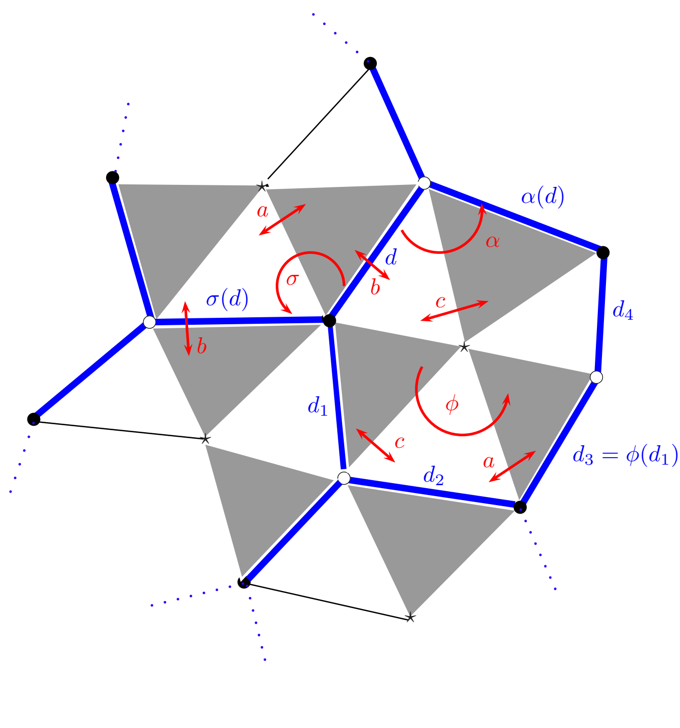
AN ELEMENTARY APPROACH TO DESSINS D'ENFANTS AND THE GROTHENDIECK-TEICHMÜLLER GROUP

PIERRE GUILLOT

ABSTRACT. We give an account of the theory of dessins d'enfants which is both elementary and self-contained. We describe the equivalence of many categories (graphs embedded nicely on surfaces, finite sets with certain permutations, certain field extensions, and some classes of algebraic curves), some of which are naturally endowed with an action of the absolute Galois group of the rational field. We prove that the action is faithful. Eventually we prove that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ embeds into the Grothendieck-Teichmüller group $\widehat{\mathcal{GT}}_0$ introduced by Drinfeld. There are explicit approximations of $\widehat{\mathcal{GT}}_0$ by finite groups, and we hope to encourage computations in this area.

Our treatment includes a result which has not appeared in the literature yet: the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on the subset of *regular* dessins – that is, those exhibiting maximal symmetry – is also faithful.



Introduction

The story of *dessins d'enfants* (children's drawings) is best told in two episodes.

The first side of the story is a surprising unification of different-looking theories: graphs embedded nicely on surfaces, finite sets with certain permutations, certain field extensions, and some classes of algebraic curves (some over $\mathbb{C}$, some over $\overline{\mathbb{Q}}$), all turn out to define equivalent categories. This result follows from powerful and yet very classical theorems, mostly from the 19th century, such as the correspondence between Riemann surfaces and their fields of meromorphic functions (of course known to Riemann himself), or the basic properties of the fundamental group (dating back to Poincaré).

One of our goals with the present paper is to give an account of this theory that sticks to elementary methods, as we believe it should. (For example we shall never need to appeal to "Weil's rigidity criterion", as is most often done in the literature on the subject; note that it is also possible, in fact, to read most of this paper without any knowledge of algebraic curves.) Our development is moreover as self-contained as is reasonable: that is, while this paper is not the place to develop the theory of Riemann surfaces, Galois extensions or covering spaces from scratch – we shall refer to basic textbooks for these – we give complete arguments from there. Also, we have striven to state the results in terms of actual equivalences of categories, a slick language which unfortunately is not always employed in the usual sources.

The term *dessins d'enfants* was coined by Grothendieck in [Gro97], in which a vast programme was laid out, giving the theory a new thrust which is the second side of the story we wish to tell. In a nutshell, some of the categories mentioned above naturally carry an action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, the absolute Galois group of the rational field. This group therefore acts on the set of isomorphism classes of objects in any of the equivalent categories; in particular one can define an action of the absolute Galois group on graphs embedded on surfaces. In this situation however, the nature of the Galois action is really very mysterious - it is hoped that, by studying it, light may be shed on the structure of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. It is the opportunity to bring some kind of basic, visual geometry to bear in the study of the absolute Galois group that makes *dessins d'enfants* – embedded graphs – so attractive.

In this paper we explain carefully, again relying only on elementary methods, how one defines the action, and how one proves that it is *faithful*. This last property is clearly crucial if we are to have any hope of studying $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ by considering graphs. We devote some space to the search for invariants of dessins belonging to the same Galois orbit, a major objective in the field.

When a group acts faithfully on something, we can usually obtain an embedding of it in some automorphism group. In our case, this leads to the *Grothendieck-Teichmüller group* $\widehat{\mathcal{GT}}$, first introduced by Drinfeld in [Dri90], and proved to contain $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ by Ihara in [Iha94]. While trying to describe Ihara's proof in any detail would carry us beyond the scope of this paper, we present a complete, elementary argument establishing that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ embeds into the slightly larger group $\widehat{\mathcal{GT}}_0$ also defined by Drinfeld. In fact we work with a group $\mathcal{GT}$ isomorphic to $\widehat{\mathcal{GT}}_0$, and which is an inverse limit

$$\mathcal{GT} = \lim_{n} \mathcal{GT}(n)\,;$$

here $\mathcal{GT}(n)$ is a certain subgroup of $Out(H_n)$ for an explicitly defined finite group H_n. So describing H_n and $\mathcal{GT}(n)$ for some n large enough gives rough information about $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ – and it is possible to do so in finite time.

In turn, we shall see that understanding H_n amounts, in a sense, to understanding all finite groups generated by two elements, whose order is less than n. We land back on our feet: from the first part of this paper, those groups are in one to one correspondence with some embedded graphs, called regular, exhibiting maximal symmetry. The classification of "regular maps", as they are sometimes called, is a classical topic which is still alive today.

⋆⋆⋆

Let us add a few informal comments of historical nature, not written by an expert in the history of mathematics.

The origin of the subjet is the study of "maps", a word meaning graphs embedded on surfaces in a certain way, the complement of the graph being a disjoint union of topological discs which may be reminiscent of countries on a map of the world. Attention has focused quickly on "regular maps", that is, those for which the automorphism group is as large as possible. For example, "maps" are mentioned in the 1957 book [CM57] by Coxeter and Moser, and older references can certainly be found. The 1978 paper [JS78] by Jones and Singerman has gained a lot of popularity; it gave the field stronger foundations, and already established bijections between "maps" and combinatorial objects such as permutations on the one hand, and also with compact Riemann surfaces, and thus complex algebraic curves, on the other hand. For a recent survey on the classification of "maps", see [Š13].

Then came the *Esquisse d'un programme* [Gro97], written by Grothendieck between 1972 and 1984. Dessins can be seen as algebraic curves over $\mathbb{C}$ with some extra structure (a morphism to $\mathbb{P}^1$ with ramification above $0, 1$ or ∞ only), and Grothendieck knew that such a curve must be defined over $\overline{\mathbb{Q}}$. Since then, this remark has been known as "the obvious part of Belyi's theorem" by people working in the field, even though it is not universally recognized as obvious, and has little to do with Belyi (one of the first complete and rigorous proofs is probably that by Wolfart [Wol97]). However, Grothendieck was very impressed by the simplicity and strength of a result by Belyi [Bel79] stating that, conversely, *any* algebraic curve defined over $\overline{\mathbb{Q}}$ can be equipped with a morphism as above (which is nowadays called a Belyi map, while it has become common to speak of Belyi's theorem to mean the equivalence of definability of $\overline{\mathbb{Q}}$ on the one hand, and the possibility of finding a Belyi map on the other hand). Thus the theory of dessins encompasses all curves over $\overline{\mathbb{Q}}$, and Grothendieck pointed out that this simple fact implied that the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on dessins must be faithful. The *esquisse* included many more ideas which will not be discussed here. For a playful exposition of many examples of the Galois action on dessins, see [LZ04].

Later, in 1990, Drinfeld defined $\widehat{\mathcal{GT}}$ in [Dri90] and studied its action on braided categories, but did not relate it explicitly to $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ although the motivation for the definition came from the *esquisse*. It was Ihara in 1994 [Iha94] who proved the existence of an embedding of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ into $\widehat{\mathcal{GT}}$; it is interesting to note that, if dessins d'enfants were the original idea for Ihara's proof, they are a little hidden behind the technicalities.

The Grothendieck-Teichmüller group has since been the object of much research, quite often using the tools of quantum algebra in the spirit of Drinfeld's original approach. See also [Fre] by Fresse, which establishes an interpretation of $\widehat{\mathcal{GT}}$ in terms of operads.

⋆⋆⋆

Here is an outline of the paper. In section 1, we introduce cell complexes, that is, spaces obtained by glueing discs to bipartite graphs; when the result is a topological surface, we have a *dessin*. In the same section we explain that dessins are entirely determined by two permutations. In section 2, we quote celebrated, classical results that establish a number of equivalences of categories between that of dessins and many others, mentioned above. In section 3 we study the regularity condition in detail. The Galois action is introduced in section 4, where we also present some concrete calculations. We show that the action is faithful. Finally in section 5 we prove that $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ embeds into the group $\mathcal{GT}$ described above.

In the course of this final proof, we obtain seemingly for free the following refinement: the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on *regular* dessins is also faithful. This fact follows mostly from a 1980 result by Jarden [Jar80] (together with known material on dessins), and it is surprising that it has not been mentioned in the literature yet. While this work was in its last stages, I have learned from Gareth Jones that the preprint [JZGD] by Andrei Jaikin-Zapirain and Gabino Gonzalez-Diez contains generalizations of Jarden's theorem while the faithfulness of the Galois action on regular dessins is explicitly mentioned as a consequence (together with more precise statements). Also in [BCG], a preprint by Ingrid Bauer, Fabrizio Catanese and Fritz Grunewald, one finds the result stated.

Acknowledgements. Nick Gill and Ian Short have followed the development of this paper from the very early stages, and I have benefited greatly from their advice. I also want to thank Gareth Jones for kind words about this work as it was reaching completion. Further corrections have been made based on comments by Olivier Guichard and Pierre de la Harpe, for which I am grateful.

Contents

1. Dessins

In this section we describe the first category of interest to us, which is that of graphs embedded on surfaces in a particularly nice way. These have been called sometimes "maps" in the literature, a term which one should avoid if possible given the other meaning of the word "map" in mathematics. We call them *dessins*.

The reader may be surprised by the number of pages devoted to this first topic, and the level of details that we go into. Would it not suffice to say that the objects we study are graphs embedded on surfaces, whose complement is a union of open discs, perhaps with just a couple of technical conditions? (A topologist would say "a CW-complex structure on a surface".)

This would not be appropriate, for several reasons. First and foremost, we aim at proving certain equivalences of categories, eventually (see next section). With the above definition, whether one takes as morphisms all continuous maps between surfaces, or restricts attention to the "cellular" ones, in any case there are simply too many morphisms taken into account (see for example [JS78]). Below, we get things just right.

Another reason is that we already present *two* categories in this section, not just one: dessins are intimately related to finite sets endowed with certain permutations. The two categories are equivalent and indeed so close that we encourage the reader to always think of these two simultaneously; we take the time to build the intuition for this.

Note also that our treatment is very general, including non-orientable dessins as well as dessins on surfaces with boundary.

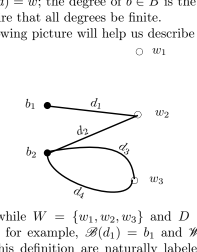
Finally, the material below is so elementary that it was possible to describe it with absolutely no reference to textbooks, an opportunity we took. We think of the objects defined in this section as the most down-to-earth of the paper, while the other categories to be introduced later are here to shed light on dessins.

1.1. **Bipartite graphs.** We start with the definition of *bipartite graphs*, or *bigraphs* for short, which are essentially graphs made of black and white vertices, such that the edges only connect vertices of different colours. More formally, a bigraph consists of

- a set B, the elements of which we call the black vertices,
- a set W, the elements of which we call the white vertices,
- a set D, the elements of which we call the *darts*,
- two maps $\mathscr{B}: D \longrightarrow B$ and $\mathscr{W}: D \longrightarrow W$.

In most examples all of the above sets will be finite, but in general we only specify a local finiteness condition, as follows. The degree of $w \in W$ is the number of darts d such that $\mathscr{W}(d) = w$; the degree of $b \in B$ is the number of darts d such that $\mathscr{B}(d) = b$. We require that all degrees be finite.

For example, the following picture will help us describe a bigraph.



Here $B = \{b_1, b_2\}$, while $W = \{w_1, w_2, w_3\}$ and $D = \{d_1, d_2, d_3, d_4\}$. The maps $\mathscr{B}$ and $\mathscr{W}$ satisfy, for example, $\mathscr{B}(d_1) = b_1$ and $\mathscr{W}(d_1) = w_2$. Note that bigraphs according to this definition are naturally labeled, even though we will often suppress the names of the vertices and darts in the pictures.

The notion of morphism of bigraphs is the obvious one: a morphism between $\mathscr{G} = (B, W, D, \mathscr{B}, \mathscr{W})$ and $\mathscr{G}' = (B', W', D', \mathscr{B}', \mathscr{W}')$ is given by three maps $B \to B'$, $W \to W'$ and $\Delta: D \to D'$ which are compatible with the maps $\mathscr{B}, \mathscr{W}, \mathscr{B}', \mathscr{W}'$. Isomorphisms are invertible morphisms, unsurprisingly. (Pedantically, one could define an unlabeled bigraph to be an isomorphism class of bigraphs.)

To a bigraph $\mathscr{G}$ we may associate a topological space $|\mathscr{G}|$, by attaching intervals to discrete points according to the maps $\mathscr{B}$ and $\mathscr{W}$; in the above example, and in all others, it will look just like the picture. To this end, take for each $d \in D$ a copy I_d of the unit interval $[0, 1]$ with its usual topology. Then consider

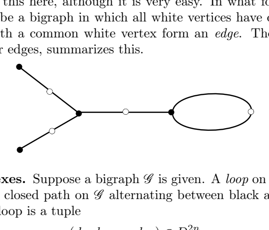
$$Y = \coprod_{d \in D} I_d$$

with the disjoint union topology, and

$$X = Y \coprod B \coprod W .$$

(Here B and W are given the discrete topology.) On X there is an equivalence relation corresponding to the identifications imposed by the maps $\mathscr{B}$ and $\mathscr{W}$. In other words, the equivalence class $[b]$ of $b \in B$ is such that $[b] \cap I_d = \{0\}$ if $\mathscr{B}(d) = b$ and $[b] \cap I_d = \emptyset$ otherwise, while $[b] \cap B = \{b\}$ and $[b] \cap W = \emptyset$; the description of the equivalence class $[w]$ when $w \in W$ is analogous, with $[w] \cap I_d = \{1\}$ precisely when $\mathscr{W}(d) = w$. All the other equivalence classes are singletons. The space $|\mathscr{G}|$ is the set of equivalence classes, with the quotient topology. Clearly, an isomorphism of graphs induces a homeomorphism between their topological realizations.

Finally, we point out that usual graphs (the reader may pick their favorite definition) can be seen as bigraphs by "inserting a white vertex inside each edge". We will not formalize this here, although it is very easy. In what follows we officially define a graph to be a bigraph in which all white vertices have degree precisely 2; a pair of darts with a common white vertex form an *edge*. The next picture, on which you see four edges, summarizes this.

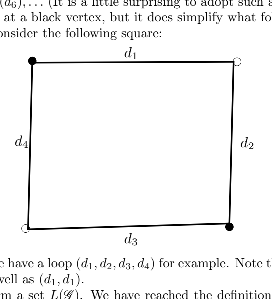


1.2. **Cell complexes.** Suppose a bigraph $\mathscr{G}$ is given. A *loop* on $\mathscr{G}$ is a sequence of darts describing a closed path on $\mathscr{G}$ alternating between black and white vertices. More precisely, a loop is a tuple

$$(d_1, d_2, \ldots, d_{2n}) \in D^{2n}$$

such that $\mathscr{W}(d_{2i+1}) = \mathscr{W}(d_{2i+2})$ and $\mathscr{B}(d_{2i+2}) = \mathscr{B}(d_{2i+3})$, for $0 \le i \le n-1$, where d_{2n+1} is to be understood as d_1. We think of this loop as starting and ending with the black vertex $\mathscr{B}(d_1)$, and visiting along the way the points $\mathscr{W}(d_2)$, $\mathscr{B}(d_3)$, $\mathscr{W}(d_4)$, $\mathscr{B}(d_5)$, $\mathscr{W}(d_6)$, … (It is a little surprising to adopt such a convention, that loops always start at a black vertex, but it does simplify what follows.)

For example, consider the following square:



On this bigraph we have a loop (d_1, d_2, d_3, d_4) for example. Note that (d_1, d_2, d_2, d_1) is also a loop, as well as (d_1, d_1).

Loops on $\mathscr{G}$ form a set $L(\mathscr{G})$. We have reached the definition of a *cell complex* (or 2-cell complex, for emphasis). This consists of

- a bigraph $\mathscr{G}$,
- a set F, the elements of which we call the *faces*,
- a map $\partial: F \to L(\mathscr{G})$, called the boundary map.

The definition of morphisms between cell complexes will wait a little.

A cell complex $\mathscr{C}$ also has a topological realization $|\mathscr{C}|$: briefly, one attaches closed discs to the space $|\mathscr{G}|$ using the specified boundary maps. In more details, for each $f \in F$ we pick a copy $\mathbb{D}_f$ of the unit disc

$$\mathbb{D} = \{z \in \mathbb{C} : |z| \le 1\} .$$

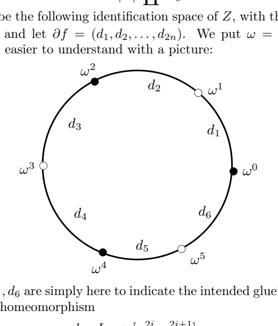
Consider then

$$Z_0 = \coprod_{f \in F} \mathbb{D}_f$$

and

$$Z = |\mathscr{G}| \coprod Z_0 .$$

We define $|\mathscr{C}|$ to be the following identification space of Z, with the quotient topology. Fix $f \in F$ and let $\partial f = (d_1, d_2, \ldots, d_{2n})$. We put $\omega = e^{\frac{2i\pi}{2n}} \in \mathbb{D}_f$. The discussion will be easier to understand with a picture:



The letters $d_1, \ldots, d_6$ are simply here to indicate the intended glueing. Let $I = [0, 1]$ and consider the homeomorphism

$$h_i: I \longrightarrow [\omega^{2i}, \omega^{2i+1}] ,$$

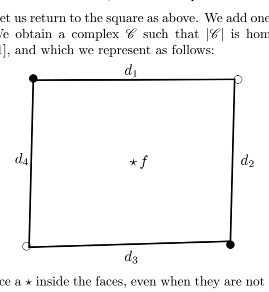
where $[\omega^{2i}, \omega^{2i+1}]$ denotes the circular arc from ω^{2i} to ω^{2i+1}, defined by $h_i(t) = \omega^{2i+t}$. We shall combine it with the continuous map

$$g_i: I \longrightarrow |\mathscr{G}|$$

which is obtained as the identification $I = I_{d_{2i+1}}$ followed by the canonical map $I_{d_{2i+1}} \to |\mathscr{G}|$ (see the definition of $|\mathscr{G}|$). We can now request, for all $t \in I$, the identification of $g_i(t)$ and $h_i(t)$, these being both points of Z.

