# UNE AUTRE APPLICATION DE LA THEORIE DES ENSEMBLES

**Patrick DEHORNOY**

Mathématiques, Université de Caen, 14 032 Caen, France
dehornoy@math.unicaen.fr

La théorie des ensembles explore la notion d'infini: il est donc peu surprenant que la plupart de ses résultats, et, partant, de ses applications classiques, comme celles présentées dans [32], concernent des objets « très grands » (groupes non dénombrables, espaces topologiques non métrisables...) ou « très compliqués » (sous-ensembles non boréliens de la droite réelle, monstres divers de la topologie générale... ), objets à l'évidence assez éloignés de ceux que la pratique mathématique aborde le plus usuellement et sur lesquels se concentre l'intérêt le plus général. La fascination initiale exercée par toute théorie des fondements laisse alors parfois la place à une déception devant des enjeux perçus comme marginaux ou artificiels. Je voudrais contribuer ici un peu à la défense de la théorie des ensembles en montrant comment celle-ci a pu mener récemment à une *autre* application, d'un type bien différent des précédentes, et dont une caractéristique est de ne mettre en jeu que des objets très naturels.

## 1. Le problème de la classification des tresses

A mille lieues des « horreurs » de la théorie des ensembles, les tresses sont des objets mathématiques des plus concrets, et elles jouent un rôle central et profond dans de nombreux développements récents (voir par exemple [4], qui inclut une bibliographie fournie, ou [25]). Le problème de la classification des tresses est une question de facture on ne peut plus classique: on considère des diagrammes de tresses, configurations formées de brins qui se croisent suivant le modèle de la Figure 1, et on cherche à reconnaître si deux tels diagrammes représentent les projections de configurations de dimension 3 isotopes, c'est-à-dire telles qu'on puisse passer de l'une à l'autre en déplaçant des brins, mais sans les faire se traverser ni détacher les extrémités—comme c'est par exemple le cas pour les deux diagrammes de la figure.

A la différence du problème général de l'isotopie des nœuds, dont il est un cas particulier[1], le problème de l'isotopie des tresses est un problème facile,

[1] Tout nœud (plongement de $S^1$ dans $\mathbb{R}^3$) peut se réaliser comme la clôture d'une tresse, obtenue en reliant les brins de sortie aux brins d'entrée (*cf.* [33]), et les clôtures de tresses isotopes sont évidemment des nœuds isotopes – mais la réciproque est fausse.

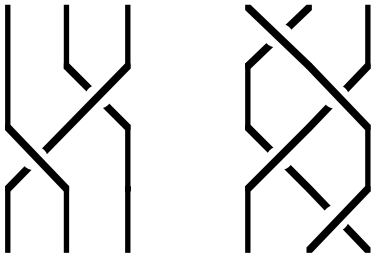
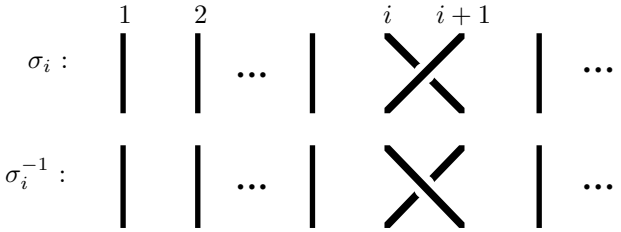


**Figure 1**: Deux tresses équivalentes

car la possibilité de composer les diagrammes de tresse à un même nombre $n$ de brins en « accrochant le second sous le premier » induit une structure de groupe sur les classes d'équivalence: le groupe des tresses à $n$ brins, introduit par Emil Artin et traditionnellement noté $B_n$. Tout diagramme de tresse peut alors clairement s'écrire comme une composition de diagrammes élémentaires à un seul croisement des types suivants:




Par exemple les décompositions des diagrammes de la Figure 1 sont $\sigma_2\sigma_1^{-1}$ et $\sigma_1^{-1}\sigma_2^{-1}\sigma_1\sigma_2$. Or Artin a observé que les relations

$$\sigma_i\sigma_j = \sigma_j\sigma_i \qquad \text{pour } |i-j| \geq 2 \tag{$\mathcal{R}_1$}$$

$$\sigma_1\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \tag{$\mathcal{R}_2$}$$

constituent une présentation pour les groupes de tresses, de sorte que le problème (topologique) de l'isotopie des tresses se réduit au problème (algébrique) de reconnaître si deux mots formés sur les générateurs $\sigma_i$ et $\sigma_i^{-1}$ deviennent égaux lorsque les relations $(\mathcal{R}_1)$ et $(\mathcal{R}_2)$ sont imposées (ainsi que $\sigma_i\sigma_i^{-1} = \sigma_i^{-1}\sigma_i = 1$).

Ce problème se trouve être assez simplement résoluble, mais, notamment à cause des liens avec les nœuds, il a été l'objet d'un intérêt soutenu au cours des décennies passées, et plusieurs solutions ont été proposées. La méthode initiale d'Artin [1], reposant sur une décomposition des groupes de tresses en produit semi-direct de groupes libres, montre que le problème est décidable, mais elle n'est guère efficace algorithmiquement. Les méthodes de Garside [22] (généralisée dans [6]), puis de ElRifai et Morton [19] ont progressé sur ce point. L'état de l'art jusque récemment était la solution proposée par Thurston en 1989 (*cf.* [20], ou [5] pour une extension à tous les groupes d'Artin de type fini), qui repose sur l'existence d'une structure de groupe automatique pour les

groupes de tresses: il existe un automate à nombre fini d'états—les $n!$ permutations des entiers 1, ..., $n$ pour les tresses à $n$ brins—qui en un certain sens calcule la loi du groupe, et fournit à la fois une inégalité isopérimétrique et un algorithme de complexité quadratique (donc efficace) pour résoudre le problème de mots. Il est aisé de mettre en œuvre l'algorithme de Thurston qui donne des résultats excellents pour les tresses de taille modérée (*cf.* implantation par J. Michel dans le système GAP), mais, à cause du facteur $n!$ ci-dessus, on ne peut aller au-delà d'un nombre de brins assez faible, typiquement une (petite) dizaine de brins. Peut-on faire mieux?

## 2. La réduction des tresses

... Oui, en utilisant la méthode qu'on va décrire maintenant.

