inria-00347424, version 1 - 15 Dec 2008

Homotopy in Cat via Paths and the Fundamental Groupoid of a Category.

François Lamarche

Loria & INRIA-Lorraine
Projet Calligramme
615, rue du Jardin Botanique
54602 Villers-lès-Nancy — France
Francois.Lamarche@loria.fr

Abstract. We construct an endofunctor of paths in the category of small category and show how to construct the standard homotopy invariants from it. We give a novel proof that the fundamental groupoid of a category is its associated universal groupoid.

1 Introduction

The Homotopy theory of small categories has been closely associated to that of simplicial set. This is no surprise, since *Cat* is naturally embedded in $Set^{\Delta^{op}}$. In particular, the geometric realization of a small category is just the geometric realization of its associated simplicial set. Also, since simplicial sets were the main example that led Quillen to his theory of model structures, it is no surprise that model structures have been the main tool for studying homotopy in *Cat*.

In this paper we present a more concrete and intuitive approach to homotopy in *Cat*, which is much closer to the historical development of topology: we will define an endofunctor of paths, and deduce the usual homotopy invariants from it, closely mimicking the standard constructions that have been done in topological spaces for about a century.

Path endo(2-)functors have been defined in the much more general context of Grothendieck toposes [3, 8, 10], based on the fact that the unit interval is an exponentiable object in that 2-category. But then the complexity of the computations makes it hard to get concrete results. What we present is very concrete and combinatorial; in particular it does not involve the real numbers or geometrical realization in any way—although the reader can certainly appeal to these for geometric insights.

We will show the validity of our approach by giving a novel proof of the standard theorem that asserts that the fundamental groupoid of a category is its associated universal groupoid.

1.1 Some conventions

We use standard notation for categorical composition, fg or $f \circ g$, with the traditional functional order; this notation has the advantages of being consistent

inria-00347424, version 1 - 15 Dec 2008

throughout, and as is as traditional as can be. But we admit that it clashes a little with the habit of Western readers of reading diagrams from left to right and top to bottom.

Given two sets X, Y, the disjoint sum (coproduct) is denoted $X + Y$ and unless circumstances prevent us we will always assume that $X \subseteq X + Y \supseteq Y$.

Given a category X (small, locally small or whatever) we write $X^{\rightarrow}$ for the category whose objects are the maps of X and maps the commutative squares.

We recall that in the world of categories a diagram is to a functor what in the world of sets an indexed family of elements of a set is to a function: logically the same thing, but notated differently because of a different emphasis.

Given a poset $(X, \leqslant)$ we recall the definition of predecessor: $x \prec y$ *x is the predecessor of y, or y is the successor of x*, when

$$x < y \text{ and } x \leqslant z \leqslant y \text{ implies } x = z \text{ or } z = y.$$

Given a natural transformation, say $\alpha\colon F \to G$ and a component of it, e.g., $\alpha_x\colon F(x) \to G(x)$, we will often drop the index.

2 A semi-abstract approach

This paper is concerned with the basics of homotopy theory in *Cat*, the category of small categories and functors.

The axioms we will present here are not sufficient to allow the derivation of the main theorem only from them, so we cannot claim to have an abstract approach to homotopy via paths in *Cat*. But the level of abstraction we provide is useful for giving an intuitive formulation and relate it to a standard topological approach, for example as presented in textbooks like Spanier or May [12].

2.1 The relevant structure

As we said we do not make use of things like Quillen model structures, but instead base everything on a path endofunctor $\mathbf{P}$ on *Cat* equipped with the following structure (where X is a generic object of *Cat*),

$$X \xrightarrow{uX} \mathbf{Q}X \subseteq \mathbf{P}X \underset{s_1X}{\overset{s_0X}{\rightrightarrows}} X\ , \tag{1}$$

natural in X. An object $\mathbf{p} \in \mathbf{P}X$ is thought of a path whose beginnning is the object $s_0(\mathbf{p})$ of X and end is $s_1(\mathbf{p})$. More precisely, a path is just an arbitrary connected sequence of maps of X, like

$$s_0(\mathbf{p}) \longrightarrow \cdot \longleftarrow \cdot \longleftarrow \cdot \longrightarrow \cdot \longrightarrow \cdot \longleftarrow \cdot \longleftarrow s_1(\mathbf{p}) \tag{2}$$

where each component map can go forward or backward, and where such a sequence can be of arbitrary length. There is nothing original about this definition:

inria-00347424, version 1 - 15 Dec 2008

it has been known forever that maps in a category correspond to 1-simplices, and it is obvious that a good general definition of path should be independent of the direction of the map itself. But the author would like to add that he got that definition through topos-theoretic considerations: given a path as above of length (number of maps) n and a decomposition of the unit interval I in n segments, one can easily construct a geometric morphism of Grothendieck toposes

$$\mathrm{Sh}(I) \longrightarrow \mathit{Set}^X .$$

But naturally topos theory is not essential at all, and the treatment in this paper is quite elementary. The real problem and novelty of the paper is not the correct definition of paths themselves, but that of the *morphisms* between two such paths. We now know that there are several possible definitions, that give rise to equivalent homotopy theories based on different path functors.

The subfunctor $\mathbf{Q}$ is certainly the most important difference with respect to ordinary topological path homotopy. The category $\mathbf{Q}X$ is seen as the full subcategory of *constant paths*; unlike what happens in a category of topological spaces, constant paths are not uniquely defined, and they can even have arbitrary length. Concretely, the constant paths are the paths of (2) whose every component map is an identity. When we need it we denote the inclusion map $\mathbf{Q}X \to \mathbf{P}X$ by iX but we try to avoid doing this. Notice that iX equalizes s_0X and s_1X.

The map $uX\colon X \to \mathbf{P}X$ takes an object of X and maps it to its *unit path*, which is the only path of length zero, the most constant path of them all. Thus in Diagram (1) above the two ways of going from X to X are the identity.

Thus, if we forget $\mathbf{Q}$ the structure $(\mathbf{P}, s_0, s_1, u)$ is an "interval object" in the functor category $[\mathit{Cat}, \mathit{Cat}]$. People interested in these things should take note that, given the universal property [6] of the cubical category $\Box$ this is enough to give every object of Cat a natural cubical set structure. But this observation is not very useful, given the need to take account of $\mathbf{Q}$ too.

This path functor is not required to have a left adjoint ("cylinder"), since it doesn't exist in our examples anyway. This means that $\mathbf{P}X$ cannot be constructed by the operation of exponentiating with a unit interval object, as happens in compactly generated spaces and related categories. As a matter of fact, the fact that Cat is cartesian closed is not used at all in this paper.

There is still a little to add to the structure we need. There is a map $jX\colon X^{\to} \to \mathbf{P}X$, natural in X. The meaning of this map: commutative squares are maps between paths of length one. Since in addition the operation of reversing paths defines a natural involution $(-)^*\colon \mathbf{P} \to \mathbf{P}$, the statement above is true regardless of the direction (remember, both direction are allowed in Diagram (2)). Also the operation of concatenating paths, turns $(\mathbf{P}, s_0, s_1, u)$ into an "internal reflexive graph with composition" (and u will act as a unit for that composition operation). But that operation is not associative in general, so we do not get a category in Cat (that is, a double category), and things look very much like they do in (say) compactly generated spaces.

inria-00347424, version 1 - 15 Dec 2008

2.2 Presenting the main construction and stating the main theorem

In the following, if we put $i = 0$ or $i = 1$ we get a superposition of two commutative squares, upper and lower, because of naturality.

$$\begin{array}{ccc} \mathbf{PP}X & \underset{s_1\mathbf{P}X}{\overset{s_0\mathbf{P}X}{\rightrightarrows}} & \mathbf{P}X \\ {\scriptstyle \mathbf{P}s_i}\downarrow & & \downarrow{\scriptstyle s_i} \\ \mathbf{P}X & \underset{s_1X}{\overset{s_0X}{\rightrightarrows}} & X \end{array}$$

We can combine for $i = 0, 1$ giving the "commutative doubled square" at the right below, and by pulling back with the inclusions $iX \times iX$, construct the object $\mathbf{H}X$ of paths between paths that are constant on the endpoints (homotopies!).

$$\begin{array}{ccccc} \mathbf{H}X & \xrightarrow{HX} & \mathbf{PP}X & \underset{s_1\mathbf{P}X}{\overset{s_0\mathbf{P}X}{\rightrightarrows}} & \mathbf{P}X \\ {\scriptstyle \langle r_0,r_1\rangle}\downarrow & & {\scriptstyle \langle \mathbf{P}s_0,\mathbf{P}s_1\rangle}\downarrow & & \downarrow{\scriptstyle \langle s_0,s_1\rangle} \\ \mathbf{Q}X \times \mathbf{Q}X & \xrightarrow[iX\times iX]{} & \mathbf{P}X \times \mathbf{P}X & \underset{s_1X\times s_1X}{\overset{s_0X\times s_0X}{\rightrightarrows}} & X \times X \end{array}$$

Obviously $iX \times iX$ equalizes the two horizontal maps at bottom right. Given that all relevant squares (see above) commute, we get that $\langle s_0, s_1\rangle$ coequalizes the two horizontal composites at the top. Thus if we construct the coequalizer (Π_1X, pX) there is a map $\langle d_0, d_1\rangle$ to $X \times X$ making the triangle commute.

$$\begin{array}{ccccc} \mathbf{H}X & \underset{s_1\mathbf{P}X\circ HX}{\overset{s_0\mathbf{P}X\circ HX}{\rightrightarrows}} & \mathbf{P}X & \xrightarrow{pX} & \Pi_1X \\ & {\scriptstyle \langle ir_0,ir_1\rangle}\searrow & \downarrow{\scriptstyle \langle s_0,s_1\rangle} & \swarrow{\scriptstyle \langle d_0,d_1\rangle} & \\ & & X \times X & & \end{array} \qquad (3)$$

We claim, taking the $\mathbf{P}$, $\mathbf{Q}$, *etc.* we will define in the next section, that the concatenation and unit on $\mathbf{P}X$ can be transferred to (Π_1, d_0, d_1), giving it an internal groupoid structure, which can be described precisely.