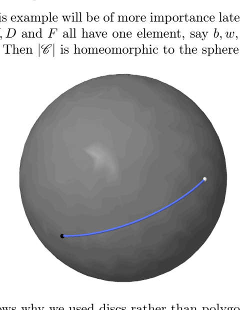
Similarly there is an identification of the arc $[\omega^{2i}, \omega^{2i-1}]$ with the image of $I_{d_{2i}}$. We prescribe no more identifications, and this completes the definition of $|\mathscr{C}|$.

Example 1.1 – Let us return to the square as above. We add one face f, with $\partial f = (d_1, d_2, d_3, d_4)$. We obtain a complex $\mathscr{C}$ such that $|\mathscr{C}|$ is homeomorphic to the square $[0, 1] \times [0, 1]$, and which we represent as follows:



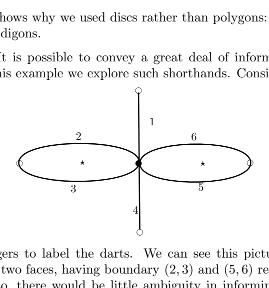
We shall often place a $\star$ inside the faces, even when they are not labeled, to remind the reader to mentally fill in a disc. The reader is invited to contemplate how the complex obtained by taking, say, $\partial f = (d_2, d_1, d_4, d_3)$ instead, produces indeed a homeomorphic realization. These two complexes ought to be isomorphic, when we have defined what isomorphisms are.

Example 1.2 – This example will be of more importance later than is immediately apparent. Let B, W, D and F all have one element, say b, w, d and f respectively; and let $\partial f = (d, d)$. Then $|\mathscr{C}|$ is homeomorphic to the sphere S^2.



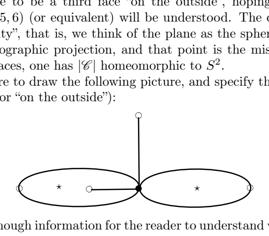
This example shows why we used discs rather than polygons: we may very well have to deal with digons.

Example 1.3 – It is possible to convey a great deal of information by pictures alone, and with this example we explore such shorthands. Consider for example:

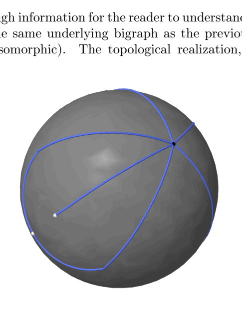


Here we use integers to label the darts. We can see this picture as depicting a cell complex with two faces, having boundary $(2, 3)$ and $(5, 6)$ respectively. Should we choose to do so, there would be little ambiguity in informing the reader that we mean for there to be a third face "on the outside", hoping that the boundary $(1, 1, 2, 3, 4, 4, 5, 6)$ (or equivalent) will be understood. The centre of that face is placed "at infinity", that is, we think of the plane as the sphere S^2 with a point removed via stereographic projection, and that point is the missing $\star$. Of course with these three faces, one has $|\mathscr{C}|$ homeomorphic to S^2.

Suppose we were to draw the following picture, and specify that there is a third face "at infinity" (or "on the outside"):



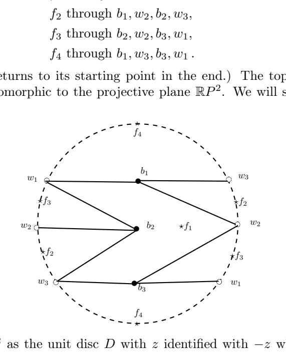
This is probably enough information for the reader to understand which cell complex we mean. (It has the same underlying bigraph as the previous one, but the cell complexes are not isomorphic). The topological realization, again a sphere, is represented below.



Example 1.4 – It is harder to draw pictures in the following case. Take $B = \{b_1, b_2, b_3\}$ and $W = \{w_1, w_2, w_3\}$, and add darts so that $\mathscr{G}$ is "the complete bipartite graph on 3 + 3 vertices" : that is, place a dart between each b_i and each w_j, for $1 \le i, j \le 3$. Since there are no multiple darts between any two vertices in this bigraph, we can designate a dart by its endpoints; we may also describe a loop by simply giving the list of vertices that it visits. With this convention, we add four faces:

f_1 through $b_1, w_2, b_3, w_3, b_2, w_1$,
f_2 through b_1, w_2, b_2, w_3,
f_3 through b_2, w_2, b_3, w_1,
f_4 through b_1, w_3, b_3, w_1.

(Each of these returns to its starting point in the end.) The topological realization $|\mathscr{C}|$ is homeomorphic to the projective plane $\mathbb{R}P^2$. We will show this with a picture:



Here we see $\mathbb{R}P^2$ as the unit disc D with z identified with $-z$ whenever $|z| = 1$; we caution that the dotted arcs, indicating the boundary of the unit circle, are not darts.

Here are some very basic properties of the geometric realization.

Proposition 1.5 – (1) *The space $|\mathscr{C}|$ is connected if and only if $|\mathscr{G}|$ is.*
(2) *The space $|\mathscr{C}|$ is compact if and only if the complex is finite (ie B, W, D and F are all finite).*

Proof. (1) It is quite easy to prove this directly, after showing that each path on $|\mathscr{C}|$ is homotopic to one lying on $|\mathscr{G}|$. The reader who has recognized that the space $|\mathscr{C}|$ is, by definition, the realization of a CW-complex, whose 1-skeleton is $|\mathscr{G}|$, will see the result as a consequence of the cellular approximation theorem ([Bre97], Theorem 11.4).

(2) By construction there is a quotient map

$$q: K = Y \coprod B \coprod W \coprod Z_0 \longrightarrow |\mathscr{C}| ,$$

where the notation is as above. Clearly K is compact if the complex is finite, so $q(K) = |\mathscr{C}|$ must be compact, too, and we have proved that the condition is sufficient.

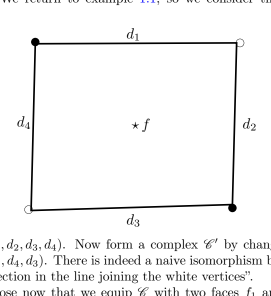
To see that it is necessary as well, one can argue that the map q is proper, or else use elementary arguments as follows. We show that the faces must be finite in number when $|\mathscr{C}|$ is compact, and the reader will do similarly with the vertices and darts.

For each $f \in F$, consider the open set $U_f \subset K$ whose complement is the union of the closed discs of radius $\frac{1}{2}$ in all the discs $D_{f'}$ for $f' \ne f$ (this complement is closed by definition of the disjoint union topology). By definition of the quotient topology $q(U_f)$ is open in $|\mathscr{C}|$, and the various open sets $q(U_f)$ form a cover of $|\mathscr{C}|$ (each $q(U_f)$ is obtained by removing a closed disc from each face of $|\mathscr{C}|$ but one). By compactness, finitely many of them will cover the space, and so finitely many of the open sets U_f will cover K. It follows that F is finite. □

1.3. **Morphisms between cell complexes; triangulations.** Let us start with a provisional definition: a *naive morphism* between $\mathscr{C} = (\mathscr{G}, F, \partial)$ and $\mathscr{C}' = (\mathscr{G}', F', \partial')$ is given by a morphism $\mathscr{G} \to \mathscr{G}'$ together with a map $\Phi: F \to F'$ such that $\partial'\Phi(f) = \Delta(\partial f)$ for $f \in F$; here the map $\Delta: D \to D'$ has been extended to the set $L(\mathscr{G})$ in the obvious way. With this definition, it is clear that naive morphisms induce continuous maps between the topological realizations.

However this definition does not allow enough morphisms. Let us examine this.

Example 1.6 – We return to example 1.1, so we consider the cell complex $\mathscr{C}$ depicted below:



Here $\partial f = (d_1, d_2, d_3, d_4)$. Now form a complex $\mathscr{C}'$ by changing only ∂ to ∂', with $\partial' f = (d_2, d_1, d_4, d_3)$. There is indeed a naive isomorphism between $\mathscr{C}$ and $\mathscr{C}'$, given by "the reflection in the line joining the white vertices".

However, suppose now that we equip $\mathscr{C}$ with two faces f_1 and f_2 (leaving the bigraph unchanged) with $\partial f_1 = (d_1, d_2, d_3, d_4) = \partial f_2$; then $|\mathscr{C}|$ is the sphere S^2. On the other hand consider $\mathscr{C}'$ having the same bigraph, and two faces satisfying $\partial f_1' = (d_1, d_2, d_3, d_4)$ and $\partial f_2' = (d_2, d_1, d_3, d_4)$. Then it is readily checked that there is no naive isomorphism between $\mathscr{C}$ and $\mathscr{C}'$.

This is disappointing, as we would like to see these two as essentially "the same" complexes. More generally we would like to think of the boundaries of the faces in a cell complex as not having a distinguished (black) starting point, and not having a particular direction.

The following better definition will be sufficient in many situations. A *lax morphism* between $\mathscr{C} = (\mathscr{G}, F, \partial)$ and $\mathscr{C}' = (\mathscr{G}', F', \partial')$ is given by a morphism $\mathscr{G} \to \mathscr{G}'$ together with a map $\Phi: F \to F'$ with the following property. If $f \in F$ with $\partial f = (d_1, \ldots, d_{2n})$, and if $\partial'\Phi(f) = (d_1', \ldots, d_{2m}')$, then

$$\Delta(\{d_1, \ldots, d_{2n}\}) = \{d_1', \ldots, d_{2m}'\} ,$$

where Δ is the map $D \to D'$. So naive morphisms are lax morphisms, but not conversely.

Example 1.7 – Resuming the notation of the last example, the identity on $\mathscr{G}$ and the bijection $F \to F'$, $f_1 \mapsto f_1'$, $f_2 \mapsto f_2'$, together define a lax isomorphism between $\mathscr{C}$ and $\mathscr{C}'$.

It is not immediate how lax morphisms can be used to induce continuous maps. Moreover, the following phenomenon must be observed.

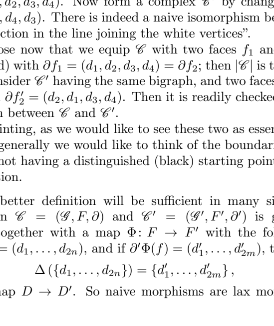
Example 1.8 – We build a bigraph $\mathscr{G}$ with only one black vertex, one white vertex, and two darts d_1 and d_2 between them; $|\mathscr{G}|$ is a circle. Turn this into a cell complex $\mathscr{C}$ by adding one face f with $\partial f = (d_1, d_2, d_1, d_2)$. The topological realization $|\mathscr{C}|$ is obtained by taking a copy of the unit disc $\mathbb{D}$, and identifying z and $-z$ when $|z| = 1$: in other words, $|\mathscr{C}|$ is the real projective plane $\mathbb{R}P^2$.

Now consider the map $z \mapsto -z$, from $\mathbb{D}$ to itself, and factor it through $\mathbb{R}P^2$; it gives a self-homeomorphism of $|\mathscr{C}|$. The latter cannot possibly be induced by a lax morphism, for it is the identity on $|\mathscr{G}|$: to define a corresponding lax isomorphism we would have to define the self maps of B, W and F to be the identity. Assuming that we had chosen a procedure to get a continuous map from a lax morphism, surely the identity would induce the identity.

However the said self-homeomorphism of $\mathbb{R}P^2$ is simple enough that we would like to see it corresponding to an isomorphism of $\mathscr{C}$.

Our troubles seem to arise when repeated darts show up in the boundary of a single face. We solve the problem by subdividing the faces, obtaining the canonical *triangulation* of our objects.

Let $\mathscr{C}$ be a cell complex. We may triangulate the faces of $|\mathscr{C}|$ by adding a point in the interior of each face (think of the point marked $\star$ in the pictures), and connecting it to the vertices on the boundary. More precisely, for each face f, with $\partial f = (d_1, \ldots, d_{2n})$, we identify $2n$ subspaces of $|\mathscr{C}|$, each homeomorphic to a triangle, as the images under the canonical quotient map of the sectors obtained on the unit disc in the fashion described on the picture below for $n = 3$. We denote them t_i^f with $1 \le i \le 2n$.



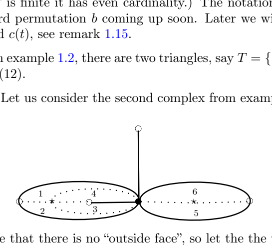
(As before the labels d_i indicate the intended gluing, while the sector bearing the name t_i^f will map to that subspace under the quotient map.) The space $|\mathscr{C}|$ is thus triangulated, yet it is not necessarily (the realization of) a simplicial complex, as distinct triangles may have the same set of vertices, as in example 1.2. This same example exhibits another relevant pathology, namely that the disc corresponding to a face might well map to something which is not homeomorphic to a disc anymore (viz. the sphere), while the triangles actually cut the space $|\mathscr{C}|$ into "easy" pieces. It also has particularly nice combinatorial properties.

We write T for the set of all triangles in the complex. We think of T as an indexing set, much like B, W, D or F. One can choose to adopt a more combinatorial approach, letting $t_1^f, \ldots, t_{2n}^f$ be (distinct) symbols attached to the face f whose boundary is $(d_1, \ldots, d_{2n})$, with T the set of all symbols. There is a map $\mathscr{D}: T \to D$ which associates t_i^f with $\mathscr{D}(t_i^f) = d_i$, there is also a map $\mathscr{F}: T \to F$ with $\mathscr{F}(t_i^f) = f$. We will gradually use more and more geometric terms when referring to the triangles, but it is always possible to translate them into combinatorial relations.

Each $t \in T$ has vertices which we may call $\bullet$, $\circ$ and $\star$ unambiguously. Its sides will be called $\bullet - \circ$, $\star - \bullet$ and $\star - \circ$. Each t also has a neighbouring triangle obtained by reflecting in the $\star - \bullet$ side; call it $a(t)$. Likewise, we may reflect in the $\star - \circ$ side and obtain a neighbouring triangle, which we call $c(t)$. In other words, T comes equipped with two permutations a and c, of order two and having no fixed points. (In particular if T is finite it has even cardinality.) The notation a, c is standard, and there is a third permutation b coming up soon. Later we will write t^a and t^c instead of $a(t)$ and $c(t)$, see remark 1.15.

Example 1.9 – In example 1.2, there are two triangles, say $T = \{1, 2\}$, and $a = c =$ the transposition (12).

Example 1.10 – Let us consider the second complex from example 1.3, that is let us have a look at



Let us first assume that there is no "outside face", so let the triangles be numbered from 1 to 6. The permutation a is then

$$a = (14)(23)(56) ,$$

while

$$c = (12)(34)(56) .$$

If one adds a face at infinity, there are six new triangles, and the permutations a and c change accordingly. We leave this as an exercise.

We have at long last arrived at the official definition of a *morphism* between $\mathscr{C} = (\mathscr{G}, F, \partial)$ and $\mathscr{C}' = (\mathscr{G}', F', \partial')$. We define this to be given by a morphism $\mathscr{G} \to \mathscr{G}'$ (thus including a map $\Delta: D \to D'$) and a map $\Theta: T \to T'$ which

(1) verifies that for each triangle t, one has $\mathscr{D}'(\Theta(t)) = \Delta(\mathscr{D}(t))$,
(2) is compatible with the permutations a and c, that is $\Theta(a(t)) = a(\Theta(t))$ and $\Theta(c(t)) = c(\Theta(t))$.

It is immediate that morphisms induce continuous maps between the topological realizations. These continuous maps restrict to homeomorphisms between the triangles.

Should this definition appear too complicated, we hasten to add:

Lemma 1.11 – *Let $\mathscr{C}$ be a cell complex such that, for each face f with $\partial f = (d_1, \ldots, d_{2n})$, the darts $d_1, \ldots, d_{2n}$ are distinct. Let $\mathscr{C}'$ be another cell complex with the same property. Then any lax morphism between $\mathscr{C}$ and $\mathscr{C}'$ defines a unique morphism, characterized by the property that $\mathscr{F}(\Theta(t)) = \Phi(\mathscr{F}(t))$ for every triangle t.*

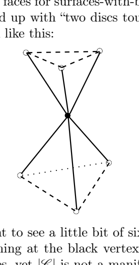
(Recall that lax morphisms have a map Φ between the sets of faces, and morphisms have a map Θ between the sets of triangles.)

Many cell complexes in practice satisfy the property stated in the lemma, and for these we specify morphisms by giving maps $B \to B'$, $W \to W'$, $D \to D'$, and $F \to F'$.

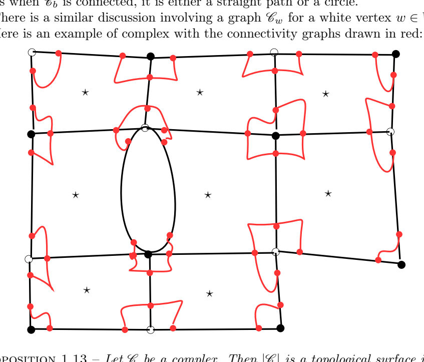
Proof. Any triangle t in $\mathscr{C}$ is now entirely determined by the face $\mathscr{F}(t)$ and the dart $\mathscr{D}(t)$; the same can be said of triangles in $\mathscr{C}'$. So $\Theta(t)$ must be defined as the only triangle t' such that $\mathscr{F}(t')$ and $\mathscr{D}(t')$ are appropriate (in symbols $\mathscr{F}(t') = \Phi(\mathscr{F}(t))$ and $\mathscr{D}(t') = \Delta(\mathscr{D}(t))$). The definition of lax morphisms guarantees the existence of t'.

