PUBLICACIONES MATEMATICAS DEL URUGUAY

VOLUMEN 10

VOLUMEN 10

PUBLICACIONES MATEMÁTICAS DEL URUGUAY
Editadas por el Centro de Matemática de la Facultad de Ciencias, Universidad de la República y el Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" Facultad de Ingeniería, Universidad de la República; con el apoyo del Programa de Desarrollo de las Ciencias Básicas.

Montevideo, Abril de 2005.

ISSN
0797-1443

PUBLICACIONES MATEMÁTICAS DEL URUGUAY

Publicaciones Matemáticas del Uruguay(PMU) fue fundada in 1987 y publicada inicialmente por el Centro de Matemática de la Universidad de la República y auspiciada por el Programa de desarrollo de las ciencias básicas (PEDECIBA). Actualmente se publica en conjunto por el Centro de Matemática de la Facultad de Ciencias y el Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" de la Facultad de Ingeniería –ambas de la Universidad de la República– y cuenta con el auspicio del PEDECIBA.

Su objetivo es el de publicar artículos de investigación en todas las áreas de la matemática. Serán considerados para su publicación, artículos originales, adelantos de resultados de investigación, así como resúmenes panorámicos en áreas importantes de la matemática. Todos los artículos serán cuidadosamente arbitrados.

Las Publicaciones Matemáticas del Uruguay aparecerán con una frecuencia semestral.

Presentación de manuscritos

El Consejo Editor de las PMU recomienda que los trabajos presentados lo sean en formato LaTeX; sin embargo trabajos presentados en cualquiera de los otros formatos TeX standardt serán también considerados. Los archivos electrónicos deben ser enviados a la siguiente dirección: pmu@cmat.edu.uy.

Si el autor prefiere enviar una versión del manuscrito en papel, se requieren tres copias. Si el artículo es aceptado, el autor deberá enviar la versión electrónica con las características mencionadas.

La dirección de recepción es la siguiente: Publicaciones Matemáticas del Uruguay, Centro de Matemática, Facultad de Ciencias, IGUA 4225, CEP 11400, Montevideo, Uruguay. Al enviar un artículo el autor está asegurando que los derechos autorales no han sido cedidos y que el artículo no ha sido publicado o enviado a publicación en otra revista. Las Publicaciones Matemáticas del Uruguay –ISSN 0797 1443– están indexadas y resumidas en Mathematical Reviews Database y Zentralblatt MATH.

Intercambios bibliográficos

Propuestas de intercambio de suscripciones de las PMU por cualquier tipo de material bibliográfico son bienvenidas y deben ser enviadas a las direcciones mencionadas.

Publicaciones Matemáticas del Uruguay was founded in 1987 and was intially published by "Centro de Matemática de la Universidad de la República" and sponsored by " Programa de desarrollo de las ciencias básicas" (PEDECIBA). Currently is published jointly by "Centro de Matemática de la Facultad de Ciencias de la Universidad de la República" and "Instituto de Matemática y Estadística "Prof. Ing. Rafael Laguardia" de la Facultad de Ingeniería" –both from "Universidad de la República"– with the sponsorship of PEDECIBA.

It endeavors to publish research articles in all areas of mathematics. Survey articles of significant areas of mathematics as well as research announcements will also be considered for publication. All articles will be carefully reviewed.

It will appear with a frequency of two issues a year.

Submission of manuscripts

Publicaciones Matemáticas del Uruguay recommends for reasons of expediency the submission of papers in LaTeX; however papers submitted in any of the other standardt TeX formats will also be considered. Electronic files should be sent to the following address: `pmu@cmat.edu.uy`.

In the case an author prefers to send hard copies, three are required and if the submitted manuscript is eventually accepted, the author will be asked to send an electronic version as described above.

The receiving address is the following: Publicaciones Matemáticas del Uruguay, Centro de Matemática, Facultad de Ciencias, IGUA 4225, CEP 11400, Montevideo, Uruguay.

The submission of a paper implies the author's assurance that it has not been copyrighted, published or submitted for publication elsewhere.

Publicaciones Matemáticas del Uruguay –ISSN 0797 1443– is indexed/abstracted in Mathematical Reviews Database and Zentralblatt MATH.

Library exchanges

Proposals for exchange of library materials in return for subscriptions to the Publicaciones Matemáticas del Uruguay are most welcome. The proposals should be sent to the above address.

Índice

Prefacio

Los cuatro primeros artículos que aparecen en este volumen son contribuciones realizadas por participantes del XI Encuentro Rioplatense de Álgebra y Geometría Algebraica realizado en Solís, Uruguay en Diciembre de 2003. Dicho encuentro fué realizado en homenaje al Prof. Dr. Alfredo Jones en ocasión de la jubilación en su cargo de Prof. del Centro de Matemática de la Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay.
Para la edición de este volumen el Consejo Editor contó con la colaboración de Héctor Merklen y Andrea Solotar.

Foreword

The first four articles of this volume are contributions by participants of the XI Encuentro Rioplatense de Álgebra y Geometría Algebraica that took place in Solís, Uruguay in December 2003. This meeting was held in honor of Prof. Alfredo Jones in ocassion of his retirement of the Centro de Matemática, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay.
For the preparation of this volume Héctor Merklen and Andrea Solotar worked together with the Editorial Board.

Partial actions of discrete groups and related structures

Fernando Abadie

ABSTRACT

We give constructions of the groupoid of a partial action of a discrete group other than the ones given in [1]. We also show that Paterson's universal groupoid of the inverse semigroup of a discrete group agrees with the groupoid associated to the "universal" partial action of the group.

RESUMEN

Se dan construcciones del grupoide asociado a una acción parcial de un grupo discreto diferentes de las presentadas en [1]. También se muestra que el grupoide universal de Paterson del semigrupo inverso de un grupo discreto coincide con el grupoide asociado a la acción parcial "universal" del grupo.

To Alfredo Jones

Introduction

We have shown in [1] how to associate a locally compact groupoid $\mathcal{G}_\theta$ with every partial action θ on a locally compact Hausdorff space. We also observed that in the case of a topologically free partial action of a discrete group, the corresponding groupoid is nothing but a sheaf groupoid of germs. The purpose of the present paper is to compare this groupoid with some other groupoids which are obtained, in the special case of discrete groups, by combining results of Exel, Nica, Paterson and Sieben. The key step consists in passing from a partial action of G to an action of an associated $\tilde{F}$–inverse semigroup, and then to construct a groupoid via Nica's theory ([4]). We show below that the groupoids thus obtained coincide with ours. We also connect the work of Nica with that of Paterson on localizations. On the other hand, Paterson shows in [5] how to associate a "universal groupoid" with any inverse semigroup. We will see that the universal groupoid of $S(G)$ is the groupoid of a suitable partial action of G.

Suppose that G is a discrete group and $\theta = (\{X_t\}_{t\in G}, \{\theta_t\}_{t\in G})$ is a partial action of G on a set X. This means that each X_t is a subset of X, with $X_e = X$

(where e is the unity element of G), $\theta_t : X_{t^{-1}} \to X_t$ is bijective, $\forall t \in G$, and θ_{st} is an extension of $\theta_s\theta_t$, $\forall s,t \in G$. Thus $\theta_e = id_X$, and $\theta_t^{-1} = \theta_{t^{-1}}$, $\forall t \in G$. If X is a topological space it is also required that each X_t is open and each θ_t is a homeomorphism.

The partial action θ on the set X can be thought of as the set of morphisms of a groupoid $\mathcal{G}_\theta$ in the following way. The set of objects of $\mathcal{G}_\theta$ is the space X, and, given $x,y \in X$, the set of morphisms $\mathcal{G}_\theta(x,y)$ from x into y is defined as $\mathcal{G}_\theta(x,y) := \{(y,t,x) \in X \times G \times X : x \in X_{t^{-1}} \text{ and } \theta_t(x) = y\}$. Composition of morphisms is given by $(z,s,y)(y,t,x) = (z,st,x)$. The fact that θ is a partial action ensures that $\mathcal{G}_\theta$ is a groupoid (see [1] for details). Note that the identity morphism at x is (x,e,x), and $(y,t,x)^{-1} = (x,t^{-1},y)$. If X is a locally compact Hausdorff space, we endow $\mathcal{G}_\theta$ with the topology inherited by the product topology on $X \times G \times X$. It turns out that $\mathcal{G}_\theta$ is a locally compact Hausdorff groupoid with this topology.

1. $\mathcal{G}_\theta$ in two steps

We briefly review some basic facts of the theory developed in [2]. The inverse semigroup $S(G)$ of a group G is defined by a set of generators $\{[t] : t \in G\}$ and the relations:

(i) $[s^{-1}][s][t] = [s^{-1}][st], \forall s,t \in G.$ (iii) $[s][e] = [s], \forall s \in G$

(ii) $[s][t][t^{-1}] = [st][t^{-1}], \forall s,t \in G.$ (iv) $[e][s] = [s], \forall s \in G.$

The semigroup $S(G)$ is characterized by the following universal property: if S is a semigroup and $f : G \to S$ is a map satisfying

$$f(s^{-1})f(s)f(t) = f(s^{-1})f(st), \forall s,t \in G \tag{1}$$

$$f(s)f(t)f(t^{-1}) = f(st)f(t^{-1}), \forall s,t \in G \tag{2}$$

$$f(s)f(e) = f(s), \forall s \in G \tag{3}$$

then there is a unique homomorphism $\tilde{f} : S(G) \to S$ such that $\tilde{f}([t]) = f(t)$, $\forall t \in G$. In particular, there is a homomorphism $\partial : S(G) \to G$ such that $\partial([t]) = t$, $\forall t \in G$. A consequence of the universal property of $S(G)$ is that there exists an involution $* : S(G) \to S(G)$ such that $[t]^* = [t^{-1}]$, $\forall t \in G$, and $\big(S(G),*\big)$ is an inverse semigroup. If $t \in G$, the element $\varepsilon_t := [t][t^{-1}]$ belongs to the semilattice $\mathcal{E}_{S(G)}$ of the idempotents of $S(G)$. Moreover, every $\sigma \in S(G)$ has a *standard form*: there exist unique disjoint subsets $\{t\}$ and $\{t_1,\ldots,t_k\}$ of G such that $\sigma = \varepsilon_{t_1}\cdots\varepsilon_{t_k}[t]$.

Every inverse semigroup S has a partial order defined by: $\sigma \leq \tau \iff \sigma\sigma^* = \tau\sigma^*$. A unital inverse semigroup S is called an $\tilde{F}$–inverse semigroup if every $\sigma \in S$, $\sigma \neq 0$, is majorized by a unique maximal element of S. If such condition holds for each $\sigma \in S$, it is said that S is an F–inverse semigroup.

PROPOSITION 1.1 *Let $S(G)$ be the inverse semigroup of the group G. We have:*

1. *If $\sigma \in S(G)$ and $t \in G$, then $\partial(\sigma) = t \iff \sigma \leq [t]$*
2. *The set of maximal elements of $S(G)$ is $\mathcal{M} = \{[t] : t \in G\}$*
3. *$S(G)$ is an F–inverse semigroup.*
4. *The maximal element that majorizes $[s][t]$ is $[st]$.*

Proof. To prove 1. note that if $\partial(\sigma) = t$ and $\sigma = \varepsilon_{s_1} \cdots \varepsilon_{s_k}[s]$, then $s = t$. Therefore $\sigma\sigma^* = \varepsilon_{s_1} \cdots \varepsilon_{s_k}\varepsilon_t = [t]\sigma^*$, so $\sigma \leq [t]$. Conversely, since homomorphisms of inverse semigroups are order preserving, we have $\partial(\sigma) \leq t$ whenever $\sigma \leq [t]$, and therefore $\partial(\sigma) = t$. Now 2., 3. and 4. follow directly from 1. □

With every pair (S, α), where S is an $\tilde{F}$–inverse semigroup S and α a left action of S on a space X, Nica associates a groupoid by using a procedure comparable with the construction of the sheaf groupoid of germs (see [1, Remark 2.1]). Let $\mathcal{M}_S$ be the set of maximal elements of S. There is a partially defined product on $\mathcal{M}_S$: if $\mu, \mu' \in \mathcal{M}_S$ are such that their product $\mu\mu'$ in S is not zero, then $\mu \cdot \mu'$ is defined to be the unique element in $\mathcal{M}_S$ that majorizes $\mu\mu'$. The groupoid of (S, α) is $\mathcal{N}_\alpha := \{(\mu, x) : \mu \in \mathcal{M}_S, x \in \mathrm{dom}(\alpha_\mu)\}$ with product $(\mu, x)(\mu', x') = (\mu \cdot \mu', x')$ whenever $\alpha_\mu(x') = x$ and $\mu\mu' \neq 0$, inversion $(\mu, x)^{-1} = (\mu^*, \alpha_\mu(x))$, and with the product topology. The reader is referred to [4] for details.

Consider now a partial action of a discrete group G on a locally compact Hausdorff space X. By [2, Theorem 4.2] θ induces a unique action (also called θ) of $S(G)$ on X, such that $\theta_t = \theta_{[t]}$, $\forall t \in G$. Following [4] we may associate with $\big(S(G), \theta\big)$ a groupoid $\mathcal{N}_\theta$. In view of parts 2. and 4. of Proposition 1.1 and the comments above we have $\mathcal{N}_\theta = \{([t], x) : t \in G, x \in X_{t^{-1}}\}$, with the product $([s], y)([t], x) = ([st], x)$, for $([s], y), ([t], x) \in \mathcal{N}_\theta$ such that $\theta_{[t]}(x) = y$.

PROPOSITION 1.2 *The groupoids $\mathcal{G}_\theta$ and $\mathcal{N}_\theta$ are naturally isomorphic.*

Proof. It is clear that the map $\psi : \mathcal{G}_\theta \to \mathcal{N}_\theta$ given by $\phi(y, t, x) = ([t], x)$ is a homeomorphism with inverse $([t], x) \mapsto (\theta_t(x), t, x)$. Moreover:

$$\psi\big((z, s, y)(y, t, x)\big) = \psi(z, st, x) = ([st], x) = ([s], y)([t], x) = \psi(z, s, y)\psi(y, t, x).$$

Therefore ψ is an isomorphism of locally compact groupoids. □

A partial action of the discrete group G on the C^*-algebra A is a partial action $\sigma = (\{D_t\}, \{\sigma_t\})$ on the set A with the requirement that every D_t is an ideal in A and every σ_t is an isomorphism of C^*-algebras. If A is commutative, then every partial action σ on A is the dual of a unique partial action θ on the spectrum X of A, that is: $D_t = C_0(X_t)$ and $\sigma_t(a) = a \circ \theta_{t^{-1}}$, $\forall t \in G$ and $a \in D_{t^{-1}}$, where X_t is an open subset of X (see [1] for details).

A covariant representation of the system (A, G, σ) on the Hilbert space H is a pair (π, u), where $\pi : A \to B(H)$ is a non-degenerate representation of A and $u : G \to B(H)$ is such that every u_t is a partial isometry on H with initial subspace $\pi(D_{t^{-1}})H$ and final space $\pi(D_t)H$, such that $u_t\pi(a)u_{t^{-1}} = \pi(\sigma_t(a))$, $\forall a \in D_{t^{-1}}$, $t \in G$, and $u_{st}h = u_s u_t h$, $\forall h \in \pi(D_{t^{-1}} \cap D_{t^{-1}s^{-1}})H$. Covariant representations of (A, G, σ) are in a bijective correspondence with non-degenerate representations of the crossed product $A \ltimes_\sigma G$. In one direction, the covariant representation (π, u) on H gives rise to a representation $\pi \times u : A \ltimes_\sigma G \to B(H)$ which is determined by $(\pi \times u)(a_t\delta_t) = \pi(a_t)u_t$, where $a_t\delta_t$ is such that $a_t \in D_t$, and $a_t\delta_t(r) = a_t$ if $r = t$, 0 otherwise (recall that the linear span of such elements is dense in $A \ltimes_\sigma G$).

In [6] Sieben associates an inverse semigroup $S_{\pi,u}$ with every covariant representation (π, u) of the system $\big(A, G, \sigma\big)$. Such inverse semigroup is given by $S_{\pi,u} = \{(\sigma_{t_1} \dots \sigma_{t_n}, u_{t_1} \dots u_{t_n}) : n \in \mathbb{N}, t_j \in G, \forall j = 1, \dots, n\}$ with the obvious operations.

THEOREM 1.3 *Let σ be a partial action of the discrete group G on the C^*-algebra A, and suppose that (π, u) is a covariant representation of the partial dynamical system $\big(A, G, \sigma\big)$ on the Hilbert space H. Then*

1. *The map $G \to S_{\pi,u}$ given by $t \mapsto (\sigma_t, u_t)$ extends uniquely to a homomorphism $S(G) \to S_{\pi,u}$, which moreover is surjective.*

2. *If the representation $\pi \times u : A \ltimes_\sigma G \to B(H)$ is faithful, then $S_{(\pi,u)}$ is an $\tilde{F}$–inverse semigroup, whose maximal set is $\mathcal{M}_{\pi,u} = \{(\sigma_t, u_t) : t \in G, \sigma_t \neq 0\}$.*

Proof. Since σ is a partial action and u is a partial representation of G (because (π, u) is a covariant representation), identities (1), (2) and (3) are satisfied by the map $t \mapsto (\sigma_t, u_t)$, so the existence and uniqueness of the claimed homomorphism is guaranteed by the universal property of $S(G)$. The surjectivity of such a homomorphism follows from the fact that $\{(\sigma_t, u_t) : t \in G\}$ generates $S_{\pi,u}$.

To prove the second assertion note first that $\forall t_1,\ldots,t_n \in G$ we have that $(\sigma_{t_1}\ldots\sigma_{t_n}, u_{t_1}\ldots u_{t_n}) \leq (\sigma_{t_1\cdots t_n}, u_{t_1\cdots t_n})$. We will show that if $0 \neq (\sigma_s, u_s) \leq (\sigma_t, u_t)$, then $s = t$. Now, $(\sigma_s, u_s) \leq (\sigma_t, u_t) \iff (id_{D_s}, u_s u_s^*) = (\sigma_t\sigma_{s^{-1}}, u_t u_s^*)$. The equality of the first coordinates implies that $D_{s^{-1}} \subseteq D_{t^{-1}}$, $D_s \subseteq D_t$ and $\sigma_t(a) = \sigma_s(a)$, $\forall a \in D_{s^{-1}}$. The equality of the second coordinates implies that the partial isometry u_t agrees with u_s on the initial space of the latter. Thus u_t^* extends the partial isometry u_s^*, so that $u_t^* u_s = u_s^* u_s$, and therefore $u_s^* u_t = u_s^* u_s$. Identify A with the image of its universal representation, and consider the normal extension of π (which will be still denoted by π) to the enveloping von Neumann algebra A'' of A. As shown in [6], if p_r is the unit element of the strong closure of D_r, then $\pi(p_r) = u_r u_r^*$, $\forall r \in G$. Since $D_s \subseteq D_t$, we have that $a\delta_t \in A \rtimes_\sigma G$, for every $a \in D_s$. Therefore for $a \in D_s$ we have:

$$(\pi \times u)(a\delta_t) = \pi(a)u_t = \pi(ap_s)u_t = \pi(a)u_s u_s^* u_t = \pi(a)u_s = (\pi \times u)(a\delta_s).$$

Since $\pi \times u$ is faithful and $D_s \neq 0$, we must have that $s = t$. Therefore $S_{\pi,u}$ is an $\tilde{F}$–inverse semigroup. □

COROLLARY 1.4 *Let X be a locally compact Hausdorff space, $A = C_0(X)$, and σ the dual of the partial action θ on X. Suppose that (π, u) is a covariant representation of (A, G, σ) such that $\pi \times u$ is faithful. Then the map $(\sigma_{t_1}\ldots\sigma_{t_n}, u_{t_1}\ldots u_{t_n}) \mapsto \theta_{t_1}\ldots\theta_{t_n}$ defines a left action of $S_{\pi,u}$ on X, and the corresponding Nica groupoid is isomorphic to $\mathcal{G}_\theta$.*

Proof. By 1.3 $S_{\pi,u}$ is an $\tilde{F}$–inverse semigroup, whose maximal set is $\mathcal{M}_{\pi,u} = \{(\sigma_t, u_t) : t \in G, \sigma_t \neq 0\}$. Since θ is a partial action, it is clear that the given map defines a left action of $S_{\pi,u}$ on X. By definition the corresponding Nica groupoid is $\mathcal{N}_{\pi,u} = \{(m, x) : m \in \mathcal{M}_{\pi,u}, x \in \mathrm{dom}(m)\}$. Therefore $\mathcal{N}_{\pi,u} = \{(\sigma_t, u_t, x) : \sigma_t \neq 0 \text{ and } x \in X_{t^{-1}}\}$. Moreover, it is routine to check that the map $\mathcal{G}_\theta \to \mathcal{N}_{\pi,u}$ given by $(y, t, x) \mapsto (\sigma_t, u_t, x)$ is an isomorphism of locally compact groupoids. □

2. The universal groupoid of $S(G)$

In [5] Paterson associates a "universal groupoid" $\Gamma_{\mathbf{u}}$ with every inverse semigroup S. The crucial property of $\Gamma_{\mathbf{u}}$ is that it determines all S–groupoids. We refer the reader to [5] for complete information. For our purposes it will be enough to recall a concrete form of $\Gamma_{\mathbf{u}}$. Suppose that S is an inverse semigroup, and let $\mathcal{E}_S$ be its semilattice of idempotents. A semicharacter of $\mathcal{E}_S$ is a non–zero homomorphism $x : \mathcal{E} \to \{0,1\}$. Denote by X_S the space of non–zero semicharacters of $\mathcal{E}_S$ endowed with the product topology. For $\sigma \in S$ consider $D_\sigma := \{x \in X_S : x(\sigma\sigma^*) = 1\}$ and $R_\sigma := \{x \in X_S : x(\sigma^*\sigma) = 1\}$. There is a right action of S on $\mathcal{E}_S$ given by: $x \mapsto x \cdot \sigma$, $\forall x \in D_\sigma$, where $x \cdot \sigma(\varepsilon) := x(\sigma\varepsilon\sigma^*)$, $\forall \varepsilon \in \mathcal{E}_S$. This map is a partial homeomorphism in X_S, with domain D_σ and range R_σ.

If $\Sigma_S := \{(x,\sigma) : \sigma \in S, x \in D_\sigma\}$, there is an equivalence relation on Σ_S: $(x,\sigma) \sim (y,\tau) \iff x = y$ and there exists $\varepsilon \in \mathcal{E}$ such that $x(\varepsilon) = 1$ and $\varepsilon\sigma = \varepsilon\tau$. Denote by $\overline{(x,\sigma)}$ the class of (x,σ).

The underlying set of the universal groupoid of S is $\Gamma_{\mathbf{u}} = \{\overline{(x,\sigma)} : \sigma \in S, x \in D_\sigma\}$. The product of two elements $\overline{(x,\sigma)}, \overline{(y,\tau)} \in \Gamma_{\mathbf{u}}$ is defined only if $y = x \cdot \sigma$, and in this case we have $\overline{(x,\sigma)}\,\overline{(y,\tau)} = \overline{(x,\sigma\tau)}$. A basis for the topology of $\Gamma_{\mathbf{u}}$ is $\mathcal{U} := \{D(U,\sigma) : U \subseteq X_S \text{ is open}, \sigma \in S\}$, where $D(U,\sigma) := \{\overline{(x,\sigma)} : x \in U\}$. The inverse of $\overline{(x,\sigma)}$ is $\overline{(x \cdot \sigma, \sigma^*)}$. See [5, Theorem 4.3.1] for details. Let us give a description of $\Gamma_{\mathbf{u}}$ for the inverse semigroup $S(G)$ in terms of a partial action of G.

LEMMA 2.1 *If $(x,\sigma),(y,\tau) \in \Sigma_{S(G)}$, then $(x,\sigma) \sim (y,\tau) \iff x = y$ and $\partial(\sigma) = \partial(\tau)$.*

Proof. It is clear that $(x,\sigma) \sim (y,\tau) \Rightarrow x = y$ and $\partial(\sigma) = \partial(\tau)$. Suppose now that $x = y$ and $\partial(\sigma) = \partial(\tau)$. Note that every $\sigma \in S(G)$ satisfies $\sigma = \sigma\sigma^*[\partial(\sigma)]$. Therefore, if we let $\varepsilon := \sigma\sigma^*\tau\tau^*$, then $x(\varepsilon) = 1$, and

$$\varepsilon\sigma = \tau\tau^*\sigma\sigma^*[\partial(\sigma)] = \tau\tau^*\sigma\sigma^*[\partial(\tau)] = \sigma\sigma^*\tau\tau^*[\partial(\tau)] = \varepsilon\tau.$$

□

In [2] and [3] several C^*-algebras are successfully described and studied in terms of invariant sets of a certain partial action. This partial action is defined as follows. Let $\Omega := \{\omega : G \to \{0,1\} / \omega(e) = 1\}$ with the product topology, where G is a discrete group. Thus Ω is a compact Hausdorff space, which we will identify with the set $\{\omega \subseteq G : e \in \omega\}$ in the obvious way. For $t \in G$ consider the compact open subset $\Omega_t := \{\omega : t \in \omega\}$ of Ω. We have a partial action $\rho = (\{\rho_t\}, \{\Omega_t\})$ of G on Ω given by $\rho_t : \Omega_{t^{-1}} \to \Omega_t$ such that $\rho_t(\omega) = t\omega$. We will show that the groupoid $\mathcal{G}_\rho$ is naturally isomorphic to the universal groupoid $\Gamma_{\mathbf{u}}$ of $S(G)$.

Suppose that $\omega \in \Omega$ and let $x^\omega : \mathcal{E}_{S(G)} \to \{0,1\}$ be given by $x^\omega(\varepsilon_{t_1} \cdots \varepsilon_{t_k}) = \chi_\omega(t_1) \cdots \chi_\omega(t_k)$, $\forall \varepsilon = \varepsilon_{t_1} \cdots \varepsilon_{t_k} \in \mathcal{E}_{S(G)}$, where χ_ω is the characteristic function of ω. Then x^ω is a semicharacter of $\mathcal{E}_{S(G)}$, and x^ω is non–zero because $x^\omega([e]) = 1$.

PROPOSITION 2.2 *The map: $h : \Omega \to X_{S(G)}$ given by $\omega \mapsto x^\omega$ is a homeomorphism, and $h(\Omega_t) = D_{[t]}$, $\forall t \in G$. Moreover, $h(\rho_{t^{-1}}\omega) = h(\omega) \cdot [t]$, $\forall t \in G, \omega \in \Omega_t$.*

Proof. For $x \in X_{S(G)}$ consider $\omega_x \subseteq G$ such that $\chi_{\omega_x}(t) = x(\varepsilon_t)$, $\forall t \in G$. Since $x([e]) = 1$, it follows that $\omega_x \in \Omega$. It is clear that $x \mapsto \omega_x$ is the inverse map of h. Note that $x \in D_{[t]} \iff x(\varepsilon_t) = 1 \iff t \in \omega_x \iff \omega_x \in \Omega_t$. Therefore $h(\Omega_t) = D_{[t]}$, $\forall t \in G$. Since both of Ω and $X_{S(G)}$ are considered with the product topologies, it is clear that h is a homeomorphism. Consider now $x \in D_{[t]}$, and let $s \in G$. By using successively the relations defining $S(G)$ we have $[t][s][s^{-1}][t^{-1}] = [ts][s^{-1}][t^{-1}][t][t^{-1}] = [ts][s^{-1}t^{-1}][t][t^{-1}] = \varepsilon_{ts}\varepsilon_t$. Thus

$$x \cdot [t](\varepsilon_s) = x([t][s][s^{-1}][t^{-1}]) = x(\varepsilon_{ts}\varepsilon_t) = x(\varepsilon_{ts})x(\varepsilon_t) = x(\varepsilon_{ts}).$$

Therefore, if $x \in D_{[t]}$, $s \in G$ we have $s \in \omega_{x\cdot[t]} \iff x(\varepsilon_{ts}) = 1 \iff \chi_{\omega^x}(ts) = 1 \iff ts \in \omega_x \iff s \in t^{-1}\omega_x$. This shows that $h(\rho_{t^{-1}}(\omega_x)) = h(\omega_x) \cdot [t]$. □

THEOREM 2.3 *Let $\Gamma_{\mathbf{u}}$ be the universal groupoid of the inverse semigroup $S(G)$. Then the map $\Phi : \mathcal{G}_\rho \to \Gamma_{\mathbf{u}}$ such that $\Phi(t\omega, t, \omega) = \overline{(h(t\omega), [t])}$ is an isomorphism of locally compact groupoids, whose inverse is Ψ, given by $\Psi\big(\overline{(x,\sigma)}\big) = \big(h^{-1}(x), \partial(\sigma), \partial(\sigma)^{-1}h^{-1}(x)\big)$*

Proof. Note that Ψ is well defined by Lemma 2.1. Let $(t\omega, t, \omega) \in \mathcal{G}_\rho$. Then:

$$\Psi\Phi(t\omega,t,\omega) = \Psi\big(\overline{(h(t\omega),[t])}\big) = \big(h^{-1}\big(h(t\omega)\big), t, t^{-1}h^{-1}\big(h(t\omega)\big)\big) = (t\omega, t, \omega)$$

Now, if $\overline{(x,\sigma)} \in \Gamma_{\mathbf{u}}$ by Lemma 2.1 we have that $\overline{(x,\sigma)} = \overline{(x, [\partial(\sigma)])}$. Therefore:

$$\Phi\Psi\big(\overline{(x,\sigma)}\big) = \Phi\big(h^{-1}(x), \partial(\sigma), \partial(\sigma)^{-1}h^{-1}(x)\big) = \overline{\big(h\big(h^{-1}(x)\big), [\partial(\sigma)]\big)} = \overline{(x,\sigma)}$$

Hence Φ is a bijection with inverse Ψ. Now, consider elements $(st\omega, s, t\omega)$ and $(t\omega, t, \omega) \in \mathcal{G}_\rho$. Since $(st\omega, s, t\omega)(t\omega, t, \omega) = (st\omega, st, \omega)$, we have that $\Phi\big((st\omega, s, t\omega)(t\omega, t, \omega)\big) = \Phi(st\omega, st, \omega) = \overline{\big(h(st\omega), [st]\big)}$. On the other hand:

$$\begin{aligned}\Phi(st\omega, s, t\omega)\Phi(t\omega, t, \omega) &= \overline{\big(h(st\omega), [s]\big)}\ \overline{\big(h(t\omega), [t]\big)} \\ &= \overline{\big(h(st\omega), [s][t]\big)} \\ &= \overline{\big(h(st\omega), [st]\big)}.\end{aligned}$$

It follows that Φ is an isomorphism of groupoids. We show next that it is also a homeomorphism. Consider, for $t, s \in G$, $a \in \{0,1\}$ the sets:

$$V_{t,s,a} := \{(t\omega, t, \omega) \in \mathcal{G}_\rho : \chi_\omega(s) = 1\} \subseteq \mathcal{G}_\rho, \qquad U_{t,s,a} := D(U_{s,a}, [t]) \subseteq \Gamma_{\mathbf{u}},$$

where $U_{s,a} := \{x \in X_{S(G)} : x(\varepsilon_s) = 1\}$. The family $\mathcal{V} := \{V_{t,s,a} : t, s \in G,\ a \in \{0,1\}\}$ is a subbasis for the topology of $\mathcal{G}_\rho$, and the family $\mathcal{U} := \{U_{t,s,a} : t, s \in G,\ a \in \{0,1\}\}$ is a subbasis for the topology of $\Gamma_{\mathbf{u}}$. Thus to see that Φ is a homeomorphism it suffices to show that $\mathcal{U} = \{\Phi(V) : V \in \mathcal{V}\}$. Now: $\Phi(V_{t,s,a}) = \{\overline{(h(t\omega), [t])} : h(\omega)(\varepsilon_s) = a)\} = \{\overline{(h(\omega) \cdot [t]^*, [t])} : (h(\omega) \cdot [t]^*) \cdot [t](\varepsilon_s) = a)\} = \{\overline{(x, [t])} : x(\varepsilon_{ts}) = a\} = \{\overline{(x, [t])} : x \in U_{ts,a}\} = U_{t,ts,a}$. Therefore $\Phi(V) \in \mathcal{U}$, $\forall V \in \mathcal{V}$. Since $U_{t,s,a} = \Phi(V_{t,t^{-1}s,a})$, we conclude that $\mathcal{U} = \{\Phi(V) : V \in \mathcal{V}\}$. □

3. $\tilde{F}$–inverse semigroups and localizations

We end the paper by showing sketchily how Nica's theory connects with the groupoids of localizations defined by Paterson ([5, page 127]). We use the notation of Section 1. Recall that a localization is a right action β of an inverse semigroup S on a space X such that $\{\mathrm{dom}(\beta_\sigma) : \sigma \in S\}$ is a basis for the topology of X. Given a localization β, Paterson considers the set $\Theta := \{(x,\sigma) \in X \times S : x \in \mathrm{dom}(\beta_\sigma)\}$ with the equivalence relation $(x,\sigma) \sim (y,\tau) \iff x = y$ and there exists $\varepsilon \in \mathcal{E}_S$ such that $x \in \mathrm{dom}(\beta_\varepsilon)$ and $\varepsilon\sigma = \varepsilon\tau$. Then he defines the locally compact groupoid $\Gamma(X,S) := \{\overline{(x,\sigma)} : \sigma \in S, x \in \mathrm{dom}(\beta_\sigma)\}$, where the product is given by $\overline{(x,\sigma)}\,\overline{(x\cdot\sigma,\tau)} := \overline{(x,\sigma\tau)}$ ($x \cdot \sigma$ denotes $\beta_\sigma(x)$), the inversion by $\overline{(x,\sigma)}^{-1} := \overline{(x\cdot\sigma,\sigma^*)}$, and a basis for the topology is given by $\{D(U,\sigma) : \sigma \in S, U$ open subset of $\mathrm{dom}(\beta_\sigma)\}$, where $D(U,\sigma) := \{\overline{(x,\sigma)} : x \in U\}$.

Suppose now that α is a (left) action of an $\tilde{F}$–inverse semigroup S on X, and consider the inverse semigroup

$$L_\alpha := \{u \in \mathrm{PHom}(X) : u \text{ is extended by some } \alpha_\sigma\},$$

where $\mathrm{PHom}(X)$ stands for the inverse semigroup of the partial homeomorphisms of X.

Proposition 3.1 *Let S be an $\tilde{F}$–inverse semigroup and $S_\alpha := \{(\sigma,u) \in S \times L_\alpha : u \leq \alpha_\sigma\}$. Then S_α is an $\tilde{F}$–inverse semigroup, and $\mathcal{M}_{S_\alpha} = \{(\mu,\alpha_\mu) : \mu \in \mathcal{M}_S\}$.*

Proof. It is clear that S_α is an inverse semigroup. Note that $(\sigma, u) \leq (\tau, v) \iff \sigma \leq \tau$ and $u \leq v$. Thus if σ is non–zero and $\sigma \leq \mu$, with $\mu \in \mathcal{M}_S$, and if $(\sigma, u) \leq (\tau, v)$, then it must be $\tau \leq \mu$, and hence $(\sigma, u) \leq (\tau, v) \leq (\mu, \alpha_\mu)$, which is obviously a maximal element. This ends the proof. □ The left action

α induces a right action β of S_α on X via $\beta_{(\sigma,u)} := u^*$ ($\beta_{(\sigma,u)}$ is the restriction of α_{σ^*} to the range of u). Then β is a localization of S on X in the sense of Paterson , and therefore it has associated a groupoid Γ_α. By Proposition 3.1 we can identify a class $\overline{(x, (\sigma, u))}$ with a pair (x, μ), where $\sigma \leq \mu$, $\mu \in \mathcal{M}_S$, and $x \in \mathrm{dom}(\alpha_\mu)$ and then we have $\Gamma_\alpha = \{(x, \mu) : \mu \in \mathcal{M}_S, x \in \mathrm{dom}(\beta_\mu)\}$, $(x, \mu)(x \cdot \mu, \mu') = (x, \mu \cdot \mu')$, $(x, \mu)^{-1} = (x \cdot \mu, \mu^*)$, and Γ_α has the product topology. It is easily checked that the map $\mathcal{N}_\alpha \to \Gamma_\alpha$ given by $(\mu, x) \mapsto (x \cdot \mu^*, \mu)$ is an isomorphism of locally compact groupoids. Thus we have:

PROPOSITION 3.2 *The Nica groupoid of the action α is naturally isomorphic to the Paterson groupoid associated with the localization induced by α.*

References

[1] F. Abadie, *On partial actions and groupoids*, Proc. Amer. Math. Soc. **132** (2003), 1037-1047.

[2] R. Exel, *Partial actions of groups and actions of inverse semigroups*, Proc. Amer. Math. Soc. **126** (1998), no. 12, 3481–3494.

[3] R. Exel, M. Laca, J. Quigg, *Partial dynamical systems and C^*-algebras generated by partial isometries*, J. Operator Theory **47** (2002), no. 1, 169–186.

[4] A. Nica, *On a groupoid construction for actions of certain inverse semigroups*, Internat. J. Math. **5** (1994), 349-372.

[5] A. L. T. Paterson, *Groupoids, Inverse Semigroups, and their Operator Algebras*, Progress in Mathematics **170**, Birkhaser, 1999.

[6] N. Sieben, *C^*–crossed products by partial actions of inverse semigroups*, J. Austral. Math. Soc. Ser. A **63** (1997), 32-46.

FERNANDO ABADIE
fabadie@cmat.edu.uy
Centro de Matemática. Facultad de Ciencias
Universidad de la República
Iguá 4225, 11400. Montevideo, Uruguay

Double Categories and Quantum groupoids

Nicolás Andruskiewitsch and Sonia Natale[1]

ABSTRACT

We give the construction of a class of weak Hopf algebras (actually, face algebras as defined by Hayashi) associated to a matched pair of groupoids and certain cocycle data. This generalizes a now well-known construction for Hopf algebras, first studied by G. I. Kac in the sixties. Our approach is based on the notion of double groupoids, as introduced by Ehresmann.

RESUMEN

Presentamos la construcción de una familia de álgebras de Hopf débiles (de hecho álgebras de caras según la definición de Hayashi) asociadas a un sistema de grupoides apareados y ciertos datos de tipo cociclo. Esto generaliza una construcción actualmente bien conocida para álgebras de Hopf, estudiada inicialmente por G. I. Kac en la década del sesenta. Nuestro enfoque está basado en la noción de grupoide doble introducida por Ehresmann.

Dedicado a Alfredo Jones en su 75 aniversario

Introduction

An *exact factorization* of a group Σ is a pair of subgroups G, F such that the multiplication map induces a bijection $m : F \times G \to \Sigma$. Given an exact factorization of a group Σ, there are a right action $\triangleleft : G \times F \to G$ and a left action $\triangleright : G \times F \to F$ defined by $sx = (s \triangleright x)(s \triangleleft x)$, for all $s \in G$, $x \in F$. These actions satisfy the compatibility conditions

$$s \triangleright xy = (s \triangleright x)((s \triangleleft x) \triangleright y), \tag{0.1}$$

$$st \triangleleft x = (s \triangleleft (t \triangleright x))(t \triangleleft x), \tag{0.2}$$

for all $s, t \in G$, $x, y \in F$. It follows that $s \triangleright 1 = 1$ and $1 \triangleleft x = 1$, for all $s \in G$, $x \in F$. Such a data of groups and compatible actions is called a *matched pair* of groups. Conversely, given a matched pair of groups F, G, one can find a group Σ together with an exact factorization $\Sigma = FG$.

[1]Partially supported by ANPCyT, CONICET, Agencia Córdoba Ciencia, Secyt (UNC) and Fundación Antorchas

Let $\Bbbk$ be a field. In the early eighties, Takeuchi achieved a construction which, starting from a matched pair F, G of finite groups, gives a (in general not commutative and not cocommutative) Hopf algebra $H := \Bbbk^G \# \Bbbk F$ [T1]. This Hopf algebra fits into an exact sequence

$$1 \longrightarrow \Bbbk^G \xrightarrow{\iota} \Bbbk^G \# \Bbbk F \xrightarrow{\pi} \Bbbk F \longrightarrow 1;$$

$\Bbbk^G \# \Bbbk F$ is called a bismash product; it is semisimple and cosemisimple if the characteristic of $\Bbbk$ is relatively prime to the order of Σ. The same construction was also presented by Majid [Mj1]. A more general instance of this construction can be done by adjoining a certain cohomological data associated to the matched pair: namely a pair of 2-cocycles $\sigma : F \times F \to (\Bbbk^G)^\times$ and $\tau : G \times G \to (\Bbbk^F)^\times$ satisfying appropriate compatibility conditions. In this way, all Hopf algebras H which fit into an exact sequence $1 \longrightarrow \Bbbk^G \xrightarrow{\iota} H \xrightarrow{\pi} \Bbbk F \longrightarrow 1$ are obtained. The compatibility conditions have an elegant description in terms of the total complex associated to a double complex that combines the group cohomologies of G, F and Σ. It turns out that this more general construction, and the cohomology theory behind it, had already been discovered by G. I. Kac [K]. Explicit computations can be done with the help of the so-called Kac exact sequence *loc. cit.* The study of this cohomology theory has been later pursued by Masuoka; see the paper [M] and references therein for details on this topic.

Quantum groups appart, this construction gave rise to one of the first genuine examples of non-commutative non-cocommutative Hopf algebras. More recently, it was shown that the Hopf algebras $\mathbb{C}^G \# \mathbb{C}F$ are exactly those having a positive basis [LYZ1].