Obviously a groupoid in *Cat* is a kind of double category [4,9, p. 44]. We recall that "maps" in these beasts (here, the morphisms of Π_1X) can be seen as squares that can be composed in two ways, traditionally called the horizontal (ordinary categorical composition in Π_1X, and the vertical (the operation inherited from path composition).

Such an "internal" approach to homotopy, in which the homotopy invariants belong more or less to the same category as the spaces under study, has been advocated before in the context of topos theory [10]. We use such an approach simply because it follows completely naturally from our simple construction of

inria-00347424, version 1 - 15 Dec 2008

Π_1, and that it obeys a simple universal property, as we will see. Subsequent work should display other advantages of this approach. It could also be done with topological spaces obeying familiar connectedness conditions, and the result would be a groupoid in the category of sheaves (étale maps) over that space (actually the maps would be locally constant, which is much stronger than just being local homeomorphisms).

Let $G\colon X \to Y$ be an arbitrary functor between small categories. We briefly recall the construction of the "comma category" $(G{\downarrow}G)$ and of its two projections $g_0, g_1 : (G{\downarrow}X) \to G$. An object is a triple (x, x', m) where $x, x' \in X$ and $m\colon Gx \to Gx'$ in Y. A map $(x, x', m) \to (y, y', n)$ is a pair $(s\colon x \to y, s' : x' \to y')$ such that $Gs' \circ m = n \circ Gs$. More succinctly, the comma category along with its projections is just the vertical map to the left of the following pullback square.

$$\begin{array}{ccc} (G \downarrow G) & \longrightarrow & Y^{\to} \\ {\scriptstyle \langle g_0, g_1\rangle}\downarrow & & \downarrow{\scriptstyle \langle d_0, d_1\rangle} \\ X \times X & \xrightarrow[G\times G]{} & Y \times Y \end{array}$$

Suppose now $Y = \mathbf{G}X$ where $\mathbf{G}X$ is the universal groupoid associated to X and $G\colon X \to \mathbf{G}X$ the universal map. The operation $(-)^{\to}$ is functorial, so there is an obvious map $X^{\to} \to (\mathbf{G}X)^{\to}$. From the pullback definitioin of $(G \downarrow G)$ and the naturality of $\langle d_0, d_1\rangle$ we immediately get a commutative triangle:

$$\begin{array}{ccc} X^{\to} & \longrightarrow & (G \downarrow G) \\ & {\scriptstyle \langle d_0, d_1\rangle}\searrow \quad \swarrow{\scriptstyle \langle g_0, g_1\rangle} & \\ & X \times X & \end{array} \tag{4}$$

Theorem 1. *In the model described below the preceding triangle is equal to*

$$\begin{array}{ccccc} X^{\to} & \xrightarrow{jX} & \mathbf{P}X & \xrightarrow{pX} & \Pi_1 X \\ & {\scriptstyle \langle d_0, d_1\rangle}\searrow & & \swarrow{\scriptstyle \langle d_0, d_1\rangle} & \\ & & X \times X & & \end{array}$$

Naturally by "equal" we mean something a little weaker than strict equality, i.e., that there exists a canonical iso connecting $(G \downarrow G)$ and $\Pi_1 X$ that makes everything commute when the two diagrams above are combined. But we will see that in practice this iso can really be thought of as an equality.

Thus, we get an internal version of the standard theorem [11] that the fundamental groupoid of a category is equivalent to to its universal associated groupoid. Notice that here we have something stronger than equivalence, since our $\Pi_1 X$ shares its objects with the points of X, as is customary in topological spaces. The traditional way that this important result is proved[1] is to look first

[1] If we extend the search to the construction of the fundamental groupoid of a simplicial set, the only really new approaches explicitly involve the construction of a free groupoid from a category [5][p.39]

inria-00347424, version 1 - 15 Dec 2008

at the geometric realization of the nerve BX of a category X and try to compute the fundamental group(oid) of that. Quillen [11] observed that the category of covering spaces over BX is equivalent to the category of functors $X \to \mathit{Set}$ that send every map of X to an iso, or, equivalently, the full subcategory of Cat/X whose objects are both discrete fibrations and discrete opfibrations, from which the result follows trivially. It turns out that BX is not necessary if one uses an abstract, topos-theoretical notion of covering space which is directly definable in Cat, that was shown to be useable in that context by Barr-Diaconescu [1] following an initial proposal by Grothendieck [7]. The abstract theory of covering spaces has been given a very general treatment [2].

3 The path endofunctor in the category of categories

3.1 Elementary paths

We will define a more general construction $\mathbf{P}(Y, X)$ that will be defined for every pair (Y, X) where X is any small category and $Y \subseteq X$ is a subcategory that shares the same objects as X[2]. The two cases that matter are when $Y = X$, which will define the ordinary path functor $\mathbf{P}X$, and $Y = |X|$, which will define $\mathbf{Q}X$, which we need to construct homotopy. The objects of $\mathbf{P}(Y, X)$ are called *elementary paths in* Y, because, unlike the maps, they are entirely determined by Y.

So let (Y, X) be as above.

Definition 1. *An* elementary path $\mathbf{p}$ *in* Y *is a quadruple*

$$\mathbf{p} \;=\; \left(\mathbb{I}, \leqslant, \sqsubseteq, (\mathbf{p}_{x,x'})_{x,x'}\right)$$

where

- $(\mathbb{I}, \leqslant)$ *is a nonempty finite totally ordered set, the* before-after order*. Thus when* $x \leqslant y$ *we say "*x *is before* y*,* y *after* x*", etc. We denote its first element by* b *(the beginning) and its last one by* e *(the end of the path).*
- $\sqsubseteq$ *is another order structure on* $\mathbb{I}$*, the* diagrammatic order*, that obey the following condition, in which* $\prec_{\leqslant}, \prec_{\sqsubseteq}$ *mean the predecessor relation on* $\leqslant, \sqsubseteq$ *respectively:*

  $$\text{if } x \prec_{\leqslant} y \text{ then either } x \prec_{\sqsubseteq} y \text{ or } y \prec_{\sqsubseteq} x.$$

- $(\mathbf{p}_{x,x'})_{x,x' \in \mathbb{I}}$ *is a diagram* $(\mathbb{I}, \sqsubseteq) \to X$*. That is, for every* $x \in \mathbb{I}$ *there is an object* $\mathbf{p}_x \in X$ *and for every* $x \sqsubseteq x'$ *there is* $\mathbf{p}_{x,x'} \colon \mathbf{p}_x \to \mathbf{p}_{x'}$*, with the usual functorial identities. In particular* $\mathbf{p}_{x,x}$ *is the identity on the object* $\mathbf{p}_x$*.*

The *length* of an elementary path $\mathbf{p}$ is $\mathrm{Card}(\mathbb{I}_{\mathbf{p}}) - 1$.

When we deal with several elementary diagrams we use subscripts to distinguish whas has to be distinguished, e.g., $\mathbb{I}_{\mathbf{p}}, V_{\mathbf{q}}$, etc..

[2] More general situations can easily be imagined but at the present time we have no idea what to do with them.

inria-00347424, version 1 - 15 Dec 2008

Thus, if $\leqslant$ is a total order, we see that $\sqsubseteq$ has the shape of a zigzag, whose "branches" are totally ordered and coincide with segments of $\leqslant$, each branch of $\sqsubseteq$ having the induced order from $\leqslant$ or its opposite.

This definition is incomplete, because we need to identify two elementary paths $\mathbf{p}, \mathbf{q}$ that differ only by the way the elements or the indexing sets $\mathbb{I}_{\mathbf{p}}, \mathbb{I}_{\mathbf{q}}$ are named. Given two arbitrary elementary paths $\mathbf{p}, \mathbf{q}$, there is a natural definition of isomorphism between the structures (bipossets) defined by the triples $\big(\mathbb{I}_{\mathbf{p}}, \leqslant, \sqsubseteq\big)$ and $\big(\mathbb{I}_{\mathbf{q}}, \leqslant, \sqsubseteq, \big)$: it's just a bijection between $\mathbb{I}_{\mathbf{p}}, \mathbb{I}_{\mathbf{q}}$ that preseves and reflects both orders. Since $\leqslant$ is a total finite order, if an iso $\alpha\colon \mathbb{I}_{\mathbf{p}} \to \mathbb{I}_{\mathbf{q}}$ exists it is unique, and we identify $\mathbf{p}, \mathbf{q}$ when we have $\mathbf{q}_{\alpha(x)} = \mathbf{p}_x$ for every $x \in \mathbb{I}_{\mathbf{p}}$. We could define the concept of elementary path without having to resort to quotienting by decreeing that $\mathbb{I}$ is always of the form $\{0, \ldots, n\}$ but this is more complicated because composing paths forces renamings. But in practice we will use natural numbers as much as we can to denote indices for paths.

Remark 1. The terminology before-after reminds one of a progress in time, which is a pretty traditional way of thinking of paths in homotopy theory as in geometry. But a category-theoretical tradition also would like us to call this order the *vertical order.* We will draw elementary paths vertically as often as we can, hoping this will not waste too much paper.