Comme pour tout problème de mot lié à une présentation d'un groupe, il suffit, pour décider de l'équivalence de deux mots quelconques, de savoir reconnaître si un mot est trivial, c'est-à-dire équivalent au mot vide représentant ici un diagramme sans croisement: le mot $u$ est équivalent au mot $v$ si le mot $uv^{-1}$ est trivial. Or pensons d'abord au cas des groupes libres: pour savoir si un mot est trivial ou non dans ce cas, il suffit de *réduire* ce mot en supprimant de proche en proche tous les motifs $aa^{-1}$ ou $a^{-1}a$ qui y apparaissent, et le mot initial est trivial si et seulement si ce processus aboutit au mot vide. Cette condition, toujours suffisante, n'est bien sûr plus nécessaire dans le cas d'un groupe non libre, précisément à cause de l'existence de relations additionnelles qui traduisent la non-liberté du groupe: par exemple, le mot de tresse $\sigma_1\sigma_2\sigma_1\sigma_2^{-1}\sigma_1^{-1}\sigma_2^{-1}$ est trivial en vertu de la relation $(\mathcal{R}_2)$, mais il ne contient aucun motif $\sigma_i\sigma_i^{-1}$ ou $\sigma_i^{-1}\sigma_i$.

Essayons néanmoins de définir une notion de réduction des mots de tresse de sorte que les mots triviaux soient à nouveau exactement ceux qui se réduisent au mot vide. Il faut alors réduire d'autres motifs que les simples motifs $\sigma_i\sigma_i^{-1}$ et $\sigma_i^{-1}\sigma_i$. Or considérons les motifs du type $\sigma_i w\sigma_i^{-1}$ ou $\sigma_i^{-1}w\sigma_i$ où $w$ est un mot quelconque ne faisant intervenir que les générateurs $\sigma_j$ ou $\sigma_j^{-1}$ avec $j > i$. Géométriquement, un tel motif correspond à une « poignée à gauche » formée par le $i+1$-ème brin, comme dans la partie gauche de la Figure 2. On peut supprimer cette poignée sans changer la classe d'isotopie en forçant le $i+1$-ème brin à contourner par la droite et non plus par la gauche les croisements voisins à droite (s'il en existe), suivant le schéma de droite de la figure.

Comme on lit sur la figure, la réduction précédente revient, en termes syntaxiques, à supprimer dans le motif $\sigma_i^{\pm1}w\sigma_i^{\mp1}$ les générateurs $\sigma_i^{\pm1}$ et à remplacer dans le facteur $w$ médian chaque générateur $\sigma_{i+1}$ par le produit $\sigma_{i+1}^{\mp1}\sigma_i\sigma_{i+1}^{\pm1}$. Il est clair que la réduction « groupe libre » est un cas particulier de la nouvelle réduction, mais que celle-ci offre de nouvelles possibilités: par exemple, on pourra réduire le mot $\sigma_1\sigma_2\sigma_1\sigma_2^{-1}\sigma_1^{-1}\sigma_2^{-1}$ en $\sigma_1\sigma_2\sigma_2^{-1}\sigma_1^{-1}\sigma_2\sigma_2^{-1}$, et, de là, en le mot vide.

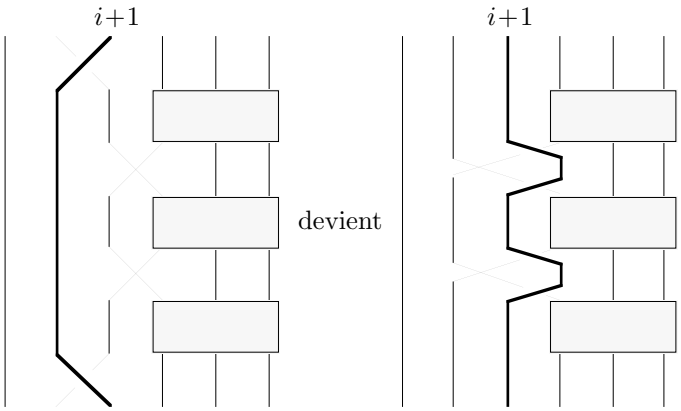



**Figure 2**: Réduction d'une poignée

**Proposition 1.** (1994, [12]) *La réduction ci-dessus résout le problème d'isotopie des tresses: un mot de tresse est trivial si, et seulement si, il se réduit au mot vide*[2].

On a donc une nouvelle solution, très simple, au problème de l'isotopie des tresses. De surcroît, cette solution se trouve être très efficace (voir le logiciel [13]), en tout cas nettement plus que les méthodes précédemment connues: le caractère purement local de la réduction la rend indifférente au nombre de brins des tresses considérées, et une implantation artisanale sur un microordinateur permet de traiter en moins d'une seconde des tresses à plusieurs milliers de croisements pour lesquelles la méthode de Thurston n'aboutit pas en 24 heures.

Or ces résultats d'énoncé élémentaire ne sont guère naturels *a priori*: tant les raisons de considérer cette réduction des mots que celles qui en assurent la validité sont peu évidentes. En particulier la *terminaison* de la réduction pose problème: la longueur des mots peut augmenter dans le processus, et, comme le montreront quelques essais sur des mots pris au hasard (voire sur les mots $(\sigma_2^2\sigma_1^2)^k\sigma_2^{-k}\sigma_3^{-1}\sigma_2^{-1}\sigma_1^{-1}$ qui sont spécialement « retors »), la réduction d'une poignée peut en faire apparaître de nouvelles d'une façon qui semble bien erratique. Il n'est pas clair en particulier que la réduction ne puisse pas parcourir une boucle *ad vitam æternam*...

Ainsi la réduction est une méthode aussi nouvelle qu'efficace pour résoudre le problème d'isotopie des tresses. Restent à en justifier l'introduction et à en comprendre les propriétés « magiques »: pourquoi considérer une telle opération, pourquoi fonctionne-t-elle?

[2] De façon précise, il faut restreindre la réduction aux motifs $\sigma_i^e w \sigma_i^{-e}$ tels que $w$ ne contienne pas lui-même une poignée de type $\sigma_{i+1}^d v \sigma_{i+1}^{-d}$: autrement dit, quand deux poignées sont imbriquées l'une dans l'autre, il faut d'abord réduire celle « de l'intérieur ».