In our previous paper [AN], we discussed braided Hopf algebras R which fit into an exact sequence $1 \longrightarrow \Bbbk^G \xrightarrow{\iota} R \xrightarrow{\pi} \Bbbk F \longrightarrow 1$. The present paper was inspired by a comment of the referee of [AN], pointing out a pictorial description of the standard basis of $\Bbbk^G \# \Bbbk F$, which gives a more compact form to the constructions. It turns out that this pictorial description, also present in [Mj2, T2, DVVV], can be stated in the language of double categories. These have been introduced by Ehresmann [E]. A double category can be defined as a $\mathcal{C}$-structured category, where $\mathcal{C}$ is the category of small categories and functors; that is, as a category object in the category of small categories.

Roughly, a small double category consists of a set of 'boxes' $\mathcal{B} = \{A, B, \dots\}$, each box having colored edges (and vertices)

$$A = l \overset{t}{\underset{b}{\square}} r;$$

boxes can be 'horizontally' and 'vertically' composed, both compositions subject to a natural interchange law. The description given in the Appendix of [AN] fits exactly into this framework: here the categories of vertical and horizontal compositions correspond to the transformation groupoids attached to the actions $\triangleleft : G \times F \to G$ and $\triangleright : G \times F \to F$, respectively. In this example the vertical edges of boxes are colored by elements of F, the horizontal edges are colored by elements of G, and it has the particularity that every box is uniquely determined by a pair of adjacent edges. Moreover, in this case there is only one coloring for the 'vertices' of boxes.

It is then natural to ask: what are the double categories that give rise to Hopf algebras in this fashion? First, we shall consider *double groupoids*– double categories where both the horizontal and vertical compositions are invertible– to have antipodes. Now, because of the positive basis Theorem in [LYZ1], we know that the answer should be the double groupoids coming from matched pairs of groups as above. Still, we can ask: what are the double groupoids that give rise to *weak Hopf algebras*?

The notion of weak Hopf algebras or quantum groupoids was recently introduced in [BNSz, BSz] as a non-commutative version of groupoids. A relevant feature is that they give rise to tensor categories. A weak Hopf algebra has an algebra and a coalgebra structure; the comultiplication is multiplicative but it does not preserve the unit. A particular but very important class of weak Hopf algebras was introduced and studied previously by T. Hayashi, see [H] and references therein.

Given a finite double groupoid, we endow the vector space with basis the set of boxes with the groupoid algebra structure of the vertical groupoid, and with the groupoid coalgebra structure of the horizontal groupoid. We found a sufficient condition to get a quantum groupoid; this is condition (2) in Proposition 2.2. It turns out that double groupoids satisfying this condition are equivalent to the *vacant* double groupoids considered by Mackenzie [Ma]: every box be determined by any pair of adjacent edges. Also, vacant double groupoids are in bijective correspondence with matched pairs of groupoids [Ma].

Our main result is that every vacant double groupoid gives rise to a weak Hopf algebra in the way described above; this weak Hopf algebra is semisimple if $\Bbbk$ has characteristic 0. The corresponding construction is also done by adjoining compatible 2-cocycle data. These 2-cocycle data is part of a first quadrant double complex, as in the group case; there is as well a "Kac exact sequence for groupoids". We point out that these weak Hopf algebras have commutative source and target subalgebras, so they are face algebras in the sense of Hayashi [H]. Also, they are involutory. Our construction may alternatively be presented without double groupoids, using just exact factorizations of groupoids. We feel however that the language of double categories is not merely accidental; it also appears again in recent work of Kerler and Lyubashenko on topological quantum field theory [KL].

The paper is organized as follows. The first section is devoted to double categories and double groupoids. For the convenience of the reader not used to the language of double categories, we include many details and proofs. In the second section we discuss vacant double groupoids. We describe vacant double groupoids in group-theoretical sense in Theorem 2.16. The third section contains the construction of semisimple quantum groupoids and a presentation of the Kac exact sequence for groupoids.

The results of this paper were reported by the first author at the University of Clermont-Ferrand, June 2003; at this occasion, he was told by S. Baaj that a topological version of the Kac exact sequence for vacant double groupoids– with set of vertices of cardinal one– has been independiently studied by G. Skandalis, S. Vaes and himself; see [BSV]. The first author also reported this work at the XV Coloquio Latinoamericano de Álgebra, Cococyoc, México, July 2003; at this occasion, G. Böhm, R. Coquereaux and K. Szlachányi pointed out the ressemblance between double groupoids and Ocneanu cells. He also reported this construction at the "Colloque Algèbres de Hopf et invariants topologiques", Luminy, March 2004.

After release of the first version of this paper, positive quasitriangular R-matrices for the weak Hopf algebras introduced here were constructed in [AA], generalizing results from [LYZ2]. See also connections with the quiver-theoretical quantum Yang-Baxter equation and face models in [A]. Explicit examples of matchedpairs of groupoids and some calculations of extensions are presented in [AM].

Acknowledgements. We thank the referee of [AN] for the key remark in her/his report. The research of the second author was done during a postdoctoral stay at the Département de mathématiques et applications of the École Normale Supérieure, Paris; she thanks Marc Rosso for the kind hospitality. Part of the work of the first author was done during a visit to the University of Reims (February - May 2003); he is very grateful to J. Alev for his generous invitation.

1. Double categories and double groupoids

1.1. Definition of double categories.

Let $\mathcal{C}$ be a category with pullbacks. Recall that a category object in $\mathcal{C}$ (or a category internal to $\mathcal{C}$) is a collection $(A, O, s, t, \mathrm{id}, m)$, where A ("arrows") and O ("objects") are objects in $\mathcal{C}$; $s, e : A \to O$ ("source" and "end = target", respectively), $\mathrm{id} : O \to A$ ("identities") and $m : A\ {}_e\!\times_s A \to A$

("composition") are arrows in $\mathcal{C}$; subject to the usual associativity and identity axioms. Similarly, a groupoid object in $\mathcal{C}$ is a category object in $\mathcal{C}$ with all "arrows" invertible, which amounts to the existence of a map $\mathcal{S} : A \to A$ with suitable properties.

Notation. Along this paper, in the case where f, g are composable arrows in a category, their composition $m(f,g)$ will be indicated by juxtaposition: $m(f,g) = fg$ (and not gf).

Definition 1.1. A (small) *double category* $\mathcal{T}$ consists of the following data:

- Four non-empty sets: $\mathcal{B}$ (boxes), $\mathcal{H}$ (horizontal edges), $\mathcal{V}$ (vertical edges) and $\mathcal{P}$ (points);
- eight boundary functions: $t, b : \mathcal{B} \to \mathcal{H}$; $r, l : \mathcal{B} \to \mathcal{V}$; $r, l : \mathcal{H} \to \mathcal{P}$; $t, b : \mathcal{V} \to \mathcal{P}$;
- four identity functions: $\mathrm{id} : \mathcal{P} \to \mathcal{H}$; $\mathrm{id} : \mathcal{P} \to \mathcal{V}$; $\mathbf{id} : \mathcal{H} \to \mathcal{B}$; $\mathbf{id} : \mathcal{V} \to \mathcal{B}$;
- four composition functions, all denoted by m:

$\mathcal{B}\,{}_b{\times}_t\mathcal{B} \to \mathcal{B}$ (vertical composition), $\mathcal{B}\,{}_r{\times}_l\mathcal{B} \to \mathcal{B}$ (horizontal composition),

$$\mathcal{H}\,{}_r\times_l \mathcal{H} \to \mathcal{H}, \quad \mathcal{V}\,{}_b\times_t \mathcal{V} \to \mathcal{V};$$

such that the following axioms are satisfied.

Axiom 0. $(\mathcal{B}, \mathcal{H}, t, b, \mathbf{id}, m)$, $(\mathcal{B}, \mathcal{V}, l, r, \mathbf{id}, m)$, $(\mathcal{H}, \mathcal{P}, l, r, \mathrm{id}, m)$, $(\mathcal{V}, \mathcal{P}, t, b, \mathrm{id}, m)$ are categories.

Axiom 1. Four identities between possible functions from $\mathcal{B}$ to $\mathcal{P}$, namely

$$tr = rt, \quad tl = lt, \quad bl = lb, \quad br = rb.$$

This axiom allows to depict graphically $A \in \mathcal{B}$ as a box

$$A = l \underset{b}{\overset{t}{\square}} r$$

where $t(A) = t$, $b(A) = b$, $r(A) = r$, $l(A) = l$, and the four vertices of the square representing A are $tl(A)$, $tr(A)$, $bl(A)$, $br(A)$. Of course, t, b, r and l mean, respectively, 'top', 'bottom', 'right' and 'left'. Most of the remaining axioms will be stated in this pictorial representation.

Warning. A box $A \in \mathcal{B}$ is, in general, *not* determined by its four boundaries t, b, r, l.

We shall write $A|B$ if $r(A) = l(B)$ (so that A and B are horizontally composable), and $\dfrac{A}{B}$ if $b(A) = t(B)$ (so that A and B are vertically composable).

The notation AB (respectively, $\frac{A}{B}$) will indicate the horizontal (respectively, vertical) compositions, whenever A and B are composable in the appropriate sense.

Axiom 2. *Consistency of boundary compositions.* Let $A = l \underset{b}{\overset{t}{\square}} r$ and $B = s \underset{c}{\overset{u}{\square}} m$ in $\mathcal{B}$.

$$\text{(1.1) If} \quad A|B, \quad \text{then} \quad AB = l \underset{bc}{\overset{tu}{\square}} m,$$

$$\text{(1.2) If} \quad \frac{A}{B}, \quad \text{then} \quad \frac{A}{B} = ls \underset{c}{\overset{t}{\square}} rm.$$

The notation $\begin{array}{c|c} A & B \\ \hline C & D \end{array}$ means that all possible horizontal and vertical products are allowed; in view of Axiom 2, this implies that $\frac{AB}{CD}$, $\begin{array}{c|c} A & B \\ C & D \end{array}$.

Axiom 3. *Interchange law between horizontal and vertical compositions.* If $\begin{array}{c|c} A & B \\ \hline C & D \end{array}$, then

$$\frac{AB}{CD} := \left\{\frac{AB}{CD}\right\} = \left\{\frac{A}{C}\right\}\left\{\frac{B}{D}\right\}. \tag{1.3}$$

A consequence of this Axiom is that, given $r \times s$ boxes A_{ij} with horizontal and vertical compositions allowed as in the following arrangement

$$\begin{array}{c|c|c|c} A_{11} & A_{12} & \dots & A_{1s} \\ \hline A_{21} & A_{22} & \dots & A_{2s} \\ \hline \dots & \dots & \dots & \dots \\ \hline A_{r1} & A_{r2} & \dots & A_{rs} \end{array},$$

then the composition

$$\begin{array}{cccc} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{array}$$

is well defined and can be computed associating in any possible way.

Axiom 4. *Horizontal and vertical identities.* The identity functions $\mathbf{id} : \mathcal{H} \to \mathcal{B}$ (vertical identity), $\mathbf{id} : \mathcal{V} \to \mathcal{B}$ (horizontal identity) satisfy

$$\mathbf{id}(g) = \mathrm{id}\ l(g) \underset{g}{\overset{g}{\square}} \mathrm{id}\ r(g), \quad g \in \mathcal{H}; \qquad \mathbf{id}(x) = \ x \underset{\mathrm{id}\ b(x)}{\overset{\mathrm{id}\ t(x)}{\square}} x \ , \quad x \in \mathcal{V}.$$

Note that, in principle, there is an ambiguity when using the notation id P for an element $P \in \mathcal{P}$; however, this ambiguity disappears in the pictorial representation. When necessary, the notation $\mathrm{id}_{\mathcal{H}} : \mathcal{P} \to \mathcal{H}$ and $\mathrm{id}_{\mathcal{V}} : \mathcal{P} \to \mathcal{V}$ for the corresponding identity maps will be used.

Axiom 5. *Horizontal and vertical identities of the identities of the points.* If $P \in \mathcal{P}$, then $\mathbf{id}\,\mathrm{id}_{\mathcal{H}} P = \mathbf{id}\,\mathrm{id}_{\mathcal{V}} P$; this box will be denoted Θ_P.

Axiom 6. *Compatibility of the identity with composition of arrows.* If $g, h \in \mathcal{V}$, $x, y \in \mathcal{H}$ are composable, then $\begin{Bmatrix} \mathbf{id}\, g \\ \mathbf{id}\, h \end{Bmatrix} = \mathbf{id}\, gh$, $\{\mathbf{id}\, x\ \mathbf{id}\, y\} = \mathbf{id}\, xy$.

The following lemma is well-known, see for example [BS].

Lemma 1.2. *A double category is a category object in the category of small categories.*

Proof. Let $\mathcal{T}$ be a category object in the category of small categories. Thus $\mathcal{T} = (\mathcal{A}, \mathcal{O}, t, b, \mathrm{id}, m)$, where $\mathcal{A}$ and $\mathcal{O}$ are small categories, $t, b : \mathcal{A} \to \mathcal{O}$, $\mathrm{id} : \mathcal{O} \to \mathcal{A}$ and $m : \mathcal{A}\,{}_b\!\times_t \mathcal{A} \to \mathcal{A}$ are functors subject to associativity and identity axioms.

Write $\mathcal{A} = (\mathcal{B}, \mathcal{V}, l, r, \mathbf{id}, m)$ and $\mathcal{O} = (\mathcal{H}, \mathcal{P}, l, r, \mathrm{id}, m)$. The functors t, b, id and m correspond, respectively, to maps

$$t, b : \mathcal{B} \rightrightarrows \mathcal{H}, \quad \mathrm{id} : \mathcal{H} \to \mathcal{B}, \quad m : \mathcal{B}\,{}_b\!\times_t \mathcal{B} \to \mathcal{B},$$

and

$$t, b : \mathcal{V} \rightrightarrows \mathcal{P}, \quad \mathrm{id} : \mathcal{P} \to \mathcal{V}, \quad m : \mathcal{V}\,{}_b\!\times_t \mathcal{V} \to \mathcal{V}.$$

The associativity and identity constraints relating the functors t, b, id and m, correspond to the fact that $(\mathcal{B}, \mathcal{H}, t, b, \mathbf{id}, m)$ and $(\mathcal{V}, \mathcal{P}, t, b, \mathrm{id}, m)$ are categories. In what follows we shall see that the functoriality of t, b, id and m, corresponds to the axioms 1–6.

The functoriality of t and b amount to the identities

$$tl = lt, \quad tr = rt, \quad bl = lb, \quad br = rb; \tag{1.4}$$

$$tm = m(t \times t), \quad bm = m(b \times b); \tag{1.5}$$

$$t\,\mathrm{id} = \mathrm{id}\,t; \quad b\,\mathrm{id} = \mathrm{id}\,b. \tag{1.6}$$

One sees that (1.4) corresponds to Axiom 1, (1.5) corresponds to (1.1) in Axiom 2, and (1.6) corresponds to the right hand side identity in Axiom 4.

The functoriality of m amounts to

$$m(l\,{}_b\!\times_t l) = lm, \quad m(r\,{}_b\!\times_t r) = rm; \tag{1.7}$$

$$m(m\,{}_b\!\times_t m) = m(m \times m); \tag{1.8}$$

$$m(\mathrm{id}\,{}_b\!\times_t \mathrm{id}) = \mathrm{id}\, m. \tag{1.9}$$

Among these, (1.7) corresponds to identity (1.2) in Axiom 2, (1.8) corresponds to Axiom 3, and (1.9) corresponds to the left hand side identity in Axiom 6.

The functoriality of id amounts to

$$\mathrm{id}\, l = l\, \mathrm{id}, \quad \mathrm{id}\, r = r\, \mathrm{id}; \tag{1.10}$$

$$\mathrm{id}\, m = m(\mathrm{id} \times \mathrm{id}); \tag{1.11}$$

$$\mathrm{id}\,\mathrm{id} = \mathrm{id}\,\mathrm{id}\,. \tag{1.12}$$

Whence, (1.10) corresponds to the left hand side identity in Axiom 4, (1.11) corresponds to the right hand side identity in Axiom 6, and (1.12) corresponds to Axiom 5. The constructions and arguments are reversible, and this finishes the proof of the lemma. □

It is customary to represent a double category in the form of four related categories

$$\begin{array}{ccc} \mathcal{B} & \rightrightarrows & \mathcal{H} \\ \downdownarrows & & \downdownarrows \\ \mathcal{V} & \rightrightarrows & \mathcal{P} \end{array}$$

subject to the above axioms. So that the vertical arrows

$$\begin{array}{cc} \mathcal{B} & \mathcal{H} \\ \downdownarrows, & \downdownarrows, \\ \mathcal{V} & \mathcal{P} \end{array}$$

correspond to the categories $\mathcal{A}$ and $\mathcal{O}$ of 'arrows' and 'objects', respectively, while the horizontal arrows

$$\mathcal{B} \rightrightarrows \mathcal{H}, \quad \mathcal{V} \rightrightarrows \mathcal{P},$$

correspond to the functors $t, b : \mathcal{A} \rightrightarrows \mathcal{O}$.

Remark 1.3. The *transpose* of a double category $\mathcal{T}$ is the double category $\mathcal{T}^t$ with the same boxes and points as $\mathcal{T}$ but interchanging the rôles of the horizontal and vertical categories. A box $B \in \mathcal{B}$ is denoted B^t when regarded in $\mathcal{T}^t$. This remark allows to deduce some "horizontal" statements from "vertical" ones (or vice versa), by passing to the transpose double category.

1.2. **Examples.**

A 2-category is just a double category where all 'vertical arrows' are identities; *i.e.*, where every element of $\mathcal{V}$ is an identity. In this case the elements of $\mathcal{H}$ are the *morphisms* of the 2-category, while the elements of $\mathcal{B}$ are the *2-cells*.

Let $\mathcal{C}$ be a small category. Then the class of all square diagrams in $\mathcal{C}$, that is all the diagrams

$$\begin{array}{ccc} X & \xrightarrow{x} & Y \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle g} \\ Z & \xrightarrow{y} & W \end{array}$$

is a double category (without assuming commutativity of the diagram!).

More examples arise considering all commutative diagrams, or all diagrams whose vertical, respectively horizontal, arrows live in a fixed subcategory. For instance, given a group Σ and two subgroups F and G, one can consider Σ as a category with only one object, and then the double category whose vertical, respectively horizontal, arrows live in F, respectively in G.

A particular case of the preceding remark is the following example, which shows a connection between double categories and some constructions in Hopf algebra theory.

Example 1.4. Let $\triangleleft : G \times F \to G$, $\triangleright : G \times F \to F$, be a matched pair of finite groups. The notation explained in [Mj2] (see also [DVVV], [T2], [AN, Appendix]) coincides with the pictorial representation of the double category defined in what follows.

Take as $\mathcal{P}$ a set with a single element: $\mathcal{P} := \{*\}$. Put also $\mathcal{B} := G \times F$, $\mathcal{H} := G$ and $\mathcal{V} := F$. We have a double category

$$\begin{array}{ccc} G \times F & \rightrightarrows & G \\ \downdownarrows & & \downdownarrows \\ F & \rightrightarrows & \mathcal{P}, \end{array}$$

defined as follows:

$F \rightrightarrows \mathcal{P}$ and $G \rightrightarrows \mathcal{P}$ are the groupoids associated to the group structure on F and G, respectively. The groupoids $G \times F \rightrightarrows G$ and $G \times F \rightrightarrows F$ are the ones corresponding to the actions $\triangleleft$ and $\triangleright$, respectively; see [R, Example 1.2.a]. More precisely, we have:

- $G \times F \rightrightarrows G$ is the category whose objects are the elements of G, the arrows are the elements of $G \times F$ and the source and target maps are defined, respectively, by

$$t := \triangleleft : G \times F \to G, \qquad b := p_1 : G \times F \to G.$$

The composition $m : (G \times F)_t \times_b (G \times F) \to G \times F$ and the identity map $\mathrm{id} : G \to G \times F$, are determined, respectively, by

$$(g, x).(h, y) := (g, xy), \qquad \mathrm{id}(g) = (g, 1),$$

for all $g, h \in G$, $x, y \in F$, such that $g \triangleleft x = h$. The inverse map is defined as $(g, x)^{-1} = (g \triangleleft x, x^{-1})$.

- $G \times F \rightrightarrows F$ is the category whose objects are the elements of F, the arrows are the elements of $G \times F$ and the structure maps are defined by

$$r := \triangleright : G \times F \to F, \qquad l := p_2 : G \times F \to F,$$

and composition $m : (G \times F)_t \times_b (G \times F) \to G \times F$ and identity $\mathrm{id} : G \to G \times F$, are determined by

$$(g, x).(h, y) := (hg, x), \qquad \mathrm{id}(x) = (1, x),$$

for all $g, h \in G$, $x, y \in F$, such that $g \triangleright x = y$. The inverse map is given in this case by $(g, x)^{-1} = (g^{-1}, g \triangleright x)$.

1.3. Basic properties.

In this section we shall prove some basic properties of double categories and introduce some terminology that will be of use in later sections.

For an element $A \in \mathcal{B}$, we shall use the notation A^h (respectively, A^v) for the horizontal (respectively, vertical) inverse of A, provided they exist; these are defined respectively by the relations

$$AA^h = \mathbf{id}\, l(A), \quad A^h A = \mathbf{id}\, r(A), \qquad \begin{matrix} A \\ A^v \end{matrix} = \mathbf{id}\, t(A), \quad \begin{matrix} A^v \\ A \end{matrix} = \mathbf{id}\, b(A).$$

Lemma 1.5. *Let $\mathcal{T}$ be a double category and let $A \in \mathcal{B}$. Suppose that $A = l \underset{b}{\overset{t}{\square}} r$ is invertible with respect to horizontal composition. Then $t = t(A), b = b(A) \in \mathcal{H}$ are invertible and we have*

$$A^h = r \underset{b^{-1}}{\overset{t^{-1}}{\square}} l\,.$$

Similarly, if A is invertible with respect to vertical composition, then $l = l(A), r = r(A) \in \mathcal{V}$ are invertible and we have

$$A^v = l^{-1} \underset{t}{\overset{b}{\square}} r^{-1}.$$

Proof. It follows from (1.1), (1.2) and Axiom 4. □

Remark 1.6. Suppose that $g \in \mathcal{V}$ is invertible. It follows from axioms 4, 5 and 6, that $\mathbf{id}\, g$ is vertically invertible and $(\mathbf{id}\, g)^v = \mathbf{id}\, g^{-1}$. Also, if $x \in \mathcal{H}$ is invertible, then $\mathbf{id}\, x$ is horizontally invertible and $(\mathbf{id}\, x)^h = \mathbf{id}\, x^{-1}$.

Lemma 1.7. *Let $\mathcal{T}$ be a double category and let $X, R \in \mathcal{B}$.*

(i) Suppose that $\dfrac{X}{R}$, *and that X and R are horizontally invertible. Then* $\begin{matrix} X \\ R \end{matrix}$ *is horizontally invertible,* $\dfrac{X^h}{R^h}$ *and* $\begin{matrix} X^h \\ R^h \end{matrix} = \begin{Bmatrix} X \\ R \end{Bmatrix}^h$.

(ii) Suppose that $X|R$, and that X and R are vertically invertible. Then XR is vertically invertible, $X^v|R^v$ and $X^vR^v = \{XR\}^v$.

Proof. We prove part (i), part (ii) being entirely similar; it can also be deduced from part (i) by going to the transpose double category. It is clear that $\dfrac{X^h}{R^h}$. On the other hand, using the interchange law, we compute

$$\begin{Bmatrix} X^h \\ R^h \end{Bmatrix} \begin{Bmatrix} X \\ R \end{Bmatrix} = \begin{Bmatrix} X^h X \\ R^h R \end{Bmatrix} = \begin{Bmatrix} \mathbf{id}\, r(X) \\ \mathbf{id}\, r(R) \end{Bmatrix} = \mathbf{id}\, r \begin{Bmatrix} X \\ R \end{Bmatrix},$$

by axioms 6 and 2. A similar computation shows that $\begin{Bmatrix} X \\ R \end{Bmatrix} \begin{Bmatrix} X^h \\ R^h \end{Bmatrix} = \mathbf{id}\, l \begin{Bmatrix} X \\ R \end{Bmatrix}$. This proves the lemma. □

Lemma 1.8. *Let $A \in \mathcal{B}$ such that A is horizontally and vertically invertible. Assume in addition that A^h is vertically invertible and A^v is horizontally invertible. Then $(A^h)^v = (A^v)^h$.*

We shall use the notation $A^{-1} := (A^h)^v = (A^v)^h$; thus $A^{-1} = r^{-1} \underset{t^{-1}}{\overset{b^{-1}}{\boxed{}}} l^{-1}$.

Proof. We have $\begin{array}{c|c} (A^v)^h & A^v \\ \hline A^h & A \end{array}$ and also $\begin{array}{c|c} (A^h)^v & A^v \\ \hline A^h & A \end{array}$. Using the axioms, we compute

$$\begin{Bmatrix} (A^v)^h & A^v \\ A^h & A \end{Bmatrix} = \begin{matrix} \{(A^v)^h A^v\} \\ \{A^h A\} \end{matrix} = \begin{matrix} \mathbf{id}\, r(A)^{-1} \\ \mathbf{id}\, r(A) \end{matrix} = \mathbf{id}\,(r(A)^{-1} r(A)) = \Theta_{br(A)}.$$

On the other hand, we have

$$\begin{Bmatrix} (A^h)^v & A^v \\ A^h & A \end{Bmatrix} = \begin{Bmatrix} (A^h)^v \\ A^h \end{Bmatrix} \begin{Bmatrix} A^v \\ A \end{Bmatrix} = \mathbf{id} b(A)^{-1} \mathbf{id} b(A) = \mathbf{id}(b(A)^{-1} b(A)) = \Theta_{rb(A)}.$$

Hence we get $\left\{\begin{matrix}(A^v)^h & A^v \\ A^h & A\end{matrix}\right\} = \left\{\begin{matrix}(A^h)^v & A^v \\ A^h & A\end{matrix}\right\}$. The horizontal cancellation of $\left\{\begin{matrix}A^v \\ A\end{matrix}\right\}$, which is licit after Remark 1.6, and the vertical cancellation of A^h imply the desired identity. □

Remark 1.9. As pointed out by the referee, the basic properties in Subsection 1.3 can also be seen as consequences of Lemma 1.2. For instance, Lemma 1.5 becomes immediate using the fact that r and l are functors $(\mathcal{B} \rightrightarrows \mathcal{H}) \underset{l}{\overset{r}{\rightrightarrows}} (\mathcal{V} \rightrightarrows \mathcal{P})$, and similarly for t and b. Also, Remark 1.6 is a consequence of the fact that $\mathbf{id} : \mathcal{H} \to \mathcal{B}$ and $\mathbf{id} : \mathcal{V} \to \mathcal{B}$ are functors. The proofs of Lemmas 1.7 and 1.8 can also be done using functoriality of the composition map.

1.4. **Double groupoids.**

A *double groupoid* [E, BS] is a double category such that all the four component categories are groupoids. Note that a double groupoid is the same thing as a groupoid object in the full subcategory **Grpd** of **Cat** whose objects are groupoids. Lemma 1.5 implies that a double category

$$\begin{matrix} \mathcal{B} & \rightrightarrows & \mathcal{H} \\ \downdownarrows & & \downdownarrows \\ \mathcal{V} & \rightrightarrows & \mathcal{P} \end{matrix}$$

is a double groupoid if and only if the component categories $\mathcal{B} \rightrightarrows \mathcal{H}$ and $\mathcal{B} \rightrightarrows \mathcal{V}$ are groupoids.

The transpose of a double groupoid is a double groupoid.

We next include some helpful technical results on general double groupoids.

Lemma 1.10. *Let $\mathcal{T}$ be a double groupoid, and let $A, B, C \in \mathcal{B}$. The following are equivalent:*

(i) $A|B|C$ and $ABC = \mathbf{id}\, t(ABC)$;

(ii) there exist $U, V \in \mathcal{B}$ such that

$$\begin{array}{c|c|c} A & U & \\ \hline & V & C \end{array}, \qquad \begin{matrix} U \\ V \end{matrix} = B, \qquad AU = \mathbf{id}\, t(AB), \qquad VC = \mathbf{id}\, b(BC);$$

(iii) there exist $U', V' \in \mathcal{B}$ such that

$$\begin{array}{c|c|c} & U' & C \\ \hline A & V' & \end{array}, \qquad \begin{matrix} U' \\ V' \end{matrix} = B, \qquad AV' = \mathbf{id}\, b(AB), \qquad U'C = \mathbf{id}\, t(BC).$$

Moreover, in (ii) and (iii) the elements U, V, U', V' are uniquely determined by A, B, C, and we have

$$\begin{array}{c|c|c} A & U & \mathbf{id}\, t(C) \\ \hline \mathbf{id}\, b(A) & V & C \end{array}, \qquad \textit{respectively} \qquad \begin{array}{c|c|c} \mathbf{id}\, t(A) & U' & C \\ \hline A & V' & \mathbf{id}\, b(C) \end{array}.$$

Proof. Let U, V as in (ii). The uniqueness of U and V follows from cancellation properties in a double groupoid. Since $VC = \mathbf{id}\, t(VC)$, $l(V) = \text{id } lt(VC)$, and on the other hand $r(\mathbf{id}\, b(A)) = \text{id } rb(A)$. Now,

$$lt(VC) = lt(V) = lb(U) = bl(U) = br(A) = rb(A),$$

since $A|U$. This shows that $\mathbf{id}\, b(A)|V$. Similarly, using that $AU = \mathbf{id}\, t(AU)$, we get $r(U) = \text{id } rt(AU)$, and

$$rt(AU) = rb(AU) = rb(U) = rt(V) = tr(V) = tl(C) = lt(C).$$

Since $l(\mathbf{id}\, t(C)) = \text{id } lt(C)$, we get $U|\mathbf{id}\, t(C)$. Then, $\begin{array}{c|c|c} A & U & \mathbf{id}\, t(C) \\ \hline \mathbf{id}\, b(A) & V & C \end{array}$, as claimed. The corresponding facts for U' and V' in (iii) are similarly established.

We shall show that (i) $\iff$ (ii). The proof of the equivalence of (i) and (iii) is similar and left to the reader.

(i) $\implies$ (ii). Let $x = t(A)t(B) \in \mathcal{H}$. Note that

$$r(A^h) = l(A) = \text{id } lt(AB) = \text{id } l(x) = l(\mathbf{id}\, x);$$

so that $A^h|\mathbf{id}\, x$. Define $U := A^h\mathbf{id}\, x$ and $V := \frac{U^v}{B}$. To see that V is well defined we compute

$$b(U^v) = t(U) = t(A^h)x = t(A)^{-1}t(A)t(B) = t(B),$$

and thus $\frac{U^v}{B}$. By definition, we have that $A|U$, $\frac{U}{V}$, $AU = \mathbf{id}\, t(AU)$ and $\frac{U}{V} = B$. Also,

$$r(V) = r(U^v)r(B) = r((A^h)^v)r(\mathbf{id}\, x)r(B) = l(A)^{-1} \text{id } r(x)r(B) = l(C),$$

the last equality since $B|C$ and $ABC = \mathbf{id}\, t(ABC)$. Thus $V|C$.

We now observe that

$$\begin{aligned} &r(U) = \text{id } rt(B) = \text{id } tr(B) = \text{id } tl(C) = \text{id } lt(C), \\ &l(V) = l(U^v)l(B) = l((A^h\mathbf{id}x)^v)l(B) = l(A^{-1}\mathbf{id}x)l(B) = r(A)^{-1}l(B) = \\ &\qquad \text{id } br(A) = \text{id } rb(A); \end{aligned}$$

hence $U|\mathbf{id}\,t(C)$ and $\mathbf{id}\,b(A)|V$. Thus $\begin{array}{c|c|c} A & U & \mathbf{id}\,t(C) \\ \hline \mathbf{id}\,b(A) & V & C \end{array}$. On the other hand $\mathbf{id}\,t(ABC) = ABC = \left\{\begin{matrix} A & U & \mathbf{id}\,t(C) \\ \mathbf{id}\,b(A) & V & C \end{matrix}\right\}$. This implies that $VC = \mathbf{id}\,t(VC)$, and part (ii) follows.

(ii) $\implies$ (i). Since $VC = \mathbf{id}\,t(VC)$, $l(V) = \mathrm{id}\; lt(VC)$. Also, since $B = \begin{matrix} U \\ V \end{matrix}$, and $A|U$, we have $l(B) = l(U)l(V) = l(U) = r(A)$; therefore $A|B$. Similarly, the assumptions $C|V$ and $AU = \mathbf{id}\,t(AU)$ imply that $r(B) = r(U)r(V) = r(V) = l(C)$; hence $B|C$.

Finally, the interchange law gives

$$ABC = \left\{\begin{matrix} A & U & \mathbf{id}\,t(C) \\ \mathbf{id}\,b(A) & V & C \end{matrix}\right\} = \mathbf{id}\,t(ABC)$$

, and (i) follows. □

The following lemma is dual to Lemma 1.10. Its proof is left to the reader.

Lemma 1.11. *Let $\mathcal{T}$ be a double groupoid, and let $A, B, C \in \mathcal{B}$. The following are equivalent:*

(i) $\begin{array}{c} A \\ \hline B \\ \hline C \end{array}$ *and* $\begin{matrix} A \\ B \\ C \end{matrix} = \mathbf{id}\,l\begin{pmatrix} A \\ B \\ C \end{pmatrix}$;

(ii) there exist $U, V \in \mathcal{B}$ such that

$$\begin{array}{c|c} A & \\ \hline U & V \\ \hline & C \end{array}, \qquad UV = B, \qquad \begin{matrix} A \\ U \end{matrix} = \mathbf{id}\,l\begin{pmatrix} A \\ U \end{pmatrix}, \qquad \begin{matrix} V \\ C \end{matrix} = \mathbf{id}\,l\begin{pmatrix} V \\ C \end{pmatrix};$$

(iii) there exist $U', V' \in \mathcal{B}$ such that

$$\begin{array}{c|c} & A \\ \hline U' & V' \\ \hline C & \end{array}, \qquad U'V' = B, \qquad \begin{matrix} A \\ V' \end{matrix} = \mathbf{id}\,l\begin{pmatrix} A \\ V' \end{pmatrix}, \qquad \begin{matrix} U' \\ C \end{matrix} = \mathbf{id}\,l\begin{pmatrix} U' \\ C \end{pmatrix}.$$

The elements U, V, U', V' in (ii) and (iii) are uniquely determined by A, B, C, and we have

$$\begin{array}{c|c} A & \mathbf{id}\,r(A) \\ \hline U & V \\ \hline \mathbf{id}\,l(C) & C \end{array}, \qquad \textit{respectively} \qquad \begin{array}{c|c} \mathbf{id}\,l(A) & A \\ \hline U' & V' \\ \hline C & \mathbf{id}\,r(C) \end{array}.$$

□

The following result is needed in the proof of Theorem 3.1.

Lemma 1.12. *The following properties hold in a double groupoid $\mathcal{T}$.*

(a) Let $A, X, Y, Z \in \mathcal{B}$ such that

$$\begin{array}{c|c|c} & X^{-1} & \\ \hline X & Y & Z \\ \hline & Z^{-1} & \end{array} \quad . \tag{1.13}$$

Then the following conditions are equivalent:

$$XYZ = A. \tag{1.14}$$

$$\left\{\begin{matrix} X^{-1} \\ Y \\ Z^{-1} \end{matrix}\right\} = A^{-1}. \tag{1.15}$$

(b) The collection $X = A = Z$, $Y = A^h$ satisfies (1.13), (1.14) *and* (1.15).

Proof. Part (b) being straightforward, we prove (a). Suppose that (1.14) holds. By assumption we have $b(X)b(Y)b(Z) = b(A)$, and $b(Y) = t(Z^{-1}) = b(Z)^{-1}$. Therefore $b(X) = b(A) = t(A^v)$. Similarly, $t(Z) = t(A) = b(A^v)$. Also, $r(X^{-1}) = l(X)^{-1} = l(A)^{-1} = l(A^v)$, and $r(A^v) = r(A)^{-1} = r(Z)^{-1} = l(Z^{-1})$. This implies that

$$\begin{array}{c|c|c} X^v & X^{-1} & A^v \\ \hline X & Y & Z \\ \hline A^v & Z^{-1} & Z^v \end{array} \quad . \tag{1.16}$$

We compute in two different ways, using the interchange law:

$$\begin{aligned} \left\{\begin{matrix} X^v & X^{-1} & A^v \\ X & Y & Z \\ A^v & Z^{-1} & Z^v \end{matrix}\right\} &= \begin{matrix} \{X^v X^{-1} A^v\} \\ \{XYZ\} \\ \{A^v Z^{-1} Z^v\} \end{matrix} = \left\{\begin{matrix} A^v \\ A \\ A^v \end{matrix}\right\} = A^v \\ &= \left\{\begin{matrix} X^v \\ X \\ A^v \end{matrix}\right\} \left\{\begin{matrix} X^{-1} \\ Y \\ Z^{-1} \end{matrix}\right\} \left\{\begin{matrix} A^v \\ Z \\ Z^v \end{matrix}\right\} = A^v \left\{\begin{matrix} X^{-1} \\ Y \\ Z^{-1} \end{matrix}\right\} A^v; \end{aligned}$$

thus $\left\{\begin{matrix} X^{-1} \\ Y \\ Z^{-1} \end{matrix}\right\} = (A^v)^h = A^{-1}$, as claimed.

Conversely, suppose that (1.15) holds. Then, in the transpose double category $\mathcal{T}^t$,

$$\begin{array}{c|c|c} & X^t & \\ \hline (X^t)^{-1} & Y^t & (Z^t)^{-1} \\ \hline & Z^t & \end{array}, \quad \text{and} \quad (X^t)^{-1} Y^t (Z^t)^{-1} = (A^t)^{-1};$$

thus $\left\{\begin{matrix} X^t \\ Y^t \\ Z^t \end{matrix}\right\} = A^t$ in $\mathcal{T}^t$ by the preceding; that is, (1.14) holds in $\mathcal{T}$. □

2. Vacant double groupoids

2.1. Definition and basic properties.

The notion of vacant double groupoids appears in [Ma, Definition 2.11].

Definition 2.1. Let $\mathcal{T}$ be a double groupoid. We shall say that $\mathcal{T}$ is *vacant* if for any $g \in \mathcal{V}$, $x \in \mathcal{H}$ such that $r(x) = t(g)$, there is exactly one $X \in \mathcal{B}$ such that $X = \overset{x}{\square}\, g$.

We give an alternative description of vacant double groupoids that we have found in the course of our research; see condition 2 below. This will be useful in Section 3.

Proposition 2.2. *Let $\mathcal{T}$ be a double groupoid. The following are equivalent.*

(1) *$\mathcal{T}$ is vacant.*

(2) *For all $R, S, P \in \mathcal{B}$ such that $\frac{R}{S}$ and $P| \begin{Bmatrix} R \\ S \end{Bmatrix}$, there exist unique $X, Y \in \mathcal{B}$ such that $\begin{array}{c|c} X & R \\ \hline Y & S \end{array}$ and $P = \frac{X}{Y}$.*

(3) *For any $f \in \mathcal{V}$, $y \in \mathcal{H}$ such that $l(y) = b(f)$, there is exactly one $Z \in \mathcal{B}$ such that $Z = f\, \underset{y}{\square}$.*

(4) *For all $T, U, Q \in \mathcal{B}$ such that $T|U$ and $\frac{Q}{TU}$, there exist unique $V, Z \in \mathcal{B}$ such that $\begin{array}{c|c} V & Z \\ \hline T & U \end{array}$ and $Q = VZ$.*

(5) *For any $f \in \mathcal{V}$, $x \in \mathcal{H}$ such that $l(x) = t(f)$, there is exactly one $Z \in \mathcal{B}$ such that $Z = f\, \overset{x}{\square}$.*

(6) *For any $g \in \mathcal{V}$, $y \in \mathcal{H}$ such that $r(y) = b(g)$, there is exactly one $Z \in \mathcal{B}$ such that $Z = \underset{y}{\square}\, g$.*

(7) *For all $A, B, X, Y \in \mathcal{B}$ such that $XY = \frac{A}{B}$, there exist unique $U, V, R, S \in \mathcal{B}$ with*

$$\begin{array}{c|c} U & V \\ \hline R & S \end{array}, \quad UV = A, \quad RS = B, \quad \frac{U}{R} = X, \quad \frac{V}{S} = Y. \tag{2.1}$$

Note that condition (4) says that the transpose double groupoid $\mathcal{T}^t$ is vacant.

Proof. (1) $\implies$ (2). Let $R, S, P \in \mathcal{B}$ such that $\frac{R}{S}$ and $P| \begin{Bmatrix} R \\ S \end{Bmatrix}$. Let $x = t(P)$, $g = l(R)$ and let X be the unique box of the form $\overset{x}{\square}\, g$. Set $Y = \frac{X^v}{P}$; then clearly $\begin{array}{c|c} X & R \\ \hline Y & S \end{array}$ and $P = \frac{X}{Y}$. Furthermore, if X', Y' are boxes with these properties then clearly X' should be of the form $\overset{x}{\square}\, g$, hence $X' = X$, by the uniqueness condition in (1); *a fortiori* $Y' = Y$.

(2) $\implies$ (1). Let $g \in \mathcal{V}$, $x \in \mathcal{H}$ such that $r(x) = t(g)$. Put $P = \mathbf{id}\, x$, $R = \mathbf{id}\, g$, $S = R^v = \mathbf{id}\, g^{-1}$; by part (2), there exist unique $X, Y \in \mathcal{B}$ such that $\begin{array}{c|c} X & R \\ \hline Y & S \end{array}$ and $P = \frac{X}{Y}$. Clearly, $X = \overset{x}{\square}\, g$.