Here is an elementary path of length 6.

$$
\begin{array}{c}
\mathbf{p}_{\mathsf{b}} = \mathbf{p}_0 \\
{\scriptstyle \mathbf{p}_{0,1}} \big\downarrow \\
\mathbf{p}_1 \\
{\scriptstyle \mathbf{p}_{1,2}} \big\downarrow \\
\mathbf{p}_2 \\
{\scriptstyle \mathbf{p}_{3,2}} \big\uparrow \\
\mathbf{p}_3 \\
{\scriptstyle \mathbf{p}_{4,3}} \big\uparrow \\
\mathbf{p}_4 \\
{\scriptstyle \mathbf{p}_{5,4}} \big\uparrow \\
\mathbf{p}_5 \\
{\scriptstyle \mathbf{p}_{5,6}} \big\downarrow \\
\mathbf{p}_{\mathsf{e}} = \mathbf{p}_6
\end{array}
\tag{5}
$$

The $\leqslant$ order is read from the top down, and so a down-arrow means that $\leqslant, \sqsubseteq$ coincide and an up-arrow the opposite. An elementary path of length zero is just an object of $X = Y$; such a path is different from all the paths of length $n > 0$ all whose $p_{x,y}$ are identities, and which constitute the other objects of $\mathbf{Q}X$.

Definition 2. *Given two elementary paths* $\mathbf{p}, \mathbf{q}$*, a map* $\mathbf{f}\colon \mathbf{p} \to \mathbf{q}$ *in* $\mathbf{P}(Y, X)$ *is a pair* $\big(G, (\mathbf{f}_{x,y})_{x,y}\big)$ *where* G *(or* $G_{\mathbf{f}}$ *when necessary) is a subset* $G \subseteq \mathbb{I}_{\mathbf{p}} \times \mathbb{I}_{\mathbf{q}}$*,* the graph of $\mathbf{f}$, *which*

inria-00347424, version 1 - 15 Dec 2008

Tot *contains both* (b,b) *and* (e,e)*, and is totally ordered for the induced order on* $(\mathbb{I}_{\mathbf{p}}, \leqslant) \times (\mathbb{I}_{\mathbf{q}}, \leqslant)$

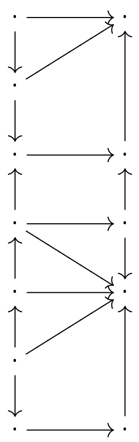
Surj *such that the two projections* $G \to \mathbb{I}_{\mathbf{p}}$ *and* $G \to \mathbb{I}_{\mathbf{q}}$ *are surjective.*

The family $(\mathbf{f}_{x,y})_{x,y}$ *assigns to every* $(x, y) \in G$*, a map* $\mathbf{f}_{x,y} \colon \mathbf{p}_x \to \mathbf{q}_y$ *in* X *such that*

Comm *everything that can possibly commute when* G *is added to the order* $(\mathbb{I}_{\mathbf{p}}, \sqsubseteq) + (\mathbb{I}_{\mathbf{q}}, \sqsubseteq)$ *does indeed commute.*

The last condition can be rephrased as follows: take the disjoint sum $\mathbb{I}_{\mathbf{f}} = \mathbb{I}_{\mathbf{p}} + \mathbb{I}_{\mathbf{q}}$. On this set put the order $\sqsubseteq_{\mathbf{f}}$ generated by $\sqsubseteq_{\mathbf{p}} \cup \sqsubseteq_{\mathbf{q}} \cup G \subseteq \mathbb{I}_{\mathbf{f}} \times \mathbb{I}_{\mathbf{f}}$ and call this relation $\sqsubseteq_{\mathbf{f}}$ (it is obviously an order and not a preorder). We are requiring that the combination of $(\mathbf{f}_{x,y})_{x,y}$ with $(\mathbf{p}_{x,x'})_{x,x'}$ and $(\mathbf{q}_{y,y'})_{y,y'}$ define a diagram $(\mathbb{I}_{\mathbf{f}}, \sqsubseteq_{\mathbf{f}}) \to X$.

An example of a map $\mathbf{f} \colon \mathbf{p} \to \mathbf{q}$ is given below (here the shape of $\mathbf{p}$ coincides with the one given in Equation (5)). As expected we draw maps horizontally.



(6)

Condition **Comm** is obviously equivalent to requiring all small triangles, squares and quadrangles in the diagram above commute.

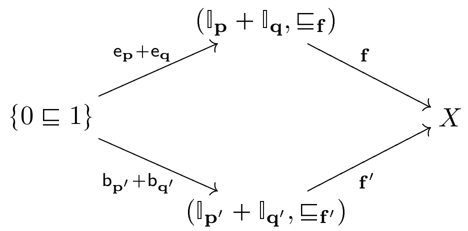
3.2 Vertical Composition

Let $\mathbf{f} \colon \mathbf{p} \to \mathbf{q}, \mathbf{f}' \colon \mathbf{p}' \to \mathbf{q}'$ be two paths such that $\mathbf{f}_{\mathsf{e},\mathsf{e}} = \mathbf{f}'_{\mathsf{b},\mathsf{b}}$. We define two paths $\mathbf{p}' * \mathbf{p}, \mathbf{q}' * \mathbf{q}$ and a map

$$\mathbf{f}' * \mathbf{f} \colon (\mathbf{p}' * \mathbf{p}) \longrightarrow (\mathbf{q}' * \mathbf{q})$$

inria-00347424, version 1 - 15 Dec 2008

in the following manner. Since we know the diagram



(7)

commutes, we define $\mathbf{f}' * \mathbf{f}$ as the universal map $\mathbb{I}_{\mathbf{f}'*\mathbf{f}} \to X \to X$ whose source is the pushout

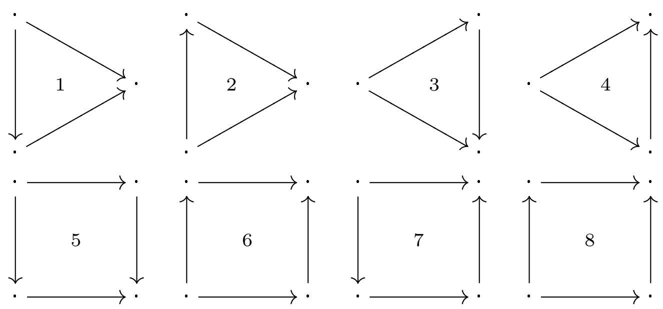
$$\mathbb{I}_{\mathbf{f}'*\mathbf{f}} \;=\; (\mathbb{I}_{\mathbf{p}} + \mathbb{I}_{\mathbf{q}}, \sqsubseteq_{\mathbf{f}}) +_{0\sqsubseteq 1} (\mathbb{I}_{\mathbf{p}'} + \mathbb{I}_{\mathbf{q}'}, \sqsubseteq_{\mathbf{f}'})\ .$$

The set $\mathbb{I}_{\mathbf{p}'*\mathbf{p}}$ is just the pushout $\mathbb{I}_{\mathbf{p}} +_0 \mathbb{I}_{\mathbf{p}'} \subseteq P$, with the diagram $\big((\mathbf{p}' * \mathbf{p})_{x,x'}\big)_{x,x'}$ obtained by restriction of $\mathbf{f}' * \mathbf{f}$. The same goes for $\mathbf{q}' * \mathbf{q}$, whose underlying set is $\mathbb{I}_{\mathbf{q}} +_1 \mathbb{I}_{\mathbf{q}'} \subseteq \mathbb{I}_{\mathbf{f}'*\mathbf{f}}$. The orders $\leqslant_{\mathbf{p}'*\mathbf{p}}$ and $\leqslant_{\mathbf{q}'*\mathbf{q}}$ come from the same pushout construction. It follows trivially that the set $G_{\mathbf{f}'*\mathbf{f}} \subseteq \mathbb{I}_{\mathbf{f}'*\mathbf{f}} \times \mathbb{I}_{\mathbf{f}'*\mathbf{f}}$, which is is the pushout of $G_{\mathbf{f}}$ and $G_{\mathbf{f}'}$, obeys the condition of Definition 2.

Thus we see that a diagrammatic representation of a vertical composite like $\mathbf{f}' * \mathbf{f}$ in the style of Equation (6) can be obtained by simply putting $\mathbf{f}$ above $\mathbf{f}'$ and identifying the last (bottom) horizontal map of $\mathbf{f}$ with the first (top) of $\mathbf{f}'$. This also says how to get diagrams for composites of paths like $\mathbf{p}' * \mathbf{p}$ in the style of Equation (5): put $\mathbf{p}$ above $\mathbf{p}'$ and connect them.

It is easy to see that vertical composition defines a category whose objects are the maps in X. Associativity comes from the associativity of the pushout operation and, given a map $f\colon a \to b$ in X, the unit $f \to f$ in that category is the map $a \to b$ in $\mathbf{P}X$ determined by f, where A, B are seen as paths of length zero.

We also see than any map in $\mathbf{P}X$ which is not a unit decomposes uniquely as vertical composition of "primitive" maps, where there are eight types (shapes) of primitive map, as follows:



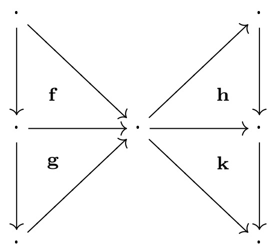
(8)

These are all the maps $\mathbf{f}$ such that $G_{\mathbf{f}}$ has exactly two elements. Now given an arbitrary map $\mathbf{g}$ we know that $G_{\mathbf{g}}$ is totally ordered for the $\leqslant$ order, and thus the decomposition is obtained by taking all pairs $\big((x, y), (x'y')\big) \in G_{\mathbf{g}} \times G_{\mathbf{g}}$ that obey the usual successor relation. Each one of these defines a primitive map.

inria-00347424, version 1 - 15 Dec 2008

3.3 Horizontal composition

We still haven't defined how maps are composed in $\mathbf{P}X$ (we use our ordinary categorical notation for this operation, either $\circ$ or juxtaposition). This needs some work, because there does not seem to be a canonical way of obtaining something like $(\mathbf{k} * \mathbf{h}) \circ (\mathbf{g} * \mathbf{f})$ in the diagram below. We have many ways to compose the "horizontal" maps here, and this creates obstacles in obtaining a graph $G_{(\mathbf{k}*\mathbf{h})\circ(\mathbf{g}*\mathbf{f})}$ which is totally ordered.