## 3. L'ordre des tresses

... Parce qu'il existe un *ordre total* sur les tresses, caractérisé en termes de décompositions « sans poignée »:

**Proposition 2.** (1992, [9]) *La relation $<$ sur les tresses telle que $\beta_1 < \beta_2$ est vraie si et seulement si la tresse $\beta_1^{-1}\beta_2$ possède une décomposition dans laquelle le générateur de plus petit indice*[2] *n'apparaît que positivement est un ordre total.*

Par exemple, dans le mot $\sigma_3^{-1}\sigma_2\sigma_4\sigma_2\sigma_3$, le générateur de plus petit indice, ici $\sigma_2$, n'apparaît que positivement (pas de $\sigma_2^{-1}$), et donc ce mot est au dessus de la tresse-unité dans l'ordre $<$. L'ordre total des tresses prolonge les ordres partiels qui avaient été considérés antérieurement, comme celui de la divisibilité ou celui d'ElRifai-Morton [19]. Une de ses propriétés remarquables est que sa restriction aux tresses admettant une décomposition où les inverses des générateurs $\sigma_i^{-1}$ n'interviennent pas est un bon ordre (Laver, [31]), de type ordinal $\omega^{\omega^{n-2}}$ pour les tresses à $n$ brins (Burckel, [3]), de sorte que toute tresse s'écrit canoniquement comme quotient de deux ordinaux—à la façon dont tout élément de $\mathbb{Z}$ ($\simeq B_2$) s'écrit comme différence de deux entiers naturels.

Revenant à la réduction des tresses, on note qu'un mot de tresse non vide et réduit, c'est-à-dire terminal vis-à-vis de la réduction du paragraphe précédent, a certainement la propriété que le générateur de plus petit indice n'y apparaît que positivement, ou que négativement, puisque, sinon, il contiendrait une « poignée » $\sigma_i^{\pm 1} w \sigma_i^{\mp 1}$ qui pourrait donner lieu à une réduction. Ce qu'affirme la proposition est la conjonction de deux résultats:
- (i) un mot de tresse réduit non vide ne peut être trivial;
- (ii) toute tresse admet une décomposition réduite.

On voit alors comment s'introduit naturellement la réduction, à savoir comme la tentative la plus naïve pour construire une décomposition réduite: puisqu'il doit exister une décomposition sans poignée, essayons de l'obtenir en éliminant de proche en proche ces dernières[2].

C'est aussi l'ordre des tresses qui explique la validité de la méthode de réduction. De façon schématique, la structure de groupe automatique des groupes de tresses fournit une borne sur les mots obtenus par réduction à partir d'un mot donné, de sorte que le point crucial reste de montrer qu'il ne peut y avoir de boucle dans les réductions. Or appelons *préfixe critique* d'un mot de tresse non réduit le début de ce mot qui va jusqu'à la première lettre de la première poignée. Le comportement du préfixe critique apparaît spécialement imprévisible au sens où, suivant le signe des occurrences du générateur $\sigma_{i+1}$ dans une poignée de type $\sigma_i^{\pm 1} w \sigma_i^{\mp 1}$, tantôt il s'allonge, tantôt

[2] c'est-à-dire $\sigma_1$, ou $\sigma_2$ s'il n'y a pas de $\sigma_1^{\pm 1}$, *etc.*

[2] D'autres méthodes ont été proposées, notamment celle, très sophistiquée, de D. Larue dans [28].

il se raccourcit. Mais, dans tous les cas, il décroît vis-à-vis de la relation $<$ (mais pas nécessairement vis-à-vis des ordres partiels que celle-ci prolonge). Le fait que la relation $<$ soit un ordre, et, plus précisément, le point (i) qui exprime qu'elle n'a pas de cycle, interdit alors toute boucle pour les préfixes critiques, et, de là, pour la réduction.

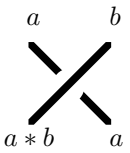
Finalement, c'est encore l'ordre des tresses qui explique l'efficacité de la méthode de réduction: en des termes heuristiques, il donne une *orientation* dans un processus qui est essentiellement la recherche d'une géodésique dans un fragment du graphe de Cayley d'un groupe $B_n$, alors que la méthode de Thurston consiste en un parcours exhaustif, ou, tout au moins, aveugle, de ce fragment faute d'une telle orientation.

Ainsi l'ordre sur les tresses est l'origine et le fondement de la réduction des tresses. Mais alors, d'où vient cet ordre?

## 4. Coloriages des tresses

... De l'étude des systèmes distributifs libres.

Pour faire le lien, partons de l'idée de colorier les brins des tresses: on fixe un ensemble quelconque, on attribue des couleurs prises dans cet ensemble aux brins des tresses « en haut » du diagramme, on propage ces couleurs vers le bas, et l'idée est de tirer des informations sur la tresse de la comparaison des couleurs d'entrée et de sortie. Si les couleurs ne se modifient pas lors des croisements, la seule information qu'on peut ainsi extraire est la permutation des brins induite par la tresse. Si par contre on autorise des changements de couleur, on peut obtenir davantage. On considère ici le type de changement suivant: dans le croisement de deux brins, le brin postérieur garde sa couleur, alors que le brin antérieur la change éventuellement, mais sa nouvelle couleur ne dépend que des couleurs précédentes des deux brins qui se croisent. C'est dire qu'on suppose donnée sur l'ensemble des couleurs une opération binaire $*$, et que les changements de couleur se font suivant le schéma




Maintenant, pour que la comparaison des couleurs d'entrée et de sortie donne des informations sur la tresse elle-même, et pas seulement sur le diagramme de tresse choisi pour la représenter, il faut que les coloriages soient compatibles avec les relations de tresse $(\mathcal{R}_1)$ et $(\mathcal{R}_2)$. Pour les premières, la compatibilité est automatique. Pour les secondes, une vérification immédiate montrera qu'elle est garantie dès que l'opération $*$ satisfait à l'identité algébrique

$$x * (y * z) = (x * y) * (x * z), \qquad (LD)$$

naturellement appelée autodistributivité à gauche. Cette condition est par exemple trivialement vérifiée dans le cas des couleurs fixes, qui correspond au choix $a * b = b$.

Appelons *LD-système* tout ensemble muni d'une opération satisfaisant à l'identité (LD): chaque LD-système est un candidat potentiel pour colorier les brins des tresses. Cette approche n'est, aux détails de la présentation près, pas nouvelle, puisqu'elle réapparaît depuis des années dans des formalismes divers et plus ou moins équivalents: « quandles » de Joyce [23], ensembles automorphes de Brieskorn [2], cristaux de Kauffman [26], « racks » de Fenn et Rourke [21]. L'idée de colorier les brins des tresses et des nœuds, ou des régions qui les séparent, remonte elle-même au moins à Alexander.