Let now X' be of the form $\overset{x}{\square}\, g$. Let $Y' := \frac{(X')^v}{P}$; then $\begin{array}{c|c} X' & R \\ \hline Y' & S \end{array}$ and $P = \frac{X'}{Y'}$. By the uniqueness in part (2), $X = X'$ (and $Y = Y'$).

(3) $\iff$ (4). This follows from the equivalence (1) $\iff$ (2) for $\mathcal{T}^t$.

(1) $\iff$ (3). If Z is of the form $f \underset{y}{\square}$, then Z^{-1} is of the form $\overset{y^{-1}}{\square}\, f^{-1}$. This remark implies the desired equivalence.

The proofs of the equivalences (1) $\iff$ (5) and (1) $\iff$ (6) are analogous, using Z^h and Z^v respectively.

(5) $\iff$ (7). A bijective correspondence between the set of quadruples (U, V, R, S) satisfying (2.1) and the set $\mathfrak{C} := \left\{ U \in \mathcal{B} : U = l(A) \overset{t(X)}{\square} \right\}$, is established by assigning to each $U \in \mathfrak{C}$ the quadruple $\left(U, U^h A, \frac{U^v}{X}, \frac{U^{-1} A^v}{Y} \right)$. The definition of $\mathfrak{C}$ guarantees that this map is well defined. Moreover, this defines a bijection, whose inverse is determined by the law $(U, V, R, S) \mapsto U$. □

Example 2.3. (i) The double category attached to a matched pair of finite groups as in in Example 1.4 is a vacant double groupoid.

(ii) Let $\mathcal{G}$ be any finite groupoid and consider the double category $\mathcal{T}$ of commuting square diagrams with vertical arrows in $\mathcal{G}$ but with horizontal arrows only identities. Then $\mathcal{T}$ is a vacant double groupoid.

We now complete the information in Lemma 1.12.

Lemma 2.4. *Let $\mathcal{T}$ be a vacant double groupoid and let $A \in \mathcal{B}$. There exists exactly one collection $X, Y, Z \in \mathcal{B}$ such that* (1.13), (1.14) *and* (1.15) *hold, namely $X = Z = A$, $Y = A^h$.*

Proof. By Lemma 1.12, the collection $X = Z = A$, $Y = A^h$ satisfies (1.13), (1.14) and (1.15). On the other hand, suppose that $X, Y, Z \in \mathcal{B}$ is any collection satisfying (1.13), (1.14) and (1.15). Then $l(X) = l(A)$ and $b(X)^{-1} = t(X^{-1}) = t(A^{-1}) = b(A)^{-1}$, hence $X = A$. Similarly, $Z = A$ and then necessarily $Y = A^h$. □

We list several technical facts about vacant double groupoids that are needed later in this paper. The straightforward proof is left to the reader.

Proposition 2.5. *Let $\mathcal{T}$ be a vacant double groupoid, $C \in \mathcal{B}$.*

(i) *If a horizontal (resp. vertical) side of C is an identity, then C is a vertical (resp. horizontal) identity.*

(ii) *The set of pairs of boxes (A, B) such that* $\begin{array}{c|c} A & B \\ \hline C & \end{array}$, $AB = \operatorname{id} t(AB)$ *and* $\dfrac{A}{C} = \operatorname{id} l \begin{pmatrix} A \\ C \end{pmatrix}$ *is either*

$$\begin{cases} \emptyset, & \text{if } C \text{ is not a horizontal identity,} \\ \{(\Theta_P, \operatorname{id} x) : \ x \in \mathcal{H},\ l(x) = P\}, & \text{if } C = \operatorname{id} g,\ P = t(g). \end{cases}$$

(iii) *The set of pairs of boxes (A, B) such that* $\begin{array}{c|c} & C \\ \hline A & B \end{array}$, $AB = \operatorname{id} b(AB)$ *and* $\dfrac{C}{B} = \operatorname{id} r \begin{pmatrix} C \\ B \end{pmatrix}$ *is either*

$$\begin{cases} \emptyset, & \text{if } C \text{ is not a horizontal identity,} \\ \{(\operatorname{id} h, \Theta_Q) : \ h \in \mathcal{V},\ r(h) = Q\}, & \text{if } C = \operatorname{id} g,\ Q = b(g). \end{cases}$$

(iv) *The set of pairs of boxes* (A, B) *such that* $\begin{array}{c|c} A & B \\ \hline B^{-1} & \end{array}$, $AB = C$ *is either*

$$\begin{cases} \emptyset, & \text{if } C \text{ is not a horizontal identity,} \\ \{(g\overset{z}{\square}, \overset{z^{-1}}{\square}g) : z \in \mathcal{H},\ l(z) = t(g)\}, & \text{if } C = \operatorname{id} g. \end{cases}$$

(v) *The set of pairs of boxes* (A, B) *such that* $\begin{array}{c|c} & A^{-1} \\ \hline A & B \end{array}$, $AB = C$ *is either*

$$\begin{cases} \emptyset, & \text{if } C \text{ is not a horizontal identity,} \\ \{(g\underset{w^{-1}}{\square}, \underset{w}{\square}g) : w \in \mathcal{H},\ r(w) = b(g)\}, & \text{if } C = \operatorname{id} g. \end{cases}$$

□

2.2. Matched pairs of groupoids.

We shall now give a characterization of vacant double groupoids in terms of matched pairs of groupoids. This characterization is due to Mackenzie [Ma].

Definition 2.6. Let $\mathcal{G}$ be a groupoid with base $\mathcal{P}$ and source and target maps $s, e : \mathcal{G} \rightrightarrows \mathcal{P}$. Let also $p : \mathcal{E} \to \mathcal{P}$ be a map. A *left action* of $\mathcal{G}$ on p is a map $\triangleright : \mathcal{G}\ {}_e\!\times_p \mathcal{E} \to \mathcal{E}$ such that

(2.2) $p(g \triangleright x) = s(g)$,

(2.3) $g \triangleright (h \triangleright x) = gh \triangleright x$,

(2.4) $\mathbf{id}\, p(x) \triangleright x = x$,

for all $g, h \in \mathcal{G}$, $x \in \mathcal{E}$ composable in the appropiate sense.

Hence, if $\mathcal{E}_b := p^{-1}(b)$, then the action of $g \in \mathcal{G}$ is an isomorphism $g \triangleright \_\_ : \mathcal{E}_{t(g)} \to \mathcal{E}_{s(g)}$. This somewhat unpleasant notation is originated by our choice of juxtaposition to denote composition.

Given actions of $\mathcal{G}$ on $p : \mathcal{E} \to \mathcal{P}$ and $p' : \mathcal{E}' \to \mathcal{P}$, a map $\phi : \mathcal{E} \to \mathcal{E}'$ is said to *intertwine* the actions if $p = \phi p'$ and $\phi(g \triangleright x) = g \triangleright \phi(x)$, for all $g \in \mathcal{G}$, $x \in \mathcal{E}$ such that $e(g) = p(x)$.

An action is *trivial* if there exists a set X such that $\mathcal{E} = \mathcal{P} \times X$, p is the first projection and $g \triangleright (e(g), x) = (s(g), x)$ for all $x \in X$, $g \in \mathcal{G}$.

Similarly, a *right action* of $\mathcal{G}$ on p is a map $\triangleleft : \mathcal{E}\,_p\times_s \mathcal{G} \to \mathcal{E}$ such that

(2.5) $p(x\triangleleft g) = e(g)$,

(2.6) $(x\triangleleft g)\triangleleft h = x\triangleleft gh$,

(2.7) $x\triangleleft \mathbf{id}\, p(x) = x$,

for all $g, h \in \mathcal{G}$, $x \in \mathcal{E}$ composable in the appropiate sense.

It is convenient to set the following notation: a *wide subgroupoid* of a groupoid $\mathcal{D}$ is a groupoid $\mathcal{V}$ provided with a functor $F : \mathcal{V} \to \mathcal{D}$ which is the identity on the objects, and induces inclusions on the hom sets. In other words, it has the same base, and (perhaps) less arrows.

The next two definitions generalize corresponding notions for finite groups, *cf.* [Ma, Definition 2.14].

Definition 2.7. A *matched pair of groupoids* is a pair of groupoids $t, b : \mathcal{V} \rightrightarrows \mathcal{P}$, $l, r : \mathcal{H} \rightrightarrows \mathcal{P}$, on the same base $\mathcal{P}$, endowed with a left action $\triangleright : \mathcal{H}\,_r\times_t \mathcal{V} \to \mathcal{V}$ of $\mathcal{H}$ on $t : \mathcal{V} \to \mathcal{P}$, and a right action $\triangleleft : \mathcal{H}\,_r\times_t \mathcal{V} \to \mathcal{H}$ of $\mathcal{V}$ on $r : \mathcal{H} \to \mathcal{P}$, satisfying

(2.8) $b(x\triangleright g) = l(x\triangleleft g)$,

(2.9) $x\triangleright fg = (x\triangleright f)((x\triangleleft f)\triangleright g)$,

(2.10) $xy\triangleleft g = (x\triangleleft(y\triangleright g))(y\triangleleft g)$,

for all $f, g \in \mathcal{V}$, $x, y \in \mathcal{H}$ such that the compositions are possible.

We claim that

$$x\triangleright\mathbf{id}\, r(x) = \mathbf{id}\, l(x), \qquad \text{for all } x \in \mathcal{H}, \tag{2.11}$$

$$\mathbf{id}\, t(g)\triangleleft g = \mathbf{id}\, b(g), \qquad \text{for all } g \in \mathcal{V}. \tag{2.12}$$

Indeed, $x\triangleright\mathbf{id}\, r(x) = x\triangleright(\mathbf{id}\, r(x)\mathbf{id}\, r(x)) = (x\triangleright\mathbf{id}\, r(x))\,((x\triangleleft\mathbf{id}\, r(x))\triangleright\mathbf{id}\, r(x)) = (x\triangleright\mathbf{id}\, r(x))\,(x\triangleright\mathbf{id}\, r(x))$, by (2.9) and (2.7). Since $t(x\triangleright\mathbf{id}\, r(x)) = l(x)$ by (2.2), (2.11) follows. Similarly (2.12) follows from (2.10), (2.4) and (2.5).

Definition 2.8. Let $\mathcal{D} \rightrightarrows \mathcal{P}$ be a groupoid. An *exact factorization* of $\mathcal{D}$ is a pair of wide subgroupoids $\mathcal{V}$, $\mathcal{H}$, such that for any $\alpha \in \mathcal{D}$, there exist unique $f \in \mathcal{V}$, $y \in \mathcal{H}$, such that $\alpha = fy$; that is, if the multiplication map $\mathcal{V}\,_b\times_l \mathcal{H} \to \mathcal{D}$ is a bijection.

Proposition 2.9. [Ma, Theorems 2.10 and 2.15] *The following notions are equivalent.*

1. *Matched pairs of groupoids.*
2. *Groupoids with an exact factorization.*
3. *Vacant double groupoids.*

If $\mathcal{V}$, $\mathcal{H}$ is a matched pair of groupoids, the groupoid arising in (2) will be denoted $\mathcal{D} = \mathcal{V} \bowtie \mathcal{H}$ and called the *diagonal* groupoid.

Proof. (1) $\Longrightarrow$ (2). Let $\triangleright : \mathcal{H}\,{}_r\!\times_t \mathcal{V} \to \mathcal{V}$, $\triangleleft : \mathcal{H}\,{}_r\!\times_t \mathcal{V} \to \mathcal{H}$ be a matched pair of groupoids on the same base $\mathcal{P}$. Let $\mathcal{D}$ be the groupoid on the base $\mathcal{P}$, with arrows $\mathcal{V}\,{}_b\!\times_l \mathcal{H}$, source and target maps $\alpha, \beta : \mathcal{V}\,{}_b\!\times_l \mathcal{H} \rightrightarrows \mathcal{P}$ given by $\alpha(f, y) = t(f)$, $\beta(f, y) = r(y)$, and composition defined by the rule: $(f, y)(h, z) = (f(y \triangleright h), (y \triangleleft h)z)$. We shall denote the arrow corresponding to $(f, y) \in \mathcal{V}\,{}_b\!\times_l \mathcal{H}$ by $\begin{array}{l|l} f & \\ \hline & y \end{array}$. A straightforward verification shows that $\mathcal{D}$ is indeed a well-defined groupoid. We identify $\mathcal{H}$, resp. $\mathcal{V}$, with the arrows of the form $\begin{array}{l|l} \mathbf{id}\, l(y) & \\ \hline & y \end{array}$, resp. $\begin{array}{l|l} f & \\ \hline & \mathbf{id}\, b(f) \end{array}$. Then the pair $\mathcal{V}, \mathcal{H}$ is an exact factorization of $\mathcal{D}$.

(2) $\Longrightarrow$ (1). Let $\mathcal{D}$ be a groupoid and let $\mathcal{V}$, $\mathcal{H}$ be an exact factorization of $\mathcal{D}$. Define $\triangleright : \mathcal{H}\,{}_r\!\times_t \mathcal{V} \to \mathcal{V}$, and $\triangleleft : \mathcal{H}\,{}_r\!\times_t \mathcal{V} \to \mathcal{H}$, by the formulas $xg = (x \triangleright g)(x \triangleleft g)$, $(x, g) \in \mathcal{H}\,{}_r\!\times_t \mathcal{V}$. The uniqueness of the factorization implies that $\triangleright$, $\triangleleft$ are well-defined. It is not difficult to see that these make $\mathcal{V}$, $\mathcal{H}$ into a matched pair of groupoids.

(3) $\Longrightarrow$ (1). Let $\mathcal{T}$ be a vacant double groupoid. Given $g \in \mathcal{V}$, $x \in \mathcal{H}$ such that $r(x) = t(g)$, we set $x \triangleleft g := b(X)$, $x \triangleright g := l(X)$ where $X \in \mathcal{B}$ is the unique box such that $X = \begin{matrix} x & \\ \square & g \end{matrix}$. That is, $X = \begin{matrix} & x & \\ x \triangleright g & \square & g \\ & x \triangleleft g & \end{matrix}$. The uniqueness gives at once that $\mathcal{V}$, $\mathcal{H}$, together with $\triangleleft$, $\triangleright$, is a matched pair of groupoids.

(1) $\Longrightarrow$ (3). Let $\mathcal{V}$, $\mathcal{H}$ be groupoids with the same set of objects $\mathcal{P}$, endowed with functions $\triangleright$, $\triangleleft$. Let $\mathcal{B} := \mathcal{H}\,{}_r\!\times_t \mathcal{V}$; we denote $X = (x, g) \in \mathcal{H}\,{}_r\!\times_t \mathcal{V}$ by $X = \begin{matrix} & x & \\ x \triangleright g & \square & g \\ & x \triangleleft g & \end{matrix}$. We leave to the reader the verification that this gives rise to a vacant double groupoid. □

Remark 2.10. Let $\mathcal{V}$, $\mathcal{H}$, be an exact factorization of a groupoid $\mathcal{D}$ and let $\mathcal{T}$ be the corresponding vacant double groupoid. Then $\mathcal{H}$, $\mathcal{V}$, is also an exact factorization of $\mathcal{D}$; the corresponding vacant double groupoid is $\mathcal{T}^t$.

2.3. Structure of vacant double groupoids.

2.3.1. *Structure of groupoids.* We first briefly recall the well-known structure of groupoids. Let $\mathcal{G} \rightrightarrows \mathcal{P}$ be a groupoid. We shall denote by $\mathcal{G}(x, y)$ the set of arrows from x to y; the set $\mathcal{G}(x, x)$ will be denoted by $\mathcal{G}(x)$.

There are two basic examples of groupoids:

- A group G, considered as the set of arrows of a category with a single object.
- An equivalence relation R on a set $\mathcal{P}$; s and e are respectively the first and the second projection, and the composition is given by $(x, y)(y, v) = (x, v)$. We shall denote by $\mathcal{P}^2$ the equivalence relation where all the elements of $\mathcal{P}$ are related; $\mathcal{P}^2$ is called the *coarse* groupoid on $\mathcal{P}$.

If $\mathcal{G} \rightrightarrows \mathcal{P}$ and $\mathcal{G}' \rightrightarrows \mathcal{P}'$ are two groupoids, then two basic operations are:

- The disjoint union $\mathcal{G} \coprod \mathcal{G}'$, a groupoid on the base $\mathcal{P} \coprod \mathcal{P}'$.
- The direct product $\mathcal{G} \times \mathcal{G}'$, a groupoid on the base $\mathcal{P} \times \mathcal{P}'$.

The structure of any groupoid can be described with the help of these basic examples and operations. Namely, let $\mathcal{G} \rightrightarrows \mathcal{P}$ be any groupoid and define an equivalence relation on $\mathcal{P}$ by $x \sim y$ iff $\mathcal{G}(x, y) \neq \emptyset$. We say that $\mathcal{G}$ is *connected* if $x \sim y$ for all $x, y \in \mathcal{P}$. The opposite case is a *group bundle*: this is a groupoid such that $x \sim y$ implies $x = y$. The trivial group bundle is $\mathcal{P} \rightrightarrows \mathcal{P}$, $s = e = \mathrm{id}$.

- If $\mathcal{G}$ is connected, then $\mathcal{G} \simeq \mathcal{G}(x) \times \mathcal{P}^2$, where x is any element of $\mathcal{P}$.

Indeed, choose an arrow $\tau_y \in \mathcal{G}(x, y)$ and define $F : \mathcal{G} \to \mathcal{G}(x) \times \mathcal{P}^2$, $F(\alpha) = (\tau_z^{-1} \alpha \tau_y, (y, z))$ if $y = s(\alpha)$, $z = t(\alpha)$. Then F is an isomorphism of groupoids.

- In general, let P be an equivalence class of $\sim$ and let $\mathcal{G}_P$ be corresponding groupoid on the base P; so that $\mathcal{G}_P(x, y) = \mathcal{G}(x, y)$, for all $x, y \in P$. Then $\mathcal{G} \simeq \coprod_{P \in \mathcal{P}/\sim} \mathcal{G}_P$.

These remarks provide the general structure of a groupoid.

2.3.2. *Structure of wide subgroupoids.* We now give a description of a wide subgroupoid of a connected groupoid in group-theoretical terms. We fix a finite non-empty set $\mathcal{P}$, a point $O \in \mathcal{P}$, and a finite group $D =: D(O)$. Let $\mathcal{D} = D(O) \times \mathcal{P}^2$ be the corresponding connected groupoid.

Lemma 2.11. *There is a bijective correspondence between the following data:*

(1) *Wide subgroupoids $\mathcal{H}$ of $\mathcal{D}$;*

(2) *collections $(\sim_H, (H_P)_{P \in \mathcal{P}}, (\overline{d_{PQ}})_{P \sim_H Q})$, where*
- *$\sim_H$ is an equivalence relation on $\mathcal{P}$,*
- *H_P is a subgroup of D, $P \in \mathcal{P}$,*

- $\overline{d_{PQ}}$ *is an element of* $H_P\backslash D/H_Q$ *such that for any representative* d_{PQ}, *the following hold:*

$$d_{PQ}H_Q = H_P d_{PQ}, \quad \text{if} \quad P \sim_H Q, \tag{2.13}$$

$$d_{PQ}d_{QR} \in H_P d_{PR}, \quad \text{if} \quad P \sim_H Q \sim_H R, \tag{2.14}$$

$$d_{PP} \in H_P, \quad P \in \mathcal{P}. \tag{2.15}$$

Proof. For each $P \in \mathcal{P}$, fix $\tau_P \in \mathcal{D}(O, P)$. The correspondence is not natural; it depends on the choice of the family $(\tau_P)_{P\in\mathcal{P}}$.

(1) $\Longrightarrow$ (2). The equivalence relation $\sim_H$ is defined by $P \sim_H Q$ iff $\mathcal{H}(P,Q) \neq \emptyset$. Given $P \in \mathcal{P}$, the subgroup H_P is defined by $H_P := \tau_P \mathcal{H}(P)\tau_P^{-1}$. If $P \sim_H Q$, choose $h_{PQ} \in \mathcal{H}(P,Q)$ and set $d_{PQ} := \tau_P h_{PQ}\tau_Q^{-1}$. Then

$$H_Q = \tau_Q\mathcal{H}(Q)\tau_Q^{-1} = \tau_Q h_{PQ}^{-1}\mathcal{H}(P)h_{PQ}\tau_Q^{-1} = d_{PQ}^{-1}\tau_P\mathcal{H}(P)\tau_P^{-1}d_{PQ} = d_{PQ}^{-1}H_P d_{PQ}.$$

This proves that condition (2.13) is satisfied. Conditions (2.14) and (2.15) are similarly verified.

If we choose another element $\widetilde{h_{PQ}} \in \mathcal{H}(P,Q)$ then $\widetilde{d_{PQ}} := \tau_P\widetilde{h_{PQ}}\tau_Q^{-1}$ has the same class in $H_P\backslash D/H_Q$ as d_{PQ}. Clearly, if (2.13), (2.14), (2.15) are true for some choice of representatives of $\overline{d_{PQ}}$ then they are true for any choice. This finishes the proof of the first implication.

(2) $\Longrightarrow$ (1). Define a wide subgroupoid $\mathcal{H}$ of $\mathcal{D}$ as follows: if $P, Q \in \mathcal{P}$, then

$$\mathcal{D}(P,Q) \supseteq \mathcal{H}(P,Q) := \begin{cases} \emptyset, & \text{if} \quad P \not\sim_H Q; \\ \tau_P^{-1}H_P d_{PQ}\tau_Q, & \text{if} \quad P \sim_H Q. \end{cases}$$

We have to check that $\mathcal{H}$ is stable under composition, inverses and identities. First, if $P \sim_H Q \sim_H R$ then

$$\begin{aligned} \mathcal{H}(P,Q)\mathcal{H}(Q,R) &= (\tau_P^{-1}H_P d_{PQ}\tau_Q)(\tau_Q^{-1}H_Q d_{QR}\tau_R) \\ &= \tau_P^{-1}H_P d_{PQ}d_{QR}\tau_R \\ &= \tau_P^{-1}H_P d_{PR}\tau_R, \end{aligned}$$

where the first equality is by definition, the second by (2.13) and the third by (2.14). Next, if $P \sim_H Q$, then

$$\mathcal{H}(P,Q)^{-1} = \tau_Q^{-1}d_{PQ}^{-1}H_P\tau_P = \tau_Q^{-1}H_Q d_{PQ}^{-1}\tau_P = \tau_Q^{-1}H_Q d_{QP}\tau_P = \mathcal{H}(Q,P),$$

using several times (2.13), (2.14) and (2.15). Similarly, $\mathrm{id}_P \in \mathcal{H}(P,P)$ by (2.15). The second implication is proved. □

2.3.3. *Double equivalence relations.* Let $\mathcal{P}$ be a finite non-empty set. Let $\sim_H$, $\sim_V$ be two equivalence relations on $\mathcal{P}$. Let $\mathcal{V} \rightrightarrows \mathcal{P}$ be the groupoid defined by the relation $\sim_V$, let $\begin{smallmatrix}\mathcal{H}\\ \downdownarrows\\ \mathcal{P}\end{smallmatrix}$ be the groupoid defined by the relation $\sim_H$, and let

$$\mathcal{B} = \left\{ \begin{pmatrix} P & Q \\ R & S \end{pmatrix} \in \mathcal{P}^{2\times 2} : P \sim_H Q,\, P \sim_V R,\, R \sim_H S,\, Q \sim_V S \right\}.$$

Let $\sim_D$ be the relation defined as follows: $P \sim_D Q$ if there exists $R \in \mathcal{P}$ such that $P \sim_H R$, $R \sim_V Q$. We shall sometimes denote this as $\begin{array}{ccc} P & - & R \\ & & | \\ & & Q \end{array}$.

Lemma 2.12. *(a). The maps* $\mathcal{B} \rightrightarrows \mathcal{H}$, $\begin{smallmatrix}\mathcal{B}\\ \downdownarrows\\ \mathcal{V}\end{smallmatrix}$ *given by* $\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \rightrightarrows \begin{matrix}(P & Q)\\(R & S)\end{matrix}$, $\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \rightrightarrows \begin{pmatrix} P \\ R \end{pmatrix}, \begin{pmatrix} Q \\ S \end{pmatrix}$, *with evident composition, define a double groupoid* $\begin{array}{ccc} \mathcal{B} & \rightrightarrows & \mathcal{H} \\ \downdownarrows & & \downdownarrows \\ \mathcal{V} & \rightrightarrows & \mathcal{P} \end{array}$*; it will be called a* double equivalence relation.

(b). The relation $\sim_D$ is an equivalence relation if and only if it is symmetric.

(c). The double equivalence relation is vacant if and only if $\sim_D$ is an equivalence relation and the following condition holds:

(2.16) *If* $R, S \in \mathcal{P}$, $R \sim_H S$, $R \sim_V S$, *then* $R = S$.

(d). Let $\begin{array}{ccc} \mathcal{B} & \rightrightarrows & \mathcal{H} \\ \downdownarrows & & \downdownarrows \\ \mathcal{V} & \rightrightarrows & \mathcal{P} \end{array}$ *be any* vacant *double groupoid and let $\sim_H$, $\sim_V$ be the equivalence relations on $\mathcal{P}$ defined by $\mathcal{H}$, $\mathcal{V}$ respectively. Then $\sim_D$ is an equivalence relation on $\mathcal{P}$.*

Proof. Part (a) is left to the reader.

(b). The relation $\sim_D$ is clearly reflexive. Assume that $\sim_D$ is symmetric. Let $P, Q, T \in \mathcal{P}$ such that $P \sim_D Q$, $Q \sim_D T$. Then there exist $R, S \in \mathcal{P}$ such that $\begin{array}{ccc} P & - & R \\ & & | \\ & & Q \end{array}$, $\begin{array}{ccc} Q & - & S \\ & & | \\ & & T \end{array}$. But then $\begin{array}{ccc} S & - & Q \\ & & | \\ & & R \end{array}$, *i. e.* $S \sim_D R$. By symmetry, there exists $V \in \mathcal{P}$ such that $\begin{array}{ccc} S & - & Q \\ | & & | \\ V & - & R \end{array}$. Hence $\begin{array}{ccc} P & - & V \\ & & | \\ & & T \end{array}$, *i. e.* $P \sim_D T$.

(c). Assume that the double equivalence relation is vacant. If $P \sim_D Q$, there exists $R \in \mathcal{P}$ such that $\begin{array}{ccc} P & — & R \\ & & | \\ & & Q \end{array}$. By vacancy, there exists $V \in \mathcal{P}$ such that $\begin{pmatrix} P & R \\ V & Q \end{pmatrix} \in \mathcal{B}$, that is, we have $\begin{array}{ccc} P & — & R \\ | & & | \\ V & — & Q \end{array}$, and thus $Q \sim_D P$. It follows that $\sim_D$ is symmetric and hence an equivalence relation, by (b). Now let $R, S \in \mathcal{P}$ be such that $\begin{array}{ccc} R & — & S \\ & & | \\ & & R \end{array}$. Then both $\begin{array}{ccc} R & — & S \\ | & & | \\ R & — & R \end{array}$ and $\begin{array}{ccc} R & — & S \\ | & & | \\ S & — & R \end{array}$ belong to $\mathcal{B}$, and by vacancy $R = S$. Thus (2.16) holds.

Conversely, assume that $\sim_D$ is an equivalence relation and (2.16) holds. Then any $\begin{array}{ccc} P & — & R \\ & & | \\ & & Q \end{array}$ can be extended to a box $\begin{array}{ccc} P & — & R \\ | & & | \\ V & — & Q \end{array}$ in $\mathcal{B}$, since $\sim_D$ is an equivalence relation. If also $\begin{array}{ccc} P & — & R \\ | & & | \\ U & — & Q \end{array}$ then clearly $\begin{array}{ccc} U & — & V \\ & & | \\ & & U \end{array}$ and hence $U = V$ by (2.16).

(d). By vacancy, the relation $\sim_D$ is symmetric, then apply (b). □

Example 2.13. Let Σ be a group, let F, G be subgroups of Σ, acting on Σ respectively on the left and on the right by multiplication. Let $\sim_H$, $\sim_V$ be the equivalence relations on Σ defined by these actions. Then the corresponding $\sim_D$ is an equivalence relation, where the associated partition of Σ is that given by the double cosets FqG, $q \in \Sigma$. Moreover, (2.16) holds if and only if $G \cap gFg^{-1} = 1$ for $g \in \Sigma$. (For instance, if the orders of F and G are relatively prime).

Definition 2.14. We shall say that a double equivalence relation is *connected* if the associated relation $\sim_D$ is so; that is, if any two elements of $\mathcal{P}$ are connected by $\sim_D$.

Let r, s be natural numbers. Let $\mathbb{X}_{rs}$ be the double equivalence relation on the set $\{1, \dots, r\} \times \{1, \dots, s\}$ and with side relations

$$(i, j) \sim_H (k, l) \iff i = k, \qquad (i, j) \sim_V (k, l) \iff j = l.$$

Clearly, $\mathbb{X}_{rs}$ is a *connected* vacant equivalence relation.

Proposition 2.15. *Any finite double equivalence relation which is vacant and connected is isomorphic to* $\mathbb{X}_{rs}$.

Proof. Let $Y_1, \dots, Y_r$ be the classes of $\sim_H$ in $\mathcal{P}$, and assume that $Y_1 = \{1, \dots, s\}$. We shall define a bijection $\phi_i : Y_1 \to Y_i$, $2 \le i \le s$ and shall prove that for $j \in Y_1$ and $k \in Y_i$, $j \sim_V k$ if and only if $k = \phi_i(j)$. So, let us fix i and set $\phi = \phi_i$. Fix $a \in Y_i$. If $j \in Y_1$, by connectedness and (2.16), there exists a unique k such that $\begin{matrix} j & \\ | & \\ k & - a \end{matrix}$. Set $\phi_i(j) = k$. We claim that ϕ_i is bijective.

Indeed, assume that $\phi_i(j) = k = \phi_i(h)$. Then $\begin{matrix} j & - j \\ | & \\ k & \end{matrix}$ and $\begin{matrix} h & - j \\ | & \\ k & \end{matrix}$; thus $j = h$ and ϕ is injective. Also, let $k \in Y_i$. Then $\begin{matrix} 1 & \\ | & \\ \phi(1) & - k \end{matrix}$, hence there exists a unique $j \in Y_1$ such that $\begin{matrix} 1 & - & j \\ | & & | \\ \phi(1) & - & k \end{matrix}$; thus $\phi(j) = k$ and ϕ is surjective. □

2.3.4. *Structure of vacant double groupoids.* Let $\mathcal{T}_1 = \begin{matrix} \mathcal{B}_1 & \rightrightarrows & \mathcal{H}_1 \\ \downdownarrows & & \downdownarrows \\ \mathcal{V}_1 & \rightrightarrows & \mathcal{P}_1 \end{matrix}$ and $\mathcal{T}_2 = \begin{matrix} \mathcal{B}_2 & \rightrightarrows & \mathcal{H}_2 \\ \downdownarrows & & \downdownarrows \\ \mathcal{V}_2 & \rightrightarrows & \mathcal{P}_2 \end{matrix}$ be double groupoids. Then we can define double groupoids

$$\mathcal{T}_1 \coprod \mathcal{T}_2 = \begin{matrix} \mathcal{B}_1 \coprod \mathcal{B}_2 & \rightrightarrows & \mathcal{H}_1 \coprod \mathcal{H}_2 \\ \downdownarrows & & \downdownarrows \\ \mathcal{V}_1 \coprod \mathcal{V}_2 & \rightrightarrows & \mathcal{P}_1 \coprod \mathcal{P}_2 \end{matrix}, \qquad \mathcal{T}_1 \times \mathcal{T}_2 = \begin{matrix} \mathcal{B}_1 \times \mathcal{B}_2 & \rightrightarrows & \mathcal{H}_1 \times \mathcal{H}_2 \\ \downdownarrows & & \downdownarrows \\ \mathcal{V}_1 \times \mathcal{V}_2 & \rightrightarrows & \mathcal{P}_1 \times \mathcal{P}_2 \end{matrix},$$

and similarly for families of double groupoids. If $\mathcal{T}_i$, $i \in I$, is a family of *vacant* double groupoids then $\coprod_{i\in I} \mathcal{T}_i$ and $\times_{i\in I}\mathcal{T}_i$ are also vacant.

Now, let $\mathcal{T}$ be a vacant double groupoid on the base $\mathcal{P}$, let $\mathcal{D}$ be the corresponding diagonal groupoid and let $\sim_D$ be the associated equivalence relation. Then $\mathcal{T} = \coprod_{P\in\mathcal{P}/\sim_D} \mathcal{T}_P$ where $\mathcal{T}_P$ is the "full" double groupoid with base the class P.

Let us say that a vacant double groupoid $\mathcal{T}$ is *connected* if the diagonal relation $\sim_D$ is transitive. The preceding remark shows that it is enough to consider connected vacant double groupoids.

As in subsection 2.3.2 above, we fix a finite non-empty set $\mathcal{P}$, a point $O \in \mathcal{P}$, and a finite group D, and set $\mathcal{D} = D(O) \times \mathcal{P}^2$.

Theorem 2.16. *Let $\mathcal{H}$, $\mathcal{V}$ be wide subgroupoids of $\mathcal{D}$ associated to data $(\sim_H, (H_P)_{P\in\mathcal{P}}, (\overline{d_{PQ}})_{P\sim_H Q})$ and $(\sim_V, (V_P)_{P\in\mathcal{P}}, (\overline{e_{PQ}})_{P\sim_H Q})$, respectively, as in Lemma 2.11.*

The following are equivalent:

(1) $\mathcal{D} = \mathcal{V}\mathcal{H}$ *is an exact factorization.*
(2) *The following conditions hold:*
 (a) *For all* $P, Q \in \mathcal{P}$, *one has*

$$D = \coprod_{R \in \mathcal{P}: P \sim_H R, R \sim_V Q} V_P \, e_{PR} d_{RQ} \, H_Q. \tag{2.17}$$

 (b) *For all* $P \in \mathcal{P}$, $V_P \cap H_P = e$.

Note that (a) implies that $\sim_D$ is an equivalence relation on $\mathcal{P}$, *cf.* subsection 2.3.3 above; this agrees with Lemma 2.12 (d).

Proof. (1) $\implies$ (2).We show (a). Since $\mathcal{T}$ is vacant, we have

$$D = \tau_P \mathcal{D}(P,Q)\tau_Q^{-1} = \tau_P \left(\coprod_{R \in \mathcal{P}: P \sim_V R, R \sim_H Q} \mathcal{V}(P,R)\mathcal{H}(R,Q) \right) \tau_Q^{-1}$$
$$= \coprod_{R \in \mathcal{P}: P \sim_V R, R \sim_H Q} V_P e_{PR} \tau_R \tau_R^{-1} H_R d_{RQ} = \coprod_{R \in \mathcal{P}: P \sim_V R, R \sim_H Q} V_P e_{PR} d_{RQ} H_Q,$$

as claimed. We show (b). Let $g \in V_P \cap H_P$. Then $\tau_P g \tau_P^{-1} \in \mathcal{V}(P) \cap \mathcal{H}(P)$; but $\mathcal{V}(P) \cap \mathcal{H}(P) = \mathrm{id}_P$ since $\mathcal{T}$ is vacant. Thus $g = e$.

(2) $\implies$ (1). Let $x \in \mathcal{H}$, $g \in \mathcal{V}$ such that $S := r(x) = t(g)$, and set $P = l(x)$, $Q = b(g)$. That is, we have $\overset{x}{\overline{\quad}}\!\rceil_g$. Now $\gamma := xg \in \mathcal{D}(P,Q) = \coprod_{R \in \mathcal{P}: P \sim_V R, R \sim_H Q} \mathcal{V}(P,R)\mathcal{H}(R,Q)$ by assumption (a). Thus there exist $R \in \mathcal{P}$ (unique!), $f \in \mathcal{V}(P,R)$ and $y \in \mathcal{H}(R,Q)$ such that $\gamma = fy$, in other words $f \overset{x}{\underset{y}{\square}} g \in \mathcal{B}$. Moreover, assume that also $h \overset{x}{\underset{w}{\square}} g \in \mathcal{B}$; note $f \in \mathcal{V}(P,R)$ and $y \in \mathcal{H}(R,Q)$. Then $z := h^{-1}f = wy^{-1} \in \mathcal{V}(R) \cap \mathcal{H}(R)$; by hypothesis (b), $z = e$. This implies that $\mathcal{T} = \mathcal{V}\mathcal{H}$ is an exact factorization. □

3. Weak Hopf algebras arising from a vacant double groupoid

Let F, G be a matched pair as in Example 1.4. As explained in many places, see *e. g.* [AN, 5.3], a bicrossed product Hopf algebra $\Bbbk^{G^\tau} \#_\sigma \Bbbk F$ admits a convenient realization in the vector space with basis $\mathcal{B}$. In this section we shall discuss a generalization of this construction.

3.1. Weak Hopf algebras (quantum groupoids).

Recall [BNSz, BSz] that a *weak bialgebra* structure on a vector space H over a field $\Bbbk$ consists of an associative algebra structure $(H, m, 1)$, a coassociative coalgebra structure (H, Δ, ε), such that the following are satisfied:

$$\Delta(ab) = \Delta(a)\Delta(b), \qquad \forall a, b \in H. \tag{3.1}$$

$$\Delta^{(2)}(1) = (\Delta(1) \otimes 1)(1 \otimes \Delta(1)) = (1 \otimes \Delta(1))(\Delta(1) \otimes 1). \tag{3.2}$$

$$\varepsilon(abc) = \varepsilon(ab_1)\varepsilon(b_2c) = \varepsilon(ab_2)\varepsilon(b_1c), \qquad \forall a, b, c \in H. \tag{3.3}$$

A weak bialgebra H is called a *weak Hopf algebra* or a *quantum groupoid* if there exists a linear map $\mathcal{S} : H \to H$ satisfying

$$m(\mathrm{id} \otimes \mathcal{S})\Delta(h) = (\varepsilon \otimes \mathrm{id})(\Delta(1)(h \otimes 1)) =: \varepsilon_t(h), \tag{3.4}$$

$$m(\mathcal{S} \otimes \mathrm{id})\Delta(h) = (\mathrm{id} \otimes \varepsilon)((1 \otimes h)\Delta(1)) =: \varepsilon_s(h), \tag{3.5}$$

$$m^{(2)}(\mathcal{S} \otimes \mathrm{id} \otimes \mathcal{S})\Delta^{(2)} = \mathcal{S}, \tag{3.6}$$

for all $h \in H$. The maps ε_s, ε_t are respectively called the source and target maps; their images are respectively called the source and target subalgebras. See [NV] for a survey on quantum groupoids. It is known that a weak Hopf algebra is a true Hopf algebra if and only if $\Delta(1) = 1 \otimes 1$.

3.2. Weak Hopf algebras arising from vacant double groupoids.

Let $\mathcal{T}$ be a *finite* double groupoid, that is, $\mathcal{B}$, $\mathcal{V}$, $\mathcal{H}$ and $\mathcal{P}$ are finite sets.

Let $\Bbbk$ be a field [2] and let $\Bbbk\mathcal{T}$ denote the $\Bbbk$-vector space with basis $\mathcal{B}$ together with the following structures.

Algebra structure. Consider the groupoid algebra structure on $\Bbbk\mathcal{T}$ corresponding to the groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$. Thus the multiplication in $\Bbbk\mathcal{T}$ is given by

$$A.B = \begin{cases} \dfrac{A}{B}, & \text{if } \dfrac{A}{B}, \\ 0, & \text{otherwise,} \end{cases}$$

for all $A, B \in \mathcal{B}$. This multiplication is associative and has a unit $\mathbf{1} := \sum_{x \in \mathcal{H}} \mathbf{id}\, x$. We shall also consider, for any $P \in \mathcal{P}$, the elements

$$_P\mathbf{1} = \sum_{x \in \mathcal{H}, l(x) = P} \mathbf{id}\, x, \qquad \mathbf{1}_P = \sum_{x \in \mathcal{H}, r(x) = P} \mathbf{id}\, x.$$

Clearly, $_P\mathbf{1}\ _Q\mathbf{1} = \delta_{P,Q}\ _P\mathbf{1}$, $\mathbf{1}_P\mathbf{1}_Q = \delta_{P,Q}\mathbf{1}_P$, for all $P, Q \in \mathcal{P}$. Hence the subalgebras $\Bbbk\mathcal{T}_s$, respectively $\Bbbk\mathcal{T}_t$, generated by $\mathbf{1}_P$, $P \in \mathcal{P}$, respectively by $_P\mathbf{1}$, $P \in \mathcal{P}$, are commutative of dimension $|\mathcal{P}|$.

[2] Most of the constructions in this section are valid over an arbitrary commutative ring.