(9)

Thus we have to make choices and eliminate some of these composites. In what follows we present such a theory of uniform choices—a notion of diagrammatic normal form. It has an slightly arbitrary character, since it can be dualized, and there is nothing that makes one system of normal forms "more canonical" than its dual.

We will do a little more than is strictly necessary and show how to compose arbitrary long sequences of maps like the following one, of length n.

$$\mathscr{D} \quad = \quad \mathbf{p}^0 \xrightarrow{\mathbf{f}^1} \mathbf{p}^1 \xrightarrow{\mathbf{f}^2} \mathbf{p}^2 \quad \cdots \quad \mathbf{p}^{n-1} \xrightarrow{\mathbf{f}^n} \mathbf{p}^n \ .$$

Given one of the $\mathbf{p}^i$ above we write $\mathbb{I}_i$ for $\mathbb{I}_{\mathbf{p}^i}$, G_i for $G_{\mathbf{f}^i}$, etc.

We first construct the disjoint sum $\mathbb{I}_{\mathscr{D}} = \sum_{0 \leqslant i \leqslant n} \mathbb{I}_i$. This set has an order $\leqslant$ inherited from the various $\leqslant_i$. It also inherits the G_i from the different $\mathbf{f}_i$.

Definition 3. *Given $\mathscr{D}$ as above, and $x \in \mathbb{I}_0, x' \in \mathbb{I}_n$, a* trail $\mathbf{x}\colon x \to x'$ *is a sequence $x = \mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n = x'$ where $(\mathbf{x}_{i-1}, \mathbf{x}_i) \in G_i$.*

Let us write $x \multimap y$ when there is i such that $x \in \mathbb{I}_{i-1}$, $y \in \mathbb{I}_i$ and $(x, y) \in G_i$.

A trail $\mathbf{x}$ defines a map $M(\mathbf{x})\colon \mathbf{p}^0_x \to \mathbf{p}^n_{x'}$ by taking

$$M(\mathbf{x}) = \mathbf{f}^n_{\mathbf{x}_{n-1},\mathbf{x}_n} \circ \mathbf{f}^{n-1}_{\mathbf{x}_{n-2},\mathbf{x}_{n-1}} \circ \ \cdots \ \circ \mathbf{f}^2_{\mathbf{x}_1,\mathbf{x}_2} \circ \mathbf{f}^1_{\mathbf{x}_0,\mathbf{x}_1}$$

There might be more than one trail $\mathbf{x}\colon x \to x'$ but

Proposition 1. *The map $M(\mathbf{x})$ depends only on the endpoints x, x' and not on the exact trail $\mathbf{x}$.*

Proof. This is a direct consequence of condition **Comm**. □

Proposition 2. *Let $\mathbf{x}\colon x \to x'$ and $\mathbf{y}\colon y \to y'$ be two trails such that $x \leqslant y$ and $y' \leqslant x'$. Then these trails cross on an object of $\mathbb{I}_{\mathscr{D}}$: there is i such that $\mathbf{x}_i = \mathbf{y}_i$.*

inria-00347424, version 1 - 15 Dec 2008

Proof. If $x = y$ we are done. So assume $x < y$. Because of condition **Tot** we have to have $\mathbf{x}_1 \leqslant \mathbf{y}_1$. If $\mathbf{x}_1 = \mathbf{y}_1$ we are done. If not, we iterate, and if it turned out we could never find j with $\mathbf{x}_j = \mathbf{y}_j$ we would end up with $\mathbf{x}_n < \mathbf{y}_n$, contradicting our assumption. □

Definition 4. *A trail* $\mathbf{x}\colon x \to x'$ *as above is said to be* normal *if it satisfies the following condition:*

- *Given another trail* $\mathbf{y}\colon y \to y'$ *such that* $y' > x'$ *we always have that* $y \geqslant x$.

There is an equivalent, symmetrical definition

Proposition 3. *For a trail* $\mathbf{x}\colon x \to x'$ *TFAE*

- *Given any* $\mathbf{y}\colon y \to y'$ *such that* $y < x$ *then we always have* $y' \leqslant x'$.
- $\mathbf{x}$ *is normal.*

Proof. Assume $\mathbf{x}$ is normal and let $\mathbf{y}\colon y \to y'$ with $y < x$. If $y' > x'$ this would contradict the normality of $\mathbf{x}$. Since $\leqslant_n$ is a total order this shows $y \leqslant x$. The proof of the converse is the exact symmetrical. □

There may be more than one normal trail $x \to x'$ (exercise).

Notice that if $\mathscr{D}$ has length one (i.e., $\mathscr{D} = \mathsf{x}_0 \xrightarrow{\mathbf{f}} \mathsf{x}_i$) then every trail is normal. This is just because the set of trails there coincides with $G_{\mathbf{f}}$ and the latter set obeys **Tot**.

Let $G(\mathscr{D}) = \{\, (x, x') \mid \text{There exists a normal trail } x \to x' \,\}$. This set inherits the $\leqslant$-order on $\mathbb{I}_0 \times \mathbb{I}_n$, and we denote it by $\leqslant_{\mathscr{D}}$.

Proposition 4. *The set* $\big(G(D), \leqslant_{\mathscr{D}}\big)$ *obeys the conditions* ***Tot*** *and* ***Surj***.

Proof. For **Tot**, the trails

$$\mathsf{b}_0 \xrightarrow{\mathbf{f}^1_{\mathsf{b},\mathsf{b}}} \mathsf{b}_1 \xrightarrow{\mathbf{f}^2_{\mathsf{b},\mathsf{b}}} \mathsf{b}_2 \cdots \xrightarrow{\mathbf{f}^n_{\mathsf{b},\mathsf{b}}} \mathsf{b}_n \quad \text{and} \quad \mathsf{e}_0 \xrightarrow{\mathbf{f}^1_{\mathsf{e},\mathsf{e}}} \mathsf{e}_1 \xrightarrow{\mathbf{f}^2_{\mathsf{e},\mathsf{e}}} \mathsf{e}_2 \cdots \xrightarrow{\mathbf{f}^n_{\mathsf{e},\mathsf{e}}} \mathsf{e}_n$$

are obviously normal. Thus $\big(G(\mathscr{D}), \leqslant\big)$ has both a top and a bottom. Let now $\mathbf{x}\colon x \to x'$ and $\mathbf{y}\colon y \to y'$ be two trails. Since $\leqslant_n$ is a total order we either have

- $y' > x'$; then by normality, we know that $y \geqslant x$ and this shows $(x, y) < (x', y')$ in $G(\mathscr{D})$.
- $x' = y'$; then since $\leqslant_0$ is a total order we either have $x < y, x = y$ or $x > y$, which gives us either $(x, x') < (y, y'), (x, x') = (y, y')$ or $(x, x') > (y, y')$ in $G(\mathscr{D})$.
- $x' < y'$; then we dualize the first argument.

We have shown that $(G(\mathscr{D}), \leqslant)$ is totally ordered.

inria-00347424, version 1 - 15 Dec 2008

Now for **Surj**: we proceed by induction, starting at $(\mathsf{b}_0, \mathsf{b}_n)$ and going downwards. So assume that $(x, x') \in G(\mathscr{D})$. Let $y \prec x, y' \prec x'$ be the predecessors of x, x' (if neither exists we are done). Then either

1. There exists a trail $y \to x'$. Then $(y, x') \in G(\mathscr{D})$. This is because any trail starting at $z < y$ has to end $\leqslant x'$ since $(x, x') \in G(\mathscr{D})$.
2. There is no trail $y \to x'$ (perhaps because x' is not defined) but there is one $x \to y'$. Let $\mathbf{z}\colon z \to z'$ be a trail with $z < x$. Since $(x, x') \in G(\mathscr{D})$ we have $z' \leqslant x$. But we cannot have $z' = x$, since then any trail $\mathbf{y}$ starting at y would have to cross $\mathbf{z}$ at some $\mathbf{z}_i$ and then we would be able to splice the two:
$$y = \mathbf{y}_0, \mathbf{y}_1, \dots \mathbf{y}_i = \mathbf{z}_i, \mathbf{z}_{i+1}, \dots \mathbf{z}_n$$
and get a trail $y \to x'$, contradicting our assumption. So $z' < x$, i.e., $z' \leqslant y'$ and this shows $(x, y') \in G(\mathscr{D})$.
3. There is neither a trail $y \to x'$ nor one $x' \to y$. By induction from 0 up let $\mathbf{y}\colon y \to \mathbf{y}_n$ be the trail such that $\mathbf{y}_{i+1}$ is always the $\leqslant_{i+1}$ greatest w such that $\mathbf{y}_i \multimap w$, and by induction from n down construct the trail $\mathbf{z}\colon \mathbf{z}_0 \to y'$, such that $\mathbf{z}_{i-1}$ is always the $\leqslant_{i-1}$-greatest element w with $w \multimap \mathbf{z}_i$. Since $(x, x') \in G(\mathscr{D})$ we have $\mathbf{y}_n \leqslant y'$ and $\mathbf{z}_0 \leqslant y$, and thus $\mathbf{y}, \mathbf{z}$ cross: let i be the least integer such that $\mathbf{y}_i = \mathbf{z}_i$. If it turned out that $i \neq 0$ we'd have $\mathbf{z}_{i-1} < \mathbf{y}_{i-1}$ and this would contradict the maximality condition that $\mathbf{z}_{i-1}$ obeys. The same argument dualized shows that the greatest i where the two trails agree has to be n, and we have shown $\mathbf{y} = \mathbf{z}$. We have constructed a trail $y \to y'$, and it is easy to show that it is normal

This construction gives us a set of normal trails which obviously satisfies condition **Surj**. □

Given any diagram $\mathscr{D}$ as above, of arbitrary length n, we can define the composition $\mathbf{f} = \mathbf{f}^n \circ \mathbf{f}^{n-1} \circ \cdots \circ \mathbf{f}^1$, where we consider $(-) \circ (-) \circ \cdots \circ (-)$ to be an n-ary operator. Obviously $\mathbb{I}_{\mathbf{f}} = \mathbb{I}_0 + \mathbb{I}_n$, and we take $G_{\mathbf{f}} = G(\mathscr{D})$, with

$$\mathbf{f}_{x,y} = M(\mathbf{x}) \text{ , where } \mathbf{x} \text{ is a normal trail } x \to y \text{ .}$$

This is defined for any $(x, y) \in G_{\mathbf{f}}$, remembering that the exact choice of $\mathbf{x}$ is unimportant. The family $(\mathbf{f}_{x,y})_{x,y}$ obeys condition **Comm**, for the same reasons as in Proposition 1.