That Θ is compatible with a and c is automatic here. Indeed $a(t)$ is the only triangle such that $\mathscr{F}(a(t)) = \mathscr{F}(t)$ and such that $\mathscr{D}(a(t))$ has the same black vertex as $\mathscr{D}(t)$. An analogous property is true of both $\Theta(a(t))$ and $a(\Theta(t))$, which must be equal. Likewise for c. □

[illegible]



[illegible]



[illegible]

Lemma 1.17 – *Let $\mathscr{C}$ be a compact dessin. Then $|\mathscr{C}|$ is connected if and only if the full cartographic group acts transitively on the set of triangles.*

Proof. Let $T_1, T_2, \ldots$ be the orbits of G in T, and let $X_i \subset |\mathscr{C}|$ be the union of the triangles in T_i. Each X_i is compact as a finite union of compact triangles, hence X_i is closed in $|\mathscr{C}|$. Also, $|\mathscr{C}|$ is the union of the X_i's, since a dessin does not have isolated vertices (condition (1) above).

Thus we merely have to prove that the X_i's are disjoint. However when two triangles intersect, they do so along an edge, and then an element of G takes one to the other. □

1.6. **Orientations.**

Proposition 1.18 – *Let $\mathscr{C}$ be a compact, connected dessin. Then the surface $|\mathscr{C}|$ is orientable if and only if it is possible to assign a colour to each triangle, black or white, in such a way that two triangles having a side in common are never of the same colour.*

Proof. We give a proof in the case when there is no boundary, leaving the general case as an exercise. We use some standard results in topology, first and foremost: $|\mathscr{C}|$ is orientable if and only if

$$H_2(|\mathscr{C}|, \mathbb{Z}) \ne 0 .$$

To compute this group we use cellular homology. More precisely, we exploit the CW-complex structure on $|\mathscr{C}|$ for which the two-cells are the triangles (of course this space also has a CW-complex in which the two-cells are the faces, but this is not relevant here). Recall that from an earlier remark that simplicial homology is not directly applicable.

We need to orient the triangles, and thus declare that the positive orientation is $\star - \bullet - \circ$; likewise, we decide to orient the 1-cells in such a fashion that $\star - \bullet$, $\bullet - \circ$ and $\circ - \star$ are oriented from the first named 0-cell to the second. Writing ∂ for the boundary in cellular homology, we have then

$$(*) \qquad \partial t = [\star - \bullet] + [\bullet - \circ] + [\circ - \star] ,$$

in notation which we hope is suggestive.

So let us assume that there is a 2-chain

$$(**) \qquad \sigma = \sum_{t \in T} n_t t \ne 0 ,$$

where $n_t \in \mathbb{Z}$, such that $\partial\sigma = 0$. Suppose t is such that $n_t \ne 0$. From (*), we know the coefficients of the neighbours of t in σ, namely

$$n_{t^a} = n_{t^b} = n_{t^c} = -n_t .$$

Since the full cartographic group acts transitively on T by the last lemma, it follows that for each $t' \in T$, the coefficient $n_{t'}$ is determined by n_t, and in fact $n_{t'} = \pm n_t$.

Now let triangles t' such that $n_{t'} > 0$ be black, and let the others be white. We have coloured the triangles as requested. The converse is no more difficult: given the colours, let $n_t = 1$ if t is black and -1 otherwise. Then the 2-chain defined by (**) is non-zero and has zero boundary, so the homology is non-zero. □

When $|\mathscr{C}|$ is orientable, we will call an *orientation* of $\mathscr{C}$ a colouring as above; there are precisely two orientations on a connected, orientable dessin. An isomorphism will be said to *preserve orientations* when it sends black triangles to black triangles. Note the following:

Lemma 1.19 – *A morphism $\mathscr{C} \to \mathscr{C}'$, where $\mathscr{C}$ and $\mathscr{C}'$ are oriented dessins, preserves the orientations if and only if Θ sends black triangles to black triangles, and white triangles to white triangles.*

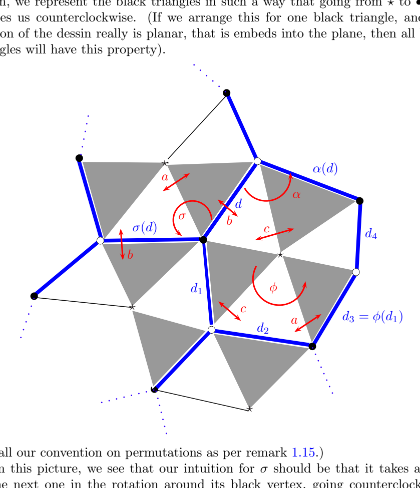
1.7. **More permutations.** Suppose that $\mathscr{C}$ is a dessin, and suppose that the surface $|\mathscr{C}|$ is oriented, and has no boundary. Then each dart is the intersection of precisely two triangles, one black and one white. The next remark is worth stating as a lemma for emphasis:

Lemma 1.20 – *When $\mathscr{C}$ is oriented, without boundary, there is a bijection between the darts and black triangles.*

Of course there is also a bijection between the darts and the white triangles, on which we comment below.

Now consider the permutations $\sigma = ab$, $\alpha = bc$ and $\phi = ca$. Each preserves the subset of T comprised by the black triangles, so we may see σ, α and ϕ as permutation of D. It is immediate that they satisfy $\sigma\alpha\phi = 1$, the identity permutation.

Let us draw a little picture to get a geometric understanding of these permutations. We adopt the following convention: when we draw a portion of an oriented dessin, we represent the black triangles in such a way that going from $\star$ to $\bullet$ to $\circ$ rotates us counterclockwise. (If we arrange this for one black triangle, and the portion of the dessin really is planar, that is embeds into the plane, then all black triangles will have this property).



(Recall our convention on permutations as per remark 1.15.)

On this picture, we see that our intuition for σ should be that it takes a dart to the next one in the rotation around its black vertex, going counterclockwise. Likewise α is interpreted as the rotation around the white vertex of the dart. As for ϕ, seen as a permutation of T, it takes a black triangle to the next one on the same face, going counterclockwise. This can be made into more than just an intuition: if $\partial f = (d_1, \ldots, d_{2n})$, and if t_i^f is black, then $\phi(d_i) = d_{i+2}$. Note that if the triangle t_i^f is white, then ϕ takes it to t_{i-2}^f. In particular if one changes the orientation of the dessin, the rotation ϕ changes direction, as do σ and α.

This is also reflected algebraically in the relation $b^{-1}\sigma b = \sigma^{-1}$ (which translates the fact that $a^2 = 1$): conjugating by b amounts to swapping the roles of the black and white triangles (or to identifying D with the white triangles instead of the blacks), and that turns σ into σ^{-1}. This remark is important in the proof of the following.

Theorem 1.21 – *Let D be a finite set endowed with three permutations σ, α, ϕ such that $\sigma\alpha\phi = 1$. Then there exists a dessin $\mathscr{C}$, oriented and without boundary, unique up to unique orientation-preserving isomorphism, such that D and σ, α, ϕ can be identified with the set of darts of $\mathscr{C}$ with the permutations described above.*

Proof. Let $T = D \times \{\pm 1\}$. We extend σ to a permutation $\tilde\sigma$ on T by the formula

$$\tilde\sigma(d, \varepsilon) = (\sigma^{\varepsilon}(d), \varepsilon) ,$$

and likewise α induces $\tilde\alpha$ on T by

$$\tilde\alpha(d, \varepsilon) = (\alpha^{\varepsilon}(d), \varepsilon) .$$

We also define a permutation b of T by

$$b(d, \varepsilon) = (d, -\varepsilon) .$$

Putting $a = \tilde\sigma b$ and $c = b\tilde\alpha$, it is immediate that a and c are of order 2 and have no fixed points.

By theorem 1.14, the set T together with a, b and c defines a dessin $\mathscr{C}$. Since b has no fixed points, $\mathscr{C}$ has no boundary. Calling the triangles in $D \times \{1\}$ black, and those in $D \times \{-1\}$ white, we see that $\mathscr{C}$ is naturally oriented.

The remaining statements are straightforward to prove. □

Remark 1.22. We point out that one may prove theorem 1.21 without appealing to theorem 1.14 first: one can identify B, resp W, resp F, with the cycles of σ, resp. α, resp ϕ, and proceed from there. We leave this to the reader.

In particular, we may identify the topological surface $|\mathscr{C}|$ easily: since it is compact, orientable, and without boundary, it is determined by its genus or its Euler characteristic. The latter is

$$\chi(|\mathscr{C}|) = n_\sigma + n_\alpha - n + n_\phi ,$$

where n is the cardinality of D (the number of darts), while n_σ, resp. n_α, resp. n_ϕ is the number of cycles of σ, resp. α, resp. ϕ.

Note that the group of permutations of D generated by σ, α and ϕ is called the *cartographic group* of $\mathscr{C}$, or sometimes the *monodromy group*.

1.8. **Categories.** Next we promote theorems 1.14 and 1.21 to equivalence of categories. We write $\mathfrak{Dessins}$ for the category whose objects are compact, oriented dessins without boundary, and whose morphisms are the orientation-preserving maps of cell complexes. Also, $\mathfrak{UDessins}$ will be the category whose objects are compact dessins without boundary (possibly on non-orientable surfaces), and whose morphisms are all morphisms of cell complexes.

We leave to the reader the task of proving the next theorem based on theorem 1.14 and corollary 1.16, as well as theorem 1.21.

Theorem 1.23 – *Consider the category $\mathfrak{Sets}_{a,b,c}$ whose objects are the finite sets T equipped with three distinguished permutations a, b, c, each of order two and having no fixed points, and whose arrows are the equivariant maps. Then the assigment $\mathscr{C} \to T$ extends to an equivalence of categories between $\mathfrak{UDessins}$ and $\mathfrak{Sets}_{a,b,c}$.*

Likewise, consider the category $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ whose objects are the finite sets D equipped with three distinguished permutations σ, α, ϕ satisfying $\sigma\alpha\phi = 1$, and whose arrows are the equivariant maps. Then the assigment $\mathscr{C} \to D$ extends to an equivalence of categories between $\mathfrak{Dessins}$ and $\mathfrak{Sets}_{\sigma,\alpha,\phi}$.

If one removes the requirement that b have no fixed point, in the first part, one obtains a category equivalent to that of compact dessins possibly with boundary.

1.9. **The isomorphism classes.** It is very easy for us now to describe the set of isomorphism classes of dessins. There are different approaches in the literature and we try to give several points of view.

Proposition 1.24 – (1) *A dessin $\mathscr{C}$ in $\mathfrak{Dessins}$ determines, and can be reconstructed from, an integer n, a subgroup G of S_n, and two distinguished generators σ and α for G. Two sets of data (n, G, σ, α) and $(n', G', \sigma', \alpha')$ determine isomorphic dessins if and only if $n = n'$ and there is a conjugation in S_n taking σ to σ' and α to α' (and in particular G to G').*
(2) *The set of isomorphism classes of connected dessins in $\mathfrak{Dessins}$ is in bijection with the set of conjugacy classes of subgroups of finite index in the free group on two generators $\langle\sigma, \alpha\rangle$.*
(3) *Any connected dessin in $\mathfrak{Dessins}$ determines, and can reconstructed from, a finite group G with two distinguished generators σ and α, and a subgroup H such that the intersection of all the conjugates of H in G is trivial. We obtain isomorphic dessins from (G, σ, α, H) and $(G', \sigma', \alpha', H')$ if and only if there is an isomorphism $G \to G'$ taking σ to σ', α to α', and H to a conjugate of H'.*

Proof. At this point this is very easy. (1) is left as an exercise. Here are some indications with (2): A connected object amounts to a finite set X with a transitive, right action of $\langle\sigma, \alpha\rangle$, so X must be isomorphic to $K\backslash\langle\sigma, \alpha\rangle$, where an isomorphism is obtained by choosing a base-point in X (whose stabilizer is K); different choices lead to conjugate subgroups. (2) follows easily.

We turn to (3). It is clear that a connected object X is isomorphic to $H\backslash G$ where G is the cartographic group and H is the stabilizer of some point; elements in the intersection of all conjugates of H stabilize all the points of X, and so must be trivial since G is by definition a subgroup of $S(X)$. Conversely any object of the form $H\backslash G$, with the actions of σ and α by multiplication on the right, can be seen in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$; it is connected since σ and α generate G; and its cartographic group must be G itself given the condition on H. What is more, there is a canonical map $f: \langle\sigma, \alpha\rangle \to G$ sending σ and α to the elements with the same name in G, and the inverse image $K = f^{-1}(H)$ is the subgroup corresponding to the dessin as in (2), while the intersection N of all the conjugates of K is the kernel of f. Thus we deduce the rest of (3) from (2). □

In §3 we shall come back to these questions (see §3.2 in particular). For the moment let us add that it is common, in the literature, to pay special attention to certain dessins for which some condition on the order of σ, α and ϕ is prescribed. For example, those interested in clean dessins very often require $\alpha^2 = 1$. Assuming that we are interested in the dessins for which, in addition, the order of σ divides a fixed integer k, and that of ϕ divides ℓ, then the objects are in bijection with the conjugacy classes of subgroups of finite index in

$$T_{k,\ell} = \langle \sigma, \alpha, \phi : \sigma^k = \alpha^2 = \phi^\ell = 1,\ \sigma\alpha\phi = 1 \rangle ,$$

usually called a *triangle group*. (We point out that, in doing so, we include more than the clean dessins, for α may have fixed points.)

The variant in the unoriented case is as follows.

Proposition 1.25 – *Consider the group $\langle a, b, c : a^2 = b^2 = c^2 = 1 \rangle = C_2 * C_2 * C_2$, the free product of three copies of the group of order 2. The isomorphism classes of connected objects in $\mathfrak{UDessins}$ are in bijection with the conjugacy classes of subgroups H of $C_2 * C_2 * C_2$ having finite index, and with the property that no conjugate of H contains any of a, b, c.*

Note that the last condition rephrases the fact that the actions of a, b and c on $H\backslash C_2 * C_2 * C_2$ (on the right) have no fixed points.

We proceed to describe a number of categories which are equivalent to the category $\mathfrak{Dessins}$ of dessins – the word dessin will henceforth mean compact, oriented dessin without boundary. These should be familiar to the reader, and there will be little need for long descriptions of the objects and morphisms.

As for proving the equivalences, it will be a matter of quoting celebrated results: the equivalence between covering spaces and sets with an action of the fundamental group, the equivalence between Riemann surfaces and their fields of meromorphic functions, the equivalence between algebraic curves and their fields of rational functions... as well as some elementary Galois theory, which we have taken from Völklein's book [Völ96]. There is a little work left for us, but we hope to convince the reader that the theory up to here is relatively easy – given the classics! What makes all this quite deep is the combination of strong theorems in many different branches of mathematics.

2.1. **Ramified covers.** Let S and R be compact topological surfaces. A map $p\colon S \to R$ is a *ramified cover* if there exists for each $s \in S$ a couple of charts, centered around s and $p(s)$ respectively, in which the map p becomes $z \mapsto z^e$ for some integer $e \geq 1$ called the *ramification index at* s (this index at s is well-defined, for p cannot look like $z \mapsto z^{e'}$ for $e' \neq e$ in other charts, as can be seen by examining how-many-to-1 the map is).

Examples are provided by complex surfaces: if S and R have complex structures, and if p is analytic (holomorphic), then it is a basic result from complex analysis that p must be a ramified cover in the above sense (as long as it is not constant on any connected component of S). However we postpone all complex analysis for a while.

Instead, we can obtain examples (and in fact all examples) by the following considerations. The set of $s \in S$ such that the ramification index e is > 1 is visibly discrete in S and closed, so it is finite by compactness. Its image in R under p is called the *ramification set* and written R_r. It follows that the restriction

$$p\colon S \smallsetminus f^{-1}(R_r) \longrightarrow R \smallsetminus R_r$$

is a finite covering in the traditional sense. Now, it is a classical result that one can go the other way around: namely, start with a compact topological surface R, let R_r denote a finite subset of R, and let $p\colon U \longrightarrow R \smallsetminus R_r$ denote a finite covering map; then one can construct a compact surface S together with a ramified cover $\bar{p}\colon S \to R$ such that U identifies with $\bar{p}^{-1}(R \smallsetminus R_r)$ and p identifies with the restriction of $\bar{p}$. The ramification set of $\bar{p}$ is then contained in R_r. See §5 of [Völ96] for all the details in the case $R = \mathbb{P}^1$ (the general case is no different).

Thus when the ramification set is constrained once and for all to be a subset of a given finite set R_r, ramified covers are in one-one correspondence with covering maps. To make this more precise, let us consider two ramified covers $p\colon S \to R$ and $p'\colon S' \to R$ both having a ramification set contained in R_r, and let us define a morphism between them to be a continuous map $h\colon S \to S'$ such that $p' \circ h = p$. Morphisms, of covering maps above $R \smallsetminus R_r$ are defined similarly. We may state:

THEOREM 2.1 – *The category of finite coverings of $R \smallsetminus R_r$ is equivalent to the category of ramified covers of R with ramification set included in R_r.*

Now let us quote a well-known result from algebraic topology:

THEOREM 2.2 – *Assume that R is connected, and pick a base point $* \in R \smallsetminus R_r$. The category of coverings of $R \smallsetminus R_r$ is equivalent to the category of right $\pi_1(R \smallsetminus R_r, *)$-sets. The functor giving the equivalence sends $p\colon U \to R \smallsetminus R_r$ to the fibre $p^{-1}(*)$ with the monodromy action.*

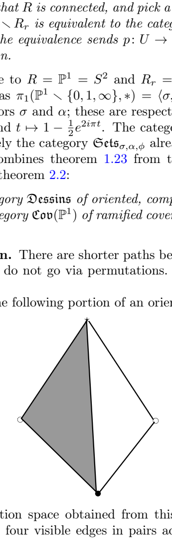
We shall now specialize to $R = \mathbb{P}^1 = S^2$ and $R_r = \{0, 1, \infty\}$. With the base point $* = \frac{1}{2}$ (say), one has $\pi_1(\mathbb{P}^1 \smallsetminus \{0,1,\infty\}, *) = \langle \sigma, \alpha \rangle$, the free group on the two distinguished generators σ and α; these are respectively the homotopy classes of the loops $t \mapsto \frac{1}{2}e^{2i\pi t}$ and $t \mapsto 1 - \frac{1}{2}e^{2i\pi t}$. The category of finite, right $\pi_1(\mathbb{P}^1 \smallsetminus \{0,1,\infty\}, *)$-sets is precisely the category $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ already mentioned.

The following result combines theorem 1.23 from the previous section, theorem 2.1 above, as well as theorem 2.2:

THEOREM 2.3 – *The category $\mathfrak{Dessins}$ of oriented, compact dessins without boundary is equivalent to the category $\mathfrak{Cov}(\mathbb{P}^1)$ of ramified covers of $\mathbb{P}^1$ having ramification set included in $\{0,1,\infty\}$.*

2.2. **Geometric intuition.** There are shorter paths between dessins and ramified covers of the sphere, that do not go via permutations. Here we comment on this approach.

