid, which is a one-object traced symmetric monoidal category (without units) when the M-monoid structure is strong.

We present results of [35] on the connection between traced monoidal categories and compact closed categories, give details of the construction of a compact closed category from the category of relations, and demonstrate how it can be restricted to the subcategory of partial bijective maps between sets.

Self-similarity considerations, and an alternative characterisation of compact closed categories, are used to define compact closed M-monoids, and hence one-object compact closed categories without units. Finally, we demonstrate how the results of Chapter 4 and 5 can be used to construct self-similar objects of the category **Rel**, and use this to construct a one-object compact closed inverse monoid, which we describe explicitly.

Applications to J-Y Girard's Geometry of Interaction 1 system [20], resolution and unification over a term language and the Geometry of Interaction 3 system [13, 22], and two-way automata [3], will be given in chapters 8, 9 and 10 respectively.

We will use 'diagrammatic reasoning', as introduced by A. Joyal, and R. Street in [33], and formally justified in [33, 34], throughout this chapter, as an illustration of the underlying processes; however, all original deductions will be justified algebraically.

6.2 The categorical trace

We present the theory of traced monoidal categories. This is due to A. Joyal, R. Street, and D. Verity, [35]; however, we consider the case of symmetric monoidal categories, rather than the full theory of balanced monoidal categories given in this paper.

Definitions 6.1

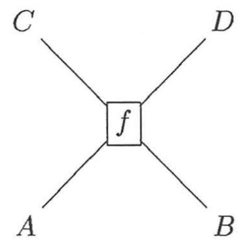
Let $(V, \otimes, s, t, \lambda, \rho, I)$ be a symmetric monoidal category. A *trace* on it is defined in [35] to be a family of functions, $Tr^U_{A,B} : V(A \otimes U, B \otimes U) \to V(A, B)$, that are natural in X, Y, U, and satisfy the following:

1. Given $f : X \otimes I \to Y \otimes I$, then $Tr^I_{X,Y}(f) = \rho f \rho^{-1} : X \to Y$.
2. Given $f : A \otimes (U \otimes V) \to B \otimes (U \otimes V)$, then
$$Tr^{U \otimes V}_{A,B}(f) = Tr^U_{A,B}(Tr^V_{A\otimes U,B\otimes U}(t_{B,U,V} f t^{-1}_{A,U,V})).$$
3. Given $f : A \otimes U \to B \otimes U$, and $g : C \to D$, then
$$Tr^U_{A,B}(f) \otimes g = Tr^U_{A\otimes C,B\otimes D}(t_{BDU}(1_B \otimes s_{D,U})t^{-1}_{BUD}(f \otimes g)t_{AUC}(1_A \otimes s_{C,U})t^{-1}_{ACU})$$
4. $Tr^U_{U,U}(s_{U,U}) = 1_U$.

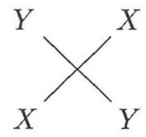
A symmetric monoidal category that has a trace is called a *traced symmetric monoidal category*.

6.2.1 Diagrammatic reasoning and the categorical trace

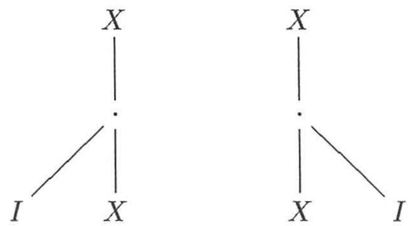
In diagrammatic reasoning, as introduced and justified by A. Joyal and R. Street in [33, 34], a morphism $f : A \otimes B \to C \otimes D$ is represented by



the commutativity morphism s_{XY} is represented by

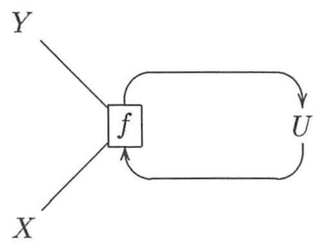


the left and right unit morphisms λ_X , ρ_X are represented by



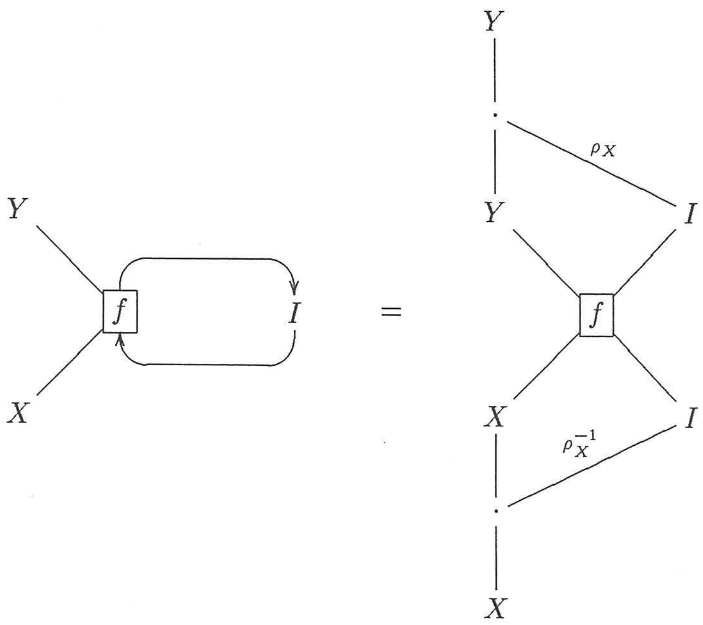
respectively. Note that there is no way to represent the associativity isomorphisms with diagrammatic reasoning.

The trace $Tr^U_{X,Y}$ on a morphism $f : X \otimes U \to Y \otimes U$ is then represented by

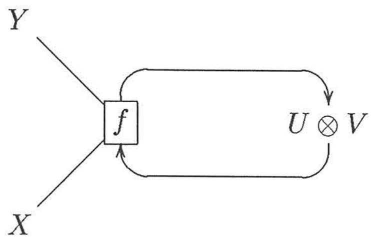


Using the above conventions, the axioms for a categorical trace are then represented as follows:

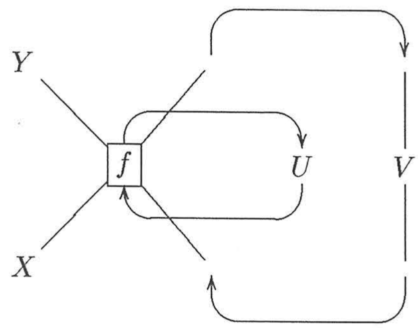
1.



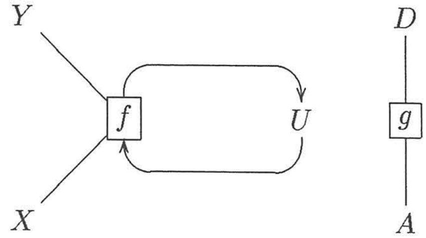
2.



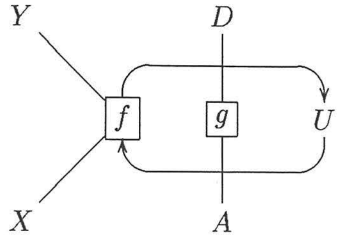
is the same as



3.

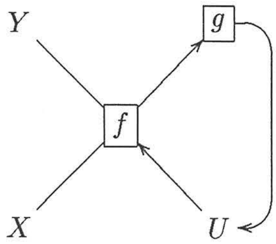


is the same as

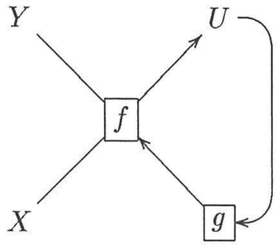


Finally, the naturality of $Tr^U_{X,Y}$ in all three variables is represented as follows:

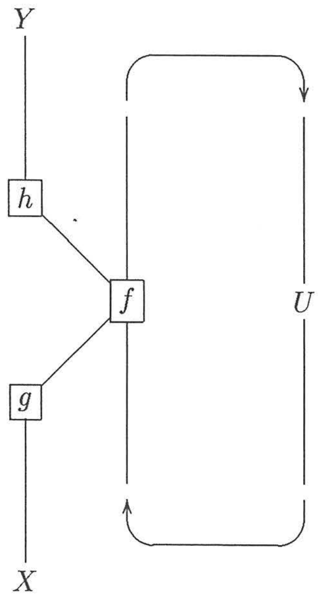
Naturality in U,



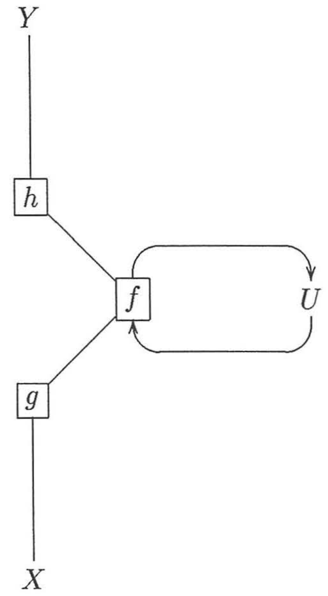
is the same as



Naturality in X and Y,



is the same as



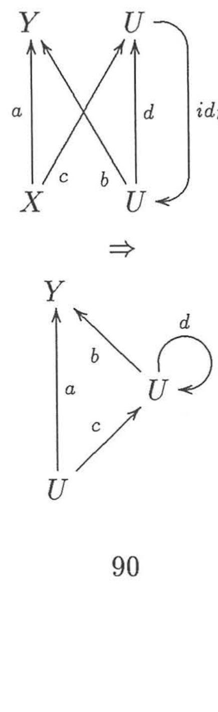
6.2.2 The trace on the category of relations

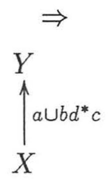
Theorem 1 *(Due to A. Joyal, R. Street, D. Verity). The category of relations,* **Rel**, *is a traced symmetric monoidal category, with trace defined as follows:*

$$Tr^U_{X,Y}(R) = a \cup bd^*c \ , \ \text{where } R = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{Rel}(X \sqcup U, Y \sqcup U).$$

Proof This is proved in [35] by A. Joyal, R. Street, and D. Verity — It is derived from elementary properties of the Kleene star on monoids of relations. □

Note that, with the intuitive idea of matrix multiplication as 'finding all possible paths through a graph', as presented in Chapter 5, Section 5.2.2, we can represent the trace in the same way, as follows:





6.2.3 The trace on the category of partial bijective maps

Recall the subcategory of the category of relations, given by the partial bijective maps, as defined in Definitions 5.5, Chapter 5, denoted **Inj**. We restrict Theorem 1 to the category of partial injective maps, and demonstrate that this subcategory is also traced.

Theorem 2 **Inj** *is closed under the trace operation inherited from the category of relations, and hence is a traced symmetric monoidal subcategory of* **Rel**.

Proof Consider $R \in \mathbf{Inj}(X \sqcup U, Y \sqcup U)$, represented by $R = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. The conditions on a, b, c, d for this matrix to represent a partial bijective map are given in Theorem 10 of Chapter 5. We seek to prove that these conditions imply that the trace $Tr^U_{X,Y} = a \cup bd^*c$ is also a partial bijective map. Hence we need to prove that the distinct terms in this union have disjoint domains / images (i.e. are disjoint).

First note that $a(bd^nc)^{-1} = ac^{-1}d^{-n}b^{-1} = 0$ for all $n \in \mathbb{N}$, since $ac^{-1} = 0$, by condition 1 of Theorem 10, Chapter 5. Also $a^{-1}bd^nc = 0$, since $a^{-1}b = 0$, by condition 3 of Theorem 10, Chapter 5. Therefore, $a \perp bd^nc$ for all $n \in \mathbb{N}$.

Secondly, $(bd^sc)(bd^tc)^{-1} = bd^scc^{-1}d^{-t}b^{-1} = bd^sd^{-t}eb^{-1}$, where $e = d^tcc^{-1}d^{-t}$. Therefore, if $t > s$, and so $s - t = -r$ for some $r \geq 1$, then $(bd^sc)(bd^tc)^{-1} = bd^{-r}eb^{-1} = 0$ since $bd^{-1} = 0$ by condition 2 of Theorem 10, Chapter 5. Alternatively, if $s > t$, and so $s - t = r$ for some $r \geq 1$, then $(bd^sc)(bd^tc)^{-1} = bd^scc^{-1}d^{-t}b^{-1} = bfd^rb^{-1}$, where $f = d^scc^{-1}d^{-s}$. However, $bfd^rb^{-1} = 0$, since $db^{-1} = 0$, by condition 2 of Theorem 10, Chapter 5. Therefore $(bd^sc)(bd^tc)^{-1} = 0$ for all $s \neq t$.

Similarly, $(bd^sc)^{-1}(bd^tc) = c^{-1}d^{-s}b^{-1}bd^tc = c^{-1}gd^{-s}d^tc$ where $g = d^{-s}b^{-1}bd^s$. So, if $t - s = -r$, where $r \geq 1$, then $(bd^sc)^{-1}(bd^tc) = c^{-1}gd^{-r}c = 0$ since $d^{-1}c = 0$, by condition 4 of Theorem 10, Chapter 5. Alternatively, $(bd^sc)^{-1}(bd^tc) = c^{-1}d^{-s}d^thc$, where $h = d^{-t}b^{-1}bd^t$. So, if $t - s = r \geq 1$, then $(bd^sc)^{-1}(bd^tc) = c^{-1}d^rhc = 0$ since $c^{-1}d = 0$, by condition 4 of Theorem 10, Chapter 5. Therefore, $(bd^sc)^{-1}(bd^tc) = 0$ for all $s \neq t$.

Hence we can deduce that $(bd^sc) \perp (bd^tc)$ for all $s \neq t$ and as we have already proved that $a \perp bd^nc$ for all $n \in \mathbb{N}$, we can deduce that all the terms in the union $Tr^U_{X,Y}(R)$ are disjoint, and so

$Tr^U_{X,Y}(R) \in \mathbf{Inj}(X, Y)$. Therefore, **Inj** is closed under the trace, and hence is a traced symmetric monoidal subcategory of **Rel**. □

6.3 The trace and self-similarity

We demonstrate how the endomorphism monoid of a self-similar object in a traced symmetric monoidal category has an operation defined on it that satisfies many of the same properties as the categorical trace, and gives a one-object traced symmetric monoidal category when the self-similarity is strong.

We then use these results to motivate the definition of a traced M-monoid, and, as a check that this is the 'correct' definition, demonstrate that the Karoubi tensor envelope of a traced M-monoid is a traced symmetric monoidal category. Finally, we show that, when the tensor category of a self-similar object is freely generated, the isomorphism between the tensor category, and the tensor envelope of the endomorphism monoid of the self-similar object preserves the trace.

6.3.1 Internalising the trace at self-similar objects

Definitions 6.2

Let N be a self-similar object of a symmetric traced monoidal category $(\mathbf{M}, \otimes)$, as defined in Definitions 4.1, Chapter 4. We define the *internalisation of the trace* to be a map $trace : \mathbf{M}(N, N) \to \mathbf{M}(N, N)$ given by $trace(f) = Tr^N_{N,N}(dfc)$ where c and d are the contraction and division morphisms of the self-similar object N. We will demonstrate that $trace$ represents the categorical trace under the internalisation process; however, we first require the following:

Lemma 3 *Let N be a self-similar object of a traced symmetric monoidal category, and let ϕ_2 be as in Definition 4.4 of Chapter 4. Then $Tr^N_{N,N}(f) = trace(\phi_2(f))$ for all $f \in \mathbf{M}(N \otimes N, N \otimes N)$.*

Proof By definition, $\phi_2(f) = cfd$, so $trace(\phi_2(f)) = Tr^N_{N,N}(dcfdc) = Tr^N_{N,N}(f)$. □

There is a natural extension of this result when $(\otimes\mathrm{N})$ is freely generated:

Theorem 4 *Let N be a self-similar object of a traced monoidal category $(\mathbf{M}, \otimes)$, and assume that $(\otimes\mathrm{N})$ is freely generated in $\mathbf{M}$, so we can define the contraction map ϕ, as in Definitions 4.3, Chapter 4. Then $trace(\phi(F)) = \phi(Tr^U_{X,Y}(F))$ for all $F \in (\otimes\mathrm{N})(X \otimes U, Y \otimes U)$.*

Proof $trace(\phi(F)) = Tr^N_{N,N}(d\phi(F)c) = Tr^N_{N,N}(dc_{Y\otimes U}Fd_{X\otimes U}c)$ by definition of *trace* and ϕ However, $c_{Y\otimes U} = c(c_Y \otimes c_U)$ and $d_{X\otimes U} = d(d_X \otimes d_U)$, by definition of the division and contraction elements. Therefore

$$trace(\phi(F)) = Tr^N_{N,N}(dc(c_Y \otimes c_U)F(d_X \otimes d_U)dc)$$

$$= Tr^N_{N,N}((c_Y \otimes c_U)F(d_X \otimes d_U)) = c_Y(Tr^N_{X,Y}((1_Y \otimes c_U)F(1_X \otimes d_U)))d_X,$$

by the naturality of the trace in X and Y, and so

$$trace(\phi(F)) = c_Y Tr^U_{X,Y}(F(1_X \otimes d_U c_U))d_X = c_Y Tr^U_{X,Y}(F(1_X \otimes 1_U))d_X = c_Y Tr^U_{X,Y}(F)d_X,$$

by the naturality of the trace in U. Therefore $trace(\phi(F)) = \phi(Tr^U_{X,Y}(F))$, by the definition of ϕ. $\Box$.

These results allow us to deduce that the internalisation of the categorical trace satisfies many similar properties to the categorical trace, as follows:

Theorem 5 *Let N be a self-similar object of a traced symmetric monoidal category* $(\mathbf{M}, \otimes)$, *let s, t, t^{-1} be as defined in Theorem 9 of Chapter 4, and let the internalisation of the trace be as defined above. Then*

(i) $trace(f(1 \oplus (1 \oplus 1))) = trace(trace(tft^{-1}))$,

(ii) $trace(f) \oplus g = trace(t(1 \oplus s)t^{-1}(f \oplus g)t(1 \oplus s)t^{-1})$,

(iii) $trace(s) = 1$,

(iv) $trace((h \oplus 1)f(g \oplus 1)) = h(trace(f))g$,

(v) $trace(f(1 \oplus g)) = trace((1 \oplus g)f)$.

Proof

(i) By definition of *trace*, t and t^{-1},

$$trace^2(tft^{-1}) = Tr^N_{N,N}(Tr^N_{N,N}(dc(c \otimes 1)t_{N,N,N}(1 \otimes d)dfc(1 \otimes c)t^{-1}_{N,N,N}(d \otimes 1)dc)c)$$

$$= Tr^N_{N,N}(Tr^N_{N,N}((c \otimes 1)t_{N,N,N}(1 \otimes d)dfc(1 \otimes c)t^{-1}_{N,N,N}(d \otimes 1))c)$$

which by the naturality of $Tr^U_{X,Y}$ in X and Y, is equal to

$$Tr^N_{N,N}(dc(Tr^N_{N\otimes N,N\otimes N}(t_{N,N,N}(1 \otimes d)dfc(1 \otimes c)t^{-1}_{N,N,N}))dc).$$

By axiom 2 for a traced symmetric monoidal category, this is equal to $Tr^{N\otimes N}_{N,N}((1 \otimes d)dfc(1 \otimes c))$ which, by the naturality of $Tr^U_{X,Y}$ in U, is equal to $Tr^N_{N,N}(dfc(1 \otimes cd))$. Finally, note that, by definition of *trace* and $\oplus$, $trace(f(1\oplus(1\oplus 1))) = Tr^N_{N,N}(dfc(1\otimes cd))$ and so $trace(f(1\oplus(1\oplus 1))) = trace(trace(tft^{-1}))$.

(ii) By definition, $trace(t(1\oplus s)t^{-1}(f\oplus g)t(1\oplus s)t^{-1}) =$

$$Tr^N_{N,N}(dc(c\otimes 1)t_{N,N,N}(1\otimes d)dc(1\otimes cs_{N,N}d)d$$

$$c(1\otimes c)t^{-1}_{N,N,N}(d\otimes 1)dc(f\otimes g)dc(c\otimes 1)t_{N,N,N}(1\otimes d)d$$

$$c(1\otimes cs_{N,N}d)dc(1\otimes c)tN,N,N^{-1}(d\otimes 1)dc)$$

$$= Tr^N_{N,N}((c\otimes 1)t_{N,N,N}(1\otimes d)(1\otimes cs_{N,N}d)(1\otimes c)t^{-1}_{N,N,N}(d\otimes 1)$$

$$(f\otimes g)$$

$$(c\otimes 1)t_{N,N,N}(1\otimes d)(1\otimes cs_{N,N}d)(1\otimes c)t^{-1}_{N,N,N}(d\otimes 1))$$

$$= Tr^N_{N,N}((c\otimes 1)t_{N,N,N}(1\otimes s_{N,N})t^{-1}_{N,N,N}(dfc\otimes g)t_{N,N,N}(1\otimes s_{N,N})tN,N,N^{-1}(d\otimes 1))$$

and by the naturality of $Tr^U_{X,Y}$ in X and Y, this is equal to

$$c(Tr^N_{N\otimes N,N\otimes N}(t_{N,N,N}(1\otimes s_{N,N})t^{-1}_{N,N,N}(dfc\otimes g)t_{N,N,N}(1\otimes s_{N,N})tN,N,N^{-1})d$$

which, by axiom 3 for a symmetric traced monoidal category, is equal to

$$c(Tr^N_{N,N}(dfc)\otimes g)d = c(trace(f)\otimes g)d = trace(f)\oplus g,$$

by definition of *trace* and $\oplus$.

(iii) Axiom 4 for a traced monoidal category states that $Tr^U_{U,U}(s_{U,U}) = 1_U$ for all $U \in Ob(\mathbf{M})$. Therefore $Tr^N_{N,N}(s_{N,N}) = 1_N$, and so $trace(s) = Tr^N_{N,N}(dsc) = Tr^N_{N,N}(s_{N,N}) = 1_N$, by definition of s. Therefore $trace(s) = 1$.

(iv) By the definition of *trace*,

$$trace((h\oplus 1)f(g\oplus 1)) = Tr^N_{N,N}(dc(h\otimes 1)dfc(g\otimes 1)dc)$$

$$= Tr^N_{N,N}((h\otimes 1)dfc(g\otimes 1)) = hTr^N_{N,N}(dfc)g$$

by the naturality of $Tr^U_{X,Y}$ in X and Y. However, this is just $h(trace(f))g$. Hence our result follows.

(v) By definition,

$$trace(f(1\oplus g)) = Tr^N_{N,N}(dfc(1\otimes g)dc) = Tr^N_{N,N}(dfc(1\otimes g)) = Tr^N_{N,N}((1\otimes g)dfc),$$

by the naturality of $Tr^U_{X,Y}$ in U. Similarly,

$$trace((1\oplus g)f) = Tr^N_{N,N}(dc(1\otimes g)dfc) = Tr^N_{N,N}((1\otimes g)dfc).$$

However, this is $trace(f(1\oplus g))$, from above. □

6.3.2 Traced M-monoids

The above results motivate the following definitions:

Definitions 6.3

Let $(M, \oplus)$ be an M-monoid, and let t, s denote the associativity and commutativity elements respectively. We say that M is *traced* if there exists a map $trace : M \to M$, which we call the *internal trace*, satisfying the following conditions:

1. $trace(f(1 \oplus (1 \oplus 1))) = trace(trace(tft^{-1}))$
2. $trace(f) \oplus g = trace(t(1 \oplus s)t^{-1}(f \oplus g)t(1 \oplus s)t^{-1})$
3. $trace(s) = 1$.
4. $trace((h \oplus 1)f(g \oplus 1)) = h(trace(f))g$
5. $trace(f(1 \oplus g)) = trace((1 \oplus g)f)$.

The axioms 1 to 3 can be thought of as the one-object case of the axioms 2 to 4 of a traced monoidal category, and the axioms 4 and 5 can be thought of as the one-object analogues of the naturality of the categorical trace. Of course, there is no analogue of axiom 1, as we do not have analogues of units elements.

To demonstrate that this is the 'correct' axiomatisation for a traced M-monoid, consider the following:

Theorem 6 *Let $(M, \oplus)$ be a traced M-monoid. Then the Karoubi tensor envelope of $(M, \oplus)$ is a traced symmetric monoidal category (without a unit object).*

Proof Recall the definition of the tensor envelope as a subcategory of the Karoubi envelope, from Definitions 4.12, Chapter 4. We have seen (Propositions 16 and 17 of Chapter 4) that this is a symmetric monoidal category (without a unit object). We define a map on $\mathbf{K}_{\mathrm{M}}^{\oplus}$ by $Tr^g_{e,f}(a) = trace(a)$ for all $a \in \mathbf{K}_{\mathrm{M}}^{\oplus}(e \otimes g, f \otimes g)$, where $trace$ is the internal trace of $(M, \oplus)$. Then $ftrace(a)e = trace((f \oplus 1)a(e \oplus 1))$, by axiom 4 for a traced M-monoid. However, $a \in \mathbf{K}_{\mathrm{M}}^{\oplus}(e \otimes g, f \otimes g)$, so $(f \oplus 1)a(e \oplus 1) = a$. Therefore, $Tr^g_{e,f}(a) \in \mathbf{K}_{\mathrm{M}}^{\oplus}(e, f)$ for all $a \in \mathbf{K}_{\mathrm{M}}^{\oplus}(e \otimes g, f \otimes g)$, and so Tr, as defined above, takes morphisms $f : X \otimes U \to Y \otimes U$ to morphisms $Tr(f) : X \to Y$, as required. We check the axioms for a traced symmetric monoidal category.

1. *This axiom is not considered, as* $\mathbf{K}^{\oplus}_{\mathbf{M}}$ *does not have a unit object.*

2. For all $a \in \mathbf{K}^{\oplus}_{\mathbf{M}}(e \otimes (g \otimes h)), f \otimes (g \otimes h))$, by definition,

$$Tr^{g}_{e,f}(Tr^{h}_{e\otimes g,f\otimes g}(t_{f,g,h}at^{-1}_{e,g,h})) = trace(trace(t_{f,g,h}at^{-1}t^{-1}_{e,g,h})).$$

Therefore, by definition of $t_{e,f,g} \in \mathbf{K}^{\oplus}_{\mathbf{M}}$ (from Proposition 16 of Chapter 4),

$$Tr^{g}_{e,f}(Tr^{h}_{e\otimes g,f\otimes g}(t_{f,g,h}at^{-1}_{e,g,h})) =$$

$$trace(trace(((f \oplus g) \oplus h)t(f \oplus (g \oplus h))a(e \oplus (g \oplus h))t^{-1}((e \oplus g) \oplus h))),$$

$$= trace(trace(t(f \oplus (g \oplus h))a(e \oplus (g \oplus h))t^{-1}))$$

and so, by axiom 1 for a traced M-monoid,

$$Tr^{g}_{e,f}(Tr^{h}_{e\otimes g,f\otimes g}(t_{f,g,h}at^{-1}_{e,g,h})) =$$

$$trace((f \oplus (g \oplus h))a(e \oplus (g \oplus h))(1 \oplus (1 \oplus 1)))$$

$$= trace((f \oplus (g \oplus h))a(e \oplus (g \oplus h))),$$

and as $a \in \mathbf{K}^{\oplus}_{\mathbf{M}}(e \otimes (g \otimes h), f \otimes (g \otimes h))$,

$$trace((f \oplus (g \oplus h))a(e \oplus (g \oplus h))) = trace(a),$$

Therefore, $Tr^{g}_{e,f}(Tr^{h}_{e\otimes g,f\otimes g}(t_{f,g,h}at^{-1}_{e,g,h})) = Tr^{g\otimes h}_{e,f}(a)$.

3. For all $a \in \mathbf{K}^{\oplus}_{\mathbf{M}}(e \otimes g, f \otimes g)$ and all $b \in K^{\oplus}_{M}(h, k)$,

$$Tr^{g}_{e\otimes h,f\otimes k}(t_{f,k,g}(1_f \otimes s_{g,k})t^{-1}_{f,g,k}(a \otimes b)(t_{e,g,h}(1_e \otimes s_{h,g})t^{-1}_{e,h,g}) =$$

$$trace(((f \oplus k) \oplus g)t(f \oplus (k \oplus g))(f \oplus s_{g,k})(f \oplus (g \oplus k))t^{-1}((f \oplus g) \oplus k)$$

$$(a \oplus b)$$

$$((e \oplus g) \oplus h)t(e \oplus (g \oplus h))(e \oplus s_{f,g})((e \oplus (h \oplus g))t^{-1}(e \oplus h) \oplus g))$$

$$= trace(((f \oplus k) \oplus g)t(f \oplus (k \oplus g))$$

$$(f \oplus (k \oplus g)s(g \oplus k))$$

$$(f \oplus (g \oplus k))t^{-1}((f \oplus g) \oplus k)$$

$$(a \oplus b)$$

$$((e \oplus g) \oplus h)t(e \oplus (g \oplus h))$$

$$(e \oplus (g \oplus h)s(h \oplus g))$$

$$(e \oplus (h \oplus g))t^{-1}((e \oplus h) \oplus g))$$

$$= trace(t(f \oplus (k \oplus h))(1 \oplus s)t^{-1}(a \oplus b)t(1 \oplus s)t^{-1}((e \oplus h) \oplus g))$$

$$= trace(t(1 \oplus s)t^{-1}((f \oplus g) \oplus k)(a \oplus b)((e \oplus g) \oplus h)t(1 \oplus s)t^{-1}$$

which equals $trace((f \oplus g)a(e \oplus g)) \oplus kbh$, (by axiom 2 for a traced M-monoid). However, $a \in \mathbf{K}^{\oplus}_{\mathbf{M}}(e \otimes g, f \otimes g)$, and $b \in \mathbf{K}^{\oplus}_{\mathbf{M}}(h, k)$, so this is $trace(a) \oplus b$. Therefore, by definition of Tr in $\mathbf{K}^{\oplus}_{\mathbf{M}}$,

$$Tr^{g}_{e \otimes h, f \otimes k}(t_{f,k,g}(1_f \otimes s_{g,k})t^{-1}_{f,g,k}(a \otimes b)t_{e,g,h}(1_e \otimes s_{h,g})t^{-1}_{e,h,g}) = Tr^{g}_{e,f}(a) \otimes b,$$

and so the third axiom for a traced symmetric monoidal category is satisfied.

4. For all $e \in Ob(\mathbf{K}^{\oplus}_{\mathbf{M}})$,

$$Tr^{e}_{e,e}(s_{e,e}) = trace((e \oplus e)s(e \oplus e)) = etrace((1 \oplus e)s(1 \oplus e))e$$

$$= etrace((e \oplus 1)s)e = e^2 trace(s)e = e.$$

by axioms 3, 4 and 5 for a traced M-monoid. Therefore, Tr satisfies axiom 4 for a traced symmetric monoidal category.