Proposition 5. *Binary composition $(-) \circ (-)$ is associative.*

Proof. Let $\mathscr{D} = \mathbf{p}^0 \xrightarrow{\mathbf{f}^1} \mathbf{p}^1 \xrightarrow{\mathbf{f}^2} \mathbf{p}^2 \xrightarrow{\mathbf{f}^3} \mathbf{p}^3$ be a diagram of length 3. We will show that

$$(\mathbf{f}^3 \circ \mathbf{f}^2) \circ \mathbf{f}^1 \ = \ \mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1 \ = \ \mathbf{f}^3 \circ (\mathbf{f}^2 \circ \mathbf{f}^1) \text{ .}$$

Let $\mathbb{I}_{\mathscr{D}}$ be constructed just as before, by taking the disjoint sum/union of the $\mathbb{I}_i$. The difference now is that this set will be used to hang more graph structures than before, namely the sets $G_{\mathbf{g}} \subseteq \mathbb{I}_{\mathscr{D}} \times \mathbb{I}_{\mathscr{D}}$ where $\mathbf{g}$ ranges over

$$\mathbf{f}^1, \mathbf{f}^2, \mathbf{f}^3,\ \mathbf{f}^2 \circ \mathbf{f}^1,\ \mathbf{f}^3 \circ \mathbf{f}^1,\ (\mathbf{f}^3 \circ \mathbf{f}^2) \circ \mathbf{f}^1,\ \mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1,\ \mathbf{f}^3 \circ (\mathbf{f}^2 \circ \mathbf{f}^1) \text{ .}$$

As a notational relief we use the $\multimap$-notation when it is unambiguous, which is in all but the last three cases in the list above. For example if $x \in \mathbb{I}_0$ and $y \in \mathbb{I}_2$ the notation $x \multimap y$ means that $(x, y) \in G_{\mathbf{f}^2 \circ \mathbf{f}^1}$, because it its the only possibility.

Let us first show first that $G_{\mathbf{f}^3 \circ (\mathbf{f}^2 \circ \mathbf{f}^1)} \subseteq G_{\mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1}$. Let $(x_0, x_3) \in G_{\mathbf{f}^3 (\circ \mathbf{f}^2 \circ \mathbf{f}^1)}$, and let y_0, y_1, y_2, y_3 be a trail in $\mathbb{I}_{\mathscr{D}}$, with $y < x_0$. If we can show that $y_3 \leqslant x_3$ we are done. We know there is x_2 with $x_0 \multimap x_2 \multimap x_3$. The shortened trail y_0, y_1, y_2 is in $\mathscr{D}_{\mathbf{f}^2 \circ \mathbf{f}^1}$ and by normality of trails for (x_0, x_2) we have $y_2 \leqslant x_2$. There are two possibilities:

- $y_2 < x_2$. Then since $x_2 \multimap x_3$ we are guaranteed that $y_3 \leqslant x_3$.
- $y_2 = x_2$. Then the same argument as in case 1. in Proposition 4 shows that $(y_0, y_2) \in G_{\mathbf{f}_2 \circ \mathbf{f}_1}$. Thus y_0, y_2, y_3 is a path in $G_{(\mathbf{f}^3 \circ \mathbf{f}^2) \circ \mathbf{f}^1}$, and because of the defining property of (x_0, x_3) we have to have $y_3 \leqslant x_3$.

For the converse, let us now show that $G_{\mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1} \subseteq G_{\mathbf{f}^3 \circ (\mathbf{f}^2 \circ \mathbf{f}^1)}$. Let (x_0, x_3) be in $G_{\mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1}$. We know there is a normal trail $x_0 \to x_3$; choose a trail $\mathbf{x} = x_0, x_1, x_2, x_3$ *(not necessarily a normal one)* such that x_2 is maximal among the w with $w \multimap x_3$. Let us show that $x_0 \multimap x_2$. Let y_0, y_1, y_2 be an arbitrary trail with $y_0 < x_0$. Suppose for a contradiction that $y_2 > x_2$. Then because of Proposition 2 we have $y_1 = x_1$; we can always find a y_3 with $y_2 \multimap y_3$, and we know $y_3 > x_3$, because if $y_3 = x_3$ the trail $x_0, x_1 = y_1, y_2, y_3$ would contradict the maximality assumption for x_2 in $\mathbf{x}$. But having $y_3 > x_3$ allows us to construct a trail y_0, y_1, y_2, y_3 that contradicts the assumption that $(x_0, x_3) \in G_{\mathbf{f}^3 \circ \mathbf{f}^2 \circ \mathbf{f}^1}$. Thus $y_2 \leqslant x_2$ and we have shown that $x_0 \multimap x_2$. Finally we show that x_0, x_2, x_3 is a normal trail, i.e., is in $G_{\mathbf{f}^3 \circ (\mathbf{f}^2 \circ \mathbf{f}^1)}$. But any trail starting strictly below x_0 and ending strictly above x_3 would have to contain x_2 and we can apply the same argument as above to show this cannot happen.

The proof of the other equation is the exact dual. □

All is left to do to get the category on $\mathbf{P}(Y, X)$ is define units. But it is easy to see that given an arbitrary path $\mathbf{p}$ the map $1 \colon \mathbf{p} \to \mathbf{p}$ defined by $G_1 = (x, x)_{x \in \mathbb{I}_{\mathbf{p}}}$ and $1_{x,x} = 1_{\mathbf{p}_x}$ will act as both a left and right unit for horizontal composition.

There are two obvious functors $s_0, s_1 \colon \mathbf{P}(Y, X) \to X$: given $\mathbf{f} \colon \mathbf{p} \to \mathbf{q}$ we take

$$\mathsf{b}_{\mathbf{p}} \xrightarrow{\mathbf{f}_{\mathsf{b},\mathsf{b}}} \mathsf{b}_{\mathbf{q}}, \quad \mathsf{e}_{\mathbf{p}} \xrightarrow{\mathbf{f}_{\mathsf{e},\mathsf{e}}} \mathsf{e}_{\mathbf{q}}$$

for $s_0(\mathbf{f}), s_1(\mathbf{f})$ respectively.

And the map $u \colon X \to \mathbf{P}(Y, X)$ that sends a map $f \colon a \to b$ to the map f between the paths of length zero a, b is also obviously functorial.

We leave it to the reader to check that the construction above gives a functor $\mathbf{P} \colon \mathit{Pair} \to \mathit{Cat}$, where an object (Y, X) in Pair is a pair of categories with $Y \subseteq X$ (with Y's objects coinciding with those of X), and a map $(Y, X) \to (Y', X')$ is a functor $X \to X'$ that factors into one $Y \to Y'$. Also the assignment $K(Y, X) = X$ is obviously a functor, and this allows us to turn the constructions s_0, s_1, u into natural transformations $\mathbf{P}(-,-) \leftrightarrow K$ in that setting.

Finally, the assignments $X \mapsto (X, X)$ and $X \mapsto (|X|, X)$ are obviously functors $\mathit{Cat} \to \mathit{Pair}$.

inria-00347424, version 1 - 15 Dec 2008

inria-00347424, version 1 - 15 Dec 2008

It is easy to see that an object in $\mathbf{Q}X$ is a diagram like Equation (5) where all the $\mathbf{p}_{i,j}$ are identity maps; thus a morphism in that category will be as in Equation (6), but all the "horizontal" maps will be forced to be equal. Thus the projection $qX : \mathbf{Q}X \to X$ is surjective on objects, and given any object $x \in X$ its fiber $q^{-1}(x)$ contains the path $\{x\} = u(x)$ of length zero: it is easy to see that it is both an initial and a terminal object in that fiber. Actually, something stronger holds: the reader can check that the pair $(uX(x), 1_x)$ is a terminal object in the category $q \downarrow x$ (often denoted q/x. Thus, because of Theorem A in [11], qX is a homotopy equivalence in the traditional sense.

Remark 2. Thus we have two notions of composition in the object of paths, vertical and horizontal, and they both define categories. But we do not get a double category, because the interchange law does not hold: the reader can check that the notion of horizontal composition we have presented gives different values for $(\mathbf{k} * \mathbf{h}) \circ (\mathbf{g} * \mathbf{f})$ and $(\mathbf{k} \circ \mathbf{g}) * (\mathbf{h} \circ \mathbf{f})$ in Diagram (9). Naturally, we *will* get a double category once we take the homotopy quotient.

Remark 3. We see that our definition of normal path could be dualized by exchanging "up" and "down" (i.e., reversing the $\leqslant$-order), giving what could be called conormal trails. This produces a different category $\mathbf{P}^{co}(Y, X)$, but the definition of homotopy we get from it is identical, since the sets of objects and maps in the pairs $(\mathbf{P}X, \mathbf{P}^{co}X)$ and $(\mathbf{Q}X, \mathbf{Q}^{co}X)$ are *identical*, the only difference being how the maps are composed. Thus the relation of homotopy between paths we have defined is independent of the exact choice of the normalization procedure used in composition.