Coalgebra structure. Dually, we consider the coalgebra structure on $\Bbbk\mathcal{T}$ dual to the algebra structure of the groupoid algebra corresponding to the groupoid $\mathcal{B} \rightrightarrows \mathcal{V}$. This means that the comultiplication of $\Bbbk\mathcal{T}$ is determined by

$$\Delta(A) = \sum B \otimes C, \qquad A \in \mathcal{B},$$

where the sum runs over all B, C with $B|C$ and $A = BC$. This comultiplication is coassociative and has counit $\varepsilon : \Bbbk\mathcal{T} \to \Bbbk$ given by

$$\varepsilon(A) = \begin{cases} 1, & \text{if } A = \mathbf{id}\, l(A), \\ 0, & \text{otherwise.} \end{cases}$$

Theorem 3.1. *$\Bbbk\mathcal{T}$ is a quantum groupoid if and only if $\mathcal{T}$ is a vacant double groupoid. If this is the case, the antipode is defined by $\mathcal{S}(A) = A^{-1}$, $\forall A \in \mathcal{B}$; and the source and target subalgebras are respectively $\Bbbk\mathcal{T}_s$, $\Bbbk\mathcal{T}_t$.*

Proof. Let $A, B \in \mathcal{B}$. It follows from the definitions that

$$\Delta(A).\Delta(B) = \sum \frac{U}{R} \otimes \frac{V}{S},$$

where the sum runs over all elements $U, V, R, S \in \mathcal{B}$, such that $\begin{array}{c|c} U & V \\ \hline R & S \end{array}$, $UV = A$ and $RS = B$. It is thus clear that $\Delta(A).\Delta(B) = 0 = \Delta(A.B)$, if A and B are not vertically composable. So assume that $\frac{A}{B}$. Then

$$\Delta(A.B) = \sum_{XY = \frac{A}{B}} X \otimes Y.$$

Since $(X \otimes Y)_{X,Y \in \mathcal{B}}$ is a basis of $\Bbbk\mathcal{T} \otimes \Bbbk\mathcal{T}$, we see that $\Delta(A).\Delta(B) = \Delta(A.B)$ if and only if

$$1 = \# \left\{ \begin{pmatrix} U & V \\ R & S \end{pmatrix} \in \mathcal{B}^4 : \begin{array}{c|c} U & V \\ \hline R & S \end{array},\ UV = A,\ RS = B,\ \frac{U}{R} = X,\ \frac{V}{S} = Y \right\}$$

for all X, $Y \in \mathcal{B}$ such that $X|Y$, $XY = \frac{A}{B}$. By Proposition 2.2, we conclude that (3.1) holds if and only if $\mathcal{T}$ is vacant.

Assume for the rest of the proof that $\mathcal{T}$ is vacant. We next prove the relationships (3.2) and (3.3). We have $\Delta(\mathbf{1}) = \sum_{A|B,\, AB = \mathbf{id}\, t(AB)} A \otimes B$. But if $AB = \mathbf{id}\, t(AB)$ then the right side of B is an identity, hence $B = \mathbf{id}\, t(B)$ and the same for A. Thus

$$\Delta(\mathbf{1}) = \sum_{x,y \in \mathcal{H} : x|y} \mathbf{id}\, x \otimes \mathbf{id}\, y.$$

Therefore, $\Delta^{(2)}(\mathbf{1}) = \sum_{x,y,z\in\mathcal{H}:x|y|z} \mathbf{id}\, x\otimes\mathbf{id}\, y\otimes\mathbf{id}\, z = (\Delta(\mathbf{1})\otimes\mathbf{1})\,(\mathbf{1}\otimes\Delta(\mathbf{1})) = (\mathbf{1}\otimes\Delta(\mathbf{1}))\,(\Delta(\mathbf{1})\otimes\mathbf{1})$. This establishes (3.2). The proof of (3.3) is similar.

We next consider the axioms of the antipode. We first treat (3.4) and (3.5). Using Proposition 2.5 (ii), we see that

$$\varepsilon_t(C) = \begin{cases} 0 & \text{if } C \text{ is not a horizontal identity,} \\ {}_P\mathbf{1}, & \text{if } C = \operatorname{id} g,\ P = t(g), \end{cases} \qquad \text{for all } C \in \mathcal{B}.$$

This coincides with the left hand side of (3.4), by Proposition 2.5 (iv). Similarly, using Proposition 2.5 (iii), we see that

$$\varepsilon_s(C) = \begin{cases} 0 & \text{if } C \text{ is not a horizontal identity,} \\ \mathbf{1}_P, & \text{if } C = \operatorname{id} g,\ P = b(g), \end{cases} \qquad \text{for all } C \in \mathcal{B}.$$

This coincides with the left hand side of (3.4), by Proposition 2.5 (v). Finally, the relation (3.6) is equivalent to the identity

$$\sum_{X,Y,Z} \begin{Bmatrix} X^{-1} \\ Y \\ Z^{-1} \end{Bmatrix} = A^{-1},$$

for all $A \in \mathcal{B}$, where the sum runs over all X, Y and Z in $\mathcal{B}$ such that

$$\begin{array}{c|c|c} & X^{-1} & \\ \hline X & Y & Z \\ \hline & Z^{-1} & \end{array}$$

and $XYZ = A$. By Lemmas 1.12 and 2.4, the left hand side equals A^{-1}. □

By construction, $\Bbbk\mathcal{T}$ is the groupoid algebra of the vertical groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$. Using the description in 2.3.1, this groupoid is isomorphic to $\coprod_{H\in\mathcal{H}/\sim_h} \mathcal{B}_H$, where $\sim_h$ is the equivalence relation in $\mathcal{H}$ defined by $\mathcal{B} \rightrightarrows \mathcal{H}$ and $\mathcal{B}_H \rightrightarrows H$ is the connected groupoid $\mathcal{B}(x) \times H^2$ on the class H, $x \in H$. Therefore, there is an isomorphism of algebras

(3.7) $$\Bbbk\mathcal{T} \simeq \oplus_{H\in\mathcal{H}/\sim_h} \Bbbk\mathcal{B}(x) \otimes M_{n(H)}(\Bbbk),$$

where $\Bbbk\mathcal{B}(x)$ is the group algebra, and $n(H) = |H|$. Similarly, we have an isomorphism of coalgebras

(3.8) $$\Bbbk\mathcal{T} \simeq \oplus_{V\in\mathcal{V}/\sim_v} \Bbbk\mathcal{B}(g)^* \otimes M_{m(V)}(\Bbbk)^*,$$

where $\sim_v$ is the equivalence relation defined by the horizontal groupoid $\mathcal{B} \rightrightarrows \mathcal{V}$, $\Bbbk\mathcal{B}(g)$ is the group algebra of $\mathcal{B}(g)$, $g \in V$, and $m(V) = |V|$.

Example 3.2. Suppose that $\Bbbk\mathcal{T}$ is simple as an algebra. Then

(1) $\mathcal{B} \rightrightarrows \mathcal{H}$ is the coarse groupoid on $\mathcal{H}$;
(2) $\mathcal{H} \rightrightarrows \mathcal{P}$ is a trivial group bundle;
(3) $\mathcal{V} \rightrightarrows \mathcal{P}$ is the coarse groupoid on $\mathcal{P}$;

(4) $\mathcal{B} \rightrightarrows \mathcal{V}$ is a trivial group bundle.

This means that as a weak Hopf algebra $\Bbbk\mathcal{T}$ is the groupoid algebra of the vertical groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$, and in particular it is cocommutative.

Proof. By (3.7), $\Bbbk\mathcal{T}$ is simple iff $|\mathcal{H}/\sim_h| = 1$ and $\mathcal{B}(x)$ is trivial for any $x \in \mathcal{H}$. Hence (1). If $x : P \to Q$ is in $\mathcal{H}$, then there is a box B connecting it to $\operatorname{id} P$; that is, there exists $g \in \mathcal{V}$ such that $B = x \triangleright g \begin{array}{c} \mathrm{id} \\ \boxed{} \\ x \end{array} g$ with $x = \operatorname{id} P \triangleleft g = \operatorname{id} Q$ by (2.12). Hence (2). If $P, Q \in \mathcal{P}$, there is a unique box connecting $\operatorname{id} P$ and $\operatorname{id} Q$; the vertical sides connect P and Q, hence (3). Finally, let $B := \begin{array}{c} x \\ \boxed{} \\ \end{array} g$ be any box in $\mathcal{B}$. Then $x = \operatorname{id}_P$ by (2), and $B = \mathbf{id}g$, by vacancy; hence (4). □

The proofs of the following statements are straightforward and are left to the reader.

Proposition 3.3. *Let $\mathcal{T}_1$, $\mathcal{T}_2$ be finite vacant double groupoids. Then there are isomorphisms of quantum groupoids $\Bbbk(\mathcal{T}_1 \coprod \mathcal{T}_2) \simeq \Bbbk\mathcal{T}_1 \times \Bbbk\mathcal{T}_2$, $\Bbbk(\mathcal{T}_1 \times \mathcal{T}_2) \simeq \Bbbk\mathcal{T}_1 \otimes \Bbbk\mathcal{T}_2$.* □

Proposition 3.4. *Let $\mathcal{T}$ be a finite vacant double groupoid and assume that $\Bbbk = \mathbb{C}$. Then $\mathbb{C}\mathcal{T}$ is a C^* quantum groupoid* [BNSz], *with the involution uniquely defined by $A^* = A^v$, $A \in \mathcal{B}$.* □

3.3. Extensions with cocycles.

We begin by recalling the following definition, which fits into the general framework of groupoid cohomology, due to Westman [We], see [R].

Definition 3.5. Let $s, e : \mathcal{G} \rightrightarrows \mathcal{P}$ be a groupoid. A *normalized 2-cocycle* on $\mathcal{G}$ with values in $\Bbbk^\times$ is a function $\sigma : \mathcal{G}\ {}_s\!\times_e \mathcal{G} \to \Bbbk^\times$ such that

(3.9) $\sigma(\alpha, \beta)\sigma(\alpha\beta, \gamma) = \sigma(\beta, \gamma)\sigma(\alpha, \beta\gamma)$;

(3.10) $\sigma(\alpha, \operatorname{id} e(\alpha)) = \sigma(\operatorname{id} s(\alpha), \alpha) = 1$,

for all composable $\alpha, \beta, \gamma \in \mathcal{G}$.

Let now $\mathcal{T}$ be a double groupoid. A *normalized vertical 2-cocycle* is a 2-cocycle on the groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$; similarly, a *normalized horizontal 2-cocycle* is a 2-cocycle on the groupoid $\mathcal{B} \rightrightarrows \mathcal{V}$. Thus, a normalized vertical 2-cocycle is a function σ on the set of all pairs (A, B) with $\frac{A}{B}$ with values in $\Bbbk^\times$ such that

(3.11) If $\dfrac{\dfrac{A}{B}}{C}$, then $\sigma(A,B)\sigma(\tfrac{A}{B}, C) = \sigma(B,C)\sigma(A, \tfrac{B}{C})$.

(3.12) If A or B is a vertical identity, then $\sigma(A, B) = 1$.

Letting $A = B^v = C$, we deduce that

(3.13) $\sigma(A, A^v) = \sigma(A^v, A)$.

Analogously, a normalized horizontal 2-cocycle is a function τ on the set of all pairs (A, B) with $A|B$, such that

(3.14) If $A|B|C$, then $\tau(A,B)\tau(AB,C) = \tau(B,C)\tau(A,BC)$.

(3.15) If A or B is a horizontal identity, then $\tau(A, B) = 1$.

Letting $A = B^h = C$, we deduce that

(3.16) $\tau(A, A^h) = \tau(A^h, A)$.

Definition 3.6. Let $\mathcal{T}$ be a double groupoid. A *normalized 2-cocycle* on $\mathcal{T}$ with values in $\Bbbk^\times$ is a pair (σ, τ), where σ is a normalized vertical 2-cocycle, τ is normalized horizontal 2-cocycle, and the following property holds:

(3.17)

If $\begin{array}{c|c} \mathrm{A} & \mathrm{B} \\ \hline \mathrm{C} & \mathrm{D} \end{array}$, then $\sigma(AB, CD)\tau\begin{pmatrix} \frac{A}{C}, \frac{B}{D} \end{pmatrix} = \tau(A,B)\tau(C,D)\sigma(A,C)\sigma(B,D)$.

Remark 3.7. (i). Let $x, y \in \mathcal{H}$, $x|y$ so that $\begin{array}{c|c} \mathbf{id}\, x & \mathbf{id}\, y \\ \hline \mathbf{id}\, x & \mathbf{id}\, y \end{array}$. By (3.12) and (3.17), we have

$$\tau(\mathbf{id}\, x, \mathbf{id}\, y) = 1. \tag{3.18}$$

(ii). Let $f, g \in \mathcal{V}$, $f|g$ so that $\begin{array}{c|c} \mathbf{id}\, f & \mathbf{id}\, f \\ \hline \mathbf{id}\, g & \mathbf{id}\, g \end{array}$. By (3.15) and (3.17), we have

$$\sigma(\mathbf{id}\, f, \mathbf{id}\, g) = 1. \tag{3.19}$$

Given a normalized vertical 2-cocycle σ and a normalized horizontal 2-cocycle τ on the double groupoid $\mathcal{T}$, one may consider the σ-*twisted* groupoid algebra structure and, dually, the τ-*twisted* groupoid coalgebra structure on the vector space $\Bbbk\mathcal{T}$ with basis $\mathcal{B}$. The following theorem asserts that, provided that $\mathcal{T}$ is vacant, the compatibility condition (3.17) guarantees that these two structures combine into a weak Hopf algebra structure.

Theorem 3.8. *Let $\mathcal{T}$ be a vacant double groupoid and let (σ, τ) be a normalized 2-cocycle on $\mathcal{T}$ with values in $\Bbbk^\times$.*

(i) Let $\Bbbk_\sigma^\tau \mathcal{T}$ be the vector space with basis $\mathcal{B}$ and multiplication and comultiplication defined, respectively by

- $A.B = \sigma(A,B)\frac{A}{B}$, *if* $\frac{A}{B}$, *and* 0 *otherwise.*
- $\Delta(A) = \sum \tau(B,C) B \otimes C$, *where the sum is over all pairs* (B,C) *with* $B|C$ *and* $A = BC$.

Then $\Bbbk_\sigma^\tau \mathcal{T}$ is a quantum groupoid with antipode defined by

$$\mathcal{S}(A) = \tau(A, A^h)^{-1}\, \sigma(A^{-1}, A^h)^{-1}\, A^{-1}. \tag{3.20}$$

The source and target subalgebras are, respectively, the subspaces spanned by $(\mathbf{1}_P)_{P\in\mathcal{P}}$ and $({}_P\mathbf{1})_{P\in\mathcal{P}}$; so they are commutative of dimension $|\mathcal{P}|$.

(ii) Let (ν, η) be another normalized 2-cocycle on $\mathcal{T}$ with values in $\Bbbk^\times$. Let $\psi : \mathcal{T} \to \Bbbk^\times$ be a map and let $\Psi : \Bbbk_\sigma^\tau \mathcal{T} \to \Bbbk_\nu^\eta \mathcal{T}$ be the linear map given by $\Psi(B) = \psi(B)B$, $B \in \mathcal{B}$. Then Ψ is an isomorphism of quantum groupoids if and only if

(3.21)
$$\psi\begin{pmatrix} A \\ B \end{pmatrix} \sigma(A,B) = \psi(A)\psi(B)\nu(A,B), \quad \textit{for all } A, B \in \mathcal{B} \textit{ such that } \frac{A}{B};$$

(3.22)
$$\psi(CD)\eta(C,D) = \psi(C)\psi(D)\tau(C,D), \quad \textit{for all } C, D \in \mathcal{B} \textit{ such that } C|D.$$

Proof. (i). Straightforward computations show that the multiplication is associative with unit 1 (because of the cocycle and unitary conditions on σ), and that the comultiplication is coassociative with counit ε (because of the cocycle and unitary conditions on τ). By (3.17), Δ is multiplicative. Now $\Delta(\mathbf{1}) = \sum_{x,y\in\mathcal{H}: x|y} \mathbf{id}\, x \otimes \mathbf{id}\, y$ by (3.18); we can conclude the validity of (3.2), using (3.19). The proof of (3.3) is similar. The proof of (3.4) and (3.5) is as in the proof of Theorem 3.1; (3.13) and (3.16) are needed. Note that the source and target maps coincide with those of $\Bbbk\mathcal{T}$. We next prove (3.6). Given $A \in \mathcal{B}$, we compute:

$$m^{(2)}(\mathcal{S} \otimes \mathrm{id} \otimes \mathcal{S})\Delta^{(2)}(A) =$$
$$m^{(2)}(\mathcal{S} \otimes \mathrm{id} \otimes \mathcal{S}) \left(\sum_{X|Y|Z, XYZ=A} \tau(X,Y)\tau(XY,Z) X \otimes Y \otimes Z \right) =$$
$$\sum \tau(X,Y)\tau(XY,Z)\tau(X,X^h)^{-1}\, \sigma(X^{-1},X^h)^{-1} \tau(Z,Z^h)^{-1}$$
$$\times\, \sigma(Z^{-1},Z^h)^{-1} \sigma(X^{-1},Y) \sigma(\tfrac{X^{-1}}{Y}, Z^{-1}) \begin{Bmatrix} X^{-1} \\ Y \\ Z^{-1} \end{Bmatrix}$$

$$= \tau(A, A^h)\tau(\mathrm{id}_{l(A)}, A)\tau(A, A^h)^{-1}\,\sigma(A^{-1}, A^h)^{-1}\tau(A, A^h)^{-1}\,\sigma(A^{-1}, A^h)^{-1} \times \sigma(A^{-1}, A^h)\sigma(\mathrm{id}_{b(A)^{-1}}, A^{-1})\, A^{-1}$$
$$= \tau(A, A^h)^{-1}\,\sigma(A^{-1}, A^h)^{-1}\, A^{-1} = \mathcal{S}(A),$$

where the second sum is over all X, Y, Z such that $\begin{array}{c|c|c} & X^{-1} & \\ \hline X & Y & Z \\ \hline & Z^{-1} & \end{array}$ and $XYZ = A$; the third equality is by Lemma 2.4; and the fourth is clear. The proof of (ii) is straightforward. □

Proposition 3.9. *Let $\mathcal{T}$ be a vacant double groupoid and let (σ, τ) be a normalized 2-cocycle on $\mathcal{T}$ with values in $\Bbbk^\times$. Then $\Bbbk^\tau_\sigma\mathcal{T}$ is a Hopf algebra if and only if $\mathcal{T}$ arises from a matched pair of groups.*

Proof. $\Delta(1) = 1 \otimes 1$ iff $\sum_{x,y\in\mathcal{H}:x|y} \mathbf{id}\, x \otimes \mathbf{id}\, y = \sum_{v,w\in\mathcal{H}} \mathbf{id}\, v \otimes \mathbf{id}\, w$. Thus, $\mathbf{id}\, x|\mathbf{id}\, y$ for any $x, y \in \mathcal{H}$. Hence $\#\mathcal{P} = 1$, *i. e.* $\mathcal{T}$ arises from a matched pair of groups. The converse is well-known. □

Proposition 3.10. *Let $\mathcal{T}$ be a vacant double groupoid and let (σ, τ) be a normalized 2-cocycle on $\mathcal{T}$ with values in $\Bbbk^\times$. Then $\Bbbk^\tau_\sigma\mathcal{T}$ is involutory. If char $\Bbbk = 0$, then $\Bbbk^\tau_\sigma\mathcal{T}$ is semisimple and cosemisimple.*

Proof. Let $A \in \mathcal{B}$. We compute

$$\mathcal{S}^2(A) = \tau(A, A^h)^{-1}\,\sigma(A^{-1}, A^h)^{-1}\,\mathcal{S}(A^{-1}) = \tau(A, A^h)^{-1}\,\sigma(A^{-1}, A^h)^{-1}\,\tau(A^{-1}, A^v)^{-1}\,\sigma(A, A^v)^{-1}\, A.$$

Now, $\begin{array}{c|c} A & A^h \\ \hline A^v & A^{-1} \end{array}$ implies that $1 = \tau(A, A^h)\tau(A^v, A^{-1})\sigma(A, A^v)\sigma(A^h, A^{-1})$. We conclude, using (3.13) and (3.16), that $\mathcal{S}^2(A) = A$. The second statement follows then from [N, Corollary 6.5]. □

The category of finite-dimensional representations of a weak Hopf algebra over $\Bbbk$ admits the structure of a $\Bbbk$-linear rigid monoidal category [NV]. Recall from [ENO] the definition of a *multifusion category*: this is a semisimple $\Bbbk$-linear rigid tensor category with finitely many isoclasses of simple objects and finite dimensional hom-spaces. Recall also that a multifusion category is called a *fusion* category if in addition the unit object is *simple.*

Proposition 3.11. *Let $\mathcal{T}$ be a vacant double groupoid and let (σ, τ) be a normalized 2-cocycle on $\mathcal{T}$ with values in $\Bbbk^\times$.*

(i). The unit object of the category $\operatorname{Rep}\Bbbk^\tau_\sigma\mathcal{T}$ *is simple if and only if $\mathcal{V} \rightrightarrows \mathcal{P}$ is connected.*

(ii). If char $\Bbbk = 0$, *then the category* $\operatorname{Rep}\Bbbk^\tau_\sigma\mathcal{T}$ *of finite dimensional* $\Bbbk^\tau_\sigma\mathcal{T}$-*modules is a multifusion category. It is fusion if and only if* $\mathcal{V} \rightrightarrows \mathcal{P}$ *is connected.*

Proof. We prove (i). We have already observed that the target subalgebra– which is the unit object of $\operatorname{Rep}\Bbbk^\tau_\sigma\mathcal{T}$ by general reasons– is the span of the elements ${}_P\mathbf{1}$. Let $\sim_V$ be the equivalence relation in $\mathcal{P}$ induced by $\mathcal{V} \rightrightarrows \mathcal{P}$. We claim that the subspaces $\sum_{P\in X} \Bbbk_P\mathbf{1}$, for X an equivalence class of $\sim_V$, are the simple subobjects of $\Bbbk^\tau_\sigma\mathcal{T}_t$. Indeed, for all $A \in \mathcal{B}$, we have

$$A._P\mathbf{1} = \epsilon_t(A\,{}_P\mathbf{1}) = \begin{cases} {}_Q\mathbf{1}, & \text{if } A = \operatorname{id} g, \text{ for } g \in \mathcal{V} : b(g) = P,\, t(g) = Q, \\ 0, & \text{otherwise.} \end{cases}$$

The claim is proved. Now (ii) follows from 3.10, general results on weak Hopf algebras and (i). □

Proposition 3.12. *Let* $\mathcal{T}$ *be a vacant double groupoid and let* (σ, τ) *be a normalized 2-cocycle on* $\mathcal{T}$ *with values in* $\Bbbk^\times$. *Then* (τ, σ) *is a normalized 2-cocycle on the transpose double groupoid* $\mathcal{T}^t$ *and the quantum groupoid* $\Bbbk^\sigma_\tau\mathcal{T}^t$ *is dual to* $\Bbbk^\tau_\sigma\mathcal{T}$.

Proof. The duality is given by the bilinear form $(B|C) = \delta_{B,C^t}$, B, $C \in \mathcal{B}$. □

3.4. The category $\operatorname{Rep}\Bbbk\mathcal{T}$.

Let $\mathcal{T}$ be a finite vacant double groupoid, and let $\Bbbk\mathcal{T}$ be the quantum groupoid associated to $\mathcal{T}$ as in Theorem 3.1. Consider the category $\mathcal{C} := \operatorname{Rep}\Bbbk\mathcal{T}$ of finite dimensional representations of $\Bbbk\mathcal{T}$. Our aim in this subsection is to sketch a combinatorial description of the category $\operatorname{Rep}\Bbbk\mathcal{T}$ in groupoid-theoretical terms. We shall follow the lines in [NV, 5.1].

Suppose that $s, e : \mathcal{G} \rightrightarrows \mathcal{P}$ is a groupoid. A $\Bbbk$-*linear* $\mathcal{G}$-*bundle*, or $\mathcal{G}$-bundle for short, is a map $p : V \to \mathcal{P}$ together with an action of $\mathcal{G}$ on p, and such that
(i) each fiber V_b ($b \in \mathcal{P}$) is a vector space over $\Bbbk$;
(ii) for all $g \in \mathcal{G}$ the map $g : V_{e(g)} \to V_{s(g)}$ is a linear isomorphism.

So one may think of a $\mathcal{G}$-bundle as a $\mathcal{P}$-graded vector space $V = \oplus_{b\in\mathcal{P}} V_b$ endowed with a linear $\mathcal{G}$-action $g : V_{e(g)} \to V_{s(g)}$, $g \in \mathcal{G}$.

Remark 3.13. The category $\mathcal{G}$-bund of (finite dimensional) $\Bbbk$-linear $\mathcal{G}$-bundles is equivalent to the category of (finite dimensional) representations of the groupoid algebra $\Bbbk\mathcal{G}$. The equivalence is defined as follows: for a $\Bbbk\mathcal{G}$-module V, we let the $\mathcal{P}$-grading on V be given by $V_b = \operatorname{id} b.V$, for all $b \in \mathcal{P}$.

Let now $\mathcal{T}$ be a vacant double groupoid. Let $\mathcal{T}$-bund be the category of $\Bbbk$-linear bundles over the vertical groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$. Thus, the objects of $\mathcal{T}$-bund are $\mathcal{H}$-graded vector spaces endowed with a left action of the vertical groupoid $\mathcal{B} \rightrightarrows \mathcal{H}$ by linear isomorphisms. There is a structure of rigid monoidal category on $\mathcal{T}$-bund:

- *Tensor product.* If V, W are $\mathcal{T}$-bundles then $V \otimes W := \oplus_{z \in \mathcal{H}} (V \otimes W)_z$, where

$$(V \otimes W)_z = \sum_{xy=z} V_x \otimes_{\Bbbk} W_y, \qquad z \in \mathcal{H}.$$

(Note that this difers from $V \otimes_{\Bbbk} W$ by the fact that we are not taking all summands $V_x \otimes_{\Bbbk} W_y$ but only those for which x and y are composable.) The action of $\mathcal{B}$ on $V \otimes W$ is given by Δ.

- *Unit object.* This is the target subalgebra $\Bbbk\mathcal{T}_t = \oplus_{P \in \mathcal{P}} \Bbbk_P \mathbf{1}$, with $\mathcal{H}$-grading defined by

$$(\Bbbk\mathcal{T}_t)_x = \begin{cases} 0, & \text{if } x \text{ is not an identity,} \\ \Bbbk_P \mathbf{1}, & \text{if } x = \operatorname{id} P, \end{cases} \qquad \text{for all } x \in \mathcal{H}.$$

  and $\mathcal{B}$-action $A._P\mathbf{1} = \epsilon_t(A\, {}_P\mathbf{1})$.

- The *dual* V^* of an object $V = \oplus_{x \in \mathcal{H}} V_x \in \mathcal{C}$ has $\mathcal{H}$-grading $(V^*)_x = (V_{x^{-1}})^*$, $x \in \mathcal{H}$; $\mathcal{B}$-action $A := (A^{-1})^* : (V^*)_{b(A)} \to (V^*)_{t(A)}$, for all $A \in \mathcal{B}$.

With remark 3.13 in mind, we can describe the monoidal structure in $\operatorname{Rep} \Bbbk\mathcal{T}$.

Proposition 3.14. *Let $\mathcal{T}$ be a vacant double groupoid. Assume that char $\Bbbk = 0$. The category $\mathcal{T}$-bund is a multifusion category over $\Bbbk$ and it is monoidally equivalent to* $\operatorname{Rep} \Bbbk\mathcal{T}$.

Proof. The expressions for the tensor product, unit object and duals are a translation of the formulas in [NV] to the language of $\mathcal{T}$-bundles. For instance, the unit isomorphism $\Bbbk\mathcal{T}_t \otimes V \to V$ is given as follows: for any $z \in \mathcal{H}$, we have $(\Bbbk\mathcal{T}_t \otimes_{\Bbbk} V)_z = \Bbbk_{l(z)}\mathbf{1} \otimes_{\Bbbk} V_z$; the isomorphism $\Bbbk\mathcal{T}_t \otimes V \to V$ is determined by its homogeneous components $(\Bbbk\mathcal{T}_t \otimes_{\Bbbk} V)_z \to V_z$, given by the action of ${}_{l(z)}\mathbf{1}$, which is the identity on V_z. The unit isomorphisms on the right and the evaluation and coevaluation maps for the duals are translated similarly from $\operatorname{Rep} \Bbbk\mathcal{T}$. □

3.5. A Kac exact sequence for matched pairs of groupoids.

We first recall the well-known definition of the groupoid cohomology via standard resolutions [We, R]. Let $s, e : \mathcal{G} \rightrightarrows \mathcal{P}$ be a groupoid. In this subsection, we shall denote by $\mathcal{G}^{(0)} := \mathcal{P}$ the base of $\mathcal{G}$, $\mathcal{G}^{(1)} := \mathcal{G}$ and

$$\mathcal{G}^{(n)} = \{(x_1, \dots, x_n) \in \mathcal{G}^n : x_1 | x_2 | \dots | x_{n-1} | x_n\}, \qquad n \geq 2.$$

Let M be an abelian group and let

$$C^n(\mathcal{G}, M) = \{f : \mathcal{G}^{(n)} \to M : f(x_1, \dots, x_n) = 0, \text{ if some } x_i \in \mathcal{G}^{(0)}\}.$$

The cohomology groups $H^n(\mathcal{G}, M)$ of $\mathcal{G}$ with coefficients in M are the cohomology groups of the complex

$$(3.23) \quad 0 \longrightarrow C^0(\mathcal{G}, M) \xrightarrow{d^0} C^1(\mathcal{G}, M) \xrightarrow{d^1} C^2(\mathcal{G}, M) \xrightarrow{d^2} \dots \\ \longrightarrow C^n(\mathcal{G}, M) \xrightarrow{d^n} C^{n+1}(\mathcal{G}, M) \longrightarrow \dots$$

where
(3.24)

$$\begin{aligned} d^0 f(x) &= f(e(x)) - f(s(x)), \\ d^n f(x_1, \dots, x_{n+1}) &= f(x_2, \dots, x_{n+1}) + \sum_{1 \leq i \leq n} (-1)^i f(x_1, \dots, x_i x_{i+1}, \dots, x_{n+1}) \\ &\quad + (-1)^{n+1} f(x_1, \dots, x_n). \end{aligned}$$

Let now $\mathcal{T}$ be a double groupoid: $\begin{matrix} \mathcal{B} & \rightrightarrows & \mathcal{H} \\ \downdownarrows & & \downdownarrows \\ \mathcal{V} & \rightrightarrows & \mathcal{P} \end{matrix}$. Let

$$\begin{aligned} \mathcal{B}^{(0,0)} &:= \mathcal{P}, \\ \mathcal{B}^{(0,s)} &:= \{(x_1, \dots, x_s) \in \mathcal{H}^s : x_1 | x_2 \dots | x_s\} = \mathcal{H}^{(s)}, \quad s > 0, \\ \mathcal{B}^{(r,0)} &:= \{(g_1, \dots, g_r) \in \mathcal{V}^r : g_1 | g_2 \dots | g_r\} = \mathcal{V}^{(r)}, \quad r > 0, \\ \mathcal{B}^{(r,s)} &:= \left\{ \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1s} \\ A_{21} & A_{22} & \dots & A_{2s} \\ \dots & \dots & \dots & \dots \\ A_{r1} & A_{r2} & \dots & A_{rs} \end{pmatrix} \in \mathcal{B}^{r \times s} : \begin{array}{c|c|c|c} A_{11} & A_{12} & \dots & A_{1s} \\ \hline A_{21} & A_{22} & \dots & A_{2s} \\ \hline \dots & \dots & \dots & \dots \\ \hline A_{r1} & A_{r2} & \dots & A_{rs} \end{array} \right\}, \end{aligned}$$

$r, s > 0$.

Let M be an abelian group and let $D^{r,s} = D^{r,s}(\mathcal{T}, M)$, $r, s \geq 0$, be defined by

$$D^{r,s} := \left\{ f : \mathcal{B}^{(r,s)} \to M : f \begin{pmatrix} A_{11} & \dots & A_{1s} \\ A_{21} & \dots & A_{2s} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{rs} \end{pmatrix} = 0, \text{ if} \begin{cases} r > 1, s > 0, A_{ij} \in \mathcal{V}, \\ \text{or } r > 1, s = 0, A_{i0} \in \mathcal{P}, \\ \text{or } r > 0, s > 1, A_{ij} \in \mathcal{H}, \\ \text{or } r = 0, s > 1, A_{0j} \in \mathcal{P}. \end{cases} \right\}$$

Let $d_H = d_H^{r,s} : D^{r,s} \to D^{r,s+1}$, $d_V = d_V^{r,s} : D^{r,s} \to D^{r+1,s}$ be, respectively, the horizontal and vertical coboundary maps defined as follows:

- If $r = 0$, d_H is as in (3.24);
- if $s = 0$, d_V is as in (3.24);
- if $r = 0$, $s > 0$,

$$d_V^{0,s} f(A_{11}, \dots, A_{1s}) = f(b(A_{11}), \dots, b(A_{1s})) - f(t(A_{11}), \dots, t(A_{1s}));$$

- if $r > 0$, $s = 0$,

$$d_H^{r,0} f \begin{pmatrix} A_{11} \\ \dots \\ A_{r1} \end{pmatrix} = f(r(A_{11}), \dots, r(A_{r1})) - f(l(A_{11}), \dots, l(A_{r1}));$$

- if $r > 0$ and $s > 0$,

$$d_V^{r,s} f \begin{pmatrix} A_{11} & \dots & A_{1s} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{rs} \\ A_{r+1,1} & \dots & A_{r+1,s} \end{pmatrix} = f \begin{pmatrix} A_{21} & \dots & A_{2s} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{rs} \\ A_{r+1,1} & \dots & A_{r+1,s} \end{pmatrix}$$

$$+ \sum_{1 \le i \le r} (-1)^i f \begin{pmatrix} A_{11} & \dots & A_{1s} \\ \dots & \dots & \dots \\ \begin{Bmatrix} A_{i1} \\ A_{i+1,1} \end{Bmatrix} & \dots & \begin{Bmatrix} A_{is} \\ A_{i+1,s} \end{Bmatrix} \\ \dots & \dots & \dots \\ A_{r+1,1} & \dots & A_{r+1,s} \end{pmatrix} + (-1)^{r+1} f \begin{pmatrix} A_{11} & \dots & A_{1s} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{rs} \end{pmatrix};$$

$$d_H^{r,s} f \begin{pmatrix} A_{11} & \dots & A_{1,s+1} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{r,s+1} \end{pmatrix} = f \begin{pmatrix} A_{12} & \dots & A_{1,s+1} \\ \dots & \dots & \dots \\ A_{r2} & \dots & A_{r,s+1} \end{pmatrix}$$

$$+ \sum_{1 \le j \le s} (-1)^j f \begin{pmatrix} A_{11} & \dots & \{A_{1,j} A_{1,j+1}\} & \dots & A_{1,s+1} \\ \dots & \dots & \dots & \dots & \dots \\ A_{r,1} & \dots & \{A_{r,j} A_{1,j+1}\} & \dots & A_{r,s+1} \end{pmatrix}$$

$$+ (-1)^{s+1} f \begin{pmatrix} A_{11} & \dots & A_{1s} \\ \dots & \dots & \dots \\ A_{r1} & \dots & A_{rs} \end{pmatrix}.$$

A straightforward computation shows that the following diagram commutes:

$$\begin{array}{ccc} D^{r+1,s} & \xrightarrow{d_H} & D^{r+1,s+1} \\ \uparrow{\scriptstyle d_V} & & \uparrow{\scriptstyle d_V} \\ D^{r,s} & \xrightarrow{d_H} & D^{r,s+1} \end{array}$$

Thus, there is a double cochain complex

$$D^{\cdot\cdot} = \begin{array}{ccccccc} \vdots & & & & & & \\ \uparrow & & & & & & \\ D^{2,0} & \xrightarrow{d_H} & \vdots\ \cdots & & & & \\ \uparrow{\scriptstyle d_V} & & \uparrow{\scriptstyle -d_V} & & & & \\ D^{1,0} & \xrightarrow{d_H} & D^{1,1} & \xrightarrow{d_H} & \vdots\ \cdots & & \\ \uparrow{\scriptstyle d_V} & & \uparrow{\scriptstyle -d_V} & & \uparrow & & \\ D^{0,0} & \xrightarrow{d_H} & D^{0,1} & \xrightarrow{d_H} & D^{0,2} & \longrightarrow & \cdots \end{array},$$

with the usual "sign trick": the vertical differential is $(-1)^s d_V^{r,s}$. We then remove the edges of this double complex setting $A^{r,s}(\mathcal{T}, M) = A^{r,s} := D^{r+1,s+1}$, $r, s \geq 0$; and denote by $E^{\cdot\cdot}(\mathcal{T}, M) = E^{\cdot\cdot}$ the double complex consisting only of the edges of $D^{\cdot\cdot}$. Compare with [M, pp. 173 ff.].

We are now ready to state a result inspired in the celebrated Kac exact sequence [K, (3.14)]. Let $\mathcal{T}$ be a *vacant* double groupoid and let $\mathcal{D} = \mathcal{V} \bowtie \mathcal{H}$ be corresponding diagonal groupoid, see Prop. 2.9.

Proposition 3.15. *There is an exact sequence*

$$\begin{aligned} &0 \to H^1(\mathcal{D}, M) \to H^1(\mathcal{H}, M) \oplus H^1(\mathcal{V}, M) \to H^0(\operatorname{Tot} A(\mathcal{T}, M)^{\cdot\cdot}, M) \\ &\to H^2(\mathcal{D}, M) \to H^2(\mathcal{H}, M) \oplus H^2(\mathcal{V}, M) \to H^1(\operatorname{Tot} A(\mathcal{T}, M)^{\cdot\cdot}, M) \\ &\to H^3(\mathcal{D}, M) \to H^3(\mathcal{H}, M) \oplus H^3(\mathcal{V}, M). \end{aligned} \tag{3.25}$$

Proof. The short exact sequence of double complexes $0 \to A^{\cdot\cdot} \to D^{\cdot\cdot} \to E^{\cdot\cdot} \to 0$ (with $A^{\cdot\cdot}$ "shifted") induces a long exact sequence in cohomology. It is clear that $H^n(\operatorname{Tot} E^{\cdot\cdot}(\mathcal{T}, M)) = H^n(\mathcal{H}, M) \oplus H^n(\mathcal{V}, M)$, for $n > 0$. We claim that

$$H^n(\operatorname{Tot} D^{\cdot\cdot}(\mathcal{T}, M))) = H^n(\mathcal{D}, M), \quad n > 0. \tag{3.26}$$

Indeed, $H^{\cdot}(\mathcal{D}, M)$ are the cohomology groups of a complex $\operatorname{Hom}_{\Bbbk\mathcal{D}}(F^{\cdot}, M)$, where $F^{\cdot}$ is some free resolution of the trivial $\mathcal{D}$-module. Now, arguing as in [M, Lemma 1.7], we see that $H^{\cdot}(\mathrm{Tot}\, D^{\cdot\cdot}(\mathcal{T}, M)))$ are also the cohomology groups of a complex $\mathrm{Hom}_{\Bbbk\mathcal{D}}(G^{\cdot}, M)$, where $G^{\cdot}$ is another free resolution of the trivial $\mathcal{D}$-module; this implies (3.26). □

If $M = \Bbbk^{\times}$, it is natural to denote

$$\mathrm{Aut}(\Bbbk\mathcal{T}) = H^0(\mathrm{Tot}\, A^{\cdot\cdot}(\mathcal{T}, \Bbbk^{\times})), \tag{3.27}$$

$$\mathrm{Opext}(\Bbbk^{\mathcal{V}}, \Bbbk\mathcal{H}) = H^1(\mathrm{Tot}\, A^{\cdot\cdot}(\mathcal{T}, \Bbbk^{\times})), \tag{3.28}$$

by Theorem 3.8 and in view of an extension theory of quantum groupoids yet to be explored. Then (3.25) has in this case the familiar expression

$$\begin{aligned} 0 &\to H^1(\mathcal{D}, \Bbbk^{\times}) \to H^1(\mathcal{H}, \Bbbk^{\times}) \oplus H^1(\mathcal{V}, \Bbbk^{\times}) \to \mathrm{Aut}(\Bbbk\mathcal{T}) \\ &\to H^2(\mathcal{D}, \Bbbk^{\times}) \to H^2(\mathcal{H}, \Bbbk^{\times}) \oplus H^2(\mathcal{V}, \Bbbk^{\times}) \to \mathrm{Opext}(\Bbbk\mathcal{T}) \\ &\to H^3(\mathcal{D}, \Bbbk^{\times}) \to H^3(\mathcal{H}, \Bbbk^{\times}) \oplus H^3(\mathcal{V}, \Bbbk^{\times}). \end{aligned} \tag{3.29}$$

3.6. **Conclusion.**

We have introduced families of quantum groupoids and *a fortiori* of tensor categories. To be sure that these tensor categories are really new, we have to explicitly compute first the $\mathrm{Opext}(\Bbbk\mathcal{T})$ groups, and second to analyze when the corresponding quantum groupoids give rise to equivalent tensor categories. We shall address both questions in subsequent work.

References

[AA] M. Aguiar and N. Andruskiewitsch, *Representations of matched pairs of groupoids and applications to weak Hopf algebras*, Contemp. Math. (to appear), `math.QA/0402118` (2004).

[A] N. Andruskiewitsch, *On the quiver-theoretical quantum Yang-Baxter equation*, `math.QA/0402269` (2004).

[AM] N. Andruskiewitsch and M. Mombelli, *Examples of weak Hopf algebras arising from vacant double groupoids*, `math.QA/0405374` (2004).

[AN] N. Andruskiewitsch and S. Natale, *Braided Hopf algebras arising from matched pairs of groups*, J. Pure Appl. Alg. **182**, 119–149 (2003).

[BSV] S. Baaj, G. Skandalis and S. Vaes, *Measurable Kac cohomology for Bicrossed Products*, preprint `math.OA/0307172`.

[BNSz] G. Böhm, F. Nill and K. Szlachányi, *Weak Hopf algebras I. Integral theory and C^*-structure*, J. Algebra **221**, 385–438 (1999).