Finally, naturality of $Tr^{U}_{X,Y}$ in all three variables follows immediately from axioms 4 and 5 for a traced M-monoid. Therefore, Tr is a categorical trace on the tensor envelope of the M-monoid $(M, \oplus)$. □

We have proved that the tensor envelope of a traced M-monoid is a traced monoidal category (without unit elements). However, recall that (from Proposition 19 of Chapter 4) the tensor envelope of a strong M-monoid M is a one-object symmetric monoidal category isomorphic to M. Therefore, we are justified in claiming that the definition of a trace given above (in the strong case) is the correct definition of the categorical trace on a one-object symmetric monoidal category without the unit object. In the weak case, we have the following expected result:

Theorem 7 *Let N be a weak self-similar object of a symmetric traced monoidal category* $(\mathbf{M}, \otimes)$, *and let* $(\otimes \mathrm{N})$ *be freely generated in* $\mathbf{M}$. *Then the isomorphism* Φ *between* $(\otimes \mathrm{N})$ *and* $\mathbf{K}^{\oplus}_{\mathbf{M}(N,N)}$ *(from Theorem 20 of Chapter 4) preserves the trace.*

Proof Recall the definition of Φ as $\Phi(X) = c_X d_X$ for all $X \in Ob(\otimes\mathrm{N})$, and $\Phi(a) = c_Y a d_X$ for all $a \in (\otimes\mathrm{N})(X, Y)$. By Theorem 4 above, $trace(\phi(a)) = \phi(Tr^U_{X,Y}(a))$, for all $a \in (\otimes\mathrm{N})(X \otimes U, Y \otimes U)$, where *trace* is the internalisation of the trace on $(\otimes\mathrm{N})$. If we denote the trace on $\mathbf{K}^{\oplus}_{\mathbf{M}}$ by Tr', then $Tr'(\Phi(F)) = \Phi(Tr(F))$ by definition of Tr'. Hence the isomorphism Φ preserves the categorical trace. □

6.4 Compact closed categories and the categorical trace

6.4.1 Introduction

Up to this point, we have beeen considering the theory of traced symmetric monoidal categories. However, as demonstrated in [35], traced monoidal categories have a very close connection with tortile monoidal categories, and, in the symmetric case, with compact closed categories. We present the definitions and basic theory of compact closed categories, which are special cases of tortile monoidal categories ([46]), *-autonomous categories ([37]), and, of course, monoidal closed categories ([37]).

Fundamental to the following sections will be a result of Joyal, Street, and Verity [35], showing that every compact closed category has a canonical trace defined on it, and every traced symmetric monoidal category $(\mathbf{M}, \otimes)$ gives rise to a compact closed category, denoted **IntM**, in which $(\mathbf{M}, \otimes)$ is canonically embedded and the trace of $(\mathbf{M}, \otimes)$ is the canonical trace. (Note that their description and construction was for tortile monoidal categories and arbitrary traced monoidal categories, of which the compact closed categories result is the symmetric case).

We will present an explicit description of the compact closed category derived from the trace on the category of relations (again due to Joyal, Street, and Verity, [35]), and show how the compact closed category derived from the subcategory of partial bijective maps (which we have proved in Theorem 2 is traced) is also an inverse category.

6.4.2 Compact closed categories

Definitions 6.4

Compact closed categories are the symmetric case of *tortile monoidal* categories, as defined in [46]. In [37], compact closed categories are defined explicitly, as follows: A *compact closed category* $\mathbf{M}$ is a symmetric monoidal category, $(\mathbf{M}, \otimes, t, s, \lambda, \rho, I)$ where there exists a self-inverse map[1] $(\ )^{\vee}$

[1] In fact, the general definition only requires the existence of left duals (and not uniqueness), and only requires that $(A^{\vee})^{\vee} \cong A$ for all A. However, all the examples we consider (i.e. all examples constructed from Joyal, Street,

that takes an object A to its *left dual* $A^\vee$. Also, for every $A \in Ob(\mathbf{M})$, there exists two morphisms, the counit map $\epsilon_A : A^\vee \otimes A \to I$, and the unit map $\eta_A : I \to A \otimes A^\vee$, that satisfy the following coherence conditions:

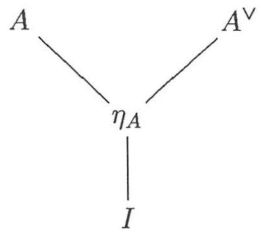
1. $\rho_A(1_A \otimes \epsilon_A)t^{-1}_{A,A^\vee,A}(\eta_A \otimes 1_A)\lambda^{-1}_A = 1_A$
2. $\lambda_{A^\vee}(\epsilon_A \otimes 1_{A^\vee})t_{A^\vee,A,A^\vee}(1_{A^\vee} \otimes \eta_A)\rho^{-1}_{A^\vee} = 1^\vee_A$

Using the definitions of the unit and counit maps, the definition of the *dual of a morphism* is given in [37], as follows:
For all $f \in \mathbf{M}(A, B)$, the dual of f is a map $f^\vee : B^\vee \to A^\vee$, defined by the composite

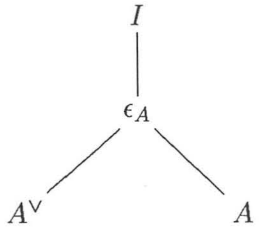
$$f^\vee = \lambda_{A^\vee}(\epsilon_B \otimes 1_{A^\vee})t_{B^\vee,B,A^\vee}(1_{B^\vee} \otimes (f \otimes 1_{A^\vee}))(1_B \otimes \eta_A)\rho^{-1}_{B^\vee}.$$

Note that [46] defines compact closed categories (as special cases of the more general tortile monoidal categories) in terms of right duals, rather than left duals; however, the two definitions are interchangable. It is proved in [46] that the correct unit and counit maps for right duals, $\epsilon'_U : U \otimes U^\vee \to I$ and $\eta'_U : I \to U^\vee \otimes U$, are given by $\epsilon' = \epsilon s_{U,U^\vee}$ and $\eta' = s_{U^\vee,U}\eta$ respectively. We use left duals in order to follow the conventions of [35].

Joyal and Street's diagrammatic reasoning, [33, 34], can also be used for compact closed categories, as proved in [34]. The unit maps are represented as follows:



and the counit maps are represented by



6.4.3 The canonical trace on a compact closed category

We present a result (due to Joyal, Street, and Verity) showing how every compact closed category has a trace defined on it, called the *canonical trace.*

---

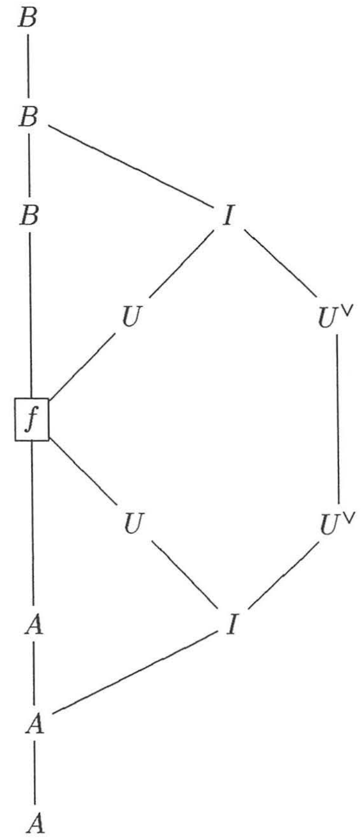
and Verity's construction, [35]) satisfy this stronger condition.

Theorem 8 *(Due to [35]) In any compact closed category* $(\mathbf{M}, \otimes, t, s, I, \lambda, \rho, \epsilon, \eta, (\ )^\vee)$ *a trace is defined by*

$$Tr^U_{A,B}(f) = \rho_B(1_B \otimes \epsilon'_{U^\vee})t^{-1}_{B,U,U^\vee}(f \otimes 1_{U^\vee})t_{A,U,U^\vee}(1_A \otimes \eta_U)\rho_A^{-1}.$$

Proof It would not be useful to reproduce this proof here; a proof can be found in [35], in terms of diagrammatic reasoning, and a formal justification of the diagram manipulations used can be found in [33, 34]. □

The above definition of the trace on a compact closed category can be represented diagramatically as follows:



Comparing this diagram with the diagram for a categorical trace shows why $U^\vee$, the dual of the object U, has been described as 'U moving in the opposite direction'.

6.4.4 Constructing compact closed categories from symmetric traced monoidal categories

We present the dual result to theorem 8 above, which is again due to Joyal, Street, and Verity, [35]. It gives a canonical way of constructing compact closed categories from symmetric traced monoidal categories (in fact, their result was a construction of tortile monoidal categories from arbitrary traced monoidal categories; we again present the symmetric case).

They use this theorem to construct a compact closed category from the trace on the category of relations; we prove that a compact closed category can be derived from the trace on the category of partial bijective maps, which is an invere subcategory.

Theorem 9 *(Due to [35], p.10-23) Let* $\mathbf{V}$ *be a symmetric traced monoidal category. A compact closed compact closed category, denoted* $\mathbf{IntV}$, *can be defined in terms of the objects and arrows of* $\mathbf{V}$, *and there exists an embedding of* $\mathbf{V}$ *into* $\mathbf{IntV}$, *where the trace on* $\mathbf{V}$ *is the canonical trace of* $\mathbf{IntV}$.

Proof The proof of this result is the symmetric case of the main theorem of [35]. □

We do not reproduce the proof here (see [35]), but give the construction of the category $\mathbf{IntV}$ in terms of the objects, morphisms, and traced monoidal structure of $\mathbf{V}$.

The construction is as follows:

Objects

The objects of $\mathbf{IntV}$ are defined to be pairs of objects of V, so

$$X, U \in Ob(\mathbf{V}) \;\Leftrightarrow\; (X, U) \in Ob(\mathbf{IntV}).$$

(In [35], it is stated that the object (X, U) should be thought of as a formalisation of $X \otimes U^{\vee}$). Also, the unit object is given by (I, I), where I is the unit object for the monoidal structure of $\mathbf{V}$.

Morphisms

The morphisms of $\mathbf{IntV}$ are defined in terms of morphisms of V, so $F : (X, U) \to (Y, V)$ in $\mathbf{IntV}$ is given by a morphism $f : X \otimes V \to Y \otimes U$ in $\mathbf{V}$. We say that F is *specified by* f. Hence there is a bijection of morphisms between $\mathbf{V}(X \otimes V, Y \otimes U)$ and $\mathbf{IntV}((X, U), (Y, V))$ denoted $f \mapsto F$, for all objects X, Y, U, V of $\mathbf{V}$.

The composition of morphisms in $\mathbf{IntV}$ is defined as follows:

Composition

Given $F : (X, U) \to (Y, V)$, and $G : (Y, V) \to (Z, W)$, specified by $f : X \otimes V \to Y \otimes U$, and $g : Y \otimes W \to Z \otimes V$ respectively, then their composite $GF : (X, U) \to (Z, W)$ is specified by

$$Tr^{V}_{X \otimes W, Z \otimes U}(C_{G,F}),$$

where $C_{G,F} : (X \otimes W) \otimes V \to (Z \otimes U) \otimes V$ is given by

$$C_{G,F} = t_{ZUV}(1_Z \otimes s_{VU})t^{-1}_{ZVU}(g \otimes 1_U)t_{YWU}(1_Y \otimes s_{UW})t^{-1}_{YUW}(f \otimes 1_W)t_{XVW}(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

Identities

For all objects (X, U), the identity at the object (X, U) is specified by $1_X \otimes 1_U$.

The monoidal functor

The monoidal functor of **IntV**, which we denote $\otimes : \mathbf{IntV} \times \mathbf{IntV} \to \mathbf{IntV}$ is defined (in terms of the monoidal functor on **V**) on objects by $(X, U) \otimes (X', U') = (X \otimes X', U' \otimes U)$ and on arrows by, for all $F : (X, U) \to (Y, V)$, and $F' : (X', U') \to (Y', V')$,

$$(F \otimes F') : (X \otimes X', U' \otimes U) \to (Y \otimes Y', V' \otimes V)$$

is specified by

$$t^{-1}((t \otimes 1) \otimes 1)((1 \otimes f') \otimes 1)(t^{-1} \otimes 1)(s \otimes s)$$

$$t^{-1}(t \otimes 1) \otimes 1)((1 \otimes f) \otimes 1)(t^{-1} \otimes 1)t(s \otimes s)$$

Where F is specified by f, and F' is specified by f'. (Note that the subscripts on the canonical isomorphisms have been omitted, for clarity).

The canonical isomorphisms

The canonical isomorphisms for **IntV** are defined as follows:

1. The commutativity isomorphism

$$S : (X, U) \otimes (X', U') \to (X', U') \otimes (X, U)$$

is specified by the following composite in **V**.

$$(s_{X,X'} \otimes s_{U',U}) : (X \otimes X') \otimes (U' \otimes U) \to (X' \otimes X) \otimes (U \otimes U').$$

2. The associativity morphism

$$T : (X, U) \otimes ((Y, V) \otimes (Z, W)) \to ((X, U) \otimes (Y, V)) \otimes (Z, W)$$

is specified by the following tensor in **V**,

$$T = t_{XYZ} \otimes t_{WVU}.$$

3. The left unit map $\Lambda : (I, I) \otimes (A, B) \to (A, B)$ is specified by

$$(\rho_B^{-1} \otimes \lambda_A)s.$$

Similarly, the right unit map $R : (A, B) \otimes (I, I) \to (A, B)$ is specified by

$$(\lambda_B^{-1} \otimes \rho_A)s.$$

The compact closed structure

The dual on objects is defined by $(X, U)^\vee = (U, X)$. The unit maps $\epsilon_{(X,U)} : (U, X) \otimes (X, U) \to (I, I)$ are then specified by $s_{U \otimes X, I}$ and the counit maps $\eta_{(X,U)} : (I, I) \to (X, U) \otimes (U, X)$ are specified by $s_{I,(X \otimes U)}$.

The dual on morphisms can then be directly defined as follows: Let $F : (X, U) \to (Y, V)$ be specified by $f : X \otimes V \to Y \otimes U$. Then $F^\vee : (U, X) \to (V, Y)$ is specified by $s_{Y,U} f s_{V,X}$.

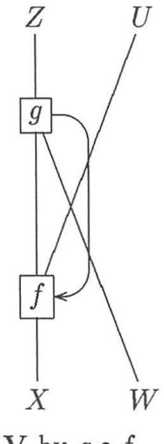
Note that this allows us to define the dual on objects without reference to the unit object or elements of **IntV**. This will be important when we come to constructing one-object compact closed categories (without units elements).

Finally, the embedding ι of **V** into **IntV** is given by

- On objects, $\iota(X) = (X, I)$,
- On morphisms, $\iota(f) = f \otimes 1_I$, for all $f : X \to Y$.

It is proved in [35] that this is an injective functor, and the trace of **V** under this functor is the canonical trace given by the compact closed structure of **IntV**.

Note that the composition (in terms of the trace on the category **V**) can be represented diagramatically, as follows:



We will denote this composite / trace in **V** by $g \circ f$.

6.4.5 The compact closed category derived from the category of relations

We apply the construction of Theorem 9 above to the traced symmetric monoidal category $(\mathbf{Rel}, \sqcup)$, to construct a compact closed category **IntRel**, in which **Rel** is embedded, and where the trace on (the embedding of) **Rel** is the canonical trace on the category. The following construction is also due to [35].

Definitions 6.5

IntRel is defined to be the category constructed by applying the construction of Theorem 9 to the traced symmetric monoidal category **Rel**. An explicit description of **IntRel** is given in [35], as follows:

Objects

The objects of **IntRel** are pairs of objects of **Rel**; that is, they are pairs of sets (X, U), where (X, U) is a formalisation of $X \otimes U^\vee$.

Morphisms

A morphism $R \in \mathbf{IntRel}((X,U),(Y,V))$ is specified by a morphism $r \in \mathbf{Rel}(X \sqcup V, Y \sqcup U)$. The morphism r can be written as a matrix $r = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where $a \in \mathbf{Rel}(X,Y)$, $b \in \mathbf{Rel}(V,Y)$, $c \in \mathbf{Rel}(X,U)$, and $d \in \mathbf{Rel}(V,U)$. So, given morphisms $F \in \mathbf{IntRel}((X,U),(Y,V))$ and $G \in \mathbf{IntRel}((Y,V),(Z,W))$ specified by $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $g = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ respectively, then their composite $GF \in \mathbf{IntRel}((X,U),(Z,W))$ is specified by $Tr^V_{X\otimes W, Z\otimes U}(C_{G,F}) \in \mathbf{Rel}(X \sqcup W, U \sqcup Z)$, where $C_{G,F} : (X \otimes W) \otimes V \to (Z \otimes U) \otimes V$ is given by

$$C_{G,F} = t_{ZUV}(1_Z \otimes s_{VU})t^{-1}_{ZVW}(g \otimes 1_U)t_{YWU}(1_Y \otimes s_{UW})t^{-1}_{YUW}(f \otimes 1_W)t_{XVW}(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

So, the composite of the two matrices in **IntRel** is specified by

$$g \circ f = \begin{pmatrix} e(bg)^*a & f \cup e(bg)^*bh \\ c \cup d(gb)^*ga & d(gb)^*h \end{pmatrix} \in \mathbf{Rel}(X \sqcup W, Z \sqcup U).$$

The monoidal functor

The definition of the tensor product in the category **IntV**, for a traced symmetric monoidal category **V** gives the following definition of the tensor in **IntRel**:

The tensor product of objects is defined by $(X,U)\otimes(X',U') = (X\sqcup X', U'\sqcup U)$, and for all $F\in \mathbf{IntRel}((X,U),(Y,V))$ and $G\in\mathbf{IntRel}((X',U'),(Y',V'))$, specified by

$$f = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{Rel}(X\sqcup V, Y\sqcup U)\ ,\ g = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix} \in \mathbf{Rel}(X'\sqcup V', Y'\sqcup U'),$$

respectively, then their tensor product is defined as follows:

$$F\otimes G \in \mathbf{IntRel}((X,U)\otimes(X',U'),(Y,V)\otimes(Y',V'))$$

is specified by

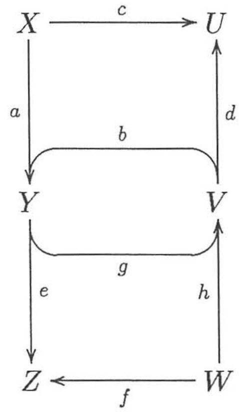
$$\begin{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a' \end{pmatrix} & \begin{pmatrix} 0 & b \\ b' & 0 \end{pmatrix} \\ \begin{pmatrix} 0 & c' \\ c & 0 \end{pmatrix} & \begin{pmatrix} d' & 0 \\ 0 & d \end{pmatrix} \end{pmatrix}$$

The dual structure

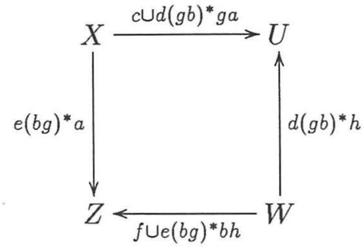
The dual on objects is given by $(U,V)^\vee = (V,U)$. The unit and counit maps are then defined as follows: $\eta_{(X,U)} : (I,I)\rightarrow (X,U)\otimes(X,U)^\vee$ is specified by $s_{I,X\otimes U}$ in **Rel**, and $\epsilon_{(X,U)}$: $(X,U)^\vee\otimes(X,U)\rightarrow(I,I)$ is specified by $s_{U\otimes X,I}$ in **Rel**.

From Theorem 9, $F^\vee$, the dual of a morphism F, is specified by $s_{Y,U} f s_{V,X} : V\sqcup X\rightarrow U\sqcup Y$ in **Rel** for all $F:(X,U)\rightarrow(Y,V)$ specified by $f: X\sqcup V\rightarrow Y\sqcup U$ in **Rel**.

In keeping with our intuitive idea of matrix multiplication and trace as 'finding all possible paths through a labelled digraph', the composition of **IntRel** can be represented as the following 'sum over all paths' construction:



becomes



6.4.6 The compact closed subcategory derived from the partial bijective maps

We have already proved that **Inj** is a traced symmetric monoidal subcategory of **Rel** (Theorem 2 of this Chapter). Therefore, we can construct the compact closed category **IntInj** as a compact closed subcategory of **IntRel**. The construction is as above. Hence the objects of **IntInj** are pairs of objects of **Inj**, so $X, Y \in Ob(\mathbf{Inj})$ iff $(X, Y) \in Ob(\mathbf{IntInj})$, and a morphism $F \in \mathbf{IntInj}((X, U), (Y, V))$ is a morphism $f : X \sqcup V \to Y \sqcup U$ in **Inj**. Finally, the tensor product, canonical morphisms, unit and counit for **IntInj** are as described above for the category **IntRel**.

We prove explicitly that **IntInj** is an inverse category.

Theorem 10 *The subcategory* **IntInj** *of* **IntRel** *is an inverse category.*

Proof *In what follows, we make use of the conditions for a matrix of partial bijective maps to represent a partial bijective map (Theorem 10 of Chapter 5).*

For every $F \in \mathbf{IntInj}((X, U), (Y, V))$, specified by $\begin{pmatrix} a & b \\ c & d \end{pmatrix} : X \sqcup V \to Y \sqcup U$ we define $F^{-1} \in \mathbf{IntInj}((Y, V), (X, U))$, to be specified by $f^{-1} = \begin{pmatrix} a^{-1} & c^{-1} \\ b^{-1} & d^{-1} \end{pmatrix} : Y \sqcup U \to X \sqcup V$. So, using the composition inherited from the category **IntRel**, $F^{-1}F \in \mathbf{IntInj}((X, U), (X, U))$ is specified by

$$f^{-1} \circ f = \begin{pmatrix} a^{-1}(bb^{-1})^* a & c^{-1} \vee a^{-1}(bb^{-1})^* bd^{-1} \\ c \vee d(b^{-1}b)^* b^{-1} a & d(b^{-1}b)^* d^{-1} \end{pmatrix}.$$

However, by condition 3 for the matrix f to represent a partial bijective map, $a^{-1}(bb^{-1})^* a = a$ and $c^{-1} \vee a^{-1}(bb^{-1})^* bd^{-1} = c^{-1}$. Also, by condition 2 for the matrix f to represent a partial bijective map, $c \vee d(b^{-1}b)^* b^{-1} a = c$ and $d(b^{-1}b)^* d^{-1} = dd^{-1}$. Therefore,

$$f^{-1} \circ f = \begin{pmatrix} a^{-1}a & c^{-1} \\ c & dd^{-1} \end{pmatrix}.$$

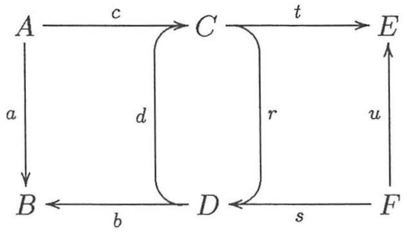
So, the composite $FF^{-1}F$ in **IntInj** is represented by

$$f \circ f^{-1} \circ f = \begin{pmatrix} a(c^{-1}c)^* a^{-1} a & b \vee a(c^{-1}c)^* c^{-1} d \\ c \vee dd^{-1}(cc^{-1})^* ca^{-1} a & dd^{-1}(cc^{-1})^* d \end{pmatrix}.$$

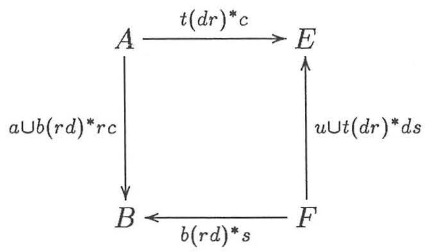
However, by condition 1 for the matrix f to represent a partial bijective map, $c \vee dd^{-1}(cc^{-1})^{*}ca^{-1}a = c$, $a(c^{-1}c)^{*}a^{-1}a = a$ and $b \vee a(c^{-1}c)^{*}c^{-1}d = b$. Finally, by condition 4 for the matrix f to represent a partial bijective map, $dd^{-1}(cc^{-1})^{*}d = d$. Therefore, we can deduce that $FF^{-1}F \in$ **IntRel**$((X,U),(Y,V))$ is specified by the matrix $f \circ f^{-1} \circ f = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = f$ and so $FF^{-1}F = F$. Therefore, in **IntInj**, every element F has a generalised inverse F^{-1} satisfying $FF^{-1}F = F$. Also, $(F^{-1})^{-1} = F$, so we can also deduce that $F^{-1}FF^{-1} = F^{-1}$. Finally, uniqueness of the generalised inverses satisfying these properties follows from the uniqueness of the generalised inverses in **Inj**. Therefore, we have proved that **IntInj** is an inverse category. □

6.4.7 An alternative composition on IntRel and IntInj

The above description of the composition of morphisms of **IntRel** in terms of finding all possible paths through a labelled digraph (from Definitions 6.5) suggests another possible composition on morphisms of **IntRel** (and hence of **IntInj**); that is,



becomes



We formalise this idea, and show that it gives an isomorphic category.

First note that the (partially defined) map $\circ : Arr(\mathbf{Rel}) \times Arr(\mathbf{Rel}) \to Arr(\mathbf{Rel})$, together with the specification of the identities at the objects, gives $Arr(\mathbf{Rel})$ a category structure, by Theorem 9, and Definitions 6.5. We define another (partial) map, as follows:

Definitions 6.6

Given morphisms $\alpha : A \sqcup C \to B \sqcup D$, and $\beta : C \sqcup F \to D \sqcup E$, represented by the matrices

$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \beta = \begin{pmatrix} r & s \\ t & u \end{pmatrix}.$$

resepctively, then the *alternative composition* (which we denote $\cdot$) is given by

$$\beta \cdot \alpha = \begin{pmatrix} a \cup b(rd)^*rc & b(rd)^*s \\ t(dr)^*c & u \cup t(dr)^*ds \end{pmatrix}.$$

Lemma 11 *Given morphisms $\alpha : A \sqcup D \to B \sqcup D$, and $\beta : C \sqcup F \to D \sqcup E$, represented by the matrices*

$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \beta = \begin{pmatrix} r & s \\ t & u \end{pmatrix}.$$

then the alternative composition satisfies $\beta \cdot \alpha = s_{EB}(s_{DE}\beta \circ s_{BC}\alpha)$.

Proof First note that $s_{XY} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} : X \sqcup Y \to Y \sqcup X$, so $s_{BC}\alpha = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$ and $s_{DE}\beta = \begin{pmatrix} t & u \\ r & s \end{pmatrix}$. Therefore, by definition of $\circ$,

$$(s\beta \circ s\alpha) = \begin{pmatrix} t(dr)^*c & u \cup t(dr)^*ds \\ a \cup b(rd)^*rc & b(rd)^*s \end{pmatrix}.$$

(We omit subscripts on commutativity morphisms, for clarity). Finally,

$$s(s\beta \circ s\alpha) = \begin{pmatrix} a \cup b(rd)^*rc & b(rd)^*s \\ t(dr)^*c & u \cup t(dr)^*ds \end{pmatrix} = \alpha \cdot \beta.$$

Therefore our result follows. □

Theorem 12 $(\mathbf{Rel}, \circ) \cong (\mathbf{Rel}, \cdot)$.

Proof The isomorphism S is defined on the underlying set of morphisms (in terms of the composition of **Rel**) as follows:

For all $f : A \sqcup X \to B \sqcup Y$, $S(f) : A \sqcup X \to Y \sqcup B$ is defined by $S(f) = s_{B,Y}f$. Then, as $s_{B,Y}$ is an isomorphism for all B, Y, we can deduce that this map is a bijection, and from Lemma 11 above, $S(f \circ g) = S(f) \cdot S(g)$. Finally, the identity arrows are those matrices of the form $\begin{pmatrix} 0 & 1_X \\ 1_U & 0 \end{pmatrix}$, and so the map S is a bijective functor. □

6.5 Self-similarity in compact closed categories

6.5.1 Introduction

We wish to develop the theory of self-similarity in compact closed categories, with a view to defining and constructing examples of one-object compact closed categories (without units). However, all the axioms for a compact closed category depend on the unit object, and as Proposition 1 of Chapter 4 shows, any one-object symmetric monoidal category with units elements has trivial monoidal structure. On the other hand, we have also seen (Theorem 9 above) how the dual on morphisms of **IntV** can be given without reference to the unit object or elements. This motivates the construction of an alternative set of axioms for a compact closed category that do not depend on unit object, and hence allows us to construct compact closed M-monoids in a similar way to the construction of traced M-monoids, as in Section 6.3.

6.5.2 An alternative characterisation of compact closed categories

We give a set of axioms (in which the unit object is not fundamental) for a category that we prove is equivalent to the set of axioms for a compact closed category.

Definitions 6.7

We define a *pairing category* to be a symmetric monoidal category $(\mathbf{T}, \otimes)$, where, for every object $A \in Ob(\mathbf{T})$, there exists a *dual object* $A^\vee$, together with morphisms that we call the *pair-creation*, and *pair destruction*, maps

- $\kappa_{XA} : X \to (A \otimes A^\vee) \otimes X$
- $\delta_{XA} : X \otimes (A^\vee \otimes A) \to X$

that are natural in X, and satisfy the following axioms

1. $\delta_{AA} t^{-1}_{AA^\vee A} \kappa_{AA} = 1_A$,
2. $\delta_{A^\vee A} s_{A^\vee \otimes A, A^\vee} t_{A^\vee AA^\vee} s_{A \otimes A^\vee, A^\vee} \kappa_{A^\vee A} = 1_{A^\vee}$,
3. $(\rho_{A \otimes A^\vee} \kappa_{IA} \otimes 1_X) \lambda_X^{-1} = \kappa_{XA}$,
4. $\rho_X (1_X \otimes \delta_{IA} \lambda^{-1}_{A^\vee \otimes A}) = \delta_{XA}$.

The condition that κ_{XA} and δ_{XA} are natural in X can be written explicitly as, for all $f : X \to Y$,

- $((1_A \otimes 1_{A^\vee}) \otimes f) \kappa_{XA} = \kappa_{YA} f$,

- $f\delta_{XA} = \delta_{YA}(f \otimes (1_{A^\vee} \otimes 1_A))$.

We demonstrate that the axioms for pairing categories and compact closed categories are equivalent.