3.4 The proof of Theorem 1

Diagram (3) tells us that $\Pi_1 X$ is obtained by a coequalizer construction. In *Cat* the object part of the target of the coequalizer is constructed just as in sets: the objects of $\Pi_1 X$ are obtained by taking the equivalence class generated by the relation $(-) \sim (-)$ where

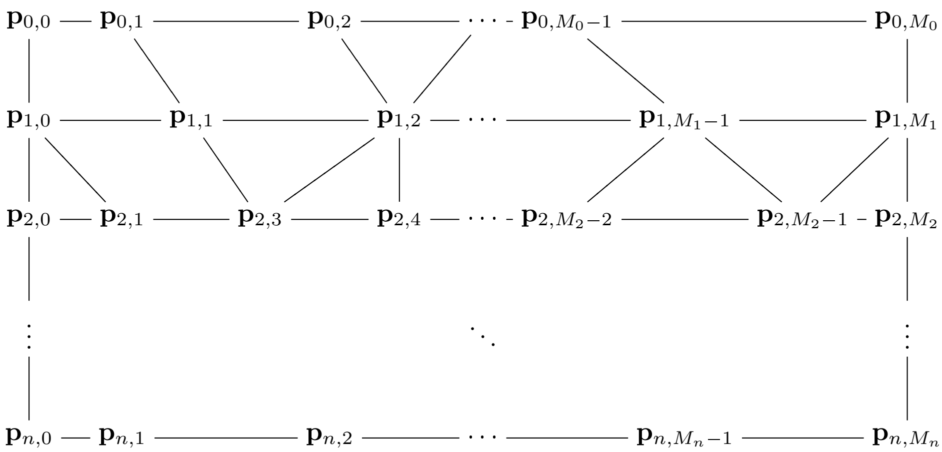
$$\mathbf{q} \sim \mathbf{r} \quad \text{if there is } \mathbf{p} \text{ in } \mathbf{H}X \text{ with } \big(s_0\mathbf{P} \circ H\big)(\mathbf{p}) = \mathbf{q} \text{ and } \big(s_1\mathbf{P} \circ H\big)(\mathbf{p}) = \mathbf{r}$$

The category $\mathbf{H}X$ is a subcategory of $\mathbf{PP}X$. In general an object $\mathbf{p}$ of that category is a path between paths; it is a diagram

$$\begin{array}{c} \mathbf{p}_0 \\ | \\ \mathbf{p}_1 \\ | \\ \mathbf{p}_2 \\ | \\ \cdots \\ | \\ \mathbf{p}_{n-1} \\ | \\ \mathbf{p}_n \end{array}$$

inria-00347424, version 1 - 15 Dec 2008

where each $\mathbf{p}_i$ is an object of $\mathbf{P}X$, and each vertical line represents a map in $\mathbf{P}X$, which may go either up or down. Supposing every path $\mathbf{p}_i$ has length M_i, we can expand these to get something like the following commutative diagram:



where the "horizontal" maps can point left or right, without any restriction, while all the maps in a row of vertical/diagonal ones point in the same direction.[3]

It should be clear that

- A path $\mathbf{p} \in \mathbf{PP}X$ is in $\mathbf{H}X$ (i.e., is mapped by $\langle \mathbf{P}s_0, \mathbf{P}s_1 \rangle$ to $\mathbf{Q}X \times \mathbf{Q}X$) iff, when given a representation as above, its leftmost and rightmost vertical column are both composed only of identities.
- Given $\mathbf{q}, \mathbf{r} \in \mathbf{P}X$ we have $\mathbf{q} \sim \mathbf{r}$ iff there is $\mathbf{p} \in \mathbf{H}X$ which, when seen as a rectangle as above, has $\mathbf{q}$ as is its top "horizontal" row and $\mathbf{r}$ as is its bottom "horizontal" row.
- Because of vertical composition in $\mathbf{PP}X$ (and the ability to reverse elementary paths) the relation $\sim$ is already an equivalence relation and there is no need to take its symmetric, transitive closure.

We also see that the constant path functor $\mathbf{Q}$ is essential to our construction. If in Diagram 2.2 we replaced $iX \times iX \colon \mathbf{Q}X \times \mathbf{Q}X \longrightarrow \mathbf{P}X \times \mathbf{P}X$ by the more traditional $\langle uX, uX \rangle \colon X \longrightarrow \mathbf{P}X \times \mathbf{P}X$ in the pullback that defines $\mathbf{H}X$, the resulting equivalence relation that constructs Π_1 would only identify identical objects.

Objects in $\mathbf{PP}X$ and $\mathbf{H}X$ can be composed in two ways (without need for the additional structure of horizontal maps).

- Ordinary vertical composition, which we denote by $(-) *_1 (-)$. It superposes two rectangles of the same "width" vertically.

[3] If paper were three-dimensional we would not hesitate to use the depth axis for what we have just drawn horizontally; we could then reserve the horizontal axis for maps in $\mathbf{PP}X$. This is why in the next few paragraphs we will enclose that word in double quotes, at the risk of appearing pedantic.

inria-00347424, version 1 - 15 Dec 2008

– The vertical composition inherited from $\mathbf{P}X$. It superposes rectangles of the same height "horizontally". We will denote it by $(-) * (-)$.

Definition 5. *Let $\mathbf{p}$ be an object of $\mathbf{P}X$ and n a natural number. There are two important ways to turn it into an object of $\mathbf{H}X$, which we denote by $1^{\uparrow}_{\mathbf{p}}$ and $1^{\downarrow}_{\mathbf{p}}$. Both elementary paths have the identity $1_{\mathbf{p}}$ as a starting point* and *as an endpoint. Both can be seen as a rectangle of height 1, but in the first case going up, and the other case going down. We use the imprecise notation $\mathbf{p}^{*n}$ to denote objects of $\mathbf{H}X$ of the form*

$$\mathbf{p}^{*n} \;=\; 1^{\eta_1}_{\mathbf{p}} *_1 1^{\eta_2}_{\mathbf{p}} *_1 \cdots *_1 1^{\eta_n}_{\mathbf{p}} \;,$$

where η_i is either $\uparrow$ or $\downarrow$. Thus seen as a rectangle its height is n, its width is Length$(\mathbf{p})$*, and we have $\mathbf{p}^{*n} = \mathbf{p}$ iff $n = 0$. The notation can be imprecise because it will always be used in contexts that will furnish the necessary contraints to force the exact values of η_i.*

Definition 6. *We will use $[\mathbf{P}X]$ and $[\Pi_1 X]$ to denote the underlying discrete categories $|\mathbf{P}X|, |\Pi_1 X|$, equipped with their natural graph structures $s_0, s_1 \colon \mathbf{P}X \to X$ and $d_0, d_1 \colon \Pi_1 X \to X$.*

Proposition 6. *The relation $\sim$ is a congruence for $*$-composition.*

Proof. Let $\mathbf{q}, \mathbf{r}$ be two parallel paths in $\mathbf{P}X$, and $\mathbf{p} \in \mathbf{H}X$ a "proof witness" of length n that shows $\mathbf{q} \sim \mathbf{r}$: in other words a rectangle of height n with $\mathbf{q}, \mathbf{r}$ as its top and bottom horizontal rows. Let $\mathbf{s}$ begin where $\mathbf{q}, \mathbf{r}$ end and $\mathbf{s}'$ end where they begin. The rectangle $\mathbf{s}^{*n} * \mathbf{p} * \mathbf{s}'^{*n}$ is a proof that $\mathbf{s} * \mathbf{q} * \mathbf{s}' \sim \mathbf{s} * \mathbf{r} * \mathbf{s}'$. ◻

Thus $(-) * (-)$ can be extended to the graph $[\Pi_1 X] = \mathbf{P}X/{\sim}$, giving it a category structure (finding the identity is left to the reader).

Now recall that the universal groupoid $\mathbf{G}X$ is constructed in the following manner: its objects are the same as those of X, and for every map $f \colon x \to y \in X$ there is a generator $f^{\circ} \colon x \to y$ and a $f^{\bullet} \colon y \to x$ in $\mathbf{G}X$. We denote by $\mathscr{F}X$ the free category obtained from these generators, i.e., the category of paths of f°s and $f^{\bullet}$s. As is well known we get $\mathbf{G}X$ by quotienting $\mathscr{F}X$ under the following relations (for clarity we will use the operator $(-) \cdot (-)$ for composition in $\mathscr{F}X$ and in $\mathbf{G}X$):

$$g^{\circ} \cdot f^{\circ} = (gf)^{\circ}, \quad g^{\bullet} \cdot f^{\bullet} = (fg)^{\bullet}, \quad f^{\bullet} \cdot f^{\circ} = 1, \quad f^{\circ} \cdot f^{\bullet} = 1 \,. \tag{10}$$

Let $K \colon [\mathbf{P}X] \to \mathscr{F}X$ be the map that sends an elementary path

$$x_0 \xrightarrow{\;f_1\;} x_1 \xrightarrow{\;f_2\;} \cdots \xrightarrow{\;f_{n-1}\;} x_{n-1} \xrightarrow{\;f_n\;} x_n$$

in X to $f_n^{\iota_n} \cdot f_{n-1}^{\iota_{n-1}} \cdot \cdots \cdot f_2^{\iota_2} \cdot f_1^{\iota_1}$, where

$$f_i^{\iota_i} = \begin{cases} f^{\circ} & \text{if } f_i \colon x_{i-1} \to x_i \\ f^{\bullet} & \text{if } f_i \colon x_i \to x_{i-1} \end{cases} .$$

inria-00347424, version 1 - 15 Dec 2008

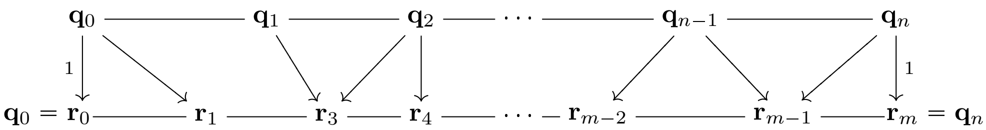
It should be obvious that K is an isomorphism between the categories $([\mathbf{P}X], *)$ and $\mathscr{F}X$. We claim that it agrees with $\sim$, i.e. that there is $\tilde{K}$ with

$$\begin{array}{ccc} [\mathbf{P}X] & \xrightarrow{K} & \mathscr{F}X \\ \downarrow & & \downarrow \\ [\Pi_1 X] & \xrightarrow[\tilde{K}]{} & \mathbf{G}X \end{array}$$

commuting.