First, let us examine the following portion of an oriented dessin:



Consider the identification space obtained from this by gluing the two white vertices into one, and the four visible edges in pairs accordingly. The result is a sphere; more precisely, we can canonically find a homeomorphism with S^2 sending $\bullet$ to 0 and $\circ$ to 1, while $\star$ is sent to ∞. Doing this for all pairs (t, t^a), where t is black, yields a single map $|\mathscr{C}| \to S^2$. The latter is the ramified cover corresponding to $\mathscr{C}$ in the equivalence of categories above.

We will not prove this last claim in detail, nor will we rely on it in the sequel. On the other hand, we do examine the reverse construction more closely. In fact let us state:

PROPOSITION 2.4 – *Let $\mathscr{C}$ correspond to $p\colon S \to \mathbb{P}^1$ in the above equivalence of categories. Then $|\mathscr{C}| \cong S$, under a homeomorphism taking $|\mathscr{G}|$ to the inverse image $p^{-1}([0,1])$.*

For the proof it will be convenient to have a modest lemma at our disposal. It gives conditions under which a ramified cover $p\colon S \to R$, which must be locally of the form $z \mapsto z^e$, can be shown to be of this form over some given open set. We will write

$$\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$$

as before, while

$$\dot{\mathbb{D}} = \{z \in \mathbb{C} : |z| < 1\},$$

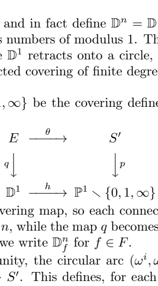
and

$$\dot{\mathbb{D}}' = \dot{\mathbb{D}} \smallsetminus \{0\}.$$

LEMMA 2.5 – *Let $p\colon S \to R$ be a ramified cover between compact surfaces. Let $x \in R_r$, and let U be an open neighbourhood of x. We assume that U is homeomorphic to a disc, and that $U \cap R_r = \{x\}$.*

Then each connected component V of $p^{-1}(U)$ contains one and only one point of the fibre $p^{-1}(x)$. Moreover, each V is itself homeomorphic to a disc and there is

[illegible]

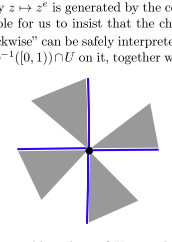


[illegible]

It remains to prove that $\mathscr{C}$ is the dessin corresponding to the ramified cover p in the equivalence of categories at hand. For this we compare the induced actions. To $\mathscr{C}$ are attached two permutations σ and α of the set D of darts. Note that D is here in bijection with the fibre $p^{-1}(\frac{1}{2})$, and taking $\frac{1}{2}$ as base point we have the monodromy action of $\pi_1(\mathbb{P}^1 \smallsetminus \{0,1,\infty\}) = \langle \sigma', \alpha' \rangle$, defining the permutations σ' and α'. We must prove that $\sigma = \sigma'$ and $\alpha = \alpha'$. Here σ' and α' are the classes of the loops defined above (where we used the notation σ and α in anticipation).

We will now use the fact (which we say more after the proof) that S can be endowed with a unique smooth structure and orientation, such that $p\colon S \to \mathbb{P}^1$ is smooth and orientation-preserving. We use this first to obtain, for each dart, a smooth parametrization $\gamma\colon [0,1] \to S$ such that $p \circ \gamma$ is the identity of $[0,1]$. Each dart belongs to two triangles, and it now makes sense to talk about the triangle *on the left* of the dart as we travel along γ. Colour it black. We will prove that this is a colouring of the type considered in §1.6.

Pick $b \in B$, and a centered chart $\mathbb{D} \to U$ onto a neighbourhood U of b, such that the map p when pulled-back to $\mathbb{D}$ is $z \mapsto z^e$. The monodromy action of $\pi_1(\dot{\mathbb{D}}')$ on the cover $\dot{\mathbb{D}}' \to \dot{\mathbb{D}}'$ given by $z \mapsto z^e$ is generated by the counterclockwise rotation of angle $\frac{2\pi}{e}$. Now it is possible for us to insist that the chart $\dot{\mathbb{D}} \to U$ be orientation-preserving, so "counterclockwise" can be safely interpreted on S as well as $\dot{\mathbb{D}}$. Let us draw a picture of U with $p^{-1}([0,1)) \cap U$ on it, together with the triangles, for $e = 4$.



The complement of the star-like subset of U given by $p^{-1}([0,1))$ falls into connected components, each contained in a face; so two darts obtained by a rotation of angle $\frac{2\pi}{e}$ are on the boundary of the same face, and must be consecutive. The symmetry a, that is the symmetry in the $\star - \bullet$ side, is now clearly seen to exchange a black triangle with a white one. What is more, calling b as usual the symmetry in the darts, the permutation $\sigma = ab$ sends a black triangle to its image under the rotation already mentioned. This is also the effect of the monodromy action, and $\sigma = \sigma'$.

Reasoning in the same fashion with white vertices, we see that c, the symmetry in the $\star - \circ$ side, also exchanges triangles of different colours. So the colouring indeed has the property that neighbouring triangles are never of the same colour. That $\alpha = \alpha'$ is observed similarly. This concludes the proof. □

EXAMPLE 2.6 (DUALITY) – The geometric intuition gained with this proposition and its proof may clarify some arguments. Let $\mathscr{C}$ be a dessin, whose sets of triangles and darts will be written T and D, so that $\mathscr{C}$ defines the object $(D, \sigma, \alpha, \phi)$ in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. Now let $p\colon S \longrightarrow \mathbb{P}^1$ correspond to $\mathscr{C}$. What is the dessin corresponding to $1/p$? And what is the object in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$?

Let us use the notation $\mathscr{C}'$, T' and D'. We can think of $\mathscr{C}$ and $\mathscr{C}'$ as being drawn on the same surface S. Zeroes of $1/p$ are poles of p and *vice-versa*, so black vertices are exchanged with face centres, while the white vertices remain in place. In fact, the most convenient property to observe is that $\mathscr{C}$ and $\mathscr{C}'$ have exactly the same triangles, as subspaces of S, and we identify $T = T'$. The $\star - \circ$ sides are promoted to darts.

The symmetries of T which we have called a, b and c become, for $\mathscr{C}'$, the symmetries $a' = a$, $b' = c$ and $c' = b$ (simply look at the definitions and exchange $\star$ and $\bullet$ throughout). It follows that $\sigma = ab$ becomes $\sigma' = a'b' = ac = \phi^{-1}$ and similarly one obtains $\alpha' = \alpha^{-1}$ and $\phi' = \sigma^{-1}$.

One must be careful, however. The object in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ defined by $1/p$, which we are after, is hidden behind one more twist. The *black* triangles in T for $\mathscr{C}$ are those mapping to the upper half plane under p; the white triangles for $\mathscr{C}$ are the black ones for $\mathscr{C}'$ as a result. Identifying darts and black triangles, we see T as the disjoint union of D and D'. While it is the case that $\mathscr{C}'$ corresponds to $(D', \phi^{-1}, \alpha^{-1}, \sigma^{-1})$ in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, this notation is confusing since we tend to think of ϕ^{-1} as a map defined on either T or D, when in fact it is the induced map on D' which is considered here (in fact we should write something like $\phi^{-1}|_{D'}$). It is clearer to use for example the map $b'\colon D \to D'$ and transport the permutations to D, which is simply a conjugation. As already observed, this "change of orientation" amounts to taking inverses for σ' and α'.

The conclusion is that *replacing p by $1/p$ takes the object $(D, \sigma, \alpha, \phi)$ to the object $(D, \phi, \alpha, \alpha^{-1}\sigma\alpha)$.*

EXAMPLE 2.7 (CHANGE OF COLOURS) – As an exercise, the reader will complete the following outline. If $\mathscr{C}$ is represented by $p\colon S \to \mathbb{P}^1$, with corresponding object $(D, \sigma, \alpha, \phi)$, then $1 - p\colon S \to \mathbb{P}^1$ corresponds to $(D, \alpha, \sigma, \alpha\phi\alpha^{-1})$. Indeed, $\mathscr{C}$ and $\mathscr{C}'$ have the same triangles, as subsets of S, and the black triangles for $\mathscr{C}$ are precisely the white ones for $\mathscr{C}'$ and *vice-versa*; the vertices of $\mathscr{C}'$ are those of $\mathscr{C}$ with the colours exchanged, while the face centres remain in place. (Informally $\mathscr{C}'$ is just that: the same as $\mathscr{C}$ with the colours exchanged.) So $c' = a$, $b' = b$ and $a' = c$, and $\sigma' = c\alpha c^{-1}$, $\alpha' = b\sigma b^{-1}$, as maps of T. As maps of D, using the bijection $b\colon D \to D'$ to transport the maps induced on D', we end up with the permutations announced.

2.3. **Complex structures.** When $p\colon S \to R$ is a ramified cover, and R is equipped with a complex structure, there is a unique complex structure on S such that p is complex analytic ([DD79], 6.1.10). Any morphism between S and S', over R, is then itself complex analytic. Conversely if S and R both have complex structures, an analytic map $S \to R$ is a ramified cover as soon as it is not constant on any connected component of S.

We may state yet another equivalence of categories. Recall that an analytic map $S \to \mathbb{P}^1$ is called a meromorphic function on S.

THEOREM 2.8 – *The category $\mathfrak{Dessins}$ is equivalent to the category $\mathfrak{Belyi}$ of compact Riemann surfaces with a meromorphic function whose ramification set is contained in $\{0,1,\infty\}$.*

(The arrows considered are the maps above $\mathbb{P}^1$.) A pair (S, p) with $p\colon S \to \mathbb{P}^1$ meromorphic, not ramified outside of $\{0,1,\infty\}$, is often called a *Belyi pair*, while p is called a *Belyi map*.

EXAMPLE 2.9 – Let us illustrate the results up to now with dessins on the sphere, so let $\mathscr{C}$ be such that $|\mathscr{C}|$ is homeomorphic to S^2. By the above, $\mathscr{C}$ corresponds to a Riemann surface S equipped with a Belyi map $p\colon S \to \mathbb{P}^1$.

By proposition 2.4, S is itself topologically a sphere. The uniformization theorem states that there is a complex isomorphism $\theta\colon \mathbb{P}^1 \longrightarrow S$, so we may as well replace S with $\mathbb{P}^1$ equipped with $F = p \circ \theta$. Then $(\mathbb{P}^1, F)$ is a Belyi pair isomorphic to (S, p).

Now $F\colon \mathbb{P}^1 \to \mathbb{P}^1$, which is complex analytic and not constant, must be given by a rational fraction, as is classical. *The bigraph $\mathscr{G}$ can be realized as the inverse image $F^{-1}([0,1])$ where $F\colon \mathbb{P}^1 \longrightarrow \mathbb{P}^1$ is a rational fraction.*

Let us take this opportunity to explain the terminology *dessins d'enfants* (children's drawings), and stress again some remarkable features. By drawing a simple picture, we may as in example 1.3 give enough information to describe a cell complex $\mathscr{C}$. Very often it is evident that $|\mathscr{C}|$ is a sphere, as we have seen in this example. What the theory predicts is that we can find a rational fraction F such that the drawing may be recovered as $F^{-1}([0,1])$. This works with pretty much any planar, connected drawing that you can think of, and gives these drawings a rigidified shape.

To be more precise, the fraction F is unique up to an isomorphism of $\mathbb{P}^1$, that is, up to precomposing with a Moebius transformation. This allows for rotation and stretching, but still some features will remain unchanged. For example the darts around a given vertex will all have the same angle $\frac{2\pi}{e}$ between them, since F looks like $z \mapsto z^e$ in conformal charts.

2.4. **Fields of meromorphic functions.** When S is a compact, connected Riemann surface, one can consider all the meromorphic functions on S, comprising a field $\mathscr{M}(S)$. When S is not assumed connected, the meromorphic functions form an *étale algebra*, still written $\mathscr{M}(S)$: in this paper an étale algebra is simply a direct sum of fields, here corresponding to the connected components of S. In what follows we shall almost always have to deal with an *étale algebra over K* where K is some field, by which we mean an étale algebra which is also a K-algebra, and which is finite-dimensional over K. (In the literature étale algebras have to satisfy a separability condition, but we work in characteristic 0 throughout the paper.)

If now $p\colon S \to R$ is a ramified cover between compact surfaces, we may speak of its degree, as the degree of the corresponding covering $p^{-1}(R \smallsetminus R_r) \to R \smallsetminus R_r$. The following is given in §6.2.4 in [DD79].

THEOREM 2.10 – *Fix a compact, connected Riemann surface R. The category of compact Riemann surfaces S with a ramified cover $S \to R$ is anti-equivalent to the category of étale algebras over $\mathscr{M}(R)$. The equivalence is given by $S \mapsto \mathscr{M}(S)$, and the degree of $S \to R$ is equal to the dimension of $\mathscr{M}(S)$ as a vector space over $\mathscr{M}(R)$.*

(Here and elsewhere, "anti-equivalent" means "equivalent to the opposite category".)

Taking $R = \mathbb{P}^1$, we get a glimpse of yet another category that could be equivalent to $\mathfrak{Dessins}$. However to pursue this, we need to translate the condition about the ramification into a statement about étale algebras (lest we should end up with a half-baked category, consisting of algebras such that the corresponding surface has a certain topological property; that would not be satisfactory). For this we reword §2.2.1 of [Völ96].

Recall that $\mathscr{M}(\mathbb{P}^1) = \mathbb{C}(x)$, where x is the identity of $\mathbb{P}^1$. So let us start with any field k at all, and consider a finite, Galois extension L of $k(x)$. We shall say that $L/k(x)$ is *not ramified* at 0 when it embeds into the extension $k((x))/k(x)$, where as usual $k((x))$ is the field of formal power series in x. In this paper we will not enter into the subtleties of the field $k((x))$, nor will we discuss the reasons why this definition makes sense. We chiefly want to mention that there is a simple algebraic statement corresponding to the topological notion of ramification, quoting the results we need.

Now take any $s \in k$. From L we construct $L_s = L \otimes_{k(x)} k(x)$, where we see $k(x)$ as an algebra over $k(x)$ *via* the map $k(x) \to k(x)$ which sends x to $x + s$; concretely if we pick a primitive element y for $L/k(x)$, so that $L \cong k(x)[y]/(P)$, then L_s is $k(x)[y]/(P_s)$ where P_s is the result of applying $x \mapsto x + s$ to the coefficients of P. When $L_s/k(x)$ is not ramified at 0, we say that $L/k(x)$ *is not ramified at s*.

Finally, using the map $k(x) \to k(x)$ which sends x to x^{-1}, we get an extension $L_\infty/k(x)$, proceeding as above. When the latter is not ramified at 0, we say that $L/k(x)$ is *not ramified at* ∞.

When the conditions above are not satisfied, for $s \in k \cup \{\infty\}$, we will of course say that L does ramify at s (or is ramified at s). That the topological and algebraic definitions of ramification actually agree is the essence of the next lemma.

LEMMA 2.11 – *Let $p\colon S \to \mathbb{P}^1$ be a ramified cover, with S connected, and assume that the corresponding extension $\mathscr{M}(S)/\mathbb{C}(x)$ is Galois. Then for any $s \in \mathbb{P}^1$, the ramification set $\mathbb{P}^1_r$ contains s if and only if $\mathscr{M}(S)/\mathbb{C}(x)$ ramifies at s in the algebraic sense.*

In particular, the ramification set in contained in $\{0,1,\infty\}$ if and only if the extension $\mathscr{M}(S)/\mathbb{C}(x)$ does not ramify at s whenever $s \notin \{0,1,\infty\}$.

This is the addendum to theorem 5.9 in [Völ96]. Now we need to get rid of the extra hypothesis that $\mathscr{M}(S)/\mathbb{C}(x)$ be Galois (a case not considered in [Völ96], strictly speaking). Algebraically, we say that an extension $L/k(x)$ does not ramify at s when its Galois closure $\tilde{L}/k(x)$ does not. To see that, with this definition, the last lemma generalizes to all ramified covers, we need to prove the following.

LEMMA 2.12 – *Let $p\colon S \to \mathbb{P}^1$ be a ramified cover, where S is connected. Let $\tilde{p}\colon \tilde{S} \to \mathbb{P}^1$ be the ramified cover such that $\mathscr{M}(\tilde{S})/\mathbb{C}(x)$ is the Galois closure of $\mathscr{M}(S)/\mathbb{C}(x)$. Then the ramification sets for S and $\tilde{S}$ are equal.*

Proof. We have $\mathbb{C}(x) \subset \mathscr{M}(S) \subset \mathscr{M}(\tilde{S})$, so we also have a factorization of $\tilde{p}$ as $\tilde{S} \to S \to \mathbb{P}^1$. From this it is clear that, if $\tilde{p}$ is not ramified at $s \in \mathbb{P}^1$, then neither is p.

The crux of the proof of the reverse inclusion is the fact that covering maps have Galois closures, usually called regular covers. The following argument anticipates the material of the next section, though it should be understandable now.

Let $\mathbb{P}^1_r$ be the ramification set for p, and let $U = p^{-1}(\mathbb{P}^1 \smallsetminus \mathbb{P}^1_r)$, so that $U \to \mathbb{P}^1 \smallsetminus \mathbb{P}^1_r$ is a finite covering map. Now let $\tilde{U} \to \mathbb{P}^1 \smallsetminus \mathbb{P}^1_r$ be the corresponding regular covering map. Here "regular" can be taken to mean that this cover has as many automorphisms as its degree; and $\tilde{U}$ is minimal with respect to this property, among the covers factoring through U. The existence of $\tilde{U}$ is standard in covering space theory, and should become very clear in the next section. Note that, if U corresponds to the subgroup H of $\pi_1(\mathbb{P}^1 \smallsetminus \mathbb{P}^1_r)$, then $\tilde{U}$ corresponds to the intersection of all the conjugates of H.

As above, we can construct a Riemann surface S' from $\tilde{U}$, and the latter does not ramify outside of $\mathbb{P}^1_r$. To prove the lemma, it is sufficient to show that S' can be identified with $\tilde{S}$.

However from basic Galois theory we see that $\mathscr{M}(S')/\mathbb{C}(x)$ must be Galois since it possesses as many automorphisms as its degree, and by minimality it must be the Galois closure of $\mathscr{M}(S)/\mathbb{C}(x)$. So S' and $\tilde{S}$ are isomorphic covers of $\mathbb{P}^1$. □

Finally, an étale algebra over $k(x)$ will be said not to ramify at s when it is a direct sum of field extensions, none of which ramifies at s. This clearly corresponds to the topological situation when $k = \mathbb{C}$, and we have established the following.

[illegible]

3. Regularity

From now on, it will be convenient to use the word *dessin* to refer to an object in any of the equivalent categories at our disposal (especially when we want to think of it simultaneously as a cell complex and a field, for example).

In this section we study regular dessins. These could have been called "Galois" instead of "regular", since the interpretation in the realm of field extensions is precisely the Galois condition, but we want to avoid the confusion with the Galois group $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ which will become a major player in the sequel.