Theorem 13 *A category that satisfies the axioms for a compact closed category also satisfies the axioms for a pairing category, and vice versa.*

Proof

($\Rightarrow$) Let $(\mathbf{T}, \otimes, \eta, \epsilon)$ be a compact closed category (we use the standard notation for the canonical associativity, commutativity, and units morphisms). We define $\kappa_{XA} = (\eta_A \otimes 1_X)\lambda_X^{-1}$ and $\delta_{XA} = \rho_X(1_X \otimes \epsilon_A)$ for all $X, A \in Ob(\mathbf{T})$. Then consider arbitrary $f : X \to Y$. By the definition of κ, and the naturality of λ_X in X,

$$(1_{A\otimes A^\vee} \otimes f)\kappa_{XA} = (1_{A\otimes A^\vee} \otimes f)(\eta_A \otimes 1_X)\lambda_X^{-1}$$

$$= (\eta_A \otimes 1_Y)(1_I \otimes f)\lambda_X^{-1} = (\eta_A \otimes 1_Y)\lambda_Y^{-1} f = \kappa_{YA} f.$$

Therefore, κ_{XA} is natural in X. Similarly, by definition of δ, and the naturality of ρ_X in X,

$$f\delta_{XA} = f\rho_X(1_X \otimes \epsilon_A) = \rho_Y(f \otimes 1_I)(1_X \otimes \epsilon_A)$$

$$= \rho_Y(1_Y \otimes \epsilon_A)(f \otimes 1_{A^\vee \otimes A}) = \delta_{YA}(f \otimes 1_{A^\vee \otimes A}).$$

Therefore, δ_{XA} is natural in X.

Also, by definition of κ and δ,

1. $\delta_{AA} t^{-1}_{AA^\vee A} \kappa_{AA} = \rho_A(1_A \otimes \epsilon_A) t^{-1}_{AA^\vee A}(\eta_A \otimes 1_A)\lambda_A^{-1} = 1_A$, by axiom 1 for a compact closed category (Definitions 6.4).

2. $\delta_{A^\vee A} s_{A^\vee \otimes A, A^\vee} t_{A^\vee A A^\vee} s_{A\otimes A^\vee, A^\vee} \delta_{A^\vee A} =$

$$\rho_{A^\vee}(1_{A^\vee} \otimes \epsilon_A) s_{A^\vee \otimes A, A^\vee} t_{A^\vee A A^\vee} s_{A\otimes A^\vee, A^\vee}(\eta_A \otimes 1_A^\vee)\lambda_{A^\vee}^{-1},$$

   by definition, and by the naturality of s_{XY} in X and Y, this is equal to

$$\rho_{A^\vee} s_{I,A^\vee}(\epsilon_A \otimes 1_{A^\vee}) t_{A^\vee A A^\vee}(1_{A^\vee} \otimes \eta_A) s_{A^\vee, I}\lambda_{A^\vee}^{-1}.$$

   Also, as $\rho_X s_{IX} = \lambda_X$, this is equal to $\lambda_{A^\vee}(\epsilon_A \otimes 1_{A^\vee}) t_{A^\vee A A^\vee}(1_{A^\vee} \otimes \eta_A)\rho_{A^\vee}^{-1}$ and by axiom 2 for a compact closed category, this is $1_{A^\vee}$.

3. $(\rho_{A\otimes A^\vee}\otimes 1_X)(\kappa_{IA}\otimes 1_X)\lambda_X^{-1} = (\rho_{A\otimes A^\vee}\otimes 1_X)(\eta_A\otimes 1_I)(\lambda_I^{-1}\otimes 1_X)$

$$= (\eta_A\otimes 1_X)(\rho_I\lambda_I^{-1}\otimes 1_X)\lambda_X^{-1} = (\eta_A\otimes 1_X)(1_I\otimes 1_X)\lambda_X^{-1} = (\eta_A\otimes 1_X)\lambda_X^{-1} = \kappa_{XA}$$

by the naturality of ρ_Z in Z, and the definition of κ_{XA}.

4. $\rho_X(1_X\otimes\delta_{IA})(1_X\otimes\lambda_{A^\vee\otimes A}^{-1}) = \rho_X(1_X\otimes\rho_I(1_I\otimes\epsilon_A))(1_X\otimes\lambda_{A^\vee\otimes A}^{-1})$

$$= \rho_X(1_X\otimes\rho_I)(1_X\otimes\lambda_I^{-1})(1_X\otimes\epsilon_A) = \rho_X(1_X\otimes\rho_I\lambda_I^{-1})(1_X\otimes\epsilon_A) = \rho_X(1_X\otimes\epsilon_A) = \delta_{XA},$$

by the definition of κ and the naturality of λ_Z in Z.

Therefore, κ_{XA} and δ_{XA} are morphisms that are natural in X, and satisfy the axioms for a pairing category. Therefore, every compact closed category has the structure of a pairing category.
($\Leftarrow$) Let $(\mathbf{T},\otimes)$ be a pairing category, so $\kappa_{XA}: X\to(A\otimes A^\vee)\otimes X$ and $\delta_{XA}:(A^\vee\otimes A)\otimes X\to X$ are as in Definitions 6.7. For all $X, A\in Ob(\mathbf{T})$, we define maps $\epsilon_A:(A^\vee\otimes A)\to I$ and $\eta_A: I\to(A\otimes A^\vee)$ by $\eta_A=\rho_{A\otimes A^\vee}\kappa_{IA}$ and $\epsilon_A=\delta_{IA}\lambda_{A^\vee\otimes A}^{-1}$.
We check the axioms for a compact closed category (Definitions 6.4).

1. $\rho_A(1_A\otimes\epsilon_A)t_{AA^\vee A}^{-1}(\eta_A\otimes 1_A)\lambda_A^{-1} = \kappa_{A,A}t_{AA^\vee A}^{-1}\delta_{AA} = 1_A$, by axiom 1 for a pairing category.

2. $\lambda_{A^\vee}(\epsilon_{A^\vee}\otimes 1_{A^\vee})t_{A^\vee AA^\vee}(1_{A^\vee}\otimes\eta_A)\rho_{A^\vee}^{-1} = \lambda_{A^\vee}(\delta_{IA^\vee}\lambda_{A\otimes A^\vee}^{-1}\otimes 1_{A^\vee})t_{A^\vee AA^\vee}(1_{A^\vee}\otimes\rho_{A\otimes A^\vee}\kappa_{IA})\rho_{A^\vee}^{-1}$ by definition of ϵ and η. However, $\lambda_{A^\vee}=\rho_{A^\vee}s_{IA^\vee}$ and $\rho_{A^\vee}^{-1}=s_{IA^\vee}\lambda_{A^\vee}^{-1}$. Therefore, this is equal to $\rho_{A^\vee}s_{IA}(\delta_{IA^\vee}\lambda_{A\otimes A^\vee}^{-1}\otimes 1_{A^\vee})t_{A^\vee AA^\vee}(1_{A^\vee}\otimes\rho_{A\otimes A^\vee}\kappa_{IA})s_{IA^\vee}\lambda_{A^\vee}^{-1}$, and by the naturality of s_{XY} in X and Y, the above is equal to

$$\rho_{A^\vee}(1_{A^\vee}\otimes\delta_{IA^\vee}\lambda_{A\otimes A^\vee}^{-1})s_{A^\vee\otimes A,A^\vee}t_{A^\vee AA^\vee}s_{A\otimes A^\vee,A^\vee}(\rho_{A\otimes A^\vee}\kappa_{IA}\otimes 1_{A^\vee})\lambda_{A^\vee}^{-1}.$$

Then by axioms 3, 4 for a pairing category, this is $\delta_{A^\vee A}s_{A^\vee\otimes A,A^\vee}t_{A^\vee AA^\vee}s_{A\otimes A^\vee,A^\vee}\kappa_{A^\vee A}$, which is $1_{A^\vee}$, by axiom 2 for a pairing category.

Finally the naturality of ϵ and η follows from the naturality of κ, δ, λ and ρ in a pairing category. Therefore, the axioms for a compact closed category are satisfied, and our result follows.

Note that the two definitions are interchangable; let $\mathbf{T}$ be a pairing category, and let ϵ and η be defined in terms of κ and δ, as above. Then the above definition of κ and δ in terms of ϵ and η gives the original κ, and δ, as follows:

The definition of κ in terms of η gives $\kappa_{XA}=(\eta_A\otimes 1_X)\lambda_X^{-1}$, and the definition of η in terms of κ gives $\kappa_{XA}=(\rho_{A\otimes A^\vee}\kappa_{IA}\otimes 1_X)\lambda_X^{-1}$. However, by axiom 3 for a pairing category, this is just

κ_{XA}, and our result follows. A similar proof holds for δ.

Conversely, let $\mathbf{T}$ be a compact closed category, and let κ and δ be defined in terms of ϵ and η. Then the definition of ϵ in terms of κ gives $\epsilon_A = \delta_{IA}\lambda^{-1}_{A^\vee \otimes A}$. However this is $\epsilon_A \rho_{A^\vee \otimes A}\lambda^{-1}_{A^\vee \otimes A} = \epsilon_A$, by the naturality of ρ. A similar proof holds for η. Therefore the two definitions are compatible. □

Convention In view of Theorem 13 above, we refer to the axioms for a pairing monoidal category (Definitions 6.7) as the *alternative axioms for a compact closed category*, and refer to the axioms of Definitions 6.4 as the standard axioms for a compact closed category. We also use the definitions interchangably, with the assumption that κ, δ and ϵ, η are related as in Theorem 13.

Note that, with the alternative axiom set for a compact closed category, the definition of the canonical trace becomes the following:

$$Tr^U_{A,B}(f) = \rho_B(1_B \otimes \epsilon'_{U^\vee})t^{-1}_{B,U,U^\vee}(f \otimes 1_{U^\vee})t_{A,U,U^\vee}(1_A \otimes \eta_U)\rho^{-1}_A$$

$$= \delta_{B,U} s_{A,(U\otimes U^\vee)} t^{-1}_{B,U,U^\vee}(f \otimes 1_{U^\vee}) t_{A,U,U^\vee} s_{A,U\otimes U^\vee} \kappa_{A,U}.$$

Therefore, the trace is definable in terms of κ and δ, without reference to the unit object or elements.

6.5.3 Self-similarity and one-object compact closed categories

We consider the rôle of self-similarity in constructing one-object compact closed categories (without unit elements). For this to make any sense, we require self-similar objects in compact closed categories that are *self-dual*; that is, they satisfy $N^\vee = N$. This is not a major restriction, as, for any object A, there exists an isomorphism between $(A \otimes A^\vee)$ and $(A \otimes A^\vee)^\vee$ (by the coherence theorem for compact closed categories, as proved in [37]), and we will construct examples later where this is a strict identity.

Definitions 6.8

Let $(\mathbf{T}, \otimes)$ be a compact closed category. We say that an object $N \in Ob(\mathbf{T})$ is *very (strongly) self-similar* if it satisfies

1. $N^\vee = N$,

2. $(X \otimes Y)^\vee = Y^\vee \otimes X^\vee$ for all $X, Y \in Ob(\otimes \mathrm{N})$.

3. N is (strongly) self-similar in the sense of Definitions 4.1, Chapter 4.

Lemma 14 *Let* $(\mathbf{T}, \otimes)$ *be a compact closed category, and let* N *be a very self-similar object of* $\mathbf{T}$*. Then* $(\otimes\mathrm{N})$ *is a compact closed category (without unit elements).*

Proof We merely need to prove that $A^\vee \in Ob(\otimes\mathrm{N})$ for all $A \in Ob(\otimes\mathrm{N})$, as the canonical morphisms will be inherited from $\mathbf{T}$. Firstly, as N is self-dual, $N^\vee = N$. Also, by axiom 2 for a very self-similar object, $(X \otimes Y)^\vee = Y^\vee \otimes X^\vee \in Ob(\otimes\mathrm{N})$ for all $X, Y \in Ob(\otimes\mathrm{N})$ satisfying $X^\vee, Y^\vee \in Ob(\otimes\mathrm{N})$. Therefore our result follows by the inductive definition of objects of $(\otimes\mathrm{N})$ (Definitions 4.3, Chapter 4). □

Theorem 15 *Let* N *be a very self-similar object of a compact closed category* $(\mathbf{M}, \otimes)$*, and let* $(\otimes\mathrm{N})$ *be freely generated in* $\mathbf{M}$*. Then for all* $X, Y \in Ob(\otimes\mathrm{N})$*,* $\kappa_{X,Y}$ *and* $\delta_{X,Y}$ *can be expressed in terms of* $\kappa_{N,N}$*,* $\delta_{N,N}$*, the canonical isomorphisms for the symmetric monoidal structure, and the division / contraction morphisms for* N*.*

We will prove this result in 6 steps, as follows:

(i) *For all* $X, A, B \in Ob(\otimes\mathrm{N})$*, there exists a definition of* $\kappa_{X,A\otimes B}$ *in terms of* κ_{XA}*, and* $\kappa_{B,A^\vee}$*.*

(ii) *For all* $X, Y \in Ob(\otimes\mathrm{N})$*, there exists a definition of* κ_{XY} *in terms of* κ_{XN}*.*

(iii) *For all* $X \in Ob(\otimes\mathrm{N})$*, there exists a definition of* κ_{XN} *in terms of* κ_{NN}*.*

(iv) *For all* $X, A, B \in Ob(\otimes\mathrm{N})$*, there exists a definition of* $\delta_{X,(A\otimes B)}$ *in terms of* δ_{XA}*,* $\delta_{B^\vee A}$*.*

(v) *For all* $X, Y \in (\otimes\mathrm{N})$*, there exists a definition of* δ_{XY} *in terms of* δ_{XN}*.*

(vi) *For all* $X \in (\otimes\mathrm{N})$*, there exists a definition of* δ_{XN} *in terms of* δ_{NN}*.*

The above 6 steps are then enough to give a definition of $\kappa_{X,Y}$ *and* $\delta_{X,Y}$ *in terms of* $\kappa_{N,N}$ *and* $\delta_{N,N}$*.*

Proof (we use the description of κ and δ in terms of ϵ and η).

(i) First note that, by the coherence theorem for compact closed categories ([37]), there exists a canonical isomorphism $u_{AB} : B^\vee \otimes A^\vee \to (A \otimes B)^\vee$, and the coherence theorem states that η_A, η_B, and $\eta_{A\otimes B}$ are related as follows:

$$(1_{A\otimes B} \otimes u_{AB}^{-1})\eta_{A\otimes B} = t_{A,B,(B^\vee\otimes A^\vee)}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})(1_A \otimes (\eta_B \otimes 1_{A^\vee}))(1_A \otimes \lambda^{-1}_{A^\vee})\eta_A.$$

However, by axiom 2 for a very self-similar object, the canonical isomorphism u_{AB} is the identity. Therefore, $\eta_{A\otimes B} = t_{A,B,(B^\vee\otimes A^\vee)}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})(1_A \otimes (\eta_B \otimes 1_{A^\vee}))(1_A \otimes \lambda^{-1}_{A^\vee})\eta_A$. Hence, for arbitrary $X \in Ob(\otimes\mathrm{N})$,

$$(\eta_{A\otimes B} \otimes 1_X)\lambda_X^{-1} = ((t_{A,B,(B^\vee\otimes A^\vee)}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})(1_A \otimes (\eta_B \otimes 1_{A^\vee}))(1_A \otimes \lambda^{-1}_{A^\vee})\eta_A) \otimes 1_X)\lambda_X^{-1}.$$

However, by definition of κ_{XY} in terms of η, λ, this becomes

$$\kappa_{X,(A\otimes B)} = t_{A,B,B^\vee\otimes A^\vee}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})(1_A \otimes (\eta_B \otimes 1_{A^\vee})\eta_A \otimes 1_X)\lambda_X^{-1}$$

$$= t_{A,B,B^\vee\otimes A^\vee}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})((1_A \otimes \kappa_{A^\vee B}) \otimes 1_X)(\eta_A \otimes 1_X)\lambda_X^{-1}$$

$$= t_{A,B,B^\vee\otimes A^\vee}(1_A \otimes t^{-1}_{B,B^\vee,A^\vee})(\kappa_{A^\vee B} \otimes 1_X)\kappa_{XA}$$

Hence we have expressed $\kappa_{X,A\otimes B}$ in terms of $\kappa_{X,A}$, $\kappa_{A^\vee,B}$, and the canonical isomorphisms t, t^{-1}.

(ii) Immediate by the inductive definition of the objects of (⊗N), and part **(i)** above.

(iii) As (⊗N) is freely generated, we can define d_X, c_X for all $X \in Ob(\otimes\mathrm{N})$ as in Definitions 4.3, Chapter 4. Then for all $X \in Ob(\otimes\mathrm{N})$, the naturality of κ_{XN} in X gives us $((1\otimes 1)\otimes d_X)\kappa_{NN}c_X = \kappa_{XN}d_Xc_X$, and by Lemma 4 of Chapter 4, $d_Xc_X = 1_X$. Therefore, $\kappa_{XN} = ((1\otimes 1)\otimes d_X)\kappa_{NN}c_X$. Hence our result follows.

(iv) By the coherence theorem for compact closed categories,

$$\epsilon_{A\otimes B} = \epsilon_B(\rho_{B^\vee}(1_{B^\vee} \otimes \epsilon_A)t_{B^\vee A^\vee A} \otimes 1_B)t_{B^\vee\otimes A^\vee,A,B},$$

or, using the definition of δ in terms of ϵ,

$$\epsilon_B(\delta_{B^\vee A}t_{B^\vee A^\vee A} \otimes 1_B)t_{B^\vee\otimes A^\vee,A,B} = \epsilon_{A\otimes B}.$$

Therefore, this implies that

$$\rho_X(1_X \otimes \epsilon_B(\delta_{B^\vee A}t_{B^\vee A^\vee A} \otimes 1_B)t_{B^\vee\otimes A^\vee,A,B}) = \rho_X(1_X \otimes \epsilon_{A\otimes B}),$$

and by definition of δ in terms of ϵ, this is

$$\delta_{XB}(1_X \otimes (\delta_{B^\vee A}t_{B^\vee A^\vee A} \otimes 1_B)t_{B^\vee\otimes A^\vee,A,B}) = \delta_{X,A\otimes B}.$$

Hence we have expressed $\delta_{X,A\otimes B}$ in terms of δ_{XB} and $\delta_{B^\vee A}$.

(v) Immediate from **(iv)** above, and the inductive definition of objects of (⊗N).

(vi) By the naturality of δ_{XN} in X, $d_X\delta_{NN}(c_X\otimes(1\otimes 1)) = d_Xc_X\delta_{XN}$, and by lemma 4 of Chapter 4, $d_Xc_X = 1_X$, and so $\delta_{XN} = d_X\delta_{NN}(c_X \otimes (1\otimes 1))$. Hence our result follows. □

Using these results, we are able to demonstate how the endomorphism monoid of a very self-similar object of a compact closed category has many of the same properties as a compact closed category.

Definitions 6.9

Let N be a very self-similar object of a compact closed category, $(\mathbf{T}, \otimes)$, and let $(\otimes\mathrm{N})$ be freely generated in $\mathbf{T}$, so the contraction map ϕ, as given in Definitions 4.3, Chapter 4, is well-defined, and the internal associativity and commutativity elements, which we denote t, t^{-1}, s satisfy $t = \phi(t_{NNN})$, $t^{-1} = \phi(t^{-1}_{NNN})$ and $s = \phi(s_{NN})$. We define the *internal pair-creation* and *internal pair-destruction* maps in terms of ϕ and κ and δ, by $\kappa = \phi(\kappa_{NN})$, $\delta = \phi(\delta_{NN})$ respectively.

Proposition 16 *Let* $(\mathbf{T}, \otimes)$ *be a compact closed category, denote the internalisation of the tensor product by* $\oplus$*, and let* ϕ*,* $s, t, t^{-1}, \kappa, \delta$ *be as defined above. Then*

(i) $\delta t^{-1}\kappa = 1$,

(ii) $\delta sts\kappa = 1$,

(iii) $((1 \oplus 1) \oplus a)\kappa = \kappa a$,

(iv) $a\delta = \delta(a \oplus (1 \oplus 1))$.

Proof Recall that the map ϕ preserves composition and maps $\otimes$ to $\oplus$, from Theorem 5 and Proposition 8 of Chapter 4. Our results are then as follows:

(i) By axiom 1 of Definitions 6.7, $\delta_{NN}t^{-1}_{NN^\vee N}\kappa_{NN} = 1_N$. However, by assumption, N is a self-dual object, and so $\delta_{NN}t^{-1}_{NNN}\kappa_{NN} = 1_N$. Therefore, we can apply ϕ to both sides of this, and deduce that $\phi(\delta_{NN}t^{-1}_{NNN}\kappa_{NN}) = \phi(1_N)$, and by definition of δ, κ, t^{-1}, this becomes $\delta t^{-1}\kappa = 1$.

(ii) By axiom 2 of Definitions 6.7, $\delta_{N^\vee N}s_{N^\vee \otimes N, N^\vee}t_{N^\vee NN^\vee}s_{N\otimes N^\vee, N^\vee}\kappa_{N^\vee N} = 1^\vee_N$, and as N is self-dual, $\delta_{NN}s_{N\otimes N,N}t_{NNN}s_{N\otimes N,N}\kappa_{NN} = 1_N$. Therefore, we can apply ϕ to both sides of the above, to deduce that $\phi(\delta_{NN}s_{N\otimes N,N}t_{NNN}s_{N\otimes N,N}\kappa_{NN}) = \phi(1_N)$ and so $\delta\phi(s_{N\otimes N,N})ts_{N\otimes N,N}\kappa = 1$, by definition of κ, δ, and t. However, $s_{N\otimes N,N} = (1 \oplus (1 \oplus 1))s((1 \oplus 1) \oplus 1)$ by the naturality of s_{XY} in X and Y, and so $\delta sts\kappa = 1$.

(iii) $((1_N \otimes 1_{N^\vee}) \otimes a)\kappa_{NN} = \kappa_{NN}a$, for arbitrary $a \in \mathbf{T}(N, N)$, by the naturality of κ_{XY} and δ_{XY} in X. Also, as N is self-dual, $((1_N \otimes 1_N) \otimes a)\kappa_{NN} = \kappa_{NN}a$, and we can apply ϕ to both sides of this equation, to deduce $\phi(((1_N \otimes 1_N) \otimes a)\kappa_{NN}) = \phi(\kappa_{NN}a)$. By definition of κ, δ, this becomes $((1 \oplus 1) \oplus a)\kappa = \kappa a$.

(iv) In a similar way to **(iii)**, $a\delta_{NN} = \delta_{NN}(a \otimes (1_{N^\vee} \otimes 1_N))$ for arbitrary $a \in \mathbf{T}(N, N)$, by the naturality of δ_{XY} in X, and as N is self-dual, $a\delta_{NN} = \delta_{NN}(a \otimes (1_N \otimes 1_N))$. Therefore, we can apply ϕ to both sides of the above to deduce that $\phi(a\delta_{NN}) = \phi(\delta_{NN}(a \otimes (1_N \otimes 1_N)))$. However, by definition of κ and δ, $a\delta = \delta(a \oplus (1 \oplus 1))$. Hence our result follows. □

Theorem 17 *Let* N *be a very strongly self-similar object of a compact closed category* $(\mathbf{T}, \otimes)$ *, and let* $(\otimes\mathrm{N})$ *be freely generated. Then the endomorphism monoid of* N *is a one-object compact* *closed category, without unit elements.*

Proof As N is strongly self-similar, the map $\oplus$ is a functor, and by Proposition 16 above, κ and δ satisfy the one-object analogues of the alternative axioms for a compact closed category. Also, the same proposition also implies the one-object analogues of the naturality of κ and δ. Hence our result follows. □

The above results motivate the following definitions:

6.5.4 Compact closed M-monoids

Definitions 6.10

We define a *compact closed M-monoid* to be an M-monoid $(M, \oplus)$ that has two distinguished maps κ and δ that satisfy, for all $a \in M$,

1. $\delta t^{-1}\kappa = 1$,
2. $\delta sts\kappa = 1$,
3. $((1 \oplus 1) \oplus a)\kappa = \kappa a$.
4. $a\delta = \delta(a \oplus (1 \oplus 1))$.

Proposition 18 *Let N be a very self-similar object of a compact closed category* $(\mathbf{T}, \otimes)$, *and let* $(\otimes \mathrm{N})$ *be freely generated in* $\mathbf{T}$. *Then the endomorphism monoid of N in* $(\otimes \mathrm{N})$ *is a compact closed M-monoid.*

Proof Immediate from the proof that $\mathbf{T}(N, N)$ is an M-monoid (Proposition 10, Chapter 4), and from Proposition 16 above. □

The following theorem demonstrates that our axiomatisation of compact closed M-monoids is the 'correct' one.

Theorem 19 *For any compact closed M-monoid $(M, \oplus)$, the tensor envelope of M is a compact closed category (without the unit object) with a very self-similar object N, satisfying* $\mathbf{K}_{\mathrm{M}}^{\oplus} \cong (\otimes \mathrm{N})$.

Proof We denote the object of $\mathbf{K}_{\mathrm{M}}^{\oplus}$ given by the idempotent $1 \in Ob(\mathbf{K}_{\mathrm{M}}^{\oplus})$ by N, for clarity. Then, by Proposition 17 and Theorem 20 of Chapter 4, N is a self-similar object of $\mathbf{K}_{\mathrm{M}}^{\oplus}$, and $(\otimes \mathrm{N}) \cong \mathbf{K}_{\mathrm{M}}^{\oplus}$.

We define left duals of the objects of $\mathbf{K}_{\mathrm{M}}^{\oplus}$ as follows:

- $1^\vee = 1$,
- $(e \oplus f)^\vee = f^\vee \oplus e^\vee$,

so by definition, N is a very self-similar object of $\mathbf{K}^{\oplus}_{\mathbf{M}}$.

Next, we define κ_{NN} and δ_{NN} by $\kappa_{NN} = \kappa$, $\delta_{NN} = \delta$, and by axiom 3 for a compact closed M-monoid, $((1 \oplus 1) \oplus 1)\kappa 1 = \kappa$ so $\kappa_{NN} \in \mathbf{K}^{\oplus}_{\mathbf{M}}(N, (N \otimes N) \otimes N)$. Similarly, by axiom 4 for a compact closed M-monoid, $1\delta = \delta = \delta(1 \oplus (1 \oplus 1)$ and so $\delta_{NN} \in \mathbf{K}^{\oplus}_{\mathbf{M}}(N \otimes (N \otimes N), N)$.

Also, $\delta_{NN} t^{-1}_{NN^\vee N} \kappa_{NN} = 1_N$ by axiom 1 for a compact closed M-monoid, and $\delta_{NN^\vee} sts \kappa_{N^\vee N} = 1^\vee_N$ by the self-duality of N, and by axiom 2 for a compact closed M-monoid. Therefore, κ_{NN} and δ_{NN} satisfy the axioms for the pair creation/annihilation morphisms of a compact closed category at N.

Finally, recall that by Theorem 15 of this Chapter, κ_{XY} and δ_{XY} are definable in terms of κ_{NN} and δ_{NN}, for all $X, Y \in (\otimes \mathrm{N})$. coherence theorem for symmetric monoidal categories, κ_{XY} and δ_{XY} satisfy the axioms of Definitions 6.7 Therefore, we can construct κ_{XY} and δ_{XY} for all $X, Y \in Ob(\mathbf{K}^{\oplus}_{\mathbf{M}})$ that satisfy the alternative axiom set for a compact closed category (without units).

Hence $\mathbf{K}^{\oplus}_{\mathbf{M}}$ is a compact closed category (without the unit object) that has a very self-similar object N satisfying $\mathbf{K}^{\oplus}_{\mathbf{M}} \cong (\otimes \mathrm{N})$, and this isomorphism preserves the compact closed structure. Hence our result follows. □

Corollary 20 *Let* $(\mathbf{T}, \otimes)$ *be a compact closed category with a very self-similar object* N *(where* $(\otimes \mathrm{N})$ *is freely generated), and let* $(M, \oplus)$ *be the M-monoid defined by* $\mathbf{T}(N, N)$*. Then* $\mathbf{K}^{\oplus}_{\mathbf{M}} \cong (\otimes \mathrm{N})$.

Proof Recall the isomorphisms $\Psi : \mathbf{K}^{\oplus}_{\mathbf{M}} \to (\otimes \mathrm{N})$ and $\Phi : (\otimes \mathrm{N}) \to \mathbf{K}^{\oplus}_{\mathbf{M}}$ of Definitions 4.13, Chapter 4. Then from Theorem 20 of Chapter 4, Φ and Ψ also preserve the monoidal structure, and it is immediate from the above that κ_{NN} in $\mathbf{K}^{\oplus}_{\mathbf{M}}$ is the same as $\kappa_{\Psi(N),\Psi(N)}$ in $(\otimes \mathrm{N})$. Therefore, our result follows from the definition of κ_{XY} and δ_{XY} in terms of κ_{NN} and δ_{NN} of Theorem 15. □

6.6 Constructing very self-similar objects of compact closed categories

Let $(\mathbf{V}, \otimes)$ be a traced symmetric monoidal category. We demonstrate a routine method of constructing very self-similar (resp. very strongly self-similar) objects of **IntV** from self-similar (resp. strongly self-similar) objects of **V**. We will use this method, together with the results from the previous section, to construct a one-object compact closed category.

For our results and construction, we first require the following:

Lemma 21 *Let* **IntV** *be the compact closed category derived from a traced symmetric monoidal category* **V** *using the method of Theorem 9 of this Chapter, and let* $F : (X, U) \to (Y, V)$ *and* $G : (Y, V) \to (Z, W)$ *be morphisms of* **IntV** *specified by morphisms* $f = a \otimes b : X \otimes V \to Y \otimes U$ *and* $g = c \otimes d : Y \otimes W \to Z \otimes V$ *in* **V**. *Then* $GF \in \mathbf{IntV}((X, U), (Z, W))$ *is specified by* $(ca \otimes bd) \in \mathbf{V}(X \otimes W, Z \otimes U)$.