Definition 7. *We denote by P the map K followed by the quotient projection, occupying the diagonal of the square above.*

To prove our claim, let $\mathbf{q}, \mathbf{r} \in \mathbf{P}X$ and $\mathbf{p}$ be a one-step proof, going (say) down



(11)

that $\mathbf{q} \sim \mathbf{r}$. In X the "horizontal" arrows can point right or left, but all the little triangles and quadrangles commute, once their directions are identified. But after we've applied P, all the "horizontal" maps now point right, and the little commutations ensures that both long "horizontal" composites are equal. Naturally this applies for proofs of arbitrary length, with steps that go up or down, and thus we have proved our claim.

It is trivial to show that $\tilde{K}$ is a functor. □

Proposition 7. *This functor is an isomorphism of categories*

$$\tilde{K} \colon ([\Pi_1 X], *) \longrightarrow (\mathbf{G}X, \cdot) \,.$$

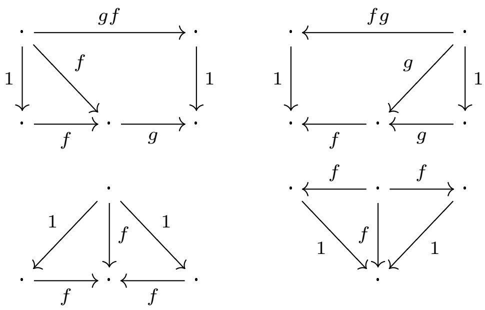
Thus as a corollary $([\Pi_1 X], *)$ is a groupoid.

Proof. Since we already know that it is a functor which is full and bijective on objects, all that is left to do is show that it is faithful. So let $\mathbf{q}, \mathbf{r}$ be parallel paths such that $P\mathbf{q} = P\mathbf{r}$. A formal proof of this equation can be represented as a sequence $\mathbf{q}_0, \mathbf{q}_1, \ldots, \mathbf{q}_n$ of parallel paths in $\mathbf{P}X$, such that

- $\mathbf{q}_0 = \mathbf{q}$ and $\mathbf{q}_n = \mathbf{r}$,
- for every $i = 1, \ldots, n$ there are paths $\mathbf{s}_i, \mathbf{s}'_i, \mathbf{a}_i, \mathbf{b}_i$ such that
  - $\mathbf{q}_{i-1} = \mathbf{s}_i * \mathbf{a}_i * \mathbf{s}'_i$ and $\mathbf{q}_i = \mathbf{s}_i * \mathbf{b}_i * \mathbf{s}'_i$
  - there is one of the four following diagrams for which either (a) $\mathbf{a}_i$ is the top row and $\mathbf{b}_i$ the bottom row, or (b) $\mathbf{a}_i$ the bottom row and $\mathbf{b}_i$ the

inria-00347424, version 1 - 15 Dec 2008

top row (compare with the equations given in Equation (10)).

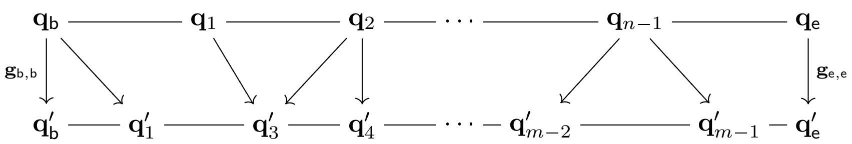


This is enough to construct a proof of $\mathbf{q} \sim \mathbf{r}$ in $\mathbf{H}X$: just do a vertical composition of n one-step proofs, each one of which is of the form $\mathbf{s}_i^{*1} * \mathbf{k}_i * \mathbf{s}_i'^{*1}$, where $\mathbf{k}_i$ is one of the four diagrams above (i.e., including the vertical maps), and is either a map $\mathbf{a}_i \to \mathbf{b}_i$ or $\mathbf{b}_i \to \mathbf{a}_i$, according to cases (a) or (b) above. □

We can now tackle the maps in Π_1. A map $\mathbf{f} \colon \mathbf{p} \to \mathbf{p}'$ in $\mathbf{PP}X$ is a diagram just as in Equation (6), but in the category $\mathbf{P}X$, so that every vertex of that graph is a path itself, and every edge a map of paths. As we have said we can imagine that that these paths (that go in what we have been calling the "horizontal" dimension) actually extend orthogonally out of the page.

Proposition 8. *Let* $\mathbf{g} \colon \mathbf{q} \to \mathbf{q}'$ *be a map in* $\mathbf{P}X$*. Then* $P\mathbf{q}' \circ G\mathbf{g}_{\mathsf{b},\mathsf{b}} = G\mathbf{g}_{\mathsf{e},\mathsf{e}} \circ P\mathbf{q}$.

Proof. Given $\mathbf{g}$ as above, we know it can be represented by something that looks like.

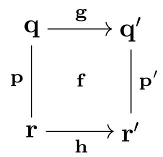


where the "horizontal" arrows can point left or right. We know that we can carry this diagram in the groupoid $\mathbf{G}X$ by applying P to the two "rows" (which makes them all point "right"), and applying plain G to the "vertical/diagonal" maps. Thus we end up with the large outer square commuting in $\mathbf{G}X$, and this translates as $P\mathbf{q}' \circ G\mathbf{g}_{\mathsf{b},\mathsf{b}} = G\mathbf{g}_{\mathsf{e},\mathsf{e}} \circ P\mathbf{q}$. □

Let us now extend the relation $\sim$ to maps: given $\mathbf{g} \colon \mathbf{q} \to \mathbf{q}'$ and $\mathbf{h} \colon \mathbf{r} \to \mathbf{r}'$ in $\mathbf{P}X$, we write $\mathbf{g} \sim \mathbf{h}$ when there exists $\mathbf{f} \colon \mathbf{p} \to \mathbf{p}'$ in $\mathbf{H}X$ such that $\big(s_0\mathbf{P} \circ H\big)(\mathbf{f}) = \mathbf{g}$ and $\big(s_1\mathbf{P} \circ H\big)(\mathbf{f}) = \mathbf{h}$. This is an equivalence relation on maps, for the same reason as it is on objects. Maps in $\mathbf{H}X$ can be written as vertical compositions of primitive maps, as given in Equation (8); so if we suppose that $\mathbf{f}$ above is primitive, in can be thought of as a one-step proof that $\mathbf{g} \sim \mathbf{h}$, and can be given

inria-00347424, version 1 - 15 Dec 2008

a generic representation by a commuting diagram in $\mathbf{P}X$ like



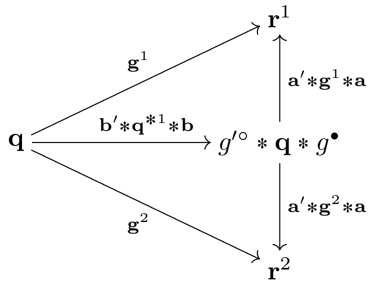
where $\mathbf{p}, \mathbf{p'}$ are maps that can go either up or down, and one of which can be an identity (but not both). Since $\mathbf{f}$ is in $\mathbf{H}X$ more is true: we know that the endpoint maps $\mathbf{p}_{\mathsf{b,b}}, \mathbf{p}_{\mathsf{e,e}}, \mathbf{p}'_{\mathsf{b,b}}, \mathbf{p}'_{\mathsf{e,e}}$ are all identifies in X (it is easy to see that this condition is necessary and sufficient to ensure that a primitive map in $\mathbf{PP}X$ is actually in $\mathbf{H}X$). The commutativity of the diagram then forces the equality of the horizontal endpoint maps:

$$\begin{aligned} \mathbf{g}_{\mathsf{b,b}} \colon \mathbf{q}_{\mathsf{b}} \to \mathbf{q}'_{\mathsf{b}} &= \mathbf{h}_{\mathsf{b,b}} \colon \mathbf{r}_{\mathsf{b}} \to \mathbf{r}'_{\mathsf{b}} \\ \mathbf{g}_{\mathsf{e,e}} \colon \mathbf{q}_{\mathsf{e}} \to \mathbf{q}'_{\mathsf{e}} &= \mathbf{h}_{\mathsf{e,e}} \colon \mathbf{r}_{\mathsf{e}} \to \mathbf{r}'_{\mathsf{e}} \,. \end{aligned}$$

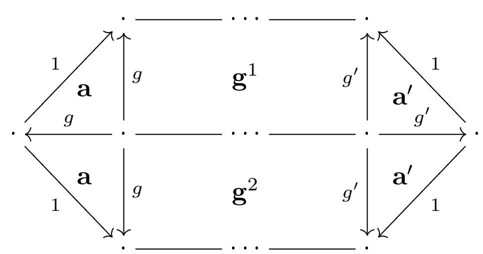
Remark 4. We remind the reader that the coequalizer $\Pi_1 X$ is the category *generated* by the graph $\mathbf{P}X/{\sim}$, since coequalizers in *Cat* are not full in general. But if we manage to show that $\sim$ is a congruence for horizontal composition, $\Pi_1 X$ will be just $\mathbf{P}X/{\sim}$. This is easy to do directly, but we will obtain that result in a more roundabout way in our proof of the main theorem.