3.1. **Definition of regularity.** An object in $\mathfrak{Dessins}$ has a *degree* given by the number of darts. In the other categories equivalent to $\mathfrak{Dessins}$, this translates in various ways. In $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, it is the cardinality of the set having the three permutations on it. In the categories of étale algebras over $\mathbb{C}(x)$ or $\overline{\mathbb{Q}}(x)$, it is the dimension of the algebra as a vector space over $\mathbb{C}(x)$ or $\overline{\mathbb{Q}}(x)$ respectively. In the category of finite coverings of $\mathbb{P}^1 \smallsetminus \{0,1,\infty\}$, it is the cardinality of any fibre.

There is also a notion of *connectedness* in these categories. A dessin $\mathscr{C}$ is connected when $|\mathscr{C}|$ is connected, which happens precisely when the corresponding étale algebras are actually fields, or when the cartographic group acts transitively (cf lemma 1.17).

In this section we shall focus on the automorphism groups of connected dessins. We are free to conduct the arguments in any category, and most of the time we prefer $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. However, note the following at once.

Lemma 3.1 – *The automorphism group of a connected dessin is a finite group, of order no greater than the degree.*

Proof. This is obvious in $\mathfrak{Etale}(\overline{\mathbb{Q}}(x))$: in fact for any finite-dimensional extension of fields L/K, basic Galois theory tells us that the automorphism group of the extension has order no greater than $[L:K]$.

A proof in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ will be immediate from lemma 3.3 below. □

A dessin will be called *regular* when it is connected and the order of its automorphism group equals its degree.

In terms of field extensions for example, then $L/\mathbb{C}(x)$ is regular if and only if it is Galois (in the elementary sense, ie normal and separable). In terms of a covering $U \to \mathbb{P}^1 \smallsetminus \{0,1,\infty\}$, with U is connected, then it is regular if and only if it is isomorphic to the cover $U \to U/G$ where G is the automorphism group (this agrees with the use of the term "regular" in covering space theory, of course).

Remark 3.2. The reader needs to pay special attention to the following convention. When X is a dessin and $h, k \in Aut(X)$, we write hk for the composition of k followed by h; that is $hk(x) = h(k(x))$, at least when we are willing to make sense of $x \in X$ (for example in $\mathfrak{Dessins}$ this will mean that x is in fact a triangle). In other words, we are letting $Aut(X)$ act on X *on the left.* While this will be very familiar to topologists, for whom it is common to see the "group of deck transformations" of a covering map act on the left and the "monodromy group" act on the right, other readers may be puzzled to see that we have treated the category of sets differently when we took the convention described in remark 1.15.

To justify this, let us spoil the surprise of the next paragraphs, and announce the main result at once: in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, a regular dessin is precisely a group G with two distinguished generators σ and α; the monodromy group is G itself, acting on the right by translations, while the automorphism group is again G itself, acting on the left by translations.

If we had taken different conventions, we would have ended up with one of these actions involving inverses, in a way which is definitely unnatural.

3.2. **Sets with permutations.** We explore the definition of regularity in the context of $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, where it is very easy to express.

Let X be a set of cardinality n, with three permutations σ, α, ϕ satisfying $\sigma\alpha\phi = 1$. Let G denote the cartographic group; recall that by definition, it is generated by σ and α as a subgroup of $S(X) \cong S_n$, acting on X on the right. We assume that G acts transitively (so the corresponding dessin is connected).

We choose a base-point $* \in X$. The map $g \mapsto *^g$ identifies $H\backslash G$ with X, where H is the stabilizer of $*$. This is an isomorphism in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, with G acting on $H\backslash G$ by right translations. As we shall insist below that the choice of base-point is somewhat significant, we shall keep the notation X and not always work directly with $H\backslash G$.

Since the morphisms in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ are special maps of sets, we can relate $Aut(X)$ and $S(X)$, where the automorphism group is taken in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, and $S(X)$ as always is the group of all permutations of X. More precisely, any $h \in Aut(X)$ can be seen as an element of $S(X)$, still written h, and there is a homomorphism $Aut(X) \to S(X)$ given by $h \mapsto h^{-1}$; our left-right conventions force us to take inverses to get a homomorphism. (In other words, $Aut(X)$ is naturally a subgroup of $S(X)^{op}$, the group $S(X)$ with the opposite composition law.) As announced, the conventions will eventually lead to a result without inverses.

Lemma 3.3 – *Let X, G, H be as above. We have the following two descriptions of $Aut(X)$.*

1. *Let $N(H)$ be the normalizer of H in G. Then for each $g \in N(H)$, the map $H\backslash G \to H\backslash G$ given by $[x] \mapsto [gx]$ is in $Aut(H\backslash G)$. This construction induces an isomorphism $Aut(X) \cong N(H)/H$.*
2. *The map $Aut(X) \to S(X)$ is an isomorphism onto the centralizer of G in $S(X)$.*

Proof. (1) The notation $[x]$ is for the class of x in $H\backslash G$, of course. To see that $[gx]$ is well-defined, let $h \in H$, then $ghx = ghg^{-1}gx$ so $[ghx] = [gx]$. The map clearly commutes with the right action of G, and so is an automorphism, with inverse given by $[x] \mapsto [g^{-1}x]$.

Conversely, any automorphism h is determined by $h([1])$, which we call $[g]$, and we must have $h([x]) = h([1]^x) = [g]^x = [gx]$ for any x; the fact that h is well-defined implies that $g \in N(H)$. So there is a surjective map $N(H) \to Aut(H\backslash G)$ whose kernel is clearly H.

(2) An automorphism of X, by its very definition, is a self-bijection of X commuting with the action of G; so this second point is obvious. □

We also note the following.

Lemma 3.4 – *$Aut(X)$ acts freely on X.*

Proof. If $h(x) = x$ for some $x \in X$, then $h(x^g) = h(x)^g = x^g$ so x^g is also fixed by h, for any $g \in G$. By assumption G acts transitively, hence the lemma. □

Proposition 3.5 – *The following are equivalent.*

1. *$Aut(X)$ acts transitively on X.*
2. *G acts freely on X.*
3. *H is normal in G.*
4. *H is trivial.*
5. *G and $Aut(X)$ are isomorphic.*
6. *G and $Aut(X)$ are both of order n.*
7. *X is regular.*

Proof. That (1) implies (2) is almost the argument we used for the last lemma, only with the roles of $Aut(X)$ and G interchanged. Condition (2) implies (4) by definition and hence (3); when we have (3) we have $N(H)/H = G/H$, and the description of the action of $N(H)/H$ on $H\backslash G$ makes it clear that (1) holds.

Condition (4) implies $N(H)/H \cong G$, so we have (5); we also have (6) since X (whose cardinality is n) can be identified with G acting on itself on the right. Conversely if we have (6), given that the cardinality of X is $n = |G|/|H|$ we deduce (4).

Finally (7), by definition, means that $Aut(X)$ has order n, so it is implied by (6). Conversely, since this group acts freely on X, having cardinality n, it is clear that (7) implies that the action is also transitive, which is (1). □

Corollary 3.6 (of the proof) – *Let X be a regular object in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ with cartographic group G. Then X can be identified with G itself with its action on itself on the right by translations. The automorphism group $Aut(X)$ can also be identified with G, acting on $X = G$ on the left by translations.*

Conversely any finite group G with two distinguished generators σ and α defines a regular object in this way.

Proof. There remains the (very easy) converse to prove. If we start with G, a finite group generated by σ and α, we can let it act on itself on the right by translations, thus defining an object in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. The cartographic group is easily seen to be isomorphic to G (in fact this is the traditional Cayley embedding of G into the symmetric group $S(G)$). The action of the cartographic group is, as a result, free and transitive, so the object is regular. □

However, some care must be taken. The identifications above are not canonical, but depend on the choice of base-point. Also, the actions of $g \in G$ on X, given by right and left multiplications, are very different-looking maps of the set X. We want to make these points crystal-clear. The letter d below is used for "dart".

Proposition 3.7 – *Suppose that X is regular. Then for each $d \in X$ there is an isomorphism*

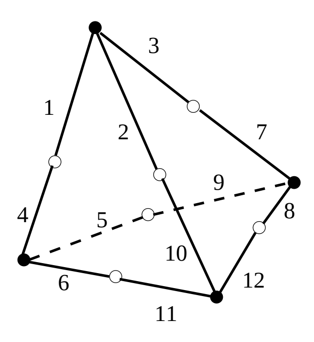
$$\iota_d \colon G \longrightarrow Aut(X)\,.$$

The automorphism $\iota_d(g)$ is the unique one taking d to d^g.

Changing d to d' amounts to conjugating, in $Aut(X)$, by the unique automorphism taking d to d'.

Proof. This is merely a reformulation of the discussion above, and we only need to check some details. We take $* = d$ as base-point. The map ι_d is clearly well-defined, and we check that it is a homomorphism: $\iota_d(gh)(d) = d^{gh} = (d^g)^h = \iota_d(g)(d)^h = \iota_d(g)(d^h) = \iota_d(g)\iota_d(h)(d)$, so the automorphisms $\iota_d(gh)$ and $\iota_d(g)\iota_d(h)$ agree at d, hence everywhere by transitivity of the action of G. □

Example 3.8 – Consider the dessin on the sphere given by the tetrahedron, as follows:



Here we have numbered the darts, for convenience (the faces, on the other hand, are implicit). There are many ways to see that this is a regular dessin. For example, one may find enough rotations to take any one dart to any other one, and apply criterion (1) of proposition 3.5. Or, we could write the permutations

$$\sigma = (123)(456)(789)(10,11,12)\,, \qquad \alpha = (14)(2,10)(37)(59)(6,11)(8,12)\,,$$

and compute the order of the group generated by σ and α, which is 12 (a computer does that for you immediately). Then appeal to criterion (6) of the same proposition. Finally, one could also determine the automorphism group of this dessin, and find that it has order 12. This is the very definition of regularity.

Take $d = 1$ as base point, and write ι for ι_1. What is $\iota(\sigma)$? This is the automorphism taking 1 to 2, which is the rotation around the black vertex adjacent to 1 and 2. The permutation of the darts induced by $\iota(\sigma)$ is

$$(123)(4,10,7)(6,12,9)(11,8,5)\,.$$

We see that σ and $\iota(\sigma)$ are not to be confused. Likewise, $\iota(\alpha)$ is the rotation taking 1 to 4, and the induced permutation is

$$(14)(8,12)(2,5)(3,6)(10,9)(11,7)\,.$$

3.3. **The distinguished triples.** From proposition 3.7, we see that each choice of dart in a regular dessin $\mathscr{C}$ defines three elements of $Aut(\mathscr{C})$, namely $\tilde\sigma = \iota_d(\sigma)$, $\tilde\alpha = \iota_d(\alpha)$, and $\tilde\phi = \iota_d(\phi)$. These are generators of $Aut(\mathscr{C})$, and they satisfy $\tilde\sigma\tilde\alpha\tilde\phi = 1$. Changing d to another dart conjugates all three generators *simultaneously.* Any such triple, obtained for a choice of d, will be called a *distinguished triple* for $\mathscr{C}$.

Lemma 3.9 – *If d and d' are darts with a common black vertex, then $\iota_d(\sigma) = \iota_{d'}(\sigma)$. Similarly if they have a common white vertex then $\iota_d(\alpha) = \iota_{d'}(\alpha)$. Finally if the black triangles corresponding to d and d' respectively lie in the same face, then $\iota_d(\phi) = \iota_{d'}(\phi)$.*

Proof. We treat the first case, for which $d' = d^{\sigma^k}$ for some k. Write $\tilde\sigma = \iota_d(\sigma)$. Since $d^{\sigma^k} = \tilde\sigma^k(d)$, we see that $\iota_{d'}(\sigma) = \tilde\sigma^k\tilde\sigma\tilde\sigma^{-k} = \tilde\sigma$. □

Thus the notation $\tilde\sigma$ makes senses unambiguously when it is understood that the possible base-darts are incident to a given black vertex. Similarly for the other types of points. We can now fully understand the fixed points of automorphisms:

Proposition 3.10 – *Let $h \in Aut(\mathscr{C})$, where $\mathscr{C}$ is regular. Suppose that the induced homeomorphism $|\mathscr{C}| \to |\mathscr{C}|$ has a fixed point. Suppose also that h is not the identity. Then the fixed point is a vertex or the centre of a face; moreover there exists an integer k such that, for any choice of dart d incident with the fixed point, we can write $h = \tilde\sigma^k$, $\tilde\alpha^k$ or $\tilde\phi^k$, according to the type of fixed point, $\bullet$, $\circ$ or $\star$.*

In particular, the subgroup of $Aut(\mathscr{C})$ comprised of the automorphisms fixing a given point of type $\bullet$ is cyclic, generated by $\tilde\sigma = \iota_d(\sigma)$ where we have chosen any dart incident with the fixed point. Likewise for the other types of fixed point.

(In this statement we have abused the language slightly, by saying that a dart is "incident" to the centre of a face if the corresponding black triangle belongs to that face.)

Proof. Let t be a triangle containing the fixed point. Note that $h(t) \neq t$: otherwise by regularity we would have h = identity. We have $t \cap h(t) \neq \emptyset$ though, and as the triangle $h(t)$ is of the same colour as t, unlike its neighbours, we conclude that $t \cap h(t)$ is a single vertex of t, and the latter is our fixed point.

Say it is a black vertex. Let d be the dart on t. Then $h(d)$ is a dart with the same black vertex as d, so $h(d) = d^{\sigma^k}$ for some integer k. In other words $h = \iota_d(\sigma^k)$. □

Thus we have a canonical generator for each of these subgroups. Here we point out, and this will matter in the sequel, that the generator $\tilde\sigma$ agrees with what Völklein calls the "distinguished generator" in Proposition 4.23 of [Völ96]. This follows from unwinding all the definitions.

The following result is used very often in the literature on regular "maps".

Proposition 3.11 – *Let $\mathscr{C}$ be a dessin, with cartographic group G, and the distinguished elements $\sigma, \alpha, \phi \in G$. Similarly, let $\mathscr{C}', G', \sigma', \alpha', \phi'$ be of the same kind. Assume that $\mathscr{C}$ and $\mathscr{C}'$ are both regular. Then the following conditions are equivalent:*

1. *$\mathscr{C}$ and $\mathscr{C}'$ are isomorphic,*
2. *there is an isomorphism $G \to G'$ taking σ to σ', α to α' and ϕ to ϕ',*
3. *there is an isomorphism $Aut(\mathscr{C}) \to Aut(\mathscr{C}')$ taking a distinguished triple to a distinguished triple.*

Proof. That (1) implies (2) is obvious, and holds without any regularity assumption. Since there are isomorphisms $G \cong Aut(\mathscr{C})$ and $G' \cong Aut(\mathscr{C}')$ taking the distinguished permutations in the cartographic group to a distinguished triple (though none of this is canonical), we see that (2) implies (3).

Finally, if we work in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, we can identify $\mathscr{C}$ with the group $Aut(\mathscr{C})$ endowed with the three elements $\tilde\sigma, \tilde\alpha, \tilde\phi$ acting by multiplication on the right, where we have picked some distinguished triple $\tilde\sigma, \tilde\alpha, \tilde\phi$. Thus (3) clearly implies (1). □

The equivalence of (1) and (3), together with corollary 3.6, reduces the classification of regular dessins to that of finite groups with two distinguished generators (or three distinguished generators whose product is 1). We state this separately as an echo to proposition 1.24. Recall that dessins are implicitly compact, oriented and without boundary here.

Proposition 3.12 –
1. *A regular dessin determines, and can be reconstructed from, a finite group G with two distinguished generators σ and α. We obtain isomorphic dessins from (G, σ, α) and (G', σ', α') if and only if there is an isomorphism $G \to G'$ taking σ to σ' and α to α'.*
2. *The set of isomorphism classes of regular dessins is in bijection with the normal subgroups of the free group on two generators. More precisely, if a connected dessin corresponds to the conjugacy class of the subgroup K as in proposition 1.24, then it is regular if and only if K is normal.*

Proof. We have already established (1). As for the first statement in (2), we only need to remark that the groups mentioned in (1) are precisely the groups of the form $G = \langle\sigma,\alpha\rangle/N$ for some normal subgroup N in the free group $F_2 = \langle\sigma,\alpha\rangle$, and that an isomorphism of the type specified in (1) between $G = F_2/N$ and $G' = F_2/N'$ exists if and only if $N = N'$.

We turn to the last statement. If a connected dessin corresponds to K, then it is isomorphic to $X = K\backslash\langle\sigma,\alpha\rangle$ in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. The action of $\langle\sigma,\alpha\rangle$ on X yields a homomorphism $f\colon \langle\sigma,\alpha\rangle \to S(X)$ whose image is the cartographic group G, and whose kernel is the intersection N of all the conjugates of K, so $G \cong \langle\sigma,\alpha\rangle/N$. Let H be the stabilizer in G of a point in X. Then $f^{-1}(H)$ is the stabilizer of that same point in $\langle\sigma,\alpha\rangle$, so it is a conjugate of K. Now, X is regular if and only if H is trivial, which happens precisely when $f^{-1}(H) = N$, which in turn occurs precisely when K is normal. □

3.4. **Regular closure & Galois correspondence.** In the discussion that follows, we restrict our attention to *connected* dessins.

When $\mathscr{C}$ and $\mathscr{C}'$ are two dessins, we call $\mathscr{C}'$ an *intermediate dessin* of $\mathscr{C}$ when there exists a morphism $\mathscr{C} \to \mathscr{C}'$. To appreciate the term "intermediate", it is best to move to categories other than $\mathfrak{Dessins}$. In $\mathfrak{Cov}(\mathbb{P}^1)$, if $\mathscr{C}$ corresponds to $p\colon S \to \mathbb{P}^1$ and $\mathscr{C}'$ corresponds to $p'\colon S' \to \mathbb{P}^1$, then $\mathscr{C}'$ is an intermediate dessin of $\mathscr{C}$ when there is a factorization of p as

$$p\colon S \xrightarrow{f} S' \xrightarrow{p'} \mathbb{P}^1\,,$$

for some map f; so $|\mathscr{C}'| = S'$ is intermediate between $|\mathscr{C}| = S$ and $\mathbb{P}^1$, if you will. In $\mathfrak{Etale}(\overline{\mathbb{Q}}(x))$, the towers $\overline{\mathbb{Q}}(x) \subset L' \subset L$ provide examples where $L'/\overline{\mathbb{Q}}(x)$ is an intermediate dessin of $L/\overline{\mathbb{Q}}(x)$, and all examples are isomorphic to one of this kind.

Of course the word "intermediate" is borrowed from field/Galois theory, where the ideas for the next paragraphs come from. Let us point out one more characterization.