Or il existe essentiellement deux exemples classiques de LD-systèmes, et les informations sur les tresses qu'on en peut tirer sont non moins classiques. Si $G$ est un groupe et que l'opération $*$ en est la conjugaison $a * b = aba^{-1}$, la comparaison des couleurs d'entrée et de sortie d'une tresse donne l'automorphisme de $G$ associé à la tresse, ce qui, dans le cas maximal d'un groupe libre, fournit la réalisation des groupes $B_n$ comme groupes d'automorphismes—en termes de nœuds, on en déduit la présentation de Wirtinger pour le groupe fondamental du complémentaire du nœud obtenu par clôture de la tresse. Si maintenant on part d'un anneau du type $R[t, t^{-1}]$ et que $*$ est la moyenne barycentrique $a * b = (1 - t)a + tb$, les couleurs de sortie sont des combinaisons linéaires des couleurs d'entrée et la matrice de passage est la matrice de Burau de la tresse—pour les nœuds, on en déduit cette fois le polynôme d'Alexander de la clôture de la tresse[3].

Il est naturel d'escompter des résultats nouveaux sur les tresses à chaque fois qu'un exemple réellement nouveau de LD-système sera explicitement décrit. Or soit $\mathfrak{f}$ le LD-système *libre* à un générateur. Un tel objet, analogue par exemple au semigroupe des entiers lorsque l'identité d'autodistributivité remplace l'identité d'associativité, existe certainement, et il est très différent des LD-systèmes décrits ci-dessus: on montre facilement que $\mathfrak{f}$ est infini, alors que tout sous-système à un générateur construit à partir de la conjugaison ou du barycentre est trivial, puisque ces opérations vérifient $a * a = a$ pour tout $a$. Maintenant on a le résultat suivant:

**Proposition 3.** (1992, [9]) *La division de $\mathfrak{f}$ n'a pas de cycle: on ne peut pas avoir $a_1 \mid a_2 \mid \ldots \mid a_n \mid a_1$, où $a \mid a'$ signifie que $a$ divise $a'$ à gauche, c'est-à-dire qu'il existe $x$ vérifiant $a' = a * x$.*

Il est alors immédiat de démontrer la Proposition 2 sur les tresses, ou, tout au moins, le point (i)[4]. Il s'agit de montrer qu'un mot de tresse dans lequel le

[3] Le second exemple n'est qu'une version linéarisée du premier comme on le voit en utilisant le calcul différentiel libre. Un autre exemple classique de LD-système est constitué par les systèmes de racines et leurs réflexions, mais il ne s'agit ici que d'une variante du barycentre $a * b = 2a - b$.

[4] Le reste de la Proposition 2, à savoir le point (ii) mentionné plus haut, provient lui aussi d'une propriété de $\mathfrak{f}$, à savoir que deux éléments de cet ensemble sont toujours comparables vis-à-vis de la divisibilité itérée: de deux éléments distincts de $\mathfrak{f}$, l'un est toujours un diviseur de l'autre, ou un diviseur d'un diviseur de l'autre, *etc.* En d'autres termes, la divisibilité itérée est un ordre total sur $\mathfrak{f}$.

générateur $\sigma_1$ apparaît au moins une fois, mais où $\sigma_1^{-1}$ n'apparaît pas, ne peut pas être trivial. Or considérons un tel mot: le diagramme associé est du type représenté dans la Figure 3. Colorions-le à l'aide de couleurs prises dans ∮: si $a_0$, $a_1$, ... sont les couleurs successives qui apparaissent sur le brin de gauche, alors par construction $a_0$ divise $a_1$, qui divise $a_2$, *etc.* Par la Proposition 3 la couleur de sortie $a_n$ ne peut être égale à $a_0$, ce qui serait forcément le cas si la tresse était triviale[5]...

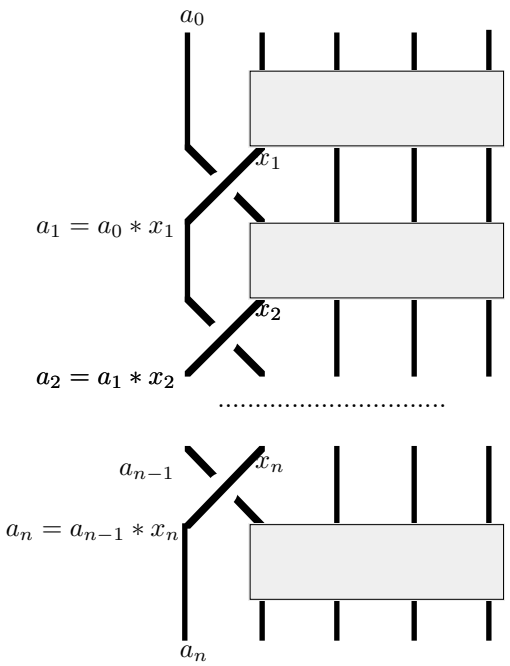



**Figure 3**: Une tresse « $\sigma_1$-positive » n'est pas triviale

On voit ainsi que les propriétés du système ∮ sont l'origine et le fondement de l'ordre sur les tresses, et, de là, de la méthode de réduction des poignées. Mais ceci, à son tour, laisse ouverte la question de l'origine de ces résultats d'algèbre: pourquoi étudier des objets aussi peu usuels que les LD-systèmes libres, pourquoi soupçonner une propriété comme l'acyclicité de la division?

[5] On passe ici sur une difficulté technique, qui n'est pas insurmontable, mais explique pourquoi on s'était limité dans la littérature à considérer les LD-systèmes classiques issus de la conjugaison des groupes: pour assurer la compatibilité des coloriages avec les relations $\sigma_i\sigma_i^{-1} = \sigma_i^{-1}\sigma_i = 1$, il faut se restreindre à des LD-systèmes dans lesquels la division à gauche est toujours possible, ce qui redonne essentiellement les exemples précédents, et exclut ∮. En fait, au prix d'un caractère partiel des coloriages, il suffit que le résultat de la division soit unique quand il existe, ce qui est le cas pour le système ∮: il n'est plus assuré que n'importe quelles couleurs initiales puissent être propagées à travers une tresse donnée, mais il reste vrai qu'il existe toujours des couleurs initiales pouvant être propagées, ce qui suffit par exemple pour la preuve ci-dessus de la Proposition 2.

## 5. L'algèbre des injections élémentaires

... Parce que la théorie des ensembles le prescrit clairement.