Proof We use the standard notation for the monoidal functor, canonical isomorphisms, and trace of **V**. Then from Theorem 9, GF is specified by the trace at V of the composite

$$C_{G,F} = t_{YWU}(1_Y \otimes s_{UW})t^{-1}_{YUW}(f \otimes 1_W)t_{XVW}(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

However, by definition, $f = a \otimes b$ and $g = c \otimes d$, so $C_{G,F} =$

$$t_{ZUV}(1_Z \otimes s_{VU})t^{-1}_{ZVU}((c \otimes d) \otimes 1_U)t_{YWU}(1_Y \otimes s_{UW})$$

$$t^{-1}_{YUW}((a \otimes b) \otimes 1_W)t_{XVW}(1_X \otimes s_{W,V})t^{-1}_{XWV}.$$

Then by axioms 1 and 2 for the canonical isomorphisms of a symmetric monoidal category, (Definitions 1.2, Chapter 1), $C_{G,F} =$

$$t_{ZUV}(1_Z \otimes s_{VU})(c \otimes (d \otimes 1_U))(1_Y \otimes s_{UW})(a \otimes (b \otimes 1_W))(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

$$= t_{ZUV}(c \otimes (1_U \otimes d))(a \otimes (b \otimes 1_W))(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

$$= t_{ZUV}(ca \otimes (b \otimes d))(1_X \otimes s_{W,V})t^{-1}_{XWV}$$

$$= ((ca \otimes 1_U) \otimes 1_V)((1_X \otimes b) \otimes d)t_{XVW}(1_X \otimes s_{WV})t^{-1}_{XWV}.$$

and the trace of this is, by the naturality of $Tr^U_{X,Y}$ in X and Y,

$$Tr(C_{G,F}) = (ca \otimes 1_U)Tr(((1_X \otimes b) \otimes d)t_{XVW}(1_X \otimes s_{WV})t^{-1}_{XWV})$$

$$= (ca \otimes 1_U)(1_X \otimes b)Tr(((1_X \otimes 1_Z) \otimes d)t_{XVW}(1_X \otimes s_{WV})t^{-1}_{XWV})$$

and by the naturality of s and t,

$$Tr(C_{G,F}) = (ca \otimes 1_U)(1_X \otimes b)Tr((t_{XVZ}(1_X \otimes s_{ZV})t^{-1}_{XZV}((1 \otimes d) \otimes 1)))$$

which, by the naturality of $Tr^U_{X,Y}$ in X and Y, gives

$$Tr(C_{G,F}) = (ca \otimes 1_U)(1_X \otimes b)Tr((t_{XVZ}(1_X \otimes s_{ZV})t^{-1}_{XZV}))(1_X \otimes d)$$

$$= (ca \otimes 1_U)(1_X \otimes bd) = (ca \otimes bd),$$

by axiom 3 for a categorical trace. Hence our result follows. □

Proposition 22 *Let N be a (strongly) self-similar object of a symmetric traced monoidal category, $\mathbf{V}$. Then $Q = (N, N)$ is a very (strongly) self-similar object of* **IntV**.

Proof Denote the contraction and division morphisms of N in $\mathbf{V}$ by $c : N \to N \otimes N$ and $d : N \otimes N \to N$ respectively, and let $C : Q \otimes Q \to Q$, and $D : Q \to Q \otimes Q$ be the maps in **IntV** specified by $d \otimes c$ and $c \otimes d$ respectively. Then, from the definition of $\otimes$ in **IntV** (from Theorem 9), $(N, N) \otimes (N, N) = (N \otimes N, N \otimes N)$, and from the characterisation of composition in Lemma 21 above, $DC = dc \otimes dc = 1_{(N \otimes N) \otimes (N \otimes N)}$, which specifies the identity at $Q \otimes Q$ in **IntV**. Similarly, if N is a strongly self-similar object, $CD = cd \otimes cd = 1_{N \otimes N}$, which specifies the identity of Q in **IntV**. Hence Q is a (strongly) self-similar object of **IntV**.

Next, consider $X, Y \in Ob(\mathbf{IntV})$, where $X = (A, B), Y = (R, S)$ for some $A, B, R, S \in Ob(\mathbf{V})$. Then from Theorem 9, $X \otimes Y = (A \otimes R, S \otimes B)$ and so, from the description of the dual on objects from Theorem 9, $(X \otimes Y)^\vee = (S \otimes B, A \otimes R)$. However, $X^\vee = (B, A)$ and $Y^\vee = (S, R)$, and so $Y^\vee \otimes X^\vee = (S \otimes B, A \otimes R) = (X \otimes Y)^\vee$. Hence for any object N of a traced symmetric monoidal category $\mathbf{V}$, $(N, N)^\vee = (N, N)$, and so $Q = (N, N)$ is self-dual. This also gives, as a special case, $(X \otimes Y)^\vee = Y^\vee \otimes X^\vee$, for all $X, Y \in (\otimes \mathrm{N})$. Hence the (strong) self-similarity of N in $\mathbf{V}$ implies that (N, N) is a very (strong) self-similar object of **IntV**. □.

The above proposition gives us a routine method of constructing very (strongly) self-similar objects of **IntV** from (strongly) self-similar objects of **V**. We use this to give an explicit example of a one-object compact closed category (without units).

6.7 An explicit description of a one-object compact closed inverse category

We demonstrate how the results of Chapter 4 on the natural numbers in the category of partial bijective maps allow us to construct very strongly self-similar objects in the category **IntRel**, and use this to construct one-object compact closed categories, and give an explicit description of the composition, monoidal structure, and duality. We will demonstrate in Chapter 8 how this is the underlying structure of Girard's cut-elimination procedure in the 'Geometry of Interaction

1' paper, [20].

For our construction, we first require the following result:

Lemma 23 *Let X be any object in the category of relations (resp. partial bijective maps). Then the tensor category of X, denoted* $(\sqcup\mathbf{X})$, *is freely generated in* **Rel** *(resp.* **Inj***).*

Proof Consider $A, B, A', B' \in Ob(\mathbf{Rel})$ satisfying $A \sqcup B = A' \sqcup B'$. Then we can write this explicitly as $(A \times \{0\}) \cup (B \times \{1\}) = (A' \times \{0\}) \cup (B' \times \{1\})$. However, it is immediate from this that $A = A'$ and $B = B'$. Therefore, as this result follows for any $A, B \in Ob(\mathbf{Rel})$, it will clearly follow for any $A, B \in Ob(\sqcup\mathbf{X})$. Hence any object of $(\sqcup\mathbf{X})$ is uniquely determined by a unique binary tree with nodes labelled by X, and so $(\sqcup\mathbf{X})$ is freely generated in **Rel**. The result for **Inj** follows immediately from the above, and the fact that **Inj** is a subcategory of **Rel**. □

Theorem 24 *The endomorphism monoid of* $(\mathbb{N}, \mathbb{N})$ *in* **IntInj** *is a one-object compact closed inverse category without units..*

Proof From Lemma 11 of Chapter 5, $\mathbb{N}$ is a strongly self-similar object of $(\mathbf{Inj}, \sqcup)$. Therefore, from Proposition 23 above, $(\mathbb{N}, \mathbb{N})$ is a very strongly self-similar object of **IntInj**. Hence, as $(\sqcup\mathbb{N})$ is freely generated in **Inj**, $(\sqcup(\mathbb{N}, \mathbb{N}))$ is freely generated in **IntInj**, and so by Theorem 17 above, the endomorphism monoid of $(\mathbb{N}, \mathbb{N})$ in **IntInj** is a one-object compact closed strong M-monoid. □

For clarity, we will denote $\mathbf{IntInj}((\mathbb{N}, \mathbb{N}), (\mathbb{N}, \mathbb{N}))$ by $\mathbb{F}$. An explicit description of $\mathbb{F}$, using Theorem 9 of this Chapter is as follows:

The elements of $\mathbb{F}$ are 2×2 matrices of partial injective maps on the natural numbers representing a partial injective map on $\mathbb{N} \sqcup \mathbb{N}$, so $f \in \mathbb{F}$ is of the form $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where elements in the same row of the matrix have disjoint images, and elements in the same column of the matrix have disjoint domains. Composition of matrices of this form is given by

$$\begin{pmatrix} e & f \\ g & h \end{pmatrix} \circ \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} e(bg)^*a & f \vee e(bg)^*bh \\ c \vee d(gb)^*ga & d(gb)^*h \end{pmatrix}$$

and the inverses are given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} a^{-1} & c^{-1} \\ b^{-1} & d^{-1} \end{pmatrix}$$

The *internal tensor homomorphism* is given as follows: Given elements of $\mathbb{F}$,

$$f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; g = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$$

then the internal tensor homomorphism, which we denote by $\oplus$ is given by

$$f \oplus g = \begin{pmatrix} p^{-1}ap \vee q^{-1}rq & q^{-1}sp \vee p^{-1}bq \\ p^{-1}tq \vee q^{-1}cp & p^{-1}up \vee q^{-1}dq \end{pmatrix}.$$

Finally, the *dual on morphisms* is defined as follows: Given

$$f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

in $\mathbb{F}$, then then $f^{\vee}$ is given by

$$f^{\vee} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}$$

The monoid $\mathbb{F}$ also has an alternative composition $\cdot$, as shown in Theorem 12. This is given by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} r & s \\ t & u \end{pmatrix} = \begin{pmatrix} a \vee b(rd)^{*}rc & b(rd)^{*}s \\ t(dr)^{*}c & u \vee t(dr)^{*}ds \end{pmatrix}.$$

If we denote the elements of $\mathbb{F}$, together with this composition by $\mathbb{F}'$, it is immediate from Theorem 12 that $\mathbb{F} \cong \mathbb{F}'$.

Chapter 7

Linear logic and the Geometry of Interaction I

7.1 Introduction

It is well-known ([38]) that Cartesian closed categories are models of typed lambda calculus and one-object Cartesian closed categories (C-monoids) are models of untyped lambda calculus, and in Chapter 4 we have shown how one-object analogues of the 'internal hom.' for a Cartesian closed category can be constructed from self-similar objects of Cartesian closed categories.

Our claim is that compact closure is the correct form of categorical closure to model a variation on lambda calculus; we turn our attention to the *Polymorphic Lambda Calculus*, or *Girard's System F* – a lambda calculus type system whose computing power lies between that of the typed and the untyped lambda calculii (see [23] for details of its construction). It is a type-based system, where the operations of application and abstraction are applicable to types, as well as functions.

In [20], J.-Y. Girard claims that *multiplicative linear logic* (see [17] for details of the full Linear Logic system[1]) has the same computing power as the polymorphic lambda calculus, and introduces a model of (restricted) multiplicative linear logic in terms of matrices of operators from $B(l^2)$; the C^*-algebra of bounded linear operators on the Hilbert space. He also claims that the polymorphic lambda calculus can be embedded in this restriction of multiplicative linear logic.

Over the next two chapters, we will show how his model is expressible in terms of the tools developed in the previous 6 chapters. However, this chapter is purely expository. We give a

[1]Linear logic itself was first introduced by J.-Y. Girard in [17]. It was originally based on a decomposition of Scott Domains (a classical model of untyped lambda calculus, see [38]); however, we follow the approach of [23], and introduce it via a restriction on the structural rules of Gentzen's sequent calculus [14].

basic outline of linear logic — in particular, the multiplicative fragment and its cut-elimination algorithm, and present the (original form of the) tools used in the 'Geometry of Interaction 1' system. In the next chapter, we will demonstrate how all the operators are defined in terms of (an embedding into $B(l^2)$ of) the disjoint closure of P_2 in $I(\mathbb{N})$, and have a close connection with the canonical isomorphisms of the two distinct M-monoid structures on $I(\mathbb{N})$ we constructed in Chapter 5. We also show how the dynamical model of the cut-elimination algorithm is an expression of the internalised trace, used to generate the composition in the one-object compact closed category $\mathbb{F}$.

Note that the following introduction to linear logic is basically the same as that found in [23].

7.2 Sequent calculus

Definitions 7.1

A *sequent* is defined to be a term of the form $A \vdash B$, where $A = a_1, a_2, \ldots, a_n$ and $B = b_1, \ldots, b_m$ are finite sequences of formulæ of some (unspecified) formal language. Intuitively, the sequent $A \vdash B$ can be considered to be the statement that the assumption of all the formulæ on the left hand side of $\vdash$ will allow us to deduce all[2] the formulæ on the right hand side of $\vdash$. So, classically, the comma on the LHS interprets as conjunction, the comma on the RHS interprets as disjunction, and the $\vdash$ interprets as implication. Finally, if A is empty, the sequent asserts the disjunction of $\{b_i\}$, if B is empty, the sequent asserts the negation of the conjunction of $\{a_i\}$; if both A and B are empty, then the sequent asserts a contradiction. Note that we adopt the convention of using upper case letters for sequences of formulæ , and using lower case letters for single formulæ .

The *sequent calculus* is a method of formally manipulating sequents of this form. It uses the symbols $\vee \ \wedge \ \Rightarrow \ \neg \ \forall \ \exists \ [\ /\ ]$. All these symbols are intended to model their informal use in mathematical deduction, as conjunction, disjunction, implication, negation, universal and existential quantification, and substitution for a free variable, respectively.

A *logical rule*, in this context, is a method of constructing new sequents from old. An application of a logical rule R to the sequent $A \vdash B$ to get the sequent $A' \vdash B'$ is represented as follows:

$$\frac{A \vdash B}{A' \vdash B'} R.$$

[2]Of course, this statement has many different interpretations, depending on the exact formalisation of 'all' and 'deduce'.

$A \vdash B$ is called the *premise*, or *assumption* of the (application of the) rule, and $A' \vdash B'$ is called the *deduction* or *conclusion* of the (application of the) rule. A *proof* of a sequent $A \vdash B$ from a set of sequents $\{X_i \vdash Y_i\}$ is then a series of applications of logical rules that have $\{X_i \vdash Y_i\}$ as assumptions, and $A \vdash B$ as conclusion. It would not be useful to present the full list of axioms for Gentzen's sequent calculus here (see [14] for details of the whole system). However, logical rules which are important to the following discussion are the left and right *weakening rules*,

$$\frac{A \vdash B}{A \vdash B, c} RW,$$

$$\frac{A \vdash B}{A, c \vdash B} LW,$$

the left and right *contraction rules*,

$$\frac{A, x, x \vdash B}{A, x \vdash B} LC,$$

$$\frac{A \vdash B, y, y}{A \vdash B, y} RC,$$

and the *Cut Rule*,

$$\frac{A \vdash C, B \quad A', C \vdash B'}{A, A' \vdash B, B'} Cut.$$

An important result of Gentzen, [14], is that any sequent that is proved using a proof that uses *Cut* can also be proved using a proof that does not use *Cut*. The procedure for doing so is referred to as the *Cut-elimination algorithm.*

7.2.1 Cut-elimination in sequent calculus

The full set of rules for cut-elimination in sequent calculus, and an algorithm for their application to any proof involving cuts, is given in [14]; however, these lead to the following:
Consider the application of the cut rule,

$$\frac{A \vdash C, B \quad D, C \vdash E}{A, D \vdash B, E} Cut.$$

This could derive from the following proof involving weakening,

$$\frac{\frac{A \vdash B}{A \vdash C, B} RW \quad \frac{D \vdash E}{D, C \vdash E} LW}{A, D \vdash B, E} Cut$$

Applying the cut-elimination algorithm to this will give us that the above proof then reduces to either

$$\frac{A \vdash B}{A, D \vdash B, E} LW, RW$$

or

$$\frac{D \vdash E}{A, D \vdash B, E} LW, RW$$

These are two very different proofs. So, cut-elimination is not deterministic. In what follows, we consider the restriction of the structural rules that give us this non-determinism (the weakening and contraction rules); this will lead to *Linear Logic.*

7.3 Introduction to linear logic

We follow the introduction to linear logic given in [23], where the basic idea behind (this approach to) linear logic is that we are forbidding the structural rules that lead to the non-deterministic behavior of cut-elimination. These are the contraction and weakening rules. The idea of (this approach to) linear logic was to construct an expanded type system in which these operations are made explicit. This was done by imposing *linearity.*

7.3.1 Linear systems

Definitions 7.2

We do not formally define linearity; however, intuitively, it can be thought of as follows:

A formal system[3] is called *linear* if each input to a process is used exactly once in producing the output of the process. In this case, enforcing linearity will require revising the weakening and contraction rules (equivalent to getting rid of the non-determinism in cut-elimination). We reconsider the operations of Gentzen's sequent calculus under these assumptions.

7.3.2 Reconsidering sequent calculus operations in terms of linearity

To start with, consider the conjunction operator, $\wedge$. There are two possible ways the conjunction is used:

1. Both components of the pair of terms input to the construction are used; this means that we can no longer use the projection operators π_1 and π_2, as this would involve discarding inputs. Also, we can no longer form the diagonal $x \mapsto x \wedge x$, as this would involve the copying of inputs. This is called *multiplicative conjunction*, written $\otimes$.

2. We only use one component of the pair given as input to the construction. So, we can use the projection operators, but we can only use one of them, and only once. The choice of which projection is taken is made by the process that takes this conjunction as its input. This conjunction is called *additive conjunction*, and is denoted &. An analogy can be made with

[3] Also another term that we do not define formally.

Heisenberg's Uncertainty Principle — a particle may have both position and momentum, but we can only observe one of these; the observer makes the choice of which one, and the act of observing one removes the possibility of observing the other[4].

The disjunction operators are defined in terms of the *linear negation*, $\perp$, which is required to behave in the same way as the classical negation, $(A^{\perp})^{\perp} = A$, and satisfy analogues of De Morgan's laws. The dual of the multiplicative conjunction is called the *tensor sum*, and is denoted $+$ (it is also sometimes denoted by an upside down &), and is defined in terms of $\otimes$ and $^{\perp}$ by

$$(A \otimes B)^{\perp} = A^{\perp} + B^{\perp} \ , \ (A + B)^{\perp} = A^{\perp} \otimes B^{\perp}.$$

Similarly, we have the dual of &, the *direct sum*, denoted $\oplus$, defined in terms of $\oplus$ and $^{\perp}$ by

$$(A \& B)^{\perp} = A^{\perp} \oplus B^{\perp},$$

$$(A \oplus B)^{\perp} = A^{\perp} \& B^{\perp}.$$

We can then define *linear implication*, the linear analogue of $\Rightarrow$, which we denote $\rightarrow$ (note that this is not a standard notation). This is defined by

$$(A \otimes B^{\perp})^{\perp} = A^{\perp} + B = A \rightarrow B$$

The interpretation of this can be thought of as 'B can be deduced in a linear manner from A'. However, as linear implication is defined in terms of the other operators, it is not explicitly modelled in Girard's Geometry of Interaction system, and hence plays no further part in this discussion.

7.3.3 The one-sided sequent convention, and logical rules

Using the linear analogues of DeMorgan's laws, we can convert a two-sided sequent $A_1, A_2, \ldots A_n \vdash B_1, B_2, \ldots B_m$ into a one-sided sequent $\vdash A_1^{\perp}, A_2^{\perp}, \ldots, A_n^{\perp}, B_1, B_2, \ldots, B_m$. This is merely done as a method of simplifying notation. The structural and logical rules can then be written in a much more concise form, as follows:

- The *exchange rule*; this is the only structural rule, and is given by

$$\frac{\vdash A, C, D, B}{\vdash A, D, C, B} \times change$$

  This can be avoided if sequents are considered to be members of a free commutative semigroup, instead of a free semigroup (however, we do not follow this approach). Alternatively,

[4] Of course, this is a gross oversimplification

the exchange rule is not a part of non-commutative linear logic, although we do not consider this.

- The *axiom rule*; this is

$$\frac{}{\vdash A, A^{\perp}} Axiom$$

- The *cut rule*;

$$\frac{\vdash C, A \quad \vdash C^{\perp}, B}{\vdash A, B} Cut$$

- The *multiplicative conjunction* rule; this is

$$\frac{\vdash C, A \quad \vdash D, B}{\vdash C \otimes D, A, B} \otimes$$

- The *additive conjunction* rule;

$$\frac{\vdash C, A \quad \vdash D, A}{\vdash C \& D, A} \&$$

- The *tensor sum* rule;

$$\frac{\vdash C, D, A}{\vdash C + D, A} +$$

- The *first additive disjunction* rule;

$$\frac{\vdash C, A}{\vdash C \oplus D, A} 1\oplus$$

- The *second additive disjunction* rule;

$$\frac{\vdash D, A}{\vdash C \oplus D, A} 2\oplus$$

We also have the *units*; these are the linear analogues of the T and F in classical logic. So we have 1 for $\otimes$, $\perp$ for $+$, $\top$ for $\&$, and 0 for $\oplus$. These satisfy

$$1^{\perp} = \perp, \quad \perp^{\perp} = 1, \quad \top^{\perp} = 0, \quad 0^{\perp} = \top.$$

The rules for their introduction are as follows:

$$\frac{}{\vdash 1} 1,$$

$$\frac{\vdash A}{\vdash \perp, A} \perp,$$

$$\frac{}{\vdash \top, A} \top.$$

Note that there is no rule for the introduction of 0.

7.3.4 Summary of LL operators

The linear logic connectives and units can be split up into two self-contained groups; the *multiplicative* and the *additive* groups.

- In the multiplicative group, we have $\otimes$, the conjunction, $+$, the disjunction, 1, the analogue of $True$, and $\perp$, the analogue of $False$.
- In the additive group, we have &, the conjunction, $\oplus$, the disjunction, $\top$, the analogue of $True$, and 0, the analogue of $False$.

(Note that both groups are closed under the linear negation operator, $(\ )^{\perp}$). The differences between the multiplicative and additive fragments can be illustrated using the conjunctions, as follows:

For the multiplicative conjunction:

$$\frac{\vdash C, A \quad \vdash D, B}{\vdash C \otimes D, A, B}\otimes$$

For the additive conjunction:

$$\frac{\vdash C, A \quad \vdash D, A}{\vdash C \& D, A}\&$$

7.3.5 The exponential operators

The concepts of weakening and contraction are not lost entirely, as mentioned earlier, they are merely made explicit each time they are used. Two extra operators are introduced, ! and ?, called *of course*, and *why not* respectively[5]. They are required to satisfy the following DeMorgan type equalities:

$$(!A)^{\perp} = ?(A^{\perp})\ ,\ (?A)^{\perp} = !(A^{\perp}),$$

and have the following logical rules:

$$\frac{\vdash A}{\vdash ?b, A}\ Weakening$$

$$\frac{\vdash ?A, b}{\vdash ?A, !b}\ !\ Introduction$$

$$\frac{\vdash ?b, ?b, A}{\vdash ?b, A}\ Contraction$$

$$\frac{\vdash b, A}{\vdash ?b, A}\ Dereliction$$

[5]Note that the same notation is used for these two logical operators, and the two monoid homomorphisms derived from the decomposition of the internalised Cartesian product at $\mathbb{N}$ (see Section 5.7 of Chapter 4). The coincidence of notation is not accidental, and it is hoped that the context will make the operator used apparent.

7.4 Multiplicative linear logic

The first two parts of the Geometry of Interaction program ([19, 20, 21, 22]) that we are seeking to model consists only of operators corresponding to the multiplicative fragment of linear logic. There are two reasons for this; firstly, the system is based on *proof nets* (see [17]), which are only properly formalised for the multiplicative connectives and constants, and secondly, the computing power of multiplicative linear logic is equivalent to the computing power of J-Y Girard's 'system F', also known as 'Polymorphic Lambda Calculus', (see [16]), which we wish to study for its categorical models.

7.4.1 Cut-elimination in MLL

We present the cut-elimination procedure for the multiplicative fragment of linear logic, as found in [17]. The exchange rule will be used without being stated explicitly each time.

The rules for cut-elimination are as follows:

1. Given a proof π_1 with conclusion $\vdash a, B$, and an axiom rule with conclusion $\vdash a^\perp, a$, then the cut between them,
$$\frac{\vdash a, B \quad \vdash a^\perp, a}{\vdash B, a}Cut$$
is replaced by the proof π_1 with conclusion $\vdash B, a$.

2. Given a proof π_1 with conclusion $\vdash b, D$, a proof π_2, with conclusion $\vdash c, E$, and a proof π_3 that ends in
$$\frac{\vdash b^\perp, c^\perp, F}{\vdash b^\perp + c^\perp, F} + Intro,$$
then a cut between
$$\frac{\vdash b, D \quad \vdash c, E}{\vdash b \otimes c, D, E} \otimes Intro$$
and the conclusion of π_3, given by
$$\frac{\vdash b \otimes c, D, E \quad \vdash b^\perp + c^\perp, F}{\vdash D, E, F}Cut$$
is replaced by the following cuts on the conclusions of π_1, π_2, π_3
$$\frac{\vdash b, D \quad \frac{\vdash c, E \quad \vdash b^\perp, c^\perp, F}{\vdash b^\perp, E, F}Cut}{\vdash D, E, F}Cut.$$

3. Given a proof π_1 ending in the following application of a $\forall$ introduction rule,

$$\frac{\vdash b, D}{\vdash \forall \alpha B, D}\forall Intro,$$

and a proof π_2 ending in the following application of an $\exists$ elimination rule,

$$\frac{\vdash b^{\perp}, E}{\vdash b^{\perp}[c/\alpha], E}\exists Elim,$$

then the following cut between their conclusions,

$$\frac{\vdash \forall \alpha b, D \quad \vdash b^{\perp}[c/\alpha], E}{\vdash D, E}$$

is replaced by the following cut

$$\frac{\vdash b, D \quad \vdash b^{\perp}, E}{\vdash D, E}Cut.$$

4. Given a proof π_1 ending in the following ! introduction rule

$$\frac{\vdash b, ?D}{\vdash !b, ?D}!intro,$$

and a proof π_2, ending in the following contraction rule

$$\frac{\vdash ?b^{\perp}, ?b^{\perp}, E}{\vdash ?b^{\perp}, E}Contraction,$$

then a cut between them, given by

$$\frac{\vdash !b, ?D \quad \vdash ?b^{\perp}, E}{\vdash ?D, E}Cut$$

is replaced by the following on the conclusions of π_1 and π_2

$$\frac{\vdash !b, ?D \quad \frac{\vdash !b, ?D \quad \vdash ?b^{\perp}, ?b^{\perp}, E}{\vdash ?b^{\perp} ?D, E}Cut}{\frac{\vdash ?D, ?D, E}{\vdash ?D, E}Contraction}Cut.$$

5. Given a proof π_1 ending in an application of the following ! introduction rule,

$$\frac{\vdash b, ?D}{\vdash !b, ?D}!intro,$$

and a proof π_2 ending in an application of the following dereliction rule,

$$\frac{\vdash b^{\perp}, E}{\vdash ?b^{\perp}, E}Dereliction,$$

then the cut between their conclusions,

$$\frac{\vdash !b, ?D \quad \vdash ?b^{\perp}, E}{\vdash ?D, E}Cut$$

is replaced by the following cut

$$\frac{\vdash b, ?D \quad \vdash b^{\perp}, E}{\vdash ?D, E}Cut.$$

6. Given a proof π_1 ending in the following application of an ! introduction rule,

$$\frac{\vdash b, ?D}{\vdash !b, ?D} !Intro,$$

and a proof π_2 ending in the following application of an ! introduction rule,

$$\frac{\vdash ?b^{\perp}, ?E, c}{\vdash ?b^{\perp} ?E, !c} !Intro$$

then the cut between their conclusions,

$$\frac{\vdash !b, ?D \quad \vdash ?b^{\perp} ?E, !c}{\vdash ?D, ?E, !c} Cut$$

is replaced by the following

$$\frac{\vdash !b, ?D \quad \vdash ?b^{\perp}, ?E, c}{\frac{\vdash ?D, ?E, c}{\vdash ?D, ?E, !c} !Intro} Cut$$

The proof that this procedure for cut-elimination terminates, and leads to a valid cut-free proof of MLL can be found in [17], where estimates for its efficiency and the number of steps before termination are also found.

7.5 The 'Geometry of Interaction' programme

7.5.1 Introduction

The Geometry of Interaction program was introduced by J-Y. Girard in [19]. Its aim was to model logical deductions, via cut-elimination in linear logic, in terms of dynamic processes, and so remove the dependence of logical systems on syntactic rules. Although this was a philosophical idea — and the paper [19] is very philosophical in approach — it was motivated by the discovery of proof nets, in which variable names are unimportant [17], and the paper [18], in which cut-elimination in proof nets (for the multiplicative case only, and without quantifiers, exponentials or constants) was modelled by means of iterations of finite permutations.

For our purposes, we present the system found in [20] purely as a formal system, without reference to the program laid out in [19]. We refer to the formal system presented in [20] as GOI1, and consider the translation of proofs from multiplicative linear logic into it.

7.5.2 MLL in The Geometry of Interaction, and its restrictions

The formal system presented in [20], which we refer to as GOI1, consists of two parts; the first part is a representation of a limited class of multiplicative linear logic proofs as (finite) matrices over the C^* algebra of bounded linear operators on a Hilbert space, and the second part is a representation of the cut-elimination process for these matrices in terms of the solution of an equation in $B(l^2)$, called the 'Resolution Formula'.

The restrictions on the types of proofs representable are found in [20], and are as follows:

1. The system is constant-free; i.e. it does not use the multiplicative analogues of True and False, 1 and $\perp$ respectively.

2. The system requires MLL proofs to memorise the formulæ that cuts are made on. So, a sequent $\vdash X$ that is proved using cuts on the formulæ c_1, c_2, c_3 is written $\vdash [c_1, c_2, c_3], X$. However, the symbols c_1, c_2, c_3 play no further part in the proof.

3. The system allows no a priori assumptions; we cannot start with an assumption $\vdash b$, and use it to deduce a conclusion; the only way of introducing formulæ is via the axiom link,

$$\overline{\vdash b^{\perp}, b}\,Axiom.$$

(This restriction is, however, common to many logical systems).

4. The context of an !-introduction rule is empty. This is a technical restriction on the form of the proofs, and its implications are discussed in [20].

The restriction on how well the resolution formula models cut-elimination is as follows:

The result of applying cut-elimination to a proof Γ *must result in a proof whose conclusion is a cut-free proof that does not contain* ? *or* $\exists$.

However, this is not as serious a restriction as it first appears; the execution formula models cut-elimination correctly when a proof involves ? or $\exists$, but not in the conclusion of the sequent. In [20], p. 239-241, this restriction is discussed, and a method of modelling the natural numbers (translated into MLL from Girard's system F) is given that does not involve ? or $\exists$ as the conclusion of a proof.

7.6 The $B(l^2)$ representation of MLL

We present the representation of MLL proofs in terms of operators from the C^*-algebra $B(l^2)$, as found in the first Geometry of Interaction paper, [20]. Of course, all proofs are subject to the above restrictions.