Lemma 3.13 – *Let $\mathscr{C}$ and $\mathscr{C}'$ correspond to the conjugacy classes of the subgroups H and H' of $\langle\sigma,\alpha\rangle$ respectively, as in proposition 1.24. Then $\mathscr{C}'$ is an intermediate dessin of $\mathscr{C}$ if and only if some conjugate of H' contains H.*

So H' is intermediate between H and the free group $\langle\sigma,\alpha\rangle$.

Proof. The object in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ corresponding to H (and also to $\mathscr{C}$) is $X = H\backslash\langle\sigma,\alpha\rangle$, and likewise for H' we can take $X' = H'\backslash\langle\sigma,\alpha\rangle$; there is a map $X \to X'$ if and only if the stabilizer of some point in X is contained in the stabilizer of some point in X', hence the lemma. □

Lemma 3.14 – *Let $\mathscr{C}$ be a connected dessin. There exists a regular dessin $\tilde{\mathscr{C}}$ such that $\mathscr{C}$ is an intermediate dessin of $\tilde{\mathscr{C}}$. Moreover, we can arrange for $\tilde{\mathscr{C}}$ to be minimal in the following sense: if $\mathscr{C}$ is an intermediate dessin of any regular dessin $\mathscr{C}'$, then $\tilde{\mathscr{C}}$ is itself an intermediate dessin of $\mathscr{C}'$. Such a minimal $\tilde{\mathscr{C}}$ is unique up to isomorphism.*

Finally, the cartographic group of $\mathscr{C}$ is isomorphic to $Aut(\tilde{\mathscr{C}})$.

We call $\tilde{\mathscr{C}}$ the *regular closure* of $\mathscr{C}$.

Proof. Leaving the last statement aside, in $\mathfrak{Etale}(\overline{\mathbb{Q}}(x))$, this is a basic result from Galois theory. Alternatively, we can rely on proposition 1.24 and the previous lemma: if $\mathscr{C}$ corresponds to the conjugacy class of H, then clearly the object corresponding to N, the intersection of all conjugates of H, suits our purpose. As for the last statement, that the cartographic group of $H\backslash\langle\sigma,\alpha\rangle$ is isomorphic to $\langle\sigma,\alpha\rangle/N$ was already observed during the proof of proposition 3.12 (and is obvious anyway). □

The fundamental theorem of Galois theory applied in $\mathfrak{Etale}(\overline{\mathbb{Q}}(x))$, or some elementary considerations with the subgroups of $\langle\sigma,\alpha\rangle$, imply:

Proposition 3.15 – *Let $\mathscr{C}$ be a regular dessin. There is a bijection between the set of isomorphism classes of intermediate dessins of $\mathscr{C}$ on the one hand, and the conjugacy classes of subgroups of $Aut(\mathscr{C})$ on the other hand. Normal subgroups corresponds to regular, intermediate dessins.*

The concepts of this section are, as usual, very easily illustrated within $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. A connected object is of the form $H\backslash G$, as we have seen, where G has two distinguished generators σ and α. The regular closure is the object G, with its right action on itself, seen in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$. Of course there is the natural map $G \to H\backslash G$. Conversely any X with a surjective, equivariant map $G \to X$ (that is, any connected, intermediate object of G) must be of the form $H\backslash G$, clearly. From this we see that whenever $\mathscr{C}$ is regular, its intermediate dessins might called its *quotient* dessins instead.

a Moebius transformation, so $\phi(0)$ must be a root of F and $\phi(1)$ must be a root of $F-1$, while $\phi(\infty)$ must be a pole of F. Since ϕ is determined by these three values, there are only finitely many possibilities. □

We shall eventually prove that any normalized fraction has its coefficients in $\overline{\mathbb{Q}}$.

Our strategy for finding a fraction $F\colon \mathbb{P}^1 \to \mathbb{P}^1$ which is a Belyi map is to pay attention to the associated fraction

$$A = \frac{F'}{F(F-1)}\,.$$

PROPOSITION 4.3 – *Let F be a Belyi fraction such that $F(\infty)=\infty$, and let A be as above. Then the following holds.*

(1) *The partial fraction decomposition of A is of the form*
$$A = \sum_i \frac{m_i}{z-w_i} - \sum_i \frac{n_i}{z-b_i}\,,$$
where the n_i's and the m_i's are positive integers, the b_i's are the roots of F, and the w_i's are the roots of $F-1$. In fact n_i is the degree of the black vertex b_i, and m_i is the degree of the white vertex w_i.

(2) *One can recover F from A as:*
$$\frac{1}{F} = 1 - \frac{\prod_i (z-w_i)^{m_i}}{\prod_i (z-b_i)^{n_i}}\,.$$

(3) *The fraction A can be written in reduced form*
$$A = \lambda \frac{\prod_i (z-f_i)^{r_i - 1}}{\prod_i (z-b_i)\prod_i (z-w_i)}\,,$$
where the f_i's are the poles of F (other than ∞), and r_i is the multiplicity of f_i as a pole of F. In fact r_i is the number of black triangles inside the face corresponding to f_i.

Conversely, let A be any rational fraction of the form given in (3), with the numbers f_i, b_i, w_i distinct. Assume that A has a partial fraction decomposition of the form given in (1); define F by (2); and finally assume that the f_i's are poles of F. Then F is a Belyi map, $A = F'/(F(F-1))$, and we are in the previous situation.

We submit a proof below. For the moment, let us see how we can use this proposition to establish the results announced above. So assume $\mathscr{C}$ is a given dessin on the sphere, and we are looking for a corresponding normalized Belyi map $F\colon \mathbb{P}^1 \to \mathbb{P}^1$. We look for the fraction A instead, and our "unknowns" are the f_i's, the b_i's, the w_i's, and λ, cf (3). Of course we the numbers r_i from counting the black triangles on $\mathscr{C}$, just as we now the number of black vertices, white vertices, and faces, giving the number of b_i's, w_i's, and f_i's (keeping in mind the pole at ∞ already accounted for).

Now comparing (3) and (1) we must have

$$(*) \qquad \lambda \frac{\prod_i (z-f_i)^{r_i-1}}{\prod_i (z-b_i)\prod_i (z-w_i)} = \sum_i \frac{m_i}{z-w_i} - \sum_i \frac{n_i}{z-b_i}$$

where the integers n_i and m_i are all known, since they are the degrees of the black and white vertices respectively, and again these can be read from $\mathscr{C}$.

Further, the f_i's must be poles of F, which is related to A by (2). Thus we must have

$$(**) \qquad \prod_i (f_j - w_i)^{m_i} = \prod_i (f_j - b_i)^{n_i}\,,$$

for all j. We also want F to be normalized so we pick indices i_0 and j_0 and throw in the equations

$$(***) \qquad b_{i_0} = 0\,, \qquad w_{j_0} = 1\,.$$

Finally we want our unknowns to be distinct. The usual trick to express this as an equality rather than an inequality is to take an extra unknown η and to require

$$(****) \qquad \eta(b_1 - b_2)(f_1 - f_2)\cdots = 1\,,$$

where in the dots we have hidden all the required differences.

LEMMA 4.4 – *The system of polynomials equations given by (*), (**), (***) and (****) has finitely many solutions in $\mathbb{C}$. These solutions are all in $\overline{\mathbb{Q}}$.*

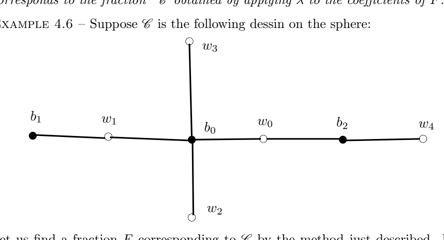
Proof. By the proposition, each solution defines a normalized Belyi map, and thus a dessin on the sphere. Define an equivalence relation on the set of solutions, by declaring two solutions to be equivalent when the corresponding dessins are isomorphic. By lemma 4.2, there are finitely many solutions in an equivalence class. However there must be finitely many classes as well, since for each n there can be only a finite number of dessins on n darts, clearly, and for all the solutions we have $n = \sum_i n_i$ darts.

It is a classical fact from either algebraic geometry, or the theory of Gröbner bases, that a system of polynomial equations with coefficients in a field K, having finitely many solutions in an algebraically closed field containing K, has in fact all its solutions in the algebraic closure of K. Here the equations have coefficients in $\mathbb{Q}$. □

We may state, as a summary of the discussion:

PROPOSITION 4.5 – *A dessin $\mathscr{C}$ on the sphere defines, and is defined by, a rational fraction F with coefficients in $\overline{\mathbb{Q}}$ which is also a Belyi map. The dessin ${}^{\lambda}\mathscr{C}$ corresponds to the fraction ${}^{\lambda}\mathscr{C}$ obtained by applying λ to the coefficients of F.*

EXAMPLE 4.6 – Suppose $\mathscr{C}$ is the following dessin on the sphere:



Let us find a fraction F corresponding to $\mathscr{C}$ by the method just described. Note that, whenever the dessin is really a planar tree, one can greatly improve the efficiency of the computations, as will be explained below, but we want to illustrate the general case.

We point out that the letters b_i and w_i above are used to label the sets B and W, and the same letters will be used in the equations which we are about to write down. A tricky aspect is that, in the equations, there is really nothing to distinguish between, say, w_2, w_3, and w_4; and we expect more solutions to our system of equations than the one we want. We shall see that some solutions will actually give a different dessin.

Here there is just one face, so F will have just the one pole at ∞; in other words F will be a polynomial. As for A, it is of the form

$$A = \frac{\lambda}{(z-b_0)(z-b_1)(z-b_2)(z-w_0)(z-w_1)(z-w_2)(z-w_2)(z-w_4)}\,.$$

The first equations are obtained by comparing this with the expression

$$A = -\frac{4}{z-b_0} - \frac{1}{z-b_1} - \frac{2}{z-b_2} + \frac{2}{z-w_0} + \frac{2}{z-w_1} + \frac{1}{z-w_2} + \frac{1}{z-w_3} + \frac{1}{z-w_4}\,.$$

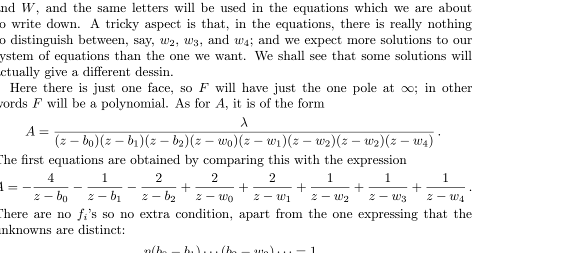
There are no f_i's so no extra condition, apart from the one expressing that the unknowns are distinct:

$$\eta(b_0 - b_1)\cdots(b_2 - w_3)\cdots = 1\,,$$

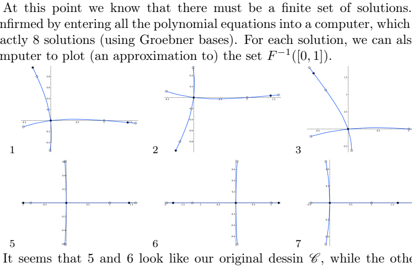
where we do not write down the 28 terms. Finally, for F to be normalized, we add

$$b_0 = 0\,, \qquad w_0 = 1\,.$$

At this point we know that there must be a finite set of solutions. This is confirmed by entering all the polynomial equations into a computer, which produces exactly 8 solutions (using Groebner bases). For each solution, we can also ask the computer to plot (an approximation to) the set $F^{-1}([0,1])$.

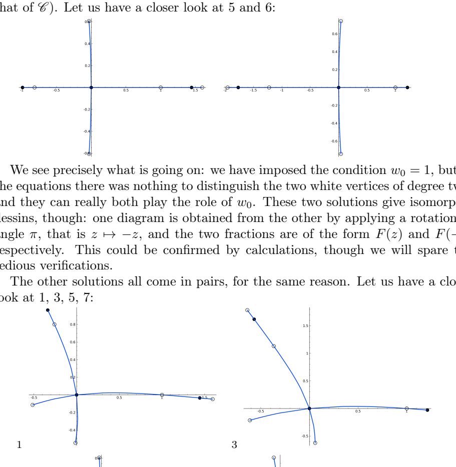


It seems that 5 and 6 look like our original dessin $\mathscr{C}$, while the other six are certainly not isomorphic to $\mathscr{C}$ (even the underlying bigraphs are not isomorphic to that of $\mathscr{C}$). Let us have a closer look at 5 and 6:



We see precisely what is going on: we have imposed the condition $w_0 = 1$, but in the equations there was nothing to distinguish the two white vertices of degree two, and they can really both play the role of w_0. These two solutions give isomorphic dessins, though: one diagram is obtained from the other by applying a rotation of angle π, that is $z \mapsto -z$, and the two fractions are of the form $F(z)$ and $F(-z)$ respectively. This could be confirmed by calculations, though we will spare the tedious verifications.

The other solutions all come in pairs, for the same reason. Let us have a closer look at 1, 3, 5, 7:



Here 1 and 3 present the same bicolored tree; 1, 5 and 7 are non-isomorphic bicolored trees. However 1 and 3 are not isomorphic dessins – or rather, they are not isomorphic as oriented dessins, as an isomorphism between the two would have to change the orientation.

Let σ, α and ϕ be the three permutations corresponding to $\mathscr{C}$. Now suppose we were to look for a dessin $\mathscr{C}'$ with permutations σ', α' and ϕ' such that σ' is conjugated to σ within S_7 (there are 7 darts here), and likewise for α' and α, and ϕ' and ϕ. Then we would write down the same equations, which only relied on the cycle types of the permutations. Thus $\mathscr{C}'$ would show up among the solutions, and conversely. So we have an interpretation of this family of four dessins.

Let us have a look at the Galois action. Here is the number b_1 in the cases 1, 3, 5, 7:

$$\frac{1}{32}\left(-2i\sqrt{5i\sqrt{7}-7\sqrt{7}} + 3i\sqrt{2}\sqrt{7} + 7\sqrt{2}\right)\sqrt{2}\,,$$

$$\frac{1}{32}\left(2\sqrt{5i\sqrt{7}+7}\sqrt{7} - 3i\sqrt{2}\sqrt{7} + 7\sqrt{2}\right)\sqrt{2}\,,$$

$$-\frac{1}{72}\left(\sqrt{8\sqrt{3}\sqrt{7}+63\sqrt{3}\sqrt{7}} - 21\sqrt{3} + 12\sqrt{7}\right)\sqrt{3}\,,$$

$$\frac{1}{72}\left(\sqrt{-8\sqrt{3}\sqrt{7}+63\sqrt{3}\sqrt{7}} + 21\sqrt{3} + 12\sqrt{7}\right)\sqrt{3}\,.$$

One can check that the minimal polynomial for b_1 in case 1 has degree 4, and that the four distinct values for b_1 in cases 1, 2, 3, 4 all have the same minimal polynomial (these are questions easily answered by a computer). Thus they are the four roots of this polynomial, which are in the same $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$-orbit. On the other hand, in cases 5, 6, 7, 8 the values for b_1 have another minimal polynomial (and they have the same one), so $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ cannot take solution 1 to any of the solutions 5, 6, 7, 8. In the end we see that the four solutions 1, 2, 3, 4 are in the same Galois orbit, in particular 1 and 3 are in the same orbit. A similar argument shows that 5 and 7 also belong to the same orbit. However these orbits are different.

Understanding the action of the absolute Galois group of $\mathbb{Q}$ on (isomorphism classes of) dessins will be a major theme in the rest of this paper.

Remark 4.7. Let us comment of efficiency issues. A seemingly anecdotal trick, whose influence on the computation is surprising, consists in grouping the vertices of the same colour and the same degree. In the last example, we would "group together" w_2, w_3 and w_4, and write

$$(z-w_2)(z-w_3)(z-w_4) = z^3 + uz^2 + vz + s\,.$$

All subsequent computations are done with the unknowns u, v and s instead of w_2, w_3 and w_4, thus *reducing the degree of the equations*.

More significant is the alternative approach at our disposal when the dessin is a planar tree. Then F is a polynomial (if we arrange for the only pole to be ∞), and F' *divides* $F(F-1)$, so $F(F-1) = PF'$, where everything in sight is a polynomial.

Coming back to the last example, we would write

$$F = cz^4(z-b_1)(z-b_2)^2$$

(incorporating $b_0 = 0$) and

$$F - 1 = c(z-1)^2(z-w_1)^2(z^3 + uz^2 + vz + s)\,,$$

the unknowns being now c, b_1, b_2, w_1, u, v and s. In the very particular case at hand, there is already a finite number of solutions to the polynomial equations resulting from the comparison of the expressions for F and $F-1$. In general though, the very easy next step is to compute the remainder in the long division of $F(F-1)$ by F', say in $\mathbb{Q}(c, b_1, b_2, w_1, u, v, s)[z]$. Since F and F' both have c as the leading coefficient, it is clear that the result will have coefficients in $\mathbb{Q}[c, b_1, b_2, w_1, u, v, s]$. These coefficients must be zero, and these are the equations to consider.

Proceeding in this way is, based on a handful of examples, several orders of magnitude faster than with the general method.

We conclude with a proof of proposition 4.3.

Proof. Let F be as in the proposition, let $A = F'/F(F-1)$, and let us write the partial fraction decomposition of A over $\mathbb{C}$:

$$A = \sum_{\alpha, r, k} \frac{\alpha}{(z-r)^k}\,.$$

Now we integrate; we do this formally, though it can be made rigorous by restricting z to lie in a certain interval of real numbers. Note that essentially we are solving the differential equation $F(F-1) = A^{-1}F'$. On the one hand:

$$\int \frac{F'(z)dz}{F(z)(F(z)-1)} = \int \frac{dF}{F(F-1)} = \int \left(\frac{-1}{F} + \frac{1}{F-1}\right) dF = \log\!\left(\frac{F-1}{F}\right),$$

up to a constant. On the other hand this must be equal to

$$\sum_{\alpha, r, k>1} \frac{\alpha}{(1-k)(z-r)^{k-1}} + \sum_{\alpha, r} \alpha \log(z-r)\,,$$

up to a constant. Thus the exponential of this last expression is a rational fraction, from which it follows that the first sum above must be zero. In other words, $k = 1$ in all the nonzero terms of the partial fraction decomposition of A. Moreover, for the same reason all α's must be integers. In the end

$$A = \sum_{\alpha, r} \frac{\alpha}{z-r}\,,$$

and

$$\frac{F-1}{F} = c\prod_{\alpha, r}(z-r)^{\alpha}\,.$$

We rewrite this

$$\frac{1}{F} = 1 - c\prod_{\alpha, r}(z-r)^{\alpha}\,.$$

Examination of this expression establishes (1) and (2) simultaneously. Indeed $F(\infty) = \infty$ implies $c = 1$ (and $\sum \alpha = 0$). Likewise, the roots of F are the numbers r's such that $\alpha < 0$, and the roots of $F-1$ are the r's such that $\alpha > 0$. The multiplicities are interpreted as degrees of vertices, as already discussed (we see that $\sum \alpha = 0$ amounts to $\sum m_i = \sum n_i$, and as a matter a fact these two sums are equal to the number n of darts, each dart joining a back vertex and a white one). Let us now use the notation b_i, w_i, n_i and m_i.