[BSz] G. Böhm and K. Szlachányi, *A coassociative C^*-quantum group with nonintegral dimensions*, Lett. in Math. Phys. **35**, 437–456 (1996).

[BS] R. Brown and C. Spencer, *Double groupoids and crossed modules*, Cahiers Topo. et Géo. Diff. **XVII**, 343–364 (1976).

[DVVV] R. Dijkgraaf, C. Vafa, E. Verlinde and H. Verlinde, *The operator algebra of orbifold models*, Commun. Math. Phys. **123**, 485–526 (1989).

[E] C. Ehresmann, *Catégories doubles et catégories structures*, C. R. Acad. Sci. Paris **256**, 1198–1201 (1963).

[ENO] P. Etingof, D. Nikshych and V. Ostrik, *On fusion categories*, preprint math.QA/0203060 (2002).

[H] T. Hayashi, *A brief introduction to face algebras*, in "New trends in Hopf Algebra Theory"; Contemp. Math. **267** (2000), 161–176.

[K] G. Kac, *Extensions of groups to ring groups*, Math. USSR Sbornik **5**, 451–474 (1968).

[KL] T. Kerler and V. Lyubashenko, *Non-Semisimple Topological Quantum Field Theories for 3-Manifolds with Corners*, Lecture Notes in Math. **1765**, Springer-Verlag, Berlin (2001).

[LYZ1] Jiang-Hua Lu, Min Yan and Yong-Chang Zhu, *On Hopf algebras with positive bases*, J. Algebra **237**, 421–445 (2001).

[LYZ2] Jiang-Hua Lu, Min Yan and Yong-Chang Zhu, *Quasi-triangular structures on Hopf algebras with positive bases*, in "New trends in Hopf Algebra Theory"; Contemp. Math. **267**, 339–356 (2000).

[Ma] K. Mackenzie, *Double Lie algebroides and Second-order Geometry, I*, Adv. Math. **94**, 180–239 (1992).

[Mj1] S. Majid, *Physics for algebraists: Non-commutative and non-cocommutative Hopf algebras by a bicrossproduct construction*, J. Algebra **130**, 17–64 (1990).

[Mj2] S. Majid, *Foundations of quantum group theory*, Cambridge Univ. Press, Cambridge (1995).

[M] A. Masuoka, *Hopf algebra extensions and cohomology*, Math. Sci. Res. Inst. Publ. **43**, 167–209 (2002).

[N] D. Nikshych, *On the structure of weak Hopf algebras*, Adv. Math. **170**, 257–286 (2002).

[NV] D. Nikshych and L. Vainerman, *Finite quantum groupoids and their applications*, Math. Sci. Res. Inst. Publ. **43**, 211–262 (2002).

[R] J. Renault, *A groupoid approach to C^*-algebras*, Lecture Notes in Math. **793**, Springer-Verlag, Berlin (1980).

[T1] M. Takeuchi,*Matched pairs of groups and bismash products of Hopf algebras*, Commun. Alg. **9**, 841–882 (1981).

[T2] M. Takeuchi, *Survey on matched pairs of groups. An elementary approach to the ESS-LYZ theory*, preprint (2001); Banach Center Publ., to appear.

[T3] M. Takeuchi, *Survey of braided Hopf algebras*, Contemp. Math. **267**, 301–324 (2000).

[We] J. Westman, *Groupoid theory in algebra, topology and analysis*, preprint (1971); Univ. of California at Irvine.

Nicolás Andruskiewitsch and Sonia Natale
andrus@mate.uncor.edu, http://www.mate.uncor.edu/andrus
natale@mate.uncor.edu, http://www.mate.uncor.edu/natale
Facultad de Matemática, Astronomía y Física
Universidad Nacional de Córdoba
CIEM-CONICET
(5000) Ciudad Universitaria. Córdoba, Argentina.

A note on aisles in a triangulated Krull Schmidt category.

Ibrahim Assem, María José Souto Salorio and Rita Zuazua[1]

ABSTRACT

We assume that $\mathcal{T}$ is a triangulated Hom-finite Krull-Schmidt k-category and that M is a strong generator such that $Hom_{\mathcal{T}}(M, M[j]) = 0$, for all $j \neq 0$. We show that the suspended subcategory $\mathcal{U}_M$ generated by M is an aisle. Further, if $\mathcal{T}$ has almost split triangles then the orthogonal $\mathcal{U}_M^{\perp}$ equals the co-aisle ${}_{\tau M}\mathcal{U}$ cogenerated by the Auslander-Reiten translate τM of M.

RESUMEN

Sea $\mathcal{T}$ una k-categoría triangulada, Hom-finita, Krull-Schmidt y M un generator fuerte, tal que $Hom_{\mathcal{T}}(M, M[j]) = 0$, para todo $j \neq 0$. Probamos que la subcategoría suspendida $\mathcal{U}_M$, generada por M es un "aisle". Además si $\mathcal{T}$ tiene triángulos casi escindidos, entonces el ortogonal $\mathcal{U}_M^{\perp}$ coincide con el "co-aisle" ${}_{\tau M}\mathcal{U}$ cogenerado por el trasladado de Auslander-Reiten τM de M.

Mathematics Subject Classification (1991): 18E30.

Introduction

The notion of triangulated category (see [V]) has proved very useful in the representation theory of algebras. In particular, there is a strong relationship between the study of t-structures and tilting theory (see, for instance, [KV, P, H, ST]). In [KV](1.1), Keller and Vossieck consider certain subcategories called aisles, and show that, if $\mathcal{U}$ is an aisle, then $(\mathcal{U}_M, \mathcal{U}_M^{\perp}[1])$ is a t-structure, and conversely any t-structure is of this form.

In this note, we give a construction procedure for aisles and hence for t-structures. We recall that, for instance, it was shown in [ST] that every perfect complex generates a t-structure on $\mathbf{D}^b(mod-A)$, where A is a Noether algebra (see also [KV](5.1)).

[1]The first and the last authors gratefully acknowledge support from the NSERC of Canada and the CONACYT of Mexico (proj. J38611-E), respectively.

We say that an object M in a triangulated category is a strong generator if $\mathcal{T}$ equals the smallest triangulated subcategory containing M and closed under direct summands. We prove the following theorem.

Theorem: Let k be a field, $\mathcal{T}$ be a triangulated Hom-finite Krull-Schmidt k-category and M be a strong generator such that $Hom_{\mathcal{T}}(M, M[j]) = 0$, for all $j \neq 0$. Then the suspended subcategory $\mathcal{U}_M$ generated by M is an aisle in $\mathcal{T}$. Dually, the cosuspended subcategory ${}_M\mathcal{U}$ cogenerated by M is a co-aisle in $\mathcal{T}$.

We next consider the case where $\mathcal{T}$ has almost split triangles. A necessary and sufficient condition for the existence of such triangles is given in [RV]. We denote by τ the Auslander-Reiten translation in $\mathcal{T}$.

Corollary: Let $\mathcal{T}$ and M be as in the theorem, and assume that $\mathcal{T}$ has almost split triangles. Then $(\mathcal{U}_M)^{\perp} =_{\tau M} \mathcal{U}$ and $\mathcal{U}_M =^{\perp} ({}_{\tau M}\mathcal{U})$.

1. The theorem.

1.1 Following [KV], we say that a full subcategory $\mathcal{U}$ of a triangulated category $\mathcal{T}$ is a **suspended subcategory** if $\mathcal{U}[1] \subset \mathcal{U}$, and it is closed under extensions (that is, if $X \to Y \to Z \to X[1]$ is a triangle in $\mathcal{T}$ and $X, Z \in \mathcal{U}$, then Y belongs to $\mathcal{U}$).

A suspended subcategory $\mathcal{U}$ is called an **aisle** in $\mathcal{T}$ if the inclusion functor $\mathcal{U} \to \mathcal{T}$ has a right adjoint functor $t_{\mathcal{U}} : \mathcal{T} \to \mathcal{U}$ (see [KV](1.1)). We define dually **co-suspended** subcategories and **co-aisles**.

Given an object M in $\mathcal{T}$, we denote by $\mathcal{U}_M$ (or ${}_M\mathcal{U}$) the smallest suspended (or cosuspended, respectively) subcategory of $\mathcal{T}$ containing M.

1.2 Let M be an object in a triangulated category $\mathcal{T}$. We define a sequence of classes of objects $(\mathcal{E}_i)_{i\geq 0}$ of $\mathcal{T}$ as follows. Let $\mathcal{E}_0 = add(\oplus_{i\in\mathbb{Z}} M\,[i])$ consist of all the summands of finite sums of copies of translates of M. Assume that $i \geq 1$, and that $\mathcal{E}_0, \mathcal{E}_1, \cdots, \mathcal{E}_{i-1}$ are already known. The class $\mathcal{E}_i$ consists of all the objects X which are direct summands of objects X' such that there is a triangle $X_0 \to X' \to X_{i-1} \to X_0[1]$, where X_0 lies in $\mathcal{E}_0$ and X_{i-1} lies in $\mathcal{E}_{i-1}$. It is not hard to show that $\cup_{i\geq 0}\mathcal{E}_i$ is the smallest triangulated subcategory of $\mathcal{T}$ closed under direct summands and containing M.

We say that M is a **strong generator** of $\mathcal{T}$ if $\mathcal{T} = \cup_{i\geq 0}\mathcal{E}_i$.

For instance, let $\mathcal{T}'$ be a triangulated compactly generated category and M be a compact generator of $\mathcal{T}'$, then M is a strong generator of the full subcategory consisting of the compact objects (see [N]).

1.3 **Proof of our theorem:** We first prove the following claim: for every

$X \in \mathcal{T}$, there is a finite set $F_X \subset \mathbb{Z}$ such that $Hom_{\mathcal{T}}(M[l], X) = 0$, for all $l \notin F_X$. Indeed, assume $X = \oplus_{i \in F} M_i^{(F_i)} \in \mathcal{E}_0$ where $M_i \simeq M[i]$ with F_i and F finite subsets of $\mathbb{Z}$. Then,

$$Hom_{\mathcal{T}}(M[t], X) = Hom_{\mathcal{T}}(M[t], \oplus_{i \in F} M_i^{(F_i)}) = 0]$$

if $t \notin F$, using the hypothesis on M. This shows our claim for X (and hence for its direct summands). Assume now $j > 0$ and $X \in \mathcal{E}_j$. Then there is a triangle

$$X_0 \xrightarrow{r} X \xrightarrow{s} X_{j-1} \to X_0[1]$$

with $X_0 \in \mathcal{E}_0$ and $X_{j-1} \in \mathcal{E}_{j-1}$.

Let $t \in \mathbb{Z}$ and $f \in Hom_{\mathcal{T}}(M\,[t]\,, X)$. If $sf = 0$, there exists a morphism $h : M[t] \to X_0$ such that $rh = f$. But we know that there are only finitely many indices l such that $Hom_{\mathcal{T}}(M[l], X_0) \neq 0$. On the other hand, the induction hypothesis says that we have only finitely many indices l such that $sf \in Hom_{\mathcal{T}}(M[l], X_{j-1})$ is non-zero. Therefore, the set $\{l \in \mathbb{Z} \mid Hom_{\mathcal{T}}(M[l], X_j) \neq 0\}$ is finite. If now Y is a direct summand of X as before, then $Hom_{\mathcal{T}}(M[l], Y)$ is a direct summand of $Hom_{\mathcal{T}}(M[l], X)$. This establishes our claim.

For an $l \in \mathbb{Z}$, let $S_l = dim_k Hom_{\mathcal{T}}(M[l], X)$ and $U_0 = \oplus_{l \in F_X \cap \mathbb{N}} M[l]^{S_l}$. Then the induced morphism

$$Hom_{\mathcal{T}}(-, U_0)_{|_{\mathcal{U}_M}} \to Hom_{\mathcal{T}}(-, X)_{|_{\mathcal{U}_M}}$$

is an epimorphism. Applying [KV](1.3) yields that $\mathcal{U}_M$ is an aisle in $\mathcal{T}$.

The second statement is proved dually. □

2. The corollary

2.1 We now assume that the triangulated category $\mathcal{T}$ has almost split triangles, or equivalently, that there is a triangulated equivalence $\tau : \mathcal{T} \to \mathcal{T}$ and an isomorphism, called the Auslander-Reiten formula,

$$\beta_{X,Y} : DHom_{\mathcal{T}}(X, Y\,[1]) \to Hom_{\mathcal{T}}(Y, \tau X),$$

functorial in both variables, X, Y in $\mathcal{T}$ (see [RV] for details.).

An example of such a situation is the case of $D^b(mod\,A)$, the derived category of bounded complexes of finitely generated (right) $A-$modules, where A is a finite dimensional $k-$algebra with finite global dimension.

For a full subcategory $\mathcal{U}$ of $\mathcal{T}$, we denote by $\mathcal{U}^{\perp}$ (or ${}^{\perp}\mathcal{U}$) the full subcategory consisting of the objects $X \in \mathcal{T}$ such that $Hom_{\mathcal{T}}(-, X)_{|_{\mathcal{U}}} = 0$ (or $Hom_{\mathcal{T}}(X, -)_{|_{\mathcal{U}}} = 0$, respectively).

2.2 The following lemma seems to be well-known. We provide its proof for the convenience of the reader. **Lemma:** The aisle $\mathcal{U}_M$ coincides with the full subcategory consisting of the objects X such that $Hom_{\mathcal{T}}(M[i], X) = 0$ for all $i < 0$. **Proof:** Let $\mathcal{S}$ be the full subcategory of $\mathcal{T}$ consisting of the objects X verifying the condition of the statement. Then $\mathcal{S}$ is closed under extensions, direct summands, positive translations and M lies in $\mathcal{S}$. Hence $\mathcal{U}_M \subseteq \mathcal{S}$.

Let $X \in \mathcal{S}$, and consider the triangle $N \to X \to B \to N[1]$ given by the definition of aisle. Applying the cohomological functor $Hom_{\mathcal{T}}(M, -)$ to the above triangle, yields $Hom_{\mathcal{T}}(M[j], B) = 0$ for all $j < 0$, because $N, X \in \mathcal{S}$. However, $Hom_{\mathcal{T}}(M[j], B) = 0$ for all $j \geq 0$, because $B \in \mathcal{U}_M{}^{\perp}$. Since M is a strong generator, $B = 0$. Hence $X \simeq N$ lies in $\mathcal{U}_M$. □

1.3 **Proof of our corollary:** Applying the above Lemma and the Auslander-Reiten formula, we get that X belongs to $\mathcal{U}_M$ if and only if, for all $j \leq 0$, $Hom_{\mathcal{T}}(X, (\tau M)[j]) \cong Hom_{\mathcal{T}}(M[j], X[1]) = 0$. This means, if and only if $X \in {}^{\perp}({}_{\tau M}\mathcal{U})$ (see [ST] (2.3)). The second statement is obtained dually. □

References

[ABM] Assem I., Beligiannis, A. and Marmaridis, N.: Right triangulated categories with right semi-equivalences, *Can. Math. Soc. Conf. Proc.*, **24** (1998), pp. 17–37.

[H] Happel, D.: Triangulated categories in the representation theory of finite-dimensional algebras. *London Mathematical Society Lecture Notes,* **119** (1988).

[KV] Keller, B. and Vossieck, D.: Aisles in derived categories, *Bull. Soc. Math. Belg.* Sér. A **40** (1988), no. 2, pp. 239–253.

[P] Parthasarathy, R.: t-structures dans la catégorie dérivée associée aux représentations d' un carquois, *C. R. Acad. Sci.* Paris 304, (1987), pp.355-357.

[N] Neeman, A.: Triangulated categories. *Annals of Mathematical Studies, Princeton University Press* (2001).

[RV] Reiten, I. and Van den Bergh, M., Noetherian hereditary abelian categories satisfying Serre duality, *J. Amer. Math. Soc.* **15,** no. 2 (2002) pp. 295-366.

[ST] Souto, M.J.; Trepode, S.: On t-structures and tilting theory, *Communications in Algebra.* vol. **31**, n 12, (2003), 6093-6114.

IBRAHIM ASSEM
ibrahim.assem@usherbrooke.ca
Univ. de Sherbrooke, Québec, Canada.

MARÍA JOSÉ SOUTO SALORIO
mariaj@udc.es
Univ. da Coruña. C. Elviña,
15071. A Coruña, España.

RITA ZUAZUA
zuazua@matmor.unam.mx
Instituto de Matemáticas.
Morelia, México.

On the Fundamental Derivations of a Finite Dimensional Algebra.

Sônia María Fernandes[1] and Eduardo do Nascimento Marcos[2]

ABSTRACT

In this paper we study the Lie structure of the vector space generated by the so called fundamental derivations. We show a relation between fundamental derivations and Galois coverings.

RESUMEN

En este trabajo estudiamos la estructura de Lie del espacio vectorial generado por las derivaciones fundamentales. Mostramos la relación entre las derivaciones fundamentales y los cubrimientos de Galois

Dedicated to Prof. Alfredo Jones on the occasion of his retirement.

Introduction In this paper, k always denotes a field and all algebras are of the form $\Lambda = k\Gamma/I$, where Γ is a finite quiver and I an admissible ideal. $Der(\Lambda)$ is the space of Λ derivations . We write $SPDer(\Lambda)$ for the span of all diagonalizable derivations of Λ. We also denote by k^+ the aditive group of k.

In this paper we study the fundamental derivations. These derivations appear naturally in the study of the Galois coverings and the fundamental group. One of the objectives of this work is to show that they are a natural object to be consider.

The fundamental derivations are particular case of diagonalizable derivations.

In [3], the authors proved that $SPDer(\Lambda)$ form a Lie ideal of $Der(\Lambda)$ whenever k has characteristic zero or is algebraically closed of positive characteristic.

The origin of the notion of fundamental derivation is the following theorem, due to Assem and De la Peña.

[1]The author thanks the Capes for a PhD Scholarship and the IME-USP.
[2]The author thanks CNPq for a research grant.

THEOREM 1 *[1] Let $\Lambda = k\Gamma/I$, a path algebra modulo relations. There is an injective map*

$$\begin{array}{ccc} Hom(\pi_1(\Gamma, I), k^+) & \longmapsto & Der(k\Gamma/I) \\ \Psi & \longmapsto & D_\Psi \end{array} .$$

Moreover each D_Ψ is a diagonalizable derivation and the induced map $Hom(\pi_1(\Gamma, I), k^+) \to HH^1(k\Gamma/I)$ is injective

The original proof of the proposition above assumes that the algebra is triangular, although the argument can be carried over to the general case. There is a proof for the general case, which appear in [4]. This former argument uses the following property of the diagonalizable derivation D_Ψ. For any pair of vertices x and y there is an undirected walk $\alpha_1^{\varepsilon(1)} \cdots \alpha_t^{\varepsilon(t)}$ such that $\sum_j \varepsilon(j)\, \lambda_j = 0$ (where $D_\Psi(\alpha_j) = \lambda_j \alpha_j$ and $\varepsilon(i) \in \{+1, -1\}$). This property turns out to be crucial in trying to characterize those diagonalizable derivations which arise from a fundamental group. These are the fundamental derivations.

In this paper $SFDer(\Lambda)$ denotes the span of all fundamental derivations of Λ.

The main result of this work, whose proof appears in section 4, is the following:

Theorem *Let Λ be a finite dimensional algebra, over a field k of characteristic zero. Then $SFDer(\Lambda)$ is a Lie ideal of the Lie algebra of derivations, $Der(\Lambda)$.*

We describe, briefly, now each section of the paper.

In section 1, we give a presentation of the fundamental group, which appears in [4].

In section 2, we give the definitions of fundamental derivation, weight function, and give a partial characterization of the fundamental derivations (see [3]).

In section 3, we establish the relation between a fundamental derivation and a certain Galois covering.

In section 4, we establish our main result, in characteristic zero, and give a counter example for a field of positive characteristic.

In section 5, we restrict our attention to the study of fundamental derivations over a finite dimensional monomial k-algebra. We show in the following theorem:

Theorem *Let Λ be a monomial algebra over a field k of characteristic zero. Then the vector space generated by all the inner derivations and by all fundamental derivations is equal $Der(\Lambda)$.*

1. Fundamental Groups

We review the definition of fundamental group, see for instance [1].

Let Γ be a finite connected quiver. We denote by Γ_0 the set of vertices and by Γ_1 the set of arrows.

For an arrow $\alpha \in \Gamma$, we denote by α^{-1} its formal inverse. A walk in Γ from x to y is a formal composition $\omega = \alpha_1^{\varepsilon(1)} \alpha_2^{\varepsilon(2)} \cdots \alpha_r^{\varepsilon(r)}$ (where $\alpha_i \in \Gamma_1$, $\varepsilon(i) \in \{+1, -1\}$ for $1 \leq i \leq t$), starting at x and ending at y. We denote by e_x the trivial path at x.

Let $\sim$ be the smallest equivalence relation on the set of all walks in Γ such that

1. For every arrows α of Γ_1, $\alpha : x \mapsto y$, $\alpha^{-1}\alpha \sim e_x$ and $\alpha\alpha^{-1} \sim e_y$.
2. If $\omega, \gamma, \omega_1, \omega_2$ are walks in Γ such that $\omega \sim \gamma$, then $\omega_1\omega\omega_2 \sim \omega_1\gamma\omega_2$, whenever both multiplications are defined.

We denote by $[u]$ the equivalence class of a walk u.

For a fixed vertex v, the classes of closed walks from v to itself has a group structure with operation the composition, this is by definition the fundamental group, $\pi_1(\Gamma)$. (Different choices of x yield isomorphic groups.)

We give now a presentation of the fundamental group, which appeared in [4]. This presentation is inspired in [1].

Fix a vertex v, the "base point". For each vertex $x \in \Gamma$, we choose a walk $\gamma_{v,x}$ from v to x, with the restriction that $\gamma_{v,v}$ is the trivial walk e_v.

The set $\gamma = \{\gamma_{v,x} : x \in \Gamma_0\}$ we will be referred to as a " parade data". If ω is any walk from x to y then we define

$$c_\gamma(\omega) = \gamma_{v,y}^{-1}\, \omega\, \gamma_{v,x}\,,$$

where $\gamma_{v,y}^{-1}$ denotes the walk of return to y from v, through of element of choice of parade data $\gamma_{v,y}$.

Remark 2 *Fix a choice of parade data γ, with base point v, the following facts are valid:*

1. *If ω is a closed walk from v to v then $c_\gamma(\omega) = \omega$;*
2. *$\pi_1(\Gamma)$ is generated by set $\{c_\gamma(\alpha) : \alpha \in \Gamma_1\}$;*
3. *Let $\gamma_{v,x} = \alpha_1^{\varepsilon(1)} \alpha_2^{\varepsilon(2)} \cdots \alpha_r^{\varepsilon(r)} \in \gamma$, with $r \geq 1$ then $[c_\gamma(\alpha_1)^{\varepsilon(1)} c_\gamma(\alpha_2)^{\varepsilon(2)} \cdots c_\gamma(\alpha_r)^{\varepsilon(r)}] = 1$ in $\pi_1(\Gamma)$.*

In [4] it is proved that $\pi_1(\Gamma)$ is the free group generated by $\{c_\gamma(\alpha) : \alpha \in \gamma_1\}$ modulo the parade walk relation obtained from γ.

Assume that I is an ideal of the path algebra $k\Gamma$ and that I is generated as an ideal by a set of relations R.
We define next the fundamental group of the presentation $k\Gamma/I$.
Assume that $\pi_1(\Gamma)$ is already described using parade data γ. Let $N(R)$ be the normal subgroup of $\pi_1(\Gamma)$ generated by $c_\gamma(p)c_\gamma(q^{-1})$, as p and q range over all paths that have non zero coefficient in a relation $\rho \in R$. Then

$$\pi_1(\Gamma, R) = \pi_1(\Gamma)/N(R).$$

We denote the canonical homomorphism from $\pi_1(\Gamma)$ to $\pi_1(\Gamma, R)$ by ξ (which depends on γ).

The fundamental group $\pi_1(\Gamma, I)$ is dependent on the choice of relations for I. This can be remedied as follows. A relation $\rho = \sum_{i=1}^{m} \lambda_i \omega_i \in I$ is called minimal if $m \geq 2$ and, for every non-empty proper subset $J \subset \{1, \cdots, m\}$, we have $\sum_{j \in J} \lambda_j \omega_j \notin I$. In [3], the authors showed that the fundamental group is the same for any two choices of R which consist of minimal and monomial relations; this group is denoted by $\pi_1(\Gamma, I)$.

2. Fundamental Derivation

In this section, we assume that I is an admissible ideal of $k\Gamma$.

The following theorem is a generalization of a theorem of Assem and De la Peña [1]. This generalization appeared in [4].

Theorem 3 *Choose a parade data for the connected quiver Γ. Then the map*

$$\begin{array}{rcl} \theta : Hom(\pi_1(\Gamma, I), k^+) & \longmapsto & SPDer(k\Gamma/I) \\ \Psi & \longmapsto & D_{\Psi,\gamma} : \alpha \mapsto \Psi[\gamma_{v,t(\alpha)}^{-1}\, \alpha\, \gamma_{v,o(\alpha)}]\, \alpha \end{array}$$

is injective and $D_{\Psi,\gamma}$ is a diagonalizable derivation.

This justifies the definition of fundamental derivation.

Definition 4 *We say that a derivation $D \in Der(\Lambda)$ is fundamental if there exists a finite quiver Γ, with an admissible ideal I of $k\Gamma$ such that $\Lambda \simeq k\Gamma/I$, and there is parade data together with some $\Psi \in Hom(\pi_1(\Gamma, I), k^+)$, so that $D = D_{\Psi\gamma}$.*

Remark 5 *If D is fundamental then*

1. *D is diagonalizable.*

2. *Exists a presentation of Λ where the images of paths in Λ are all eigenvectors. Consequently $D(rad\Lambda) \subset rad\Lambda$.*

We recall now the useful notion of weight (see [5]). Suppose that Γ is a finite quiver and H is a group with identity element e. A weight function for Γ with values in H is an assignment W from the arrows of Γ to H. We extend W multiplicatively so that vertices have weight $e \in H$ and $W(\alpha_1^{\varepsilon(1)} \ldots \alpha_t^{\varepsilon(t)}) = W(\alpha_1)^{\varepsilon(1)} \ldots W(\alpha_t)^{\varepsilon(t)}$ for every walk $\alpha_1^{\varepsilon(1)} \ldots \alpha_t^{\varepsilon(t)}$.

The weight function induces an H-grading on $k\Gamma$. We say that an ideal I of $k\Gamma$ is homogeneous for W provided it is homogeneous with respect to this grading. For such ideal, the weight induces a grading on $k\Gamma/I$. In the case of fundamental groups, we can consider the weight with values in $\pi_1(\Gamma, I)$ such that it sends $\alpha \in \Gamma_1$ to $\xi(c_\gamma \alpha)$ (see [5]). Therefore each parade data γ induces a $\pi_1(\Gamma, I)$-grading on $k\Gamma/I$ and we have the following corollary.

COROLLARY 6 *The algebra $\Lambda = k\Gamma/I$ is graded by its fundamental group $\pi_1(\Gamma, I)$.*

From now on, given the weight W we denote by H the subgroup generated by set $\{W(\alpha) : \alpha \in \Gamma_1\}$, which we call group generated by the weights.

We give a partial characterization of fundamental derivations. This characterization will be used various times in this work.

THEOREM 7 *Assume that I is an admissible ideal of $k\Gamma$ and D is $k\Gamma/I$ derivation such that:*

1. *D vanishes on the images of vertices, and for each arrow α, there is a scalar $\omega(\overline{\alpha})$, such that $D(\overline{\alpha}) = \omega(\overline{\alpha})\overline{\alpha}$.*

2. *There is a vertex v such that for every other vertex x there exists a walk $\gamma_{v,x} : \alpha_1^{\varepsilon(1)} \ldots \alpha_t^{\varepsilon(t)}$ from v to x such that $\sum_j \varepsilon(j)\omega(\alpha_j) = 0$.*

Then D is a fundamental derivation.

3. Galois covering and Fundamental Derivation

In this section, we show the relation between Galois coverings and fundamental derivations. The definitions of covering and graph, which we will use, can be found, for instance, in [5] and [6].

Let Γ be a locally finite directed graph, which we also call quiver (as usual in the theory of representations). We fix an embedding of Γ in real 3-space.

We also assume that this embedding satisfies the axioms of a graphs given in [6, Chapter 6, § 2].

Let $F : \widetilde{\Gamma} \mapsto \Gamma$ be a covering projection of graphs (in the topological sense,(see [6], chapters 5,6]). If p is a path in $\widetilde{\Gamma}$ then $F(p)$ is a path in Γ.

Let $L : \Gamma_0 \mapsto \widetilde{\Gamma}_0$ such that $L(x) \in F^{-1}(x)$. We call L a lifting. By uniqueness of path lifting, if p is a path (possibly undirected) with origin x and terminus y in Γ, then we denote by $L(p)$ the unique path in $\widetilde{\Gamma}$ originating at $L(x)$ such that $F(L(p)) = p$. Note that $L(p)$ in general does not terminate at $L(y)$. If $t = \sum \mu_i p_i$ is a k-linear combination of directed paths in $\widetilde{\Gamma}$, we let $L(t)$ denote $\sum_{i=1}^{m} \mu_i L(p_i)$.

Finally if $F : \widetilde{\Gamma} \mapsto \Gamma$ is a covering and $\widetilde{v} \in \widetilde{\Gamma}_0$, we let $\pi_1(\widetilde{\Gamma}, \widetilde{v})$ the fundamental group of $\widetilde{\Gamma}$ and $F_* : \pi_1(\widetilde{\Gamma}, \widetilde{v}) \mapsto \pi_1(\Gamma, F(\widetilde{v}))$ be the map induced by F (see [6], chapter 2]). We say $F : \widetilde{\Gamma} \mapsto \Gamma$ is a regular covering if $F_*(\pi_1(\widetilde{\Gamma}, \widetilde{v}))$ is a normal subgroup of $\pi_1(\Gamma, F(\widetilde{v}))$.

Recall that a relation r in $k\Gamma$ is a linear combination of paths, of length bigger than one, all starting at the same vertex $o(r)$ and ending at the same vertex $t(r)$. A quiver with relations (Γ, R) (also called a bounded quiver) is a pair, where Γ is a quiver and R is a set of relations in $k\Gamma$. Given (Γ, R) and (Γ', R') two quiver with relations, a morphism F between them is a morphism of quivers such that if $r = \sum \lambda_i \gamma_i \in R$ is a relation then $F(r) = \sum \lambda_i F(\gamma_i) \in R'$. In this faction, we can define the category of quivers with relations.

If $t \in k\Gamma$ and $u, v \in \Gamma_0$ the (u, v)-component of t, $c_{u,v}(t) = \sum_{j=1}^{l} \mu_{i_j} p_{i_j}$, where p_{i_j} is the subset of the set of paths $\sum_{j=1}^{m} \mu_{i_j} p_{i_j}$, which starts at u and ends at v. In other words $c_{u,v}(t) = utv$.

Given $(\widetilde{\Gamma}, \widetilde{R})$ and (Γ, R) quivers with relations and a morphism $F : (\widetilde{\Gamma}, \widetilde{R}) \mapsto (\Gamma, R)$, then F is a Galois covering if $F : \widetilde{\Gamma} \mapsto \Gamma$ is a regular covering of graphs such that

1. $\widetilde{R} = \{L(t) : L : \widetilde{\Gamma}_0 \mapsto \Gamma_0 \text{ is a lifting and } t \in R\}$;

2. If $t \in \widetilde{r}$ and $u, v \in \Gamma_0$ then there exist $\widetilde{u}, \widetilde{v}$ such that $F(c_{\widetilde{u},\widetilde{v}}(t)) = c_{u,v}(F(t))$.

Definition 8 *Let W be a weight function. The weight W is semi-connected if for all x and $y \in \Gamma_0$ and for all $h \in H$ exists a walk p from x to y such that $W(p) = h$ (see [5]).*

It is common, in theory of representation algebras, to consider path algebras $k\Gamma$ as k-categories, where the objects are the vertices and the morphism space between two vertices x and y is the vector space generated by paths from x to y which is denoted by $\Gamma(x, y)$.

Let $\Lambda = k\Gamma/I$ be an algebra and $W : \Gamma_1 \mapsto H$ a weight function, we can define a k-category $k\widetilde{\Gamma}$ whose objects are the elements of

$$\widetilde{\Gamma}_0 = \Gamma_0 \times H = \{(x, h) = x^h : (x, h) \in \Gamma_0 \times H\}$$

and the set of morphisms is

$$\widetilde{\Gamma}_1(x^h, y^g) = \{\omega^h : \omega \in \Gamma(x, y) \text{ and } W(\omega) = gh^{-1}\}.$$

The algebra $k\widetilde{\Gamma}(x, y)$ is isomorphic to the smash product $k\Gamma\#G$, associated with the grading defined by the weight. This was used to give a generalization of coverings to the setting of categories over a ring in [2]

Let $F : k\widetilde{\Gamma} \mapsto k\Gamma$ be a morphism of quivers with relations define by $F(x^h) = x$ and

$$\begin{array}{rcl} F : \widetilde{\Gamma}_1(x^h, y^g) & \mapsto & \Gamma(x, y) \\ \omega^h & \mapsto & \omega. \end{array}$$

The next theorem shows the connection between semi-connected weight and Galois coverings.

Theorem 9 *The weight W is semi-connected if and only if $F : (\widetilde{\Gamma}, \widetilde{I}) \mapsto (\Gamma, I)$ is a Galois covering with $\widetilde{I} = \{\sum \lambda_\gamma \gamma \in k\widetilde{\Gamma} : \sum \lambda_\gamma F(\gamma) \in I\}$.*

Proof. ($\Leftarrow$) Let $h \in H$ be and $u, v \in \Gamma_0$. Since Γ is a connected quiver, it follows from the definition that $\widetilde{\Gamma}$ is connected. Therefore given u^e and $v^h \in \widetilde{\Gamma}_0$ exists a walk ε^e from u^e to v^h. So ε is a walk from u to v such that $W(\varepsilon) = h$, thus W is semi-connected.

($\Rightarrow$) Let u^{h_1} and $v^{h_2} \in \widetilde{\Gamma}_0$. Using the hypothesis that W is semi-connected we see that there exists a walk p in Γ such that $W(p) = h_2 h_1^{-1} \in H$. It follows that $L(p)$ is a walk from u^{h_1} to v^{h_2} and thus the quiver $\widetilde{\Gamma}$ is connected.

By construction of $\widetilde{\Gamma}$ we have v^h is a vertex in $\widetilde{\Gamma}$ and p is a walk in Γ with origin in v then exists a unique walk with origin in v^h such that $F(p^h) = p$.

Using this observation, and the construction of the coverings we see that $F : \widetilde{\Gamma} \mapsto \Gamma$ is a regular covering.

Since I is a homogeneous ideal with respect W, in order to prove that F is Galois covering, we need to show that given $v^h \in \widetilde{\Gamma}$ we have that $F_*(\pi_1(\widetilde{\Gamma}, v^h))$

is a normal subgroup $\pi_1(\Gamma, v)$. This is consequence of the fact that if p is a walk closed in Γ, then each lifting is closed if and only if $W(p) = e$. □

In the next theorem we show a relation between Galois covering and fundamental derivation.

THEOREM 10 *Let D be a derivation of $Der(\Lambda)$ associated to a weight W. The derivation D is fundamental if and only if W is semi-connected.*

Proof.

($\Leftarrow$) The weight W is associated to D thus $D(\alpha) = W(\alpha)\alpha$ for all $\alpha \in \Gamma_1$. We fix a vertex x, since W is semi-connected, for each vertex y there exist a walk p such that $W(p) = e$. We now apply the Theorem 7 and get that D is fundamental.

($\Rightarrow$) We need to show that for all $x, y \in \Gamma_0$, and for all $h \in H$ there is a walk p from x to y such that $W(p) = h$. Since D is fundamental we fix a vertex v then there is a choose of parade data $\gamma = \{\gamma_{v,w} | w \in \Gamma_1\}$. Since $h \in H$ there is a walk $g = \alpha_1^{\epsilon(1)} \dots \alpha_t^{\epsilon(t)}$ such that $\sum \epsilon(i) W(\alpha_i) = h$.

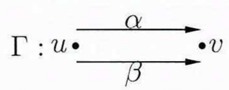
Take $p = \gamma_{v,y}\, \gamma_{v,t(g)}^{-1}\, g\, \gamma_{v,o(g)}\, \gamma_{v,x}^{-1}$. Thus we have $W(p) = W(g) = h$ and hence we conclude that W is semi-connected. □

COROLLARY 11 *Let $D \in Der(\Lambda)$ be a non zero fundamental derivation associated to a presentation $\Lambda = k\Gamma/I$. Then for all $x \in \Gamma_0$ there exist $\gamma_{x,x} = \alpha_1^{\epsilon(1)} \dots \alpha_t^{\epsilon(t)}$ such that:*

$$\sum \epsilon(i) W(\alpha_i) \neq e$$

COROLLARY 12 *Let D be a diagonalizable derivation of $Der(\Lambda)$ associated to a weight W, and $\widetilde{\Gamma}$ the covering associated with the weight W. Then D is fundamental if and only if $\widetilde{\Gamma}$ is connected.*

EXAMPLE 13 *Let Γ be the following quiver*



$\Lambda = k\Gamma$ and $D \in Der(\Lambda)$ defined by $D(\alpha) = \alpha$ e $D(\beta) = -\beta$.

As we will see in the figure 1, the quiver $\widetilde{\Gamma}$ is disconnected thus the derivation D is not fundamental.

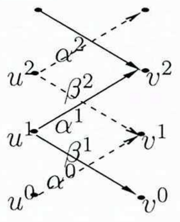


Figure 1:

In the next proposition we give a very natural condition, which is equivalent of a diagonalizable derivation to be associated to a weight. Recall that given a set of primitive idempotents a derivation is called normalized if it vanishes in this set of idempotents.

PROPOSITION 14 *Let Λ be finite-dimensional k-algebra and let D be a diagonalizable normalized derivation. The following two conditions are equivalent:*

1. *There exist an epimorphism of algebras $\pi : k\Gamma \longmapsto \Lambda$, and a weight function $W : \Gamma_1 \mapsto H$ such that $D(\pi(\omega)) = W(\omega)\pi(\omega)$ for all $\omega \in \Gamma_1$.*
2. $D(rad(\Lambda)) \subset rad(\Lambda)$.

Proof. We show only that the second condition implies the first, since the other implication is clear.

Since D is diagonalizable we can write $\Lambda = \Lambda_0 \coprod \Lambda_1$, where Λ_0 is the eigenspace associated to eigenvalue 0 and Λ_1 is the sum of eigenspace associated to non zero eigenvalues.

CLAIM 15 $\Lambda_1 \subset rad\Lambda$

Proof of the Claim. *Let α be an eigenvector associated to an eigenvalue $\lambda \neq 0$.*

$\Lambda = E \coprod rad(\Lambda)$, *thus* $\alpha = \lambda_1 v + \omega$ *where* $v \in E$, $\lambda \in k$ *and* $\omega \in rad(\Lambda)$,

$$D(\alpha) = D(\omega) = \lambda\alpha = \lambda\lambda_1 v + \lambda\omega.$$

$D(rad(\Lambda)) \subset rad(\Lambda)$ then $\lambda_1 = 0$ and therefore $\alpha \in rad(\Lambda)$. This finishes the proof of the claim.

Let $x \in rad\Lambda$, then $x = x_0 + x_1$ where $x_0 \in \Lambda_0$ and $x_1 \in \Lambda_1$. Since $\Lambda_1 \subset rad\Lambda$ we see that $x_0 \in rad\Lambda$ and we can make a restriction of D to $rad\Lambda$. We denote by $D|_{rad\Lambda}$ the restriction of D to $rad\Lambda$. This is diagonalizable linear map.

For each arrow α_i, let λ_i be a scalar such that $D(\alpha_i) = \lambda_i\alpha_i$.

It is clear that the set of classes modulo $rad^2(\Lambda)$ of $\{\overline{\alpha_1}, \ldots, \overline{\alpha_t}\}$ is a generator set of $rad\Lambda/rad^2\Lambda$. Reordering the elements if necessary, we can suppose that $\{\overline{\alpha_1}, \ldots, \overline{\alpha_l}\}$ is a basis of $rad\Lambda/rad^2\Lambda$.

Thus, there is a quiver Γ and an epimorphism $\pi : k\Gamma \longmapsto \Lambda$ such that the images of arrows form the set $\{\overline{\alpha_1}, \ldots, \overline{\alpha_l}\}$. This prove our proposition.

□

COROLLARY 16 *Let D be a diagonalizable normalized derivation and k a field of characteristic zero then D is associated to a weight.*

Proof. Let α be an eigenvector associated to an eigenvalue $\lambda \neq 0$. Thus, α commutes with $D(\alpha)$. We conclude that for every $n \in \mathbb{N}$,

$$D(\alpha^n) = n\alpha^{n-1}D(\alpha) = n\lambda\alpha^n.$$

Thus, α^n is eigenvector associated to $n\lambda$. Observe that all $n\lambda$ are distinct. Since the algebra is finite dimensional and $charack = 0$, we have $\alpha^n = 0$ to some n. It follows that α is nilpotent and thus $\alpha \in rad\Lambda$.

Hence we have proved that $D(rad(\Lambda)) \subset rad(\Lambda)$, using proposition 14 we get our result.

□

The corollary 16 is not valid if k is a field of positive characteristic. As we see in the next example.