Comme on le rappelait au début, celle-ci se concentre sur la notion d'infini, notion qui se partage vite en des notions d'infini, ou *grands cardinaux*. On sait maintenant que ces notions constituent la clé d'une classification exhaustive des extensions du système de Zermelo-Fraenkel suivant le programme proposé par Gödel (*cf.* [24], ou [8] pour une introduction). Pour introduire de tels grands ensembles, le schéma naturel et usuel consiste à isoler les propriétés par lesquelles un ensemble infini dépasse tout ensemble fini, puis à renforcer ces propriétés pour définir de possibles ensembles « très infinis ». Par exemple, un ensemble infini est tellement grand qu'il est semblable à une de ses parties strictes, autrement dit, il existe une injection non bijective de cet ensemble dans lui-même, propriété que n'ont évidemment pas les ensembles finis. Si on considère le cas des entiers et de l'injection $n \mapsto n+1$, on voit que la similitude entre le tout et la partie peut inclure la préservation d'une structure additionnelle: ici l'ordre est préservé (mais évidemment pas la structure algébrique: le successeur de la somme de deux entiers n'est pas la somme des successeurs de ces entiers). Il devrait apparaître comme relativement naturel d'introduire comme « très infini » (on dira plutôt ici « autosimilaire ») un ensemble $A$ tel qu'il existe une injection non bijective de $A$ dans $A$ qui préserve *toute* propriété, c'est-à-dire tel qu'aucune propriété ne puisse distinguer la partie du tout. Dans le contexte de la théorie des ensembles, il faut entendre par propriété toute propriété *définissable* à partir des opérations ensemblistes de base, et on parle alors d'injection *élémentaire*. Il n'est guère difficile de voir qu'un ensemble autosimilaire doit être très grand. Par exemple, il ne saurait être dénombrable: si $I$ est une injection élémentaire non bijective de $A$ dans lui-même, et qu'une numérotation définissable des éléments de $A$ a été fixée, l'image par $I$ du premier élément de $A$ est forcément le premier élément, car $I$ doit préserver la propriété « être le premier élément », de même pour le second, ... Le même type d'argument montre que cardinal ne peut être $\aleph_1$, ou un quelconque cardinal $\aleph_\alpha$ avec $\alpha < \aleph_\alpha$.

Dans l'élaboration de la théorie des ensembles, des ensembles particuliers appelés *rangs* jouent un rôle technique privilégié car ils ont la propriété bien peu intuitive d'être clos à la fois par appartenance itérée (les éléments des éléments sont encore des éléments) et par formation d'ensembles (tout ensemble d'éléments d'un rang $R$ indexé par un élément de $R$ est encore un élément de $R$). Ces étranges propriétés entraînent que, si $I$ est une fonction définie sur un rang $R$, alors $I$ est essentiellement aussi un élément de $R$[6]. Supposons alors que $R$ soit un rang autosimilaire, et que $I$, $J$ soient deux injections élémentaires de $R$ dans lui-même. L'injection $I$ peut être appliquée à tout élément de $R$. Mais on vient de dire que $J$ elle-même est, ou peut être considérée comme, un élément de $R$: donc $I$ peut être *appliquée* à $J$, fournissant un nouvel objet $I(J)$.

[6] Pour de bénignes raisons d'« ensemble de tous les ensembles », ceci n'est pas littéralement vrai: $I$ est plutôt limite inductive d'éléments de $R$ qui l'approximent. Ceci ne change rien à l'idée.

Et comme, par définition, $I$ préserve toutes les propriétés définissables, $I(J)$ hérite toutes les propriétés définissables de $J$: $I(J)$ est encore une injection non triviale de $R$ dans lui-même, de surcroît élémentaire (l'élémentarité est une propriété définissable). Ainsi l'ensemble $E_R$ des injections élémentaires de $R$ dans lui-même se trouve muni d'une opération binaire, l'application (à ne pas confondre avec la composition, évidemment aussi possible, mais bien plus triviale).

L'opération d'application sur $E_R$ a une propriété algébrique évidente. En effet, si $x$, $y$ et $f$ sont trois éléments de $R$ tels que $f$ est une fonction et que $y$ est l'image de $x$ par $f$, et si $I$ est une injection élémentaire de $R$ dans lui-même, certainement $I(f)$ sera une fonction, et $I(y)$ sera l'image de $I(x)$ par $I(f)$, toujours parce qu'être une fonction ou être l'image de quelque chose par une fonction sont des notions définissables en théorie des ensembles. Cette relation vaut en particulier quand $f$ et $x$ sont eux-mêmes des injections élémentaires (ou plutôt d'abord des approximations convenables de celles-ci), ce qui se traduit par le fait que l'identité

$$I(J(K)) = I(J)(I(K))$$

est valable dans toute structure $E_R$. A la notation de l'opération près, on y reconnaît l'identité (LD) du paragraphe précédent: donc, pour chaque rang autosimilaire $R$, le système algébrique formé par $E_R$ avec l'application est un LD-système. En particulier, si $I$ est une injection élémentaire non bijective de $R$ dans lui-même, on peut introduire le sous-système $S(I)$ de $E_R$ engendré par $I$, c'est-à-dire la famille formée par toutes les itérées $I$, $I(I)$, $I(I(I))$, $I(I)(I)$, *etc.* Par construction $S(I)$ est un LD-système à un générateur.

L'étude des injections élémentaires et de leurs itérations est à la fois un sujet classique et un développement récent de la théorie des ensembles. Dans la mesure où les injections élémentaires sont la notion qui, dans le contexte de la théorie des ensembles, répond à celle d'homomorphisme en algèbre, il n'est guère surprenant de lui voir jouer un rôle fondamental dans la théorie moderne, notamment dans l'étude des grands cardinaux et la plupart des résultats majeurs obtenus depuis 1975. D'un autre côté, la situation envisagée ici des injections élémentaires d'un rang dans lui-même n'est pas la seule possible, et l'étude purement algébrique des structures $S(I)$ et du rôle de l'identité (LD) n'a débuté qu'assez récemment. Or l'un des premiers résultats a été

**Proposition 4.** (Laver, 1989, [29]) *Si $I$ est une injection élémentaire (non triviale) d'un rang dans lui-même, alors la division du LD-système $S(I)$ n'a pas de cycle.*

La preuve est *facile*: son principe est de tirer parti du bon ordre des ordinaux. Tout rang inclut un début de la suite des ordinaux (qui commence par les entiers naturels), et, comme, à nouveau, le fait d'être un ordinal est une propriété définissable, toute injection élémentaire $I$ d'un rang $R$ dans lui-même doit envoyer les ordinaux de $R$ sur d'autres ordinaux de $R$: 0 est certainement envoyé sur lui-même car « être 0 » est définissable, de même que 1, 2, ..., puis