We have shown that

$$A = \lambda \frac{B}{\prod_i (z-b_i)(z-w_i)}$$

where B is a monic polynomial. It remains, in order to prove (3), to find the roots of B together with their multiplicities, knowing that B does not vanish at any b_i or any w_i.

For this write $F = P/Q$ with P, Q coprime polynomials, so that

$$A = \frac{P'Q + PQ'}{P(P-Q)}\,.$$

If f_i is a root of Q, with multiplicity r_i, then it is a root of $P'Q + PQ'$ with multiplicity $r_i - 1$. Also, it is not a root of $P(P-Q)$, so in the end f_i is a root of B of multiplicity $r_i - 1$.

Finally, from the expression $A = F'/F(F-1)$ we know that the roots of A are to be found among the roots of F' and the poles of $F(F-1)$, that is the roots of Q. So a root of A which is not a root of Q would have to be a root of F'. Now we use the fact that F is a Belyi map: a root of F' is taken by F to 0 or 1, so it is among the b_i's and the w_i's. These are not roots of B, as observed, so we have proved (3).

Now we turn to the converse, so we let A have the form in (3), we suppose that (1) holds and define F by (2). From the arguments above it is clear that $A = F'/F(F-1)$.

Is F a Belyi map? For z_0 satisfying $F'(z_0) = 0$, we need to examine whether the value $F(z_0)$ is among $0, 1, \infty$. Suppose $F(z_0)$ is neither 0 nor 1. Then it is not a root of $F(F-1)$, so it is a root of A. If we throw in the assumption that the roots of A are poles of F, it follows that $F(z_0) = \infty$. □

4.3. **Examples in genus 1; faithfulness of the action.** Let us briefly discuss the Galois action in the language of curves, as in §2.6. A dessin defines a curve C, which can be taken to be defined by homogeneous polynomial equations $P_i = 0$ in projective space, where P_i has coefficients in $\overline{\mathbb{Q}}$. Also C comes equipped with a map $F\colon C \to \mathbb{P}^1$, or equivalently $F \in \mathscr{M}(C)$, and F can be written as a quotient $F = P/Q$ where P and Q are homogeneous polynomials of the same degree, again with coefficients in $\overline{\mathbb{Q}}$. Conversely such a curve, assuming that F does not ramify except possibly at 0, 1 or ∞, defines a dessin.

It is then easy to show (though we shall not do it here) that ${}^{\lambda}\mathscr{C}$ corresponds to the curve ${}^{\lambda}C$ obtained by applying λ to the coefficients of each P_i; it comes with a Belyi map, namely ${}^{\lambda}F$, which we again obtain by applying λ to the coefficients of F. (Note in particular that ${}^{\lambda}C$, as a curve without mention of a Belyi map, is obtained from λ and C alone, and F does not enter the picture.)

We illustrate this with dessins in degree 1. An *elliptic curve* is a curve C given in $\mathbb{P}^2$ by a "Weierstrass equation", that is, one of the form

$$y^2z - x^3 - axz^2 - bz^3 = 0\,.$$

Assuming we work over $\overline{\mathbb{Q}}$ or $\mathbb{C}$, the surface $C(\mathbb{C})$ is then a torus. One can show conversely that whenever $C(\mathbb{C})$ has genus 1, the curve is an elliptic curve.

The equation is of course not uniquely determined by the curve. However one can prove that

$$j = 1728(4a)^3/16(4a^3 + 27b^2)$$

depends only on C up to isomorphism. (The notation is standard, with 1728 emphasized.) What is more, over an algebraically closed field we have a converse: the number j determines C up to isomorphism. Further, each number $j \in K$ actually corresponds to an elliptic curve over K. These are all classical results, see for example [Sil09].

Now we see that, in obvious notation, $j({}^{\lambda}C) = {}^{\lambda}j(C)$, with the following consequence. Given $\lambda \in \mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ which is not the identity, there is certainly a number $j \in \overline{\mathbb{Q}}$ such that ${}^{\lambda}j \neq j$. Considering the (unique) curve C such that $j(C) = j$, we can use Belyi's theorem to make sure that it possess a Belyi map F (it really does not matter which, for our purposes), producing at least one dessin $\mathscr{C}$. It follows that ${}^{\lambda}\mathscr{C}$ is not isomorphic to $\mathscr{C}$, and we see that the action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on dessins is *faithful*.

As it happens, one can show that the action is faithful even when restricted to genus 0, and even to plane trees. What is more, the argument is easy and elementary, see the paper by Schneps [Sch94], who ascribes the result to Lenstra.

We note for the record:

THEOREM 4.8 – *The action of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ on dessins is faithful. In fact, the action on plane trees is faithful, as is the action on dessins of genus 1.*

In this statement it is implicit that the image of a plane tree under the Galois action is another plane tree. Theorem 4.11 below proves this, and more.

4.4. **Invariants.** We would like to find common features to the dessins $\mathscr{C}$ and ${}^{\lambda}\mathscr{C}$, assumed connected for simplicity. First and foremost, if $L/\overline{\mathbb{Q}}(x)$ corresponds to $\mathscr{C}$, one must observe that there is the following commutative diagram:

$$\begin{array}{ccc} \overline{\mathbb{Q}}(x) & \xrightarrow{\ \lambda\ } & \overline{\mathbb{Q}}(x) \\ \downarrow & & \downarrow \\ L = L \otimes_{id} \overline{\mathbb{Q}}(x) & \xrightarrow{\ y \otimes a \mapsto y \otimes \lambda(a)\ } & L \otimes_{\lambda} \overline{\mathbb{Q}}(x) = {}^{\lambda}L\,. \end{array}$$

Here both horizontal arrows are isomorphisms of fields (but the bottom one is not an isomorphism of $\overline{\mathbb{Q}}(x)$-extensions, of course). It follows that there is an isomorphism

$$\lambda^*\colon \mathrm{Gal}(L/\overline{\mathbb{Q}}(x)) \longrightarrow \mathrm{Gal}({}^{\lambda}L/\overline{\mathbb{Q}}(x))\,,$$

obtained by conjugating by the bottom isomorphism (this is the approach taken in [Völ96]). Alternatively, the existence of a homomorphism λ^* between these groups is guaranteed by the functoriality of the Galois action; while the fact that λ^* is a bijection is established by noting that its inverse is $(\lambda^{-1})^*$. These two definitions of λ^* agree, as is readily seen.

For the record, we note:

LEMMA 4.9 – *If $\mathscr{C}$ is regular, so is ${}^{\lambda}\mathscr{C}$.*

Proof. It is clear that $\mathscr{C}$ and ${}^{\lambda}\mathscr{C}$ have the same degree, and their automorphism groups are isomorphic under λ^*, so the lemma is obvious. □

Using curves, we can guess a property of λ^* which is essential (a rigorous argument will be given next). Let C be a curve in projective space corresponding to $\mathscr{C}$. It is a consequence of the material in §2 that $C(\mathbb{C})$ is homeomorphic to $|\mathscr{C}|$. The automorphism $\tilde{\sigma}$ must then correspond to a self-map $C \to C$, and the latter must

5.3. **Action of $Out(H_n)$ on dessins.** We seek to define a down-to-earth description of an action of $\lim_n Out(H_n)$ on (isomorphism classes of) dessins. In fact we only define an action on *connected* dessins in what follows, and will not recall that assumption. (It is trivial to extend the action to all dessins if the reader wishes to do so.)

We work in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$, in which a typical (connected) object is $K\backslash G$, where G is a finite group with two generators σ and α and K is a subgroup. Assume that G has order $\leq n$. Then there is a surjective map $p\colon H_n \to G$, sending σ and α to the elements bearing the same name. We let $N = \ker(p)$ and $\bar{K} = p^{-1}(K)$.

Now suppose γ is an automorphism of H_n. We can consider ${}^{\gamma}G = H_n/\gamma(N)$, which we see as possessing the distinguished generators σ and α, the images under $H_n \to H_n/\gamma(N)$ of the elements with the same name. We certainly *do not* take $\gamma(\sigma)$ and $\gamma(\alpha)$ as generators; on the other hand γ induces an isomorphism of groups $G \to {}^{\gamma}G$ which is *not* compatible with the distinguished generators. Finally ${}^{\gamma}G$ has the subgroup ${}^{\gamma}K$, the image of $\gamma(\bar{K})$ under $H_n \to H_n/\gamma(N)$. The object ${}^{\gamma}K\backslash{}^{\gamma}G$ in $\mathfrak{Sets}_{\sigma,\alpha,\phi}$ is the result of applying γ to $K\backslash G$. Clearly this defines an action of $Out(H_n)$ on isomorphism classes of dessins whose cartographic group has order $\leq n$.

LEMMA 5.5 – *Suppose $\gamma \in Out(H_n)$ is of the form $\gamma = \Gamma_n(\lambda)$ for some $\lambda \in$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$. Then the action of γ on (isomorphism classes of) dessins agrees with that of λ.*

Proof. We keep the notation introduced above, and write $\mathscr{C}$ for the regular dessin defined by the finite group H_n with its two canonical generators. The dessin $X = K\backslash G$ considered is the intermediate dessin of $\mathscr{C}$ corresponding to the subgroup $\bar{K}$ of $Aut(\mathscr{C}) = H_n$. Thus ${}^{\lambda}X$ corresponds to the subgroup $\lambda^*(\bar{K})$ of $Aut({}^{\lambda}\mathscr{C}) = \lambda^*(H_n)$. Picking an isomorphism ι between $\mathscr{C}$ and ${}^{\lambda}\mathscr{C}$ as before, we see that ${}^{\lambda}X$ is isomorphic $\gamma(\bar{K})\backslash H_n$ as requested. □

LEMMA 5.6 – *The actions defined above are compatible as n varies and can be combined into a single action of $\lim_n Out(H_n)$ on the isomorphism classes of dessins.*

Proof. It suffices to prove that, for any integers n, s, if we pick $\gamma_{n+s} \in Out(H_{n+s})$ and let γ_n be its image under the projection $Out(H_{n+s}) \to Out(H_n)$, then for any dessin X whose cartographic group G has order $\leq n$ the dessins ${}^{\gamma_{n+s}}X$ and ${}^{\gamma_n}X$ are isomorphic. However this follows easily from the fact that the projection $p_{n+s}\colon H_{n+s} \to G$ factors as $p_n \circ \pi_{n+s}$, where we write $\pi_{n+s}\colon H_{n+s} \to H_n$ for the natural map. □

Now we seek to prove that the action of $\lim_n Out(H_n)$ on dessins is faithful.

THEOREM 5.7 – *The group $\lim_n Out(H_n) \cong Out(\hat{F}_2)$ acts faithfully on the set of regular dessins.*

Proof. Let $\beta \in Aut(\hat{F}_2)$ correspond to $\gamma = (\gamma_n)_{n\geq 1} \in \lim_n Out(H_n)$. If the action of this element is trivial on the set of all regular dessins, then the automorphism β must fix all open, normal subgroups of $\hat{F}_2$. However the theorem of Jarden already used in the proof of lemma 5.3 implies then that β is an inner automorphism of $\hat{F}_2$. As a result, $\gamma_n = 1$ for all n. □

Here it was necessary to see $\lim_n H_n$ as $Out(\hat{F}_2)$ to conduct the proof (or more precisely, to be able to apply Jarden's theorem which is stated in terms of $\hat{F}_2$).

COROLLARY 5.8 – *The group* Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ *acts faithfully on the set of regular dessins.*

EXAMPLE 5.9 – Suppose γ is an automorphism of H_n for which you have an explicit formula, say

$$\gamma(\sigma) = \alpha\phi\alpha^{-1}\,, \qquad \gamma(\alpha) = \alpha\,.$$

What is the effect of γ on dessins, explicitly? Discussing this for regular dessins for simplicity, say you have G, a finite group of order $\leq n$ with two distinguished generators written as always σ and α. Can we compute the effect of γ on (G, σ, α) immediately?

The answer is that some care is needed. Looking at the definitions, we write $G = H_n/N$ for some uniquely defined N, and the new dessin is $(H_n/\gamma(N), \sigma, \alpha)$. If we want to write this more simply, according to the principle that "applying γ gives the same group with new generators", we exploit the isomorphism of groups

$$G = H_n/N \longrightarrow H_n/\gamma(N)$$

which is induced by γ. Transporting the canonical generators of $H_n/\gamma(N)$ to G *via* this isomorphism gives is fact $(G, \gamma^{-1}(\sigma), \gamma^{-1}(\alpha))$ (note the inverses!).

In our case we compute $\gamma^{-1}(\sigma) = \phi$, $\gamma^{-1}(\alpha) = \alpha$. In short

$${}^{\gamma}(G, \sigma, \alpha) = (G, \phi, \alpha)$$

with, as ever, $\phi = (\sigma\alpha)^{-1}$. Incidentally, if we compare this with example 2.6, we see that the action of γ is to turn a dessin into its "dual".

5.4. **The coarse Grothendieck-Teichmüller group.** Let us give a list of conditions describing a subgroup of $\lim Out(H_n)$ containing the image of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$.

LEMMA 5.10 – *Let $\gamma = \Gamma_n(\lambda) \in Out(H_n)$, for some $\lambda \in$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$. Then γ can be represented by an element of $Aut(H_n)$, still written γ for simplicity, and enjoying the following extra properties: there exists an integer k prime to the order of H_n, and an element $f \in [H_n, H_n]$, the commutator subgroup, such that*

$$\gamma(\sigma) = \sigma^k \quad \text{and} \quad \gamma(\alpha) = f^{-1}\alpha^k f\,.$$

Moreover $\gamma(\sigma\alpha)$ is conjugated to $(\sigma\alpha)^k$.

Proof. This follows from proposition 4.10 (the branch cycle argument) applied to L_n. More precisely, let us write σ for $\tilde{\sigma}$ and α for $\tilde{\alpha}$, etc. Then there is an isomorphism ι between L_n and ${}^{\lambda}L_n$, under which $\sigma_\lambda \in Aut({}^{\lambda}L_n)$ is identified with $\sigma \in H_n$, and similarly for α_λ and ϕ_λ ; as for λ^*, it becomes $\Gamma_n(\lambda)$ when viewed in $Out(H_n)$. Thus a simple translation of the notation shows that $\gamma(\sigma)$ is conjugated to σ^k, where k is determined by the action of λ on roots of unity, while $\gamma(\alpha)$ is conjugated to α^k and $\gamma(\sigma\alpha)$ is conjugated to $(\sigma\alpha)^k$. By composing with an inner automorphism, we may thus assume that $\gamma(\sigma) = \sigma^k$.

Let $g \in H_n$ be such that $\gamma(\alpha) = g^{-1}\alpha^k g$. Every element of the abelian group $H_n/[H_n, H_n]$ can be written $\alpha^j\sigma^i$ for some integers i, j, so let us write $g = \alpha^j\sigma^i c_1$ for some $c_1 \in [H_n, H_n]$. Further put $\sigma^i c_1 = c_2 c_1 \sigma^i$; here c_2 is a commutator, so that $f = c_2 c_1 \in [H_n, H_n]$. Thus $g = \alpha^j f \sigma^i$ and

$$\gamma(\alpha) = g\alpha^k g^{-1} = (\sigma^{-i} f^{-1}\alpha^{-j})\alpha^k(\alpha^j f\sigma^i) = \sigma^{-i}(f\alpha^k f^{-1})\sigma^i\,.$$

By composing γ with conjugation by σ^i, we obtain a representative which is of the desired form. □

For each n there is an automorphism δ_n of H_n satisfying $\delta_n(\sigma) = \alpha\phi\alpha^{-1} = \sigma^{-1}\alpha^{-1}$, $\delta_n(\alpha) = \alpha$, $\delta_n(\phi) = \sigma$. We write $\delta = (\delta_n)_{n\geq 1}$ for the corresponding element of $\lim_n Out(H_n)$. The letter δ is for duality, as the next lemma explains.

LEMMA 5.11 – (1) *The dessin ${}^{\delta}\mathscr{C}$ resulting from the application of δ to an arbitrary dessin $\mathscr{C}$ is its "dual". If $\mathscr{C}$ corresponds to the surface S endowed with the Belyi map $F\colon S \to \mathbb{P}^1$, then ${}^{\delta}\mathscr{C}$ corresponds to S endowed with $1/F$.*
(2) *If $\gamma = \Gamma(\lambda) \in \lim_n Out(H_n)$ for $\lambda \in$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$, then γ and δ commute.*

Note that δ squares to conjugation by α. Thus in $Out(H_n)$, it is equal to its inverse, and the letter ω is often used in the literature for δ^{-1}.

Proof. (1) follows from the computations in example 2.6 and example 5.9.

Since the Galois action proceeds by the effect of $\lambda \in$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ on the coefficients of the equations defining S as a curve, and the coefficients of the rational fraction F, the first point implies that ${}^{\lambda\delta}\mathscr{C} \cong {}^{\delta\lambda}\mathscr{C}$ for any dessin $\mathscr{C}$. Since the action of $\lim_n Out(H_n)$ on isomorphism classes of dessins is faithful, this implies $\lambda\delta = \delta\lambda$. □

Note that we have relied on the point of view of algebraic curves in this argument.

Now we turn to the study of the automorphism of H_n usually written θ_n which satisfies $\theta_n(\sigma) = \alpha$ and $\theta_n(\alpha) = \sigma$. We write $\theta = (\theta_n)_{n\geq 1}$ for the corresponding element of $\lim_n Out(H_n)$.

LEMMA 5.12 – (1) *The dessin ${}^{\theta}\mathscr{C}$ resulting from the application of θ to an arbitrary dessin $\mathscr{C}$ is simply obtained by changing the colours of all the vertices in $\mathscr{C}$. If $\mathscr{C}$ corresponds to the surface S endowed with the Belyi map $F\colon S \to \mathbb{P}^1$, then ${}^{\theta}\mathscr{C}$ corresponds to S endowed with $1 - F$.*
(2) *If $\gamma = \Gamma(\lambda) \in \lim_n Out(H_n)$ for $\lambda \in$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$, then γ and θ commute.*

Proof. As the previous proof, based on example 2.7. □

We come to the definition of the *coarse Grothendieck-Teichmüller group*, to be denoted $\mathcal{GT}$. In fact, we start by defining the subgroup $\mathcal{GT}(n)$ of $Out(H_n)$ comprised of all the elements γ such that:

(GT0) γ has a representative in $Aut(H_n)$, say $\tilde{\gamma}$, for which there exists an integer k_n prime to the order of H_n, and an element $f_n \in [H_n, H_n]$, such that

$$\tilde{\gamma}(\sigma) = \sigma^{k_n} \quad \text{and} \quad \tilde{\gamma}(\alpha) = f_n^{-1}\alpha^{k_n} f_n\,.$$

(GT1) γ commutes with θ_n.
(GT2) γ commutes with δ_n,

Remark that conditions (GT2) and (GT0) together imply that $\tilde{\gamma}(\sigma\alpha)$ is conjugated to $(\sigma\alpha)^{k_n}$.