The translation is given inductively; the representation of axiom links is given, and the procedures for representing the application of logical rules to proofs are given. Note that all matrices are finite; however, the formulæ modelling the cut-elimination procedure compose matrices of different orders; so (for the purposes of this chapter), we consider all finite matrices over $B(l^2)$ to be infinite matrices with a finite number of non-zero entries. To represent the logical rules as matrices over $B(l^2)$, we first require the following operation:

Definitions 7.3

Given two $n \times n$ matrices over $B(l^2)$,

$$A = \begin{pmatrix} a_{00} & a_{01} & \dots & a_{0,n-2} & a_{0,n-1} \\ a_{10} & a_{11} & & a_{1,n-2} & a_{1,n-1} \\ \dots & & & & \dots \\ a_{n-2,0} & a_{n-2,1} & & a_{n-2,n-2} & a_{n-2,n-1} \\ a_{n-1,0} & a_{n-1,1} & \dots & a_{n-1,n-2} & a_{n-1,n-1} \end{pmatrix}$$

and

$$B = \begin{pmatrix} b_{00} & b_{01} & \dots & b_{0,n-2} & b_{0,n-1} \\ b_{10} & b_{11} & & b_{1,n-2} & b_{1,n-1} \\ \dots & & & & \dots \\ b_{n-2,0} & b_{n-2,1} & & b_{n-2,n-2} & b_{n-2,n-1} \\ b_{n-1,0} & b_{n-1,1} & \dots & b_{n-1,n-2} & b_{n-1,n-1} \end{pmatrix}$$

then their *shuffle*, written $Sh(A, B)$ is defined by $Sh(A, B) =$

$$\begin{bmatrix} a_{00} & 0 & a_{01} & 0 & \dots & a_{0,n-2} & 0 & a_{0,n-1} & 0 \\ 0 & b_{00} & 0 & b_{01} & \dots & 0 & b_{0,n-2} & 0 & b_{0,n-1} \\ a_{10} & 0 & a_{11} & 0 & & a_{1,n-2} & 0 & a_{1,n-1} & 0 \\ 0 & b_{10} & 0 & b_{11} & & 0 & b_{1,n-2} & 0 & b_{1,n-1} \\ \dots & \dots & & & & & & \dots & \dots \\ a_{n-2,0} & 0 & a_{n-2,1} & 0 & & a_{n-2,n-2} & 0 & a_{n-2,n-1} & 0 \\ 0 & b_{n-2,0} & 0 & b_{n-2,1} & & 0 & b_{n-2,n-2} & 0 & b_{n-2,n-1} \\ a_{n-1,0} & 0 & a_{n-1,1} & 0 & \dots & a_{n-1,n-2} & 0 & a_{n-1,n-1} & 0 \\ 0 & b_{n-1,0} & 0 & b_{n-1,1} & \dots & 0 & b_{n-1,n-2} & 0 & b_{n-1,n-1} \end{bmatrix}$$

When the two matrices are of different sizes, then an outer row/column of zeros is put onto the smaller, until they are of the same size.

The logical rules are then represented in the following manner:

First, a fixed representation of P_2 in $B(l^2)$ is chosen. We will abuse notation, and (in this chapter) identify the elements of P_2 with their images under this embedding. This representation is required to be weak; that is, it is required to satisfy $p_1^{-1}p_1 + p_2^{-1}p_2 = e < 1$. (Note that we denote the generators of the weak embedding by p_1 and p_2, rather than p and q. This is to emphasise that we are using a weak embedding).

Then a fixed orthogonal basis set $\{b_i\}_{i=0}^{\infty}$ for l^2 is chosen, along with a bijection $[\ ,\ ]$ from $\mathbb{N}\times\mathbb{N}$ to $\mathbb{N}$. This then induces a map $\beta : l^2 \times l^2 \to l^2$ by $\beta(b_i, b_j) = b_{[i,j]}$, and the system then requires a map $\otimes : B(l^2) \times B(l^2) \to B(l^2)$ that satisfies $(F \otimes G)(\beta(x, y)) = \beta(F(x), G(y))$. However [20] does not give an explicit description of this.

The logical rules are then represented as follows:

1. An axiom link is represented as the 2×2 antidiagonal matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

2. Given matrices S, T, of orders $(a \times a)$ and $(b \times b)$ respectively, representing proofs of $\vdash [X], A, b$, and $\vdash [Y], C, d$, then an application of the $\otimes$ rule, to get a matrix representing the proof ending in the deduction

$$\frac{\vdash [X], A, b \quad \vdash [Y], C, d}{\vdash [X, Y], A, C, b \otimes d} \otimes Intro$$

is given by shuffling the matrices S, T, and applying the map $C : M_{a+b}(B(l^2)) \to M_{a+b-1}(B(l^2)$ as found in Chapter 3, Corollory 8, using the weak embedding of P_2 in $B(l^2)$ specified above.

3. Given a matrix S representing a proof of $\vdash [X], A, b, c$, then the application of the $+$ rule, to get a matrix representing the proof ending in the deduction

$$\frac{\vdash [X], A, b, c}{\vdash [X], A, b + c} + Intro$$

is given by applying the contraction map $C : M_n(B(l^2)) \to M_{n-1}(B(l^2))$, as found in Chapter 3, Corollory 8, (again using our weak embedding of P_2) to the matrix S.

4. Given a matrix R representing a proof of a sequent $\vdash [X], a, B$, and a matrix S representing a proof of a sequent $\vdash [Y], a^\perp, C$, then the result of applying the cut rule to get a proof ending in

$$\frac{\vdash [X], a, B \quad \vdash [Y], a^\perp, C}{\vdash [X, Y, a], B, C} Cut$$

is represented by $Sh(R, S)$, given by shuffling the matrices R and S.

5. Given an $n \times n$ matrix R representing a proof of a sequent $\vdash [X], A, b$, then the result of applying the dereliction rule to get a proof ending in

$$\frac{\vdash [X], A, b}{\vdash [X], A, ?b} Dereliction$$

is given by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ & & & \dots & \\ 0 & 0 & 0 & & p_1^{-1} \end{pmatrix} R \begin{pmatrix} 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ & & & \dots & \\ 0 & 0 & 0 & & p_1 \end{pmatrix}$$

6. Given a matrix

$$R = \begin{pmatrix} r_{00} & r_{01} & \dots & r_{0,n-2} & r_{0,n-1} \\ r_{10} & r_{11} & & r_{1,n-2} & r_{1,n-1} \\ \dots & & & & \dots \\ r_{n-2,0} & r_{n-2,1} & & r_{n-2,n-2} & r_{n-2,n-1} \\ r_{n-1,0} & r_{n-1,1} & \dots & r_{n-1,n-2} & r_{n-1,n-1} \end{pmatrix}$$

representing a proof of a sequent $\vdash [X], ?A, b$, then the result of applying an ! introduction rule to get a proof ending in

$$\frac{\vdash [X], ?A, b}{\vdash [X], ?A, !b} !Intro$$

is given by the matrix

$$\begin{pmatrix} \tau^{-1}(1 \otimes r_{00})\tau & \tau^{-1}(1 \otimes r_{01})\tau & \dots & \tau^{-1}(1 \otimes r_{0,n-1}) \\ \tau^{-1}(1 \otimes r_{10})\tau & \tau^{-1}(1 \otimes r_{11})\tau & & \tau^{-1}(1 \otimes r_{1,n-1}) \\ \dots & & & \dots \\ \tau^{-1}(1 \otimes r_{n-2,0})\tau & \tau^{-1}(1 \otimes r_{n-2,1})\tau & & \tau^{-1}(1 \otimes r_{n-2,n-1}) \\ (1 \otimes r_{n-1,0})\tau & (1 \otimes r_{n-1,1})\tau & \dots & (1 \otimes r_{n-1,n-1}) \end{pmatrix}$$

where $\otimes : B(l^2) \times B(l^2) \to B(l^2)$ is the operator specified in Definitions 7.3, and τ is an invertible element of $B(l^2)$ that is required to satisfy $\tau(u \otimes (v \otimes w)) = ((u \otimes v) \otimes w)\tau$ and $\tau\tau^{-1} = 1 = \tau^{-1}\tau$.

7. Given a matrix R representing a proof of a sequent $\vdash [X], A$, then the application of a weakening rule to get a proof ending in

$$\frac{\vdash [X], A}{\vdash [X], A, ?b} Weakening$$

is given by adding an extra row / column of zeros to the outside of R.

8. Given a matrix R representing a proof of a sequent $\vdash [X], A, ?b, ?b$, then the result of applying the contraction rule to get a proof ending in

$$\frac{\vdash [X], A, ?b, ?b}{\vdash [X], A, ?b} Contraction$$

is given by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & & 0 & 0 \\ 0 & 1 & 0 & & 0 & 0 \\ 0 & 0 & 1 & & 0 & 0 \\ & & & \dots & & \\ 0 & 0 & 0 & & (p_1^{-1} \otimes 1) & (p_2^{-1} \otimes 1) \end{pmatrix} R \begin{pmatrix} 1 & 0 & 0 & & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ & & & \dots & \\ 0 & 0 & 0 & & (p_1 \otimes 1) \\ 0 & 0 & 0 & & (p_2 \otimes 1) \end{pmatrix}$$

Note that this is the contraction map $C : M_n(B(l^2)) \to M_{n-1}(B(l^2))$ given in Corollory 8, Chapter 3. However, it uses the embedding of P_2 into $B(l^2)$ generated by $(p_1 \otimes 1)$ and $(p_2 \otimes 1)$ in place of p_1 and p_2, as used in the representation of the + rule.

9. The applications of $\exists$ and $\forall$ rules on proofs are represented by the identity maps on matrices.

10. Given a matrix R representing a proof of a sequent

$$\vdash [X], a_1, a_2, \ldots, a_i, a_{i+1}, \ldots, a_n$$

where $|X| = x$, then the result of applying the exchange rule to get a proof ending in the deduction

$$\vdash [X], a_1, a_2, \ldots, a_{i+1}, a_i, \ldots, a_n$$

is given by exchanging the rows / columns $2x + i - 1$ and $2x + i$ respectively.

7.7 Cut-elimination in GOI1

The above system gives a (restricted) representation of MLL proofs in terms of matrices over $B(l^2)$. However, an important part of the system is the representation the cut-elimination procedure (what J.-Y. Girard refers to as 'the dynamics'). This is given in terms of two formulæ, the execution formula and the resolution formula[6]. These formulæ are defined as follows:

[6]This is actually called the execution formula in [20]; however, we follow the conventions of [21] and [22], where it (or a very similar formula) is referred to as the resolution formula.

Definitions 7.4

Consider a matrix R of the *GOI*1 system, representing a proof of a sequent $\vdash [X], A$, where $x = |X|$. The *feedback matrix*, σ, is defined to be the $2x \times 2x$ matrix

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

that is, it is the disjoint union (in the sense of Definitions 3.2 of Chapter 3) of x antidiagonal matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. The *Execution Formula* is then defined to be the central part of the resolution formula, given by

$$Ex(R, \sigma) = (1 - \sigma R)^{-1} = R + R\sigma R + R\sigma R\sigma R + R\sigma R\sigma R\sigma R + \ldots,$$

and the *Resolution Formula* is defined to be a projection of the execution formula, given by

$$Res(R, \sigma) = (1 - \sigma^2) Ex(R, \sigma)(1 - \sigma^2),$$

or equivalently, $Res(R, \sigma) = (1 - \sigma^2)R(1 - \sigma R)^{-1}(1 - \sigma^2)$. Note that these two formulæ are only well-defined if the element σR is *nilpotent*; that is, it satisfies $(\sigma R)^n = 0$ for some $n \in \mathbb{N}$. It is proved in [20] that all matrices constructed from the GOI1 representation of MLL proofs satisfy this property.

Theorem 1 *The result of applying the resolution formula to R and σ, as defined above, gives the GOI1 representation of the proof found by applying the cut-elimination procedure to the proof represented by R (up to an unspecified series of applications of the exchange rule).*

The proof of this is presented in [20] p.234 - 243, where it is also proved that it is well-defined for any matrix arising from the geometry of interaction 1 system. In fact, the result of applying the cut-elimination formula to R, σ gives the translation of the cut-free proof required, with an extra $2n$ rows / columns of zeros at the beginning, where n is the number of cut variables. However, this is equivalent to the matrix required, up to a number of applications of the exchange rule, and our assumption that all matrices of $GOI1$ are infinite, with a finite number of non-zero entries.

The proof given in [20] is long and unenlightening, so it is not presented here; however, in the next chapter, we study the resolution formula, and the logical operations algebraically and categorically, in terms of the results of Chapters 1 to 7.

Chapter 8

Analysis of the Geometry of Interaction I

8.1 Introduction

In this chapter, we construct algebraic and categorical models of the operations of GOI1, as presented in the last chapter. This requires the construction of an inverse monoid isomorphic to $I(\mathbb{N})$ whose elements are infinite matrices of elements of $I(\mathbb{N})$, as an infinitary version of the self-embedding results of Chapter 5 on the category of partial bijective maps. We demonstrate how the GOI1 system can be represented in this monoid, and hence in $I(\mathbb{N})$, and how the logical operations depend fundamentally on the construction of M-monoids from the self-embedding results on P_2 and P_∞, from Chapters 2 and 5. We also demonstrate how the cut and cut-elimination procedures depend on the internalised trace, as given in Chapter 6, and the definition of composition in the one-object compact closed inverse category $\mathbb{F}$, presented in Chapter 6.

8.2 Inverse semigroup preliminaries

Note that all the operations of GOI1 from Chapter 7 were defined in terms of (infinite matrices over) an embedding of $I(\mathbb{N})$ into $B(l^2)$. Therefore, in this chapter we abuse notation, and work entirely in (infinite matrices over) $I(\mathbb{N})$. This is justified by the embedding $l : I(\mathbb{N}) \to B(l^2)$, of Theorem 18, Chapter 2.

8.2.1 Prerequisites on polycyclic monoids

We will require several algebraic tools for our representation of the system presented in Chapter 7. We require the strong embedding of P_2 into $I(\mathbb{N})$ of Definitions 2.2, Chapter 2, generated by $p^{-1}(n) = 2n$, $q^{-1}(n) = 2n+1$ (the interleaving embedding). We identify elements of P_2 with their images under this embedding.

We also require the embedding of P_∞ into P_2, and hence into $I(\mathbb{N})$, as given in Definitions 2.3, Chapter 2 by $p_i^{-1} = q^{-i}p^{-1}$ for all $i \in \mathbb{N}$ (the right-associative embedding). We again identify elements of P_∞ with their images under this embedding.

This then gives us (from Lemma 4, Chapter 2) a bijection $[\ ,\ ]$ from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$, as required.

The system presented in Chapter 7 also requires a *weak* embedding of P_2. We choose the embedding generated by p_1 and p_2, together with their generalised inverses, for convenience; however, any pair of generators from the above embedding of P_∞ will suffice.

We also require $\oplus : I(\mathbb{N}) \times I(\mathbb{N}) \to I(\mathbb{N})$, the internalisation of the disjoint union in **Inj** at $\mathbb{N}$ (as defined in Theorem 12, Chapter 5 to be $a \oplus b = p^{-1}ap \vee q^{-1}bq$), together with its associativity and commutativity morphisms (denoted t, s repectively, and proved in Theorem 12 of Chapter 5 to be defined by $t = p^{-2}p \vee p^{-1}q^{-1}pq \vee q^{-1}q^2$ and $s = q^{-1}p \vee p^{-1}q$).

Similarly, we require $\otimes : I(\mathbb{N}) \times I(\mathbb{N}) \to I(\mathbb{N})$, the internalisation of the cartesian product of **Inj** at $\mathbb{N}$, which is proved in Proposition 16, Chapter 5 to be

$$(u \otimes v) = \bigvee_{i=0}^{\infty} p_{u(i)}^{-1} v p_i.$$

We also require its associativity and commutativity morphisms, which are denoted τ, σ respectively. It is proved in Proposition 17, Chapter 5 that

$$\tau = \bigvee_{i=0}^{\infty} (p_i^{-1} \otimes 1) p_i,$$

and in Proposition 19, Chapter 5 that σ is not a member of the disjoint closure of P_∞ in $I(\mathbb{N})$.

8.2.2 Summary of algebra used

From the above prerequisites, the system requires the following:

- The strong interleaving embedding of P_2 into $I(\mathbb{N})$, given by $p^{-1}(n) = 2n$, $q^{-1}(n) = 2n+1$.
- The strong right-associative embedding of P_∞ into P_2, given by $p_i^{-1} = q^{-i}p^{-1}$.
- The internalised coproduct, generated by the embedding of P_2, denoted $\oplus$, together with its canonical elements t, s.

- The internalised cartesian product, generated by the embedding of P_∞, denoted $\otimes$, together with its canonical elements, τ, σ.

8.2.3 An infinite matrix semigroup isomorphic to $I(\mathbb{N})$

In Chapter 7, we considered operations defined in terms of infinite matrices over $I(\mathbb{N})$ (or rather, an embedding of $I(\mathbb{N})$ into the C^*–algebra $B(l^2)$). This motivates the construction of an injective map from a monoid of infinite matrices over the monoid of relations $B(\mathbb{N})$ into $B(\mathbb{N})$, which we then restrict to $I(\mathbb{N})$, as an inverse submonoid of $B(\mathbb{N})$.

Definitions 8.1

We define $M_\infty(B(\mathbb{N}))$ to be the set of all infinite matrices over the monoid of relations on the natural numbers, together with the composition given as follows:
Given $A, B \in M_\infty(B(\mathbb{N}))$

$$(AB)_{i,k} = \bigcup_{j=0}^{\infty} A_{ij}B_{jk}.$$

It is immediate that this composition is associative, so $M_\infty(B(\mathbb{N}))$ is a semigroup. Also, the element I satisfying $I_{a,b} = \delta_{ab}$ is clearly an identity with respect to this composition, so $M_\infty(B(\mathbb{N}))$ is a monoid. Then assume the existence of a strong embedding of P_∞ into $B(\mathbb{N})$, and consider the matrices

$$v_\infty^t = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \dots \end{pmatrix}, \quad h_\infty = \begin{pmatrix} p_0^{-1} & p_1^{-1} & p_2^{-1} & \dots \end{pmatrix},$$

(we extend the above composition on $M_\infty(B(\mathbb{N}))$ to these matrices in the natural way). We then have the following useful result:

Lemma 1 *Let v_∞^t, h_∞ be as defined above. Then $h_\infty v_\infty^t = 1 \in B(\mathbb{N})$, $v_\infty^t h_\infty = I \in M_\infty(B(\mathbb{N}))$.*

Proof From the definition of matrix multiplication above, $[v_\infty^t h_\infty]_{ij} = p_i p_j^{-1} = \delta_{ij}$, and so $v_\infty^t h_\infty = I \in M_\infty(B(\mathbb{N}))$. Similarly, $h_\infty v_\infty^t = \bigcup_{i=0}^{\infty} p_i^{-1} p_i$ and as the embedding of P_∞ specified is strong, this is the identity of $B(\mathbb{N})$. □

Definitions 8.2

We define maps $F : B(\mathbb{N}) \to M_\infty(B(\mathbb{N}))$ and $G : M_\infty(B(\mathbb{N})) \to B(\mathbb{N})$ by $G(M) = h_\infty M v_\infty^t$ and $F(r) = v_\infty^t r h_\infty$ respectively.

Proposition 2 *F and G are mutually inverse monoid isomorphisms.*

Proof First note that, for all $r, s \in B(\mathbb{N})$,

$$F(r)F(s) = v_\infty^t r h_\infty v_\infty^t s h_\infty = v_\infty^t r h_\infty v_\infty^t s h_\infty = v_\infty^t r 1 s h_\infty = F(rs),$$

by Lemma 1 above. Similarly,

$$G(M)G(N) = h_\infty M v_\infty^t h_\infty N v_\infty^t = h_\infty M I N v_\infty^t = G(MN),$$

by Lemma 1 above. To see that they are both monoid homomorphisms,

$$G(I) = h_\infty v_\infty^t = 1 \in B(\mathbb{N}) \ , \ F(1) = v_\infty^t h_\infty = I \in M_\infty(B(\mathbb{N})),$$

by Lemma 1 above. Finally, $GF(r) = v_\infty^t h_\infty r v_\infty^t h_\infty = 1r1 = r$ and $FG(M) = h_\infty v_\infty^t M h_\infty v_\infty^t = IMI = M$. Therefore, G and F are mutually inverse, and our result follows. □

Definitions 8.3

We denote the restriction of F to $I(\mathbb{N})$ by Ψ and denote the image of Ψ by $I_\infty(\mathbb{N})$. Similarly, we denote the restriction of G to $I_\infty(\mathbb{N})$ by Φ. As F and G are mutually inverse monoid isomorphisms (Proposition 2), $I_\infty(\mathbb{N})$ is an inverse semigroup isomorphic to $I(\mathbb{N})$, and Φ and Ψ are mutually inverse monoid isomorphisms. The definition of Φ in inverse semigroup theoretic terms is

$$\Phi(M) = \bigvee_{i=0}^{\infty} \bigvee_{j=0}^{\infty} p_i^{-1} m_{ij} p_j,$$

and it is immediate from the construction of $I_\infty(\mathbb{N})$ that this double infinite join is well-defined. The construction of $I_\infty(\mathbb{N})$ can be thought of as the infinitary analogue of the conditions for a 2×2 matrix over $I(X)$ to represent a partial bijective map in $I(X \sqcup X)$, as found in Theorem 10, Chapter 5. In what follows, we prove that all matrices of GOI1 are of the form $\Psi(a)$, for some $a \in DC_{\mathbb{N}}(P_2)$, and hence, as Ψ is an isomorphism, the Geometry of Interaction 1 system can be represented in $DC_{\mathbb{N}}(P_2)$.

8.3 Representing the basic operations of GOI1

In this context, we take 'basic' to be the matrix manipulations used in GOI1, rather than the actual logical rules. We will indicate, along with each basic operation, which logical rules use it.

Note that our operations differ in the following way; all the matrix operations of GOI1 (apart from the cut / cut-elimination process) are applied to the lower right rows / columns of matrices; however, we consider them to be applied to the upper left rows / columns of matrices. The resulting system is of course equivalent, because of the exchange rule, which allows unrestricted (finite) permutations of rows / columns.

8.3.1 The 'permute rows / columns' operation

This operation is used explicitly in the representation of the exchange rule, and implicitly in all other operations.

Definitions 8.4

We define this operation to be the result of applying a finite permutation ρ of $\mathbb{N}$ to the indices of some (infinite) matrix M. So, the result of applying this operation will give a matrix M', where $M'_{\rho(i),\rho(j)} = M_{i,j}$.

Proposition 3 *Let ρ be a permutation of $\mathbb{N}$, and Let M' be derived from M, as above. Then*

$$M' = \Psi(?(\rho^{-1}))M\Psi(?(\rho)),$$

where $?(a) = (a \otimes 1)$, as defined in Definitions 5.6, Chapter 5.

Proof From the definition of the ? homomorphism

$$?(\rho) = \bigvee_{i=0}^{\infty} p_{\rho(i)}^{-1} p_i \ , \ ?(\rho^{-1}) = \bigvee_{j=0}^{\infty} p_j^{-1} p_{\rho(j)}$$

and so $[\Psi(?(\rho))]_{ij} = \delta_{\rho(i),j}$ from the definition of Ψ. Similarly, $[\Psi(?(\rho^{-1}))]_{ij} = \delta_{i,\rho(j)}$, and so

$$[\Psi(?(\rho))M\Psi(?(\rho^{-1}))]_{\rho(i),\rho(j)} = [M]_{i,j}.$$

Hence our result follows. □

8.3.2 The 'shuffle' operation

This operation is used in the representations of the $\otimes$ and Cut rules.

Let M and N be two matrices over $I(\mathbb{N})$. Then their shuffle is as defined in Definitions 7.3, Chapter 7 (or rather, the natural extension of this to infinite matrices).

Proposition 4 *Let M and N be as above. Then*

$$Sh(M,N) = \Psi(?(p^{-1}))M\Psi(?(p)) \vee \Psi(?(q^{-1}))N\Psi(?(q)),$$

where $?(a) = (a \otimes 1)$, as defined in Definitions 5.6, Chapter 5.

Proof By definition of $? : I(\mathbb{N}) \to I(\mathbb{N})$,

$$?(p^{-1}) = \bigvee_{i=0}^{\infty} p_{p^{-1}(i)}^{-1} p_i \ , \ ?(p) = \bigvee_{i=0}^{\infty} p_i^{-1} p_{p^{-1}(i)}.$$

However, as we are using the interleaving embedding of P_2 into $I(\mathbb{N})$, we can write $p(i)$ and $p^{-1}(i)$ explicitly, and so

$$?(p^{-1}) = \bigvee_{i=0}^{\infty} p_{2i}^{-1} p_i \ , \ ?(p) = \bigvee_{i=0}^{\infty} p_i^{-1} p_{2i}.$$

Similarly,

$$?(q^{-1}) = \bigvee_{i=0}^{\infty} p_{2i+1}^{-1} p_i \ , \ ?(q) = \bigvee_{i=0}^{\infty} p_i^{-1} p_{2i+1}.$$

Therefore,

$$[\Psi(?(p^{-1}))]_{ij} = \delta_{2i,j} \ , \ [\Psi(?(p))]_{ij} = \delta_{i,2j}.$$

and similarly

$$[\Psi(?(q^{-1}))]_{ij} = \delta_{2i+1,j} \ , \ [\Psi(?(q))]_{ij} = \delta_{i,2j+1}.$$

This implies that

$$[\Psi(?(p^{-1})M\Psi^{-1}(?(p))]_{ij} = \begin{cases} [M]_{i/2,j/2} & i,j \text{ are even,} \\ 0 & \text{otherwise.} \end{cases}$$

and in a similar way

$$[\Psi(?(q^{-1}))N\Psi(?(q))]_{ij} = \begin{cases} [N]_{(i-1)/2,(j-1)/2} & i,j \geq 1 \text{ are odd,} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore, the join $\Psi(?(p^{-1})M\Psi(?(p)) \vee \Psi(?(q^{-1}))N\Psi(?(q))$ is clearly defined, and equal to $Sh(M,N)$. □

8.3.3 The 'contraction by p_1, p_2' operation

This operation is used in the representation of the ⊗ and + rules.

Definitions 8.5

Let M be a matrix over $I(\mathbb{N})$, and let p, q, p_1, p_2 be as defined in Section 8.2.2 above. Then the result of applying the *contraction by* p_1, p_2 map, which contracts the first two rows / columns into a single row / column, to a matrix M gives a matrix M', where

$$M' = \begin{pmatrix} p_1^{-1} & p_2^{-1} & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ & & & & \ddots \end{pmatrix} M \begin{pmatrix} p_1 & 0 & 0 & \\ p_2 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ & & & \ddots \end{pmatrix}.$$

Proposition 5 *Let M and M' be as above. Then $M' = \Psi(t')M\Psi(t'^{-1})$, where*

$$t' = p^{-1}p_1^{-1}p \vee p^{-1}p_2^{-1}pq \vee q^{-1}q^2 \ , \ t'^{-1} = q^{-2}q \vee q^{-1}p^{-1}p_2p \vee p^{-1}p_1p.$$

Proof Denote $\Psi(t')$ by T, for clarity. Then by definition of Ψ, $T_{ij} = p_it'p_j^{-1}$. Therefore,

$$T_{00} = p_0(p^{-1}p_1^{-1}p \vee p^{-1}p_2^{-1}pq \vee q^{-1}q^2)p_0^{-1}.$$

However, $p_0 = p$, by definition of the right-associative embedding of P_∞ into P_2, and so $T_{00} = pp^{-1}p_1^{-1}pp^{-1} = p_1^{-1}$. Also, $T_{01} = p(p^{-1}p_1^{-1}p \vee p^{-1}p_2^{-1}pq \vee q^{-1}q^2)p_1^{-1}$ and $p_1^{-1} = q^{-1}p^{-1}$ by definition of the right-associative embedding of P_∞ into P_2. Therefore, $T_{01} = p(p^{-1}p_1^{-1}p \vee p^{-1}p_2^{-1}pq \vee q^{-1}q^2)q^{-1}p^{-1} = p_2^{-1}$. Next, $T_{ij} = p_i^{-1}q^{-2}qp_j^{-1}$ for all $i \neq 0$ and $j \neq 0, 1$, and as $p_i = pq^i$ and $p_j^{-1} = q^{-j}p^{-1}$,

$$T_{ij} = pq^iq^{-1}q^2q^{-j}p^{-1} = pq^{i-1}q^2q^{-j}p^{-1} = pq^{i+1}q^{-j}p^{-1} = \delta_{i+1,j}.$$

Therefore,

$$\Psi(t') = \begin{pmatrix} p_1^{-1} & p_2^{-1} & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ & & & & \cdots \end{pmatrix},$$

and a similar proof gives that

$$\Psi(t'^{-1}) = \begin{pmatrix} p_1 & 0 & 0 & \\ p_2 & 0 & 0 & \\ 0 & 1 & 0 & \\ & & & \cdots \end{pmatrix}.$$

Hence our result follows. □

8.3.4 The 'conjugate row/column by p_1' operation

This operation is used in the representation of the dereliction rule.

Definitions 8.6

Let M be a matrix over $I(\mathbb{N})$. Then the result of applying the *conjugate row / column by p_1* map to M is the matrix M', where

$$M' = \begin{pmatrix} p_1^{-1} & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ & & & \cdots \end{pmatrix} M \begin{pmatrix} p_1 & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ & & & \cdots \end{pmatrix}.$$

Proposition 6 *Let M and M' be as above. Then $M' = \Psi(p_1^{-1} \oplus 1)M\Psi(p_1 \oplus 1)$.*

Proof $p^{-1} \oplus 1 = p^{-1}p_1^{-1}p \vee q^{-1}q$, from the definition of $\oplus$. Then, from the definition of Ψ, $[\Psi(p^{-1} \oplus 1)]_{ij} = p_i(p^{-1}p_1^{-1}p \vee q^{-1}q)p_j^{-1}$, and from the definition of the right-associative embedding of P_∞ into P_2,

$$[\Psi(p^{-1} \oplus 1)]_{ij} = \begin{cases} p^{-1} & i = j = 0 \\ \delta_{ij} & \text{otherwise.} \end{cases}$$

A similar proof gives that

$$[\Psi(p \oplus 1)]_{ij} = \begin{cases} p & i = j = 0 \\ \delta_{ij} & \text{otherwise,} \end{cases}$$

and hence our result follows. □

8.3.5 The 'apply ! to each element' operation

This operation is used in the representation of the !-introduction rule

Definitions 8.7

Let M denote the matrix over $I(\mathbb{N})$ given by

$$\begin{pmatrix} m_{00} & m_{01} & \\ m_{10} & m_{11} & \\ & & \dots \end{pmatrix}.$$

Then the result of applying the *apply* ! *to each element* map to M is the matrix M' satisfying

$$M' = \begin{pmatrix} !(m_{00}) & !(m_{01}) & \\ !(m_{10}) & !(m_{11}) & \\ & & \dots \end{pmatrix}.$$

Proposition 7 *Let M and M', as given above, be in $I_\infty(\mathbb{N})$. Then $\Phi(M') = ?(\sigma)!(\Phi(M))?(\sigma)$.*

Proof We denote $\Phi_\infty(M)$ by a, and $\Phi_\infty(M')$ by a'. Then from the definition of the Φ homomorphism,

$$a = \bigvee_{i,j=0}^{\infty} p_i^{-1}m_{ij}p_j \ , \ b = \bigvee_{i,j=0}^{\infty} p_i^{-1}!(m_{ij})p_j.$$

Hence, by definition of $! : I(\mathbb{N}) \to I(\mathbb{N})$,

$$!(a) = \bigvee_{i,j,k=0}^{\infty} p_k^{-1}p_i^{-1}m_{ij}p_jp_k,$$

and by definition of $\sigma \otimes 1$,

$$(\sigma \otimes 1)!(a)(\sigma \otimes 1) = \bigvee_{i,j,k=0}^{\infty} p_i^{-1} p_k^{-1} m_{ij} p_k p_j = \bigvee_{i,j=0}^{\infty} p_i^{-1}!(m_{ij}) p_j,$$

Therefore, $b = ?(\sigma)!(a)?(\sigma)$, and so our result follows. □

8.3.6 The 'conjugate all rows/columns but one with τ' operation

This operation is used in the !-introduction rule.