EXAMPLE 17 *Let $\Lambda = k[\ x\ ]/\langle x^2\rangle$, where k is a field of characteristic two, and $D(x) = x + 1$. The derivation D is a diagonalizable normalized but it is not associated to a weight because $D(rad(\Lambda)) \subsetneq rad(\Lambda)$.*

4. The Lie Structure of the Space of Fundamental Derivations

In this section we look at the Lie structure of the vector space generated by the fundamental derivations.

We state now a useful lemma which appeared in [3].

LEMMA 18 *[3] Let Λ be a finite dimensional algebra over a field k of characteristic zero. If D and $E \in Der(\Lambda)$ satisfying $[D, E] = \mu E$ for $\mu \neq 0$ and E is nilpotent then:*

$$\exp(\frac{1}{\mu}E)^{-1}\, D\, \exp(\frac{1}{\mu}E) = D + E.$$

Using this lemma we get the following result for the fundamental derivations.

THEOREM 19 *Let Λ be a finite dimensional over a field k of characteristic zero. Then $SFDer(\Lambda)$ is an Lie ideal of $Der(\Lambda)$.*

Proof. It is sufficient to prove that if D is a fundamental derivation then $[D, E] \in SFDer(\Lambda)$ for all $E \in Der(\Lambda)$.

Since D is a diagonalizable derivation, adD is also diagonalizable where $adD : Der(\Lambda) \to Der(\Lambda)$ is the linear transformation which associates to each derivation F the derivation $[D, F] = DF - FD$. Therefore, we can write $E = \sum_\mu E(\mu)$ where $E(\mu)$ is an eigenvector of adD associated to eigenvalue μ. Then

$$[D, E] = \sum_\mu \mu E(\mu).$$

We see that $E(\mu)$ is nilpotent if $\mu \neq 0$. Let $\alpha \in rad\Lambda$. Since Λ is a finite dimensional algebra, we can consider $n \in \mathbb{N}$ such that $\alpha^n = 0$ and $\alpha^{n-1} \neq 0$.

If $E(\mu)(\alpha) = 1$ then $E(\mu)(\alpha^n) = n\alpha^{n-1} = 0$. This is a contradiction.

Thus, if $\alpha \in rad\Lambda$, it follows that $E(\mu)^n(\alpha) \in rad^n\Lambda$. Therefore $E(\mu)$ is nilpotent, since Λ is a finite dimensional algebra.

By Lemma 18, $D + E(\mu) = P^{-1}\, D\, P$ where

$$P = \exp(\frac{1}{\mu}E(\mu)).$$

Let α be an eigenvector of D associated to eigenvalue λ_α. It follows that

$$P^{-1}DP\,(P^{-1}\alpha) = P^{-1}D(\alpha) = \lambda_\alpha P^{-1}(\alpha).$$

Thus, $P^{-1}(\alpha)$ is an eigenvector of $P^{-1}DP$ associated also to eigenvalue λ_α.

Let us assume first that $D+E(\mu)$ is fundamental for all $\mu \neq 0$, and conclude that $SFDer(\Lambda)$ is a Lie ideal of $Der(\Lambda)$. In this case

$$\mu E(\mu) = \mu(\,(D + E(\mu)) - D) \in SFDer(\Lambda).$$

It follows that $[D,E] = \sum_\mu \mu\, E(\mu) \in SFDer(\Lambda)$.

It remains to be shown that $D + E(\mu)$ is fundamental for all $\mu \neq 0$. Since D is fundamental, it exists a presentation (Γ, I) of Λ such that given a vertex x, for another vertex y it exists a walk

$$\gamma_{x\,y} : \alpha_1^{\varepsilon(1)} \ldots \alpha_t^{\varepsilon(t)} \text{ such that } \sum_j \varepsilon(j) W(\alpha_j) = e. \tag{1}$$

We define $I^\star = \{P^{-1}(r) : r \in I\}$ and $\Gamma^\star = (\Gamma_0^\star, \Gamma_1^\star)$, where $\Gamma_0^\star = \{P^{-1}(v) : v \in \Gamma_0\}$ and $\Gamma_1^\star = \{P^{-1}(\alpha) : \alpha \in \Gamma_1\}$.

It follows from the definition of P^{-1} that $k\Gamma/I \simeq k\Gamma^\star/I^\star$.

Given a presentation $(\Gamma^\star, I^\star)$, we fix a vertex $x^* \in \Gamma_0^*$. Then for any other vertex y^*, there are $x, y \in \Gamma_0$ such that $P^{-1}(x) = x^*$ and $P^{-1}(y) = y^*$. By (1), it exists a walk

$$P^{-1}(\gamma_{x\,y}) = P^{-1}(\alpha_1)^{\varepsilon(1)} \ldots P^{-1}(\alpha_t)^{\varepsilon(t)}$$

from x^* to y^* such that

$$\sum_j \varepsilon(j) W(P^{-1}(\alpha_j)) = \sum_j \varepsilon(j) W(\alpha_j) = e.$$

Thus, the derivation $D + E(\mu)$ is fundamental for all $\mu \neq 0$, which ends the proof of the theorem. □

REMARK 20 *It was proved in [3] that if the field k has characteristic zero or is algebraically closed then $SPDer(\Lambda)$ is a Lie ideal of $Der(\Lambda)$. Inspired by this result we asked ourselves the same question arises naturally. Is the subspace $SFDer(\Lambda)$ a Lie ideal of $Der(\Lambda)$? This is not true, in general. We give the following example*

EXAMPLE 21 *Let $\Lambda = k[\ x\ ]/\langle x^2 \rangle$, where k is a field of characteristic two.*

Let D and $E \in Der(\Lambda)$ defined by

$D(x) = x$ and $E(x) = 1$.

The derivation D is fundamental and $[D,E](x) = 1$. However $[D,E]$ is not a fundamental derivation because it does not take $rad(\Lambda)$ in $rad(\Lambda)$.

In order to give a partial positive answer for the result in characteristic zero, we need to increase the set of fundamental derivations adding also the inner derivations.

5. Fundamental derivation of monomial algebra

We remember that the set of inner derivations is a Lie ideal of $Der(\Lambda)$. We denote by $\overline{SFDer(\Lambda)}$ the vector space generated by the union of the inner and the fundamental derivations.

Recall that an ideal I is called monomial if it can be generated by a set of paths. An algebra Λ is called monomial if it is isomorphic to an algebra of the form $k\Gamma/I$ where I is monomial.

Given a path algebra $k\Gamma$, an arrow α in $k\Gamma[u,v]$ and an element $\rho \in k\Gamma[u,v]$ we denote by $\rho\frac{\partial}{\partial\alpha}$ the derivation of $k\Gamma$ which vanishes on the vertex set and in all arrows except α, and it associates to α the element ρ, to define it on a general path we use the Leibniz's rule.

THEOREM 22 *Let $\Lambda = k\Gamma/I$ be a finite dimensional monomial k-algebra and D a derivation of $k\Gamma/I$ which vanishes on images of vertices and for which the images of arrows are eigenvectors then $D \in \overline{SFDer(\Lambda)}$.*

Proof. We know that for each $\alpha \in \Gamma_1$, there is a scalar λ_α such that $D(\alpha) = \lambda_\alpha\alpha$. Since Λ is monomial, $\alpha\frac{\partial}{\partial\alpha}$ is an element of $Der(\Lambda)$ for all $\alpha \in \Gamma_1$. It follows that

$$D = \sum_{\alpha\in\Gamma_1} \lambda_\alpha\alpha\frac{\partial}{\partial\alpha}.$$

We show that $\alpha\frac{\partial}{\partial\alpha}$ is a fundamental or inner derivation for all arrow α of Γ_1. We need to consider the two cases.

1. α belongs to a circuit.

   a) α is a loop, that is α begins and ends in the vertex v. Given another vertex x distinct of v, it is always possible to obtain a walk ε from v to x such that it does not pass through α or α^{-1} and thus $\alpha\frac{\partial}{\partial\alpha}(\varepsilon) = 0$. If we take the walk $\alpha\alpha^{-1}$ from vertex x to vertex x, the weight of this walk is also trivial. By Theorem 7, in this case $\alpha\frac{\partial}{\partial\alpha}$ is a fundamental derivation.

   b) We assume now that α is not a loop but it belongs to a circuit

   $$v_1 \xrightarrow{\alpha} v_2 \xrightarrow{\alpha_2} \cdots \to v_t \xrightarrow{\alpha_t} v_1$$

   Where there are no repeated arrows.

   We fix the vertex v_1. The walk $\gamma_{v_1v_i} = \alpha_t^{-1}\cdots\alpha_i^{-1}$ is such that $\alpha\frac{\partial}{\partial\alpha}(\gamma_{v_1v_i}) = 0$ for all $i \in \{2,\cdots,t\}$.

Given a vertex x such that $x \notin \{v_1, \cdots, v_t\}$. It is always possible to obtain a walk ε from v_1 to x such that, it does not pass through α and thus $\alpha\frac{\partial}{\partial\alpha}(\varepsilon) = 0$. It follows from Theorem 7 that $\alpha\frac{\partial}{\partial\alpha}$ is fundamental.

2. α not part of a circuit.

   Define the following set
   $\Gamma_{o(\alpha)} = \{v \in \Gamma_0 :$ it exists a walk linking v to $o(\alpha)$ without to pass through $\alpha\}$. It is a full subquiver of Γ.

   If β is an arrow beginning at the vertex v_1 and ending at the vertex v_2 with $v_1 \neq v_2$ and both vertex are different of $o(\alpha)$ then

   $[\sum_{v\in\Gamma_{o(\alpha)}} v, \beta] = [v_1 + v_2, \beta] = 0$

   If β is an arrow beginning and ending at the vertex v_1 then also we have

   $[\sum_{v\in\Gamma_{o(\alpha)}} v, \beta] = [v_1, \beta] = 0.$

   Finally, $[\sum_{v\in\Gamma_{o(\alpha)}} v,\ \alpha] = [o(\alpha),\ \alpha] = \alpha.$

   Thus $[\sum_{v\in\Gamma_{o(\alpha)}} v,\ -] = \alpha\frac{\partial}{\partial\alpha}$, because they coincide in the vertices and in the arrows. We have proved in this case, $\alpha\frac{\partial}{\partial\alpha}$ is an inner derivation.

   We have shown that $\alpha\frac{\partial}{\partial\alpha}$ is always a fundamental or an inner derivation for all element of Γ_1.

□

From now on, k will denote a field of characteristic zero. We recall some facts from [4].

Let Λ be a finite dimensional positively graded algebra such that

$$\Lambda = \Lambda_0 \amalg \Lambda_1 \amalg \cdots \amalg \Lambda_t$$

where $\Lambda_0 = ke_1 \amalg \cdots \amalg ke_n$ with e_j idempotent orthogonal for all $j \in \{1, \cdots, n\}$ and Λ is generated, as an algebra, by Λ_0 and Λ_1.

The Euler derivation associated to a graduation is defined by

$$E(x) = mx \text{ for all } x \in \Lambda_m.$$

Since E is diagonalizable, adE also is diagonalizable. Then it exists a basis of derivations $D_1, \cdots, D_t$ of $Der(\Lambda)$ such that $[E, D_i] = \mu_i D_i$ with $\mu_i \in k$ for all $i \in \{1, \cdots, t\}$.

We denote by $\mathcal{W}$ the Lie algebra consisting of the derivations which vanishes in each idempotent e_i, and by $\mathcal{W}_\lambda$ the eigenspace for adE associated to eigenvalue λ.

LEMMA 23 *Using the previous notations it is prove in [3] that*

$$\mathcal{W} = \mathcal{W}_0 \amalg \cdots \amalg \mathcal{W}_{t-1}$$

with $\mathcal{W}_i(\Lambda_j) \subset \Lambda_{i+j}$ *and* $[\mathcal{W}_i, \mathcal{W}_j] \subset \mathcal{W}_{i+j}$.

We consider now $\Lambda = k\Gamma/I$, where the graded of Λ is induced from the grading in $k\Gamma$ given by length of paths.

LEMMA 24 *Let* $D \in \mathcal{W}$ *be fundamental. Let* $D = A + B$ *be with* $A \in \mathcal{W}_0$ *and* $B \in \sum_{j\geq 1} \mathcal{W}_j$*, then* A *is fundamental.*

Proof. By hypothesis D is fundamental then exist a presentation (Γ, I) of Λ such that $D(\alpha) = \lambda_\alpha \alpha$ for every $\alpha \in \Gamma_1$ where $\lambda_\alpha \in k$. Thus $A(\alpha) = \lambda_\alpha \alpha$ for every $\alpha \in \Gamma_1$ hence $D = A + B$ and $B \in \sum_{j\geq 1} \mathcal{W}_j$.

Then the weight associated the derivation D is the same associated the derivation A. Using Theorem 7, we see that A is fundamental. □

LEMMA 25 *Using the previous notations we have*

$$\overline{SF(\mathcal{W})} = \overline{SF(\mathcal{W}_0)} \amalg \mathcal{W}_1 \amalg \cdots \amalg \mathcal{W}_{t-1}.$$

Proof. Let $D \in \mathcal{W}_j$ with $j \neq 0$. Since the Euler derivation E is equal to $\sum_{\alpha\in\Gamma_1} \alpha \frac{\partial}{\partial \alpha}$ using Theorem 22 we get that $E \in \overline{SF(\mathcal{W})}$.

Applying the Theorem 19, and the fact that $Inn(\Lambda)$ is a Lie ideal of $\mathcal{W}$, we get that $\frac{1}{j}[E, D] = D \in \overline{SF(\mathcal{W})}$.

To finish our proof we need to prove that $\overline{SF(\mathcal{W})} \bigcap \mathcal{W}_0 \subset \overline{SF(\mathcal{W}_0)}$.

Let $B \in \mathcal{W}_0 \bigcap \overline{SF(\mathcal{W})}$ then $B = \sum_i (A_i + D_i)$ where $A_i + D_i$ is fundamental or inner derivation in $\mathcal{W}$ for all i with $A_i \in \mathcal{W}_0$ and $D_i \in \sum_{j\geq 1} \mathcal{W}_j$ hence $B = \sum_i A_i$.

If $A_i + D_i$ is fundamental, by lemma 24, A_i is fundamental.

If $A_i + D_i$ is inner then there is $s = s_0 + s_1 + \cdots + s_n \in S$ with $s_i \in \Lambda_i$, $\forall\, i \in \{1, \cdots, n\}$ such that $(A_i + D_i)(x) = xs - sx$, for all $x \in \Lambda$.

Let $x \in \Lambda_j$ then $(A_i + D_i)(x) = (xs_0 - s_0 x) + (xs_1 - s_1 x) + \cdots + (xs_n - s_n x)$.

Since $\mathcal{W}_i(\Lambda_j) \subset \Lambda_{i+j}$ we have $A_i(x) = xs_0 - s_0 x$. It follows that $A_i + D_i$ is inner therefore A_i is inner.

We can conclude that $\overline{SF(\mathcal{W}_0)} = \overline{SF(\mathcal{W})} \bigcap \mathcal{W}_0$, which finishes the proof.
□

The following result appears also in [3].

LEMMA 26 *[3] Let $k\Gamma/I$ be a finite dimensional monomial algebra. Then the set $\{\beta\frac{\partial}{\partial\alpha} : \alpha,\ \beta \in \Gamma_1$ and $\beta\frac{\partial}{\partial\alpha}(I) \subset I\}$ generates $\mathcal{W}_0$, as vector space.*

We are now in conditions to prove the following theorem.

THEOREM 27 *Let $k\Gamma/I$ a finite dimensional monomial algebra. Then $\overline{SF(\mathcal{W})} = \mathcal{W}$.*

Proof. By Lemma 25, it is sufficient to prove that $\overline{SF(\mathcal{W}_0)} = \mathcal{W}_0$.

A derivation of form $\alpha\frac{\partial}{\partial\alpha} \in \mathcal{W}_0$, with $\alpha \in \Gamma_1$, are such that given $\omega \in I$, $\alpha\frac{\partial}{\partial\alpha}(\omega)$ is zero or is a multiple of ω. From this fact and Lemma 26, it follows that $\mathcal{W}_0$ has a generating set consisting of all derivations of the form $\alpha\frac{\partial}{\partial\alpha}$, and some derivations of the form $\beta\frac{\partial}{\partial\alpha}$.

Using Theorem 22, we see that $\alpha\frac{\partial}{\partial\alpha} \in \overline{SF(\mathcal{W}_0)}$ for all $\alpha \in \Gamma_1$. Using the Theorem 19 and remembering that $Inn(\Lambda)$ is a Lie ideal of $Der(\Lambda)$ we have

$$[\beta\frac{\partial}{\partial\alpha}, \alpha\frac{\partial}{\partial\alpha}] = \beta\frac{\partial}{\partial\alpha} \in \overline{SF(\mathcal{W})}.$$

Therefore $\overline{SF(\mathcal{W})} = \mathcal{W}$. □

References

[1] I. Assem and J. A. de la Peña. The fundamental groups of a triangular algebra. *Comm. in algebra*, 24(1):187–208, 1996.

[2] Claude Cibils and Eduardo Marcos. Skew category, galois covering and smash of category over a ring. *pre-print.*

[3] Daniel R. Farkas, Christof Geiss, Edward L. Green, and Eduardo Marcos. Diagonalizable derivations of finite-dimensional algebras i. *Israel Journal of Mathematics*, (117):157–181, 2000.

[4] Daniel R. Farkas, Edward L. Green, and Eduardo Marcos. Diagonalizable derivations of finite-dimensional algebras ii. *Pacific journal of mathematics*, 196(2), 2000.

[5] Edward L. Green. Graphs with relations, coverings and group-graded algebras. *Transactions of the American Mathematical society*, 279(1), 1983.

[6] William S. Massey. Algebraic topology: an introduction. *Harcourt, Brace & World, New york*, 1967.

Sônia Maria Fernandes
somari@ufv.br
Departamento de Matemática
Universidade Federal de Viçosa
Viçosa-MG, CEP 36570-000, Brasil.

Eduardo do Nascimento Marcos
enmarcos@ime.usp.br
IME-USP Departamento de Matemática
Universidade de São Paulo,
CP 66281 CEP 05315-970,
São Paulo-SP, Brasil.

Random sampling for continuous time Markov additive processes

Claudio Macci[1]

ABSTRACT

Let $((J_t, S_t))$ be a continuous time Markov additive process with cadlag paths and let (T_n) be a sequence of random variables such that $(T_n - T_{n-1})$ are i.i.d. and positive, with $T_0 = 0$. Then, if $((J_t, S_t))$ and (T_n) are independent, we prove that $((J_{T_n}, S_{T_n}))$ is a discrete time Markov additive process and we provide an expression for the corresponding MA transition function. Moreover we concentrate our attention on the finite environment's state space case and, under further quite general hypotheses, we present some large deviations results.

Keywords: Markov additive processes; large deviations; level crossing probabilities; Lundberg parameter

1. Introduction

In this paper we concentrate our attention on randomly sampled continuous time Markov additive processes with cadlag paths. As pointed out in condition **RS1** presented below, roughly speaking we have a continuous time Markov additive process sampled at independent random times such that the inter-observations times are i.i.d. and positive.

The paper is organized as follows. In Section 2 we deal with a general environment's state space case and we prove that the random sampled process is a discrete time Markov additive process. In Section 3 we concentrate our attention on the finite environment's state space case. After recalling some preliminaries, we consider some quite general hypotheses pointed out in condition **RS2** and we prove the results. Our main results concern large deviations theory and one result provides an estimate for level crossing probabilities described in terms of the so-called Lundberg parameter (other terms are Lundberg exponent or adjustment coefficient). The Lundberg parameter is the unique positive solution of the Lundberg equation (for random walks and Lévy processes see e.g. [1], Chapter XII, Section 5, page 269); this terminology is standard in risk theory (see Preface in [5], page vi). Anyway our

[1]This work has been partially supported by Murst Project "Processi Stocastici e applicazioni a Filtraggio, Controllo, Simulazione e Finanza Matematica".

result concerning this item allows to say that, for any light tail distribution for the inter-observations times, the Lundberg parameter of the continuous time Markov additive process coincides with the Lundberg parameter of the randomly sampled discrete time Markov additive process.

Finally we point out that all the Markov processes in this paper are homogeneous.

2. Preliminaries and random sampling

2.1. Preliminaries

Let us start with some preliminaries on Markov additive processes in [6]; see also [2] (Chapter 2, Section 5, pp. 39-47) and the references cited therein. Given a measurable space $(E, \mathcal{E})$, a Markov additive process (MAP for short) is a $E \times \mathbb{R}$ valued Markov process $((J_t, S_t))$ with transition function

$$q(t,(j,s);A\times\Gamma) = P((J_{u+t}, S_{u+t}) \in A\times\Gamma | (J_u, S_u) = (j,s))$$

(with $t, u \geq 0$, $(j,s) \in E\times\mathbb{R}$, $A \in \mathcal{E}$ and $\Gamma \in \mathcal{B}(\mathbb{R})$, where $\mathcal{B}(\mathbb{R})$ is the Borel $\sigma-$algebra on $\mathbb{R}$) such that (J_t) is a Markov process (called environment) and we have the following identity:

$$q(t,(j,s);A\times(\Gamma+s)) \equiv q(t,(j,0);A\times\Gamma);$$

in such a case

$$Q_t(j;A\times\Gamma) \equiv q(t,(j,0);A\times\Gamma)$$

is called MA transition function of $((J_t, S_t))$. The process (S_t) is called additive part.

This presentation can be considered for both continuous time case and discrete time case. In particular for the discrete time case it is enough to consider $u = n$ and $t = 1$ for each $n \in \mathbb{N}$; thus the transition function for a discrete time MAP $((J_n^{(d)}, S_n^{(d)}))$ is

$$q^{(d)}((j,s);A\times\Gamma) \equiv P((J_{n+1}^{(d)}, S_{n+1}^{(d)}) \in A\times\Gamma | (J_n^{(d)}, S_n^{(d)}) = (j,s)),$$

we have the identity

$$q^{(d)}((j,s);A\times(\Gamma+s)) \equiv q^{(d)}((j,0);A\times\Gamma)$$

and the MA transition function is

$$Q^{(d)}(j;A\times\Gamma) \equiv q^{(d)}((j,0);A\times\Gamma).$$

2.2. Random sampling

Let us start with the following condition.

Condition RS1. Let $((J_t, S_t))$ be a continuous time MAP with cadlag paths and (T_n) a sequence of positive random variables such that: $T_0 = 0$ and $(T_n - T_{n-1})$ are i.i.d.; $((J_t, S_t))$ and (T_n) are independent.
Moreover let $((J_n^{(d)}, S_n^{(d)}))$ be the discrete time process defined by $(J_n^{(d)}, S_n^{(d)}) = (J_{T_n}, S_{T_n})$.

The results in this subsection show that $((J_n^{(d)}, S_n^{(d)}))$ in **RS1** is a MAP and its MA transition function can expressed in terms of the MA transition function of $((J_t, S_t))$ and the common distribution of the random variables $(T_n - T_{n-1})$. The first result (Lemma 2.1) generalizes Theorem 3.1 in [9] which concerns the case when the set E is at most countable; the proof is similar and therefore omitted.

LEMMA 2.1 *Let F be a Polish space and let $\mathcal{F}$ be the Borel σ-field. Let (X_t) be a F valued Markov process with cadlag paths and with transition function*

$$p(t, x; A) = P(X_{u+t} \in A | X_u = x)$$

(with $t, u \geq 0$, $x \in F$, $A \in \mathcal{F}$). Let (T_n) be a sequence of positive random variables such that: $T_0 = 0$ and $(T_n - T_{n-1})$ are i.i.d.; (X_t) and (T_n) are independent.
Then the discrete time process $(X_n^{(d)})$ defined by $X_n^{(d)} = X_{T_n}$ is a Markov process with transition function

$$p^{(d)}(x; A) = \mathbb{E}[p(T_1, x; A)].$$

PROPOSITION 2.2 *Assume that **RS1** holds and let $Q_t(j; A \times \Gamma)$ be the MA transition function of $((J_t, S_t))$. Then $((J_n^{(d)}, S_n^{(d)}))$ is a discrete time MAP with MA transition function*

$$Q^{(d)}(j; A \times \Gamma) \equiv \mathbb{E}[Q_{T_1}(j; A \times \Gamma)].$$

Proof. First of all, by Lemma 2.1, $((J_n^{(d)}, S_n^{(d)}))$ and $(J_n^{(d)})$ are discrete time Markov processes and, if we denote the transition function of $((J_t, S_t))$ by $q(t, (j, s); A \times \Gamma)$, the transition function of $((J_n^{(d)}, S_n^{(d)}))$ is

$$q^{(d)}((j, s); A \times \Gamma) \equiv \mathbb{E}[q(T_1, (j, s); A \times \Gamma)].$$

Furthermore, since $((J_t, S_t))$ is a continuous time MAP, we have the identity

$$q(t, (j, s); A \times (\Gamma + s)) \equiv q(t, (j, 0); A \times \Gamma) \equiv Q_t(j; A \times \Gamma);$$

then we also have

$$q^{(d)}((j,s);A\times(\Gamma+s))\equiv\mathbb{E}[q(T_1,(j,s);A\times(\Gamma+s))]\equiv$$
$$\equiv\mathbb{E}[q(T_1,(j,0);A\times\Gamma)]\equiv q^{(d)}((j,0);A\times\Gamma).$$

In conclusion $((J_n^{(d)},S_n^{(d)}))$ is a discrete time MAP with MA transition function

$$Q^{(d)}(j;A\times\Gamma)\equiv q^{(d)}((j,0);A\times\Gamma)\equiv\mathbb{E}[q(T_1,(j,0);A\times\Gamma)]\equiv\mathbb{E}[Q_{T_1}(j;A\times\Gamma)].\ \square$$

3. The finite environment's state space case

Now let us consider the case in which E is finite with cardinality s and, without loss of generality, we can consider $E=\{1,\dots,s\}$. In such a case the structure of the MAP is completely understood (see e.g. [2], pp. 39-40). In order to have a simpler presentation we always assume that the additive components starts at zero with probability 1. In what follows we present some large deviations principles as a consequence of Gärtner Ellis Theorem (see e.g. [4], Theorem 2.3.6, page 45); in particolar the concept of essentially smooth function is needed (see e.g. [4], Definition 2.3.5, page 44).

3.1. Preliminaries on the continuous time case

Let $((J_t,S_t))$ be a continuous time MAP. Then $((J_t,S_t))$ is a $E\times\mathbb{R}$ valued continuous time Markov process and (J_t) is a continuous time Markov chain with state space E and intensity matrix $\Lambda=(\lambda_{ij})$; moreover, assuming to have cadlag paths, conditionally on (J_t) we have the following situation:
• when $J_t=i\in E$ for $t\in[t_1,t_2)$ the process $(S_t)_{t\in[t_1,t_2)}$ evolves like a Lévy process with drift $\mu_i\in\mathbb{R}$, variance $\sigma_i^2\geq 0$ and Lévy measure ν_i;
• each transition of (J_t) from $i\in E$ to $j\neq i$ gives rise to a jump for (S_t) at the same time which is distributed with a specified cumulant κ_{ij}, namely

$$i,j\in E,\ i\neq j\Rightarrow\log\mathbb{E}[e^{\alpha(S_t-S_{t-})}|J_{t-}=i,\ J_t=j]\equiv\kappa_{ij}(\alpha).$$

In view of what follows it is useful to consider the cumulants $(\kappa_i:i\in E)$ defined by

$$\kappa_i(\alpha)\equiv\mu_i\alpha+\frac{\sigma^2}{2}\alpha^2+\int_{\mathbb{R}}(e^{\alpha x}-1)\nu_i(dx).$$

The matrices $G(\alpha)=(G_{ij}(\alpha))$ (varying $\alpha\in\mathbb{R}$) defined by

$$G_{ij}(\alpha)\equiv\begin{cases}\lambda_{ij}e^{\kappa_{ij}(\alpha)} & i\neq j\\ \lambda_{ii}+\kappa_i(\alpha) & i=j\end{cases}$$

play a crucial role; moreover we point out that $G(0) = \Lambda$ and, for each exponential matrix $e^{tG(\alpha)} = ((e^{tG(\alpha)})_{ij})$ (for $t > 0$), we have

$$(e^{tG(\alpha)})_{ij} = \mathbb{E}[e^{\alpha S_t} 1_{J_t=j} | J_0 = i]. \tag{1}$$

Assume that (J_t) is irreducible. Then, whatever is the distribution of J_0, there exists the limit

$$\kappa(\alpha) = \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}[e^{\alpha S_t}] \; (\forall \alpha \in \mathbb{R})$$

as an extended real value; then, if such a function κ is finite in a neighbourhood of $\alpha = 0$, essentially smooth and lower semicontinuous, by Gärtner Ellis Theorem $(\frac{S_t}{t})$ satisfies the large deviations principle with rate function κ^* defined by

$$\kappa^*(x) = \sup_{\alpha\in\mathbb{R}} [x\alpha - \kappa(\alpha)].$$

Moreover, when the matrix $G(\alpha)$ has finite entries, we can apply Perron Frobenius Theorem to the matrices $e^{tG(\alpha)}$ (for each $t > 0$) and we have

$$\sum_{j\in E} (e^{tG(\alpha)})_{ij} h_j(\alpha) = e^{t\kappa(\alpha)} h_i(\alpha) \; (\forall i \in E) \tag{2}$$

where $e^{t\kappa(\alpha)}$ is a simple and positive eigenvalue of $e^{tG(\alpha)}$ which is equal to its spectral radius and $(h_1(\alpha), \ldots, h_s(\alpha))$ is a positive eigenvector.

3.2. Preliminaries on the discrete time case

Let $((J_n^{(d)}, S_n^{(d)}))$ be a discrete time MAP. Then $((J_n^{(d)}, S_n^{(d)}))$ is a $E \times \mathbb{R}$ valued discrete time Markov process and $(J_n^{(d)})$ is a discrete time Markov chain with state space E and transition matrix $P = (p_{ij})$; moreover conditionally on $(J_n^{(d)})$ we have the following situation:
the random variables $(S_n^{(d)} - S_{n-1}^{(d)})$ are independent and, as far as the conditional distribution of such random variables given $J = (J_n^{(d)})$, we have

$$\log \mathbb{E}[e^{\alpha(S_n^{(d)} - S_{n-1}^{(d)})} | J] = \log \mathbb{E}[e^{\alpha(S_n^{(d)} - S_{n-1}^{(d)})} | J_{n-1}^{(d)}, J_n^{(d)}] = [\kappa_{ij}^{(d)}(\alpha)]_{(i,j)=(J_{n-1}^{(d)}, J_n^{(d)})}$$

where $\kappa_{ij}^{(d)}$ is a specified cumulant of a real random variable.

The matrices $F^{(d)}(\alpha) = (F_{ij}^{(d)}(\alpha))$ (varying $\alpha \in \mathbb{R}$) defined by

$$F_{ij}^{(d)}(\alpha) = \mathbb{E}[e^{\alpha S_1^{(d)}} 1_{J_1^{(d)}=j} | J_0^{(d)} = i] = e^{\kappa_{ij}^{(d)}(\alpha)} p_{ij} \tag{3}$$

play a crucial role; moreover we point out that $F^{(d)}(0) = P$.

Assume that $(J_n^{(d)})$ is irreducible. Then, whatever is the distribution of $J_0^{(d)}$, there exists the limit

$$\kappa_{(d)}(\alpha) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[e^{\alpha S_n^{(d)}}] \ (\forall \alpha \in \mathbb{R}) \tag{4}$$

as an extended real value; then, if such a function $\kappa_{(d)}$ is finite in a neighbourhood of $\alpha = 0$, essentially smooth and lower semicontinuous, by Gärtner Ellis Theorem $(\frac{S_n^{(d)}}{n})$ satisfies the large deviations principle with rate function $\kappa_{(d)}^*$ defined by

$$\kappa_{(d)}^*(x) = \sup_{\alpha\in\mathbb{R}} [x\alpha - \kappa_{(d)}(\alpha)].$$

Moreover, when the matrix $F^{(d)}(\alpha)$ has finite entries, we can apply Perron Frobenius Theorem to $F^{(d)}(\alpha)$ and then $e^{\kappa_{(d)}(\alpha)}$ is a simple and positive eigenvalue of $F^{(d)}(\alpha)$ which is equal to its spectral radius.

3.3. Random sampling for continuous time MAP

Let us consider the following condition which plays a crucial role in what follows.

Condition RS2. Let $((J_t, S_t))$ be a continuous time MAP with cadlag paths and such that $E = \{1, \ldots, s\}$ and (J_t) is irreducible. Moreover let (T_n) be a sequence of positive random variables such that: $((J_t, S_t))$ and (T_n) are independent; $T_0 = 0$ and $(T_n - T_{n-1})$ are i.i.d. and the common cumulant

$$\Lambda_T(\beta) \equiv \log \mathbb{E}[e^{\beta T_1}]$$

is such that the function

$$(\alpha, \beta) \in \mathbb{R}^2 \mapsto \Lambda_T(\kappa(\alpha) + \beta)$$

is finite in a neighbourhood of $(\alpha, \beta) = (0, 0)$, essentially smooth and lower semicontinuous.
Then let $((J_n^{(d)}, S_n^{(d)}))$ be the discrete time process defined by $(J_n^{(d)}, S_n^{(d)}) = (J_{T_n}, S_{T_n})$.

When **RS2** holds, the common distribution of the random variables $(T_n - T_{n-1})$ has to be light tail, namely the function Λ_T has to be finite in a neighbourhood of $\beta = 0$. Condition **RS2** plays a crucial role below when we deal with large deviations results. Anyway it is interesting to present the next proposition which relates the continuous time structure and the discrete time structure when **RS2** holds (even if not all the conditions in **RS2** are needed).

PROPOSITION 3.1 *Assume* **RS2** *holds. Then for all* $i, j \in E$ *we have*

$$p_{ij} = \mathbb{E}[(e^{T_1\Lambda})_{ij}] \tag{5}$$

and, for $\alpha \in \mathbb{R}$ *such that* $\kappa_{ij}^{(d)}(\alpha) < \infty$*, we have*

$$\kappa_{ij}^{(d)}(\alpha) = \log(\frac{\mathbb{E}[(e^{T_1 G(\alpha)})_{ij}]}{\mathbb{E}[(e^{T_1\Lambda})_{ij}]}) \tag{6}$$

Proof. Let $i, j \in E$ and $\alpha \in \mathbb{R}$ be arbitrarily fixed, with $\kappa_{ij}^{(d)}(\alpha) < \infty$. Then, since $S_1^{(d)} = S_{T_1}$, by (3) we have

$$e^{\kappa_{ij}^{(d)}(\alpha)} p_{ij} = \mathbb{E}[e^{\alpha S_1^{(d)}} 1_{J_1^{(d)}=j} | J_0^{(d)} = i] = \mathbb{E}[e^{\alpha S_{T_1}} 1_{J_{T_1}=j} | J_0 = i] = \frac{\mathbb{E}[e^{\alpha S_{T_1}} 1_{J_{T_1}=j} 1_{J_0=i}]}{P(J_0 = i)};$$

moreover, since $((J_t, S_t))$ and T_1 are independent, we obtain

$$e^{\kappa_{ij}^{(d)}(\alpha)} p_{ij} = \frac{\mathbb{E}[\mathbb{E}[e^{\alpha S_t} 1_{J_t=j} 1_{J_0=i}]_{t=T_1}]}{P(J_0 = i)} =$$

$$= \frac{\mathbb{E}[\mathbb{E}[e^{\alpha S_t} 1_{J_t=j} | J_0 = i] P(J_0 = i)_{t=T_1}]}{P(J_0 = i)} = \mathbb{E}[\mathbb{E}[e^{\alpha S_t} 1_{J_t=j} | J_0 = i]_{t=T_1}].$$

In conclusion, by (1), we have

$$e^{\kappa_{ij}^{(d)}(\alpha)} p_{ij} = \mathbb{E}[(e^{T_1 G(\alpha)})_{ij}] \tag{7}$$

whence we obtain (5) by setting $\alpha = 0$ and by taking into account $G(0) = \Lambda$; moreover we obtain (6) by putting (5) in (7) and noting that the value p_{ij} is strictly positive by (5). □

3.4. Random sampling and large deviations

When **RS2** holds let us consider the functions $(t \cdot \kappa)^*$ and Λ_T^* defined by

$$(t \cdot \kappa)^*(x) \equiv \sup_{\alpha \in \mathbb{R}} [x\alpha - t\kappa(\alpha)] \text{ and } \Lambda_T^*(t) \equiv \sup_{\beta \in \mathbb{R}} [t\beta - \Lambda_T(\beta)].$$

LEMMA 3.2 *Assume that* **RS2** *holds. Then* $((J_n^{(d)}, S_n^{(d)}))$ *is a discrete time MAP with* $E = \{1, \dots, s\}$*,* (J_n) *is an irreducible Markov chain on* E *and we have the following identity concerning* $\kappa_{(d)}$*:*

$$\kappa_{(d)}(\alpha) \equiv \Lambda_T(\kappa(\alpha)). \tag{8}$$

Proof. First of all $((J_n^{(d)}, S_n^{(d)}))$ is a discrete time MAP with $E = \{1, \ldots, s\}$ by Proposition 2.2; moreover (J_n) is obviously an irreducible Markov chain on E (in particular this follows from the first statement in Proposition 3.1, eq. (5)).
Now, in view of what follows, let $\alpha \in \mathbb{R}$, $t \geq 0$ and the distribution of J_0 be arbitrarily fixed. Then

$$\mathbb{E}[e^{\alpha S_t}] = \sum_{i,j \in E} P(J_0 = i)(e^{tG(\alpha)})_{ij}$$

follows from (1) and, by also taking into account (2), we obtain

$$\mathbb{E}[e^{\alpha S_t}] \geq \sum_{i,j \in E} P(J_0 = i)(e^{tG(\alpha)})_{ij} \frac{h_j(\alpha)}{\max_{k \in E} h_k(\alpha)} = e^{t\kappa(\alpha)} \frac{\sum_{i \in E} P(J_0 = i) h_i(\alpha)}{\max_{k \in E} h_k(\alpha)}$$

and

$$\mathbb{E}[e^{\alpha S_t}] \leq \sum_{i,j \in E} P(J_0 = i)(e^{tG(\alpha)})_{ij} \frac{h_j(\alpha)}{\min_{k \in E} h_k(\alpha)} = e^{t\kappa(\alpha)} \frac{\sum_{i \in E} P(J_0 = i) h_i(\alpha)}{\min_{k \in E} h_k(\alpha)};$$

thus we have

$$c_1(\alpha) e^{t\kappa(\alpha)} \leq \mathbb{E}[e^{\alpha S_t}] \leq c_2(\alpha) e^{t\kappa(\alpha)} \tag{9}$$

with

$$c_1(\alpha) = \frac{\sum_{i \in E} P(J_0 = i) h_i(\alpha)}{\max_{k \in E} h_k(\alpha)} \text{ and } c_2(\alpha) = \frac{\sum_{i \in E} P(J_0 = i) h_i(\alpha)}{\min_{k \in E} h_k(\alpha)}.$$

Now let $n \in \mathbb{N}$ be arbitrarily fixed; thus we obtain

$$c_1(\alpha) \mathbb{E}[e^{T_n \kappa(\alpha)}] \leq \mathbb{E}[e^{\alpha S_n^{(d)}}] \leq c_2(\alpha) \mathbb{E}[e^{T_n \kappa(\alpha)}]$$

since we have $\mathbb{E}[e^{\alpha S_n^{(d)}}] = \mathbb{E}[e^{\alpha S_{T_n}}] = \mathbb{E}[\mathbb{E}[e^{\alpha S_t}]_{t=T_n}]$ and by (9). Then we have

$$c_1(\alpha) e^{n\Lambda_T(\kappa(\alpha))} \leq \mathbb{E}[e^{\alpha S_n^{(d)}}] \leq c_2(\alpha) e^{n\Lambda_T(\kappa(\alpha))}$$

whence we obtain the limit

$$\Lambda_T(\kappa(\alpha)) = \lim_{n \to \infty} \frac{1}{n} \log \mathbb{E}[e^{\alpha S_n^{(d)}}]$$

so that (8) follows from (4). □

REMARK 3.3 *Assume that* **RS2** *holds and let us consider the derivative of* $\kappa_{(d)}$ *in (8):*

$$\kappa'_{(d)}(\alpha) \equiv \Lambda'_T(\kappa(\alpha))\kappa'(\alpha). \tag{10}$$

Thus, by putting $\alpha = 0$ *in (10) and by* $\kappa(0) = 0$, *we have*

$$\kappa'_{(d)}(0) = \Lambda'_T(0)\kappa'(0).$$

Then this formula has some analogy with Wald's identity; indeed $\Lambda'_T(0) = \mathbb{E}[T_1]$ *and, if the distribution of* J_0 *is the stationary distribution of* J, $\kappa'(0)$ *and* $\kappa'_{(d)}(0)$ *are the means of* S_1 *and* $S_1^{(d)} = S_{T_1}$ *respectively.*

PROPOSITION 3.4 *Assume that* **RS2** *holds. Then:*
(i) $((\frac{S_n^{(d)}}{n}, \frac{T_n}{n}))$ *satisfies the large deviations principle with rate function*

$$I^{(d)}_{(S,T)}(x,t) = (t\cdot\kappa)^*(x) + \Lambda^*_T(t); \tag{11}$$

(ii) as far as the large deviations rate function $\kappa^*_{(d)}$ *concerning* $(\frac{S_n^{(d)}}{n})$, *we have*

$$\kappa^*_{(d)}(x) = \inf_{t\geq 0}[(t\cdot\kappa)^*(x) + \Lambda^*_T(t)]. \tag{12}$$

Proof. Part (ii) immediately follows from (i) and contraction principle (see e.g. [4], Theorem 4.2.1, page 110), indeed $\Lambda^*_T(t) = \infty$ for $t < 0$. Thus we only need to prove (i) and the idea is to employ Gärtner Ellis Theorem. Let $\alpha, \beta \in \mathbb{R}$ be arbitrarily fixed and we have

$$\mathbb{E}[e^{\alpha S_n^{(d)} + \beta T_n}] = \mathbb{E}[\mathbb{E}[e^{\alpha S_t}]_{t=T_n} e^{\beta T_n}]$$

whence we obtain

$$c_1(\alpha)\mathbb{E}[e^{T_n(\kappa(\alpha)+\beta)}] \leq \mathbb{E}[e^{\alpha S_n^{(d)} + \beta T_n}] \leq c_2(\alpha)\mathbb{E}[e^{T_n(\kappa(\alpha)+\beta)}]$$

by (9) (as in the proof of Lemma 3.2). In conclusion we have

$$c_1(\alpha)e^{n\Lambda_T(\kappa(\alpha)+\beta)} \leq \mathbb{E}[e^{\alpha S_n^{(d)} + \beta T_n}] \leq c_2(\alpha)e^{n\Lambda_T(\kappa(\alpha)+\beta)}$$

and we obtain the following limit

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[e^{\alpha S_n^{(d)} + \beta T_n}] = \Lambda_T(\kappa(\alpha) + \beta).$$

Then, by **RS2**, $((\frac{S_n^{(d)}}{n}, \frac{T_n}{n}))$ satisfies the large deviations principle with rate function $I_{(S,T)}^{(d)}$ defined by

$$I_{(S,T)}^{(d)}(x,t) = \sup_{(\alpha,\beta)\in\mathbb{R}^2} [x\alpha + t\beta - \Lambda_T(\kappa(\alpha)+\beta)]$$

as a consequence of Gärtner Ellis Theorem and, in conclusion, we have to check the identity (11). In order to do that let $(x,t) \in \mathbb{R}^2$ be arbitrarily fixed and let us check the two following complementary inequalities. The first one is

$$I_{(S,T)}^{(d)}(x,t) = \sup_{(\alpha,\beta)\in\mathbb{R}^2} [x\alpha - t\kappa(\alpha) + t(\kappa(\alpha)+\beta) - \Lambda_T(\kappa(\alpha)+\beta)] \leq$$

$$\leq \sup_{(\alpha,\beta)\in\mathbb{R}^2} [x\alpha - t\kappa(\alpha)] + \sup_{(\alpha,\beta)\in\mathbb{R}^2} [t(\kappa(\alpha)+\beta) - \Lambda_T(\kappa(\alpha)+\beta)] = (t\cdot\kappa)^*(x) + \Lambda_T^*(t).$$

For the second inequality there exist two sequences (α_n) and (β_n) such that

$$(t\cdot\kappa)^*(x) = \lim_{n\to\infty} x\alpha_n - t\kappa(\alpha_n) \text{ and } \Lambda_T^*(t) = \lim_{n\to\infty} t\beta_n - \Lambda_T(\beta_n);$$

then, for each fixed $n \in \mathbb{N}$, we have

$$I_{(S,T)}^{(d)}(x,t) \geq x\alpha_n + t(\beta_n - \kappa(\alpha_n)) - \Lambda_T(\kappa(\alpha_n) + (\beta_n - \kappa(\alpha_n))) =$$

$$= x\alpha_n - t\kappa(\alpha_n) + t\beta_n - \Lambda_T(\beta_n)$$

and, by taking the limit as $n \to \infty$, we obtain

$$I_{(S,T)}^{(d)}(x,t) \geq (t\cdot\kappa)^*(x) + \Lambda_T^*(t). \ \square$$

3.5. Random sampling and level crossing probabilities

Assume **RS2** holds. Furthermore let us consider the level crossing probabilities

$$\psi(b) = P(\sup_{t\geq 0} S_t > b) \ (\forall b > 0).$$

We have trivial cases when (S_t) has nonincreasing paths or when (S_t) diverges to $+\infty$; indeed, for all $b > 0$, we have $\psi(b) = 0$ or $\psi(b) = 1$ respectively. Then, in order to avoid these trivial cases, it is useful to refer to the following condition:

$$\text{there exists } w > 0 \text{ such that } \kappa(w) = 0 \text{ and } \kappa'(w) > 0. \qquad (13)$$

We have a similar situation for $((J_n^{(d)}, S_n^{(d)}))$. Thus we can consider the level crossing probabilities

$$\psi_{(d)}(b) = P(\sup_{n \in \mathbb{N}} S_n^{(d)} > b) \ (\forall b > 0)$$

and the following condition:

$$\text{there exists } w_{(d)} > 0 \text{ such that } \kappa_{(d)}(w_{(d)}) = 0 \text{ and } \kappa'_{(d)}(w_{(d)}) > 0. \tag{14}$$

Then the following limits provide an exponential decay for the level crossing probabilities (and the values w and $w_{(d)}$ are the so-called Lundberg parameters):

$$\lim_{b \to \infty} \frac{1}{b} \log \psi(b) = -w \text{ when (13) holds}$$

and

$$\lim_{b \to \infty} \frac{1}{b} \log \psi_{(d)}(b) = -w_{(d)} \text{ when (14) holds;}$$

see Theorem 3.1 in [7] for the discrete time case and the proof can be adapted to the continuous time case (see [8], Proposition 5.1).