$\omega$, le plus petit ordinal infini, puis son successeur $\omega + 1$, et ainsi de suite aussi longtemps qu'on ne rencontre que des ordinaux qui ont une définition. Pourtant il doit exister des ordinaux qui ne sont pas fixes par $I$, sinon $I$ entière serait l'identité. Il existe donc un plus petit ordinal qui est bougé par $I$, et qu'on appelle l'*ordinal critique* de $I$. Par les propriétés usuelles des bons ordres, l'ordinal critique de $I$ ne peut appartenir à l'image de $I$. Il est alors trivial que la division du LD-système $S(I)$ ne peut avoir de cycle de longueur 1, c'est-à-dire que toute égalité $J = J(K)$ est impossible: en effet, par élémentarité de $J$, l'ordinal critique de $J(K)$ est l'image par $J$ de l'ordinal critique de $K$, il appartient donc à l'image de l'application $J$, et ne peut être égal à l'ordinal critique de $J$. Un argument à peine plus compliqué montre que les cycles de longueur 2 sont impossibles, et la preuve générale utilise le même type de raisonnement sur les ordinaux critiques.

On voit alors immédiatement l'origine de la Proposition 3 concernant le LD-système libre ƒ: comme tout LD-système à un générateur est par définition quotient de ƒ, c'est en particulier le cas des systèmes $S(I)$ associés aux injections élémentaires des rangs dans eux-mêmes; or tout cycle pour la division dans ƒ se projetterait en un cycle analogue dans tout LD-système à un générateur. Ainsi c'est la Proposition 4 qui a mené à la Proposition 3, et, de là, aux Propositions 2, puis finalement 1. On a donc une filiation directe[7] entre un pur résultat de théorie des ensembles et un pur résultat de topologie des tresses.

## 6. Une application de la théorie des ensembles?

S'il y a bien trois années d'écart entre la preuve initiale de la Proposition 4 et celle de la Proposition 3 qui en semble un corollaire immédiat d'après l'argument ci-dessus, c'est que l'existence d'au moins un rang autosimilaire, sans lequel ce passage ne peut se faire, n'est qu'une *hypothèse* de la théorie des ensembles, de surcroît indémontrable[8]. Le résultat de Laver a créé une situation étrange, en donnant une preuve d'un résultat algébrique très simple et ne concernant que des objets très petits[9] à partir d'une hypothèse ensembliste indémontrable et portant sur l'existence d'objets « très grands »: le seul corollaire qu'on puisse tirer de la Proposition 4 est une version conditionnelle de la Proposition 3: « S'il existe un rang autosimilaire, alors la division du système ƒ n'a pas de cycle ». Eliminer cette hypothèse, si cela était possible, en construisant une preuve directe de la Proposition 3 devenait dès lors une tâche

[7] et effectivement conforme à l'enchaînement chronologique et heuristique des travaux: la description présente n'est pas une reconstruction *a posteriori*!

[8] La seule chose qu'on puisse éventuellement démontrer, comme pour tout axiome affirmant l'existence de grands ensembles, serait qu'elle soit contradictoire. Par contre, à cause du théorème d'incomplétude de Gödel, sa non-contradiction ne peut pas être établie.

[9] Il en allait de même pour des corollaires qu'on savait découler de la propriété d'acyclicité, notamment la décidabilité du problème de mot pour l'identité (LD), c'est-à-dire la question de reconnaître algorithmiquement si deux expressions formelles construites avec des variables et une opération binaire deviennent égales lorsque l'opération satisfait l'identité en question.

évidente[10], et c'est ce qui a été fait dans [9]: la preuve se fait par l'introduction d'outils purement algébriques, notamment l'introduction d'un groupe qui décrit en un certain sens la géométrie propre de l'identité (LD) et qui, se trouvant (de façon non fortuite) être une extension des groupes de tresses, est l'origine des applications ultérieures. De la sorte nulle hypothèse ensembliste ne vient plus grever les résultats affirmés dans les Propositions 1, 2 ou 3[11].

Dès lors, on pourrait être tenté de dénier aux résultats sur les tresses le caractère d'applications du résultat de Laver en théorie des ensembles. A la différence de certains logiciens qui auraient préféré que l'axiome « exotique » s'avère indispensable, je ne pense pas que son élimination ultérieure change rien au rôle essentiel joué ici par la théorie des ensembles. En particulier il existe un très grand écart de complexité entre la démonstration de la propriété d'acyclicité de la division pour les systèmes $S(I)$, et celle de la même propriété pour le système $\mathfrak{f}$ dans un contexte de pure algèbre, c'est-à-dire sans les outils spécifiques de la théorie des ensembles. Il est fort peu probable qu'on ait trouvé la motivation pour construire une preuve comme celle de [9] sans l'indication apportée par le résultat paradoxal de Laver. Plus précisément, on peut retenir qu'ici le rôle de la théorie des ensembles a été de désigner comme digne d'étude et source de questions profondes un objet, le LD-système libre $\mathfrak{f}$, qui n'a rien d'ensembliste et qui n'avait aucune raison particulière de recevoir cet intérêt[12]. Par ailleurs, elle a *révélé* une propriété-clé (la propriété d'acyclicité de la division de $\mathfrak{f}$), et en a établi, sinon la vérité, du moins la plausibilité, en les faisant dériver d'hypothèses indémontrables mais parties d'une approche globale cohérente et d'un programme jusqu'à présent exempt de contradiction.

De plus, les quelques détails qui ont été donnés sur l'apparition des structures algébriques $S(I)$ devraient persuader que ce rôle de révélateur joué par la théorie des ensembles ne doit rien au hasard: c'est précisément le contexte et les outils spécifiques de cette branche qui ont permis d'« accueillir » une intuition plutôt commune (celle de l'autosimilarité) et à l'élaborer en des énoncés précis et relativement profonds: s'il est assez naturel d'associer l'identité d'autodistributivité à l'idée d'autosimilarité, il semble peu évident de rendre ces liens directement exploitables sans le formalisme des rangs, objets à la fois naturels dans le contexte de la théorie des ensembles et complètement artificiels dans tout autre contexte; de même et surtout il aurait été bien difficile de faire apparaître l'acyclicité de la division sans les propriétés des ordinaux,

[10] Il n'était absolument pas clair *a priori* que cette élimination soit possible: on sait qu'il existe des résultats, comme la détermination des sous-ensembles projectifs de la droite réelle, qui nécessitent des hypothèses ensemblistes (*cf.* [8]), et rien n'aurait interdit (*cf.* [35]) qu'il en soit de même pour les résultats sur les LD-systèmes libres.