Definitions 8.8

Let M denote a matrix over $I(\mathbb{N})$. Then the result of applying the *conjugate all rows / colums but one with* τ map is given by the matrix M', where

$$M' = \begin{pmatrix} 1 & 0 & 0 & \\ 0 & \tau^{-1} & 0 & \\ 0 & 0 & \tau^{-1} & \\ & & & \cdots \end{pmatrix} M \begin{pmatrix} 1 & 0 & 0 & \\ 0 & \tau & 0 & \\ 0 & 0 & \tau & \\ & & & \cdots \end{pmatrix}$$

Proposition 8 *Let M and M' be as above. Then $M' = \Psi(1 \oplus !(\tau^{-1})) M \Psi(1 \oplus !(\tau))$.*

Proof From the definition of $\oplus$,

$$1 \oplus !(\tau^{-1}) = p^{-1} p \vee q^{-1} !(\tau^{-1}) q = p^{-1} p \vee q^{-1} \left(\bigvee_{i=0}^{\infty} p_i^{-1} \tau^{-1} p_i \right) q.$$

However, $p_i^{-1} = q^{-i} p^{-1}$, by the definition of the right-associative embedding of P_∞ into P_2, and so

$$1 \oplus !(\tau^{-1}) = p^{-1} p \vee \bigvee_{i=1}^{\infty} p_i^{-1} \tau^{-1} p_i.$$

Then it is immediate by definition of Ψ that

$$[\Psi(1 \oplus !(\tau^{-1}))]_{ij} = \begin{cases} 1 & i = j = 0 \\ \tau^{-1} \delta_{ij} & \text{otherwise.} \end{cases}$$

An almost identical proof gives that

$$[\Psi(1 \oplus !(\tau))]_{ij} = \begin{cases} 1 & i = j = 0 \\ \tau \delta_{ij} & \text{otherwise.} \end{cases}$$

Hence our result follows. □

8.3.7 The 'add inner zeros' operation

Let M be a matrix over $I(\mathbb{N})$, and let M' denote the matrix that results from adding an inner row/column of zeros to M.

Proposition 9 *Let M and M' be as above. Then M' satisfies $M' = \Psi(q^{-1})M\Psi(q)$ and hence $\Phi(M') = 0 \oplus \Phi(M)$ when $M \in I_\infty(\mathbb{N})$.*

Proof $[\Psi(q^{-1})]_{ij} = p_i q^{-1} p_j^{-1}$ from the definition of Ψ, and from the definition of the right-associative embedding of P_∞ into P_2, $q^{-1}p_j^{-1} = q^{-1}q^{-j}p^{-1} = q^{-j-1}p^{-1} = p_{j+1}^{-1}$. Therefore $[\Psi(q^{-1})]_{ij} = \delta_{i,j+1}$, and similarly, $[\Psi(q)]_{ij} = \delta_{i+1,j}$. This implies that

$$[\Psi(q^{-1})M\Psi(q)]_{ij} = \begin{cases} M_{i-1,j-1} & i,j \geq 1 \\ 0 & \text{otherwise.} \end{cases}$$

Hence the first part of our result follows. The second follows by definition of $\oplus$. □

8.3.8 The 'contract rows / columns using $?(p_1), ?(p_2)$' operation

This operation is used in the representation of the contraction rule.

Definitions 8.9

Let M be a matrix over $I(\mathbb{N})$. Then the result of applying the *contract rows / columns using* $?(p_1), ?(p_2)$ operation to M is given by the matrix M', where

$$M' = \begin{pmatrix} ?(p_1^{-1}) & ?(p_2^{-1}) & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 1 & \\ & & & & \ldots \end{pmatrix} M \begin{pmatrix} ?(p_1) & 0 & 0 & \\ ?(p_2) & 0 & 0 & \\ 0 & 1 & 0 & \\ 0 & 0 & 1 & \\ & & & \ldots \end{pmatrix}.$$

Proposition 10 *Let M and M' be as above. Then $M' = \Psi(t'')M\Psi(t''^{-1})$, where*

$$t'' = p^{-1}?(p_1^{-1})p \vee p^{-1}?(p_2^{-1})pq \vee q^{-1}q^2 \ , \ t''^{-1} = q^{-2}q \vee q^{-1}p^{-1}?(p_2)p \vee p^{-1}?(p_1)p.$$

Proof The proof of this is practically identical to the proof of Proposition 5; it makes no difference to the calculations of the proof to use the weak embedding of P_2 generated by $(p_1 \otimes 1)$ and $(p_2 \otimes 1)$ in place of that generated by p_1 and p_2. □

8.4 Representing GOI1 in $DC_{\mathbb{N}}(P_2)$

Lemma 11 *The matrix* $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *representing the axiom link satisfies* $\Phi(A) = t^{-1}(s \oplus 0)t$.

Proof $\Phi(A) = p_0^{-1}p_1 \vee p_1^{-1}p_0$ by definition of Φ, and $\Phi(A) = p^{-1}pq \vee q^{-1}p^{-1}p$, by the definition of the embedding $\theta_\infty : P_\infty \to P_2$. Therefore, by definition of the associativity elements for the M-monoid $(I(\mathbb{N}), \oplus)$,

$$t\Phi(A)t^{-1} = (p^{-2}p \vee p^{-1}q^{-1}pq \vee q^{-1}q^2)(p^{-1}pq \vee q^{-1}p^{-1}p)(q^{-2}q \vee q^{-1}p^{-1}qp \vee p^{-1}p^2)$$

$$= (p^{-2}pq \vee p^{-1}q^{-1}p)(q^{-2}q \vee q^{-1}p^{-1}qp \vee p^{-1}p^2) = p^{-2}qp \vee p^{-1}q^{-1}p^2.$$

However, by definition of the commutativity element for the M-monoid $(I(\mathbb{N}), \oplus)$, and the definition of the internalised disjoint union in terms of P_2, $s \oplus 0 = p^{-1}(p^{-1}q \vee q^{-1}p)p \vee q^{-1}0q = p^{-2}qp \vee p^{-1}q^{-1}p^2$. Hence $t\Phi(A)t^{-1} = s \oplus 0$, and so $\Phi(A) = t^{-1}(s \oplus 0)t$, as $t^{-1}t = 1 = tt^{-1}$. □

Theorem 12 *Every matrix* M *arising from the GOI1 system satisfies* $M = \Psi_\infty(a)$, *for some* $a \in DC_{\mathbb{N}}(P_2)$ *satisfying* $a^{-1} = a$.

Proof It is immediate that the axiom link satisfies this property, from Lemma 11 above. Also, for arbitrary $f \in I(\mathbb{N})$, $?(f) \in DC_{\mathbb{N}}(P_2)$ by Proposition 24 of Chapter 5. Finally, given $a, b, f \in I(\mathbb{N})$, where $a^{-1} = a$, $b^{-1} = b$, then

- $(faf^{-1})^{-1} = faf^{-1}$
- $(a \vee b)^{-1} = a^{-1} \vee b^{-1} = a \vee b$
- $!(a)^{-1} = !(a^{-1}) = !(a)$
- $?(a)^{-1} = ?(a^{-1}) = ?(a)$
- $(a \oplus b)^{-1} = a^{-1} \oplus b^{-1} = a \oplus b$

Therefore, the set of self-inverse elements of $I(\mathbb{N})$ is closed under conjugation by arbitrary members of $I(\mathbb{N})$, disjoint join, and the ! and ? homomorphisms. Hence, as all the operations of Propositions 4 to 10 above are constructed in this way, our result follows by induction. □

8.5 Cut-elimination as the internal trace of an M-monoid

We claim that the resolution formula, which models cut-elimination in the GOI1 system, is given by the internalisation of the trace (at $\mathbb{N}$) in the category of partial bijective maps (up to an associativity and a commutativity isomorphism). We first construct an explicit description of the internalisation of the trace at the strongly self-similar object $\mathbb{N}$ in the category $(\mathbf{Inj}, \sqcup)$, give the 'opposite' construction to it, and then demonstrate how the resolution formula is given in terms of this.

8.5.1 The internal trace at $\mathbb{N}$

Recall the definition of the trace on the category **Inj** from Theorem 2 of Chapter 6, as follows: Given a morphism $F \in \mathbf{Inj}(X \sqcup U, Y \sqcup U)$ represented by the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $Tr^U_{X,Y}(F) = a \vee bd^*c$, and we have proved (Theorem 2, Chapter 6) that this is also a partial bijective map. However, as $\mathbb{N}$ is a self-similar object of **Inj** we can construct the internalisation of the categorical trace in the endomorphism monoid of $\mathbb{N}$, as shown in Definitions 6.2, Chapter 6, and an explicit description of this is as follows:

Proposition 13 *For all $f \in I(\mathbb{N})$, the internal trace of f is given by*

$$Trace(f) = pfp^{-1} \vee pf((0 \oplus 1)f)^*((0 \oplus 1)f)p^{-1}.$$

Proof Recall the definition of the map from $I(\mathbb{N})$ to $I(\mathbb{N} \sqcup \mathbb{N})$ of Theorem 12, Chapter 5, as

$$f \mapsto \begin{pmatrix} pfp^{-1} & pfq^{-1} \\ qfp^{-1} & qfq^{-1} \end{pmatrix}$$

Then by the definition of the trace in the category of partial bijective maps,

$$Tr^{\mathbb{N}}_{\mathbb{N},\mathbb{N}} \begin{pmatrix} pfp^{-1} & pfq^{-1} \\ qfp^{-1} & qfq^{-1} \end{pmatrix} = pfp^{-1} \vee pfq^{-1}(qfq^{-1})^* qfp^{-1},$$

and so, by definition of the internalisation of a trace, (Definitions 6.2, Chapter 6), $trace(f) = pfp^{-1} \vee pfq^{-1}(qfq^{-1})^* qfp^{-1}$. However, $q^{-1}q = 0 \oplus 1$, by definition of $\oplus$, so

$$trace(f) = pfp^{-1} \vee pf\left(\bigvee_{i=1}^{\infty} ((0 \oplus 1)f)^i\right)p^{-1} = pfp^{-1} \vee pf((0 \oplus 1)f)^*((0 \oplus 1)f)p^{-1}.$$

Hence our result follows. □

Definitions 8.10

We define the *opposite trace*, which we denote $OpTr$, as follows:

Let $F \in \mathbf{Inj}(U \sqcup X, U \sqcup Y)$ be represented by the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then $OpTr^U_{X,Y} = d \vee ca^*b$.

It is immediate that $OpTr^U_{X,Y} = Tr^U_{X,Y}(s_{U,Y}Fs_{U,X})$ so the theory of the opposite trace follows immediately from the theory of the trace. In particular, at a self-similar object, we can define the *internalised opposite trace*, which we denote $OpTrace$ in the same way as the internalised trace (Definitions 6.2, Chapter 6), and it is immediate that $OpTrace(f) = trace(sfs)$, where s is the internalisation of the commutativity morphism (as found in Theorem 9, Chapter 4). Also, we can construct an explicit description of the internalised opposite trace as a corollary of Proposition 13 above.

Corollary 14 *The internalisation of the opposite trace is given by, for all $f \in I(\mathbb{N})$,*

$$OpTrace(f) = qfq^{-1} \vee q[f(1 \oplus 0)f]^*(1 \oplus 0)fq^{-1}.$$

Proof Immediate from Proposition 13, and the definition of the opposite internal trace. □

8.5.2 Connecting the resolution formula and the internalised trace

Proposition 15 *Given U, a matrix of GOI1, where $\Phi(U) = u$, for some $u \in I(\mathbb{N})$, and a matrix $S = \Psi(s)$, where s is the commutativity element for the M-monoid $(I(\mathbb{N}), \oplus)$, then*

$$Res(U, S) = \Psi(0 \oplus OpTrace((s \oplus 1)u)).$$

Proof Let us denote $(s \oplus 1)u$ by α, for clarity, so

$$OpTrace(\alpha) = q\alpha q^{-1} \vee q[\alpha(1 \oplus 0)\alpha]^*(1 \oplus 0)\alpha q^{-1}.$$

However, $\alpha = (p^{-1}sp \vee q^{-1}q)u$, by definition of $\oplus$, and so

$$p\alpha p^{-1} = spup^{-1}, \quad p\alpha q^{-1} = spuq^{-1},$$
$$q\alpha p^{-1} = qup^{-1}, \quad q\alpha q^{-1} = quq^{-1}.$$

Hence $OpTrace(\alpha) = quq^{-1} \vee qup^{-1}(spup^{-1})^*spuq^{-1}$. Also, $p^{-1}sp = s \oplus 0$, by definition of $\oplus$, and so $OpTrace(\alpha) = quq^{-1} \vee qu(s \oplus 0)(u(s \oplus 0))^*q^{-1}$. Therefore,

$$(0 \oplus OpTrace(\alpha)) = q^{-1}quq^{-1}q \vee q^{-1}qu(s \oplus 0)(u(s \oplus 0))^*q^{-1}q.$$

Next, note that $q^{-1}q = 0 \oplus 1$ and so

$$(0 \oplus OpTrace(\alpha)) = (0 \oplus 1)u(0 \oplus 1) \vee (0 \oplus 1)u(s \oplus 0)(u(s \oplus 0))^*(0 \oplus 1))$$

$$= \bigvee_{i=0}^{\infty} (0 \oplus 1)u((s \oplus 0)u)^i(0 \oplus 1).$$

Finally, $(s \oplus 0)^2 = (1 \oplus 0)$, and $(0 \oplus 1)^{\perp} = 1 \oplus 0$. Therefore, $(1 \oplus 0) = ((s \oplus 0)^2)^{\perp}$, (where $(\ )^{\perp}$ is as defined in Definitions 1.9, Chapter 1), and as Ψ is an isomorphism,

$$\Psi(0 \oplus Optrace(\alpha)) = \Psi(\bigvee_{i=0}^{\infty}(0 \oplus 1)(u((s \oplus 0)u)^i(0 \oplus 1)).$$

As $u = \Phi(U)$ and $s \oplus 0 = \Phi(S)$, we can write this as $\Psi(0 \oplus Optrace(\alpha)) = (S^2)^{\perp}U(SU)^*(S^2)^{\perp}$, and under the embedding of $I(\mathbb{N})$ into $B(l^2)$, this is the definition of the resolution formula of Definitions 7.4, Chapter 7. Therefore, we have proved that, for any matrix U in GOI1, and matrix S, as above, $Res(U, S)$ is in $I_\infty(\mathbb{N})$, and hence the resolution formula can also be represented in $I(\mathbb{N})$. □

8.5.3 Cut-elimination as the double internal trace

We consider the exact form of the categorical trace used in the Resolution formula, as found in Chapter 7. For this, we require the following result:

Lemma 16 *The map $T : I(\mathbb{N}) \to I(\mathbb{N})$ defined by $T(a) = t^{-1}at$, where t is the associativity element for the strong M-monoid $(I(\mathbb{N}), \oplus)$, is an M-monoid isomorphism.*

Proof Note that $tt^{-1} = 1 = t^{-1}t$. Therefore $T(a)T(b) = t^{-1}att^{-1}bt = t^{-1}abt = T(ab)$ and $T(1) = t^{-1}t = 1$. Also, it is immediate that $T(a \vee b) = T(a) \vee T(b)$, and so the M-monoid structure is also preserved. Therefore our result follows. □

Theorem 17 *Let M be a matrix of GOI1 representing a proof π of MLL with a single cut variable (subject to the restrictions on MLL proofs given by Section 7.5.2 of Chapter 7), and let π' be the proof of MLL given by applying the cut-elimination procedure to π. Then the matrix M' representing the translation of the proof π into GOI1 satisfies*

$$\Phi(M') = OpTrace^2(\Phi(M(S \sqcup I))) = trace(s(trace(s(\Phi(U)t^{-1}(s \oplus 1)t)s))s).$$

Proof First note that the feedback matrix S used in the elimination of a single cut variable of the GOI1 system is the same matrix as used in the axiom link, and we have proved (Lemma 11) that this matrix is given by $S = \Psi(t^{-1}(s \oplus 0)t)$. Therefore, given a matrix M representing a proof in GOI1, and the feedback matrix S, the result of applying the resolution formula satisfies

$$T(\Psi(Res(M, S))) = (0 \oplus OpTrace(T(\Phi(M))(s \oplus 1))) = 0 \oplus (Optrace(T(\Phi(M)\Phi(S)))).$$

However, $Optrace(t^{-1}at) = OpTrace(OpTrace(a))$, by the dual result to axiom 1 for a traced strong M-monoid (Definitions 6.3, Chapter 6). Therefore,

$$T(\Psi(Res(M,S))) = (0 \oplus (OpTrace(OpTrace(\Phi(M(S \sqcup I))))).$$

Next, it is immediate that $T^{-1}(0 \oplus x) = 0 \oplus (0 \oplus x)$, for all $x \in I(\mathbb{N})$. Therefore,

$$\Psi(Res(M,S)) = 0 \oplus (0 \oplus (OpTrace^2(\Phi(M(S \sqcup I))))).$$

Finally, recall from Theorem 1, Chapter 7 that the resolution formula gives the representation of the cut-free proof, with an extra $2n$ upper left rows / columns of zeros, where n is the number of cut variables. Therefore, the translation of the cut proof is given by $q^2(\Psi(Res(M,S))q^{-2}$ and, by above, this is equal to $OpTrace^2(\Phi(M(S \sqcup I)))$. Therefore, by definitions of $OpTrace$ and of S, the translation of the cut-free proof π' is given by $trace(s(trace(s(\Phi(U)t(s \oplus 1)t^{-1})s))s)$. Hence our result follows. □

8.6 The cut / cut-elimination process as composition in $\mathbb{F}$

We consider the consequences of combining the cut and cut-elimination processes (on a single variable) to construct a binary operation on GOI1 matrices. In what follows, we abuse notation by considering the cut / cut-elimination procedure on 2×2 matrices, representing morphisms in $I(\mathbb{N} \sqcup \mathbb{N})$. However, in view of the isomorphisms between $I(\sqcup^{\alpha}\mathbb{N})$, for countable α, this is not a restriction.

Definitions 8.11

We define the *computation* operation to be a binary operation $C : I(\mathbb{N} \sqcup \mathbb{N}) \times I(\mathbb{N} \sqcup \mathbb{N}) \to I(\mathbb{N} \sqcup \mathbb{N})$ that corresponds to applying the cut rule to a pair of 2×2 matrices over $I(\mathbb{N})$, and then applying the cut-elimination process (up to applications of the exchange rule), as given by the resolution formula. Our aim is to demonstrate how this operation can be written in terms of the one-object compact closed inverse category $\mathbb{F}$, defined in Section 6.7, Chapter 6.

Theorem 18 *Let F and G be two square matrices over $I(\mathbb{N})$, as above. Then $C(F,G)$ is given by $C(F,G) = F^{\vee} \cdot G$, where $(\ )^{\vee}$ is the dual on the elements of $\mathbb{F}$ and $\cdot$ is the alternative composition, as given in Section 6.7, Chapter 6.*

Proof Let F and G be represented by the matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} e & f \\ g & h \end{pmatrix}$ respectively. The result of shuffling them (which, by Section 7.6 of Chapter 7 is the representation of the cut rule) is given by

$$Sh(F,G) = \begin{pmatrix} a & 0 & b & 0 \\ 0 & e & 0 & f \\ c & 0 & d & 0 \\ 0 & g & 0 & h \end{pmatrix}.$$

We then apply the cut-elimination procedure, as given in Theorem 17 above.

First, we premultiply by the image of $S \sqcup I$, given by

$$S \sqcup I = \Psi(t(s \oplus 1)t^{-1}) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

to get

$$(S \sqcup I)Sh(F,G) = \begin{pmatrix} 0 & e & 0 & f \\ a & 0 & b & 0 \\ c & 0 & d & 0 \\ 0 & g & 0 & h \end{pmatrix},$$

and then apply the (opposite) trace twice to get the matrix M' representing the cut-free proof we require (Theorem 17 above). Alternatively, we can contract the above into a (2×2) matrix, which we denote

$$M = \begin{pmatrix} \begin{pmatrix} 0 & a \\ e & 0 \end{pmatrix} & \begin{pmatrix} 0 & f \\ b & 0 \end{pmatrix} \\ \begin{pmatrix} c & 0 \\ 0 & g \end{pmatrix} & \begin{pmatrix} d & 0 \\ 0 & h \end{pmatrix} \end{pmatrix},$$

and then (by the naturality of the categorical trace), apply a single (opposite) trace, to get the matrix M'. If we define $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where

$$A = \begin{pmatrix} 0 & a \\ e & 0 \end{pmatrix} \quad B = \begin{pmatrix} 0 & f \\ b & 0 \end{pmatrix}$$
$$C = \begin{pmatrix} c & 0 \\ 0 & g \end{pmatrix} \quad D = \begin{pmatrix} d & 0 \\ 0 & h \end{pmatrix},$$

Then the application of the opposite trace to M will give us $M' = D \vee CA^*B$. However, it is immediate that $A^* = \bigcup_{i=0}^{\infty} \begin{pmatrix} (ea)^i & e(ae)^i \\ a(ea)^i & (ae)^i \end{pmatrix}$ and so

$$M' = \begin{pmatrix} d & 0 \\ 0 & h \end{pmatrix} \vee \bigvee_{i=0}^{\infty} \left[\begin{pmatrix} c & 0 \\ 0 & g \end{pmatrix} \begin{pmatrix} (ea)^i & e(ae)^i \\ a(ea)^i & (ae)^i \end{pmatrix} \begin{pmatrix} 0 & f \\ b & 0 \end{pmatrix} \right].$$

Therefore,

$$M' = \begin{pmatrix} d \vee c(ea)^*eb & c(ea)^*f \\ g(ae)^*b & h \vee g(ae)^*af \end{pmatrix}.$$

The definition of the dual on elements of $\mathbb{F}$ is given in Section 6.7 of Chapter 6 as $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{\vee} = \begin{pmatrix} d & c \\ b & a \end{pmatrix}$. It is then immediate from this, and from the definition of the alternative composition on $\mathbb{F}$ that

$$F^{\vee} \cdot G = \begin{pmatrix} d & c \\ b & a \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} d \vee c(ea)^*eb & c(ea)^*f \\ g(ae)^*b & h \vee g(ae)^*af \end{pmatrix}.$$

Hence our result follows. □

This then makes immediate what J.-Y. Girard refers to as the 'essential case' of the cut-elimination theorem ([20], p.235), as follows:

Corollary 19 *The matrix representing the application of the cut-elimination procedure to a matrix representing a cut between two axiom links is another matrix representing an axiom link.*

Proof In the monoid $(\mathbb{F}, \circ)$, the identity matrix I is $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and this is clearly self-dual. Also, under the isomorphism $S : \mathbb{F} \to \mathbb{F}'$, $S(I) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is the representation of the axiom link. Therefore, the result follows immediately from $I \cdot I = I$ in $(\mathbb{F}, \cdot)$. □

8.7 Full linear logic, and the GOI3 system

The translation of the system presented in [20] and [21] into the closure of P_2 in $I(\mathbb{N})$, and the description of the resolution formula in terms of the categorical trace and the monoid $\mathbb{F}$ is enough to model multiplicative linear logic (subject to the conditions presented in Section 7.6.2, Chapter

7), and Girard's System F. However, in [22], J.-Y. Girard claims a dynamical model of cut-elimination for the whole of linear logic using the resolution formula (which we have seen is the internalised trace of an M-monoid) and modelling the logical operators using the *clause algebra* – a structure that we study in the next chapter. In particular, we prove that this is also an inverse semigroup, and give a representation in terms of the inverse semigroup of partial bijective maps on a term language, derived from the semilattice structure of a term language.

The system presented in [22] depends on the extension of Proof Nets (see [17]) to the whole of linear logic, which is an unpublished result of J.-Y. Girard, and involves major changes to the logical operations used and the cut-eliminations procedure. In view of this, we do not study the clause algebra in terms of its applications to linear logic. We do, however, study its structure in terms of inverse semigroup theory, and its resolution formula in terms of the categorical trace.

Chapter 9

The clause semigroup and its applications

9.1 Introduction

In this chapter, we demonstrate how the operations of unification, resolution and substitution, defined on term–languages, can be represented in terms of semilattice and inverse semigroup theory. The construction of the clause algebra, which is defined and used to model the third part of the Geometry of Interaction system in [22], is given in terms of the inverse semigroup of partial bijective maps on a term language, and its structure, and action on the term language is studied.

The conditions on a term language required by J.-Y. Girard for a representation of full linear logic are also considered, and the Resolution formula, introduced for the clause algebra in [22] and used to model computation, is proved to be the categorical trace in the category of partial bijective maps.

9.2 Term languages, substitution, and unification

The following definitions are taken from [13].

Definitions 9.1

A *ranked alphabet* is defined to be a set Σ, together with a countably infinite subset of *variable symbols*, written $V \subseteq \Sigma$, and an *arity* or *rank* function $r : \Sigma \to \mathbb{N}$. Variables have rank 0, and non-variable symbols of rank 0 are called *constants.* A Σ*–tree* is defined to be a tree T (in the graph–theoretic sense), together with a function t from the set of nodes of T, written $dom(T)$, to the ranked alphabet Σ. We define the *domain function*, $d : dom(T) \to \mathbb{N}$, that gives the number of outgoings of the node n, and require t to satisfy $d(n) = r(t(n))$. Σ–trees are also referred to as Σ–terms, or just terms. The set of variables in a term T is called the *free variables* of T, and denoted $FV(T)$. A *ground term* is a term G that contains no variable symbols, so that $FV(G) = \emptyset$. The set of all Σ–trees is called the *term–language* given by Σ, written L_Σ, or just L. By convention, the tree structure is represented by bracketing.

A *substitution* is defined to be a map $\sigma : V \to L$ that assigns a member of L to each variable. The set of all substitutions on a term language L is denoted $Subst(L)$. The set of variables changed by a substitution σ is called the *support* of σ, written $supp(\sigma)$. The set of all terms of L over the a subset of variables $X \subseteq V$ is denoted $L(X)$. In [13], p.342-344, it is proved that there exists a unique extension of a substitution $\sigma : V \to L$ to $\hat{\sigma} : L \to L$, the (recursively defined) function that replaces each occurence of a member of the support of σ occurring in a Σ–term T with its image under the substitution σ. We abuse notation, and refer to the substitution $\sigma : L \to L$, unless the distinction is important. Let L be the term language defined by the ranked alphabet Σ, and consider an arbitrary subset $S \subseteq L$. A substitution $\sigma : L \to L$ is called a *unifier* of the set $S \subseteq L$ if $\sigma(s) = \sigma(s')$ for all $s, s' \in S$. The term $ci_\sigma(S) = \sigma(s) = \sigma(s')$ is called the *common instance* of the set S, generated by σ. A unifier σ is called a *most general unifier* of the set S if, given an arbitrary unifier σ', there exists a substitution θ such that $\sigma' = \theta \circ \sigma$. We write $\sigma \in mgu(S)$. A common instance generated by a most general unifier is called a *most common instance*, denoted $\sigma(s) \in mci(S)$.

In [13], p.383, it is proved that, unless a set of terms is non-unifiable, then it has a most general unifier, and most general unifiers are unique up to isomorphism (that is, renaming of variables). There also exists a deterministic, terminating algorithm that can be proved to find a most general unifier (if one exists) of a set of members of a term language. This algorithm (in LISP), along with a proof of this result, can be found in [13] p.383-385.

We demonstrate how an inverse semigroup can be constructed from a semilattice structure derived from the operations of unification of a term language. This inverse semigroup, along with an action on the ground terms of a term language is used by J.-Y. Girard in [22] to construct a C^*–algebra of bounded linear operators acting on a Hilbert space of square-summable sequences of complex numbers indexed by members of the term language. This construction first requires an analysis of partial orders and preorders that can be defined on a term language using the operations of substitution and unification, as follows:

9.2.1 Partial orders and preorders on term languages

Definitions 9.2

We define a binary relation $<$ on the term language L_Σ as follows:
For arbitrary terms $a, b \in L$, we say $a < b$ iff there exists a substitution σ satisfying $a = \sigma(b)$.

Lemma 1 *$<$ is a preorder on the term language L, but is not a partial order.*

Proof Given $a < b$ and $b < c$ in L, we can find substitutions σ, τ such that $a = \sigma(b)$, $b = \tau(c)$. Therefore, $a = \sigma\tau(c)$, and so $a < c$. Hence $<$ is a preorder. However, $x < y$ and $y < x$ does not imply that $x = y$, as x and y may differ by a renaming substitution. Therefore, $<$ is not a partial order. □

Definitions 9.3

Given $a \in L$, we define its *substitution class* by $[a] = \{\sigma(a) : \sigma \in Subst(L)\}$. We also define a binary relation $\leq$ on substitution classes by $[a] \leq [b]$ iff for all $x \in [a]$, there exists $y \in [b]$ satisfying $x < y$.

Lemma 2 *$\leq$ is a partial order on the set of substitution classes of L.*

Proof Given $[a] \leq [b]$ and $[b] \leq [c]$, then for all $x \in [a]$ we can find $y \in [b]$ such that $x = \sigma(b)$, for some substitution σ. Also, by definition of $[b] \leq [c]$, we can find $z \in [c]$ such that $y = \tau(z)$, for some substitution τ. Therefore, $x = \sigma\tau(z)$, and so the relation τ is transitive. Also, given $[a] \leq [b]$, and $[b] \leq [a]$, then a and b must differ by a renaming substitution, and it is immediate that $[a] = [b]$. Therefore, $\leq$ is a partial order. □

Lemma 3 *Let $a, b \in L$ satisfy $FV(a) \cap FV(b) = \emptyset$. Then $[a] \cap [b] = [mci(a, b)]$, when a most common instance of a and b exists.*

Proof (We assume that $[a] \cap [b] \neq \emptyset$). From the definition of substitution classes, it is immediate that $x \in [a] \cap [b]$ implies that x is a common instance of a and b. Also, by [13] p.383, for any (unifiable) pair of terms there is a unique (up to a renaming substitution) most common instance, which we denote $a \wedge b$. Then by definition of most common instance of a pair of terms, x is a common instance of a and b implies that there exists a substitution κ satisfying $x = \kappa(a \wedge b)$. Therefore, $x \in [a] \cap [b]$ iff $x \in [a \wedge b]$. Hence our result follows. □

9.2.2 Relations on a term language, and the clause semigroup

Definitions 9.4

We define $LP \subseteq L_\Sigma \times L_\Sigma$, the set of *linear pairs of L*, by $(u,v) \in LP$ iff $FV(u) = FV(v)$, and define a relation $<$ on LP by $(a,b) < (c,d)$ iff there exists exists a substitution σ satisfying $a = \sigma(c)$ and $b = \sigma(d)$.