Proposition 9. *Let* $\mathbf{g}^1 \colon \mathbf{q} \to \mathbf{r}^1$ *and* $\mathbf{g}^2 \colon \mathbf{q} \to \mathbf{r}^2$ *be such that their endpoint maps obey* $\mathbf{g}^1_{\mathsf{b,b}} = \mathbf{g}^2_{\mathsf{b,b}}$ *and* $\mathbf{g}^1_{\mathsf{e,e}} = \mathbf{g}^2_{\mathsf{e,e}}$*. Then* $\mathbf{g}^1 \sim \mathbf{g}^2$*.*

Proof. Let us first rename $g = \mathbf{g}^i_{\mathsf{b,b}}$ and $g' = \mathbf{g}^i_{\mathsf{e,e}}$. A proof witness that $\mathbf{g}^1 \sim \mathbf{g}^2$ is given by

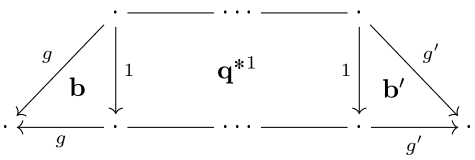


where the “vertical” part of this diagram is



inria-00347424, version 1 - 15 Dec 2008

and the map $\mathbf{b}' * \mathbf{q}^{*1} * \mathbf{b}$ is



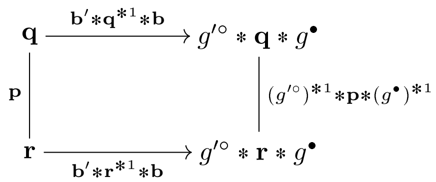
Checking that everything commutes is trivial. □

Proposition 10. *Let* $\mathbf{q}$ *be an object of* $\mathbf{P}X$ *and* $g\colon \mathbf{q}_{\mathsf{b}} \to x, g'\colon \mathbf{q}_{\mathsf{e}} \to x'$ *two maps in* X*. Then there is an extension* $\mathbf{g}\colon \mathbf{q} \to \mathbf{q}'$ *in* $\mathbf{P}X$ *above* (g, g') *(this means that* $\mathbf{g}_{\mathsf{b},\mathsf{b}} = g, \mathbf{g}_{\mathsf{e},\mathsf{e}} = g'$*). Furthermore, if* $\mathbf{r} \sim \mathbf{q}$ *is given, then we can find* $\mathbf{h}\colon \mathbf{r} \to \mathbf{r}'$ *also above* (g, g') *such that there is a proof witness of* $\mathbf{g} \sim \mathbf{h}$

Proof. For the first part just take $\mathbf{g} = \mathbf{b}' * \mathbf{g}^{*1} * \mathbf{b}$, where this is interpreted exactly as in the previous proposition. For the second part, we choose a proof witness $\mathbf{p}$ of $\mathbf{q} \sim \mathbf{r}$. Let us first assume that

- $\mathbf{p}$ is a one-step proof, direction irrelevant. Thus it is a rectangle of arbitrary "width", but with height 1, just as in Diagram (11).
- the two primitive cells at both ends of $\mathbf{p}$ *are not triangles,* i.e., do not have shape 1,2,3 or 4.

We can now construct



and the fact that no triangles appear at the ends of $\mathbf{p}$ guarantees that the square will commute, giving the desired proof witness.

We can now tackle the general case, and assume $\mathbf{p}$ is an arbitrary n-step proof witness. Let $\mathbf{c}, \mathbf{c}'$ be two paths of length one made with an identity morphism, such that $\mathbf{c}' * \mathbf{q} * \mathbf{c}$ is defined (e.g., $\mathbf{c} = 1^{\circ}_{\mathbf{q}_{\mathsf{b}}}$ $(= 1^{\circ}_{\mathbf{r}_{\mathsf{b}}})$, $\mathbf{c}' = 1^{\circ}_{\mathbf{q}_{\mathsf{e}}}$ $(= 1^{\circ}_{\mathbf{r}_{\mathsf{e}}})$. Now "pad" the proof witness $\mathbf{p}$ with unit squares made with $\mathbf{c}, \mathbf{c}'$, and add triangles at the ends:

$$
\begin{array}{ccc}
\mathbf{q} & \searrow & \\
\downarrow {\scriptstyle \triangle * \mathbf{q}^{*1} * \triangle} & & \\
\mathbf{c}' * \mathbf{q} * \mathbf{c} & \xrightarrow{\mathbf{b}' * \mathbf{c}'^{*1} * \mathbf{q}^{*1} * \mathbf{c}^{*1} * \mathbf{b}} & g'^{\circ} * \mathbf{c}' * \mathbf{q} * \mathbf{c} * g^{\bullet} \\
\Big| {\scriptstyle \mathbf{c}'^{*1} * \mathbf{p} * \mathbf{c}^{*1}} & & \Big| {\scriptstyle (g'^{\circ})^{*1} * \mathbf{p} * (g^{\bullet})^{*1}} \\
\mathbf{c}' * \mathbf{r} * \mathbf{c} & \xrightarrow[\mathbf{b}' * \mathbf{c}'^{*1} * \mathbf{r}^{*1} * \mathbf{c}^{*1} * \mathbf{b}]{} & g'^{\bullet} * \mathbf{c}' * \mathbf{r} * \mathbf{c} * g^{\circ} \\
\uparrow {\scriptstyle \triangledown * \mathbf{r}^{*1} * \triangledown} & & \\
\mathbf{r} & \nearrow &
\end{array}
$$

where $\triangle$, *etc.* denote the obvious corresponding triangle composed only of identities. The square in the middle commutes by the just-proved special case applied n times, the two triangles commute by definition, and the whole diagram gives the desired result. □

Theorem 2. *Let* $\mathbf{g}\colon \mathbf{q} \to \mathbf{q}'$ *and* $\mathbf{h}\colon \mathbf{r} \to \mathbf{r}'$ *be maps in* $\mathbf{P}X$. *Then TFAE:*

(i) $\mathbf{g}_{\mathsf{b},\mathsf{b}} = \mathbf{h}_{\mathsf{b},\mathsf{b}}, \mathbf{g}_{\mathsf{e},\mathsf{e}} = \mathbf{h}_{\mathsf{e},\mathsf{e}}$ *and* $\mathbf{q} \sim \mathbf{r}$

(ii) $\mathbf{g}_{\mathsf{b},\mathsf{b}} = \mathbf{h}_{\mathsf{b},\mathsf{b}}, \mathbf{g}_{\mathsf{e},\mathsf{e}} = \mathbf{h}_{\mathsf{e},\mathsf{e}}$ *and* $\mathbf{q}' \sim \mathbf{r}'$,

(iii) $\mathbf{g} \sim \mathbf{h}$,

Proof. Showing the equivalence between (i) and (ii) is very easy: we already know because of Proposition 8 that

$$G\mathbf{g}_{\mathsf{e},\mathsf{e}} = G\mathbf{g}_{\mathsf{b},\mathsf{b}} \circ P\mathbf{q}' \quad \text{and} \quad P\mathbf{r} \circ G\mathbf{h}_{\mathsf{e},\mathsf{e}} = G\mathbf{h}_{\mathsf{b},\mathsf{b}} \circ P\mathbf{r}' \,. \tag{12}$$

Because of Proposition 7 we also know that $\mathbf{q} \sim \mathbf{r}$ iff $P\mathbf{q} = P\mathbf{q}$ and $\mathbf{q}' \sim \mathbf{r}'$ iff $P\mathbf{q}' = P\mathbf{q}'$. Substituting one of of these equations in (12) and cancelling (we are working in a groupoid, remember) gives us the other one.

Let us now show the equivalence between (i) and (iii). Supposing the former, we can apply Proposition (10) and extend two new maps $\mathbf{g}'\colon \mathbf{q} \to \mathbf{q}''$ and $\mathbf{h}'\colon \mathbf{r} \to \mathbf{r}''$ such that $\mathbf{g}' \sim \mathbf{h}'$. But applying Proposition 9 twice, we get us that $\mathbf{g} \sim \mathbf{g}'$, $\mathbf{r} \sim \mathbf{r}'$ and we are done. The reverse implication is trivial.

The proof of Theorem 1 is now a formality: the previous Proposition shows that there is a canonical bijection between the maps of $\mathbf{P}X/{\sim}$ and those of $G \downarrow G$. It is easy to see that this bijection respects composition.

References

1. M. Barr and R. Diaconescu. On locally simply connected toposes and their fundamental groups. *Cahiers Top. Géom. Diff.*, XXII(3):301–314, 1981.

inria-00347424, version 1 - 15 Dec 2008

2. F. Borceux and G. Janelidze. *Galois Theories*, volume 72 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2001.
3. M. Bunge and I. Moerdijk. On the constructin of the Grothendieck fundamental group of a topos by paths. *J. Pure App. Algebra*, 116:99–119, 1997.
4. Ch. Ehresmann. *Catégories et Structures*. Dunod, 1965.
5. P. G. Goerss and J. F. Jardine. *Simplicial Homotopy Theory*, volume 174 of *Progress in Mathematics*. Birkhäuser, 1999.
6. M. Grandis and L. Mauri. Cubical sets and their site. *Theory Appl. Categ.*, 11(8):185–211, 2003.
7. A. Grothendieck. *Revêtements étales et groupe fondamental*, volume 224 of *Lecture Notes in Math.* Springer, 1981.
8. A. Joyal and I. Moerdijk. Toposes as homotopy groupoids. *Advances in Mathematics*, 80:22–38, 1990.
9. Saunders Mac Lane. *Categories for the Working Mathematician*. Springer-Verlag, Berlin, 1971.
10. I. Moerdijk and G. Wraith. Connected locally connected toposes are path-connected. *Trans. Am. Math. Soc.*, 295(2):849–859, June 1986.
11. D. Quillen. Higher algebraic K-theory. In H. Bass, editor, *Higher K-Theory*, volume 147 of *Springer Lect. Notes Math.*, pages 85–147. 1973.
12. E. M. Spanier. *Algebraic Topology*. McGraw-Hill Series in Higher Mathematics. McGraw-Hill, 1966. Reprinted by Springer Verlag.

inria-00347424, version 1 - 15 Dec 2008