[11] On peut même éliminer à son tour la Proposition 3 en construisant une preuve directe (et plus simple) de la Proposition 2, ainsi que l'a fait D. Larue dans [27].

[12] Des résultats comme ceux de [7] avaient démontré que les propriétés purement algébriques (par opposition à logiques) des systèmes $S(I)$, et en particulier le fait qu'ils satisfassent à l'identité (LD), sont non triviales car responsables de conséquences logiques fortes. Par ailleurs Laver montre dans [29] que les systèmes $S(I)$ sont, s'ils existent, libres, donc isomorphes à $\mathfrak{f}$, et ce dernier système est bien ainsi désigné comme l'objet d'étude principal.

comme le montrent les contorsions techniques nécessaires dans [15] pour essayer de mimer très partiellement les systèmes $S(I)$ dans un contexte combinatoire « élémentaire ».

On notera que le rôle joué ici par la théorie des ensembles paraît très similaire à celui que joue par exemple la physique théorique lorsqu'elle introduit sur des bases heuristiques des énoncés ou des formules qu'il s'agit ensuite pour le mathématicien de justifier rigoureusement, c'est-à-dire sans faire appel à aucun des principes additionnels qui les ont fait apparaître. Dans le cas présent, l'introduction d'un axiome ensembliste supplémentaire, non démontré et même non démontrable, a permis d'isoler une propriété nouvelle, pour laquelle ensuite on a pu donner une démonstration « purifiée », c'est-à-dire en quelque sorte rendue rigoureuse par l'élimination des outils non fondés d'abord utilisés. Notons qu'introduire un axiome ensembliste additionnel est précisément s'autoriser un type de raisonnement logique supplémentaire, dont on conçoit bien qu'il puisse permettre de « démontrer » de nouvelles propriétés ou d'en rendre la preuve plus facile. Dans cette optique, la discussion sur le caractère vrai, plausible, intuitif, ou simplement naturel de l'axiome considéré devient secondaire devant la question principale qui est plutôt celle de sa richesse potentielle comme principe de démonstration: de ce dernier point de vue, on peut s'attendre à ce que les axiomes les plus improbables soient justement les plus féconds[13].

Il n'y a pas lieu de surestimer l'importance d'un exemple qui, de surcroît, reste isolé[14]: le problème de la classification des tresses est résolu depuis longtemps, et les améliorations qu'on a décrites ne démontrent aucune conjecture fameuse[15]... Pourtant, en montrant que des recherches sur des objets certainement non-dénombrables et non-définissables comme les rangs autosimilaires peuvent mener, d'une façon certes indirecte et imprévisible au départ, à des applications incontestablement concrètes, il me semble que l'existence d'une application comme celle décrite ici peut contribuer à dissiper certains doutes (légitimes) sur l'opportunité de telles recherches et la signification des résultats qu'elles apportent. Certes les objets de la théorie des ensembles « existent » probablement moins que d'autres, mais c'est précisément le caractère flexible et vague de leur faible notion d'existence qui les rend susceptibles de se plier à des contraintes proches de la contradiction, et par là permet d'élaborer des intuitions qui n'entreraient pas d'emblée dans des contextes plus rigides.

---

[13] Il se trouve que l'axiome « il existe un rang autosimilaire » est, du point de vue de la force logique, l'un des plus forts de ceux qui aient été considérés, et donc l'un des plus improbables, ou au moins l'un de ceux qu'on pourrait envisager avec le plus de suspicion.

[14] [11] propose une autre application de l'ordre des tresses, cette fois à la représentation de Burau.

[15] La comparaison avec la physique théorique porte sur le mécanisme d'interaction, et ne saurait bien sûr porter sur la profondeur des résultats acquis à ce jour! De même, il ne faut pas perdre de vue que, sur le plan de la profondeur mathématique intrinsèque, l'exemple décrit ici est probablement très anecdotique par rapport à des résultats majeurs de la théorie des ensembles comme ceux de Martin, Steel et Woodin sur les sous-ensembles de la droite réelle (*cf.* [8]).

Il ne reste qu'à espérer que de nouveaux exemples pourront dans l'avenir être développés suivant le schéma décrit ici, faisant à nouveau jouer à la théorie des ensembles le rôle d'un révélateur original et profond...

## 7. Envoi

... Et on terminera avec une question de combinatoire dont l'énoncé élémentaire ne doit pas masquer la probable très grande difficulté—et qui a bien des chances d'être la prochaine application « concrète » de la théorie des ensembles.

Sur l'ensemble $\{1, \ldots, N\}$, cherchons à construire un produit $*$ qui en fasse un LD-système en partant des valeurs aux bords

$$p * 1 = p + 1 \pmod{N}.$$

Il est facile de voir qu'il existe au plus une solution au problème, qu'on obtient par récurrence double « de gauche à droite et de bas en haut » en appliquant l'autodistributivité à partir de la dernière ligne, et, pour chaque ligne, de la première colonne. La table ainsi construite est alors effectivement un LD-système si, et seulement si, l'entier $N$ est une puissance de 2 ([34], [16]). Notons $A_n$ la table correspondant à l'intervalle $\{1, \ldots, 2^n\}$. Par exemple on pourra vérifier que la table $A_2$ est

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| 1 | 2 | 4 | 2 | 4 |
| 2 | 3 | 4 | 3 | 4 |
| 3 | 4 | 4 | 4 | 4 |
| 4 | 1 | 2 | 3 | 4 |

On montre que chacune des lignes des tables $A_n$ se présente comme la répétition périodique d'un motif de base (variable) formé d'une suite de valeurs croissant jusqu'à $2^n$, motif dont la longueur est elle-même une puissance de 2. Soit $\pi(n)$ la période de la première ligne de la table $A_n$, c'est-à-dire le nombre de valeurs distinctes qui y apparaissent: par exemple $\pi(2)$ est égal à 2. Comme la projection *modulo* $2^n$ donne un morphisme surjectif de $A_{n'}$ sur $A_n$ pour $n' \geq n$, la fonction $\pi$ est croissante au sens large.