Lemma 4 *$<$ is a preorder on LP but is not a partial order.*

Proof Given $(a,b) < (c,d)$ and $(c,d) < (e,f)$ in LP, we can find substitutions σ, τ such that $a = \sigma(c)$, $b = \sigma(d)$, and $c = \tau(e)$, $d = \tau(f)$. Therefore, $a = \sigma\tau(e)$, $b = \sigma\tau(f)$ and so $(a,b) < (e,f)$. Hence $<$ is a preorder on LP. However, $(u,v) < (x,y)$ and $(x,y) < (u,v)$ does not imply that $(u,v) = (x,y)$, as (x,u) and (y,v) may differ by a renaming substitution. Hence $<$ is not a partial order. □

Definitions 9.5

Given a linear pair $(x,y) \in LP$, we define its *substitution class* by

$$[x,y] = \{(\sigma(x), \sigma(y)) : \sigma \in Subst(L)\} = \{(u,v) : (u,v) < (x,y)\}$$

We then define a relation $\leq$ on substitution classes of members of LP by

$$[a,b] \leq [c,d] \Leftrightarrow \forall (u,v) \in [a,b]\, \exists (x,y) \in [c,d] \ s.t.\ (u,v) < (x,y).$$

Lemma 5 *$\leq$ is a partial order on substitution classes of LP.*

Proof Given $[a,b] \leq [c,d]$ and $[c,d] \leq [e,f]$, then for all $(x,y) \in [a,b]$ we can find $(z,t) \in [c,d]$ such that $(x,y) < (z,t)$. Also, by definition of $[c,d] \leq [e,f]$, we can find $(u,v) \in [e,f]$ such that $(z,t) < (u,v)$. Therefore, $(x,y) < (u,v)$, and so the relation τ is transitive. Also, given $[a,b] \leq [c,d]$, and $[c,d] \leq [a,b]$, then a and c, and b and d must differ by (the same) renaming substitution, and it is immediate that $[a,b] = [c,d]$. Therefore, $\leq$ is a partial order.□

Note that each substitution class of a linear pair is a relation on L. We consider the set of substitution classes of LP, together with the usual composition of relations,

$$SR = \{(z,x) : \exists y\ s.t.(z,y) \in S, (y,x) \in R\}.$$

In terms of substitution classes of linear pairs, this becomes

$$[u,v][w,x] = \{(c,a) : \exists b \in L_\Sigma\ s.t.\ (c,b) \in [u,v], (b,a) \in [w,x]\}.$$

We will demonstrate that the set of substitution classes of linear pairs (together with the empty relation, denoted 0) is closed under this composition, and forms an inverse monoid. However, we first require the following:

Lemma 6 *For all $a, b \in L_\Sigma$, there exists $u, v \in L_\Sigma$ satisfying $[a] = [u]$, $[b] = [v]$, and $FV(u) \cap FV(v) = \emptyset$.*

Proof Denote the set of variables of Σ by $\{x_i\}_{i=0}^{\infty}$, and consider arbitrary $A, B \subseteq \{x_i\}_{i=0}^{\infty}$. Consider the renaming substitutions $\alpha(x_i) = x_{2i}$ and $\beta(x_j) = x_{2j+1}$. Then it is immediate that $\alpha(A) \cap \beta(B) = \emptyset$, and so the extensions of the maps α, β to the whole term language (which we also denote by α and β) satisfy $FV(\alpha(a)) \cap FV(\beta(b)) = \emptyset$, and our result is an application of this. □

Theorem 7 *The set of substitution classes of linear pairs (including the empty relation 0), together with the above composition, is an inverse submonoid of $I(L)$.*

Proof For any $(x, y) \in [u, v]$, $x = \sigma(u)$, $y = \sigma(v)$, for the same substitution σ, by definition of substitution classes of linear pairs. Therefore, as u, v have the same free variables, any element $(x, y') \in [u, v]$ must satisfy $y' = \sigma(v) = y$. Similarly, any element (x', y) must satisfy $x' = \sigma(u) = x$. Therefore, $[u, v]$ is a partial bijective relation. and so the set of substitution classes of members of P is a subset of the set of partial bijective maps on L.

Consider linear pairs (a, b) and (c, d) satisfying $[a, b][c, d] \neq 0$. We can assume without loss of generality (by Lemma 6), that $FV(b) \cap FV(c) = \emptyset$. Then

$$\{s : (r, s) \in [a, b], (s, t) \in [c, d]\} = \{s : s = \sigma(b) = \tau(c)\} = [b] \cap [c] = [b \wedge c].$$

So, we can find substitutions γ, δ satisfying $\gamma(b) = b \wedge c$ and $\delta(c) = b \wedge c$. This implies that $FV(\gamma(a)) = FV(\delta(d))$ and so $(r, t) \in [a, b][c, d]$ iff $(r, t) < (\delta(a), \gamma(d))$ Therefore $[a, b][c, d] = [mgu(b, c)(a), mgu(c, b)(d)]$ and hence the set of substitution classes of members of LP is closed under composition.

So, we have seen that the substitution classes of members of LP are partial bijective relations on the set L_Σ, and are closed under composition. Also, it is immediate that $[a, b]^{-1} = [b, a]$, from the description in terms of partial bijective relations. Finally, to see that the identity of $I(L)$ is a substitution class of a linear pair, note that for an arbitrary variable symbol x, $t < x$ for every term $t \in L$. Therefore, $[a, b][x, x] = [a, b] = [x, x][a, b]$, and hence our result follows. □

Definitions 9.6

Following the terminology of [22], we call the set of substitution classes of linear pairs, together with the above composition, the *clause semigroup on* L, which we denote $Cl(L)$. This semigroup (or rather, the contracted semigroup ring $\mathbb{Z}Cl(L)$) is used by J.-Y. Girard in [22] to construct a dynamical model of the whole of linear logic (with significant modifications); see [22] for more details on how linear logic is represented.

Note that this semigroup has also been constructed, as a specific example of a general construction of inverse semigroups from category actions on sets, in [40], where a description of the polycyclic monoids in the same terms is also given.

9.3 The clause semigroup in the Geometry of Interaction

We study how the clause semigroup is used in the third part of the Geometry of Interaction series of papers, [22]. We do not study the representations of the logical operators, for the following reasons:

(i) The representation of linear logic is not standard — in particular, the representation of the ? rule requires significant modifications of the syntax of the sequent calculus, and applications of the cut rule require the introduction of a new constant, $\flat$.

(ii) The algebra of clauses can be thought of as the 'operations behind the Geometry of Interaction' as all the operations of linear logic (if the claim of [22] is correct) can be modelled using the following structures, assuming certain conditions on the term language, that we will consider in what follows.

9.3.1 The action of the clause semigroup on the term language

Definitions 9.7

In [22], J-Y Girard defines the (partial) action of $Cl(L)$ on G, the set of ground terms of L, as follows: Consider $[P, Q] \in Cl(L)$ and $g \in G$. The action is defined by $[P, Q](g) = \theta(P)$, where $\theta = mgu(g, Q)$. Note that this is also a ground term, since if g unifies with Q, then $\theta = mgu(g, Q)$ is also a unifier of P and Q. Hence $\theta(P)$ is a ground term. However, J.-Y. Girard does not prove that this action is independent of the representative of $[P, Q]$. To prove this, we require the following:

First note that by definition, there is a bijection between substitution classes of members of L, and idempotents of $Cl(L)$, given by $[T] \mapsto [T,T]$. Next, for an arbitrary inverse semigroup, the following result is classical (see [30]).

Lemma 8 *For any $e, x \in S$ where S is an inverse semigroup, and $e^2 = e$, there exists $e' \in S$ such that $xe = e'x$.*

Proof Define $e' = xex^{-1}$, so that $e'x = xex^{-1}x = xx^{-1}xe = xe$. □

The application of this result to the clause semigroup gives the following:

Proposition 9 *$[A,B][t,t] = [t',t'][A,B]$, where $[t',t'] = [mgu(t,B)(A), mgu(t,B)(A)]$, for any $[t,t], [A,B] \in Cl(L)$.*

Proof $[A,B][t,t] = [A,B][t,t][A,B]^{-1}[A,B]$. However,

$$[A,B][t,t][B,A] = [A,B][mgu(t,B)(t), mgu(t,B)(A)] = [mgu(t,B)(A), mgu(t,B)(A)] = [t',t'].$$

Therefore, our result follows. □

Definitions 9.8

We use the above result to define the *action of $Cl(L)$ on substitution classes* by $[A,B]([T]) = [mgu(T,B)(A)]$ and from above, this is well-defined, and unique.

Proposition 10 *The action of the clause semigroup $Cl(L)$ on the ground terms G, as defined by J.-Y. Girard in [22] to be $[A,B](g) = mgu(B,g)(A)$ is independent of the representative of $[A,B]$ chosen.*

Proof First note that, by definition of substitution classes of a term language, $[g] = \{g\}$, for all ground terms g. Next note that, from above, the action of $Cl(L)$ on substitution classs of a term language is well-defined, and restricts to J.-Y. Girard's definition for ground terms. Our result then follows from the bijection between ground terms, and the substitution classes they determine. □

Finally, note that the action of $Cl(L)$ on substitution classes of L (and hence Girard's action of $Cl(L)$ on ground terms) is compatible with composition; that is:

Proposition 11 *For all $[A,B], [C,D] \in Cl(L)$ and $T \in L$, denote $[A,B]([T])$ by $[T']$. Then $[C,D]([T']) = ([C,D][A,B])([T])$.*

Proof Denote $[A,B]$, $[C,D]$, $[T,T]$ and $[T',T']$ by x, y, e and e' respectively, for clarity. Then $ye'y^{-1} = yxex^{-1}y$, and so $[C,D]([T']) = ([C,D][A,B])([T])$. □

9.3.2 Self-similarity and the clause semigroup

In [22], J.-Y. Girard specifies two conditions on a term language that he requires for his representation of full linear logic. These are as follows:

1. The term language is required to have a binary operation $\odot$, and two constant symbols g and d.

2. The ranked alphabet Σ is required to have a countably infinite number of terms of rank 2, which we denote $A_i(\ , \ )$ for $i \in \mathbb{N}$. (We use a different notation to [22], to avoid a conflict with previously defined concepts).

We demonstrate how these two conditions relate to the ideas of self-similarity (although not in a well-defined categorical sense) as presented in previous chapters.

Proposition 12 *For any term language L_Σ, where Σ has a binary function symbol $\odot$ and two constant symbols, g, d, there exists an embedding of P_2 into $Cl(L)$.*

Proof Given a variable symbol x, we define clauses $p = [x, g \odot x]$ and $q = [x, d \odot x]$. Their generalised inverses are given by $p^{-1} = [g \odot x, x]$ and $q^{-1} = [d \odot x, x]$. From the definition of composition in the clause semigroup we can see that $pp^{-1} = [x, x] = qq^{-1}$, and the clause $[x, x]$ has been shown to be the identity of $Cl(L)$; also, $pq^{-1} = 0 = qp^{-1}$ since $g \odot x$ and $d \odot x$ are non-unifiable. Therefore, the elements p^{-1}, q^{-1} satisfy the axioms for the generators of P_2, and as polycyclic monoids are congruence-free, they generate an inverse submonoid of $Cl(L)$ that is isomorphic to P_2. □

Corollary 13 *For any term language satisfying the above condition, there exists an embedding of P_α into $Cl(L)$, for any countable α.*

Proof Immediate by the above result, and the embeddings of P_n (for all $n \in \mathbb{N}$) and P_∞ into P_2, as found in Theorems 8 and 9 of Chapter 1. □

The second condition then gives us the following result:

Proposition 14 *Given a term language L_Σ, where Σ has a countably infinite number of terms of rank 2, which we denote $A_i(\ , \ )$ for $i \in \mathbb{N}$, then there exists a countably infinite number of distinct injective functions from $L \times L$ to L with disjoint images.*

Proof Consider a binary term A_i, and define a map $\langle\ ,\ \rangle_i : L \times L \to L$ by $\langle a, b\rangle_i = A_i(a, b)$. Then by definition of a term language, $A_i(a, b) = A_i(a', b')$ implies that $a = a'$ and $b = b'$. Therefore, the map $\langle\ ,\ \rangle_i$ is injective for all $i \in \mathbb{N}$. Next note that the A_i are distinct, so $A_i(x, y) = A_j(x', y')$ implies that $i = j$, and from above, $x = x', y = y'$. Therefore, the functions $\langle\ ,\ \rangle_i$ for $i \in \mathbb{N}$ all have distinct images. □

9.3.3 Girard's execution and resolution formulæ and the clause semigroup

We study how computation is modelled in the clause semigroup by J.Y. Girard, in the third part of the Geometry of Interaction series ([22]). This is by means of the 'execution' and 'resolution' formulae, which are defined in a very similar way to Definitions 7.4 of Chapter 7. However, we use the embedding of the clause semigroup in $I(L)$, even though [22] (implicitly) uses the contracted semigroup ring $\mathbb{Z}(Cl(L))$. This is because

(i) We are interested in the action of the clause semigroup on the term language L,

(ii) The resolution formula and the categorical trace may not always be defined in $\mathbb{Z}(Cl(L))$ (because they may involve infinite sums) although they are always defined in $I(L)$ (because infinite disjoint unions are well-defined).

In their original form (taken from [22]), the resolution and execution formulæ were defined on the contracted semigroup ring of the clause semigroup, in an analogous way to the execution and resolution formulæ used in the first two Geometry of Interaction papers (see Definitions 7.4 and Theorem 1 of Chapter 7 for an exposition of this), as follows:

The resolution formula has 2 inputs, U, σ, formal sums of pairwise disjoint elements of members of $Cl(L)$, that are required to satisfy the following conditions

- σ is a partial symmetry; that is, $\sigma^{-1} = \sigma$.
- σU is nilpotent; that is, there exists $n \neq 0$ satisfying $(\sigma U)^n = 0$.

The Execution formula is then defined to be the following sum in $\mathbb{Z}Cl(L)$

$$Ex(U, \sigma) = \sum_{k=0}^{n-1} U(1 - \sigma U)^{-1} = U \sum_{k=0}^{n-1} (\sigma U)^k,$$

and the Resolution formula is defined to be the following projection of the Execution formula

$$Res(U, \sigma) = (1 - \sigma^2)Ex(U, \sigma)(1 - \sigma^2) = \sum_{k=0}^{n-1} (1 - \sigma^2)U(\sigma U)^k(1 - \sigma^2).$$

As we are working in the monoid of relations on a term language L, we make the following definitions, by analogy with the above constructions in $\mathbb{Z}Cl(L)$:

Definitions 9.9

The *execution formula* is defined by $Ex(U,\sigma) = U(\sigma U)^*$, and the *resolution formula* is a projection of this, defined by $Res(U,\sigma) = (\sigma^2)^\perp U(\sigma U)^*(\sigma^2)^\perp$. We wish to study these formulæ in terms of the categorical trace. However, as there is no M-monoid structure apparent on (the disjoint closure of) the clause semigroup, we take the following naive approach to the categorical trace on $I(L)$: Recall that $CL(L)$ is an inverse submonoid of $I(L)$, from Theorem 7 above. Then given $f \in I(L)$, together with two disjoint subsets, $G, D \subseteq L$, we can construct a map $\Lambda_{G,D}$ from $I(L)$ to $I(G \sqcup D)$ by $\Lambda_{G,D}(f) = \begin{pmatrix} 1_G f 1_G & 1_G f 1_D \\ 1_D f 1_G & 1_D f 1_D \end{pmatrix}$.

Proposition 15 *$\Lambda_{G,D}$, as defined above, is an inverse monoid homomorphism, and is an isomorphism when $G \cup D = L$.*

Proof Consider the local submonoid of $I(L)$ determined by the idempotent $e = 1_G \vee 1_D$. Then it is immediate from Proposition 5 and Theorem 6 of Chapter 5 that Λ is a map from $I(L)$ to (a monoid isomorphic to) $eI(L)e$, given by $f \mapsto efe$. Hence, it is immediate that Λ_{DG} is a monoid homomorphism, and the second part of our result follows by $e = 1$, when $G \cup D = L$. □

We have seen in Theorem 2 of Chapter 6 that the categorical trace can be applied to partial bijective maps written in matrix form to give another partial bijective map in $I(G) \leq I(L)$, as follows:

$$Tr^D_{G,G}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \vee \bigvee_{i=0}^{\infty} bd^ic \in I(G).$$

This then gives us the categorical interpretation of the Resolution formula:

Theorem 16 *Consider $U, \sigma \in I(L)$, where $\sigma = \sigma^{-1}$. We denote the partial identity on $D = dom(\sigma)$ by π, and the partial identity of its complement $G = L \setminus dom(\sigma)$ by $\pi^\perp$. Then the resolution formula satisfies $Res(U,\sigma) = Tr^D_{G,G}(\Lambda_{\pi^\perp,\pi}((\sigma \vee \pi^\perp)U)$.*

Proof First note that

$$\Lambda(\sigma \vee \pi^\perp) = \begin{pmatrix} \pi^\perp\pi^\perp & 0 \\ 0 & \pi\sigma\pi \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sigma \end{pmatrix}, \ \Lambda(U) = \begin{pmatrix} \pi^\perp U\pi^\perp & \pi^\perp U\pi \\ \pi U\pi^\perp & \pi U\pi \end{pmatrix}.$$

Therefore,

$$\Lambda((\sigma \vee \pi^\perp)U) = \begin{pmatrix} \pi^\perp U\pi^\perp & \pi^\perp U\pi \\ \sigma\pi U\pi^\perp & \pi\sigma U\pi \end{pmatrix},$$

and so $Tr^D_{G,G}(\Lambda_{\pi^\perp,\pi}((\sigma \vee \pi^\perp)U) = \pi^\perp U \pi^\perp \vee \pi^\perp U(\pi\sigma U\pi)^*\sigma\pi U\pi^\perp$, and as $dom(\sigma) = im(\sigma) = \pi$, this is equal to $\pi^\perp U\pi^\perp \vee \pi^\perp U(\sigma U)^* r(\sigma U)\pi^\perp = \pi^\perp U(\sigma U)^*\pi^\perp = Res(U, \sigma)$. Therefore our result follows. □

Chapter 10

Applications of the trace to automata

10.1 Introduction

We apply the results of Chapter 6 on the trace on the category of relations, and the compact closed category **IntRel**, to the theory of two-way automata. We show how the global transition relations for a two-way automaton[1] can be derived in a categorical way, in terms of the construction of the compact closed monoidal category **IntRel** from the traced symmetric monoidal category **Rel**. We then show how the action of a two-way automaton can be reconstructed from the action of the left and right moving one-way automata determined by it, using a semigroup-theoretic version of Girard's resolution formula, similar to that found in Chapter 9.

The results of this Chapter can either be thought of as a concrete application of the trace, compact closure, and the resolution formula, or can be thought of as an automata-theoretic interpretation of (some of) the constructions presented in Chapters 7 to 9.

10.2 The basic theory of automata

Definitions 10.1

An *automaton* A is specified by the following data: $A = (Q, \Sigma, \circ, q_0, F)$, where

- Q is the *set of states*,
- Σ is the *input alphabet*,

[1]We use a slightly different model of 2-way automata to [3], who in turn uses a slightly different model to [28]. However the equivalences of the models of [3] and [28] is proved in [3], and the equivalence of our model to that of [3] is almost trivial.

- $\circ : \Sigma \times Q \to P(Q)$ is the *next state function*,
- $q_0 \in Q$ is the *initial state*,
- $F \subseteq Q$ is the *set of terminal states.*

If $x \circ q$ is a single element set for all $q \in Q$, $x \in \Sigma$, the automaton A is called *deterministic*, and we refer to the next state function $\circ : \Sigma \times Q \to Q$.

The function $\circ : \Sigma \times Q \to P(Q)$ can be extended to to $\Sigma^* \times Q \to P(Q)$ (it is conventional to use the same notation for both) by

- $\lambda \circ q = q$, where λ denotes the empty word, and
- $wx \circ q = w \circ (x \circ q)$ for all $w \in \Sigma^*$ and $x \in \Sigma$.

This allows us to construct the *transition function*, $t : \Sigma^* \to B(Q)$ as follows:

$$(q', q) \in t(x) \ \Leftrightarrow\ q' \in x \circ q,$$

$$(q'', q) \in t(wx) \ \Leftrightarrow\ \exists q' \in x \circ q \ s.t.\ (q'', q') \in t(w).$$

Clearly, t is a monoid homomorphism from the free monoid Σ^* to the monoid of relations on Q, $B(Q)$, that satisfies $t(u) = t(v)$ iff $u \circ q = v \circ q$ for all $q \in Q$. The *transition congruence* on Σ^*, denoted $\sim$, is defined by

$$u \sim v \ \Leftrightarrow t(u) = t(v),$$

and the *transition monoid* of the automaton A is defined to be

$$\frac{\Sigma^*}{\sim} \ \cong\ Im(t) \ \leq\ B(Q).$$

A word $w \in \Sigma^*$ is said to be *recognised by* the automaton A if $(f, q_0) \in t(w)$ for some $f \in F$. The set of all words recognised by an automaton A is called the *language recognised by* A, written $Rec(A)$, and a language $L \subseteq \Sigma^*$ is called recognisable if $L = Rec(A)$ for some automaton A. The theory of recognisable languages of finite-state automata is well-known; see, for example, [28, 29].

We are interested in studying transitions of automata, rather than the languages recognised by them, so for the remainder of this chapter, we will refer to the automaton $A = (Q, \Sigma, \circ)$. An automaton specified solely by this data is uniquely determined by a monoid homomorphism from Σ^* into $B(Q)$.

10.2.1 Dual automata

Definitions 10.2

Given an automaton $A = (Q, \Sigma, \circ)$, we define its dual automaton $\Delta(A) = (Q, \Sigma, \bar{\circ})$, where $\bar{\circ} : Q \times \Sigma \to P(Q)$ is defined by, for all $x \in \Sigma$, $q \in Q$, $q\bar{\circ}x = x \circ q$. As before, there is a natural extension of this map to $\bar{\circ} : Q \times \Sigma^* \to P(Q)$ by

$$q\bar{\circ}\lambda = q,$$

$$q\bar{\circ}xw = (q\bar{\circ}x)\bar{\circ}w.$$

(Again, we use the same notation for both). This allows us to construct the *transition antihomomorphism* $\bar{t} : \Sigma^* \to B(Q)$, where $\bar{t}(x) = t(x)$ and $\bar{t}(wx) = \bar{t}(x)\bar{t}(w)$, for all $x \in \Sigma$ and $w \in \Sigma^*$. This antihomomorphism satisfies $\bar{t}(u) = \bar{t}(v)$ iff $q\bar{\circ}u = q\bar{\circ}v$ for all $q \in Q$.

Intuitively, we can think of an automaton A as receiving a word w of Σ^* written on a tape, and reading each symbol of w, from right to left, and applying the appropriate transition to its set of states. Conversely, its dual automaton $\Delta(A)$ can be though of as receiving a word w written on a tape, and reading each symbol of w from left to right, and applying the appropriate transition to its set of states. We refer to A as a *left moving automaton* and to $\Delta(A)$ a *right moving automaton.*

10.3 2-way automata

Note that the following definition of 2-way automata is that of J. Birget, [4]. However, he proves in [3] that his model of 2-way automata is equivalent to the model of 2-way automata found in [28].

Definitions 10.3

A *two-way automaton* is specified by the following data: $\mathbb{A} = (Q = Q_l \cup Q_r, \Sigma, \circ, q_0, F)$, where

- Q_l is the set of *left moving states* and Q_r is the set of *right moving states* (Note that we do *not* assume that $Q_l \cap Q_r = \emptyset$),
- Σ is the *input alphabet,*
- $\circ : \Sigma \times Q \to P(Q)$ is the transition function,
- $q_0 \in Q$ is the initial state,
- $F \subseteq Q$ is the set of terminal states.

As before, we are interested in the transitions of 2-way automata, rather than the languages recognised by them, so we will refer to 2-way automata as being of the form $\mathbb{A} = (Q = Q_l \cup Q_r, \Sigma, \circ)$. A *configuration* of the automaton $\mathbb{A}$ is given by a set of expressions of the form $u\ ^{(q)}\ v$, where $u, v \in \Sigma^*$, $q \in Q_l \cup Q_r$. This is drawn as a tape, together with a read head, labelled by a state, with the read head on the boundary between cells on the tape. The transition function has the following action on the set of configurations of the automaton:
Given a configuration

$$\dots a_{k-1}a_k\ ^{(q)}\ a_{k+1}a_{k+2}\ \dots$$

with $a_i \in \Sigma$, $q \in Q$, then the set of next configurations is given by the union of

$$\{a_{k-1}a_ka_{k+1}\ ^{(q')}\ a_{k+2}\ ,\ q' \in a_{k+1} \circ q : q \in Q_r\}$$

and

$$\{a_{k-1}\ ^{(q')}\ a_ka_{k+1}a_{k+2}\ ,\ q' \in a_k \circ q : q \in Q_l\}.$$

Two-way automata have the (possibly more appropriate) description as Turing machines that do not write. However, they have the same recognising power as one-way automata [45], so are usually studied in the context of automata theory.

10.3.1 Algebraic models of 2-way automata

The following algebraic method of studying transitions of 2-way automata is taken from [3].

Definitions 10.4

Given a 2-way automaton $\mathbb{A}$, as above, then for every word $w \in \Sigma$, there are four associated relations, as follows:

- $[\rightleftharpoons w] \in B(Q)$, given by $(q', q) \in [\rightleftharpoons w]$ if and only if there exists a computation of $\mathbb{A}$, starting in configuration $^{(q)}w$ and finishing in configuration $^{(q')}w$, with $q' \in Q_l$ and $q \in Q_r$.
- $[-w \rightarrow] \in B(Q)$, given by $(q', q) \in [-w \rightarrow]$ if and only if there exists a computation of $\mathbb{A}$, starting in configuration $^{(q)}w$ and finishing in configuration $w^{(q')}$, with $q, q' \in Q_r$.
- $[\leftarrow w-] \in B(Q)$, given by $(q', q) \in [\leftarrow w-]$ if and only if there exists a computation of $\mathbb{A}$, starting in configuration $w^{(q)}$ and finishing in configuration $^{(q')}w$, with $q, q' \in Q_l$.
- $[w \rightleftharpoons] \in B(Q)$, given by $(q', q) \in [w \rightleftharpoons]$ if and only if there exists a computation of $\mathbb{A}$, starting in configuration $w^{(q)}$ and finishing in configuration $w^{(q')}$, with $q \in Q_l$ and $q' \in Q_r$.

These relations were first introduced explicitly in [3], where they are considered in terms of the monoid $B(Q_r \sqcup Q_l)$; however, they feature implicitly in the earlier work [45].

10.3.2 The composition of global transition relations

The composition of global transition relations of a two-way automaton, determined by words in Σ^*, was first described in [4], where the following theorem is proved:

Theorem 1 *Given a two-way automaton* $\mathbb{A} = (Q = Q_l \cup Q_r, \Sigma, \circ)$, *and* $u, v \in (\Sigma)^*$, *then the global transition relations for the composite* uv *are defined in terms of the global transition relations of* u *and* v, *as follows:*

- $[-uv \rightarrow] = [-v \rightarrow]([u \rightleftharpoons][\rightleftharpoons v])^*[-u \rightarrow]$,
- $[uv \rightleftharpoons] = [v \rightleftharpoons] \cup [-v \rightarrow][u \rightleftharpoons]([\rightleftharpoons v][u \rightleftharpoons])^*[\leftarrow v-]$,
- $[\rightleftharpoons uv] = [\rightleftharpoons u] \cup [\leftarrow u-][\rightleftharpoons v]([u \rightleftharpoons][\rightleftharpoons v])^*[-u \rightarrow]$,
- $[\leftarrow uv-] = [\leftarrow u-]([\rightleftharpoons v][u \rightleftharpoons])^*[\leftarrow v-]$.

Proof Found in [4]. □

Definitions 10.5

Note that any word $w \in (\Sigma)^*$ determines 4 relations $a, b, c, d \in B(Q)$, where $a = [\leftarrow w-]$, $b = [\rightleftharpoons w]$, $c = [w \rightleftharpoons]$, and $d = [-w \rightarrow]$. Then from Proposition 5 of Chapter 5, these relations determine a single relation in $B(Q \sqcup Q)$, which we write as $[w] = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and refer to as *the global transition relation determined by* w.

Also, by Definitions 6.5 of Chapter 6, any relation $R \in \mathbf{Rel}(X \sqcup V, Y \sqcup U)$ uniquely determines, and is uniquely determined by, a morphism $R \in \mathbf{IntRel}((X, U), (Y, V))$. Therefore, the global transition relation $[w]$ uniquely determines a morphism from (Q, Q) to (Q, Q) in **IntRel**. This gives us the following result.

Theorem 2 *Given* $u, v \in (\Sigma)^*$, *then* $[vu] = [v] \circ [u]$, *where* $\circ$ *is the representation in* **Rel** *of the composition of* **IntRel** *given in Definitions 6.5 of Chapter 6.*

Proof Given

$$[u] = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad [v] = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

then from the composition rules for global transition relations given above,

$$[vu] = \begin{pmatrix} e(bg)^*a & f \cup e(bg)^*bh \\ c \cup d(gb)^*ga & d(gb)^*h \end{pmatrix}.$$

Therefore, from the definition of the composition in the category **IntRel**, $[uv] = [u] \circ [v]$, and so our result follows. □

Corollary 3 *For all global transition relations* $[u], [v], [w]$ *of some 2-way automaton,*

$$([u] \circ [v]) \circ [w] = [u] \circ ([v] \circ [w]).$$

Proof By definition, $\mathbf{IntRel}((Q,Q),(Q,Q))$ is the endomorphism monoid of the object (Q,Q) of **IntRel**. Therefore, as **IntRel** is a category, the composition is associative. □

Definitions 10.6

We refer to the image of Σ^* under the map $[\ ] : \Sigma^* \to \mathbf{IntRel}((Q,Q),(Q,Q))$ as the *global transition monoid* of the two-way automaton $\mathbb{A}$.

Theorem 4 *Finite state 2-way automata can be simulated by finite state one-way automata. — Note that by 'simulated by', we mean that for every finite state 2-way automaton, we can construct a finite state one-way automaton with an isomorphic transition monoid.*

Proof By the construction of **IntRel** (Definitions 6.5, Chapter 6), $\mathbf{IntRel}((Q,Q),(Q,Q))$ is finite, For every finite set Q. Also, by [30], there is a representation of every finite monoid in the monoid of relations on a finite set. Therefore, the global transition monoid of a 2-way automaton is uniquely determined by a homomorphism from Σ^* into a finite monoid of relations on a set. Therefore, our result follows by the definition of a one-way automaton (Definitions 10.1). □

10.4 An alternative model of 2-way automata

We introduce a slightly different model of 2-way automata to that given in [3]. However, we demonstrate that the two definitions are equivalent.