We let $\mathcal{GT} = \lim_n \mathcal{GT}(n)$. The contents of this section may thus be summarized as follows, throwing in the extra information we have from proposition 4.10:

THEOREM 5.13 – *There is an injective homomorphism*

$$\Gamma\colon \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \mathcal{GT}\,.$$

Moreover, for $\gamma = \Gamma(\lambda)$, the integer k_n can be taken to be any integer satisfying

$$\lambda(\zeta_N) = \zeta_N^{k_n}\,.$$

Here N is the order of H_n, and $\zeta_N = e^{\frac{2i\pi}{N}}$.

5.5. **Variants.** It should be clear that the groups H_n are not the only ones we could have worked with. In fact, let $\mathscr{N}$ be a collection of subgroups of F_2 with the following properties:

(i) each $N \in \mathscr{N}$ has finite index in F_2 ,
(ii) each $N \in \mathscr{N}$ is characteristic (and in particular normal),
(iii) for any normal subgroup K in F_2, there exists $N \in \mathscr{N}$ such that $N \subset K$.
(iv) for each $N \in \mathscr{N}$, the group $G = F_2/N$ has the following property: given two pairs of generators (g_1, g_2) and (h_1, h_2) for G, there exists an automorphism of G taking g_i to h_i, for $i = 1, 2$.

So far we have worked with $\mathscr{N}$ = the collection of all subgroups $F_2^{(n)}$ (for $n \geq 1$). Other choices include:

- For $n \geq 1$, let $F_2^{[n]}$ = the intersection of all normal subgroups of F_2 of order *dividing* n. Then take $\mathscr{N}$ = the collection, for all $n \geq 1$, of all the groups $F^{[n]}$.
- For G a finite group, let N_G = the intersection of all the normal subgroups N of F_2 such that F_2/N is isomorphic to G (the group G *not* having distinguished generators). Then take $\mathscr{N}$ = the collection of all N_G, where G runs through representatives for the isomorphism classes of finite groups which can be generated by two elements.

To establish condition (iv) in each case, one proves a more "universal" property analogous to (2) of proposition 5.1 for H_n.

The reader will check that all the preceding material is based only on these four conditions, and the results below follow *mutatis mutandis.* First, as in §5.1 we have

$$\hat{F}_2 \cong \lim_{N\in\mathscr{N}} F_2/N\,,$$

and

$$Out(\hat{F}_2) \cong \lim_{N\in\mathscr{N}} Out(F_2/N)\,.$$

In particular we have maps Gal$(\overline{\mathbb{Q}}/\mathbb{Q}) \to Out(F_2/N)$ for N running through $\mathscr{N}$, and any non-trivial element of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ has non-trivial image in some $Out(F_2/N)$.

Let us introduce the notation $\mathcal{GT}(K)$, for any characteristic subgroup K of finite index in F_2, to mean the subgroup of $Out(F_2/K)$ of those elements satisfying (GT0) - (GT1) - (GT2). Note that N being characteristic, it makes sense to speak of δ and θ as elements of $Out(F_2/N)$. In the same fashion we define $\mathcal{GT}(K)$, as a subgroup of $Out(\hat{F}_2/K)$, when K is open and characteristic in $\hat{F}_2$.

With this terminology, one proves that the elements of $Out(F_2/N)$ coming from elements of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ must in fact lie in $\mathcal{GT}(N)$. If we let $\mathcal{GT}(\mathscr{N})$ denote the inverse limit of the groups $\mathcal{GT}(N)$ for $N \in \mathscr{N}$, then it is isomorphic to a subgroup of $Out(\hat{F}_2)$ and we have an injection of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ into $\mathcal{GT}(\mathscr{N})$.

The next lemma then proves that $\mathcal{GT}(\mathscr{N})$ is independent of $\mathscr{N}$:

LEMMA 5.14 – *Let $\beta \in Out(\hat{F}_2)$. Then β lies in $\mathcal{GT}(\mathscr{N})$ if and only if for each open, characteristic subgroup K of $\hat{F}_2$, the induced element of $Out(\hat{F}_2/K)$ is in $\mathcal{GT}(K)$.*

In particular, the group $\mathcal{GT}(\mathscr{N})$, as a subgroup of $Out(\hat{F}_2)$ is independent of the choice of $\mathscr{N}$.

Proof. The condition is clearly sufficient, as we see by letting K run through the closures of the elements of $\mathscr{N}$.

To see that it is necessary, we only need to observe that K contains the closure of an element $N \in \mathscr{N}$, so $\hat{F}_2/K$ is a quotient of F_2/N and the automorphism induced by β on $\hat{F}_2/K$ is also induced by an element of $\mathcal{GT}(N)$; thus it must lie in $\mathcal{GT}(K)$.

This characterization of elements of $\mathcal{GT}(\mathscr{N})$ visibly does not make any reference to $\mathscr{N}$. □

In theory, all choices for $\mathscr{N}$ are equally valid, and in fact no mention of *any* choice is necessary: one may state all the results of this section in terms of $Out(\hat{F}_2)$, for example defining $\mathcal{GT}$ by the characteristic property given in the lemma. In practice however, choosing a collection $\mathscr{N}$ allows us to compute $\mathcal{GT}(N)$ explicitly for some groups $N \in \mathscr{N}$, and that is at least a baby step towards a description of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$. The difficulty of the computations will depend greatly on the choices we make. For example, with the groups $F^{(n)}$, the order of H_n increases very rapidly with n, but the indexing set is very simple; with $F^{[n]}$, the order of $F_2/F_2^{[n]}$ is much less than the order of H_n, but the inverse limits are more involved. In a subsequent publication, computations with the family $\mathscr{N}$ of all the groups of the form N_G will be presented.

We conclude with yet another definition of $\mathcal{GT}$ which does involve choosing a collection $\mathscr{N}$. This is the traditional definition.

5.6. **Taking coordinates; the group $\widehat{\mathcal{GT}}_0$.** We start with a couple of observations about H_n.

LEMMA 5.15 – *If k_1 and k_2 are integers such that σ^{k_1} and σ^{k_2} are conjugate in H_n, then $k_1 \equiv k_2$ mod n. Similarly for α.*

Proof. We use the map $H_n \to C_n = \langle x\rangle$, where C_n is the cyclic group of order n, sending both σ and α to x. The image of σ^{k_i} is x^{k_i} (for $i = 1, 2$), and conjugate elements of C_n are equal, so $k_1 \equiv k_2$ mod n. □

COROLLARY 5.16 – *Let $\gamma \in \mathcal{GT}(n)$. For $i = 1, 2$, let $\tilde{\gamma}_i$ be a representative for γ in $Aut(H_n)$ such that $\tilde{\gamma}_i(\sigma)$ is conjugate to σ^{k_i}. Then $k_1 \equiv k_2$ mod n. This defines a homomorphism*

$$\mathcal{GT}(n) \longrightarrow (\mathbb{Z}/n)^\times\,,$$

which we write $\gamma \mapsto k(\gamma)$ (or sometimes $k_n(\gamma)$ for emphasis).

Letting n vary, we obtain a homomorphism

$$k\colon \mathcal{GT} \longrightarrow \hat{\mathbb{Z}}^\times\,.$$

Here $\hat{\mathbb{Z}} = \lim_n \mathbb{Z}/n\mathbb{Z}$ is the profinite completion of the ring $\mathbb{Z}$.

PROPOSITION 5.17 – *Let $\gamma \in \mathcal{GT}$. Then γ has a lift $\beta \in Aut(\hat{F}_2)$ satisfying*

$$\beta(\sigma) = \sigma^{k(\gamma)}\,, \qquad \beta(\alpha) = f^{-1}\alpha^{k(\gamma)}f\,,$$

for some $f \in [\hat{F}_2, \hat{F}_2]$, the commutator subgroup. The element f is unique, and as a result, so is β.

Proof. Start with any lift β_0. The elements $\beta_0(\sigma)$ and $\sigma^{k(\gamma)}$ are conjugate in every group H_n, so $\beta_0(\sigma)$ is in the closure of the conjugacy class of $\sigma^{k(\gamma)}$. However this class is closed (the map $x \mapsto x\sigma^{k(\gamma)}x^{-1}$ is continuous and its image must be closed since its source $\hat{F}_2$ is compact). So $\beta_0(\sigma)$ is conjugated to $\sigma^{k(\gamma)}$, and likewise $\beta_0(\alpha)$ is conjugated to $\alpha^{k(\gamma)}$. Now, argue as in lemma 5.10 to obtain the existence of a representative β as stated.

We turn to the uniqueness. If f' can replace f, then $f = c_1 f' c_2$ where c_2 centralizes σ and c_1 centralizes α. However the centralizer of σ in $\hat{F}_2$ is the (closed) subgroup generated by σ and likewise for α. Since f and f' are assumed to be both commutators, we can reduce mod $[\hat{F}_2, \hat{F}_2]$ and obtain a relation $c_1c_2 = 1$; the latter must then hold true in any finite, abelian group on two generators σ and α, and this is clearly only possible if $c_1 = c_2 = 1$ in $\hat{F}_2$. □

We observe at once:

COROLLARY 5.18 – *The injection $\Gamma\colon$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q}) \to Out(\hat{F}_2)$ lifts to an injection $\tilde{\Gamma}\colon$ Gal$(\overline{\mathbb{Q}}/\mathbb{Q}) \to Aut(\hat{F}_2)$. In particular, an element of* Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ *can be entirely described by a pair $(k, f) \in \hat{\mathbb{Z}}^\times \times [\hat{F}_2, \hat{F}_2]$.*

Proof. Let $\tilde{\Gamma}(\lambda)$ be the lift of $\Gamma(\lambda)$ described in the proposition. The composition of two automorphisms of $\hat{F}_2$ of this form is again of this form, so $\tilde{\Gamma}(\lambda)\tilde{\Gamma}(\mu)$ must be the lift of $\Gamma(\lambda)\Gamma(\mu) = \Gamma(\lambda\mu)$, that is, it must be equal to $\tilde{\Gamma}(\lambda\mu)$. □

We want to describe a group analogous to $\mathcal{GT}$ in terms of the pairs (k, f). There is a subtlety here, in that if we pick $k \in \hat{\mathbb{Z}}^\times$ and $f \in [\hat{F}_2, \hat{F}_2]$ arbitrarily, the self-homomorphism β of $\hat{F}_2$ satisfying

$$(*) \qquad \beta(\sigma) = \sigma^k\,, \qquad \beta(\alpha) = f^{-1}\alpha^k f$$

may not be an automorphism. Keeping this in mind, we define a group $\widehat{\mathcal{GT}}_0$ now – the notation is standard, and the index "0" is not to be confused with our writing $\mathcal{GT}(n)$ for $n = 0$; moreover the notation does not refer to a profinite completion of some underlying group $\mathcal{GT}_0$. So let $\widehat{\mathcal{GT}}_0$ be the group of all pairs $(k, f) \in \hat{\mathbb{Z}}^\times \times [\hat{F}_2, \hat{F}_2]$ such that :

- Let β be the self-homomorphism defined by (*); then β is an automorphism.
- β commutes with δ in $Out(\hat{F}_2)$.
- β commutes with θ in $Out(\hat{F}_2)$.

The composition law on $\widehat{\mathcal{GT}}_0$ is defined *via* the composition of the corresponding automorphisms of $\hat{F}_2$; one may recover k and f from β, and indeed $\widehat{\mathcal{GT}}_0$ could have been defined as a subgroup of $Aut(\hat{F}_2)$, though that is not what has been traditionally done in the literature.

The definition of $\widehat{\mathcal{GT}}_0$ was given by Drinfeld in [Dri90]. The reader who is familiar with *loc. cit.* may not recognize $\widehat{\mathcal{GT}}_0$ immediately behind our three conditions, so let us add:

LEMMA 5.19 – *This definition of $\widehat{\mathcal{GT}}_0$ agrees with Drinfeld's.*

Proof. This follows from [Sch97], §1.2, last theorem, stating that "conditions (I) and (II)" are equivalent with the commutativity conditions with θ and δ respectively (the author using the notation ω for an inverse of δ in $Out(\hat{F}_2)$). □

The natural map $Aut(\hat{F}_2) \to Out(\hat{F}_2)$ induces a map $\widehat{\mathcal{GT}}_0 \to \mathcal{GT}$. The existence and uniquess statements in proposition 5.17 imply the surjectivity and injectivity of this map, respectively, hence:

PROPOSITION 5.20 – *$\widehat{\mathcal{GT}}_0$ and $\mathcal{GT}$ are isomorphic.*

One may rewrite the main theorem of this section, theorem 5.13, as follows:

THEOREM 5.21 – *There is an injective homomorphism of groups*

$$\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \longrightarrow \widehat{\mathcal{GT}}_0\,.$$

Composing this homomorphism with the projection $\widehat{\mathcal{GT}}_0 \to \hat{\mathbb{Z}}^\times$ gives the cyclotomic character of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$.

We conclude with a few remarks about the (real) *Grothendieck-Teichmüller group*. This is a certain subgroup of $\widehat{\mathcal{GT}}_0$, denoted $\widehat{\mathcal{GT}}$, also defined by Drinfeld in [Dri90]. It consists of all the elements of $\widehat{\mathcal{GT}}_0$ satisfying the so-called "pentagon equation" (or "condition (III)").

Ihara in [Iha94] was the first to prove the existence of an injection of Gal$(\overline{\mathbb{Q}}/\mathbb{Q})$ into $\widehat{\mathcal{GT}}$. His method is quite different from ours, and indeed proving the pentagon equation following our elementary approach would require quite a bit of extra work.

Another noteworthy feature of Ihara's proof (beside the fact that it refines ours by dealing with $\widehat{\mathcal{GT}}$ rather than $\widehat{\mathcal{GT}}_0$) is that it does not, or at least not explicitly, refer to dessins d'enfants. It is pretty clear that the original ideas stem from the material in the *esquisse* [Gro97] on dessins, but the children's drawings have disappeared from the formal argument. We hope to have demonstrated that the elementary methods could be pushed quite a long way.

References

[BCG] Ingrid Bauer, Fabrizio Catanese, and Fritz Grunewald, *Faithful actions of the absolute Galois group on connected components of moduli spaces*, preprint.

[Bel79] G. V. Belyĭ, *Galois extensions of a maximal cyclotomic field*, Izv. Akad. Nauk SSSR Ser. Mat. **43** (1979), no. 2, 267–276, 479. MR 534593 (80f:12008)

[Bre97] Glen E. Bredon, *Topology and geometry*, Graduate Texts in Mathematics, vol. 139, Springer-Verlag, New York, 1997, Corrected third printing of the 1993 original. MR 1700700 (2000b:55001)

[CM57] H. S. M. Coxeter and W. O. J. Moser, *Generators and relations for discrete groups*, Springer-Verlag, Berlin, 1957. MR 0088489 (19,527d)

[DD79] Régine Douady and Adrien Douady, *Algèbre et théories galoisiennes. 2*, CEDIC, Paris, 1979, Théories galoisiennes. [Galois theories]. MR 595328 (82b:12024b)

[Dri90] V. G. Drinfel′d, *On quasitriangular quasi-Hopf algebras and on a group that is closely connected with* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, Algebra i Analiz **2** (1990), no. 4, 149–181. MR 1080203 (92f:16047)

[Fre] Benoît Fresse, *Homotopy of operads and Grothendieck-Teichmüller groups*, Research monograph in preparation.

[GH94] Phillip Griffiths and Joseph Harris, *Principles of algebraic geometry*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1994, Reprint of the 1978 original. MR 1288523 (95d:14001)

[Gro97] Alexandre Grothendieck, *Esquisse d'un programme*, Geometric Galois actions, 1, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, With an English translation on pp. 243–283, pp. 5–48. MR 1483107 (99c:14034)

[Iha94] Yasutaka Ihara, *On the embedding of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *into* $\widehat{\mathrm{GT}}$, The Grothendieck theory of dessins d'enfants (Luminy, 1993), London Math. Soc. Lecture Note Ser., vol. 200, Cambridge Univ. Press, Cambridge, 1994, pp. 289–321. MR 1305402 (96b:14014)

[Jar80] Moshe Jarden, *Normal automorphisms of free profinite groups*, J. Algebra **62** (1980), no. 1, 118–123. MR 561120 (81b:20024)

[JS78] Gareth A. Jones and David Singerman, *Theory of maps on orientable surfaces*, Proc. London Math. Soc. (3) **37** (1978), no. 2, 273–307. MR 0505721 (58 #21744)

[JZGD] Andrei Jaikin-Zapirain and Gabino Gonzalez-Diez, *The absolute Galois group acts faithfully on regular dessins and on Beauville surfaces*, preprint.

[LZ04] Sergei K. Lando and Alexander K. Zvonkin, *Graphs on surfaces and their applications*, Encyclopaedia of Mathematical Sciences, vol. 141, Springer-Verlag, Berlin, 2004, With an appendix by Don B. Zagier, Low-Dimensional Topology, II. MR 2036721 (2005b:14068)

[Sch94] Leila Schneps, *Dessins d'enfants on the Riemann sphere*, The Grothendieck theory of dessins d'enfants (Luminy, 1993), London Math. Soc. Lecture Note Ser., vol. 200, Cambridge Univ. Press, Cambridge, 1994, pp. 47–77. MR 1305393 (95j:11061)

[Sch97] ———, *The Grothendieck-Teichmüller group* $\widehat{\mathrm{GT}}$*: a survey*, Geometric Galois actions, 1, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, pp. 183–203. MR 1483118 (99a:14043)

[Sil09] Joseph H. Silverman, *The arithmetic of elliptic curves*, second ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009. MR 2514094 (2010i:11005)

[Völ96] Helmut Völklein, *Groups as Galois groups*, Cambridge Studies in Advanced Mathematics, vol. 53, Cambridge University Press, Cambridge, 1996, An introduction. MR 1405612 (98b:12003)

[Š13] Jozef Širáň, *How symmetric can maps on surfaces be?*, Surveys in combinatorics 2013, London Math. Soc. Lecture Note Ser., vol. 409, Cambridge Univ. Press, Cambridge, 2013, pp. 161–238.

[Wol97] Jürgen Wolfart, *The "obvious" part of Belyi's theorem and Riemann surfaces with many automorphisms*, Geometric Galois actions, 1, London Math. Soc. Lecture Note Ser., vol. 242, Cambridge Univ. Press, Cambridge, 1997, pp. 97–112. MR 1483112 (99a:14036)