**Proposition 5.** (Laver, [30]) *S'il existe un rang autosimilaire, alors la fonction $\pi$ tend vers l'infini.*

**Problème ouvert:** Donner de ce résultat une preuve « élémentaire » n'utilisant aucune hypothèse ensembliste.

On ne peut imaginer énoncé de nature plus finitiste que le précédent. Pourtant la seule preuve connue du caractère non borné de $\pi$ est celle de Laver qui réalise les tables $A_n$ comme quotients du système $S(I)$ associé à une injection

élémentaire[16] et traduit la propriété de $\pi$ en une propriété non triviale des ordinaux critiques. Bien que techniquement virtuoses, les efforts pour déterminer « à la main » les valeurs de la fonction $\pi$ n'ont pour le moment abouti qu'à des résultats très partiels [17]. On sait seulement que la première valeur de $n$ pour laquelle $\pi(n)$ dépasse 16 est un nombre gigantesque [14]—probablement le plus grand entier pour lequel on ait jamais donné une définition aussi simple.

Les tables $A_n$ sont des objets algébriques naturels qui jouent, vis à vis des LD-systèmes, un rôle semblable à celui des $\mathbb{Z}/n\mathbb{Z}$ pour les groupes [18]. On peut penser que leur étude n'est qu'un amusant passe-temps, mais il s'agit aussi de LD-systèmes d'un type nouveau et qui a l'avantage de se prêter aux calculs[17], et on peut espérer en tirer un jour des propriétés moins « futiles », par exemple par le biais de coloriages de tresses. Dans tous les cas, il s'agira là à nouveau d'applications de la théorie des ensembles...

## Références


[1] E. Artin, *Theory of Braids*, Ann. of Math. **48** (1947) 101–126.

[2] E. Brieskorn, *Automorphic sets and braids and singularities*, Braids, Contemporary Maths AMS **78** (1988) 45–117.

[3] S. Burckel, *The Wellordering on Positive Braids*, J. Pure Appl. Algebra, à paraître.

[4] P. Cartier, *Développements récents sur les groupes de tresses, applications à la topologie et à l'algèbre*, Séminaire Bourbaki, exposé 716 (novembre 1989).

[5] R. Charney, *Artin groups of finite type are biautomatic*, Math. Ann. **292-4** (1992) 671–683.

[7] P. Dehornoy, $\Pi^1_1$*-complete families of elementary sequences*, Ann. P. Appl. Logic **38** (1988) 257–287.

[8] —, *La détermination projective (d'après Martin, Steel et Woodin)*, Séminaire Bourbaki, Astérique **177–178** (1989) 261–276.

[9] —, *Braid Groups and Left Distributive Operations*, Trans. Amer. Math. Soc **345-1** (1994) 115–151.

[10] —, *From Large Cardinals to Braids via Distributive Algebra*, J. Knot Theory & Ramifications **4-1** (1995) 33–79.

[11] —, *Weak Faithfulness Properties for the Burau Representation*, Topology and its Applic., à paraître.

[12] —, *A Fast Method for Comparing Braids*, Advances in Math., à paraître.

[13] —, *Handle Reduction of Braids*, Logiciel de démonstration pour Macintosh, disponible sur http://www.math.unicaen.fr/.


[16] C'est par ce biais que Laver a, le premier, introduit les tables $A_n$

[17] La Proposition 5 exprime en des termes « naïfs » le fait que les tables $A_n$ convergent vers une structure libre, à savoir que le sous-système engendré par $(1, 1, \ldots)$ dans la limite projective des $A_n$ est libre: le calcul dans les tables $A_n$ est donc l'approximation naturelle d'un calcul dans le LD-système libre ƒ qui, lui, reste souvent inaccessible.

[6] P. Deligne, *Les immeubles des groupes de tresses généralisés*, Invent. Math. **17** (1972) 273–302.

[14] R. Dougherty, *Critical points in an algebra of elementary embeddings*, Ann. P. Appl. Logic **65** (1993) 211–241.

[15] R. Dougherty & T. Jech, *Finite left-distributive algebras and embedding algebras*, Advances in Math., à paraître.

[16] A. Drápal, *Homomorphisms of Primitive Left Distributive Groupoids*, Comm. in Algebra **22-7** (1994) 2579–2592.

[17] —, *Persistence of Left Distributive Algebras*, J. Pure Appl. Algebra, à paraître.

[18] —, *Finite Left Distributive Algebras with One Generator*, preprint (1995);

[19] E. A. Elrifai & H. R. Morton, *Algorithms for positive braids*, Quart. J. Math. Oxford **45-2** (1994) 479–497.

[20] D. Epstein & *al.*, *Word Processing in Groups*, Jones & Barlett Publ. (1992).

[21] R. Fenn & C. P. Rourke, *Racks and links in codimension 2*, J. of Knot Theory and its Ramifications (1992) 343–406.

[22] F. A. Garside, *The Braid Group and other Groups*, Quart. J. Math. Oxford **20** No.78 (1969) 235–254.

[23] D. Joyce, *A classifying invariant of knots: the knot quandle*, J. of Pure and Appl. Algebra **23** (1982) 37–65;

[24] A. Kanamori, *The Higher Infinite*, Springer Verlag (1994).

[25] C. Kassel, *Invariants des nœuds, catégories tensorielles et groupes quantiques*, Gazette des Mathématiciens **56** (1993) 63–80.

[26] L. Kauffman, *Knots and Physics*, World Scientific (1991).

[27] D. Larue, *On Braid Words and Irreflexivity*, Algebra Univ. **31** (1994) 104–112.

[28] —, *Left-Distributive and Left-Distributive Idempotent Algebras*, Ph D Thesis, University of Colorado, Boulder (1994).

[29] R. Laver, *The left distributive law and the freeness of an algebra of elementary embeddings*, Advances in Math. **91-2** (1992) 209–231.

[30] —, *On the algebra of elementary embeddings of a rank into itself*, Advances in Math. **110** (1995) 334–346.

[31] —, *Braid group actions on left distributive structures and well-orderings in the braid group*, J. Pure Appl. Algebra **108-1** (1996) 81–98.

[32] J. Roitman, *The Uses of Set Theory*, Math. Intelligencer **14–1** (1992) 63–69.

[33] P. Vogel, *Representation of links by braids: A new algorithm*, Comment. Math. Helvetici **65** (1990) 104–113.

[34] F. Wehrung, *Gerbes primitives*, C. R. Acad. Sci. Paris **313-I** (1991) 357–362.

[35] H. Woodin, *Large Cardinal Axioms and Independence: The Continuum Problem Revisited*, Math. Intelligencer **16-3** (1994) 31–35.