Definitions 10.7

We define a two-way automaton to be specified by the data $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$, where Q is the set of states, Σ is the input alphabet, configurations are as in Definitions 10.3, and we have two next-state functions, $\circ_l : \Sigma \times Q \to P(Q)$, and $\overline{\circ_r} : Q \times \Sigma \to P(Q)$. These act on the set of configurations as follows:
Given a configuration $\ldots x_{i-2}x_{i-1} \ ^{(q)}\ x_i x_{i+1} \ldots$, then the next set of configurations is given by

$$\{ \ldots x_{i-2} \ ^{(q')}\ x_{i-1}x_i \ldots : q' \in x_{i-1} \circ_l q\} \cup \{ \ldots x_{i-1}x_i \ ^{(q'')}\ x_{i+1} \ldots : q'' \in q\overline{\circ_r}x_i\}.$$

The equivalence between the model of 2-way automata due to J.-C. Birget (Definitions 10.3), and the above model of 2-way automata above is as follows:
It is clear that the above model contains Birget's definition as a special case; given an automaton $(Q = Q_l \cup Q_r, \Sigma, \circ)$ as defined by Birget, we can define an automaton as above by

$$(Q, \Sigma, \circ_l : \Sigma \times Q \to P(Q), \overline{\circ_r} : Q \times \Sigma \to P(Q))$$

where

$$x \circ_l q = \begin{cases} x \circ q & x \in Q_l \\ \emptyset & \text{otherwise} \end{cases}$$

and

$$q \circ_r x = \begin{cases} x \circ q & x \in Q_r \\ \emptyset & \text{otherwise} \end{cases}$$

Conversely, given an automaton as defined above, we define $Q_l = \{q \in Q : \Sigma \circ_l q \neq \emptyset\}$ and $Q_r = \{q \in Q : q\overline{\circ_r}\Sigma \neq \emptyset\}$. However, note that in our model we do not have $x \circ_l q = q \circ_r x$, in general. This corresponds to the following:
Consider $q \in Q_l \cap Q_r$, together with the computations

$$^{(q)}x \;\Rightarrow\; \{x^{(q_i)}\} \text{ and } x^{(q)} \;\Rightarrow \{\, ^{(q_j)}x\}$$

In Birget's model, $\{q_i\}$ must be the same as $\{q_j\}$; in our model they may differ.

The translation is then as follows: Given an automaton $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$, we define an automaton (as specified by Birget) by $(Q_l \sqcup Q_r, \Sigma, \circ : \Sigma \times Q \to Q)$, where Q_l and Q_r are as above, $Q_l \sqcup Q_r = Q_l \times \{0\} \cup Q_r \times \{1\}$, and $\circ$ is defined by

$$x \circ (q, i) = \begin{cases} x \circ_l q & i = 0 \\ q \circ_r x & i = 1 \end{cases}$$

This then gives an equivalent two-way automaton where the left-moving states are distinct from the right-moving states, so the above problem does not arise. This is an example of the use of a construction of J.-C. Birget that is used for constructing deterministic two-way automata from non-deterministic two-way automata. See [3] for details of this.

In our model, $A_l = (Q, \Sigma, \circ_l)$ is the restriction of $\mathbb{A}$ to its left-moving action, and $A_r = (Q, \Sigma, \circ_r)$ is the restriction of $\mathbb{A}$ to its right-moving action. So, A_l is a (one-way) automaton that reads an input word from right to left, and A_r is a (one-way) dual automaton that reads an input word from left to right. Note that this definition of two-way automata allows us to extend the definition of the direction symmetry operator Δ to two-way automata, and in this case, $\Delta^2(\mathbb{A}) = \mathbb{A}$.

This operator also has a categorical interpretation. Recall the definition of the left dual of a morphism in **IntRel**, from Definitions 6.5, Chapter 6. This allows us to describe the Δ operator categorically.

Proposition 5 *Let* $[w]$ *be the global transition relation of a word* w *for the two-way automaton* $\mathbb{A}$. *Then the global transition relation of the word* w *for the dual automaton* $\Delta(\mathbb{A})$ *is given by* $[w]^\vee$, *the left dual of* $[w]$.

Proof Immediate from the description of the global transition relations of a two-way automaton, and the description of the Δ operator as interchanging the rôles of the left and right moving states. □.

10.5 The one-way automata associated with a 2-way automaton

The alternative definition of a 2-way automaton given above allows a natural decomposition of a two-way automaton into two one-way automata, as follows:

Definitions 10.8

Given a two-way automaton, $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$, we define the *associated left and right moving automata* A_l and A_r, by $A_l = (Q, \Sigma, \circ_l)$ and $A_r = (Q, \Sigma, \overline{\circ_r})$. We can also construct $t_l : \Sigma^* \rightarrow B(Q)$, the transition homomorphism of A_l, and $\overline{t_r} : \Sigma^* \rightarrow B(Q)$, the transition antihomomorphism of A_r.

10.6 Global transition relations, and Girard's resolution formula

In what follows, we demonstrate how the global transition monoid of $\mathbb{A}$ can be reconstructed from the next-state functions of A_l and A_r. First note that the action of A_l is uniquely determined by the action of t_l on Σ, and the action of A_r is uniquely determined by the action of $\overline{t_r}$ on Σ. Also, as we know that the composition of global transition relations is the composition of **IntRel** at the monoid $\mathbf{IntRel}((Q,Q),(Q,Q))$, the action of $\mathbb{A}$ is uniquely determined by the images of the elements of Σ under the $[\,] : \Sigma^* \rightarrow \mathbf{IntRel}((Q,Q),(Q,Q))$ map.

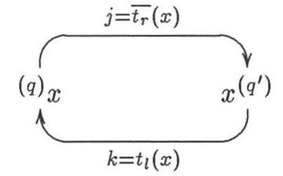
Definitions 10.9

Given a two-way automaton $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$, we define the *left and right projection maps* to be $\pi_l, \pi_r : Q \to Q$, where π_l is the partial identity on $Q_l \subseteq Q$, and π_r is the partial identity on $Q_r \subseteq Q$.

Proposition 6 *Let $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$, be a two-way automaton, and let $x \in \Sigma$. Then if we denote $t_l(x)$ by k, and $\overline{t_r}(x)$ by j, the global transition of x is given by*

$$[x] = \begin{pmatrix} \pi_l k(jk)^* \pi_l & \pi_l (kj)(kj)^* \pi_r \\ \pi_r (jk)(jk)^* \pi_l & \pi_r j(kj)^* \pi_r \end{pmatrix}.$$

Proof First note that the possible transitions on a singleton word x can be summarised by the following diagram:



Then all the possible transitions from configurations of the form ${}^{(q)}x$ to configurations of the form ${}^{(q')}x$ are

$$kj, (kj)^2, (kj)^3, \ldots.$$

Therefore, $[\rightleftharpoons x]$ is the intersection of $(kj)(kj)^*$ with $Q_l \times Q_r$, and so $[\rightleftharpoons x] = \pi_l(kj)(kj)^*\pi_r$. A similar proof gives $[x \rightleftharpoons] = \pi_r(jk)(jk)^*\pi_l$. Also, all possible transitions from configurations of the form ${}^{(q)}x$ to $x^{(q')}$ are

$$j, jkj, j(kj)^2, j(kj)^3, \ldots$$

so $[-x \rightarrow]$ is the intersection of $j(kj)^*$ with $Q_r \times Q_l$. Therefore $[-x \rightarrow] = \pi_r j(kj)^* \pi_l$, and a similar proof then gives $[\leftarrow x-] = \pi_l k(jk)^* \pi_r$. Therefore, the global transition relation of x is as above. So we have deduced the required formula for $[x]$. □

Theorem 7 *Let $\mathbb{A} = (Q, \Sigma, \circ_l, \overline{\circ_r})$ be a two-way automaton. Then the global transition relation of a member of Σ is given by the following version of Girard's resolution formula in the monoid of relations on $Q \times Q$:*

$$[x] = Res(U, \sigma) = \pi U(1 - \sigma U)^{-1}\pi = \pi U(\sigma U)^* \pi,$$

where

$$\pi = \begin{pmatrix} \pi_l & 0 \\ 0 & \pi_r \end{pmatrix}, \quad U = \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

(As before, we define $j = \overline{t_r}(x)$, $k = t_l(x)$*).*

Proof First note that

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} = \begin{pmatrix} 0 & j \\ k & 0 \end{pmatrix},$$

so

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \right]^{2n} = \begin{pmatrix} (jk)^n & 0 \\ 0 & (kj)^n \end{pmatrix},$$

and

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \right]^{2n+1} = \begin{pmatrix} 0 & (jk)^n j \\ (kj)^n k & 0 \end{pmatrix}.$$

Therefore,

$$\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \right]^{*} = \begin{pmatrix} (jk)^* & 0 \\ 0 & (kj)^* \end{pmatrix} \cup \begin{pmatrix} 0 & j \cup (jk)^* k \\ k \cup (kj)^* j & 0 \end{pmatrix}$$

$$= \begin{pmatrix} (jk)^* & (jk)^* j \\ (kj)^* k & (kj)^* \end{pmatrix},$$

and so

$$\begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \right]^{*} = \begin{pmatrix} k(jk)^* & (kj)(kj)^* \\ (jk)(jk)^* & j(kj)^* \end{pmatrix}.$$

Therefore, $Res(U, \sigma) =$

$$\begin{pmatrix} \pi_l & 0 \\ 0 & \pi_r \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} k & 0 \\ 0 & j \end{pmatrix} \right]^{*} \begin{pmatrix} \pi_l & 0 \\ 0 & \pi_r \end{pmatrix}$$

$$= \begin{pmatrix} \pi_l k (jk)^* \pi_l & \pi_l (kj)(kj)^* \pi_r \\ \pi_r (jk)(jk)^* \pi_l & \pi_r j (kj)^* \pi_r \end{pmatrix},$$

and we have seen from above that this is the global transition relation of x. □

Therefore, we have proved that the global transition relations of a two-way automaton can be reconstructed from its associated left and right moving automata, using a version of Girard's resolution formula, and the composition on the endomorphism monoid of a compact closed category.

Appendix A

Critique, and ideas for future work

We review the previous 10 Chapters, with specific reference to further work that can be developed from the concepts introduced.

A.1 Introduction

The stated aim of this thesis was to understand the algebra and category theory arising from the Geometry of Interaction series of papers ([20, 21, 22]). The attempt to do so led to a number of interesting new structures, both algebraic and categorical, so it can be considered a worthwhile project. However, many of the new structures, and results, are tangentially related to the original program — so much so that we have not had the time or the space to develop them for their own sake. In light of this, our review of the thesis mainly takes the form of ideas for extra work, as well as a critique of the way we chose to study the concepts introduced.

A.2 Review of Chapters

Chapter 1

As an introduction of basic concepts in algebra and category theory, the first few sections are basically self-contained. However, the disjoint closure (in its various forms) is new, and can be considered to be the inverse semigroup analogue of the power set construction. Extra work is needed to consider the properties of this map; in particular, the relation of the natural partial order to set-theoretic inclusion. To study the categorical trace, as presented in Chapter 6, it is also worth studying the conditions for terms built up from elements of an inverse semigroup, together with the Kleene star, to be members of the disjoint closure in a symmetric inverse monoid.

Potential for extra work also arises with the introduction of P_2, and its embeddings into symmetric inverse monoids. Our work mainly considered the monoids P_2 and P_∞; however, considering the strong embeddings of P_n into $I(X)$, given by a strong embedding of P_a into $I(X)$ would be an interesting problem; we can reasonably expect the number of distinct embeddings to be given by the enumeration of distinct binary trees (i.e. the Catalan numbers) when $a = 2$, and presumably the general case would best be considered in terms of coding theory.

Chapter 2

At this point, we chose to develop the thesis in terms of embeddings into the natural numbers, rather than the Cantor set, for consistency with (some of) the structures of the Geometry of Interaction series; however, the Cantor set can be considered a more natural (and graphical) representation of self-similarity.

The self-embedding results on the natural numbers are basically self-contained, although they can of course be written in terms of the M-monoid structures of the endomorphism monoid of $\mathbb{N}$ in **Inj**. (For example, the embedding of P_2 into $I(\mathbb{N})$ determined by an embedding of P_∞ into $I(\mathbb{N})$ has a natural definition in terms of the ? operator introduced in Chapter 5).

The connection between the Cantor set and the natural numbers representations of polycyclic monoids is not as straightforward as it might appear; although the constructions of topologies on the Cantor set and $\mathbb{N}$ from an embedding of P_2 gives a bijection between the basic open sets of each, the underlying set of the Cantor set is uncountable, whereas that of the natural numbers, by definition, is countable. There is also a connection between the embeddings of P_2 chosen for the Cantor set, and the natural numbers; consider binary representations of the natural numbers. Let w be a word in $\{0, 1\}$ representing a natural number. Then $2 \times w$ is written as $w0$, $2 \times w + 1$ is written as $w1$, and the connection with the action of P_2 on the Cantor set is apparent.

Finally, the results generated by the embedding of P_2 into C^* algebras — in particular e^z as a fixed point of p, q, the connection with differentiation, and Euler's identity — are well worth further study. However this would require additional expertise in functional analysis.

Chapter 3

The results on ring theory are interesting concrete examples of self-similarity considerations. Although they follow almost trivially as corollories of categorical self-similarity, as presented in Chapter 4, they form a motivating example that is self-contained. However, all the results also seem to follow for semi-rings, so perhaps this chapter would have been better phrased in terms of additive categories, rather than just rings.

Also, we have not given any concrete examples of constructing the K_0 group of a ring using an embedding of P_2 — perhaps this could be used (along with the embedding of P_2 into $B(l^2)$ of Chapter 2) to demonstrate that the K_0 group of $B(l^2)$ is trivial.

Chapter 4

The concept of self-similarity in its categorical form needs a great deal of extra study. In particular, we would like to be able to prove that when the tensor category of a self-similar object is an X – category, (where X is symmetric monoidal, Cartesian closed, traced, compact closed, or whatever), the endomorphism monoid of the self-similar object is an X M-monoid, and an X – one-object category (without unit conditions) when the self-similarity is strong. Of course, the difficulty with this is formalising the idea precisely.

Also, some analogue of a coherence condition for a symmetric monoidal category with self-similarity morphisms at some object would be useful. This is presumably related to proving that the (weakly) self-similar category **S** is inverse, and not just regular – this is why we did not attempt to prove this, although we expect it to be true.

More concretely, we could develop the connections between self-similarity and embeddings of polycyclic monoids, in light of the result (from [38]) that idempotents of a Karoubi envelope are splitting idempotents. This would be connected with studying the endomorphism monoid of the distinguished self-similar element of the self-similar category **S** in the weak and strong cases. This would hopefully shed light on the category theory, via the congruence-freeness of polycyclic monoids.

Finally, it would be interesting, and probably enlightening, to study Dana Scott's original construction of a C-monoid in terms of categorical self-similarity. This would also involve finding the conditions for the tensor category of an object of a Cartesian closed category to be a Cartesian closed subcategory (without unit elements). A good starting point for this would be the interleaving embedding of P_2 he implicitly uses (again, see [38]).

Chapter 5

This Chapter is concerned with providing examples of previously defined concepts (although,as we later show, it also forms the algebraic framework for our models of Geometry of Interaction). It would be useful to find an entirely categorical interpretation of the various connections between the two distinct strong M-monoid structures on $I(\mathbb{N})$. A start has been made on this with the identification of the $!(a) = (1 \otimes a)$ map as a right fixed point homomorphism for $(I(\mathbb{N}), \oplus)$, however, much remains to be done.

Also, it is somewhat surprising that the commutativity morphism σ for $\otimes$ is not representable in terms of the disjoint closure of P_2 in $I(\mathbb{N})$, and a proof of this that did not depend on constructing a topology on the underlying set would clearly be more general. Similarly, it is surprising to find a naturally arising right fixed point homomorphism in $DC(P_2)$, but not a left one (although we have not proved that one does not exist). These results are presumably related.

Chapter 6

This Chapter introduces a large number of new structures, and as such, a great deal of further work remains to be done. In particular, we restricted our consideration to symmetric traced monoidal categories — and hence to compact closed categories, via the construction of Joyal, Street, and Verity. The natural thing to do is to consider the general case of braided traced monoidal categories — and hence tortile monoidal categories, via the same construction. This would necessarily depend heavily on the use of diagrammatic reasoning, and correctly defining braided M-monoids.

Also, a formalisation of the 'summing over all possible paths' explanation we gave for matrix multiplication, the trace, and composition in **IntRel** seems feasible. A method of computing such sums (or rather, unions and disjoint unions) seems possible using analogues of Kauffman's state summation in bracket polynomials, for splitting up diagrams — in fact, he presents a method of calculating matrix determinants from state summations in [36].

Work also remains to be done on understanding our alternative characterisation of compact closed categories – in particular, whether the axiom set can be simplified. Ideally, we would find an equivalent set of axioms that do not mention the units elements at all.

In terms of self-similarity, the most obvious omission is a method of constructing compact closed M-monoids directly from traced M-monoids. However, the proofs associated with the construction of compact closed categories from traced symmetric monoidal categories were phrased in terms of diagrammatic reasoning, so this would probably require an analogue of diagrammatic reasoning for M-monoids, and a proof of its validity. Also, we did not study the case when an object N was a very self-similar object of a compact closed category $(\mathbf{T}, \otimes)$ but $(\otimes \mathbf{N})$ was not freely generated in $\mathbf{T}$. This is because the specific case we required was freely generated, and this simplified calculations considerably.

Finally, the monoid $\mathbb{F}$ we introduced requires a great deal of extra study. Categorically, we have not given explicit descriptions of the κ and δ maps for it (although the dual on elements was enough for our analysis of cut-elimination in GOI1). In terms of semigroup theory, computing Green's relations, the semilattice structure of the idempotents, its representations in symmetric inverse monoids, and other standard inverse semigroup theoretic tools would shed more light on its structure. Also, it appears that (two copies of) the disjoint closure of P_2 in $I(\mathbb{N})$ is closed under all the operations presented, and perhaps this would give the correct definition of the inverse monoid behind the Geometry of Interaction 1.

On a more speculative level, the diagrammatic reasoning representation of the object / dual object creation and destruction operators of compact closed categories are very reminiscent of Feynman diagrams — as is the 'summing over all possible paths' interpretation. However, a great deal of speculation has been done on the connections between linear logic and theoretical physics without producing any concrete results (see [19] for comments on this).

Chapter 7

Although this Chapter is an exposition of pre-existing work, it does raise several interesting points:

Firstly, the question of the exact computing power of the GOI1 system remains open. Although it is an interpretation of (a variant of) multiplicative linear logic, and as such, Girard's system F is expressible in it, he also makes the comment 'Actually, an interpretation of pure lambda calculus seems at hand, since our system forgets types' — [20], and a similar, although not identical interpretation of untyped lambda calculus in terms of the same algebraic tools is given in [9]. On the other hand, we later demonstrate that the correct categorical model of GOI1 seems to be that of (one-object) compact closed categories, and the standard model of pure lambda calculus is one-object Cartesian closed categories. Also, it is well-known that Turing Machines have the same computing power as pure lambda calculus (and as such, should have the same categorical models), and we have used compact closed categories to model Turing machines that do not write (i.e. 2-way automata) in Chapter 10. Clearly, much work remains to be done.

Secondly, there is a concrete interpretation of the members of P_2 in terms of the basic operations of computers; the generators p, q, p^{-1}, q^{-1} can be considered to be the operations

- Push 0 onto a stack,
- Push 1 onto a stack,
- Pop the first member of a stack, and check for equality to 0,
- Pop the first member of a stack, and check for equality to 1,

respectively. These operations are fundamental to all computer design, and there is a similar way of representing words of P_2 in their canonical form, and their composition. Also, matrix multiplication is the standard method of implementing parallel computation, and as most the entries of matrices from this system are zero (we can simplify computation considerably by the result that $M_{ij} = M_{ji}^{-1}$, for all matrices in this system), the matrices are 'sparse arrays', for which fast parallel multiplication algorithms exist. Finally, although the ! operator is used, which seems to imply infinite copies of words, its characterisation as a fixed point operator allows a finite (and indeed, simple) description of its conjugation by words of P_2. In view of this, it seems that the GOI1 system can be though of as a very concrete specification of a low-level parallel computing system.

Chapter 8

Although the introduction to this Chapter was phrased in terms of categorical models of lambda calculi, it is really about the Geometry of Interaction 1 system; categorical models of polymorphism have been constructed (see, for example, [31]), and using the Geometry of Interaction system does not appear to be the right approach. A direct approach to this seems better.

The main feature that this chapter requires, and does not have, is a categorical interpretation of the (multiplicative) operations of Linear Logic. Even though the GOI1 operations are representable in terms of the canonical elements of the M-monoid structures on $I(\mathbb{N})$ (or variations of them), the reason for their use in this form is not clear (except perhaps for the axiom link, and the cut / cut-elimination process). If the GOI1 system is the 'right' representation (and some doubt is cast on this claim by the restrictions on, and non-standard treatment of some of the operations), we would expect canonical categorical interpretations of the logical operators in terms of the canonical morphisms of one-object compact closed categories. However, this may just be because we lack details of the interaction of the two distinct M-monoid structures on $I(\mathbb{N})$, and the M-monoid structures they induce on $\mathbb{F}$.

More concretely, as the interpretaion of the cut-elimination process is phrased in terms of the opposite trace, and the alternative composition on the monoid $\mathbb{F}$, it would probably have been better to write Chapter 6 in terms of right duals, rather than left ones.

Also, the 'infinite matrices with a finite number of non-zero elements' interpretation of GOI1 matrices may not be the correct one; however, it is difficult to see how an alternative interpretation would allow the unrestricted compositions required.

Chapter 9

In terms of interpreting the GOI3 system, this chapter must be considered to be only just started, as opposed to requiring extra work — the only concrete results on the GOI3 system are that it is modelled in terms of inverse semigroups, and the resolution formula is defined in terms of the categorical trace. The difficulty with the self-similarity approach (which is again only just started) is finding a natural choice of category in which the clause semigroups are the endomorphism monoids. What was hoped for, and does not seem apparent, is an M-monoid structure on the clause semigroups used, hopefully involving compact closure. However, work has not really started on analysing the representations of the logical rules of full linear logic in this system — possibly the structures we are looking for will be apparent then.

In terms of an algebraic approach, the representation of the clause semigroup in the symmetric inverse monoid of partial injective maps on a term language, and its action on substitution classes, uses a lot of semilattice theoretic constructions, and these have not been studied in great depth.

Also, in [22], the connection between the clause algebra and the computer language Prolog is commented on. However, this is not the standard representation of computation in Prolog (which is given in terms of the Kleene star on the monoid of relations on a term language [6]), so a comparison of the two would be useful, possibly in terms of the usual sequent-calculus interpretation of substitution, unification, and composition in $B(L)$.

Finally, when the logical operations have been analysed, it would seem reasonable to attempt to model GOI3 computations in Prolog, or a related language. However, this is a long way off at the moment.

Chapter 10

The natural next step from this chapter is to consider algebraic models of automata that rewrite the tape they are reading, and use the method of 'sticking together' two one-way automata, to construct categorical models of Turing machines, in terms of compact closed categories (however, see the comments made on this in the critique of Chapter 7 above). Work has already started on this, by requiring symmetry between the set of states, and the input alphabet in the definition of a (one-way) automaton. The correct algebraic model of this seems to be Rees matrix semigroups, and work is progressing on the best way to 'stick together' two automata of this form, by analogy with the compact closed category construction for standard two-way automata.

An alternative extension of this chapter would be to attempt an automata theoretic interpretation of the way the trace and compact closed category composition are used in GOI1. This would require automata-theoretic interpretations of such things as inverse transition monoids, global symmetry between left and right parts of a two-way moving automaton, and the contraction maps (not to mention fixed point homomorphisms and associativity and commutativity operators) — it is unclear at the moment whether this approach would shed any light on either the Geometry of Interaction, or on the theory of two-way automata. However, polycyclic monoids are used in modelling pushdown automata, [15] so perhaps this is reasonable after all. At the very least, it would appear that 2-way pushdown automaton would have analogues of global transition relations that look very much like (fragments of) the Geometry of Interaction 1 system.

Bibliography

[1] **M. Barr, C. Wells**, *Category Theory for Computing Science*, Prentice Hall International (1990)

[2] **J. Berstel, D. Perrin**, *Theory of Codes*, Academic Press inc., Harcourt Brace Jovanovich (1985)

[3] **J.-C. Birget**, Basic Techniques for Two-way Finite Automata, in : J.E. Pin (ed.), *Formal Properties of Finite Automata and Applications*, SLNCS, Springer-Verlag (1989) 56-64

[4] **J.-C. Birget**, Concatenation of inputs in a two-way automaton, *Theoretical Computer Science 63* (1989) 141-156

[5] **G. Birkhoff, S. MacLane**, *A survey of Modern Algebra*, Collier - Macmillan, London (1965)

[6] **W.D. Burnham, A.R. Hall**, *Prolog Programming and Applications*, Macmillan London (1985)

[7] **R. Brown**, *Elements of Modern Topology*, McGraw-Hill, London (1968)

[8] **J. Cuntz**, Simple C^*-algebras generated by isometries, *Comm. Math. Phys.* **57** *(1977) 173-185*

[9] **V. Danos, L. Regnier**, Local and Asynchronous Beta Reduction, (an analysis of Girard's execution formula), *Proc. Eighth Ann. IEEE Symposium on Logic in Computer Science*, IEEE Comp. Soc. Press (1993) 296-306

[10] **V. Danos, L. Regnier**, Proof Nets and the Hilbert Space, *Advances in Linear Logic, LMS Lecture notes 222*, Cambridge University Press

[11] **M. Dunnet**, with **R. Minio**, *Elements of Intuitionism*, Clarendon Press, Oxford (1977)

[12] **C. Faith**, *Algebra I Rings, Modules, and Categories*, Springer-Verlag, Grundleher der mathematischen Wissenschaften 190 (1981)

[13] **J.H. Gallier**, Logic for Computer Science, *John Wiley and Sons inc.* (1987)

[14] **G. Gentzen**, *Collected works of Gerhard Gentzen*, (Ed. M. E. Szabo) North-Holland (1969)

[15] **R. Gilman**, Formal Languages and Infinite Groups, *DIMACS series in Discrete mathematics and Theoretical Computer Science Volume 00 (1991)*

[16] **J-Y. Girard**, Une extension de l'interprétation de Goedel á l'analyse, et son application 'a l'élimination des coupres dans l'analyse et la théorie des types, *J.E. Fenstad (ed), Proc. Scandinavian Logic Symposium, North-Holland (1971)*

[17] **J-Y. Girard**, Linear Logic, *Theor. Comp. Sci. 50, (1987) 1-102*

[18] **J-Y. Girard**, Multiplicatives, *Rendiconti del Seminario Matematico dell'Università e Politecnico di Torino, Special Issue on Logic and Computer Science, (1988)*

[19] **J-Y. Girard**, Towards a Geometry of Interaction, *Contemporary Mathematics (92), (1989) 69-108*

[20] **J-Y. Girard**, Geometry of interaction 1, *Proceedings Logic Colloquium '88, (1989) 221-260, North Holland.*

[21] **J-Y. Girard**, Geometry of interaction 2, *Proceedings of COLOG 88, Martin-Lof & Mints, 76-93 SLNCS 417*

[22] **J-Y. Girard**, Geometry of interaction 3, *Proceedings of the Workshop on Linear Logic 1994, MIT Press, To appear.*

[23] **J-Y. Girard, Y. Lafont, P. Taylor**, *Proofs and Types*, Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, (1989)

[24] **K. R. Goodearl**, *Notes on Real and Complex C^*-algebras*, Shiva Mathematical series

[25] **J.A. Green**, On the structure of semigroups, *Annals of mathematics* **54**, (1951) 163-172.

[26] **J. B. Hickey**, On variants of a semigroup, *Bull. Austral. Math. Soc.* **34** *(1986), 447-459*

[27] **P. Hines, M.V. Lawson**, An application of polycyclic monoids to rings, *Semigroup Forum*, to appear.

[28] **J. Hopcroft, J. Ullman**, *Introduction to Automata Theory, Languages and Computation,* Addison-Wesley (1979)

[29] **J. Howie**, *Automata and Languages*, Oxford Science Publications, Oxford University Press, (1991)

[30] **J. Howie**, *Fundamentals of Semigroup Theory*, Clarendon Press, Oxford, (1995)

[31] **M. Hyland**, A small complete category, *Annals of Pure and Applied Logic 40* (1989)

[32] **N. Jacobson**, Some remarks on one-sided inverses, *Proc. Amer. Math. Soc. 1 (1950), 352-355*

[33] **A. Joyal, R. Street**, The geometry of tensor calculus I, *Advances in Math. 102 (1993) 20-78*

[34] **A. Joyal, R. Street**, The Geometry of Tensor calculus II, *to appear.*

[35] **A. Joyal, R. Street, D. Verity**, Traced Monoidal categories, *Math. Proc. Camb. Phil. Soc., (1996) 425-446.*

[36] **L. Kauffman**, *On Knots*, Annals of Mathematical studies, Princeton University Press (1987)

[37] **G. Kelly, M. Laplaza**, Coherence for Compact Closed Categories, *Journal of Pure and Applied Algebra 19 (1980) 193-213*

[38] **J. Lambek, P.J. Scott**, *Introduction to Higher Order Categorical Logic*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, (1986)

[39] **S. Lang**, *Algebra*, 3^{rd} Edition, Addison-Wesley, (1993)

[40] **M.V.Lawson** Constructing inverse semigroups from category actions, *Journal of Pure and Applied Algebra, to appear.*

[41] **S. MacLane**, *Categories for the Working Mathematician*, Springer-Verlag New York, (1971)

[42] **M. Nivat**, Ann. Inst. Fourier, *Grenoble, 18, (1968) 339-456*

[43] **M. Nivat, J.-F. Perrot**, Une Generalisation du Monoide bicyclique, *C. R. Acad. Sc. Paris, t.271 (1970)*

[44] **M. Petrich**, *Inverse Semigroups*, Pure and Applied Mathematics, Wiley-Interscience Series, John Wiley and Sons. (1984)

[45] **J. C. Shepherdson**, The reduction of two-way automata to one-way automata, *IBM Journal of Research and Development (1959) 115-125*

[46] **Mei Chee Shum**, Tortile tensor categories, *Journal of Pure and Applied Algebra 93 (1994) 57-110*

[47] **G.F. Simmons**, *Introduction to Topology and Modern Analysis*, McGraw-Hill (1963)

[48] **N. E. Wegge-Olsen**, *K-Theory and C^*-algebras*, Oxford University Press, (1993)