Now let us illustrate a technique called importance sampling for the estimation of the level crossing probabilities by Monte Carlo simulations. We shall refer to the discrete time case (see e.g. subsection 2.1 in [7] for a complete description) and we have a similar description for the continuous time case (see e.g. section 5 in [8]); moreover the reader can find a general description of the use of large deviations theory for the estimation the probability of rare events by importance sampling in the first part of Introduction in [3].

Let us consider some independent simulations under P and we want to estimate the level crossing probability $\psi_{(d)}(b)$ by the relative frequency of the level crossings, which is an unbiased estimator; then the number of replications needs to grow exponentially with b to keep a fixed relative precision. Moreover the simulation time is not finite when the level crossing does not occur.

In order to overcome these troubles the importance sampling technique suggests to consider independent simulation under another law Q and another unbiased estimator which is a suitable weighted mean. Moreover Q has to be chosen in a class of admissible laws such that the level crossing occurs with probability 1 and therefore the simulation time is always finite. In general, for any admissible law Q, we have

$$\liminf_{b \to \infty} \frac{1}{b} \log \eta_Q(b) \geq -2w_{(d)}$$

where $\eta_Q(b)$ is a suitable second moment associated to Q. Then there exists a unique admissible law Q^* such that the lower bound is attained (see Theorem 3.2 in [7]) and in such a case the variance is minimized in the sense of our interest. More precisely, if we simply write $\eta_{(d)}(b)$ in place of $\eta_{Q^*}(b)$, we can conclude with the following statement (see eq. (3.4) in [7]):

$$\lim_{b\to\infty}\frac{1}{b}\log\eta_{(d)}(b) = -2w_{(d)} \text{ when (14) holds.}$$

An analogous procedure can be considered for the continuous time case and, if we denote by $\eta(b)$ the second moment which plays the role of $\eta_{(d)}(b)$ in the discrete time case, we have (see e.g. [8], Proposition 5.2)

$$\lim_{b\to\infty}\frac{1}{b}\log\eta(b) = -2w \text{ when (13) holds.}$$

Now we can prove the results concerning level crossing probabilities. Roughly speaking we show that the Lundberg parameter of the continuous time Markov additive process coincides with the Lundberg parameter of the randomly sampled discrete time Markov additive process. In some sense this fact is not surprising; indeed the idea is that the exponential decay of $\psi_{(d)}$ as $b\to\infty$ is determined by the behaviour of the continuous time process (S_t) only, while the distribution of the inter-observations times $(T_n - T_{n-1})$ has no influence.

PROPOSITION 3.5 *Assume that* **RS2** *and (13) hold. Then we have*

$$\lim_{b\to\infty}\frac{1}{b}\log\psi_{(d)}(b) = -w \text{ and } \lim_{b\to\infty}\frac{1}{b}\log\eta_{(d)}(b) = -2w.$$

Proof. By taking into account the results above and the preliminaries on level crossing probabilities in this subsection, it sufficies to show that (14) holds with $w_{(d)} = w$. This can be immediately checked by setting $\alpha = w$ in (8) and in (10); indeed we have

$$\kappa_{(d)}(w) = \Lambda_T(\kappa(w)) = \Lambda_T(0) = 0 \text{ and } \kappa'_{(d)}(w) = \Lambda'_T(0)\kappa'(w) > 0$$

by (13) and by noting that $\Lambda'_T(0) = \mathbb{E}[T_1] > 0$. □

Acknowledgements

I thank Brian McGurk and an anonymous referee. Brian McGurk sent me the file of his Ph. D. thesis cited in the references and the comments of the referee led to an improvement of the presentation of the paper.

References

[1] S. Asmussen, Applied Probability and Queues, John Wiley and Sons, Chichester, 1987.

[2] S. Asmussen, Ruin Probabilities, World Scientific Publishing Co., Inc., River Edge, NJ, 2000.

[3] P. Baldi, M. Piccioni, Importance sampling for continuous time Markov chains and applications to fluid models. Methodol. Comput. Appl. Probab. 1 (1999), no. 4, 375–390.

[4] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications, Jones and Bartlett Publishers, Boston, MA, 1993.

[5] J. Grandell, Aspects of Risk Theory, Springer-Verlag, New York, 1991.

[6] I. Iscoe, P. Ney, E. Nummelin, Large deviations of uniformly recurrent Markov additive processes, Adv. in Appl. Math. 6 (1985), no. 4, 373–412.

[7] T. Lehtonen, H. Nyrhinen, On asymptotically efficient simulation of ruin probabilities in a Markovian environment, Scand. Actuarial J., 1992, no. 1, 60–75.

[8] C. Macci, Continuous time Markov additive processes: composition of large deviations principles and comparison between exponential rates of convergence, J. Appl. Probab. 38 (2001), no. 4, 917–931.

[9] B. McGurk, Random sampling and large deviations, Ph. D. thesis, School of Mathematics, University of Dublin, 2000.

Claudio Macci
macci@mat.uniroma2.it
Dipartimento di Matematica
Università degli Studi di Roma "Tor Vergata"
Via della Ricerca Scientifica, I-00133 Roma, Italia.

Teoría de conjuntos según von Neumann

Antonio Montalbán

ABSTRACT

We prove that inside the theory of von Neumann (VN), there is a natural model for Zermelo-Fraenkel system (ZF) without the foundations action and we show that neither the foundation axiom nor its negation are therems in this model. We also obtain that the consistency of VN is a consequence of the consistency of $ZF + (choice\ ) + (\exists unaccesible\ cardinals)$.

RESUMEN

Probamos que, dentro de la teoría de von Neumann (VN), hay un modelo natural de la teoría de Zermelo-Fraenkel (ZF) sin el axioma de fundación y vemos que ni el axioma de fundación ni su negación son teoremas en ese modelo. Por último obtenemos que la consistencia de VN está implicada por la consistencia de $ZF + (elección) + (\exists cardinal\ inaccesible)$.

1. Introducción

En este articulo estudiamos la teoría axiomática de conjuntos que propuso von Neumann en [Neu25]. Es una axiomática muy diferente de la de Zermelo-Fraenkel. Aunque la teoría de von Neumann tiene ideas nuevas con respecto a la teoría de Zermelo-Fraenkel, la primera tiene dentro de ella un modelo natural de ZFC^- y probablemente fue pensada para garantizar la existencia de este modelo. Un estudio comparativo de la teoría de von Neumann con respecto a la teoría de Zermelo-Fraenkel se puede encontrar en [Mor98]. Una particularidad es que VN está basada en el concepto de función, mientras que ZF está basada en el concepto de pertenencia. Esto no crea diferencias profundas, ya que un conjunto se puede expresar cómo una función característica y una función se puede expresar cómo un conjunto de pares. Una de las cualidades de esta axiomática es que tiene una cantidad finita de axiomas (de hecho tiene dieciocho), mientras que ZF no es finitamente axiomatizable. Otra ventaja es que se pueden manejar clases dentro de la teoría. Estas ventajas se conservaron en la teoría, hoy llamada NGB (von Neumann, Gödel, Bernays), la cual es usada actualmente. Los fundamentos básicos de esta teoría están desarrollados, por ejemplo, en [Men97].

En la sección 2 presentamos formalmente a la teoría de von Neumann y demostramos algunos teoremas en ella. Demostramos el teorema de reducción normal que nos permitirá representar cierto tipo de formulas del lenguaje con objetos de la teoría.

En la sección 3 mostramos un modelo natural (**I**) de la teoría de Zermelo-Fraenkel con el axioma de elección y sin el de fundación (ZFC^-) dentro de la teoría de von Neumann. Esto nos muestra que en VN podemos hacer, por lo menos, todo lo que podemos hacer en ZFC^-. Hay quienes piensan que a partir de ZFC^- se puede desarrollar toda la matemática, por lo tanto, toda la matemática se podría desarrollar en VN.

En la sección 4 empezamos construyendo, en la teoría de Zermelo-Fraenkel con el axioma de elección y con el que afirma la existencia de algun cardinal inaccesible, un modelo de la teoría de von Neumann. De donde obtendremos que si esta extención de la teoría de Zermelo-Fraenkel es consistente entonces tambinén VN lo es. Una pregunta que es natural hacerse es si se puede demostrar el axioma de fundación en el modelo **I** de $\widetilde{ZFC^-}$ dentro de VN, que analizamos en la sección 3. Para esto construiremos, en ZFC, dos modelos de VN usando el modelo que ya habíamos construido. En uno de ellos se verificará fundación y en el otro se verificará su negación, obteniendo así que ni fundación ni su negación son demostrables en la estructura **I** dentro de VN.

Todo lo que presentamos en este artículo esta detalladamente explicado en [Mon00].

2. Teoría de conjuntos de von Neumann

Desarrollamos la teoría de conjuntos de von Neumann en el cálculo de predicados de primer orden con igualdad. El lenguaje, que notaremos con $\mathcal{L}^{VN}$, contiene un símbolo de predicado $\mathbf{I}(\cdot)$, dos símbolos de constante $\bot$ y $\top$, y dos símbolos de función $\langle\cdot,\cdot\rangle$ y $\cdot[\cdot]$. Cuando escribimos $\mathbf{I}(x)$ decimos "x es un uno-objeto". Intuitivamente estos uno-objetos van a ser los argumentos y los resultados de las funciones, y también van a ser funciones. Los objetos que no son uno-objetos se caracterizan por ser, en cierto sentido, muy grandes, como por ejemplo lo son las clases propias en la teoría de conjuntos usual. Llamamos *bottom* y *top* a los símbolos de constantes $\bot$ y $\top$, que representan falso y verdadero respectivamente. Cuando escribimos $\langle x, y\rangle$ decimos "el par x y" y lo interpretamos como el par ordenado formado por x e y. Cuando escribimos $x[y]$ decimos "y aplicado a x" o "x evaluado en y", a y le llamamos el *argumento* de la aplicación, y lo interpretamos como la función x evaluada en el valor y.

Los axiomas de la teoría de conjuntos de von Neumann están divididos en cinco grupos:

$VN.I$: axiomas introductorios

$VN.II$: axiomas de construcciones aritméticas
$VN.III$: axiomas de construcciones lógicas
$VN.IV$: axioma sobre los uno-objetos
$VN.V$: axiomas de infinitud.

Describimos a continuación cada uno de los grupos de axiomas.

2.1. Grupo I. Axiomas introductorios

Utilizamos $\forall^{\mathbb{M}}x\,\varphi$ en lugar de $\forall x(\mathbb{M}(x) \to \varphi)$ y $\exists^{\mathbb{M}}x\varphi$ en lugar de $\exists x(\mathbb{M}(x) \wedge \varphi)$.

$VN.I,1 : \mathbf{I}(\bot) \wedge \mathbf{I}(\top) \wedge \bot \neq \top$
$VN.I,2 : \forall f \forall^{\mathbf{I}} x(\mathbf{I}(f[x]))$
$VN.I,3 : \forall^{\mathbf{I}} xy(\mathbf{I}\langle x, y\rangle)$
$VN.I,4 : \forall fg(\forall^{\mathbf{I}} x(f[x] = g[x]) \to f = g)$

Se podría decir que los primeros tres axiomas se dedican a definir "el tipo" de las funciones del lenguaje. El cuarto axioma caracteriza la igualdad: dos elementos son iguales si son iguales como funciones. Este axioma es el que nos permite pensar a los objetos como funciones, y al símbolo de función $\cdot[\cdot]$ como la aplicación. El segundo y el cuarto axioma permiten decir que todos los objetos son funciones que van desde los uno-objetos a los uno-objetos.

En todos los axiomas de la teoría, cada vez que aparece una aplicación su segunda entrada es un uno-objeto y cada vez que aparece un par sus dos entradas son uno-objetos. Esto implica que en ningún resultado relevante de la teoría aparecerán aplicaciones o pares fuera de estas condiciones. Para la definición de estos términos consideraremos un conjunto de variables X y definiremos la clase de los términos que están, en el sentido de la oración anterior, bien formados cuando suponemos que todos los elementos de X son uno-objetos.

La definición formal de esta clase de términos es la siguiente:

Definición 2.1. Sea X un conjunto de variables, definimos por inducción al conjunto de términos *bien formados sobre* X (*tbf*(X)):

- $\bot$ y $\top$ son *tbf*(X) para cualquier conjunto de variables X.
- x es un *tbf*(X) si $x \in X$.
- $\langle \mathfrak{u}_1, \mathfrak{u}_2\rangle$ y $\mathfrak{u}_1[\mathfrak{u}_2]$ son *tbf*(X) si $\mathfrak{u}_1$ y $\mathfrak{u}_2$ son *tbf*(X).
- $y[\mathfrak{u}]$ es un *tbf*(X) si $\mathfrak{u}$ es un *tbf*(X) e y es una variable cualquiera.

Teorema 2.2. Dado $\mathfrak{u}$ un *tbf*($\{x_1, \ldots, x_n\}$), tenemos que

$$VN.I \vdash \forall^{\mathbf{I}} x_1 \ldots x_n \mathbf{I}(\mathfrak{u})$$

DEMOSTRACIÓN: Se demuestra fácilmente por inducción en u, un *tbf*(X). □

Observemos que todo término es bien formado sobre su conjunto de variables, pero lo interesante es que un término puede ser bien formado sobre conjuntos menores.

2.2. Grupo II. Axiomas de construcciones aritméticas

$VN.II{,}1 : \exists a(\forall^{\mathbf{I}} x(a[x] = x))$
$VN.II{,}2 : \forall^{\mathbf{I}} u \exists a \forall^{\mathbf{I}} x(a[x] = u)$
$VN.II{,}3 : \exists a \forall^{\mathbf{I}} xy(a[\langle x, y\rangle] = x)$
$VN.II{,}4 : \exists a \forall^{\mathbf{I}} xy(a[\langle x, y\rangle] = y)$
$VN.II{,}5 : \exists a \forall^{\mathbf{I}} xy(a[\langle x, y\rangle] = x[y])$
$VN.II{,}6 : \forall ab \exists c \forall^{\mathbf{I}} x(c[x] = \langle a[x], b[x]\rangle)$
$VN.II{,}7 : \forall ab \exists c \forall^{\mathbf{I}} x(c[x] = a[b[x]])$

El grupo II de axiomas asegura la existencia de ciertas funciones que son la base para construir -como veremos más adelante- muchas otras. Estos axiomas aseguran la existencia de las funciones: identidad, constante igual a u para cada uno-objeto u, las funciones proyección que aplicadas a un par devuelven una de sus componentes, la función aplicación que dado un par de uno-objetos da la primera componente aplicada a la segunda, las funciones que devuelven pares dando las funciones coordenadas y la función compuesta de dos funciones.

De los axiomas $VN.II{,}3$ y $VN.II{,}4$ se deduce que dos pares son iguales si y solo si sus componentes lo son. Esto nos permite pensar a $\langle x, y\rangle$ como lo que usualmente es el par ordenado x y.

Definiremos un símbolo de función o de constante para cada axioma del grupo II. Estos axiomas nos permiten definir estos símbolos sin agregar nada nuevo a la teoría . Luego definiremos, como si programáramos en un lenguaje de programación funcional, nuevas funciones a partir de las ya obtenidas. Recordamos que para definir un símbolo de función f tal que $\forall x \forall a(a = \mathsf{f}(x) \leftrightarrow \varphi(x, a))$ es necesario tener probado $\forall x \exists! a \varphi(x, a)$. Para definir un símbolo de constante c tal que $\forall a(a = \mathsf{c} \rightarrow \varphi(a))$ sólo es necesario tener probado $\exists a \varphi(a)$.

Definición 2.3. Definimos los símbolos de función id, cte$(\cdot)$, par$(\cdot,\cdot)$ y comp$(\cdot,\cdot)$ y los símbolos de constantes pr_1, pr_2 y ap tales que satisfacen:
- $\forall a(a = \mathsf{id} \leftrightarrow \forall^{\mathbf{I}} x(a[x] = x))$
- $\forall^{\mathbf{I}} u \forall a(a = \mathsf{cte}(u) \leftrightarrow \forall^{\mathbf{I}} x(a[x] = u))$
- $\forall^{\mathbf{I}} xy(\mathsf{pr}_1[\langle x, y\rangle] = x)$
- $\forall^{\mathbf{I}} xy(\mathsf{pr}_2[\langle x, y\rangle] = y)$
- $\forall^{\mathbf{I}} xy(\mathsf{ap}[\langle x, y\rangle] = x[y])$

- $\forall bc \forall a(a = \mathsf{par}(b, c) \leftrightarrow \forall^{\mathbf{I}} x(a[x] = \langle b[x], c[x] \rangle))$
- $\forall bc \forall a(a = \mathsf{comp}(b, c) \leftrightarrow \forall^{\mathbf{I}} x(a[x] = b[c[x]]))$

A continuación fijamos algunas notaciones que nos permitirán considerar n-uplas y funciones de n argumentos.

Definición 2.4.
$\langle x_1, \ldots, x_n \rangle := \langle \langle x_1, \ldots, x_{n-1} \rangle, x_n \rangle$
$a[x_1, \ldots, x_n] := a[\langle x_1, \ldots, x_n \rangle]$

Un importante resultado que se deduce a partir de $VN.I, II$ es que dada una expresión $\mathfrak{u}(y_1, \ldots, y_n)$ construida a partir de las variables $y_1, \ldots, y_n, z_1, \ldots, z_m$, de las constantes $\bot$ y $\top$ y usando las funciones $\langle \cdot, \cdot \rangle$ y $\cdot[\cdot]$, tenemos que existe un objeto a tal que, para cualesquiera uno-objetos $y_1, \ldots, y_n$, $\mathfrak{u}(y_1, \ldots, y_n) = a[y_1, \ldots, y_n]$. Este término $\mathfrak{u}$ tiene que estar formado correctamente.

Teorema de Reducción 2.5. Dado $\mathfrak{u}$, un *tbf*$(\{y_1, \ldots, y_m, x_1, \ldots, x_n\})$ tenemos que:

$$VN.I, II \vdash \forall^{\mathbf{I}} y_1 \ldots y_m \exists a \forall^{\mathbf{I}} x_1 \ldots x_n (a[x_1, \ldots, x_n] = \mathfrak{u})$$

DEMOSTRACIÓN: La demostración se hace por inducción en $\mathfrak{u}$, un *tbf* y usando las funciones que definimos en 2.3 □

2.3. Grupo III. Axiomas de construcciones lógicas

$VN.III,1 : \exists a \forall^{\mathbf{I}} xy(a[\langle x, y \rangle] \neq \bot \leftrightarrow x = y)$
$VN.III,2 : \forall b \exists a \forall^{\mathbf{I}} x(a[x] \neq \bot \leftrightarrow \forall^{\mathbf{I}} y(b[\langle x, y \rangle] = \bot))$
$VN.III,3 : \forall b \exists a \forall^{\mathbf{I}} x(\exists!^{\mathbf{I}} y(b[\langle x, y \rangle] \neq \bot) \rightarrow b[\langle x, a[x] \rangle] \neq \bot)$

Los axiomas del grupo III nos aseguran la existencia de funciones que nos permiten asignar a cada frase un objeto que se comporte de la misma forma. Para interpretar estos axiomas pensemos que bottom significa falso y ser distinto de bottom verdadero. El primer axioma asegura la existencia de una función que dice si dos elementos son iguales o no. El segundo asegura la existencia de una función que se comporta como un "para todo no". Finalmente, el tercero permite que el y que obtenemos por una condición implícita $b[\langle x, y \rangle] \neq \bot$ se obtenga explícitamente como $y = a[x]$. Así como los axiomas del grupo II, los del grupo III nos permiten introducir nuevos símbolos de función. Por ejemplo, el axioma $VN.III,1$ nos permite definir al símbolo de constante ig de *igualdad* que satisface: $\forall^{\mathbf{I}} xy(x = y \leftrightarrow \mathsf{ig}[x, y] \neq \bot)$

Antes de ver que se deduce de este grupo definiremos algunos predicados.

Definición 2.6.

$x \epsilon a := \mathbf{I}(x) \wedge a[x] \neq \bot$ $\qquad$ $a \sqsubseteq b := \forall^{\mathbf{I}} x(x \epsilon a \rightarrow x \epsilon b)$
$clase(a) := \forall^{\mathbf{I}} x(a[x] = \bot \vee a[x] = \top)$ $\qquad$ $a \sim b := a \sqsubseteq b \wedge b \sqsubseteq a$
$conj(a) := \mathbf{I}(a) \wedge clase(a)$ $\qquad$ $a \sqsubset b := a \sqsubseteq b \wedge \neg a \sim b$

La definición de $x \epsilon a$ significa intuitivamente que x está en el dominio de a, y $a \sqsubseteq b$ significa que el domino de a está incluido en el de b. Decimos que un objeto es una clase si solamente toma valores bottom o top, lo que estamos haciendo es representar las clases con su función característica. Y un objeto es un conjunto si su función característica es un uno-objeto.

Demostraremos que para una cierta clase de fórmulas existe una clase (en el sentido de la definición 2.6) que contiene a todos los uno-objetos que satisfacen dicha fórmula.

A continuación definimos el conjunto de fórmulas que podremos representar con objetos de la teoría.

Definición 2.7. Definimos por inducción al conjunto de fórmulas **I**-*reducibles sobre* X, donde X es un conjunto de variables:

- $\mathfrak{u}_1 = \mathfrak{u}_2$ es **I**-reducible sobre X si $\mathfrak{u}_1$ y $\mathfrak{u}_2$ son *tbf*(X).
- $\mathbf{I}(\mathfrak{u})$ es **I**-reducible sobre X si $\mathfrak{u}$ es un *tbf*(X).
- $\varphi \vee \psi$ es **I**-reducible sobre X si φ y ψ son **I**-reducibles sobre X.
- $\neg\varphi$ es **I**-reducible sobre X si φ es **I**-reducible sobre X.
- Si φ es **I**-reducible sobre X, entonces $\exists^{\mathbf{I}} x\varphi$ es **I**-reducible sobre $X \setminus \{x\}$.

Obsérvese que toda fórmula $\varphi \in \mathcal{L}^{VN}$ relativizada a la clase **I** es **I**-reducible sobre *FreeVars*(φ), el conjunto de variables libres de φ, pero también puede serlo sobre algún conjunto menor.

Teorema de reducción normal 2.8. Dada φ una fórmula **I**-reducible sobre el conjunto de variables $\{y_1, \ldots, y_m, x_1, \ldots, x_n\}$ se tiene que:

$$VN.I, II, III \vdash \forall^{\mathbf{I}} y_1 \ldots y_m \exists a \forall^{\mathbf{I}} x_1, \ldots, x_n(a[x_1, \ldots, x_n] \neq \bot \leftrightarrow \varphi)$$

Demostración: La demostración se hace por inducción en φ, fórmula **I**-reducible, usando el teorema de reducción. Además hay que usar la función ig para las fórmulas de las formas $\mathfrak{u}_1 = \mathfrak{u}_2$ o $\psi \vee \chi$ y la función que define el axioma $VN.III$,2 para las fórmulas de las formas $\neg\psi$ o $\exists^{\mathbf{I}} x\psi$. □

Para el próximo teorema usaremos el axioma $VN.III$,3, que no ha sido usado hasta ahora.

Teorema 2.9. Dada φ una fórmula **I**-reducible sobre el conjunto de variables $\{y_1,\ldots,y_m,x_1,\ldots,x_n,y\}$ se tiene que:

$$VN.I,II,III \vdash \forall^{\mathbf{I}} y_1\ldots y_m(\forall^{\mathbf{I}} x_1,\ldots,x_n\exists!^{\mathbf{I}} y(\varphi)) \rightarrow \\ \exists a\forall^{\mathbf{I}} x_1,\ldots,x_n,y(a[x_1,\ldots,x_n]=y\leftrightarrow\varphi)$$

Demostración: Notamos con $\vec{x}$ a $x_1,\ldots,x_n$ y sean $y_1,\ldots,y_m$ uno-objetos.

Por el teorema anterior tenemos que existe b tal que$\forall^{\mathbf{I}}\vec{x},y(b[\vec{x},y]\neq\bot\leftrightarrow\varphi)$. Por la hipótesis tenemos que $\forall^{\mathbf{I}}\vec{x}\exists!^{\mathbf{I}}y(b[\vec{x},y]\neq\bot)$, luego usando el axioma $VN.III$,3 tenemos que existe a tal que $\forall^{\mathbf{I}}\vec{x},y(a[\vec{x}]=y\leftrightarrow b[\vec{x},y]\neq\bot)$. Ese a es el requerido. □

Corolario 2.10. $VN.I,II,III\vdash\forall a\exists!\bar{a}\forall^{\mathbf{I}}x((\bar{a}[x]=\bot\leftrightarrow a[x]=\bot)\wedge(\bar{a}[x]=\top\leftrightarrow a[x]\neq\bot))$

Demostración: Como $\bar{a}[x]$ está definido para cualquier uno-objeto x, la unicidad está dada por el axioma $VN.I$,4. Para la existencia consideremos la fórmula $\varphi(x,y):\ (y=\bot\leftrightarrow a[x]=\bot)\wedge(y=\top\leftrightarrow a[x]\neq\bot)$. Como necesariamente $a[x]=\bot\vee a[x]\neq\bot$ tenemos que $\forall^{\mathbf{I}}x\exists!^{\mathbf{I}}y(\varphi(x,y))$, luego $\bar{a}$ esta dada por el teorema. □

Dado un objeto a, pensamos a su dominio como la clase de todos los uno-objetos x tales que $a[x]\neq\bot$. Este corolario nos permite definir al dominio de una función de la siguiente manera:

Definición 2.11. El símbolo de función unario $\mathsf{Dom}(\cdot)$ se define tal que:

$$\forall a\forall^{\mathbf{I}}x((\mathsf{Dom}(a)[x]=\bot\leftrightarrow a[x]=\bot)\wedge(\mathsf{Dom}(a)[x]=\top\leftrightarrow a[x]\neq\bot))$$

Ya que de las definiciones de Dom y *clase* se obtiene que:

$$\forall a(clase(\mathsf{Dom}(a))\wedge\forall^{\mathbf{I}}x(x\in\mathsf{Dom}(a)\leftrightarrow a[x]\neq\bot)),$$

queda claro que lo que hace $\mathsf{Dom}(a)$ es darnos el dominio de la función a.

El siguiente teorema nos asegura la existencia de una clase a de n-uplas de uno-objetos que satisfacen una propiedad φ.

Teorema 2.12. Dada φ una fórmula **I**-reducible sobre el conjunto de variables $\{y_1,\ldots,y_m,x_1,\ldots,x_n\}$, se tiene que:

$$VN.I,II,III\vdash\forall^{\mathbf{I}}y_1\ldots y_m\exists!^{clase}a\forall^{\mathbf{I}}z(z\in a\leftrightarrow\exists^{\mathbf{I}}\vec{x}(z=\langle\vec{x}\rangle\wedge\varphi))$$

Demostración: Como $a[z]$ esta definido para cualquier uno-objeto z, el axioma $VN.I$,4 garantiza la unicidad. Por el teorema de reducción normal 2.8 tenemos que existe b tal que $\forall^{\mathbf{I}} z(z \in b \leftrightarrow \exists^{\mathbf{I}} \vec{x}(z = \langle \vec{x} \rangle \wedge \varphi))$. Alcanza con tomar $a := \mathsf{Dom}(b)$, luego $x \in a \leftrightarrow x \in b \leftrightarrow \varphi$ y $clase(a)$. □

Definición 2.13. Llamamos a esa clase $[\![\varphi]\!]_{\vec{x}}$, o sea que $[\![\varphi]\!]_{\vec{x}}$ es tal que $clase([\![\varphi]\!]_{\vec{x}})$ y $\forall^{\mathbf{I}} \vec{x}(\langle \vec{x} \rangle \in [\![\varphi]\!]_{\vec{x}} \leftrightarrow \varphi(\vec{x}))$. Escribimos simplemente $[\![\varphi]\!]$ si esta claro cuales son $\vec{x}$.

Ahora el trabajo con clases y conjuntos se torna mucho más simple.

Definición 2.14. Notamos $[\![\mathbf{I}]\!]$ a $[\![\mathbf{I}(x)]\!]_x$, la clase de todos los uno-objetos. Va a cumplir que $\forall^{\mathbf{I}} x(x \in [\![\mathbf{I}]\!])$.

Definimos $\mathsf{Im}(b, a) := [\![\exists^{\mathbf{I}} y(y \in a \wedge b[y] = x)]\!]_x$

2.4. Grupo IV. Uno-objetos

$VN.IV,1 : \forall a(\neg \mathbf{I}(a) \leftrightarrow \exists b \forall^{\mathbf{I}} x \exists^{\mathbf{I}} y(a[y] \neq \bot \wedge b[y] = x))$

El axioma $VN.IV$,1 es muy importante en la teoría ya que nos da una condición necesaria y suficiente para que un objeto sea o no un uno-objeto. Intuitivamente dice que un objeto, mirado como función, es un uno-objeto si y sólo si su dominio no es demasiado grande.

Escribamos el axioma del grupo IV usando los símbolos de función que definimos:

$$VN.IV,1 : \ \forall a(\neg \mathbf{I}(a) \leftrightarrow \exists b(\mathsf{Im}(b, a) = [\![\mathbf{I}]\!]))$$

Lo que dice es que a no es un uno-objeto si y sólo si existe una función $b : a \rightarrow [\![\mathbf{I}]\!]$ que sea sobreyectiva, que se puede interpretar como que el cardinal de a (o de $\mathsf{Dom}(a)$) es igual al de la clase de todos los uno-objetos.

2.5. Grupo V. Axiomas de infinitud

$VN.V,1 : \exists^{\mathbf{I}} a(\exists^{\mathbf{I}} x(x \in a) \wedge \forall^{\mathbf{I}} x(x \in a \rightarrow \exists^{\mathbf{I}} y(x \sqsubset y \wedge y \in a)))$

$VN.V,2 : \forall^{\mathbf{I}} b \exists^{\mathbf{I}} a \forall^{\mathbf{I}} x(\exists^{\mathbf{I}} y(x \in y \wedge y \in b) \rightarrow x \in a)$

$VN.V,3 : \forall^{\mathbf{I}} b \exists^{\mathbf{I}} a \forall^{\mathbf{I}} x(x \sqsubseteq b \rightarrow \exists^{\mathbf{I}} y(y \sim x \wedge y \in a))$

El primer axioma asegura la existencia de un uno-objeto con dominio infinito. El segundo asegura la existencia de un uno-objeto que está relacionado con la unión usual de conjuntos, es un uno-objeto cuyo dominio es la unión de los dominios de los elementos que están en el dominio de un cierto uno-objeto. El tercero asegura la existencia de un uno-objeto que está relacionado con el conjunto de partes usual de un conjunto.

3. Relación con la teoría de Zermelo-Fraenkel

El objetivo de esta sección es mostrar como existe un modelo natural de la teoría de Zermelo-Fraenkel dentro de la teoría de von Neumann.

Empezamos enunciando los axiomas de ZFC. Luego demostramos que los axiomas de ZFC^- se verifican en el modelo construido de donde deducimos un resultado de consistencia.

3.1. Teoría de Zermelo-Fraenkel

La axiomática de Zermelo-Fraenkel, que escribimos ZF, es una de las que más se usa en la actualidad. En este sección veremos un variación de ZF, que escribimos $\widetilde{ZF}$, en la cual se permite la existencia de urelementos*, usaremos ZF en la última sección.

Desarrollamos la teoría de conjuntos de Zermelo-Fraenkel en el cálculo de predicados de primer orden con igualdad. El lenguaje, que notamos con $\mathcal{L}^{\widetilde{ZF}}$, contiene dos símbolos de predicados: $\cdot \in \cdot$ que es un predicado de aridad dos que llamamos pertenece, y un símbolo un-ario conj que nos servirá para identificar a los elementos que son conjuntos.

Escribamos los axiomas:

$\widetilde{ZF}$,0 : $\exists x(x = x)$
$\widetilde{ZF}$,1 :*Extensionalidad:* $\forall^{\mathsf{conj}} xy(\forall z(z \in x \leftrightarrow z \in y) \to x = y)$
$\widetilde{ZF}$,2 :*Fundaci'on:* $\forall^{\mathsf{conj}} x(\exists y \in x \to \exists y \in x \forall z \in x(z \notin y))^{**}$
$\widetilde{ZF}$,3_φ :*Comprensi'on:* $\forall^{\mathsf{conj}} a \exists^{\mathsf{conj}} z \forall w(w \in z \leftrightarrow w \in a \wedge \varphi)^{***}$
$\widetilde{ZF}$,4 : *Pareo:* $\forall xy \exists^{\mathsf{conj}} z(x \in z \wedge y \in z)$
$\widetilde{ZF}$,5 :*Uni'on:* $\forall^{\mathsf{conj}} x(\forall y \in x(\mathsf{conj}(y)) \to \exists^{\mathsf{conj}} u \forall y \in x(y \subset u))$
$\widetilde{ZF}$,6_φ :*Reemplazo:* $\forall^{\mathsf{conj}} a((\forall x \in a \exists! y \varphi(x,y)) \to \exists^{\mathsf{conj}} b \forall x \in a \exists y \in b \varphi(x,y))$
$\widetilde{ZF}$,7 :*Infinito:* $\exists^{\mathsf{conj}} x(\emptyset \in x \wedge \forall y \in x(S(y) \in x))$
$\widetilde{ZF}$,8 :*Potencia:* $\forall^{\mathsf{conj}} a \exists^{\mathsf{conj}} b \forall^{\mathsf{conj}} y(y \subset a \to y \in b)$
$\widetilde{AC}$:*Elecci'on:* $\forall^{\mathsf{conj}} a \exists R(R \text{ bien ordena } a)$

En los axiomas $\widetilde{ZF}$,5, $\widetilde{ZF}$,7 y $\widetilde{ZF}$,8 aparecen los símbolos $\subset$, S y $\emptyset$ que no son parte del lenguaje. Definimos $x \subset y$ como una abreviación de $\forall w \in x(w \in y)$. De los axiomas $\widetilde{ZF}$,0, 1, 3 se demuestra que $\exists!^{\mathsf{conj}} c \forall x(x \notin c)$, a este

---

*Se le llama *urelementos* a los objetos de la teoría que no son conjuntos.

**Escribimos $\exists x \in a\varphi$ en lugar de $\exists x(x \in a \wedge \varphi)$ y escribimos $\forall x \in a\varphi$ en lugar de $\forall x(x \in a \to \varphi)$.

***Comprensión, al igual que Reempazo, no es un axioma solo sino que es una familia de axiomas. Hay un axioma por cada fórmula del lenguaje.

c le llamamos *conjunto vacío* y lo notamos con $\emptyset$. De los axiomas $\widetilde{ZF}$,1, 3, 4, 5 se demuestra que podemos definir $S(x) := \{x\} \cup x$.

Al axioma de elección no lo escribimos como fórmula del lenguaje porque es muy complicado de escribir y entender.

Notamos con $\widetilde{ZFC^-}$ a los axiomas de $\widetilde{ZF}$ sin el axioma 2 y con el axioma $\widetilde{AC}$.

La única diferencia entre ZF y $\widetilde{ZF}$ es que en ZF no aparece el símbolo de predicado conj. O se podría decir que sí aparece, pero que cumple que $\forall x(\mathsf{conj}(x))$, por lo tanto se trivializa y no es necesario ponerlo. Dicho de otra forma:

$$ZF = \widetilde{ZF} + \forall x(\mathsf{conj}(x))$$

3.2. $\widetilde{ZF}$ en VN

Sea $\mathbf{I}$ la estructura $< \mathbf{I}(\cdot), conj, \epsilon >$ para $\mathcal{L}^{\widetilde{ZF}}$, es decir que $\mathbf{I}$ es la estructura que tiene como dominio a los uno-objetos y que interpreta a los símbolos conj y $\in$ con los predicados $conj$ y ϵ que definimos en 2.6. Probaremos que es un modelo de $\widetilde{ZFC^-}$ en VN. Obsérvese la simplicidad de la estructura $\mathbf{I}$. El hecho de que $\mathbf{I}$ es un modelo de $\widetilde{ZFC^-}$ ya había sido enunciado por von Neumann en [Neu25] y es de suponer que él armó la teoría para que esto ocurra.

Teorema 3.1. $\mathbf{I}$ es un modelo de $\widetilde{ZFC^-}$ en VN.

DEMOSTRACIÓN: Para probar este teorema tenemos que probar, en VN, la relativización a $\mathbf{I}$ de cada uno de los axiomas de $\widetilde{ZFC^-}$.

El axioma de extensionalidad relativizado a $\mathbf{I}$, que notamos $\widetilde{ZF}$,$1^{\mathbf{I}}$, queda

$$\forall^{conj} xy(\forall^{\mathbf{I}} z(x[z] \neq \bot \leftrightarrow y[z] \neq \bot) \to x = y)$$

lo cual se deduce del axioma $VN.I$,4.

Para probar pareo hay que definir la clase $[\![z = x \vee z = y]\!]_z$ y probar que es un uno-objeto.

Las familias de axiomas Comprensión y Reemplazo se deducen de $VN.IV$,1 y de los teoremas 2.12 y 2.9.

Potencia, Unión e Infinito se deducen de los axiomas del quinto grupo. Las formulaciones de estos axiomas no son exactamente las mismas que las formulaciones de los axiomas de Potencia, Unión e Infinito de la teoría de von Neumann, por lo tanto también va a ser necesario usar la teoría que tenemos desarollada en VN y la que conocemos a partir de los axiomas de $\widetilde{ZF}$ que ya probamos.

La prueba de $\widetilde{AC}^{\mathbf{I}}$ es la más interesante: La teoría de ordinales se puede desarrollar en VN sin ninguna dificultad. Si notamos con ON a la clase (clase segun la definición 2.6) de todos los ordinales, la paradoja de Buralli-Forti nos dice $\neg conj(ON)$ y por lo tanto $\neg\mathbf{I}(ON)$. Ahora del axioma $VN.IV$,1 se deduce que existe un objeto b que como función restringida a ON es sobreyectiva sobre $[\![\mathbf{I}]\!]$. Esto nos permite definir función $a : [\![\mathbf{I}]\!] \to ON$ inyectiva y de esta forma tenemos un buen orden en $[\![\mathbf{I}]\!]$ y por lo tanto en todos sus subconjuntos. □

Corolario 3.2. $CON(VN) \Rightarrow CON(\widetilde{ZFC^-})$

Se prueba que $< \mathbb{WF}, \mathfrak{conj}, \in>$ es un modelo de ZFC dentro de $\widetilde{ZFC^-}$, donde $\mathbb{WF}$ es la clase de los conjuntos bien formados (ver [Kun80], Ch.IV §4). Por tanto $CON(\widetilde{ZFC^-}) \Rightarrow CON(ZFC)$. Si juntamos este resultado con el corolario anterior obtenemos el siguiente corolario:

Corolario 3.3. $CON(VN) \Rightarrow CON(ZFC)$

4. Modelos de la teoría de von Neumann

En esta sección demostraremos algunos resultados de consistencia relativa. Usaremos, al igual que en la sección anterior, el método de relativización.

En todo la sección trabajaremos en la teoría ZFC.

4.1. Algunas definiciones

En esta sección presentaremos algunas definiciones y demostraremos algunas proposiciones en ZFC que nos serán de utilidad más adelante.

Usaremos los cardinales inaccesibles [Kun80] para construir modelos que contengan objetos grandes como clases propias. Recordamos que en la sección 2.4 dijimos que, intuitivamente, un objeto es un uno-objeto si y sólo si su dominio no es demasiado grande, usaremos a los cardinales inaccesibles para modelar la noción de "demasiado grande".

Definición 4.1.

- Decimos que un conjunto A es *acotado* en un cardinal κ si existe un ordinal $\alpha < \kappa$ tal que $A \subset \alpha$.
- Un cardinal κ es *regular* si para todo $A \subset \kappa$ tal que $\|A\| < \kappa$, A es acotado en κ.
- Un cardinal κ es *límite fuerte* si $\forall^{\mathbb{ON}}\alpha < \kappa(\|\mathcal{P}(\alpha)\| < \kappa)$

- Decimos que un cardinal es *(fuertemente) inaccesible* si es regular, límite fuerte y no numerable.
- Notamos con EI a la frase $\exists\kappa(\kappa$ *inaccesible*$)$.

Observación 4.2.

1. Un cardinal infinito κ es regular sii para todo conjunto A tal que $\|A\| < \kappa$ y $\forall x \in A(\|x\| < \kappa)$ se tiene que $\|\bigcup A\| < \kappa$.
2. Un cardinal κ es límite fuerte sii para todo conjunto A tal que $\|A\| < \kappa$ se tiene que $\|\mathcal{P}(A)\| < \kappa$.

A estos cardinales se los llama inaccesibles porque no se pueden construir a partir de cardinales menores usando la potenciación o la unión. De hecho, se puede probar que la existencia de estos cardinales no se deduce de ZFC (el lector puede encontrar una demostración usando métodos de relativización en [Kun80], pág. 133). También se sabe que no se puede encontrar una prueba, expresable en ZFC,[****] de $CON(ZFC) \Rightarrow CON(ZFC + EI)$ (En [Kun80], pág 145, hay una demostración que usa fuertemente el teorema de Incompletitud de Gödel). A pesar de esto, en general, los matemáticos creen en la consistencia de $ZFC + EI$.

Ahora definiremos al conjunto de funciones κ-casi nulas. La idea de considerar el soporte de las funciones y de definir funciones casi nulas es bastante común en muchas áreas de la matemática.

Definición 4.3.

- $\mathcal{F}(A, B) := \{f \subset A \times B; f$ función con dominio A y codominio $B\}$
- Dada $f \in \mathcal{F}(A, B)$ llamamos *soporte* de f a $sop(f) := \{x \in A : f(x) \neq \emptyset\}$.
- Decimos que $f \in \mathcal{F}(A, B)$ es *κ-casi nula* si $\|sop(f)\| < \kappa$. Notamos $\mathcal{F}_\kappa(A, B)$ al conjunto de las funciones κ-casi nulas de $\mathcal{F}(A, B)$

Proposición 4.4. Para todo κ inaccesible, $\|\mathcal{F}_\kappa(\kappa, \kappa)\| = \kappa$

Demostración: El hecho de que $\kappa \leq \|\mathcal{F}_\kappa(\kappa, \kappa)\|$ es fácil de probar. Probaremos que $\|\mathcal{F}_\kappa(\kappa, \kappa)\| \leq \kappa$. Si tenemos $A, B \subset \kappa$ y $f \in \mathcal{F}(A, B)$ definimos $\overline{f} \in \mathcal{F}(\kappa, \kappa)$ de la siguiente forma:

$$\overline{f}(x) := \begin{cases} f(x) & \text{si } x \in A \\ \emptyset & \text{si } x \in \kappa \setminus A \end{cases}$$

[****] Todos los elementos del cálculo de predicados se pueden expresar codificados en ZFC, por lo tanto también podemos expresar resultados y demostraciones en la metateoría en ZFC.

Observemos que si $f \in \mathcal{F}_\kappa(\kappa,\kappa)$, existe un ordinal $\lambda < \kappa$ tal que $f|_\lambda \in \mathcal{F}(\lambda,\lambda)$ y $\overline{f|_\lambda} = f$. Esto es porque: como $f \in \mathcal{F}_\kappa(\kappa,\kappa)$, tenemos que $\|sop(f)\| < \kappa$ y como κ es regular existe $\lambda_1 < \kappa$ tal que $sop(f) \subset \lambda_1$, y también tenemos que $\|\{f(x) : x \in sop(f)\}\| \leq \|sop(f)\| < \kappa$ y como κ es regular existe $\lambda_2 < \kappa$ tal que $\{f(x) : x \in sop(f)\} \subset \lambda_2$. Sea $\lambda := max(\lambda_1, \lambda_2)$, luego tenemos que $f|_\lambda \in \mathcal{F}(\lambda,\lambda)$ y que $\overline{f|_\lambda} = f$. Por lo tanto, tenemos que $\mathcal{F}_\kappa(\kappa,\kappa) = \bigcup_{\lambda\in\kappa}\{\overline{f} : f \in \mathcal{F}(\lambda,\lambda)\}$. Como la unión esta tomada sobre κ, alcanza con demostrar que $\|\mathcal{F}(\lambda,\lambda)\| \leq \kappa$ para todo $\lambda < \kappa$, lo cual es inmediato de que κ es límite fuerte y de que $\|\mathcal{F}(\lambda,\lambda)\| = \|\mathcal{P}(\lambda)\|$. □

4.2. El Modelo $M(\mathcal{K},\sigma)$

En esta sección construimos la estructura $M(\mathcal{K},\sigma)$ en $ZFC + EI$ y probamos que es un modelo de VN.

Sea κ un cardinal inaccesible, sea $\mathcal{K}$ un conjunto de cardinal κ que contenga a $\emptyset$,[*****] sea $\ll\cdot,\cdot\gg : \mathcal{K}\times\mathcal{K} \to \mathcal{K}$ inyectiva y T un elemento de $\mathcal{K}$ distinto de $\emptyset$. De la proposición 4.4 deducimos que existe una función $\sigma : \mathcal{F}_\kappa(\mathcal{K},\mathcal{K}) \to \mathcal{K}$ biyectiva. Usaremos estos objetos para construir $M(\mathcal{K},\sigma)$. Notaremos con *modelo*$(\mathcal{K},\sigma,\ll,\gg,T)$ al predicado: "$\|\mathcal{K}\| = \kappa$ *es un cardinal inaccesible*, $\sigma : \mathcal{F}_\kappa \to \mathcal{K}$ *biyectiva*, $\ll\cdot,\cdot\gg : \mathcal{K}\times\mathcal{K} \to \mathcal{K}$ *inyectiva*, $\emptyset \in \mathcal{K}$, $T \in \mathcal{K}$ y $T \neq \emptyset$". Tenemos que EI implica que existen $\mathcal{K}$, σ, $\ll,\gg$ y T tales que (*modelo*$(\mathcal{K},\sigma,\ll,\gg,T)$). Notamos $\mathcal{F} := \mathcal{F}(\mathcal{K},\mathcal{K})$ y $\mathcal{F}_\kappa := \mathcal{F}_\kappa(\mathcal{K},\mathcal{K})$.

En la teoría de von Neumann todos los objetos se comportan como funciones y los uno-objetos se comportan, además, como argumentos y resultados de esas funciones. El hecho de tener una biyección entre $\mathcal{F}_\kappa$ y $\mathcal{K}$ nos permite identificar a los elementos de $\mathcal{F}_\kappa(\mathcal{K},\mathcal{K})$ con los de $\mathcal{K}$, o sea que podemos pensar a las funciones de $\mathcal{F}_\kappa$ como argumentos o resultados de las funciones de $\mathcal{F}$ (y de $\mathcal{F}_\kappa$). Siguiendo esta idea construimos el modelo $M(\mathcal{K},\sigma)$ de VN en $ZFC + EI$. Lo que hacemos es considerar $\mathcal{F}$ como nuestro universo de elementos de VN. Incluido en $\mathcal{F}$ está $\mathcal{F}_\kappa$, que usamos para representar a los uno-objetos. Representamos a bottom y top con los elementos de $\mathcal{F}_\kappa$ que se corresponden mediante σ con $\emptyset$ y con T. Para definir el pareo de elementos de $\mathcal{F}_\kappa$ usamos $\ll\cdot,\cdot\gg : \mathcal{K}\times\mathcal{K} \to \mathcal{K}$ y para definir la aplicación de un elemento de $\mathcal{F}(\mathcal{K},\mathcal{K})$ sobre un elemento de $\mathcal{F}_\kappa$ usamos la identificación entre los elementos de $\mathcal{F}_\kappa$ y los de $\mathcal{K}$. La definición formal es la siguiente:

Definición 4.5. Definimos la estructura

$$M(\mathcal{K},\sigma,\ll,\gg,T) := < \mathcal{F},\ \mathbf{I}^M(\cdot), \bot^M, \top^M, \langle\cdot,\cdot\rangle^M, \cdot[\cdot]^M >$$

[*****] Podríamos no pedir que $\emptyset \in \mathcal{K}$, elegir un elemento de $\mathcal{K}$ y usarlo en lugar de $\emptyset$ en la definición de función casi nula.

de la siguiente forma:

- $\mathbf{I}^M(x) := \quad x \in \mathcal{F}_\kappa(\mathcal{K},\mathcal{K})$
- $\bot^M := \sigma^{-1}(\emptyset)$
- $\top^M := \sigma^{-1}(T)$
- $\langle x,y\rangle^M := \sigma^{-1}(\ll\sigma(x),\sigma(y)\gg)$
- $x[y]^M := \sigma^{-1}(x(\sigma(y)))$

Eliminamos $\ll,\gg$ y T de la notación y escribimos $M(\mathcal{K},\sigma)$, simplemente M cuando es claro quienes son $\mathcal{K}$ y σ.

Estamos definiendo $\langle x,y\rangle^M$ sólo para $x,y \in \mathcal{F}_\kappa$ y $x[y]^M$ para $x \in \mathcal{F}$ y $y \in \mathcal{F}_\kappa$ cuando deberíamos definirlos para todos $x,y \in \mathcal{F}$. Solamente nos interesaran estos casos, por lo tanto alcanza con definir a las funciones $\langle\cdot,\cdot\rangle^M$ y $\cdot[\cdot]^M$ de forma arbitraria en el resto de los casos.

Es fácil verificar que en todas las definiciones σ y σ^{-1} están aplicadas a objetos de su dominio. Para definir $M(\mathcal{K},\sigma)$ como estructura tenemos que verificar que $\bot^M \in \mathcal{F}$, $\top^M \in \mathcal{F}$, $\langle x,y\rangle^M \in \mathcal{F}$ para todos $x,y \in \mathcal{F}_\kappa$ y que $x[y]^M \in \mathcal{F}$ para $x \in \mathcal{F}$ e $y \in \mathcal{F}_\kappa$, lo cual se prueba observando que $ran(\sigma^{-1}) = \mathcal{F}_\kappa \subset \mathcal{F}$.

Al relativizar la teoría VN se relativizan los símbolos de predicado y de función que definimos en ella. Notamos a los símbolos relativizados agregándole el supraíndice M. Por ejemplo, $conj(x)$ significa $\mathbf{I}(x)\wedge\forall^{\mathbf{I}}z(x[z]=\bot\vee x[z]=\top)$ luego $conj^M(x)$ significa $x \in \mathcal{F}_\kappa \wedge \forall^{\mathcal{F}_\kappa}z(x[z]^M = \bot^M \vee x[z]^M = \top^M)$ que es equivalente a $x \in \mathcal{F}_\kappa \wedge \forall^{\mathcal{K}}w(x(w)=\emptyset \vee x(w)=T)$.

De la misma forma obtenemos que $x \,\epsilon^M\, y$ es equivalente a $\sigma(x) \in sop(y)$.

Teorema 4.6. $M(\mathcal{K},\sigma)$ es un modelo de VN en ZFC+*modelo*$(\mathcal{K},\sigma,\ll,\gg,T)$.

Demostración: Para probar este teorema tenemos que probar, en ZFC, las frases que resultan de relativizar los axiomas de VN y de usar las definiciones de los símbolos $\bot^M, \top^M, \langle\cdot,\cdot\rangle^M, \cdot[\cdot]^M, \mathbf{I}^M(\cdot)$.

Para probar $VN.I{,}1^{M(\mathcal{K},\sigma)}$, $VN.I{,}2^{M(\mathcal{K},\sigma)}$ y $VN.I{,}3^{M(\mathcal{K},\sigma)}$ sólo hay que observar que la imagen de σ^{-1} es $\mathcal{F}_\kappa$ y σ es biyectiva. Mientras que para probar $VN.I{,}4^{M(\mathcal{K},\sigma)}$ hay que usar que dos funciones iguales sii lo son punto a punto.

Todos los axiomas del grupo II , relativizados a $M(\mathcal{K},\sigma)$, enuncian la existencia de ciertas funciones de $\mathcal{K}$ en $\mathcal{K}$. En todos los casos es obvio que estas funciones existen en $\mathcal{F}$. Con los axiomas del tercer grupo ocurre lo mismo que con las del segundo.

$VN.IV,1^{M(\mathcal{K},\sigma)}$ es $\forall^{\mathcal{F}}a(a \notin \mathcal{F}_\kappa \leftrightarrow \exists^{\mathcal{F}}b\forall^{\mathcal{F}_\kappa}x\exists^{\mathcal{F}_\kappa}y(\sigma(y) \in sop(a) \wedge b(\sigma(y)) = \sigma(x)))$ que es equivalente a $\forall^{\mathcal{F}}a(\neg(\|sop(a)\| < \kappa) \leftrightarrow \exists^{\mathcal{F}}b(\{b(x) : x \in sop(a)\} = \mathcal{K}))$. Es claro que si $\|sop(a)\| < \kappa$ entonces $\forall^{\mathcal{F}}b\ \|\{b(x) : x \in sop(a)\}\| \leq \|sop(a)\| < \kappa = \|\mathcal{K}\|$, luego es imposible que $\{b(x) : x \in sop(a)\} = \mathcal{K}$. Por otro lado, si $\|sop(a)\| = \kappa = \|\mathcal{K}\|$, existe una biyección $b \in \mathcal{F}$ tal que $\{b(x) : x \in sop(a)\} = \mathcal{K}$

Las fórmulas $VN.V,1^{M(\mathcal{K},\sigma)}$, $VN.V,2^{M(\mathcal{K},\sigma)}$ y $VN.V,3^{M(\mathcal{K},\sigma)}$ se deducen de que κ es no numerable, regular y límite fuerte respectivamente. □

Lo importante de este teorema es que de él se deduce el siguiente corolario.

Corolario 4.7. $CON(ZFC + EI) \Rightarrow CON(VN)$

DEMOSTRACIÓN: Sabemos que en $ZFC + EI$ existen $\mathcal{K}$, σ, $\ll,\gg$ y T tales que *modelo*$(\mathcal{K}, \sigma, \ll, \gg, T)$. Luego podemos construir el modelo $M(\mathcal{K}, \sigma)$ de VN en $ZFC + EI$. □

Tenemos probado que la consistencia de VN "está entre" la de ZFC y la de $ZFC + EI$, es decir $CON(ZFC + EI) \Rightarrow CON(VN)$ y $CON(VN) \Rightarrow CON(ZFC)$.

4.3. Sobre el axioma de fundación en la teoría de von Neumann

Recordamos que en la sección 3.2 probamos que $VN \vdash \widetilde{ZFC^-}^{\mathbf{I}}$, o sea que probamos que $\mathbf{I}$ es un modelo de $\widetilde{ZFC^-}$, pero no se hizo nada sobre si $\widetilde{ZF},2$ se satisface o no en ese modelo, donde $\widetilde{ZF},2$ es el axioma de fundación. En esta sección construimos dos modelos de VN y luego estudiamos al modelo $\mathbf{I}$ dentro de los modelos de VN, en uno de ellos se satisface $\neg\widetilde{ZF},2$ y en el otro se satisface $\widetilde{ZF},2$. De aquí deducimos que si $ZFC + EI$ es consistente, entonces ni $\widetilde{ZF},2^{\mathbf{I}}$ ni $\neg\widetilde{ZF},2^{\mathbf{I}}$ son demostrables en VN.

4.3.1. Consistencia de $VN + \neg\mathbf{I}$

Para probar que $\widetilde{ZF},2^{\mathbf{I}}$ no se deduce de VN usaremos como modelo un caso particular de $M(\mathcal{K}, \sigma)$. Lo que hacemos es cambiar la biyección σ de forma que exista algún conjunto que se pertenezca a si mismo.

Teorema 4.8. Dados κ un cardinal inaccesible, $\mathcal{K}$ tal que $\|\mathcal{K}\| = \kappa$ y $\emptyset \in \mathcal{K}$, $\ll\cdot,\cdot\gg : \mathcal{K} \times \mathcal{K} \to \mathcal{K}$ inyectiva y $T \in \mathcal{K}$ distinto de $\emptyset$ existe $\widetilde{\sigma} : \mathcal{F}_\kappa \to \mathcal{K}$ biyectiva tal que existe $x \in \mathcal{F}_\kappa$ que verifica $conj^{\widetilde{M}}(x) \wedge x \epsilon^{\widetilde{M}} x$. Donde escribimos $\widetilde{M}$ en lugar de $M(\mathcal{K}, \widetilde{\sigma})$.

Demostración: Sea $\sigma : \mathcal{F}_\kappa \to \mathcal{K}$ una función biyectiva que sabemos que existe por 4.4. Sean $x, y \in \mathcal{F}_\kappa$, distintos de $\bot^M$ y de $\top^M$, tales que $conj^M(x)$ y $y \in^M x$.

Sea $\tau : \mathcal{F}_\kappa \to \mathcal{F}_\kappa$ tal que $\tau(y) = x$, $\tau(x) = y$ y $\tau(z) = z$ si $z \neq x \wedge z \neq y$. Ahora definimos $\widetilde{\sigma} : \mathcal{F}_\kappa \to \mathcal{K}$, $\widetilde{\sigma} = \sigma \circ \tau$, luego, para todo $z \in \mathcal{F}_\kappa$:

$$x[z]^{\widetilde{M}} = \widetilde{\sigma}^{-1}(x(\widetilde{\sigma}(z))) = \tau^{-1}(\sigma^{-1}(x(\sigma(\tau(z))))) = \tau(x[\tau(z)]^M)$$

Como $\bot^{\widetilde{M}} = \tau(\bot^M) = \bot^M$, $\top^{\widetilde{M}} = \tau(\top^M) = \top^M$ y $\forall^{\mathcal{F}_\kappa} y(x[y]^M = \bot^M \text{ o } \top^M)$ tenemos que $\forall^{\mathcal{F}_\kappa} y(x[y]^{\widetilde{M}} = \bot^{\widetilde{M}} \text{ o } \top^{\widetilde{M}})$, y como $\mathbf{I}^{\widetilde{M}}(x)$ tenemos que $conj^{\widetilde{M}}(x)$. Como $\tau(x) = y$, $x[y]^M = \top^M$ y $\tau(\top^M) = \top^{\widetilde{M}}$ tenemos que $x[x]^{\widetilde{M}} = \top^{\widetilde{M}}$, luego $x \in^{\widetilde{M}} x$. □

Corolario 4.9. $CON(ZFC + EI) \Rightarrow CON(VN + \neg\widetilde{ZF},2^{\mathbf{I}})$

Demostración: Es claro que $\exists^{conj} x(x \in x) \to \neg\widetilde{ZF},2^{\mathbf{I}}$. Entonces $ZFC + EI$ podemos construir el modelo $M(\mathcal{K}, \widetilde{\sigma})$ como en el teorema, tal que se puede probar $(VN + \neg\widetilde{ZF},2^{\mathbf{I}})^{M(\mathcal{K},\widetilde{\sigma})}$. □

Corolario 4.10. Suponiendo que $ZFC + EI$ es consistente,

$$VN \nvdash \widetilde{ZF},2^{\mathbf{I}}$$

De aquí también se deduce un resultado conocido de la teoría de conjuntos que dice que el axioma de fundación no se deduce de los demás:

Corolario 4.11. $CON(ZFC + EI) \Rightarrow CON(\widetilde{ZFC}^- + \neg\widetilde{ZF},2)$

Demostración: En $ZFC + EI$ tenemos la estructura $M(\mathcal{K}, \widetilde{\sigma})$ y dentro tenemos la estructura **I**. Como $VN \vdash \widetilde{ZFC}^{-\mathbf{I}}$, tenemos probado $((\widetilde{ZFC}^- + \neg\widetilde{ZF},2)^{\mathbf{I}})^{M(\mathcal{K},\widetilde{\sigma})}$, entonces obtenemos lo que queríamos. □

4.3.2. Consistencia de $VN + \widetilde{ZF},2^{\mathbf{I}}$

En esta sección construimos, en $ZFC + EI$, un modelo de $VN + \widetilde{ZF},2^{\mathbf{I}}$. La idea de esta construcción es parecida a la de la construcción de la clase $\mathbb{WF}$ (Well Founded) (Ver [Kun80].).

Definición 4.12.

- Sean $\perp\!\!\!\perp := \emptyset$, $\mathbb{T} := \{\emptyset\}$ y $\mathbb{T}' := \{\prec\perp\!\!\!\perp, \mathbb{T}\succ\}$.
- Notamos con $\mathcal{F}^P(A, B)$ al conjunto de las *funciones parciales* de A en B, es decir $\mathcal{F}^P(A, B) := \bigcup_{X \subset A} \mathcal{F}(X, B)$.

- Definamos al conjunto $\mathcal{D}_\alpha$ por recursión en $\alpha \in \mathbb{ON}$:
  - $\mathcal{D}_0 := \{\perp\!\!\!\perp\}$
  - $\mathcal{D}_{\beta+1} := \mathcal{F}^P(\mathcal{D}_\beta, \mathcal{D}_\beta \setminus \{\perp\!\!\!\perp\}) \cup \{\mathbb{T}\} \setminus \{\mathbb{T}'\}$
  - $\mathcal{D}_\alpha := \bigcup_{\xi<\alpha} \mathcal{D}_\xi$ en el caso α ordinal límite.

Los elementos de $\mathcal{D}_{\alpha+1}$ son las funciones parciales que se pueden construir con elementos de $\mathcal{D}_\alpha$ sin $\perp\!\!\!\perp$ en la imagen, y sustituyendo $\mathbb{T}'$ por $\mathbb{T}$. En realidad lo que queremos es $\mathbb{T} = \{\prec\perp\!\!\!\perp, \mathbb{T}\succ\}$, lo cual es imposible en ZFC, por eso definimos $\mathbb{T}' = \{\prec\perp\!\!\!\perp, \mathbb{T}\succ\}$ y luego cada vez que aparece $\mathbb{T}'$ lo sustituimos por $\mathbb{T}$.

Si calculamos $\mathcal{D}_1$ obtenemos $\{\perp\!\!\!\perp, \mathbb{T}\}$, ya que $\perp\!\!\!\perp = \emptyset$ es la función vacía que es una función parcial de A en B para todo A y B. Es fácil ver que $\mathcal{D}_2$ es

$$\{\ \perp\!\!\!\perp,\ \mathbb{T},\ \{\prec\mathbb{T}, \mathbb{T}\succ\},\ \{\prec\perp\!\!\!\perp, \mathbb{T}\succ, \prec\mathbb{T}, \mathbb{T}\succ\}\ \}$$

no explicitamos $\mathcal{D}_3$ porque tiene 256 elementos.

Empezamos demostrando que los conjuntos $\mathcal{D}_\alpha$ forman una sucesión estrictamente creciente.

Lema 4.13. $\forall \xi < \alpha(\mathcal{D}_\xi \subsetneq \mathcal{D}_\alpha)$.

DEMOSTRACIÓN: Demostraremos $\forall \xi < \alpha(\mathcal{D}_\xi \subset \mathcal{D}_\alpha)$ por inducción en $\alpha \in \mathbb{ON}$.

Es claro cuando $\alpha = 0$, $\alpha = 1$ o α es ordinal límite. Veamos el caso $\alpha = \beta + 1$ con $\beta \geq 1$. Sea $x \in \mathcal{D}_\beta$, si $x = \mathbb{T}$ luego $x \in \mathcal{D}_\alpha$, si no $x \in \mathcal{F}^P(\mathcal{D}_\xi, \mathcal{D}_\xi \setminus \{\perp\!\!\!\perp\}) \setminus \{\mathbb{T}'\}$ para algún $\xi < \beta$ (si β es sucesor, entonces $\xi = \beta - 1$ y si es límite ξ es tal $x \in \mathcal{D}_{\xi+1}$), por hipótesis inductiva $\mathcal{D}_\xi \subset \mathcal{D}_\beta$, luego $x \in \mathcal{F}^P(\mathcal{D}_\beta, \mathcal{D}_\beta \setminus \{\perp\!\!\!\perp\}) \setminus \{\mathbb{T}'\} \subset \mathcal{D}_\alpha$. De donde concluimos que $\mathcal{D}_\beta \subset \mathcal{D}_\alpha$, entonces para todo $\xi < \alpha$: $\mathcal{D}_\xi \subset \mathcal{D}_\beta \subset \mathcal{D}_\alpha$.

Ahora demostremos $\forall \xi < \alpha(\mathcal{D}_\xi \neq \mathcal{D}_\alpha)$ por inducción en α.

Es trivial cuando $\xi = 0$, así que consideremos $0 < \xi < \alpha$. Sea $f \in \mathcal{F}(\mathcal{D}_\xi, \mathcal{D}_\xi \setminus \{\perp\!\!\!\perp\})$, tal que $f(x) = \mathbb{T}$ para todo $x \in \mathcal{D}_\xi$. Supongamos que $f \in \mathcal{D}_\xi$, luego existe $\xi\prime < \xi$ tal que $f \in \mathcal{F}^P(\mathcal{D}_{\xi\prime}, \mathcal{D}_{\xi\prime} \setminus \{\perp\!\!\!\perp\})$, luego $\mathcal{D}_\xi = dom(f) \subset \mathcal{D}_{\xi\prime} \subset \mathcal{D}_\xi$, pero por hipótesis inductiva $\mathcal{D}_{\xi\prime} \neq \mathcal{D}_\xi$. Concluimos que $f \in \mathcal{D}_\alpha \setminus \mathcal{D}_\xi$, luego $\mathcal{D}_\xi \neq \mathcal{D}_\alpha$. □

Este lema nos induce a hacer la siguiente definición:

Definición 4.14. Sea $Rank : \mathcal{D}_\kappa \to \kappa$ definida de la siguiente forma:

$$Rank(x) := min\{\alpha \in \kappa : x \in \mathcal{D}_\alpha\}$$

Observación 4.15.

- $Rank(\perp\!\!\!\perp) = 0$ y $Rank(\top\!\!\!\top) = 1$.
- $\mathcal{D}_\alpha = \{x \in \mathcal{D}_\kappa : Rank(x) \leq \alpha\}$
- $Rank(x)$ nunca es un ordinal límite.

Demostración: Para la primera parte hay que observar que $\mathcal{D}_0 = \{\perp\!\!\!\perp\}$ y que $\mathcal{D}_1 = \{\perp\!\!\!\perp, \top\!\!\!\top\}$. En la segunda parte las dos inclusiones son fáciles de demostrar. Para la tercera parte hay que observar que si $x \in \mathcal{D}_\alpha$ con α ordinal límite, entonces existe $\xi < \alpha$ tal que $x \in \mathcal{D}_\xi$. □

En el siguiente lema estudiamos la cardinalidad de $\mathcal{D}_\alpha$

Lema 4.16. Para todo $\alpha < \kappa$, $\|\alpha\| \leq \|\mathcal{D}_\alpha\| < \kappa$ y $\|\mathcal{D}_\kappa\| = \kappa$.

Demostración: Los conjuntos $\mathcal{D}_{\xi+1} \setminus \mathcal{D}_\xi$ con $\xi < \alpha$ son disjuntos y no vacíos por el lema anterior, luego $\|\mathcal{D}_\alpha\| = \|\bigcup_{\xi<\alpha} \mathcal{D}_{\xi+1} \setminus \mathcal{D}_\xi\| \geq \|\alpha\|$.

Ahora probaremos $\|\mathcal{D}_\alpha\| < \kappa$ por inducción en $\alpha \in \kappa$. Si $\alpha = 0$ es trivial. Si $\alpha = \beta + 1$ y $\|\mathcal{D}_\beta\| < \kappa$ luego $\mathcal{D}_\alpha \subset \mathcal{P}(\mathcal{D}_\beta \times \mathcal{D}_\beta) \cup \{\top\!\!\!\top\}$, como κ es límite fuerte y $\|\mathcal{D}_\beta\| < \kappa$ tenemos que $\kappa > \|\mathcal{P}(\mathcal{D}_\beta \times \mathcal{D}_\beta) \cup \{\top\!\!\!\top\}\| \geq \|\mathcal{D}_\alpha\|$. Si α es ordinal límite, de la observación 4.2 y de la hipótesis inductiva obtenemos que $\|\bigcup_{\xi\in\alpha} \mathcal{D}_\xi\| < \kappa$.

Para probar la última afirmación tenemos que observar que $\|\mathcal{D}_\kappa\| = sup\{\|\mathcal{D}_\alpha\| : \alpha < \kappa\}$ y que para todo $\alpha < \kappa$ tenemos que $\|\alpha\| \leq sup\{\|\mathcal{D}_\xi\| : \xi < \kappa\} \leq \kappa$. De donde concluimos $\|\mathcal{D}_\kappa\| = \kappa$. □

Ahora veremos que hay una forma natural de identificar a las funciones κ-casi nulas de $\mathcal{D}_\kappa \to \mathcal{D}_\kappa$ con los elementos de $\mathcal{D}_\kappa$. Considerar esta biyección particular, en la creación del modelo de VN, es lo que nos va a permitir probar $\widetilde{ZF}$,2.

Definición 4.17. Dada $f \in \mathcal{D}_\kappa$, definimos $\bar{f} \in \mathcal{F}(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$ tal que dado $x \in \mathcal{D}_\kappa$

$$\bar{f}(x) := \begin{cases} f(x) & \text{si } f \neq \top\!\!\!\top \text{ y } x \in dom(f) \\ \perp\!\!\!\perp & \text{si } f \neq \top\!\!\!\top \text{ y } x \in \mathcal{D}_\kappa \setminus dom(f) \\ \top\!\!\!\top & \text{si } f = \top\!\!\!\top \text{ y } x = \perp\!\!\!\perp \\ \perp\!\!\!\perp & \text{si } f = \top\!\!\!\top \text{ y } x \neq \perp\!\!\!\perp \end{cases}$$

Definimos $\overline{\mathcal{D}_\kappa} := \{\bar{f} : f \in \mathcal{D}_\kappa\}$.

Observemos que $\emptyset = \perp\!\!\!\perp \in \mathcal{D}_\kappa$ luego podemos hablar de $\mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$. Y también observemos que $sop(\bar{f}) = dom(f)$.

Es fácil verificar que la función de $\mathcal{D}_\kappa \to \overline{\mathcal{D}_\kappa}$ que lleva a f en $\bar{f}$ es biyectiva y su inversa es la función ϱ que se obtiene restringiendo a las funciones de $\overline{\mathcal{D}_\kappa}$ a su soporte. Hay que tener cuidado con la identificación que hicimos de $\mathbb{T}$ con $\mathbb{T}'$, la definición de $\varrho : \overline{\mathcal{D}_\kappa} \to \mathcal{D}_\kappa$ es:

$$\varrho(g) = \begin{cases} g|_{sop(g)} & \text{si } g|_{sop(g)} \neq \mathbb{T}' \\ \mathbb{T} & \text{si } g|_{sop(g)} = \mathbb{T}' \end{cases}$$

Lema 4.18. $\overline{\mathcal{D}_\kappa} = \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$.

DEMOSTRACIÓN: ($\overline{\mathcal{D}_\kappa} \subset \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$): Dada $f \in \mathcal{D}_\kappa$, $\bar{f} \in \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$ porque $\|sop(\bar{f})\| = \|\,dom(f)\| \leq \|\mathcal{D}_\alpha\| < \kappa$ con $\alpha < \kappa$.

($\overline{\mathcal{D}_\kappa} \supset \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$): Si $g \in \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$, sea $f = \varrho(g)$, o sea que $g = \bar{f}$. Si $f = \mathbb{T}$, $g = \bar{f} \in \overline{\mathcal{D}_\kappa}$. Supongamos ahora que $f \neq \mathbb{T}$ y probemos que $f \in \mathcal{D}_\kappa$. Consideramos $\{Rank(x); x \in dom(f)\} \subset \kappa$, como κ es regular y $\|\{Rank(x); x \in dom(f)\}\| \leq \|\,dom(f)\| = \|sop(g)\| < \kappa$ tenemos que $\exists \lambda_1 \in \kappa$ tal que $\{Rank(x); x \in dom(f)\} \subset \kappa \subset \lambda_1$, o sea que $dom(f) \subset \mathcal{D}_{\lambda_1}$. De la misma forma probamos que existe $\lambda_2 \in \kappa$ tal que $ran(f) \subset \mathcal{D}_{\lambda_2}$. Sea $\lambda = max(\lambda_1, \lambda_2)$. Tenemos que $f \in \mathcal{F}^P(\mathcal{D}_\lambda, \mathcal{D}_\lambda \setminus \{\perp\!\!\!\perp\}) \setminus \{\mathbb{T}'\} \subset \mathcal{D}_{\lambda+1} \subset \mathcal{D}_\kappa$ y $g = \bar{f} \in \overline{\mathcal{D}_\kappa}$. □

Sea $\ll\cdot,\cdot\gg : \mathcal{D}_\kappa \times \mathcal{D}_\kappa \to \mathcal{D}_\kappa$ una función inyectiva cualquiera. Del lema anterior y del comentario anterior al lema concluimos que ϱ es una biyección de $\mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa) \to \mathcal{D}_\kappa$, luego tenemos $modelo(\mathcal{D}_\kappa, \varrho, \ll, \gg, \mathbb{T})$. Por lo tanto, por el teorema 4.6, tenemos que se verifica $VN^{M(\mathcal{D}_\kappa,\varrho)}$, sólo nos queda probar $(\widetilde{ZF},2^{\mathbf{I}})^{M(\mathcal{D}_\kappa,\varrho)}$. Para esto necesitamos el siguiente lema.

Lema 4.19. Si $f, g \in \mathcal{D}_\kappa$ y $\bar{f}(g) \neq \perp\!\!\!\perp$, entonces $Rank(g) < Rank(f)$

DEMOSTRACIÓN: Sea $\alpha = Rank(f)$, entonces $\alpha = 0$ o $\alpha = \beta + 1$. Si $\alpha = 0$, luego $f = \perp\!\!\!\perp = \emptyset$ y para todo $g \in \mathcal{D}_\kappa$, $\bar{f}(g) = \perp\!\!\!\perp$. Entonces $\alpha = \beta + 1$. Si $f = \mathbb{T}$, como $\bar{f}(g) \neq \perp\!\!\!\perp$, tenemos que $g = \perp\!\!\!\perp$ y $Rank(g) = 0 < \alpha$. Supongamos que $f \neq \mathbb{T}$, entonces $f \in \mathcal{F}^P(\mathcal{D}_\beta, \mathcal{D}_\beta \setminus \{\perp\!\!\!\perp\})$, luego $g \in dom(f) \subset \mathcal{D}_\beta$, de donde obtenemos $Rank(g) \leq \beta < \alpha$. □

Teorema 4.20. En $ZFC + EI$ se prueba $(VN + \widetilde{ZF},2^{\mathbf{I}})^{M(\mathcal{D}_\kappa,\varrho)}$.

DEMOSTRACIÓN: Ya sabemos $VN^{M(\mathcal{D}_\kappa,\varrho)}$. Notaremos $\mathcal{F} = \mathcal{F}(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$ y $\mathcal{F}_\kappa = \mathcal{F}_\kappa(\mathcal{D}_\kappa, \mathcal{D}_\kappa)$.

La fórmula $(\widetilde{ZF},2^{\mathbf{I}})^{M(\mathcal{K},\varrho)}$ es:

$$\forall^{conj^M} x(\exists^{\mathcal{F}_\kappa} y \,\epsilon^M\, x \to \exists^{\mathcal{F}_\kappa} y \,\epsilon^M\, x \forall z \,\epsilon^M\, x(z \not\epsilon^M y))$$

Sea $x \in \mathcal{F}_\kappa$ tal que $conj^M(x)$ y tal que $\exists^{\mathcal{F}_\kappa} y \in^M x$. Sea $y \in \{\varrho^{-1}(w); w \in sop(x)\}$, o lo que es lo mismo $y \in^M x$, tal que $Rank(\varrho(y)) = min\{Rank(w); w \in sop(x)\}$. Sea $z \in^M x$, probaremos que $z \notin^M y$. Tenemos que $Rank(\varrho(z)) \geq Rank(\varrho(y))$, luego, del lema anterior , deducimos que $\neg(y(\varrho(z)) \neq \perp\!\!\!\perp)$, aplicando ϱ^{-1} en ambos lados de la desigualdad obtenemos $\neg(y[z]^M \neq \perp^M)$, o sea que $z \notin^M y$. □

Corolario 4.21. $CON(ZFC + EI) \Rightarrow CON(VN + \widetilde{ZF},2^{\mathbf{I}})$

Corolario 4.22. Suponiendo que $ZFC + EI$ es consistente:

$$VN \nvdash \neg\widetilde{ZF},2^{\mathbf{I}}$$

Agadecimientos

Quiero agradecer a Paula Severi y a Walter Ferrer que, además de haberme ayudado con este trabajo, fueron muy importantes en lo que va de mi carrera.

Referencias

[Hei67] Jean van Heijenoort. *From Frege to Gödel. A source book in mathematical logic.* Harvard University Press, Cambridge, 1967.

[Kun80] Kenneth Kunen. *Set Theory. An Introduction to Independence Proofs.* North Holland, Amsterdam, 1980.

[Men97] Elliott Mendelson. *Introduction to Mathematical Logic.* Chapman y Hall, Londres, fourth edition, 1997.

[Mon00] Antonio Montalbán. Teoría de conjuntos según von Neumann, 2000. Monografía de licenciatura orientada por Paula Severi. UDELAR, Montevideo.

[Mor98] Walter Moreira. Comentarios sobre la teoría axiomática de conjuntos. Trabajo monográfico para la materia Epistemología, 1998.

[Neu25] John von Neumann. Eine Axiomatisierung der Mengenlehre. *Journal für die reine und angewandte Mathematik*, 154:219–240, 1925. Traducido al inglés en [Hei67], 393-413.

Antonio Montalbán
antonio@math.cornell.edu
Mathematics Department, Cornell University, Ithaca, NY 14850

Impreso y Encuadernado en
MASTERGRAF SRL
Gral. Pagola 1727 - CP 11800 - Tel.: 203 4760*
Montevideo - Uruguay
E-mail: mastergraf@netgate.com.uy

Depósito Legal 335.961 - Comisión del Papel
Edición Amparada al Decreto 218/96